

Министерство образования и науки Российской Федерации  
Сибирский государственный университет науки и технологий  
имени академика М.Ф. Решетнева

На правах рукописи

Иванов Илья Андреевич

**Проектирование нейросетевых систем глубинного обучения эволюционными  
алгоритмами для задачи человеко-машинного взаимодействия**

05.13.01 – Системный анализ, управление и обработка информации  
(космические и информационные технологии)

Диссертация на соискание ученой степени  
кандидата технических наук

Научный руководитель:  
кандидат технических наук,  
доцент Е.А. Сопов

Красноярск – 2017

## Содержание

|                                                                                                                                                |           |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Введение.....</b>                                                                                                                           | <b>5</b>  |
| <b>Глава 1. Системный анализ проблемы применения методов машинного обучения и оптимизации в задачах человеко-машинного взаимодействия.....</b> | <b>14</b> |
| § 1.1 Обзор современных методов машинного обучения, классификации и оптимизации.....                                                           | 14        |
| § 1.2 Задача человеко-машинного взаимодействия и обзор существующих подходов к ее решению.....                                                 | 36        |
| § 1.3 Постановка задачи распознавания эмоций. Используемые базы данных и подходы на основе машинного обучения.....                             | 45        |
| Выводы.....                                                                                                                                    | 50        |
| <b>Глава 2. Разработка коллективного самоконфигурируемого эволюционного алгоритма многокритериальной оптимизации.....</b>                      | <b>52</b> |
| § 2.1 Эволюционные алгоритмы однокритериальной и многокритериальной оптимизации.....                                                           | 52        |
| § 2.2 Разработка и реализация самоконфигурируемого эволюционного алгоритма многокритериальной оптимизации.....                                 | 64        |
| § 2.3 Исследование эффективности самоконфигурируемого алгоритма на репрезентативном наборе тестовых задач оптимизации                          | 68        |
| Выводы.....                                                                                                                                    | 78        |
| <b>Глава 3. Разработка многокритериального подхода к проектированию ансамбля классификаторов и отбору информативных признаков.....</b>         | <b>80</b> |
| § 3.1 Настройка параметров и проектирование ансамблей алгоритмов машинного обучения.....                                                       | 80        |
| § 3.2 Разработка и реализация многокритериального подхода к отбору информативных признаков.....                                                | 86        |

|                                                                                                                                                                                                                                     |            |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| § 3.3 Разработка и реализация многокритериального подхода к проектированию ансамбля нейросетевых классификаторов.....                                                                                                               | 91         |
| § 3.4 Исследование эффективности многокритериального подхода к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов.....                                                                           | 96         |
| Выводы.....                                                                                                                                                                                                                         | 108        |
| <b>Глава 4. Разработка гибридного алгоритма обучения конволюционной нейронной сети с применением эволюционного алгоритма оптимизации.....</b>                                                                                       | <b>111</b> |
| § 4.1 Конволюционная нейронная сеть и суть методов глубинного обучения.....                                                                                                                                                         | 111        |
| § 4.2 Достоинства и недостатки алгоритма обратного распространения ошибки и эволюционного алгоритма.....                                                                                                                            | 116        |
| § 4.3 Разработка и реализация гибридного алгоритма обучения конволюционной нейронной сети.....                                                                                                                                      | 121        |
| § 4.4 Исследование эффективности гибридного алгоритма обучения конволюционной нейронной сети на тестовых задачах анализа изображений.....                                                                                           | 124        |
| Выводы.....                                                                                                                                                                                                                         | 132        |
| <b>Глава 5. Разработка обобщенного метода для решения задач анализа гетерогенных данных на основе разработанных алгоритмов.....</b>                                                                                                 | <b>134</b> |
| § 5.1 Разработка и исследование эффективности метода слияния аудио-видео информации на уровне данных и на уровне классификаторов в рамках задачи распознавания эмоций.....                                                          | 134        |
| § 5.2 Разработка обобщенного метода для решения задач анализа гетерогенных данных на основе слияния данных, многокритериального отбора признаков и оптимизации алгоритмов машинного обучения, и конволюционных нейронных сетей..... | 140        |

|                                                                                                                                                                                                      |            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| § 5.3 Исследование эффективности разработанного обобщенного метода на задаче распознавания эмоций.....                                                                                               | 143        |
| Выводы.....                                                                                                                                                                                          | 145        |
| <b>Заключение.....</b>                                                                                                                                                                               | <b>146</b> |
| <b>Список использованных источников.....</b>                                                                                                                                                         | <b>148</b> |
| <b>Список публикаций автора.....</b>                                                                                                                                                                 | <b>163</b> |
| <b>Приложение А. Сравнение разработанного алгоритма SelfCOMOGA с алгоритмами-победителями соревнования CEC по метрике IGD.....</b>                                                                   | <b>166</b> |
| <b>Приложение Б. Результаты экспериментов по исследованию эффективности многокритериального подхода к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов.....</b> | <b>173</b> |
| <b>Приложение В. Результаты экспериментов по исследованию эффективности гибридного алгоритма обучения конволюционной нейронной сети на задаче распознавания эмоций.....</b>                          | <b>183</b> |
| <b>Приложение Г. Результаты экспериментов по исследованию эффективности обобщенного метода анализа гетерогенных данных на задаче распознавания эмоций.....</b>                                       | <b>190</b> |

## Введение

**Актуальность темы.** Основная цель машинного обучения заключается в построении модели по имеющейся базе данных в соответствии с некоторым алгоритмом. В общем случае, алгоритмы машинного обучения не позволяют добиться высокой точности решения задачи без предварительной настройки параметров. Настройка параметров алгоритмов вручную может оказаться очень затратной по времени. Кроме того, эксперт в области машинного обучения должен обладать необходимыми знаниями о настраиваемом алгоритме и свойствах процесса обучения данного алгоритма.

Данная диссертация посвящена проблеме проектирования нейросетевых систем машинного обучения эволюционными алгоритмами при решении задач человеко-машинного взаимодействия.

Поиск по сетке является простейшим улучшением ручной настройки алгоритмов. Дальнейшим улучшением является использование алгоритмов однокритериальной и многокритериальной оптимизации. Tušar в своей работе использует дифференциальную эволюцию для многокритериальной оптимизации совместно с алгоритмом машинного обучения. Kohavi и John вели поиск подходящих параметров алгоритма C4.5 для построения деревьев решений. Согласно результатам, оптимизированные значения параметров алгоритма в большинстве случаев обеспечили лучшую либо не уступающую точность классификации. Похожие эксперименты проводились Младеничем для поиска параметров при решении задачи пост-пруннинга дерева решений. Оптимизируемым критерием выступала точность классификации дерева решений, вычисленная по 10-кратной кросс-валидации. Bohanec и Bratko представили алгоритм OPT, который на каждой итерации искал дерево решений, обеспечивающее наибольшую точность классификации среди всех деревьев того же размера. Bergstra использовал случайный поиск и алгоритм Древовидная оценка Парзена для поиска параметров нейронных сетей.

Работы некоторых авторов посвящены оптимизации параметров метода опорных векторов (support vector machine, SVM). Rossi и Carvalho провели сравнение 4 алгоритмов оптимизации параметров данного метода: генетический алгоритм, алгоритм клонируемой селекции, муравьиный алгоритм, алгоритм роя частиц. В некоторых случаях алгоритм SVM с параметрами по умолчанию оказался более эффективен, чем с оптимизированными параметрами. Lessmann, Stahlbock и Crone оптимизировали параметры алгоритма SVM с помощью генетического алгоритма. В сравнении с поиском по решетке генетический алгоритм обеспечил лучшие и более стабильные результаты.

В работах Almeida и Leung использовались эволюционные алгоритмы для инициализации параметров нейронных сетей. Красноярская научная школа Семенкина Е.С. также активно занимается темой разработки эволюционных алгоритмов. Ахмедовой Ш.А.К. был разработан коллективный алгоритм оптимизации, комбинирующий в себе различные бионические алгоритмы. Сергиенко Р.Б. разработал коэволюционный алгоритм многокритериальной оптимизации. Разрабатываемые данной научной школой эволюционные алгоритмы многокритериальной оптимизации используются для оптимизации параметров алгоритмов машинного обучения, таких как нейронные сети (Ахмедова Ш.А.К., Брестер К.Ю.), нечеткая логика (Сергиенко Р.Б.), генетическое программирование (Сопов Е.А.) и др. Эти алгоритмы используются для решения различных практических задач: задача распознавания эмоций человека по аудиозаписи и видеозаписи лица (Сидоров М.Ю.), задача выбора эффективных вариантов системы управления космическим аппаратом (Семенкина М.Е.) и др.

Несмотря на то, что тема исследована большим количеством ученых и специалистов, исчерпывающего решения проблемы не предложено. Более того, появляются новые задачи, методы и модели машинного обучения, для которых также требуется разработка методов автоматизированного проектирования. Таким образом, разработка методов автоматизированной настройки алгоритмов

машинного обучения в целом, и нейронных сетей в частности, является **актуальной научно-технической задачей.**

**Объектами исследования** данной работы выступают конволюционная нейронная сеть, нейронная сеть прямого распространения и эволюционные алгоритмы оптимизации.

**Предмет исследования** - оценка эффективности проектирования нейросетевых систем глубинного обучения и нейронных сетей прямого распространения эволюционными алгоритмами для решения задачи человеко-машинного взаимодействия.

**Целью работы** является совершенствование методов проектирования нейросетевых систем глубинного и машинного обучения.

Исходя из цели, были сформулированы следующие **задачи исследования:**

1. Провести анализ основных подходов к решению задачи распознавания эмоций, алгоритмов оптимизации и машинного обучения, включая методы глубинного обучения.
2. Разработать коэволюционный алгоритм многокритериальной оптимизации, программно его реализовать и исследовать его эффективность.
3. Разработать и программно реализовать многокритериальный подход к проектированию ансамбля классификаторов и отбору информативных признаков, провести сравнительный анализ эффективности с однокритериальным подходом.
4. Разработать конволюционную нейронную сеть с гибридным алгоритмом обучения на основе эволюционного алгоритма оптимизации, программно ее реализовать и исследовать эффективность.
5. Разработать обобщенный метод решения задач классификации, включающих использование гетерогенных аудио-видеоданных, провести его апробацию на задаче распознавания эмоций.
6. Провести анализ и сделать вывод об эффективности синтеза эволюционных алгоритмов оптимизации и алгоритмов глубинного обучения в целом и в рамках задачи распознавания эмоций.

**Методы исследования.** Для решения поставленных задач были использованы методы системного анализа, теории вероятности и математической статистики, эволюционных алгоритмов, машинного обучения и анализа скрытых закономерностей в данных.

**Научная новизна** диссертационной работы состоит в следующем:

1. Предложен новый коэволюционный алгоритм многокритериальной оптимизации, отличающийся от известных методов оценкой эффективности работы входящих в него коэволюционирующих алгоритмов-компонент.
2. Разработан новый многокритериальный подход к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов, отличающийся от известных подходов алгоритмом слияния классификаторов в ансамбль.
3. Разработан новый гибридный алгоритм обучения конволюционной нейронной сети, сочетающий в себе эволюционный алгоритм оптимизации и алгоритм обратного распространения ошибки, отличающийся от известных использованием F-меры в качестве оптимизируемого критерия для эволюционного алгоритма.
4. Предложен новый подход к слиянию аудиоинформации с видеоинформацией применительно к задаче распознавания эмоций, отличающийся от известных тем, что в нем осуществляется слияние информации как на уровне данных, так и на уровне нейросетевых классификаторов.
5. Впервые предложен обобщенный метод для решения задач классификации, включающих использование гетерогенных аудио-видеоданных, на основе многокритериального подхода к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов, а также конволюционной нейронной сети с гибридным алгоритмом обучения, отличающийся от известных совместным применением количественных аудио-видео признаков и цифровых изображений в качестве входных данных.



**Теоретическая значимость.** В результате выполнения диссертационной работы был разработан новый коэволюционный алгоритм многокритериальной оптимизации, проведено его экспериментальное сравнение с другими популярными эволюционными алгоритмами многокритериальной оптимизации.

Получены новые знания о способах отбора признаков в задачах машинного обучения на основе алгоритмов многокритериальной оптимизации, проведено сравнение данного подхода с другими способами отбора признаков и снижения размерности данных. Исследованы различные методы настройки гиперпараметров нейронных сетей прямого распространения.

Разработан гибридный алгоритм обучения конволюционной нейронной сети, основанный на эволюционном алгоритме оптимизации и алгоритме обратного распространения ошибки, проведена апробация разработанного метода на задаче распознавания эмоций, задаче распознавания рукописных цифр и задаче распознавания объектов. Разработанный гибридный алгоритм превзошел по эффективности стандартный алгоритм обратного распространения ошибки по критерию точности классификации и по критерию F-меры, в то время как генетический алгоритм, работающий отдельно от алгоритма обратного распространения ошибки, оказался неэффективен ввиду слишком высокой размерности пространства поиска.

Предложен обобщенный метод решения задач анализа гетерогенных данных, проведена его апробация на задаче распознавания эмоций. Алгоритм многокритериальной оптимизации SelfCOMOGA в составе обобщенного метода обеспечил наибольшую эффективность по сравнению с другими рассмотренными алгоритмами оптимизации. Мета-классификация оказалась самым эффективным методом слияния классификаторов в коллектив.

**Практическая значимость.** На основе предложенных алгоритмов и подходов разработаны программные системы (ПС), которые могут быть использованы исследователями в данной области как база для проведения собственных исследований, а также для совершенствования разработанных

методов. В ПС "Коэволюционный алгоритм однокритериальной оптимизации" реализован алгоритм однокритериальной оптимизации, а также решение с его помощью задач отбора признаков и оптимизации параметров нейронных сетей. ПС "Эволюционные алгоритмы многокритериальной оптимизации" объединяет в себе алгоритмы VEGA (Vector Evaluated Genetic algorithm), SPEA (Strength Pareto Evolutionary algorithm), NSGA-2 (Non-dominated Sorting Genetic algorithm-2) и SelfCOMOGA (Self-configuring Coevolutionary Multi-objective Genetic algorithm), также имеется возможность тестирования данных алгоритмов на большом наборе тестовых задач многокритериальной оптимизации, в том числе на задачах конкурса CEC, на задачах отбора признаков, проектирования ансамбля нейросетевых классификаторов. Также в данную ПС входит реализация обобщенного подхода для анализа гетерогенных данных на примере задачи распознавания эмоций. ПС "Конволюционная нейронная сеть с гибридным алгоритмом обучения" позволяет решать задачи анализа и классификации изображений, проводить исследования эффективности гибридного алгоритма обучения по критериям точности классификации и F-меры. Разработанные программные системы могут быть использованы в лабораторных практикумах по предметам "Эволюционные методы оптимизации", "Методы машинного обучения и анализа данных".

**Реализация результатов работы.** Предложенная система слияния аудио-видео информации применительно к задаче распознавания эмоций была разработана в рамках работы в международном проекте "Dialog Speech Systems" в Университете г. Ульм (ФРГ). На основе полученных результатов была подготовлена статья в соавторстве с коллегами из Университета Ульма, которая была представлена и опубликована в сборнике международной конференции ICINCO (г. Колмар, Франция, 2015 г.).

Также, вышеупомянутая система слияния аудио-видео информации, и предложенный многокритериальный подход к проектированию ансамбля нейросетевых классификаторов и отбору информативных признаков были

разработаны в рамках исследования, выполненного по гранту конкурса УМНИК. По результатам, полученным в рамках исследования, были опубликованы пять научных работ в Российских научных изданиях, а также в сборниках международных конференций.

Подходы и методы, предложенные в данной диссертационной работе, были применены в рамках выполнения НИР по следующим грантам и проектам:

1. Грант Президента РФ № МК-3285.2015.9, проект "Самоконфигурируемая метаэвристика решения задач нестационарной оптимизации стохастическими поисковыми алгоритмами".

2. "Автоматическая сегментация левого желудочка сердца на снимках магнитно-резонансной томографии на основе кластерного подхода", грант РФФИ (Российского фонда фундаментальных исследований) и Правительства Красноярского края № 16-41-243036.

3. Российско-германский проект: "Разработка эффективного алгоритмического обеспечения для автоматизированного проектирования распределенных мультилингвистических систем поддержки электронного документооборота на облачных вычислениях" - в рамках реализации мероприятия № 1.2.2 Проведение научных исследований научными группами под руководством кандидатов наук, госконтракт N14.740.12.1341.

4. Проектная часть государственного задания "Разработка и исследование самоконфигурируемых гиперэвристик решения сложных задач нестационарной мультимодальной оптимизации бионическими алгоритмами", №2.1676.2017/ПЧ.

В ходе выполнения работы были реализованы 4 программные системы, зарегистрированные в федеральной службе по интеллектуальной собственности (Роспатент).

#### **Основные защищаемые положения:**

1. Предложенный коэволюционный алгоритм многокритериальной оптимизации превосходит в среднем по метрике IGD алгоритмы оптимизации,

входящие в него в качестве компонент, а также другие алгоритмы оптимизации, участвовавшие в конкурсе СЕС, на тестовых задачах конкурса СЕС.

2. Результаты тестирования многокритериального подхода к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов превосходят результаты тестирования однокритериального подхода по точности классификации, полученной на наборе тестовых задач классификации.

3. Результаты тестирования гибридного алгоритма обучения конволюционной нейронной сети, сочетающей в себе эволюционный алгоритм оптимизации и алгоритм обратного распространения ошибки, превосходят результаты тестирования алгоритма обратного распространения ошибки по точности классификации, полученной на наборе тестовых задач классификации изображений, гибридный алгоритм предотвращает стагнацию процесса обучения сети.

4. Предложенный подход к слиянию аудиоинформации с видеоинформацией при слиянии соответствующих нейросетевых классификаторов обеспечил лучшую точность классификации, чем классификаторы, использующие только аудиоданные или только видеоданные.

5. Предложенный обобщенный метод для решения задач классификации, включающих использование гетерогенных аудио-видеоданных, позволяет улучшить точность классификации для задачи распознавания эмоций.

**Апробация работы.** Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих научно-практических конференциях: Пятая международная конференция "Системный анализ и информационные технологии" САИТ-2013, СФУ, Красноярск, 2013; XVIII Международная научная конференция "Решетневские чтения", СибГАУ, Красноярск, 2014; 11th International Conference on Informatics in Control, Automation and Robotics, ICINCO-2014, Вена, Австрия, 2014; Всероссийская научно-практическая конференция студентов, аспирантов и молодых

специалистов, посвященная дню авиации и космонавтики, СибГАУ, Красноярск, 2015; 12th International Conference on Informatics in Control, Automation and Robotics, ICINCO-2015, Колмар, Франция, 2015; IEEE Symposium Series on Computational Intelligence, Кейптаун, ЮАР, 2015; XIX Международная научная конференция "Решетневские чтения", СибГАУ, Красноярск, 2015; International Workshop on Mathematical Models and its Applications, IWMMMA-2015, СибГАУ, Красноярск, 2015; XV Международная научная конференция бакалавров, магистрантов, аспирантов и молодых ученых "Молодежь. Общество. Современная наука, техника и инновации", СибГАУ, Красноярск, 2016; V Всероссийская научно-методическая конференция с международным участием "Информационные технологии в математике и математическом образовании", КГПУ, Красноярск, 2016.

Кроме того, отдельные результаты работы обсуждались на научно-техническом семинаре исследовательской группы диалоговых систем при университете г. Ульм (ФРГ) в рамках программы Эйлера, научных семинарах кафедры системного анализа и исследования операций СибГАУ.

**Публикации.** По результатам диссертационной работы опубликовано тринадцать печатных научных работ, среди которых 5 в рецензируемых журналах из перечня ВАК, 3 проиндексированы в Scopus, 2 - в базе Web of Science. Список публикаций приведен в конце диссертации.

**Структура работы.** Диссертация состоит из введения, пяти глав, заключения, списка использованных источников и четырех приложений.

# **Глава 1. Системный анализ проблемы применения методов машинного обучения и оптимизации в задачах человеко-машинного взаимодействия**

## **§ 1.1 Обзор современных методов машинного обучения, классификации и оптимизации**

Машинное обучение - обширный раздел искусственного интеллекта, посвященный разработке алгоритмов для обучения машин решению практических задач [4, 26]. Машинное обучение находится на стыке дисциплин, таких как математическая статистика, методы оптимизации, информатика. Кроме того, практическая направленность машинного обучения связывает его со многими другими областями человеческих знаний, на первый взгляд никак не связанными с математикой и вычислениями. К примеру, медицинская информационная система, способная автоматически ставить диагноз пациента по входным симптомам, относится к приложениям машинного обучения, но для создания такой системы наряду со знаниями в области математических алгоритмов требуются также знания в предметной области решаемой задачи - медицине. На сегодняшний день сфера применения алгоритмов машинного обучения стала столь широка, что данная дисциплина стала связана с большим количеством технических и гуманитарных отраслей человеческой деятельности.

Различают два типа машинного обучения - индуктивное и дедуктивное. Задача индуктивного обучения состоит в выявлении общих закономерностей в данных с целью их систематичного описания, либо с целью прогнозирования будущих данных. При дедуктивном обучении создается некоторая общая модель на основании знаний экспертов предметной области, которая используется для вывода конечных заключений. Приложения дедуктивного обучения относят к отдельной области экспертных систем [14, 27], поэтому на практике под машинным обучением обычно понимают индуктивное обучение.

Индуктивное обучение, которое далее мы будем называть просто машинным обучением, как следует из его определения, неразрывно связано с данными. Данные, как правило, представляют собой некоторые прецеденты из предметной области, организованные в виде таблиц, в которых каждая строка представляет собой вектор, описывающий отдельный прецедент. Например, в задаче медицинской диагностики данные представляют собой таблицу симптомов пациентов, в которой каждая строка соответствует отдельному пациенту, а каждый столбец - отдельному симптому. Таким образом, машинное обучение тесно связано с другой развивающейся дисциплиной - интеллектуальным анализом данных (data mining) [7, 12], главной целью которого является обнаружение в данных ранее неизвестных и нетривиальных закономерностей.

Машинное обучение по типу обучения подразделяется на несколько подклассов:

1. Обучение с учителем (supervised learning).
2. Обучение без учителя (unsupervised learning).
3. Частичное обучение (semi-supervised learning).
4. Обучение с подкреплением (reinforcement learning).
5. Динамическое обучение (online learning).
6. Активное обучение (active learning).

Обучение с учителем является наиболее распространенным вариантом машинного обучения в современных практических приложениях. Данный тип обучения работает с данными, организованными в виде структуры "объект - метка". Задача состоит в обучении алгоритма, восстанавливающего некую зависимость между признаками объекта и метками. При этом исходные данные разбиваются на непересекающиеся выборки - обучающую и тестовую. Различают несколько типов задач обучения с учителем:

1. Задача классификации (classification) [15]. В данной задаче конечное множество возможных меток объектов, называемых метками классов, или просто классами. Задача алгоритма обучения состоит в правильном отнесении объекта к

одному из классов. Качество работы алгоритма определяется ошибкой классификации, то есть долей объектов тестовой выборки, отнесенных к неверному классу.

2. Задача регрессии (regression) [31] отличается от задачи классификации тем, что меткой каждого объекта служит действительное число, следовательно, множество возможных меток неограниченно. Алгоритм обучения аппроксимирует некоторую функциональную зависимость числовой метки объекта от его признаков.

3. Задача прогнозирования (forecasting) [132]. Объектами являются значения некоторого параметра (вектора параметров), расположенные по оси времени. Совокупность таких объектов называют временным рядом, а саму задачу - прогнозированием временных рядов. Задача алгоритма обучения - на основании имеющихся объектов сделать прогноз на будущее.

Обучение без учителя использует данные, в которых не заданы метки объектов, то есть каждый объект представляет собой вектор значений признаков, либо вектор расстояний в признаковом пространстве до остальных объектов выборки. Цель алгоритмов обучения без учителя - поиск зависимостей между объектами на основании данных об их признаках. Различают следующие задачи обучения без учителя:

1. Задача кластеризации (clustering) [16] заключается в разбиении выборки объектов на группы таким образом, чтобы объекты внутри группы были схожи по некоторым признакам, а объекты разных групп отличались по этим признакам. Так как в данной задаче нет меток классов, как в задаче классификации, критерий качества кластеризации может быть задан как отношение среднего межкластерного и среднего внутрикластерного расстояния между объектами. Чем больше среднее расстояние между кластерами и чем меньше среднее расстояние между объектами одного кластера, тем лучше алгоритм кластеризации разделил объекты.



2. Задача фильтрации выбросов (outlier detection) [18] состоит в поиске нетипичных объектов выборки, отличных от других. Данная задача может быть как конечной, так и вспомогательной при решении задач обучения с учителем. Например, в практической задаче обнаружения бракованных деталей на предприятии поиск объектов выборки с нетипичными признаками является самоцелью, тогда как при решении задачи классификации на основе неточных, ошибочных данных поиск и исключение из выборки выбросов может послужить повышению точности конечной системы обучения с учителем.

3. Задача сокращения размерности (dimensionality reduction) [2] заключается в применении некоторых преобразований над данными, переводящих исходные признаки к меньшему числу новых признаков без потери информации об объектах выборки. Алгоритмы, решающие данную задачу, объединены под названием алгоритмов факторного анализа. Задача сокращения размерности может также быть отнесена к обучению с учителем, так как есть подкласс задач сокращения размерности, называемый отбором признаков, в котором исходные признаки не трансформируются в новые, а лишь выбираются наиболее информативные из числа имеющихся.

4. Задача заполнения пропусков в данных (missing values imputation) [119] актуальна при работе с выборками, в которых присутствуют пропущенные значения. Цель данной задачи состоит в прогнозировании пропущенных значений по имеющимся данным.

Частичное обучение (semi-supervised learning) имеет дело с частично размеченными данными, то есть метки даны лишь для некоторых объектов выборки [57]. Например, в базе данных клиентов указан возраст одной трети людей, необходимо предсказать возраст остальной части людей, для которых он не указан.

Обучение с подкреплением (reinforcement learning) отличается тем, что в нем роль объектов играют пары "состояние - действие" (state - action), а меткой является реакция окружающей среды на произведенное действие,

характеризующая правильность действия [136]. Обучение с подкреплением является более сложной версией обучения с учителем, так как реакция среды может быть не мгновенной, а достаточно отдаленной во времени, что усложняет прогнозирование. Обучение с подкреплением нашло обширное применение в робототехнике, где робот (агент) в начале своего обучения ничего не знает об окружающей среде, но постепенно учится эффективному взаимодействию с ней, получая и анализируя реакцию среды на свои действия.

В динамическом обучении (online learning) объекты поступают один за другим, а не все сразу. В связи с этим алгоритму приходится обрабатывать каждый объект по отдельности, и дообучаться с учетом новых знаний о новом объекте. Модели динамического обучения крайне важны с точки зрения практики, так как в реальных задачах, как правило, данные не даны в полном объеме сразу, а поступают порциями с течением времени.

Наконец, при активном обучении (active learning) объекты поступают на вход алгоритма не случайно, а в определенной последовательности, заложенной в алгоритм, которая позволяет алгоритму эффективнее обучаться. Этот тип обучения связан с областью планирования эксперимента [1].

В данной работе акцент делается на исследование алгоритмов обучения с учителем и на решение задач классификации, поэтому остановимся на этой теме более детально.

**Постановка задачи классификации.** Формальная математическая постановка задачи классификации выглядит следующим образом [2]. Пусть  $X$  - множество объектов,  $Y$  - конечное множество меток классов этих объектов. Пусть существует некоторая неизвестная зависимость  $f: X \rightarrow Y$ , связывающая объекты и их метки. Данная зависимость известна лишь для объектов конечной обучающей выборки  $(X_{\text{обуч}} \in X) = \{(\bar{x}_1, y_1), \dots, (\bar{x}_n, y_n)\}$ . Задача - построить алгоритм, называемый также решающим правилом, способный классифицировать произвольный объект  $\bar{x} \in X, \bar{x} \notin X_{\text{обуч}}$ .

Алгоритмы классификации неразрывно связаны с данными, на которых они строятся и обучаются. Данные чаще всего представлены в виде признакового описания объектов классификации. При этом признаки могут быть в количественной, порядковой, либо номинальной шкале. Другой вариант представления данных, использующийся значительно реже - с помощью матрицы расстояний между объектами. Всегда существует возможность перехода от одного типа представления данных к другому путем использования какой-либо метрики, то есть способа вычисления расстояния между объектами на основе их признаков.

Ключевым понятием, неразрывно связанным с классификацией, является понятие признакового пространства. Если задан набор признаков  $f_1, \dots, f_m$ , которыми описывается объект, то множество  $X = D_{f_1} \times \dots \times D_{f_m}$  называют признаковым пространством,  $D_{f_i}$  - областью определения (множество допустимых значений) признака  $f_i$ .

Задачи классификации различают по типам и количеству классов:

1. Двухклассовая классификация.
2. Мультиклассовая классификация [137] - в задаче больше 2 классов.
3. Пересекающиеся/непересекающиеся классы - каждый объект может относиться только к одному классу, или же к нескольким классам.
4. Нечеткая классификация [90, 143] - требуется определить степень принадлежности объекта каждому из классов.

Существует несколько общепринятых способов оценки качества алгоритмов классификации:

1. Разделение выборки на обучающую и тестовую - модель строится по данным из обучающей выборки и тестируется по тестовой выборке. Критерием качества выступает ошибка классификации, либо среднеквадратичная ошибка в случае задачи регрессии, либо иной критерий.

2. Кросс-проверка (cross-validation) - исходные данные случайно разделяются на  $m$  подвыборок, процедура обучения и тестирования модели

проводится  $m$  раз: в первый раз модель обучается на всех подвыборках, кроме 1-ой, тестируется на 1-ой; затем обучается на всех подвыборках, кроме 2-ой, тестируется на 2-ой; и так далее. Такой процесс называется  $m$ -кратной кросс-проверкой.

3. Скользящий экзамен (leave-one-out experiment) - является  $n$ -кратной кросс-проверкой, где  $n$  - объем выборки. То есть сперва модель строится на всех объектах выборки, кроме 1-го, проверяется на 1-ом; затем данная процедура повторяется для всех объектов выборки.

**Обзор алгоритмов классификации.** Существующие алгоритмы классификации могут быть объединены в несколько групп по принципу их действия:

1. Классификаторы на основе сходства объектов.
2. Алгоритмы статистической классификации.
3. Классификаторы на основе разделимости классов в признаковом пространстве.
4. Алгоритмы логической классификации.
5. Нейронные сети.

Алгоритмы первой группы основаны на гипотезе компактности, в которой предполагается, что объекты одного класса чаще всего похожи друг на друга, а объекты разных классов отличны. Схожесть объектов в  $n$ -мерном признаковом пространстве может вычисляться по различным метрикам:

1. Евклидово расстояние

$$d_E(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (1.1)$$

2. Манхэттенское расстояние

$$d_{Man}(x, y) = \sum_{i=1}^n |x_i - y_i|. \quad (1.2)$$

3. Чебышевское расстояние

$$d_{ch}(x, y) = \max_i (|x_i - y_i|). \quad (1.3)$$

4. Расстояние Минковского  $p$ -го порядка

$$d_{Mink}(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}. \quad (1.4)$$

Наиболее известными алгоритмами этой группы являются следующие:

1. Метод построения эталонов. По обучающей выборке вычисляются точки-эталон, являющиеся центрами объектов каждого класса в признаковом пространстве, по формуле:

$$E_i^k = \frac{1}{n^k} \sum (x_i: f(\bar{x}) = k); i = \overline{1, m}; k = \overline{1, K}, \quad (1.5)$$

где  $E^k$  - эталон  $k$ -го класса,  $n^k$  - число объектов выборки  $k$ -го класса,  $m$  - размерность признакового пространства,  $K$  - число классов. По определенной метрике рассчитывается расстояние от классифицируемого объекта до каждого из эталонов, объект относят к тому классу, расстояние до эталона которого минимально.

2. Метод  $k$  ближайших соседей [38]. Для этого алгоритма данные должны быть представлены в виде матрицы расстояний между объектами выборки, вычисленных по определенной метрике. Рассматриваются классы  $k$  объектов, ближайших к классифицируемому. Объект относится к тому классу, который наиболее часто встречается среди его соседей. При решении практических задач данным методом очень важно выбрать правильную метрику вычисления расстояния между объектами, а также значение параметра  $k$ . При слишком маленьком значении параметра (например,  $k = 1$ ) алгоритм становится подвержен негативному влиянию выбросов, при слишком высоком значении  $k$  в расчеты включаются слишком много соседних объектов, что также может негативно сказаться на качестве классификации, так как среди соседей классифицируемого объекта может оказаться много объектов другого класса.

3. Метод потенциальных функций [3]. Данный метод построен на физическом принципе потенциала электрического поля заряженной частицы.

Рассчитывается расстояние от классифицируемого объекта до каждого объекта обучающей выборки. Решающее правило строится как в методе ближайших соседей, разница состоит в том, что объект выборки обладает некоторой метой важности ("зарядом") относительно классифицируемого объекта.

Алгоритмы статистической классификации основаны на оценивании плотности распределения классов по выборке. В зависимости от способа оценивания различают алгоритмы с параметрическим и непараметрическим оцениванием плотности, а также оцениванием плотности как смеси параметрических распределений. Известные алгоритмы этой группы:

1. Наивный байесовский классификатор [116]. Данный метод основан на предположении, что признаки, которыми описываются объекты выборки, статистически независимы. Данное предположение существенно облегчает задачу оценивания плотности распределения, так как вместо  $n$ -мерной плотности необходимо оценить  $n$  одномерных плотностей. Плотности могут оцениваться как параметрическим, так и непараметрическим способом. По правилу Байеса находятся апостериорные вероятности каждого из  $K$  классов при условии измерения признака  $x$  классифицируемого объекта:

$$P(i|x) = \frac{f(x|i) * P(i)}{\sum_{j=1}^m f(x|j) * P(j)}, i = \overline{1, K}, \quad (1.6)$$

где  $f(x|i)$  - оценка условной плотности распределения признака  $x$  для  $i$ -го класса,  $P(i)$  - оценка априорной вероятности класса. Решающее правило для классифицируемого объекта  $x$  выглядит следующим образом:

$$i^* = \underset{i}{\operatorname{argmax}} P(i|x), i = \overline{1, K}. \quad (1.7)$$

2. Метод парзеновского окна использует непараметрическое оценивание плотности [17] распределения классов по имеющейся выборке, следовательно, в нем не выдвигаются гипотезы о структуре функции плотности распределения. Решающее правило классификации объекта  $x$  выглядит следующим образом:

$$i^* = \operatorname{argmax}_i \lambda_i \sum_{j=1}^n [y_j = i] K\left(\frac{d(x, x_j)}{h}\right), i = \overline{1, K}, \quad (1.8)$$

здесь  $\lambda_i$  - цена правильного ответа для класса  $i$ ,  $n$  - объем выборки,  $y_j$  - класс  $j$ -го объекта,  $K(t)$  - ядерная функция,  $d(x, x_j)$  - расстояние между классифицируемым объектом  $x$  и объектом  $x_j$ ,  $h$  - ширина окна.

3. ЕМ-алгоритм (expectation-maximization) [65] оценивает плотность как смесь параметрических распределений. В данном алгоритме итеративно выполняются два этапа: этап оценки (estimation), на котором рассчитывается ожидаемое значение функции правдоподобия, и этап максимизации (maximization), на котором вычисляются параметры функции правдоподобия, доставляющие ей максимум.

Следующую группу составляют классификаторы на основе разделимости классов. Данные алгоритмы строят разделяющую поверхность в признаковом пространстве, разделяющую объекты на непересекающиеся классы. Наиболее известные методы данной группы следующие:

1. Линейный дискриминант Фишера [125], также известный как линейный дискриминантный анализ, применим, если выборка удовлетворяет следующим гипотезам: классы распределены по нормальному закону, и матрицы ковариаций классов равны. Линейный дискриминант Фишера является упрощением квадратичного дискриминанта. В случае двух классов в двумерном пространстве разделяющей поверхностью, построенной с помощью этого метода, будет прямая. В случае большего числа классов разделяющая поверхность будет кусочно-линейной.

2. Логистическая регрессия [85]. Для случая 2 классов строится линейный алгоритм классификации с решающим правилом вида:

$$\logReg(x, w) = \operatorname{sign}\left(\sum_{j=1}^m w_j x_j - w_0\right) = \operatorname{sign}((x, w)), \quad (1.9)$$

где  $w_j$  - вес  $j$ -го признака,  $w_0$  - порог принятия решения,  $w$  - вектор весов,  $(x, w)$  - скалярное произведение вектора весов и признаков объекта. Задача обучения алгоритма логистической регрессии состоит в нахождении оптимального вектора весов  $w$ , минимизирующего функцию потерь вида:

$$L(w) = \sum_{i=1}^n \ln(1 + \exp(-y_i(x, w))). \quad (1.10)$$

3. Метод опорных векторов (support vector machine) [49] является одним из наиболее популярных методов обучения с учителем по нескольким причинам:

- а. Наиболее быстрый метод нахождения решающего правила.
- б. Сводится к решению задачи квадратичного программирования в выпуклой области, которая всегда имеет единственное решение.
- в. Находит разделяющую поверхность классов с разделяющей полосой максимальной ширины, что способствует более уверенной классификации.

В случае двух классов и линейно разделимой выборки решающее правило алгоритма SVM принимает вид:

$$y_i((w, x_i) + b) \geq 1, i = \overline{1, n}; \quad (1.11)$$

где  $y_i$  - метка класса  $i$ -го объекта выборки,  $w$  - вектор весовых коэффициентов,  $n$  - объем выборки. Вектор  $w$  ищется в виде:

$$w = \sum_{i=1}^n \lambda_i y_i x_i; \quad (1.12)$$

В эту сумму с ненулевыми коэффициентами  $\lambda_i$  входят только те объекты выборки, которые лежат на разделяющей поверхности. Эти объекты называют опорными векторами. В случае линейно неразделимой выборки пространство признаков  $R^n$  переводится в пространство большей размерности  $N$  с помощью отображения  $\varphi(x)$ , в котором выборка становится линейно разделимой. В этом случае решающее правило ищется в виде:

$$y_i((w, \varphi(x_i)) + b) \geq 1, i = \overline{1, n}; \quad (1.13)$$



$$w = \sum_{i=1}^n \lambda_i y_i \varphi(x_i). \quad (1.14)$$

Существует также мультиклассовый метод опорных векторов, который сводится к разбиению задачи на несколько задач бинарной классификации по схеме "один против всех".

Принцип работы алгоритмов следующей группы, называемых алгоритмами логической классификации, основан на построении иерархических композиций простых правил. Наиболее ярким представителем этой группы алгоритмов является дерево решений [109]. Структура дерева состоит из так называемых ветвей и листьев, в ветвях записаны признаки, от которых зависит целевая функция, а в листьях - значения целевой функции. Деревья решений пригодны как для решения задач классификации, так и решения задач регрессии. В случае задачи классификации в листах дерева записывается класс, которому принадлежит объект классификации, в случае задач регрессии - действительное число.

Схема построения дерева выглядит следующим образом:

1. Выбирается очередной признак  $f_i(x)$ .
2. Значения выбранного признака разбиваются на несколько подгрупп, тем самым разделяя признаковое пространство на несколько подпространств.
3. Процедура рекурсивно повторяется, пока не будет достигнут один из критериев останова - максимальная глубина дерева, количество листьев, либо максимальный размер каждого листа.

В зависимости от того, по какому принципу выбирается очередной признак для разбиения, и как он разбивается, существует несколько вариантов данного алгоритма:

1. Алгоритм ID3 - в нем выбор очередного признака производится по критерию прироста информации.
2. Алгоритм C4.5 [108] является улучшенной версией ID3, выбор признака по критерию нормализованного прироста информации.

### 3. Алгоритм CART (classification and regression tree) [52].

На практике во избежание эффекта переобучения после построения дерева решений некоторые его ветви усекают для поддержания лучшей обобщающей способности, данная процедура называется pruning.

Наконец, последняя группа алгоритмов классификации, называемая нейронные сети [30], основана на принципе связи большого числа простых элементов, построенных в виде оптимальной структуры. В качестве простых элементов выступают искусственные нейроны - математические конструкции, создание которых было вдохновлено изучением нервных клеток головного мозга человека. Понятие нейронной сети было формализовано Маккалоком и Питтсом в 1943 году, а первый алгоритм обучения предложен Хеббом в 1949 году. В 1958 году Розенблатт изобрел персептрон, первую нейронную сеть, которая использовалась для распознавания образов. В те годы в научных кругах наблюдался высокий интерес к нейронным сетям, однако после того как Минский в 1969 году опубликовал формальное доказательство ограниченности персептрона при распознавании сложных, вариативных изображений, интерес к нейронным сетям в целом, и к персептрону в частности, резко снизился. Немного позднее, в 1974 году, был изобретен алгоритм обратного распространения ошибки, а в 1982 году предложены нейронные сети с обратными связями, обучение которых проводилось на основе минимизации так называемой функции энергии сети [76]. В 1986 году интерес к направлению нейронных сетей возобновился, после того как был существенно доработан алгоритм обратного распространения ошибки. Позднее данный алгоритм был с успехом применен для обучения нового поколения глубоких нейронных сетей.

Каждый искусственный нейрон представляет собой упрощенную математическую модель реального нейрона. На вход нейрона поступают выходные сигналы с других нейронов, у каждого связи есть свой вес, внутри нейрона входящий сигнал суммируется и проходит через активационную функцию, в результате чего получается выходной сигнал (рисунок 1.1):

$$a = f(\theta + \sum_{i=1}^m w_i x_i); \quad (1.15)$$

здесь  $a$  - выход нейрона,  $x$  - вектор входов,  $w$  - вектор входных весов,  $\theta$  - дополнительная компонента, называемая смещением,  $m$  - число входов,  $f(x)$  - нелинейная активационная функция, в качестве которой обычно берут функции насыщения, например сигмоиду или гиперболический тангенс.

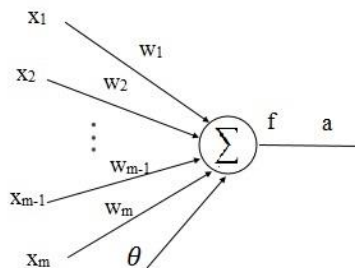


Рисунок 1.1. Модель нейрона

Помимо персептрона существует также полносвязная многослойная нейронная сеть прямого распространения [30] (рисунок 1.2). Ее название объясняется тем, что каждый нейрон промежуточного (скрытого) слоя связан с каждым нейроном предыдущего и следующего слоев, а сеть представляет собой ориентированный граф, в котором все ребра направлены в одну сторону, от входа к выходу, без обратных и циклических связей.

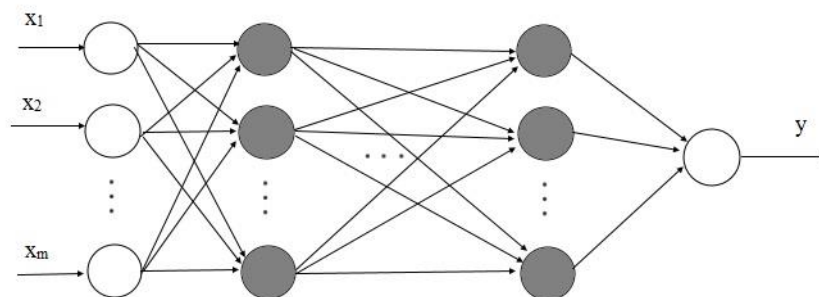


Рисунок 1.2. Структурная схема многослойной нейронной сети прямого распространения; серым цветом выделены нейроны скрытых слоев

Недостатком нейронных сетей является тот факт, что они решают задачу распознавания образов по принципу черного ящика, то есть не предоставляют понятных для человека логических пояснений, почему конкретный объект был классифицирован конкретным образом. В связи с этим возникает проблема выбора оптимальной структуры сети. Чаще всего структура выбирается методом проб и ошибок для каждой конкретной задачи распознавания. Что касается настройки параметров сети, таких как число нейронов на каждом скрытом слое, число итераций обучения сети, были предложены алгоритмы автоматической настройки данных параметров, основанные на алгоритмах оптимизации.

**Обзор алгоритмов оптимизации.** Ниже приведем постановку задачи оптимизации и обзор основных классов алгоритмов оптимизации.

Оптимизацией называют выбор наилучшего варианта из множества заданных альтернатив. В формальной математической постановке задача однокритериальной безусловной оптимизации задается следующим образом: пусть дано множество  $X \subset R^n$ , задана целевая функция (ЦФ) - отображение  $f: X \rightarrow R$ , и критерий максимизации либо минимизации. Тогда решить задачу оптимизации  $\min_{\bar{x} \in X} f(\bar{x})$  означает найти  $\bar{x}^* \in X: f(\bar{x}^*) = \min_{\bar{x} \in X} f(\bar{x})$ .

Приведем классификацию задач оптимизации.  $f(\bar{x})$  представляет собой набор критериев размерности  $k$ , которые должны быть оптимизированы одновременно. Если  $k = 1$ , то задача оптимизации называется однокритериальной, если  $k > 1$ , задача оптимизации называется многокритериальной (векторной).

Вектор неизвестных  $\bar{x}$  может содержать одну компоненту ( $n = 1$ ), такая задача называется одномерной (линейной), либо больше одной компоненты ( $n > 1$ ), такая задача называется многомерной.

Если пространство  $X$  не имеет каких-либо ограничений, то задача оптимизации безусловная, в противном случае условная.

Целевая функция и функция, задающая ограничения, могут быть заданы не аналитически, а в виде компьютерной программы, имитационной модели или как выход объекта. Такие задачи называются задачами с неявными функциями, или поисковыми задачами оптимизации. Если же все функции, входящие в постановку задачи, записаны в явном аналитическом виде, такая задача называется задачей математического программирования [24]. Чаще всего математическим программированием называют задачи с ограничениями, подчеркивая их отличие от задач безусловной оптимизации с явно заданной ЦФ. Кроме того, в математическое программирование не включают задачи, имеющие более одного локального экстремума.

Различают задачи локальной и глобальной оптимизации. В задачах локальной оптимизации задача состоит в нахождении локального экстремума в некоторой подобласти, в задачах глобальной оптимизации ищется глобальный экстремум во всей области определения функции. В случае оптимизации унимодальных функций имеется всего один экстремум, который одновременно является локальным и глобальным.

Если ЦФ и функция ограничений являются линейными относительно независимых переменных, такая задача называется задачей линейного программирования. Если ЦФ квадратичная, а ограничения линейные, то это задача квадратичного программирования. Если ЦФ выпуклая, а функции ограничений образуют выпуклую допустимую область, то имеем задачу выпуклой оптимизации. В случае функций общего вида говорят об общей задаче нелинейного программирования.

Алгоритмы оптимизации также имеют свою классификацию. Они подразделяются на детерминистские и стохастические. Детерминистские методы представляют собой алгоритмы с четкой последовательностью шагов, в которых исключена случайность. Стохастические алгоритмы, как следует из названия, включают в себя элемент случайности в той или иной форме. Преимущество детерминистских методов в том, что они "гарантируют" нахождение оптимума,

если он существует, и если были соблюдены все нужные предпосылки для его нахождения. Такими предпосылками могут быть структура модели, ее параметры, а также начальная область поиска. Недостаток детерминистских алгоритмов оптимизации в том, что почти все они являются алгоритмами локальной оптимизации, и непригодны для поиска глобального оптимума. Для стохастических алгоритмов нет гарантий сходимости к оптимуму, однако, как показывает практика, данный тип алгоритмов более пригоден именно для глобальной оптимизации, так как за счет введения случайности в процесс поиска алгоритмы данного типа не застревают в локальных экстремумах.

По типу работы с ЦФ методы оптимизации делятся на методы нулевого, первого (градиентные) и второго (Ньютоновские) порядка [6]. Методы нулевого порядка наименее требовательны к виду оптимизируемой функции, требуются лишь значения оптимизируемой функции в точках оптимизации, и не требуется аналитическое задание функции или ее производной. Функция при этом может быть задана любым образом, в том числе алгоритмически в виде компьютерной программы. В методах первого порядка вычисляются первые частные производные целевой функции от входных переменных, а в методах второго порядка - вторые частные производные, то есть гессиан целевой функции. Следовательно, для применения методов первого порядка необходимо, чтобы ЦФ была дифференцируема по каждой из входных переменных, для использования методов второго порядка необходимо, чтобы она была дважды дифференцируема.

Различают методы прямого поиска, линейной и квадратичной аппроксимации ЦФ, называемые методами спуска. Если метод не предполагает за функцией никаких свойств, а лишь вычисляет ее значения в точке и переходит в другую точку на основе информации о значении ЦФ, такой метод называют методом прямого поиска.

По степени математической обоснованности методы делят на рациональные и эвристические. Рациональные методы ориентированы на некоторую математическую модель ЦФ и обладают строгим доказательством сходимости к

точке экстремума и оценкой скорости сходимости. Эвристические методы не используют модель ЦФ, строгие доказательства их сходимости и теоретические оценки скорости сходимости обычно отсутствуют. Критерием остановки алгоритма обычно служит такое условие как малое приращение ЦФ на нескольких последовательных шагах оптимизации, что характерно при приближении к точке экстремума, либо иной критерий.

Среди простейших методов одномерной оптимизации можно выделить метод деления отрезка пополам, метод золотого сечения, метод Фибоначчи, метод квадратичной аппроксимации.

Большинство методов многомерной оптимизации основаны на следующей рекуррентной схеме:

$$\overline{x}^{k+1} = \overline{x}^k + \lambda^k \overline{S}^k; \quad (1.16)$$

где  $\overline{x}^k$  - вектор независимых переменных на итерации  $k$  алгоритма,  $\lambda^k$  - длина шага,  $\overline{S}^k$  - направление поиска. Различные поисковые алгоритмы отличаются друг от друга способом выбора длины шага и направления поиска на каждой итерации алгоритма. Методы многомерной оптимизации сводятся к методам одномерной оптимизации и предполагают ту или иную степень гладкости ЦФ. Они не дают гарантий сходимости к глобальному экстремуму, а в лучшем случае сходятся к одному из локальных.

Наиболее известный пример методов прямого поиска - метод покоординатного спуска, называемый также методом Гаусса-Зейделя. Его суть заключается в следующем: в исходной точке поиска  $\overline{x}^0$  все ее компоненты, кроме одной (например,  $x_1$ ), фиксируются, и осуществляется одномерная оптимизация ЦФ по  $x_1$  любым из перечисленных выше алгоритмов. В результате оптимизации получаем новое значение координаты  $x_1^1$ . В этой точке фиксируем все координаты, кроме второй, и повторяем процедуру одномерной оптимизации. В результате циклического перебора всех координат получаем новое приближение  $\overline{x}^1$ . Далее указанная процедура продолжается, пока не будут выполнены условия остановки:

$$|f(\overline{x^{k+1}}) - f(\overline{x^k})| \leq \varepsilon; \quad (1.17)$$

$$|\overline{x^{k+1}} - \overline{x^k}| \leq \delta. \quad (1.18)$$

Помимо этого, к методам прямого поиска также относятся метод Хука-Дживса (поиск по образцу), метод деформируемого многогранника (метод Нелдера-Мида), метод вращающихся координат (метод Розенброка), метод параллельных касательных (метод Пауэлла).

Еще один класс методов оптимизации называется случайным поиском [29]. В случае простых ЦФ методы случайного поиска менее эффективны, чем любая другая регулярная оптимизационная процедура. Однако они полезны, когда ЦФ настолько сложна, что невозможно заранее установить ее свойства, позволяющие рационально выбрать направление поиска. Эти методы пригодны для оптимизации любой ЦФ независимо от того, является она унимодальной или нет. Известно, что для любого регулярного алгоритма можно построить такую функцию, на которой он не будет работать. Методы случайного поиска позволяют оптимизировать любую функцию, но с различной, зачастую низкой эффективностью. Методы случайного поиска зачастую применяют в сочетании с методами других типов.

Общая формула методов случайного поиска выглядит следующим образом:

$$\overline{x^{k+1}} = \overline{x^k} + \overline{\xi^k}; \quad (1.19)$$

где  $\overline{\xi^k}$  - многомерная случайная величина. Различные методы случайного поиска отличаются вероятностным распределением этой случайной величины и ее изменением на каждом шаге алгоритма.

К простейшим методам случайного поиска относят алгоритм с парной пробой, алгоритм с возвратом при неудачном шаге, алгоритм наилучшей пробы, поиск на гиперсфере. Существует множество алгоритмов случайного поиска, отличающихся друг от друга соотношением случайности/регулярности, способом выбора случайного направления, поведением при успешном и неудачном шаге, и т.д.



В отличие от методов прямого поиска, градиентные методы безусловной оптимизации используют первые производные ЦФ по независимым переменным, и являются методами линейной аппроксимации, то есть ЦФ на каждом шаге алгоритма заменяется касательной гиперплоскостью. На очередном шаге градиентного алгоритма переход из предыдущей точки в следующую осуществляется по общей формуле, сходной с формулой (16). Различие заключается в способе выбора направления поиска  $\bar{S}^k$ , который вычисляется по формуле:

$$\bar{S}^k = -\nabla f(\bar{x}^k). \quad (1.20)$$

Градиент ЦФ в любой точке есть вектор, направленный в сторону наибольшего возрастания функции, следовательно антиградиент является направлением наискорейшего спуска в точке  $\bar{x}^k$ . Антиградиент дает только направление спуска, но не величину шага. Процедура спуска повторяется, пока не выполнится условие  $\nabla f(\bar{x}^k) = \bar{0}$ .

Известными представителями градиентных методов оптимизации являются метод градиента с фиксированным шагом, метод наискорейшего спуска, метод Флетчера-Ривса, алгоритм Полака-Рибьера.

В Ньютоновских методах направление поиска, соответствующее наискорейшему спуску, связано с аппроксимацией ЦФ. При разложении ЦФ в ряд Тейлора отбрасываются члены 3-го и более высших порядков. В методе Ньютона направление поиска имеет вид:

$$f(\bar{x}^k + \bar{S}) = f(\bar{x}^k) + \nabla^T f(\bar{x}^k) \cdot \bar{S} + \frac{1}{2} \bar{S}^T \cdot G(\bar{x}^k) \cdot \bar{S}, \quad (1.21)$$

где  $\nabla f(\bar{x}^k)$  - градиент ЦФ в точке  $\bar{x}^k$ ,  $G(\bar{x}^k)$  - гессиан в точке  $\bar{x}^k$ . Направление поиска, определенное этим направлением, называется Ньютоновским направлением. Минимум ЦФ достигается при следующих условиях:

$$G(\bar{x}^k) \cdot \bar{S} = -\nabla f(\bar{x}^k); \quad (1.22)$$

$$\bar{S} = -G^{-1}(\bar{x}^k) \cdot \nabla f(\bar{x}^k). \quad (1.23)$$

Кроме метода Ньютона среди Ньютоновских методов можно выделить метод Давидона-Флетчера-Пауэлла, который основан на использовании Ньютоновских направлений, но не требует затратного вычисления обратной матрицы Гессе  $G^{-1}$  на каждом шаге.

Задачи оптимизации нелинейных функций с нелинейными ограничениями называются задачами нелинейного программирования. В общем виде постановка таких задач имеет следующий вид:

$$f(\bar{x}) \rightarrow \max_{\bar{x}}/\min, \quad (1.24)$$

$$g_i(\bar{x}) \leq b_i; \quad i = \overline{1, k}, \quad (1.25)$$

$$g_i(\bar{x}) = b_i; \quad i = \overline{k+1, m}. \quad (1.26)$$

В Евклидовом пространстве система ограничений (1.25)-(1.26) определяет область допустимых решений задачи. В отличие от задач линейного программирования, в которых функции ограничений и ЦФ имеют линейный вид, данная область не всегда является выпуклой.

К методам нелинейного программирования [5] относятся метод множителей Лагранжа, метод Франка-Вульфа, метод штрафных функций, метод Эрроу-Гурвица. Однако данные методы являются градиентными, а, следовательно, гарантированно находят лишь локальный экстремум.

Методы глобальной оптимизации в теоретическом плане развиты гораздо слабее, чем методы локальной оптимизации, однако на практике представляют больший интерес.

К детерминированным алгоритмам глобальной оптимизации относятся метод ветвей и границ и метод покрытий. В методе ветвей и границ исходная область поиска разбивается на подобласти, в которых производится оценка верхней и нижней границы ЦФ. Рассчитанные для разных подобластей оценки сравниваются, и производится отсев подобластей поиска, заведомо не содержащих глобального экстремума.

Среди стохастических алгоритмов глобальной оптимизации простейшим является мултистарт алгоритмов локального поиска. Позднее были также

предложены некоторые его модификации. Другой класс данной группы методов - алгоритмы адаптивного случайного поиска [28], которые могут быть использованы как для решения локальных, так и глобальных задач оптимизации. Принцип их работы заключается в следующем: генерируется множество начальных точек поиска, в процессе случайного поиска, описываемого формулой (19), образуются новые точки, которые заменяют собой худшие точки, на предыдущем этапе.

Отдельно среди стохастических алгоритмов глобальной оптимизации стоит отметить подкласс эволюционных алгоритмов (ЭА), одним из представителей которого является генетический алгоритм (ГА) [13]. В нем использована идея естественного отбора, заимствованная из эволюции. Преимущество данного алгоритма состоит в том, что в нем не требуется никакой априорной информации об оптимизируемой ЦФ, он относится к группе алгоритмов прямого поиска, однако при этом в нем присутствуют механизмы, направляющие поиск в сторону глобального экстремума и препятствующие попаданию в локальные экстремумы. Главный недостаток алгоритма, как и любого другого стохастического метода, заключается в том, что нет гарантий нахождения решения задачи, а также в высоких вычислительных затратах. Поэтому в случаях, когда имеется некоторая информация об оптимизируемой ЦФ, либо при решении простых задач ГА не используют.

В отличие от задач однокритериальной оптимизации, задачи многокритериальной оптимизации являются гораздо более сложными, и не существует методов, способных гарантированно находить Парето-оптимальные решения для произвольной задачи. Наиболее часто применяющийся подход при решении задач оптимизации с несколькими критериями заключается в сведении задачи к однокритериальной оптимизации. Делается это с помощью взвешенного суммирования критериев, где каждому критерию соответствует некоторый вес в соответствии с его важностью. Однако такой подход имеет свои недостатки:

1. При настройке весов критериев требуется участие эксперта, обладающего знаниями о природе решаемой задачи.

2. Данный подход выдает единственное решение, которое напрямую зависит от выбранных весов критериев.

Альтернативой является применение эволюционных алгоритмов многокритериальной оптимизации. Данный класс алгоритмов получил значительное развитие. В частности, были предложены алгоритмы, основанные на концепциях прямого Парето-доминирования, которые на выходе выдают не одно оптимальное решение, а множество Парето-оптимальных решений. Стратегия одновременной оптимизации популяции решений, используемая в эволюционных алгоритмах, хорошо подошла при поиске множества и фронта Парето.

## **§ 1.2 Задача человеко-машинного взаимодействия и обзор существующих подходов к ее решению**

Человеко-машинное взаимодействие (ЧМВ) - это процесс связи между устройством и его пользователем. Задача ЧМВ появилась достаточно давно, с появлением первых технических устройств. Однако существенное развитие она получила после изобретения компьютеров. Появилось отдельное направление данной задачи, называемое человеко-компьютерным взаимодействием (human-computer interaction, HCI).

Область человеко-машинного взаимодействия связана с концепцией философии техники, в которой утверждается, что технические средства, разрабатываемые человеком, являются продолжением его органов чувств, призванным расширить их возможности. В связи с этим, проблема создания эффективного человеко-машинного интерфейса является крайне важной как с методологической, так и с практической точек зрения.

На заре компьютерной техники взаимодействие между компьютером и пользователем (оператором) было крайне затруднено. Например, команды для

первой электронной машины ENIAC (1943 г.) передавались с помощью переключателей и кабелей со штекерами, которые нужно было вставлять в нужные разъемы. В электронной вычислительной машине (ЭВМ) Mark I (1945 г.) стали применяться перфоленты с нанесенной на них программой. В 1960 г. Ликлайдером была выдвинута идея объединения человеческого интеллекта и вычислительных мощностей ЭВМ. Для воплощения данной идеи в жизнь им были сформулированы промежуточные цели. К краткосрочным целям относились следующие:

1. Электронный ввод/вывод текстовой и графической информации.
2. Разделение компьютерного времени между отдельными пользователями.
3. Создание систем хранения и поиска информации.
4. Создание интерактивных систем обработки информации и программирования.

Среднесрочные цели:

1. Обучение ЭВМ распознаванию речи оператора и рукописных текстов.
2. Координация объединений разработчиков при проектировании и программировании крупных систем.
3. Создание светового пера - устройства ввода координат на дисплее.

Долгосрочные цели:

1. Обучение ЭВМ распознаванию естественного языка произвольного пользователя.
2. Создание интеллектуальных программ, обладающих гибкостью и эвристичностью.

На сегодняшний день краткосрочные цели, сформулированные Ликлайдером, полностью выполнены. Среднесрочные цели также можно считать выполненными. Сегодня сложно представить работу на компьютере без такого устройства, как компьютерная мышь, благодаря сети Интернет существует множество средств совместной разработки больших систем и приложений

программистами, находящимися сколь угодно далеко друг от друга, а также создаются системы распознавания рукописных текстов и человеческой речи благодаря прогрессу в областях искусственного интеллекта и машинного обучения. Что касается долгосрочных целей, они пока находятся на стадии реализации, хотя в последние годы наметились перспективы существенного сдвига в этих областях.

Прогресс в создании систем человеко-компьютерного взаимодействия обусловлен как созданием новых технических средств, так и совершенствованием их алгоритмического обеспечения.

Сегодня персональный компьютер в силу постоянного удешевления его производства уже не является роскошью. Процессоры с каждым годом становятся все производительнее и компактнее, на смену большим стационарным компьютерам пришли ноутбуки, планшеты, смартфоны. Эти причины обусловили широкое распространение вычислительной компьютерной техники среди широких масс населения, не являющихся профессионалами в области информационных технологий. В связи с этим возникает потребность в создании интуитивно понятных интеллектуальных пользовательских интерфейсов.

Наряду с аппаратным обеспечением, создание эффективных систем человеко-машинного взаимодействия требует внедрения передовых разработок алгоритмического обеспечения. В последнее время при создании приложений на основе искусственного интеллекта на практике используются различные алгоритмы машинного обучения, классификации и оптимизации, описанные в § 1.1. Одним из новых направлений решения задачи человеко-машинного взаимодействия стало создание диалоговых систем, способных вести интеллектуальный диалог с пользователем.

Диалоговая система состоит из следующих блоков (рисунок 1.3):

1. Блок сбора сведений о пользователе: пол, возраст, эмоциональное состояние, прочее.
  - а. Сбор сведений в интерактивном формате "вопрос-ответ".

- б. Автоматический сбор сведений по входящей информации, и их уточнение в процессе диалога.
2. Блок распознавания входной информации.
  - а. Распознавание речи.
  - б. Распознавание рукописного текста
  - в. Распознавание жестов.
3. Блок управления диалогом - хранит историю и текущее состояние диалога, управляет общим ходом диалога.
4. Блок хранения знаний предметной области - знания могут храниться в виде базы правил, базы данных, либо в ином виде. Данных блоков может быть несколько, к ним в процессе работы обращается блок управления диалогом.
5. Блок генерации ответа системы.
  - а. Генератор естественного языка (текст).
  - б. Преобразование текста в речь (аудио).
  - в. Генератор жестов, мимики (антропоморфные роботы).

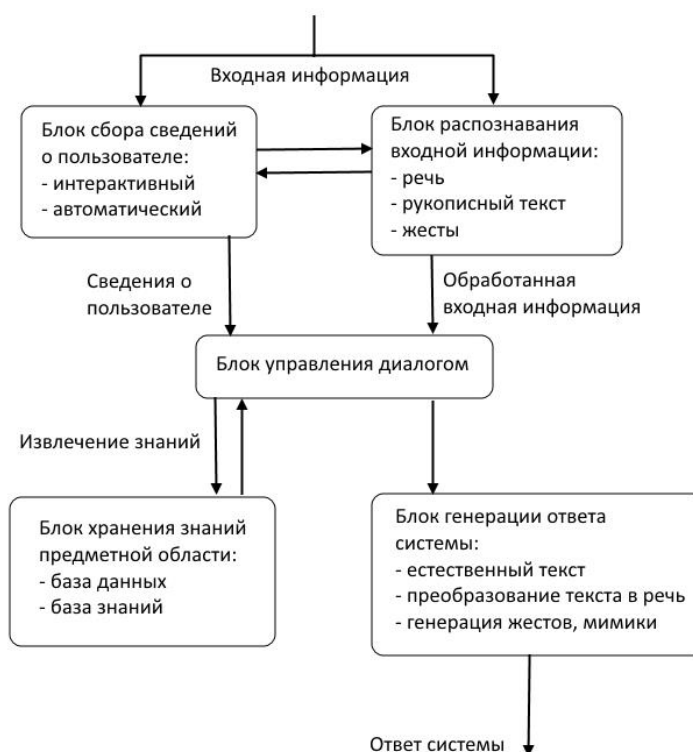


Рисунок 1.3. Принципиальная схема диалоговой системы

Блок сбора сведений о пользователе осуществляет сбор информации, которая может быть использована для построения интеллектуального диалога с пользователем, а также для более эффективной работы блока распознавания входной информации. Сведения о пользователе могут включать в себя как статические параметры в рамках процесса диалога, такие как пол, возраст, национальность, так и динамические, например эмоциональное состояние, отвлечение внимания. Знание этих параметров позволяет системе вовремя реагировать на изменения состояния пользователя и принимать необходимые проактивные действия. Например, в автоматизированных колл-центрах может производиться отслеживание эмоционального состояния клиента по голосу, и в случае, если система обнаружит, что клиент раздражен либо обозлен, она переключит его на человека-оператора. Или, распознав национальность говорящего, система может автоматически сменить язык общения с пользователем.

Статические сведения о пользователе могут быть запрошены в интерактивном режиме и записаны системой. Динамические же параметры необходимо отслеживать в процессе ведения диалога в режиме реального времени с определенным шагом по оси времени, например каждые 1-2 секунды, следовательно их необходимо оценивать автоматически.

Данная проблема порождает целый спектр задач машинного обучения и системного анализа данных. Сюда входят такие задачи, как автоматический анализ и распознавание голоса, анализ лица пользователя, нахождение различных частей лица и отслеживание их изменений во времени.

Каждая из этих подзадач глобальной задачи ЧМВ сама по себе является крайне сложной задачей, и в вычислительном, и в алгоритмическом плане. Классические математические алгоритмы классификации, как правило, либо вовсе не работают на таких задачах, либо выдают неудовлетворительный результат. Это обусловлено тем, что традиционные алгоритмы классификации



разрабатывались для решения простых задач с малым числом параметров. Некоторые из них предусматривали ограничения, такие как линейная разделимость классов. Однако при переходе от простых задач классификации к сложным, более приближенным к реальности, линейное увеличение числа параметров задачи классификации влечет нелинейное увеличение сложности задачи.

Приведем обзор прикладных задач машинного обучения, решаемых в рамках построения диалоговых систем, а также предложенных методов решения этих задач.

Наиболее важной задачей при построении диалоговой системы является задача распознавания речи (speech recognition). Задачи распознавания речи подразделяются на следующие типы:

1. По зависимости от диктора: дикторозависимые (распознавание речи одного конкретного диктора) и дикторонезависимые (распознавание речи произвольного диктора).
2. По типу речи: распознавание слитной и раздельной речи.
3. По зависимости от словаря: распознавание слов, входящих в словарь, либо распознавание произвольных слов.
4. По типу структурной единицы распознавания: фонема, слово, словосочетание, предложение.
5. По принципу выделения структурных единиц: по шаблону, по лексической и синтаксической схеме.

Скрытые марковские модели (СММ) являются классическим инструментом решения задачи распознавания речи [88, 110], до сих пор используемым на практике. В последние годы для решения этой задачи стали использоваться алгоритмы глубинного обучения, такие как рекуррентные нейронные сети, глубокие нейронные сети с предобучением, конволюционные нейронные сети [61, 82, 87].

Не менее важной является задача распознавания рукописного текста (handwriting recognition). Данная задача делится на 2 типа:

1. Автономное (off-line) распознавание - на вход поступает изображение с текстом, который переводится системой в машиночитаемые символы.
2. Интерактивное (on-line) распознавание - текст распознается в процессе написания пользователем на специальном экране. Этот тип данных известен под названием электронные чернила.

Общий процесс распознавания рукописного текста, как и любую другую задачу машинного обучения, можно разбить на 3 этапа:

1. Предобработка входных данных: цель данного этапа - отсеять шумовые воздействия. Предобработка обычно включает в себя нормализацию данных, сглаживание и удаление шумов.
2. Извлечение признаков: из предобработанных данных извлекаются признаки, релевантные с точки зрения решения задачи. В случае офф-лайн распознавания это могут быть скопления пикселей определенной формы, в случае он-лайн распознавания - скорость или изменение направления электронного пера.
3. Классификация: на основе извлеченных признаков производится классификация рукописных символов с помощью одного из алгоритмов классификации.

Проводились исследования по применению рекуррентных нейронных сетей, скрытых марковских моделей и метода опорных векторов для решения задачи распознавания рукописного текста [39, 68, 78], а также их комбинаций [44]. Существует целое научное сообщество по решению данной задачи, среди крупнейших конференций по данной тематике стоит отметить конференции ICFHR (International conference on frontiers in handwriting recognition) и ICDAR (International conference on document analysis and recognition). За последние десятилетия были сформированы базы данных рукописного текста на английском и других языках для тестирования и сравнения новых подходов с уже

имеющимися, а также проводились многочисленные конкурсы по созданию эффективных систем распознавания текста [98, 99].

С появлением устройств захвата движения появилась возможность распознавания и анализа жестов и движений человека, например движения кисти и пальцев руки. Помимо уже упомянутых алгоритмов, используемых для распознавания речи и рукописного текста [96, 105], для распознавания жестов также используются методы дискриминантного анализа [103], фильтр частиц, конечные автоматы и другие методы [101].

Интерактивный сбор сведений о пользователе прост в реализации, а также является более точным и предпочтительным инструментом добычи вспомогательных сведений для диалоговой системы. Однако данный вид сбора сведений предполагает дополнительную нагрузку на пользователя, которому приходится отвечать на вопросы диалоговой системы касательно его возраста, предпочтительного языка общения и пр. Кроме того, некоторую информацию о пользователе требуется узнавать и обновлять много раз в процессе диалога, например эмоциональное состояние пользователя или степень его отвлеченности от системы. В этих случаях интерактивный сбор сведений неприменим, вместо него необходимо использовать автоматизированную процедуру.

Среди задач автоматического сбора сведений о пользователе можно выделить задачи распознавания следующих признаков:

1. Пол. Задача определения пола по голосу, по изображению, либо их комбинации на сегодняшний день достаточно эффективно решается. Это обусловлено тем, что данная задача достаточно проста, так как имеется всего 2 класса, а также потому что было проведено достаточно большое количество исследований на данную тему [41, 129].

2. Возраст. Данная задача зачастую решается в комбинации с определением пола. Так как возраст у женщин и мужчин проявляется по-разному, обычно сперва проводят классификацию пользователя по полу, а затем относят

его к одному из возрастных классов, или же предсказывают точный возраст [47, 64].

3. Эмоции. Данный вид задач наиболее сложен, так как эмоциональное состояние человека зачастую неявно выражено, и даже человеческий интеллект не всегда справляется с этой задачей. Подробный обзор подходов к решению задачи распознавания эмоций будет дан в § 1.3.

4. Личность. Задача распознавания личности по аудио-визуальным данным крайне важна с практической точки зрения. Например, на секретных объектах производится автоматическое распознавание и учет лиц, имеющих/не имеющих доступ к определенным помещениям. Данная задача отличается от перечисленных выше тем, что для нее требуется заранее заданная база данных пользователей. Для решения данной задачи используются марковские сети [36], алгоритмы дискриминантного анализа, учитывающие не только лицо человека, но и другие его признаки, такие как одежда [93].

Помимо анализа входной информации диалоговая система должна быть способна выводить ответ в удобном виде для восприятия пользователем. Ответ системы может выводиться в текстовом, звуковом, либо визуальном формате. Данные типы вывода информации порождают 3 задачи: генерация естественного текста [114], синтез речи [104], а также сопровождающих речь жестов и мимики (например, при разработке человекоподобных роботов).

Помимо механизмов обработки и вывода информации, центральным компонентом любой диалоговой системы является блок управления диалогом, в который поступает обработанная входная информация от пользователя и генерируется ответ системы. Данный блок тесно взаимодействует с блоком хранения знаний о предметной области.

Проблема человеко-машинного взаимодействия в широком понимании не ограничивается лишь построением интерфейсов между человеком и ЭВМ. Эта область гораздо шире, она лежит на стыке таких дисциплин, как техника, психология, экономика, и связана с понятиями эргономики и удобства

использования (usability). Под термином "машина" в данном контексте может пониматься практически любое устройство, созданное человеком. Однако, в силу исторических причин, именно компьютер занял основополагающее место среди технических средств, с которыми люди взаимодействуют больше всего, как в профессиональной деятельности, так и в повседневной жизни. По этой причине, а также по причине того, что компьютеры обрели массовую доступность, область создания эффективных человеко-компьютерных интерфейсов получила наибольшее развитие. Так как аппаратная составляющая этих интерфейсов (компьютерная мышь, эргономичная клавиатура и прочее) уже создана и не может быть существенно улучшена, следующий этап состоит в создании программного обеспечения для ЭВМ, имитирующего и воспроизводящего тип человеческого мышления. В настоящее время наилучшим инструментом достижения этой цели являются методы машинного обучения и оптимизации.

### **§ 1.3 Постановка задачи распознавания эмоций. Используемые базы данных и подходы на основе машинного обучения**

Задача распознавания эмоций является наиболее сложной из перечисленных в предыдущем параграфе с точки зрения программной реализации на ЭВМ. Это обусловлено несколькими причинами:

1. Эмоции - очень сложный психо-физический процесс, проявляющийся в виде системной комбинации выражения лица и взаимного расположения различных частей тела.
2. Проявление эмоций у разных людей происходит по-разному. Оно зависит от пола, возраста, культурной принадлежности, социального статуса и других особенностей.
3. Зачастую эмоции являются скрытыми, поэтому невозможно судить о внутренних эмоциях человека по внешним физическим проявлениям.

Человеческий мозг справляется с задачей классификации эмоций окружающих людей автоматически, однако для этого у него имеется ряд преимуществ перед ЭВМ:

1. Структура мозга сложнее, чем структура процессора ЭВМ, используемый в мозге механизм распределенных активаций нейронов лучше подходит для решения задач распознавания.
2. Мозг обучается несколько лет перед тем, как становится способен распознавать эмоции.
3. В процессе обучения мозг обрабатывает большое количество изображений лиц людей в различных контекстах.
4. Мозг обладает более широким спектром входной информации от органов чувств.

Несмотря на указанные преимущества, мозг распознает эмоции отнюдь не с 100% точностью, а у некоторых людей данный показатель может быть существенно ниже вплоть до точности случайного угадывания [67]. Цель создания систем автоматического распознавания эмоций заключается в достижении точности распознавания эмоций человеком, поэтому если точность разработанной системы будет превосходить точность случайного угадывания, либо другого тривиального метода (например, постоянное предсказание наиболее часто встречающегося класса эмоций), такой результат можно назвать удовлетворительным.

В зависимости от конечных практических целей задача распознавания эмоций может иметь разные постановки:

1. Два класса эмоций - "злой", "не злой". Практическое применение - автоматизированные колл-центры, системы слежения за соблюдением порядка.
2. Определение степени возбужденности пользователя. Практическое применение - определение сонного состояния водителя автомобиля с целью предотвращения дорожно-транспортного происшествия.

3. Два класса эмоций - "нейтральная", "не нейтральная". Данная система может использоваться как вспомогательная для выявления факта проявления любой эмоции.

4. Семь классов основных эмоций - "нейтральная", "счастье", "злость", "грусть", "удивление", "страх", "отвращение". Данный тип задачи может использоваться в диалоговых системах общего пользования, например в смартфонах, планшетах, персональных компьютерах, а также в автоматизированных системах сбора мнений.

Для распознавания эмоций ЭВМ может анализировать аудио или видео информацию, такую как голос пользователя, изображение его лица и всей фигуры в целом, а также комбинацию аудио-видео информации (multimodal information). Методы слияния мультимодальной аудио-видео информации представляют собой предмет исследования [54, 56, 127, 134, 144, 145].

Предложенные различными исследователями подходы к автоматическому распознаванию эмоций можно разделить на подгруппы по следующим категориям:

1. Используемые входные признаки: интенсивности пикселей входного изображения; координаты движения основных частей лица - глаз, носа, рта, лба, щек [55, 144]; данные электроэнцефалограммы, реакция зрачков, дальность взгляда [134]; движение тела, жесты [56]; параметры голоса [127].

2. Используемые методы снижения размерности: метод главных компонент, линейный дискриминантный анализ [81], дисперсионный анализ [134],

3. Используемые алгоритмы классификации и распознавания: метод опорных векторов [134], Байесовский классификатор [54, 56, 81, 127], алгоритм k ближайших соседей [54], скрытые марковские модели [62, 126], нейронные сети [89], алгоритм адаптивного бустинга AdaBoost [48].

Для тестирования и сравнения эффективности различных подходов к распознаванию эмоций был создан ряд баз данных, в которых собраны аудио- и

видеозаписи людей, изображающих определенные эмоции, либо запечатлены естественные эмоции. Информация по наиболее известным базам данных представлена в таблице 1.1.

Таблица 1.1. Базы данных эмоций; А - аудио, В - видео, АВ - аудио-видео; И - наигранные эмоции, Е - естественные эмоции; ЭЭГ - данные электроэнцефалограммы, ЭД - данные электродермальной активности; 6 основных эмоций - злость, отвращение, страх, счастье, грусть, удивление

| База данных<br>(год создания)                      | Аудио<br>/ Видео<br>(А, В,<br>АВ) | Тип<br>записываемых<br>эмоций (И, Е) | Содержание<br>базы данных                           | Классы эмоций                                                                                                                   |
|----------------------------------------------------|-----------------------------------|--------------------------------------|-----------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| DES - Danish<br>Emotional<br>Speech (1996)<br>[69] | А                                 | И                                    | 10 минут, 2 актера<br>и 2 актрисы; 9<br>предложений | Злость, счастье, грусть,<br>удивление, спокойствие                                                                              |
| Chen-Huang<br>(2000) [60]                          | АВ                                | И                                    | 100 людей, 9900<br>выражений                        | 6 основных эмоций, скука,<br>разочарование,<br>заинтересованность,<br>замешательство                                            |
| AIBO (2004)<br>[40]                                | А                                 | Е                                    | 110 диалогов,<br>29200 слов                         | Злость, скука,<br>настойчивость,<br>беспомощность, ирония,<br>радость, забота, выговор,<br>расслабленность,<br>удивление, обида |
| Adult<br>Attachment<br>Interview<br>(2004) [118]   | АВ                                | Е                                    | 60 людей, каждое<br>интервью 30-60<br>минут         | 6 основных эмоций,<br>презрение, смущение, стыд                                                                                 |
| Berlin (2005)<br>[53]                              | А                                 | И                                    | 493 предложения,<br>5 актеров и 5<br>актрис         | Злость, скука, отвращение,<br>страх, счастье, грусть,<br>спокойствие                                                            |



Продолжение таблицы 1.1.

| <b>База данных<br/>(год создания)</b>                                                | <b>Аудио<br/>/ Видео<br/>(А, В,<br/>АВ)</b> | <b>Тип<br/>записываемых<br/>эмоций (И, Е)</b> | <b>Содержание<br/>базы данных</b>                       | <b>Классы эмоций</b>                                                                                  |
|--------------------------------------------------------------------------------------|---------------------------------------------|-----------------------------------------------|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| UT Dallas<br>(2005) [107]                                                            | В                                           | Е                                             | 229 людей                                               | 6 основных эмоций, скука,<br>недоверие, веселье,<br>замешательство                                    |
| BU-3DFE<br>(2006) [142]                                                              | В                                           | И                                             | 100 людей                                               | 6 основных эмоций с 4<br>уровнями интенсивности                                                       |
| FABO face and<br>body gesture<br>(2006) [79]                                         | В                                           | И                                             | 23 человека, 210<br>видео                               | 6 основных эмоций,<br>тревога, скука,<br>спокойствие,<br>неуверенность                                |
| Busso-<br>Narayanan<br>(2007) [55]                                                   | АВ                                          | И                                             | 612 предложений,<br>1 актриса                           | Злость, счастье, грусть,<br>спокойствие                                                               |
| SAVEE -<br>Surrey Audio-<br>Visual<br>Expressed<br>Emotions<br>(2009) [81]           | АВ                                          | И                                             | 480 предложений,<br>4 диктора                           | 6 основных эмоций,<br>спокойствие                                                                     |
| RECOLA -<br>Remote<br>Collaborative<br>and Affective<br>Interactions<br>(2013) [115] | АВ +<br>ЭЭГ,<br>ЭД                          | Е                                             | 46 дикторов: 27<br>женщин, 19<br>мужчин; 9.5 ч<br>видео | Степень возбуждения,<br>степень психологической<br>валентности<br>(привлекательность /<br>отвращение) |

Несмотря на то, что на некоторых из представленных в таблице 1.1 данных была достигнута высокая экспериментальная точность распознавания эмоций, пока не было предложено универсального эффективного подхода для решения

этой задачи в условиях, более приближенных к реальным. В реальных условиях при распознавании эмоций по видеоряду как правило не имеется данных электроэнцефалограммы человека, точек-маркеров на лице, а лицо зачастую вообще не видно, если человек стоит к камере спиной. В связи с этим актуальной задачей является создание дополнительных баз данных эмоций, содержащих видео, записанное в реальных условиях, а также создание базы данных, являющейся стандартом для унифицированного сравнения различных подходов к распознаванию эмоций.

Среди представленных баз данных для дальнейшего использования в рамках данной диссертации была выбрана БД SAVEE. Ее выбор обусловлен тем, что она включает в себя видеозаписи эмоций обычных людей, а не профессиональных актеров, а следовательно более приближена к реальности. Кроме того, в ней равномерно представлены 7 основных классов эмоций. Эта база данных включает в себя 4 дикторов. В следующих главах, при упоминании задачи распознавания эмоций, подразумевается ее решение с использованием БД SAVEE.

## **Выводы**

На сегодняшний день проблема человеко-машинного взаимодействия в основном сосредоточена на создании программных систем для компьютеров, способствующих более эффективному взаимодействию человека и компьютера. Подзадачами человеко-машинного взаимодействия являются такие задачи, как автоматизированный сбор сведений о пользователе (пол, возраст, эмоциональное состояние, национальность), распознавание рукописного текста, естественной человеческой речи, а также генерация соответствующего ответа системы в удобном для человеческого восприятия формате. В частности, задача автоматического распознавания эмоций человека интересна с точки зрения построения эффективных диалоговых систем.

Каждая из перечисленных подзадач человеко-машинного взаимодействия крайне сложна в вычислительном и концептуальном плане. Наиболее развитый подход в решении такого рода задач называется машинным обучением. Он основан на использовании данных, на основе которых ЭВМ в соответствии с некоторым алгоритмом обучается решать поставленную задачу. Для различных задач человеко-машинного взаимодействия, и в частности для задачи распознавания эмоций, было собрано множество баз данных для создания систем машинного обучения.

С точки зрения системного анализа решение задач человеко-машинного взаимодействия с помощью методов машинного обучения и оптимизации оправдано, так как алгоритмы обучения в большинстве своем основаны на разбиении исходной задачи на более мелкие и простые задачи. Кроме того, многие модели машинного обучения, например нейронные сети и деревья решений, имеют распределенную структуру, отражающую внутренние закономерности в данных.

Задачи оптимизации являются неотъемлемой частью задач распознавания (оптимизация структуры и параметров модели машинного обучения, оптимальный выбор примеров из обучающей выборки), поэтому методы оптимизации являются вспомогательным средством при решении задач человеко-машинного взаимодействия.

В следующей главе будет рассмотрен класс эволюционных алгоритмов однокритериальной и многокритериальной оптимизации, а также описан разработанный коэволюционный алгоритм многокритериальной оптимизации.

## Глава 2. Разработка коллективного самоконфигурируемого эволюционного алгоритма многокритериальной оптимизации

### § 2.1 Эволюционные алгоритмы однокритериальной и многокритериальной оптимизации

Эволюционные алгоритмы - класс стохастических алгоритмов оптимизации, симулирующих процесс естественной эволюции. Несмотря на простой принцип работы, эти алгоритмы зарекомендовали себя как робастные и мощные методы поиска общего назначения [100]. Эволюционные алгоритмы подразделяются по типу решаемых задач оптимизации на однокритериальные и многокритериальные.

Ниже представлена схема стандартного генетического алгоритма (ГА) однокритериальной оптимизации. Предполагается, что задана некоторая задача в виде оптимизируемой целевой функции (ЦФ), и преобразование решений-кандидатов в двоичную строку.

#### Стандартный генетический алгоритм

##### *Параметры:*

1.  $t$  - количество итераций (поколений) алгоритма;
2.  $n$  - размер популяции;
3.  $m$  - количество бит в хромосоме, может быть задано непосредственно, либо вычислено, исходя из желаемой точности кодирования решения;
4.  $p_c$  - вероятность скрещивания;
5.  $p_m$  - вероятность мутации.

**Вход:** случайно сгенерированная популяция  $n$  решений, каждое решение - хромосома из  $m$  бит.

**Выход:** хромосома, соответствующая наилучшему найденному решению, и соответствующее ей значение ЦФ.

**Основной цикл** (выполняется  $t$  раз):

1. Вычислить пригодность  $f(x)$  каждой хромосомы  $x$  в текущей популяции; пригодность может совпадать с целевой функцией в случае задачи максимизации, либо быть обратной ей - в случае задачи минимизации.

2. Повторять следующие операции, пока не будет сгенерировано  $n$  решений-потомков:

а. Выбрать пару хромосом-родителей из текущей популяции таким образом, что вероятность выбора определенной хромосомы тем выше, чем выше пригодность этой хромосомы. Селекция проводится "с замещением", т.е. одна и та же хромосома может быть выбрана несколько раз.

б. С вероятностью  $p_c$  скрестить выбранную пару хромосом в случайно выбранной точке, чтобы сформировать 2 хромосомы-потомка. Если скрещивание не производится, то хромосомы-потомки являются копиями хромосом-родителей.

в. С вероятностью  $p_m$  произвести мутацию хромосом-потомков (случайное изменение битов), поместить получившиеся хромосомы в новую популяцию.

3. Заменить текущую популяцию новой.

Помимо операторов стандартного ГА существуют также другие:

1. Несколько типов селекции хромосом-родителей:

а. Пропорциональная - вероятность выбора решения-индивида пропорциональна значению его пригодности (рисунок 2.1);

б. Ранговая - индивидам популяции присваиваются ранги на основе значений их пригодности: индивид с максимальной пригодностью по популяции получает ранг 1, следующий по пригодности - ранг 2, и т.д. Индивидам с одинаковой пригодностью присваивается среднеарифметический ранг этих индивидов в отсортированной по пригодности популяции. Далее выбор индивидов проводится по аналогии с пропорциональной селекцией, но вместо значений пригодности используются ранги индивидов (рисунок 2.2);

в. Турнирная - из популяции выбираются  $T$  индивидов, где  $T$  - размер турнира, являющийся параметром данного оператора селекции. Из данной группы выбирается индивид с наибольшим значением функции пригодности (рисунок 2.3).

2. Несколько типов скрещивания хромосом-родителей:

а. Одноточечное - в случайном месте хромосом-родителей производится разрыв между позициями битов, хромосомы обмениваются частями, в результате образуются две новые хромосомы (рисунок 2.4);

б. Двухточечное - операция аналогична одноточечному скрещиванию, но разрыв между битами хромосом-родителей производится в двух случайных местах (рисунок 2.5);

в. Равномерное - хромосома-потомок состоит из случайно выбранных битов одного из родителей на соответствующих позициях (рисунок 2.6).

3. Оператор элитизма, гарантированно переводящий решение с лучшей пригодностью в следующее поколение



Рисунок 2.1. Схема пропорциональной селекции; пригодности индивидов нормируются для получения вероятностей выбора, генерируется случайное число из интервала  $[0, 1]$ , выбирается соответствующий индивид

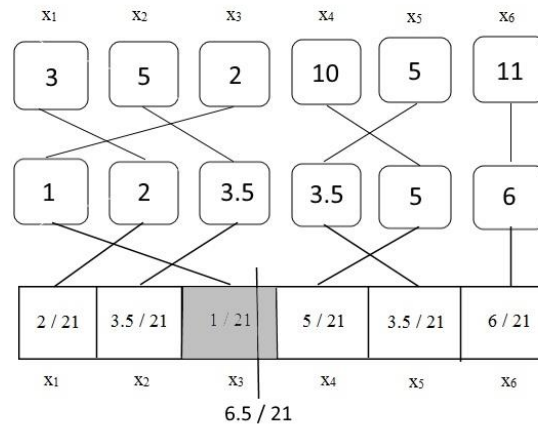


Рисунок 2.2. Схема ранговой селекции; индивиды популяции ранжируются по пригодности, каждому индивиду присваивается ранг; ранги нормируются в интервал  $[0, 1]$ , генерируется случайное число из этого интервала, выбирается соответствующий индивид

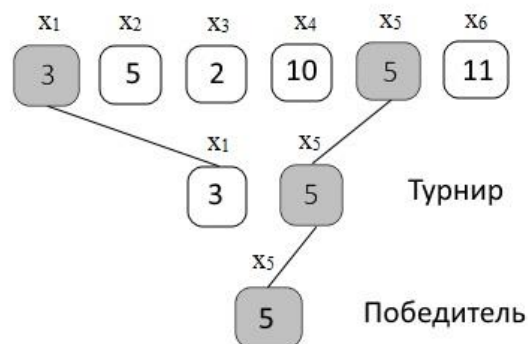


Рисунок 2.3. Схема турнирной селекции; размер турнира - 2

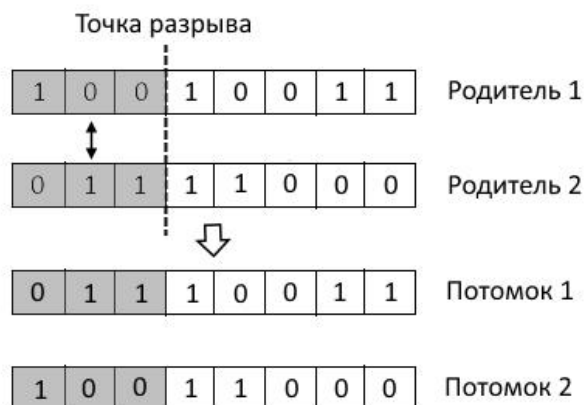


Рисунок 2.4. Схема одноточечного скрещивания

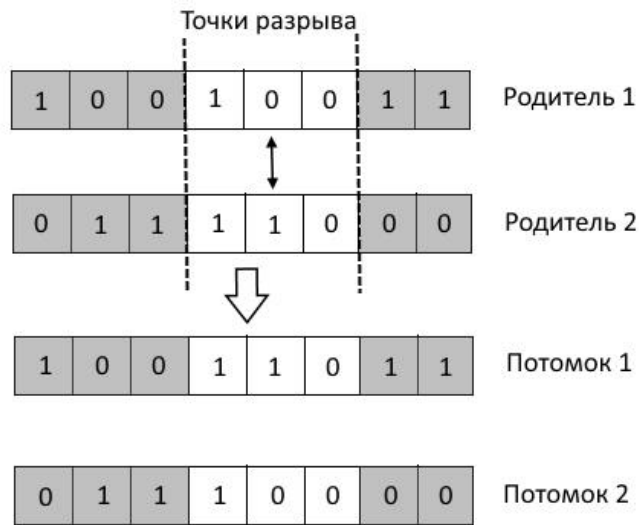


Рисунок 2.5. Схема двухточечного скрещивания

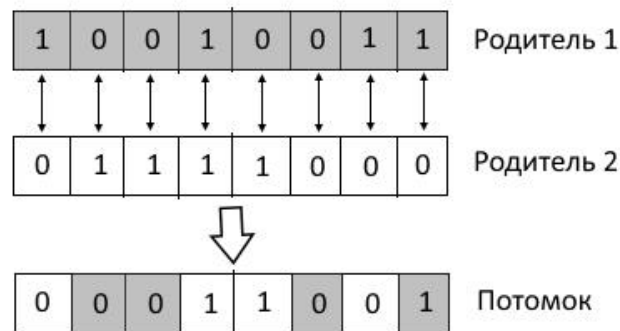


Рисунок 2.6. Равномерное скрещивание; каждый бит потомка случайно берется у одного из родителей на соответствующих позициях

Было изобретено множество операторов и модификаций стандартного ГА, и проводились исследования их эффективности [11, 19, 21, 25]. При таком обилии параметров и операторов возникает проблема настройки ГА под конкретную задачу оптимизации. Тонкая настройка параметров ГА крайне важна для успешного решения сложных задач, однако на практике сделать это очень трудно, так как заранее не ясно, какой набор параметров окажется эффективным для решения конкретной задачи оптимизации.



Для решения этой проблемы развивается направление самоконфигурируемых эволюционных алгоритмов [22, 32]. В частности, был создан коэволюционный генетический алгоритм, являющийся самонастраивающейся модификацией стандартного ГА. Базовый принцип, лежащий в основе коэволюционного ГА, основан на островной модели, в которой несколько популяций решений развиваются, не сообщаясь между собой. В определенные моменты (на определенных итерациях) происходит обмен наилучшими найденными решениями между популяциями (миграция решений). Смысл состоит в том, что каждая популяция развивается по различному сценарию. Таким образом, коэволюционный ГА состоит из нескольких популяций, каждая из которых оптимизируется стандартным ГА с различными параметрами. Этот подход уменьшает влияние выбора параметров стандартного ГА на исход решаемой задачи, а также позволяет нескольким ГА за счет коллективных усилий находить решения, которые отдельно взятый стандартный ГА найти не может.

Эволюционные алгоритмы также хорошо подходят для решения задач многокритериальной оптимизации. Благодаря внутреннему параллелизму, связанному с одновременной эволюцией популяции решений, они позволяют найти множество Парето-оптимальных решений за один запуск алгоритма [8, 10, 20]. Так как в сложных задачах многокритериальной оптимизации зачастую невозможно точно найти множество Парето-оптимальных решений, цель решения задачи может быть сформулирована в виде трех критериев:

1. Минимизация расстояния от найденного алгоритмом недоминируемого фронта решений (в критериальном пространстве) до истинного Парето-оптимального фронта;
2. Равномерное распределение решений по фронту более предпочтительно;
3. Значения критериев найденных решений должны охватывать как можно больший диапазон значений.

Исходя из этих трех критериев возникают два вопроса, которые необходимо решать при применении эволюционных алгоритмов:

1. Как рассчитывать пригодность индивидов и проводить селекцию, чтобы вести поиск в направлении Парето-оптимального множества?
2. Как обеспечивать разнообразие популяции в процессе поиска, чтобы избежать преждевременной сходимости алгоритма?

В отличие от задач однокритериальной оптимизации, где целевая функция и функция пригодности зачастую совпадают, в задачах многокритериальной оптимизации данные функции не совпадают, и расчет пригодностей и селекция проводятся с учетом нескольких критериев. Различают три подхода к селекции, используемые в многокритериальных эволюционных алгоритмах: селекция по отдельным критериям, агрегированная селекция, и прямое использование концепции Парето доминирования.

В алгоритмах первого подхода селекция индивидов проводится по отдельно взятым критериям, доля индивидов в общей популяции, выбранных по каждому критерию, одинакова. Допустим, решается задача оптимизации с пятью критериями, размер популяции 100. Следовательно, будут выбраны 20 индивидов, оптимальных по первому критерию, 20 индивидов, оптимальных по второму критерию, и т.д.

В алгоритмах с *агрегированной селекцией* используется классическая техника агрегирования критериев в параметризованную целевую функцию. Параметры изменяются в процессе работы алгоритма, и могут быть либо закодированы в хромосомы индивида, либо выбраны случайным образом. Это позволяет вести оптимизацию одновременно в различных направлениях.

В алгоритмах, *напрямую использующих концепцию Парето-доминирования*, каждому индивиду популяции присваивается ранг. В популяции находят недоминируемые решения, им присваивается ранг 1, они временно удаляются из популяции. Среди оставшихся индивидов также находят недоминируемых, им присваивается ранг 2. Процедура продолжается до тех пор, пока всем индивидам

не будет присвоен ранг, который и используется в качестве пригодности при проведении селекции.

Одна из проблем эволюционных алгоритмов заключается в преждевременной сходимости, когда один наилучший индивид (зачастую неоптимальный) заполняет собой всю популяцию, приводя поиск в тупик. Чтобы избежать этой проблемы, необходимо искусственно поддерживать разнообразие популяции. Ниже представлены механизмы поддержания разнообразия, наиболее часто используемые в эволюционных алгоритмах.

*Разделение пригодностей*, также именуемое *механизмом ниши*, основано на идее штрафов для индивидов, находящихся близко друг другу в пространстве критериев (пространстве решений). Вводится параметр *радиус ниши*  $\sigma$ , определяющий пороговое значение близости индивидов. Индивиды, расстояние между которыми меньше порогового, считаются соседними. Значение пригодности каждого индивида пересчитывается по формуле:

$$F(x_i) = \frac{F'(x_i)}{\sum_{x_j \in P} s(d(x_i, x_j))}, \quad (2.1)$$

$$s(d(x_i, x_j)) = \begin{cases} 1 - (\frac{d(x_i, x_j)}{\sigma})^\alpha, & \text{если } d(x_i, x_j) < \sigma \\ 0, & \text{иначе} \end{cases} \quad (2.2)$$

Здесь  $F'(x_i)$  - вычисленная на предыдущем этапе пригодность индивида  $x_i$ ,  $P$  - популяция индивидов,  $d(x_i, x_j)$  - расстояние между индивидами  $x_i$  и  $x_j$ , вычисленное по определенной метрике,  $s(d(\cdot))$  - функция разделения пригодностей.

При *ограниченном скрещивании* двум индивидам позволено скрещиваться только если они находятся на определенном расстоянии друг от друга (определяется параметром  $\sigma_{mate}$ ).

В некоторых алгоритмах используется кластеризация индивидов популяции, кластеры близко лежащих индивидов заменяются одним индивидом, соответствующим центру кластера. Данный механизм признан одним из самых

эффективных для поддержания разнообразия популяции и равномерности покрытия фронта Парето.

Одним из первых разработанных эволюционных алгоритмов многокритериальной оптимизации был алгоритм *Vector Evaluated Genetic Algorithm* (VEGA) [124], предложенный Шаффером. В данном алгоритме используется селекция по отдельно взятым критериям (рисунок 2.7). Доля индивидов, выбранных по каждому из критериев, одинакова.

### **Алгоритм: Расчет пригодностей и селекция в алгоритме VEGA**

**Вход:**

1.  $P_t$  - популяция индивидов на итерации  $t$ .
2.  $N$  - размер популяции.
3.  $k$  - число критериев задачи оптимизации.

**Выход:**  $P'$  - популяция индивидов для скрещивания.

1. Инициализировать популяцию для скрещивания  $P' = \emptyset$ .
2. Цикл по критериям  $i = \overline{1, k}$ :

а. Для каждого индивида  $x_j \in P_t$  рассчитать пригодность  $F(x_j) = f_i(x_j)$ .

б. Выбрать  $N/k$  индивидов, лучших по пригодности и поместить их в популяцию для скрещивания  $P'$ .

Здесь  $f_i(x)$  - функция  $i$ -го критерия. После селекции над популяцией  $P'$  применяются стандартные операторы скрещивания и мутации.

Шринивас и Деб разработали алгоритм *Nondominated Sorting Genetic Algorithm* (NSGA) [63, 135], в котором напрямую использована концепция Парето-доминирования для расчета пригодностей, и разделение пригодностей для поддержания разнообразия популяции (рисунок 2.8).

### **Алгоритм: Расчет пригодностей в NSGA**

**Вход:**

1.  $P_t$  - популяция индивидов на итерации  $t$ .
2.  $\sigma$  - радиус ниши.

3.  $N$  - размер популяции.

**Выход:**  $F$  - значения пригодностей.

1. Инициализировать популяцию оставшихся индивидов  $P_{\text{ост}} = P_t$ , инициализировать значения пригодностей  $F_{\text{нач}}(x_j) = N$  значениями по умолчанию,  $x_j \in P_t$ .

2. Пока  $P_{\text{ост}} \neq \emptyset$ :

а. Найти подмножество недоминируемых решений  $P_{\text{недом}}$  в  $P_{\text{ост}}$ .

б. Индивидам  $x_j \in P_{\text{недом}}$  присвоить пригодность  $F(x_j) = F_{\text{нач}}(x_j)$ , проделать процедуру разделения пригодностей в пространстве решений, только для  $P_{\text{недом}}$ .

в. Уменьшить значения пригодностей по умолчанию  $F_{\text{нач}}$  таким образом, чтобы они были меньше чем наименьшая пригодность в  $P_{\text{недом}}$ :  
 $0 < F_{\text{нач}} < \min\{F(x_j) \mid x_j \in P_{\text{недом}}\}$ .

г.  $P_{\text{ост}} = P_{\text{ост}} - P_{\text{недом}}$ .

Здесь большие значения пригодности соответствуют большей вероятности воспроизводства при селекции.

Еще один эффективный алгоритм многокритериальной оптимизации под названием Strength Pareto Evolutionary Algorithm (SPEA) был предложен Цицлером [150, 151]. В этом алгоритме найденные на текущей итерации недоминируемые решения хранятся во внешнем множестве. Пригодности индивидов вычисляются, исходя из концепции Парето доминирования (рисунок 2.9). Для уменьшения хранимых во внешнем множестве решений используется процедура кластеризации.

### **Алгоритм: расчет пригодностей в SPEA**

**Вход:**

$P_t$  - популяция индивидов на итерации  $t$ .

$\overline{P}_t$  - внешнее множество индивидов на итерации  $t$ .

$N$  - размер популяции.

**Выход:**  $F$  - значения пригодностей индивидов популяции и внешнего множества.

1. Для каждого индивида  $x_j \in \overline{P}_t$  рассчитывается значение  $S(x_j) \in [0,1)$ , называемое силой; сила индивида внешнего множества  $S(x_j)$  пропорциональна количеству индивидов популяции  $x_i \in P_t$ , которых он доминирует  $x_j \succ x_i$ ;  $x_i \in P_t$ :

$$S(x_j) = \frac{|\{x_i | x_i \in P_t \wedge x_j \succ x_i\}|}{N + 1}. \quad (2.3)$$

Пригодность индивида  $x_j$  из внешнего множества равна его силе  $F(x_j) = S(x_j)$ .

2. Пригодность индивида популяции  $x_i \in P_t$  вычисляется путем суммирования сил всех индивидов внешнего множества, которые доминируют  $x_i$ .

$$F(x_i) = 1 + \sum_{x_j \in \overline{P}_t, x_j \succ x_i} S(x_j); F(x_i) \in [1, N). \quad (2.4)$$

Один добавляется для того, чтобы гарантировать, что пригодность любого индивида внешнего множества меньше пригодности любого индивида популяции (в данном случае меньшие значения пригодности соответствуют большей вероятности воспроизводства).

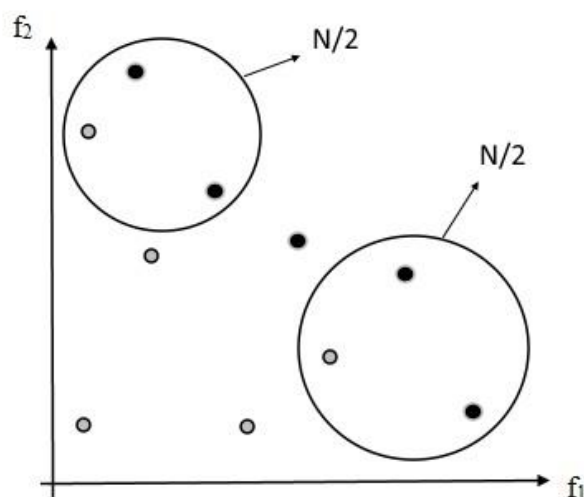


Рисунок 2.7. Схема селекции по отдельным критериям, используемая в алгоритме VEGA; доля индивидов, отбираемых по каждому критерию, одинакова

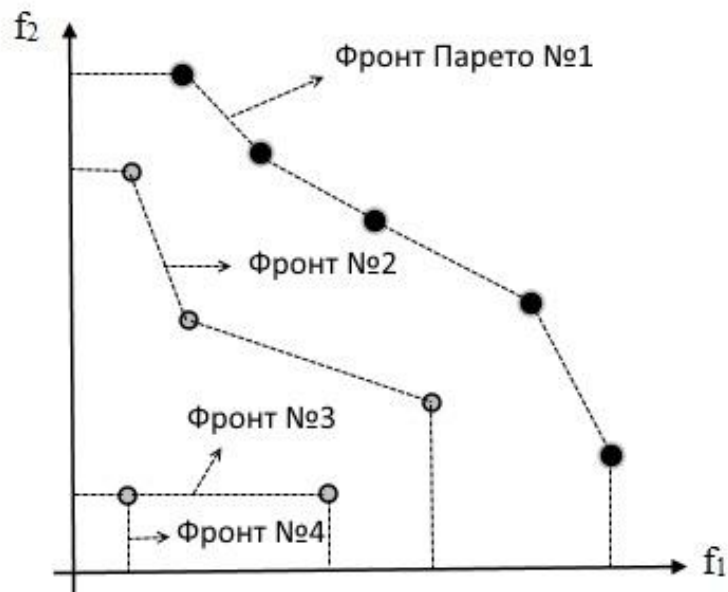


Рисунок 2.8. Схема расчета пригодностей индивидов в алгоритме NSGA, основанного на недоминируемой сортировке



Рисунок 2.9. Схема расчета пригодностей индивидов в алгоритме SPEA; черные точки - элементы внешнего множества, серые точки - элементы популяции

## **§ 2.2 Разработка и реализация самоконфигурируемого эволюционного алгоритма многокритериальной оптимизации**

Каждый из вышеописанных эволюционных алгоритмов многокритериальной оптимизации обладает своими преимуществами и недостатками. В частности, алгоритм VEGA обладает лучшей сходимостью, но не имеет механизма обеспечения равномерности покрытия фронта Парето. Алгоритм NSGA, его модифицированная версия NSGA-2, а также алгоритм SPEA обеспечивают хорошее покрытие фронта Парето, но требуют больше вычислительных затрат, а механизмы поддержания разнообразия решений зачастую выносят решения за пределы множества Парето.

На базе алгоритмов VEGA, SPEA и NSGA-2 автором данной диссертации был создан коэволюционный генетический алгоритм многокритериальной оптимизации SelfCOMOGA [23], в котором данные алгоритмы объединяются в коллектив. Данный алгоритм является развитием метода селективных гипер-эвристик с онлайн обучением, предложенных Соповым Е.А. [Sopov, E.A. Self-configuring Metaheuristic for Control of Multi-Strategy Evolutionary Search // Advances in Swarm and Computational Intelligence, Lecture Notes in Computer Science, Volume 9142, 2015, pp. 29-37].

В диссертации Сергиенко Р. был предложен коэволюционный алгоритм многокритериальной оптимизации, в котором в качестве коэволюционирующих компонент использовались алгоритмы VEGA, FFGA, NPGA, NSGA, SPEA и SPEA-2.

В описываемом далее алгоритме количество коэволюционирующих компонент было сокращено до трех алгоритмов - SPEA, VEGA и NSGA-2. Данные алгоритмы были выбраны ввиду того, что в них используются различные подходы к селекции индивидов популяции, что в процессе многокритериального поиска позволяет избежать стагнации и обеспечить более равномерное покрытие множества недоминируемых решений. Отличительной особенностью



предлагаемого коэволюционного алгоритма также является этап оценки эффективности работы коэволюционирующих алгоритмов.

Общая структура алгоритма SelfCOMOGA схожа со структурой стандартного генетического алгоритма, описанной в §2.1.

На рисунке 2.10 представлена схема алгоритма SelfCOMOGA:

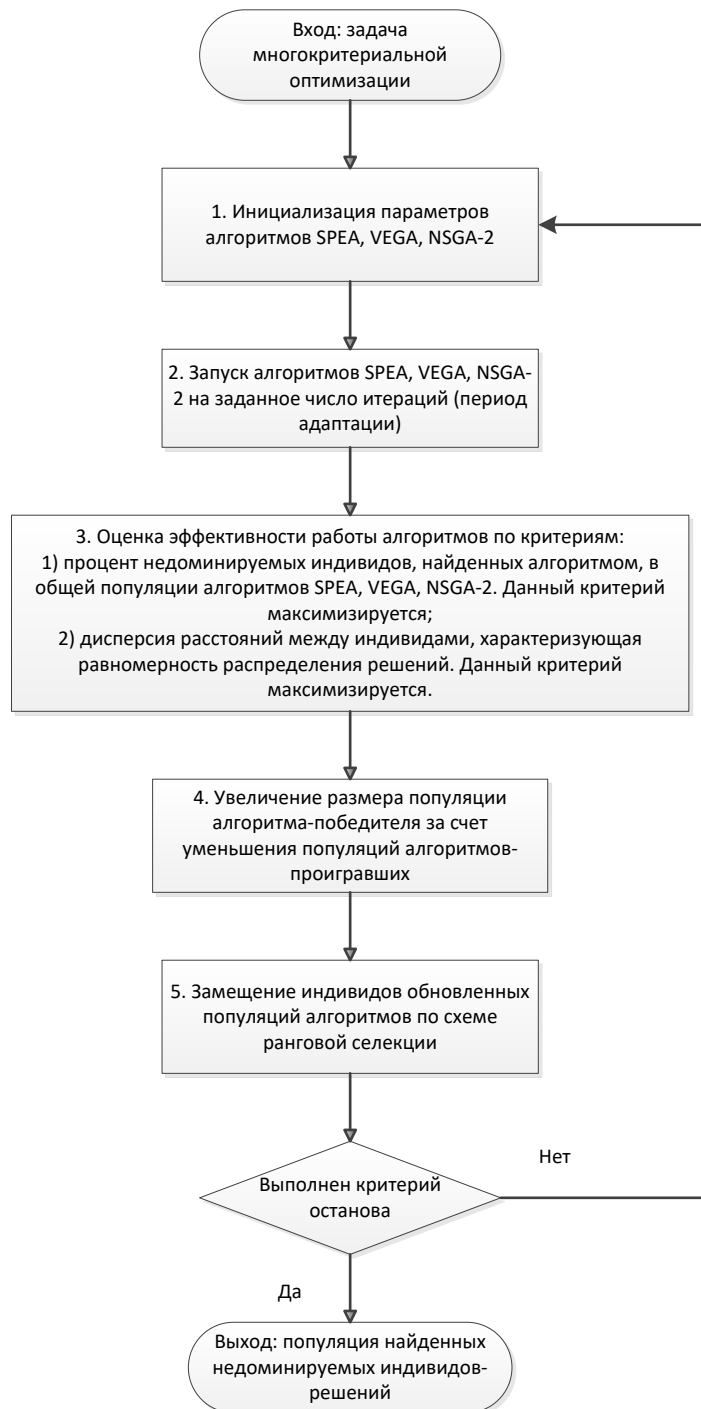


Рисунок 2.10. Схема алгоритма SelfCOMOGA

Алгоритм SelfCOMOGA включает в себя следующие основные этапы:

1. Инициализация параметров алгоритмов SPEA, VEGA, NSGA-2, участвующих в коэволюции.
2. Параллельная работа алгоритмов SPEA, VEGA, NSGA-2 заданное число итераций (период адаптации).
3. Оценка эффективности работы алгоритмов по критериям (процент недоминируемых индивидов в общей популяции коэволюционирующих алгоритмов, равномерность распределения решений в пространстве критериев).
4. Изменение размеров популяций коэволюционирующих алгоритмов (увеличение размера популяции алгоритма-победителя за счет уменьшения популяций алгоритмов-проигравших).
5. Замещение индивидов в популяциях коэволюционирующих алгоритмов по схеме ранговой селекции, переход к этапу 2.

Рассмотрим этапы подробнее.

**На первом этапе** создаются экземпляры каждого алгоритма - VEGA, NSGA-2 и SPEA. Существуют 3 возможные стратегии инициализации параметров каждого из алгоритмов:

1. Детерминированный выбор алгоритмов, свойства которых известны и могут быть эффективно использованы для решения поставленной задачи.
2. Включение всех возможных комбинаций параметров алгоритмов.
3. Случайный выбор некоторого числа алгоритмов из множества всех возможных комбинаций алгоритмов и их параметров.

Первый вариант требует привлечения знаний об алгоритмах и решаемой задаче, что неприменимо для сложных задач, и противоречит идее самонастройки ГА. Второй подход – вычислительно затратный, так как число комбинаций может быть большим (сотни вариантов). В данной работе использована последняя стратегия: создается один экземпляр каждого алгоритма - VEGA, NSGA-2, SPEA.

Параметры алгоритмов выбираются случайно из множества заранее заданных вариантов:

1. Тип скрещивания: одноточечное, двухточечное, равномерное.
2. Вероятность мутации: высокая, средняя, низкая.
3. Вероятность скрещивания: высокая, средняя, низкая.
4. Максимальный размер множества Парето (для SPEA): 30, 50, 70.

**На втором этапе** проинициализированные алгоритмы SPEA, VEGA, NSGA-2 параллельно отрабатывают заданное число итераций. Данное число итераций именуется периодом адаптации и является одним из параметров коэволюционного алгоритма SelfCOMOGA.

**На третьем этапе** производится оценка эффективности коэволюционирующих алгоритмов по некоторым критериям качества. Поскольку на каждой итерации алгоритма имеется не одно наилучшее решение, а некоторая подпопуляция недоминируемых решений, нет возможности прямого сравнения алгоритмов и известные подходы, применимые для однокритериальных ГА, не могут быть использованы. В различных работах предложены следующие основные критерии сравнения: близость найденного множества решений к истинному множеству Парето и равномерность распределения найденных решений. Очевидно, что для практических задач первый критерий непригоден, так как истинное множество Парето неизвестно, его требуется найти. В данной работе для оценки индивидуальных алгоритмов в SelfCOMOGA используется взвешенная сумма следующих критериев (веса критериев приняты равными 0.7 и 0.3 соответственно):

1.  $K_1$  – процент недоминируемых решений. Создается общий пул решений коэволюционирующих алгоритмов, который подвергается недоминируемой сортировке. Далее из отсортированного массива выбирается множество наилучших решений в количестве, равном размеру популяции, и вычисляется доля решений каждого из коэволюционирующих алгоритмов, попавших в данное множество.

2.  $K_2$  – равномерность распределения недоминируемых решений коэволюционирующего алгоритма. Равномерность вычисляется как дисперсия расстояний между недоминируемыми решениями в пространстве критериев.

**На четвертом этапе** производится перераспределение ресурсов между алгоритмами. Размер популяции алгоритма-победителя, выявленного на основании взвешенной суммы вышеописанных критериев, увеличивается за счет уменьшения размеров популяций всех остальных алгоритмов. Размер популяции алгоритма-проигравшего уменьшается на некоторый процент от текущего размера популяции - размер штрафа, который является параметром коэволюционного алгоритма. Размер популяции алгоритма не может быть уменьшен ниже минимального гарантированного размера популяции, который также является параметром коэволюционного алгоритма.

**На пятом этапе** обновленные популяции алгоритмов заполняются решениями из общего пула, отсортированного с помощью недоминируемой сортировки, по схеме ранговой селекции.

Критерием останова коэволюционного алгоритма многокритериальной оптимизации является заранее заданное число итераций алгоритма.

### **§ 2.3 Исследование эффективности самоконфигурируемого алгоритма на репрезентативном наборе тестовых задач оптимизации**

Для оценки эффективности разработанного самоконфигурируемого коэволюционного алгоритма многокритериальной оптимизации были использованы 10 тестовых сложных задач многокритериальной оптимизации, предложенные в рамках конкурса CEC, приуроченного к конференции CEC (Congress on evolutionary computation) [146]. CEC - одна из наиболее авторитетных конференций по эволюционным алгоритмам. На рассматриваемом наборе тестовых задач многокритериальной оптимизации CEC тестировалось большое количество современных эволюционных алгоритмов оптимизации,

следовательно, на данных задачах может быть проведено объективное сравнение разработанного алгоритма SelfCOMOGA как с входящими в него алгоритмами-компонентами SPEA, VEGA, NSGA-2, так и с другими современными алгоритмами оптимизации.

1. Тестовая задача №1:

$$\begin{cases} f_1 = x_1 + \frac{2}{|J_1|} \sum_{j \in J_1} (x_j - \sin(6\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \\ f_2 = 1 - \sqrt{x_1} + \frac{2}{|J_2|} \sum_{j \in J_2} (x_j - \sin(6\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \end{cases} \quad (2.5)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ .

Область определения:  $\bar{x} \in [0,1] \cdot [-1,1]^{n-1}$ .

Множество Парето:  $0 \leq x_1 \leq 1$ ,  $x_j = \sin(6\pi x_1 + \frac{j\pi}{n})$ ,  $j = \overline{2, n}$ .

Фронт Парето:  $f_2 = 1 - \sqrt{f_1}$ ,  $0 \leq f_1 \leq 1$ .

2. Тестовая задача №2:

$$\begin{cases} f_1 = x_1 + \frac{2}{|J_1|} \sum_{j \in J_1} y_j^2 \rightarrow \min, \\ f_2 = 1 - \sqrt{x_1} + \frac{2}{|J_2|} \sum_{j \in J_2} y_j^2 \rightarrow \min, \end{cases} \quad (2.6)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ .

$$y_j = \begin{cases} x_j - \left(0.3x_1^2 \cdot \cos\left(24\pi x_1 + \frac{4j\pi}{n}\right) + 0.6x_1\right) \cdot \cos\left(6\pi x_1 + \frac{j\pi}{n}\right), j \in J_1, \\ x_j - \left(0.3x_1^2 \cdot \cos\left(24\pi x_1 + \frac{4j\pi}{n}\right) + 0.6x_1\right) \cdot \sin\left(6\pi x_1 + \frac{j\pi}{n}\right), j \in J_2. \end{cases} \quad (2.7)$$

Область определения:  $\bar{x} \in [0,1] \cdot [-1,1]^{n-1}$ .

Множество Парето:  $0 \leq x_1 \leq 1$ ,

$$x_j = \begin{cases} \left(0.3x_1^2 \cdot \cos\left(24\pi x_1 + \frac{4j\pi}{n}\right) + 0.6x_1\right) \cdot \cos\left(6\pi x_1 + \frac{j\pi}{n}\right), j \in J_1, \\ \left(0.3x_1^2 \cdot \cos\left(24\pi x_1 + \frac{4j\pi}{n}\right) + 0.6x_1\right) \cdot \sin\left(6\pi x_1 + \frac{j\pi}{n}\right), j \in J_2. \end{cases} \quad (2.8)$$

Фронт Парето:  $f_2 = 1 - \sqrt{f_1}$ ,  $0 \leq f_1 \leq 1$ .

3. Тестовая задача №3:

$$\begin{cases} f_1 = x_1 + \frac{2}{|J_1|} (4 \cdot \sum_{j \in J_1} y_j^2 - 2 \cdot \prod_{j \in J_1} \cos\left(\frac{20y_j\pi}{\sqrt{j}}\right) + 2) \rightarrow \min, \\ f_2 = 1 - \sqrt{x_1} + \frac{2}{|J_2|} (4 \cdot \sum_{j \in J_2} y_j^2 - 2 \cdot \prod_{j \in J_2} \cos\left(\frac{20y_j\pi}{\sqrt{j}}\right) + 2) \rightarrow \min, \end{cases} \quad (2.9)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ ,  $y_j = x_j - x_1^{0.5(1+\frac{3(j-2)}{n-2})}$ ,  $j = \overline{2, n}$ .

Область определения:  $\bar{x} \in [0, 1]^n$ .

Множество Парето:  $0 \leq x_1 \leq 1$ ,  $x_j = x_1^{0.5(1+\frac{3(j-2)}{n-2})}$ ,  $j = \overline{2, n}$ .

Фронт Парето:  $f_2 = 1 - \sqrt{f_1}$ ,  $0 \leq f_1 \leq 1$ .

4. Тестовая задача №4:

$$\begin{cases} f_1 = x_1 + \frac{2}{|J_1|} \sum_{j \in J_1} h(y_j) \rightarrow \min, \\ f_2 = 1 - x_1^2 + \frac{2}{|J_2|} \sum_{j \in J_2} h(y_j) \rightarrow \min, \end{cases} \quad (2.10)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ ,  $y_j = x_j - \sin\left(6\pi x_1 + \frac{j\pi}{n}\right)$ ,  $j = \overline{2, n}$ ,  $h(t) = \frac{|t|}{1+e^{2|t|}}$ .

Область определения:  $\bar{x} \in [0, 1] \cdot [-2, 2]^{n-1}$ .

Множество Парето:  $0 \leq x_1 \leq 1$ ,  $x_j = \sin\left(6\pi x_1 + \frac{j\pi}{n}\right)$ ,  $j = \overline{2, n}$ .

Фронт Парето:  $f_2 = 1 - f_1^2$ ,  $0 \leq f_1 \leq 1$ .

5. Тестовая задача №5:

$$\begin{cases} f_1 = x_1 + \left(\frac{1}{2N} + \varepsilon\right) \cdot |\sin(2N\pi x_1)| + \frac{2}{|J_1|} \sum_{j \in J_1} h(y_j) \rightarrow \min, \\ f_2 = 1 - x_1 + \left(\frac{1}{2N} + \varepsilon\right) \cdot |\sin(2N\pi x_1)| + \frac{2}{|J_2|} \sum_{j \in J_2} h(y_j) \rightarrow \min, \end{cases} \quad (2.11)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ ,  $y_j = x_j - \sin\left(6\pi x_1 + \frac{j\pi}{n}\right)$ ,  $j = \overline{2, n}$ ,  $h(t) = 2t^2 - \cos(4\pi t) + 1$ ,  $\varepsilon = 0.1$ ,  $N = 10$ .

Область определения:  $\bar{x} \in [0, 1] \cdot [-1, 1]^{n-1}$ .

Фронт Парето:  $\left(\frac{i}{2N}, 1 - \frac{i}{2N}\right)$ ,  $i = \overline{0, 2N}$ .

6. Тестовая задача №6:

$$\left\{ \begin{array}{l} f_1 = x_1 + \max \left\{ 0, 2 \left( \frac{1}{2N} + \varepsilon \right) \sin(2N\pi x_1) \right\} + \\ + \frac{2}{|J_1|} \left( 4 \cdot \sum_{j \in J_1} y_j^2 - 2 \cdot \prod_{j \in J_1} \cos \left( \frac{20y_j\pi}{\sqrt{j}} \right) + 2 \right) \rightarrow \min, \\ f_2 = 1 - x_1 + \max \left\{ 0, 2 \left( \frac{1}{2N} + \varepsilon \right) \sin(2N\pi x_1) \right\} + \\ + \frac{2}{|J_2|} \left( 4 \cdot \sum_{j \in J_2} y_j^2 - 2 \cdot \prod_{j \in J_2} \cos \left( \frac{20y_j\pi}{\sqrt{j}} \right) + 2 \right) \rightarrow \min, \end{array} \right. \quad (2.12)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ ,  $y_j = x_j - \sin \left( 6\pi x_1 + \frac{j\pi}{n} \right)$ ,  $j = \overline{2, n}$ ,  $\varepsilon = 0.1$ ,  $N = 10$ .

Область определения:  $\bar{x} \in [0, 1] \cdot [-1, 1]^{n-1}$ .

Фронт Парето:  $(0, 1)$  – изолированная точка и  $N$  разрывных частей:  $f_2 = 1 - f_1$ ,  $f_1 \in \bigcup_{i=1}^N \left( \frac{2i-1}{2N}, \frac{2i}{2N} \right)$ .

7. Тестовая задача №7:

$$\left\{ \begin{array}{l} f_1 = \sqrt[5]{x_1} + \frac{2}{|J_1|} \sum_{j \in J_1} y_j^2 \rightarrow \min, \\ f_2 = 1 - \sqrt[5]{x_1} + \frac{2}{|J_2|} \sum_{j \in J_2} y_j^2 \rightarrow \min, \end{array} \right. \quad (2.13)$$

где  $n = 30$ ,  $J_1 = \{j | j - \text{нечетное}, 2 \leq j \leq n\}$ ,  $J_2 = \{j | j - \text{четное}, 2 \leq j \leq n\}$ ,  $y_j = x_j - \sin \left( 6\pi x_1 + \frac{j\pi}{n} \right)$ ,  $j = \overline{2, n}$ .

Область определения:  $\bar{x} \in [0, 1] \cdot [-1, 1]^{n-1}$ .

Множество Парето:  $0 \leq x_1 \leq 1$ ,  $x_j = \sin \left( 6\pi x_1 + \frac{j\pi}{n} \right)$ ,  $j = \overline{2, n}$ .

Фронт Парето:  $f_2 = 1 - f_1$ ,  $0 \leq f_1 \leq 1$ .

8. Тестовая задача №8:

$$\left\{ \begin{array}{l} f_1 = \cos(0.5x_1\pi) \cos(0.5x_2\pi) + \\ + \frac{2}{|J_1|} \sum_{j \in J_1} (x_j - 2x_2 \cdot \sin(2\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \\ f_2 = \cos(0.5x_1\pi) \cos(0.5x_2\pi) + \\ + \frac{2}{|J_2|} \sum_{j \in J_2} \left( x_j - 2x_2 \cdot \sin\left(2\pi x_1 + \frac{j\pi}{n}\right) \right)^2 \rightarrow \min, \\ f_3 = \sin(0.5x_1\pi) + \frac{2}{|J_3|} \sum_{j \in J_3} \left( x_j - 2x_2 \cdot \sin\left(2\pi x_1 + \frac{j\pi}{n}\right) \right)^2 \rightarrow \min, \end{array} \right. \quad (2.14)$$

где  $J_1 = \{j | (j-1) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_2 = \{j | (j-2) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_3 = \{j | j - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $n = 30$ .

Область определения:  $\bar{x} \in [0,1]^2 \cdot [-2,2]^{n-2}$ .

Множество Парето:  $0 \leq x_1, x_2 \leq 1, x_j = 2x_2 \sin\left(2\pi x_1 + \frac{j\pi}{n}\right), j = \overline{3, n}$ .

Фронт Парето:  $f_1^2 + f_2^2 + f_3^2 = 1, 0 \leq f_1, f_2, f_3 \leq 1$ .

9. Тестовая задача №9:

$$\left\{ \begin{array}{l} f_1 = 0.5(\max\{0, (1+\varepsilon)(1-4(2x_1-1)^2)\} + 2x_1)x_2 + \\ + \frac{2}{|J_1|} \sum_{j \in J_1} (x_j - 2x_2 \sin(2\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \\ f_2 = 0.5(\max\{0, (1+\varepsilon)(1-4(2x_1-1)^2)\} + 2x_1)x_2 + \\ + \frac{2}{|J_2|} \sum_{j \in J_2} (x_j - 2x_2 \sin(2\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \\ f_3 = 1 - x_2 + \frac{2}{|J_3|} \sum_{j \in J_3} (x_j - 2x_2 \sin(2\pi x_1 + \frac{j\pi}{n}))^2 \rightarrow \min, \end{array} \right. \quad (2.15)$$

где  $J_1 = \{j | (j-1) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_2 = \{j | (j-2) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_3 = \{j | j - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $n = 30, \varepsilon = 0.1$ .

Область определения:  $\bar{x} \in [0,1]^2 \cdot [-2,2]^{n-2}$ .

Множество Парето:  $x_1 \in [0, 0.25] \cup [0.75, 1], 0 \leq x_2 \leq 1, x_j = 2x_2 \sin\left(2\pi x_1 + \frac{j\pi}{n}\right), j = \overline{3, n}$ .

Фронт Парето:  $0 \leq f_3 \leq 1, 0 \leq f_1 \leq \frac{1}{4}(1-f_3), f_2 = 1 - f_1 - f_3$  и  $0 \leq f_3 \leq 1, \frac{3}{4}(1-f_3) \leq f_1 \leq 1, f_2 = 1 - f_1 - f_3$ .

10. Тестовая задача №10:



$$\left\{ \begin{array}{l} f_1 = \cos(0.5x_1\pi) \cos(0.5x_2\pi) + \\ + \frac{2}{|J_1|} \sum_{j \in J_1} (4y_j^2 - \cos(8\pi y_j) + 1) \rightarrow \min, \\ f_2 = \cos(0.5x_1\pi) \cos(0.5x_2\pi) + \\ + \frac{2}{|J_2|} \sum_{j \in J_2} (4y_j^2 - \cos(8\pi y_j) + 1) \rightarrow \min, \\ f_3 = \sin(0.5x_1\pi) + \frac{2}{|J_3|} \sum_{j \in J_3} (4y_j^2 - \\ - \cos(8\pi y_j) + 1) \rightarrow \min, \end{array} \right. \quad (2.16)$$

где  $J_1 = \{j | (j-1) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_2 = \{j | (j-2) - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $J_3 = \{j | j - \text{кратно } 3, 3 \leq j \leq n\}$ ,  $y_j = x_j - 2x_2 \sin\left(2\pi x_1 + \frac{j\pi}{n}\right)$ ,  $j = \overline{3, n}$ ,  $n = 30$ .

Область определения:  $\bar{x} \in [0, 1]^2 \cdot [-2, 2]^{n-2}$ .

Множество Парето:  $0 \leq x_1, x_2 \leq 1$ ,  $x_j = 2x_2 \sin\left(2\pi x_1 + \frac{j\pi}{n}\right)$ ,  $j = \overline{3, n}$ .

Фронт Парето:  $f_1^2 + f_2^2 + f_3^2 = 1$ ,  $0 \leq f_1, f_2, f_3 \leq 1$ .

Сравнение эффективности алгоритмов проводилось по метрике IGD (indicator of generational distance):

$$IGD(A, P^*) = \frac{\sum_{v \in P^*} d(v, A)}{|P^*|}, \quad (2.17)$$

где  $P^*$  – истинный фронт Парето,  $A$  – аппроксимация фронта Парето в результате работы алгоритма оптимизации,  $v$  – точка истинного фронта Парето,  $d(v, A)$  – минимальное расстояние между точкой  $v \in P^*$  и множеством  $A$ , вычисленное по евклидовой метрике. Чем меньше значение метрики IGD, тем эффективнее решена задача оптимизации.

Так как алгоритмы SPEA, VEGA, NSGA-2 и SelfCOMOGA являются рандомизированными, проводилось 20 запусков каждого из вышеперечисленных алгоритмов на каждой из 10 тестовых задач. По результатам прогонов фиксировалось значение метрики IGD и время работы программной реализации алгоритма в секундах. На основе набранной статистики была проверена значимость отличий метрики IGD между рассматриваемыми алгоритмами с помощью W-критерия Вилкоксона.

Значения параметров генетических алгоритмов, использованные в экспериментах, представлены в таблице 2.1.

Таблица 2.1. Значения параметров эволюционных алгоритмов многокритериальной оптимизации, использованные при тестировании на задачах конкурса CEC

| Алгоритм                       | Параметр                                                                                                       | Значение                   |
|--------------------------------|----------------------------------------------------------------------------------------------------------------|----------------------------|
| SelfCOMOGA, SPEA, VEGA, NSGA-2 | $N_p$ – размер популяции, для коэволюционного алгоритма - общий размер всех подпопуляций                       | 600                        |
|                                | Число поколений                                                                                                | 500                        |
|                                | Тип скрещивания                                                                                                | Одноточечное               |
|                                | Вероятность скрещивания                                                                                        | 1                          |
|                                | Вероятность мутации                                                                                            | $\frac{1}{k * N_p}, k = 3$ |
|                                | Точность представления решения                                                                                 | 0.0001                     |
| SelfCOMOGA                     | Размер периода адаптации коэволюционирующих алгоритмов (в поколениях)                                          | 10                         |
|                                | Размер штрафа (% от текущего размера популяции коэволюционирующего алгоритма)                                  | 10                         |
|                                | Минимальный гарантированный размер популяции (% от начального размера популяции коэволюционирующего алгоритма) | 10                         |
| SPEA                           | Максимальный размер внешнего множества                                                                         | $N_p$                      |

Результаты тестирования алгоритмов на тестовых задачах (выборочное среднее, среднеквадратичное отклонение, минимум и максимум метрики IGD и времени выполнения) представлены в таблицах 2.2 – 2.5. В таблице 2.6 представлены результаты применения W-критерия для определения значимости отличий между метрикой IGD алгоритма SelfCOMOGA и алгоритмами SPEA, VEGA, NSGA-2.

Таблица 2.2. Результаты тестирования алгоритма SPEA на тестовых задачах конкурса CEC

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |        |        |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|--------|--------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min    | max    |
| 1            | 0,110                | 0,015    | 0,097 | 0,140 | 92.2                  | 17.0     | 76.9   | 119.5  |
| 2            | 0,086                | 0,025    | 0,026 | 0,155 | 611.8                 | 116.6    | 475.1  | 788.8  |
| 3            | 0,355                | 0,107    | 0,214 | 0,579 | 49.8                  | 19.7     | 27.6   | 66.1   |
| 4            | 0,070                | 0,029    | 0,035 | 0,114 | 61.9                  | 4.8      | 57.6   | 68.2   |
| 5            | 0,833                | 0,084    | 0,801 | 1,039 | 53.6                  | 1.7      | 52.2   | 56.5   |
| 6            | 0,202                | 0,042    | 0,101 | 0,241 | 54.8                  | 3.7      | 51.3   | 60.5   |
| 7            | 0,096                | 0,026    | 0,051 | 0,123 | 505.4                 | 45.0     | 455.3  | 561.7  |
| 8            | 0,439                | 0,108    | 0,271 | 0,546 | 589.3                 | 166.5    | 399.4  | 817.7  |
| 9            | 0,284                | 0,048    | 0,212 | 0,343 | 1317.7                | 52.1     | 1289.3 | 1410.7 |
| 10           | 0,698                | 0,137    | 0,379 | 0,942 | 447.9                 | 195.7    | 266.0  | 708.9  |

Таблица 2.3. Результаты тестирования алгоритма VEGA на тестовых задачах конкурса CEC

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |      |      |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|------|------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min  | max  |
| 1            | 0,560                | 0,053    | 0,455 | 0,578 | 34.2                  | 2.0      | 32.9 | 37.8 |
| 2            | 0,707                | 0,090    | 0,637 | 0,872 | 33.1                  | 0.3      | 32.7 | 33.5 |
| 3            | 0,981                | 0,112    | 0,821 | 1,168 | 30.5                  | 0.8      | 29.6 | 31.6 |
| 4            | 0,453                | 0,113    | 0,259 | 0,715 | 26.2                  | 1.1      | 24.6 | 27.7 |
| 5            | 0,936                | 0,053    | 0,850 | 0,964 | 44.5                  | 2.0      | 41.0 | 46.2 |

Продолжение таблицы 2.2.

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |      |      |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|------|------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min  | max  |
| 6            | 0,478                | 0,049    | 0,363 | 0,553 | 37.5                  | 5.4      | 32.6 | 44.9 |
| 7            | 0,274                | 0,115    | 0,103 | 0,419 | 36.1                  | 2.2      | 33.4 | 39.0 |
| 8            | 1,314                | 0,657    | 0,823 | 2,893 | 37.5                  | 3.2      | 33.5 | 41.9 |
| 9            | 2,247                | 0,567    | 0,887 | 3,101 | 40.1                  | 3.6      | 37.7 | 45.8 |
| 10           | 2,996                | 1,729    | 0,466 | 5,272 | 37.1                  | 0.5      | 36.3 | 37.5 |

Таблица 2.4. Результаты тестирования алгоритма NSGA-2 на тестовых задачах конкурса CEC

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |       |       |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|-------|-------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min   | max   |
| 1            | 0,107                | 0,012    | 0,083 | 0,128 | 241.0                 | 6.2      | 232.9 | 248.5 |
| 2            | 0,051                | 0,012    | 0,023 | 0,070 | 235.9                 | 6.2      | 225.4 | 241.7 |
| 3            | 0,326                | 0,047    | 0,282 | 0,407 | 300.3                 | 29.6     | 257.5 | 321.8 |
| 4            | 0,123                | 0,054    | 0,012 | 0,193 | 228.5                 | 16.4     | 218.0 | 257.4 |
| 5            | 0,842                | 0,043    | 0,755 | 0,927 | 355.0                 | 48.0     | 272.1 | 393.7 |
| 6            | 0,166                | 0,037    | 0,116 | 0,192 | 243.0                 | 5.5      | 234.5 | 249.0 |
| 7            | 0,092                | 0,030    | 0,037 | 0,165 | 242.6                 | 5.8      | 233.4 | 249.5 |
| 8            | 0,932                | 0,165    | 0,760 | 1,041 | 261.6                 | 7.3      | 253.4 | 270.5 |
| 9            | 0,227                | 0,107    | 0,087 | 0,358 | 266.4                 | 11.4     | 252.8 | 284.0 |
| 10           | 0,743                | 0,360    | 0,314 | 1,253 | 271.2                 | 8.6      | 264.0 | 285.1 |

Таблица 2.5. Результаты тестирования алгоритма SelfCOMOGA на тестовых задачах конкурса CEC

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |       |       |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|-------|-------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min   | max   |
| 1            | 0,066                | 0,017    | 0,021 | 0,090 | 173.9                 | 5.4      | 165.9 | 179.7 |
| 2            | 0,027                | 0,010    | 0,023 | 0,049 | 190.5                 | 7.6      | 182.6 | 200.4 |
| 3            | 0,187                | 0,082    | 0,041 | 0,309 | 226.2                 | 15.0     | 211.9 | 244.2 |
| 4            | 0,040                | 0,011    | 0,028 | 0,054 | 190.5                 | 16.4     | 177.4 | 217.4 |

Продолжение таблицы 2.5.

| Номер задачи | Значение метрики IGD |          |       |       | Время выполнения, сек |          |       |       |
|--------------|----------------------|----------|-------|-------|-----------------------|----------|-------|-------|
|              | $\bar{x}$            | $\sigma$ | min   | max   | $\bar{x}$             | $\sigma$ | min   | max   |
| 5            | 0,677                | 0,050    | 0,618 | 0,782 | 150.4                 | 96.6     | 75.6  | 275.9 |
| 6            | 0,101                | 0,032    | 0,058 | 0,168 | 203.7                 | 16.4     | 191.5 | 231.7 |
| 7            | 0,053                | 0,018    | 0,005 | 0,080 | 201.9                 | 4.9      | 194.0 | 207.4 |
| 8            | 0,854                | 0,207    | 0,335 | 1,336 | 223.0                 | 9.2      | 211.2 | 235.5 |
| 9            | 0,159                | 0,058    | 0,097 | 0,236 | 243.3                 | 49.4     | 211.1 | 328.9 |
| 10           | 0,507                | 0,115    | 0,377 | 0,578 | 230.3                 | 25.5     | 212.1 | 275.3 |

Таблица 2.6. Результат применения W-критерия Вилкоксона для определения значимости отличий между метрикой IGD алгоритма SelfCOMOGA и алгоритмами SPEA, VEGA, NSGA-2 (уровень значимости  $p = 0.05$ ,  $\approx$  - отличие незначимо)

| Номер задачи | SelfCOMOGA против SPEA | SelfCOMOGA против VEGA | SelfCOMOGA против NSGA-2 |
|--------------|------------------------|------------------------|--------------------------|
| 1            | Превосходит            | Превосходит            | Превосходит              |
| 2            | Превосходит            | Превосходит            | Превосходит              |
| 3            | Превосходит            | Превосходит            | Превосходит              |
| 4            | Превосходит            | Превосходит            | Превосходит              |
| 5            | Превосходит            | Превосходит            | Превосходит              |
| 6            | Превосходит            | Превосходит            | Превосходит              |
| 7            | Превосходит            | Превосходит            | Превосходит              |
| 8            | Уступает               | Превосходит            | $\approx$                |
| 9            | Превосходит            | Превосходит            | $\approx$                |
| 10           | Превосходит            | Превосходит            | Превосходит              |

Из результатов видно, что на 9 тестовых задачах из 10 алгоритм SelfCOMOGA статистически значимо (при уровне значимости  $p = 0.05$ ) превосходит алгоритмы SPEA, VEGA и NSGA-2 по критерию IGD, уступая алгоритму SPEA на одной задаче. Алгоритм SelfCOMOGA быстрее алгоритмов SPEA и NSGA-2 по времени выполнения, и уступает по этому параметру только

алгоритму VEGA, при этом существенно превосходя его по критерию IGD, который в данном случае является ключевым.

Полученные результаты доказывают эффективность разработанного самонастраивающегося коэволюционного алгоритма и целесообразность объединения в коэволюцию нескольких эволюционных алгоритмов многокритериальной оптимизации с различными механизмами селекции решений.

В приложении А приведено сравнение средней метрики IGD алгоритма SelfCOMOGA с алгоритмами-победителями конкурса CEC для каждой из тестовых задач. Алгоритм SelfCOMOGA на различных задачах занял различные позиции среди алгоритмов-победителей, от 4 до 14. Однако на всех 10 тестовых задачах качество работы разработанного алгоритма сравнимо с качеством работы алгоритмов-победителей, что свидетельствует об универсальности алгоритма SelfCOMOGA применительно к решению разноплановых сложных задач многокритериальной оптимизации. Помимо этого, алгоритм SelfCOMOGA является самонастраиваемым, следовательно исключается необходимость тонкой настройки его параметров.

## **Выводы**

В данной главе описан коллективный самоконфигурируемый эволюционный алгоритм многокритериальной оптимизации (SelfCOMOGA). Идея алгоритма основана на идее коэволюции, то есть параллельной эволюции нескольких популяций решений и обмене решениями между ними (миграция). В качестве параллельно работающих алгоритмов, входящих в коэволюцию, были выбраны алгоритмы SPEA, VEGA и NSGA-2. Выбор именно этих алгоритмов обусловлен тем, что в них используются принципиально разные подходы к селекции индивидов и поддержанию разнообразия популяции, а следовательно,

объединение данных алгоритмов в коллектив способно дать общее улучшение эффективности многокритериальной оптимизации.

Численное сравнение коэволюционного алгоритма с алгоритмами SPEA, VEGA и NSGA-2 по метрике IGD было проведено на тестовом наборе задач многокритериальной оптимизации CEC, включающем в себя 10 задач с двумя либо тремя оптимизируемыми критериями.

В результате проведения серии вычислительных экспериментов было выявлено, что алгоритм SelfCOMOGA превосходит по метрике IGD входящие в него алгоритмы-компоненты, либо не уступает им в 9 из 10 тестовых задач, уступая алгоритму SPEA на одной задаче.

Качество работы коэволюционного алгоритма SelfCOMOGA на наборе тестовых задач CEC сравнимо с качеством работы алгоритмов-победителей одноименного конкурса. Разработанный алгоритм показал высокие результаты на всех 10 тестовых задачах. При этом данный алгоритм является самонастраиваемым, что исключает необходимость тонкой настройки его параметров. Это свидетельствует о его универсальности и автономности.

В следующей главе описанные алгоритмы многокритериальной оптимизации, в том числе и алгоритм SelfCOMOGA, будут применены в рамках многокритериального подхода для решения задач отбора информативных признаков и проектирования оптимального коллектива нейронных сетей для ряда тестовых и практических задач машинного обучения.

## **Глава 3. Разработка многокритериального подхода к проектированию ансамбля классификаторов и отбору информативных признаков**

### **§ 3.1 Настройка параметров и проектирование ансамблей алгоритмов машинного обучения**

Алгоритмы машинного обучения используются для решения задач классификации (обучение с учителем), кластеризации (обучение без учителя), обучения с подкреплением и пр. Обучение заключается в использовании имеющихся данных для настройки варьируемых параметров алгоритма с целью подгонки алгоритма под специфику решаемой задачи.

Гиперпараметрическая оптимизация (задача выбора модели) - это задача поиска набора гиперпараметров алгоритма обучения, при которых достигается наибольшая эффективность работы алгоритма при решении конкретной задачи.

Выбор подходящего метода и его настройка наиболее сильно влияют на успех решения поставленной задачи. Работоспособность алгоритмов машинного обучения зависит от параметров алгоритма и выбранной структуры модели. Настройка алгоритма под конкретную решаемую задачу, связанная с выбором оптимального набора параметров алгоритма, представляет собой нетривиальную задачу.

Задача выбора структуры модели традиционно отдавалась на откуп исследователя, так как выбор структуры модели - задача менее формализованная, и более важная для построения эффективной системы машинного обучения.

Яркий пример - параметрическая оценка регрессии. Допустим, у нас есть выборка из 100 точек  $(x, y)$  в двумерном пространстве. Задача состоит в восстановлении функции регрессии  $y = f(x)$  по имеющимся точкам. Проблема состоит в выборе вида функции  $f(\cdot)$ . Она может быть линейная, полиномиальная, показательная, и т.д. Посмотрев на график имеющихся точек, исследователь



может сделать вывод о возможном виде функции регрессии. Однако на практике реальные данные редко идеально подходят под определенный класс функций, кроме того, входная переменная  $x$ , как правило, является вектором, а не скаляром. В этом случае полноценно визуализировать данные для их оценки экспертом уже невозможно. Добавляя к списку проблем возможные пропуски в данных и разреженность данных, задача правильного выбора модели сильно усложняется.

Для решения проблемы оптимального выбора модели был предложен ряд методов, которые можно разделить в 2 группы: алгоритмы гиперпараметрической оптимизации и ансамблевые методы, в которых несколько алгоритмов машинного обучения параллельно используются для коллективного решения поставленной задачи. Хотя ни один из предложенных алгоритмов не является панацеей и не гарантирует успешного решения поставленной задачи, были получены экспериментальные доказательства их эффективности [74, 149].

**Метод бустинга** основан на следующей идее: если несколько слабых моделей объединить в коллектив, получится сильная модель. Под слабой моделью понимается модель, которая решает задачу лучше случайного угадывания, но обладает слабой обобщающей способностью. Сильная модель обладает хорошей обобщающей способностью и правильно классифицирует объекты, которые не использовались при ее построении.

Наиболее известным алгоритмом бустинга является алгоритм AdaBoost [72]. На вход алгоритма подается обучающая выборка  $(x_1, y_1), \dots, (x_n, y_n)$ , каждый объект  $x_i$  которой принадлежит некоторой области определения  $X$ , а каждая метка класса  $y_i$  принадлежит множеству  $Y$ . Предположим, что  $Y = \{-1, +1\}$ , т. е. решается задача бинарной классификации. AdaBoost вызывает заданный слабый алгоритм обучения последовательно некоторое количество итераций  $t = \overline{1, T}$ . Одна из основных идей алгоритма - присвоение весовых коэффициентов объектам обучающей выборки. Вес объекта  $x_i$  на итерации  $t$  обозначим  $D_t(x_i)$ . Изначально все веса устанавливаются одинаковыми, но на каждой итерации веса неверно классифицированных объектов увеличиваются, поэтому алгоритм обучения

вынужден фокусироваться на объектах выборки, наиболее трудных для классификации.

Задача слабого классификатора заключается в поиске слабого правила  $h_t : X \rightarrow \{-1, +1\}$ , подходящего под распределение весов  $D_t$ . Пригодность слабого правила вычисляется по формуле ошибки:

$$e_t = \Pr_{i \sim D_t}[h_t(x_i) \neq y_i] = \sum_{i: h_t(x_i) \neq y_i} D_t(x_i). \quad (3.1)$$

Ошибка вычисляется исходя из распределения весов  $D_t$ , на которых был обучен классификатор. На практике, в качестве слабого классификатора может выступать алгоритм, использующий веса  $D_t$  объектов обучающей выборки, либо для обучения алгоритма может быть использовано подмножество объектов обучающей выборки, отобранных случайно по распределению  $D_t$ . Ниже представлена схема алгоритма AdaBoost.

Дано:  $(x_1, y_1), \dots, (x_n, y_n)$ , где  $x_i \in X$ ,  $y_i \in Y = \{-1, +1\}$ .

Задать  $D_1(x_i) = \frac{1}{n}$ .

For  $t = \overline{1, T}$ :

1. Обучить слабый классификатор, используя распределение  $D_t$ .
2. Найти слабое правило  $h_t : X \rightarrow \{-1, +1\}$  с ошибкой

$$e_t = \sum_{i: h_t(x_i) \neq y_i} D_t(i). \quad (3.2)$$

3. Вычислить вес классификатора  $t$ :

$$\alpha_t = \frac{1}{2} \ln \left( \frac{1 - e_t}{e_t} \right). \quad (3.3)$$

4. Обновить веса объектов обучающей выборки:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t}, & \text{если } h_t(x_i) = y_i; \\ e^{\alpha_t}, & \text{если } h_t(x_i) \neq y_i; \end{cases} \quad (3.4)$$

где  $Z_t$  - нормирующий коэффициент, используемый для того, чтобы  $D_{t+1}$  было распределением.

Вывести конечное правило классификации:

$$H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(x) \right). \quad (3.5)$$

**Метод бэггинга (Bootstrap aggregating)** был предложен Лео Брейманом [50] для улучшения эффективности классификации путем комбинирования классификаторов, построенных на случайных подвыборках заданной обучающей выборки.

Пусть имеется обучающая выборка  $X$  объема  $n$ . Процедура бэггинга генерирует  $m$  новых обучающих выборок  $X_i$ , каждая размером  $n'$ , путем равномерного отбора с замещением из изначальной выборки. Таким образом, элементы выборок  $X_i$  могут содержать дубликаты. Если  $n' = n$ , тогда для больших  $n$  выборка  $X_i$  будет содержать  $\approx 63.2\%$  уникальных экземпляров исходной выборки  $X$ , остальные элементы будут дубликатами. Такой тип отбора элементов выборки называется *bootstrap*.  $m$  полученных выборок используются для обучения  $m$  моделей, которые в свою очередь объединяются в коллектив. В задачах регрессии выходы классификаторов усредняются, в задачах классификации применяется метод голосования.

Bagging приводит к улучшению *нестабильных алгоритмов*, таких как искусственные нейронные сети и деревья CART (classification and regression trees). Благодаря применению бэггинга в некоторых работах было отмечено улучшение в распознавании прообразов [121, 130]. С другой стороны, бэггинг ухудшает эффективность работы стабильных методов, например  $k$  ближайших соседей.

Простейший способ проведения гиперпараметрической оптимизации - **поиск по решетке (grid search)**, который является методом полного перебора по заданному множеству точек в пространстве гиперпараметров алгоритма обучения. Оптимизируемым критерием выступает метрика эффективности алгоритма обучения, обычно вычисляемая по кросс-валидации по обучающей выборке, или по тестовой выборке [86].

Так как пространство параметров алгоритма обучения теоретически может включать вещественные или бесконечные значения для некоторых параметров,

может быть необходимо искусственно задавать границы изменения параметров и точность их дискретизации перед применением метода grid search.

Например, в алгоритме обучения дерево решений (Classification and regression tree, CART) имеется ряд параметров:

1. Максимальная глубина дерева,  $D \in [1; \infty)$ .
2. Минимальный размер листа,  $B \in [1; \infty)$ .
3. Минимальный размер узла для разбиения,  $N \in [2, \infty)$ .

Все эти параметры вещественны, поэтому для применения grid search выбирается ограниченное множество "подходящих" значений, которые может принимать каждый параметр (на пространство поиска накладывается сетка). Затем для каждого набора значений параметров  $(D, B, N)$  строится дерево решений. Для каждого построенного дерева вычисляется его эффективность, либо по тестовой выборке, либо по процедуре кросс-валидации. В итоге, grid search выдает набор значений параметров, по которым было построено наиболее эффективное дерево.

Недостатком алгоритма grid search является "проклятие размерности", в то же время он может быть легко разделен на отдельные подзадачи для параллельных вычислений, так как отдельные наборы значений параметров независимы друг от друга [45].

**Байесовская оптимизация** - это метод глобальной оптимизации функций, заданных моделью черного ящика с шумом. Применительно к гиперпараметрической оптимизации, Байесовская оптимизация подразумевает построение статистической модели функции, описывающей зависимость критерия эффективности построенной модели от значений параметров модели. Предполагается, что зависимость между параметрами модели и критерием эффективности может быть описана некоторой гладкой функцией, включающей в себя шумовые помехи. В байесовской оптимизации целью является подбор значений параметров алгоритма обучения таким образом, чтобы минимизировать количество раз оценивания построенной модели, и при этом получить как можно

больше информации о искомой функции зависимости критерия эффективности от параметров, и в частности о точке оптимума этой функции. Значения гиперпараметров выбираются таким образом, чтобы одновременно вести поиск, при этом используя области пространства поиска, соответствующие лучшим значениям критерия.

На практике было показано, что Байесовская оптимизация позволяет получать лучшие результаты за меньшее количество экспериментов, чем grid search и случайный поиск (random search), за счет возможности прогнозировать эффективность экспериментов по тестированию значений гиперпараметров [133, 138].

В связи с тем, что поиск по сетке является методом полного перебора, и, следовательно, требует много вычислительных ресурсов, были предложены альтернативные методы.

Одним из таких методов является **случайный поиск**, в котором производится случайный выбор набора значений параметров фиксированное число раз. Было показано, что данный метод эффективнее метода полного перебора в задачах с высокой размерностью пространства поиска [45].

Для определенных алгоритмов машинного обучения могут быть использованы специализированные алгоритмы оптимизации гиперпараметров. Например, в работе Chapelle используется алгоритм градиентного спуска для минимизации ошибки классификации метода опорных векторов [58].

Создание ансамблей алгоритмов машинного обучения также является эффективным подходом для повышения точности систем искусственного интеллекта и анализа данных. Среди алгоритмов машинного обучения, объединяемых в ансамбль, прежде всего стоит отметить нейронные сети [80, 147] и деревья решений [37, 117]. Алгоритмы оптимизации применяются для настройки параметров ансамблей алгоритмов машинного обучения. В частности, эволюционные алгоритмы, описанные в предыдущем параграфе, применяются

для автоматизированного проектирования и настройки ансамблей нейронных сетей [92, 139].

Преимущество таких автоматизированных систем в том, что исследователь освобождается от необходимости проектирования алгоритмов машинного обучения вручную. Однако, такие системы, как правило, могут использоваться для решения лишь узкого круга практических задач.

### § 3.2 Разработка и реализация многокритериального подхода к отбору информативных признаков

Отбор информативных признаков является неотъемлемым этапом процесса машинного обучения. Данный этап предшествует непосредственному обучению алгоритма классификации, поэтому от того, насколько качественно выполнен этот этап, зависит эффективность всей системы машинного обучения. Был разработан и исследован метод, в котором применяется многокритериальный подход для отбора признаков (рисунок 3.1).

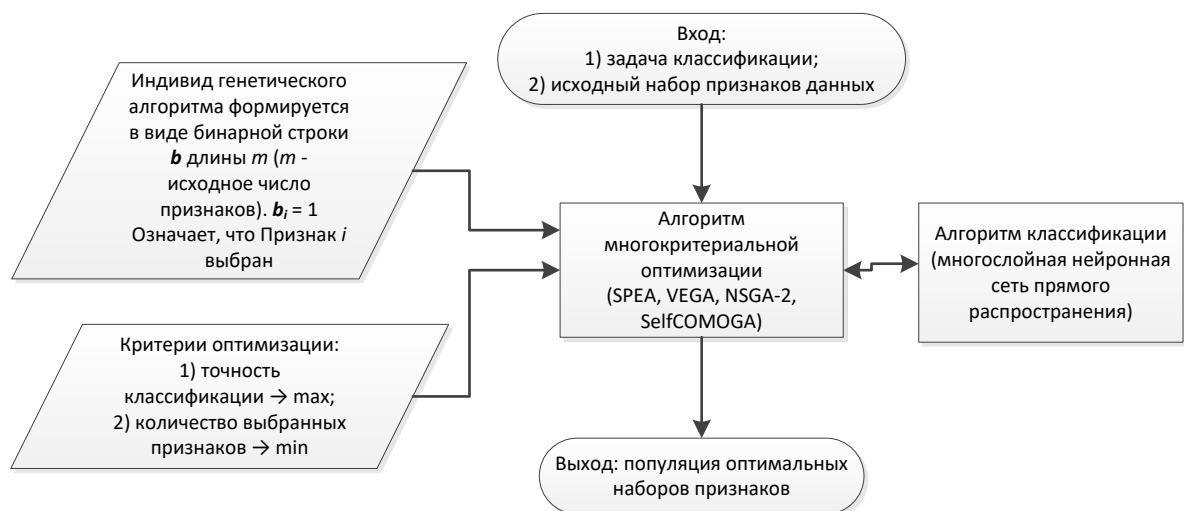


Рисунок 3.1. Схема многокритериального подхода к отбору информативных признаков

Предложенный метод отбора признаков, использующий многокритериальную оптимизацию, относится к классу оберточных методов. Было проведено его сравнение с методом главных компонент (principal component analysis, PCA) и однокритериальным отбором признаков. Методы отбора признаков, основанные на оптимизации, были спроектированы следующим образом. Входящие переменные представляют собой бинарный вектор длины  $m$ , где  $m$  - это исходное число признаков выборки. Каждый бит такого вектора принимает значение 1 либо 0, где 1 означает, что соответствующий признак выбран для дальнейшего включения в модель, 0 - не выбран. В однокритериальной постановке точность классификации является максимизируемым критерием, и задается следующим образом:

$$R = (N_c/N) \times 100 \%, \quad (3.6)$$

где  $N_c$  - число верно классифицированных объектов выборки,  $N$  - общее число объектов выборки,  $R$  - точность классификации.

В многокритериальной постановке добавляется второй (минимизируемый) критерий - число выбранных признаков. Идея данного критерия в том, что модели, построенные на выборках с меньшим количеством признаков, проще, следовательно обладают лучшей обобщающей способностью.

Мы выбрали класс эволюционных алгоритмов оптимизации для решения задачи отбора признаков в вышеприведенной постановке, потому что алгоритмы данного класса эффективно находят глобальный оптимум, и для их использования не требуется информация о поверхности оптимизируемой функции.

В однокритериальной постановке был использован коэволюционный генетический алгоритм (КГА). Данный алгоритм объединяет несколько стандартных ГА, у которых различаются значения параметров. Стандартные ГА параллельно работают некоторое количество итераций, после чего обмениваются индивидами между собой (миграция индивидов) и сохраняют лучшие найденные решения (элитизм). Такой алгоритм освобождает исследователя от настройки

параметров стандартного ГА, неправильная настройка которых зачастую приводит к неудовлетворительным результатам оптимизации.

В многокритериальной постановке были использованы алгоритмы SPEA, VEGA, NSGA-2 и SelfCOMOGA с низкой вероятностью скрещивания и мутации. Количественно данная вероятность вычисляется по следующей формуле:

$$p = 1/(k \times |P|), \quad (3.7)$$

где  $k = 3$  в проведенных экспериментах, но может принимать и другие действительные неотрицательные значения,  $|P|$  - размер популяции,  $p$  - значение вероятности.

Для проведения экспериментов по решению задачи отбора информативных признаков была создана программная система на языке C# в среде разработки Microsoft Visual Studio 2012, в которой были реализованы эволюционные алгоритмы многокритериальной оптимизации SPEA, VEGA, NSGA-2 и SelfCOMOGA. Также для проведения сравнительного исследования был программно реализован коэволюционный алгоритм однокритериальной оптимизации в виде отдельной программной системы. В графическом пользовательском интерфейсе первой программной системы предусмотрена возможность выбора используемого алгоритма многокритериальной оптимизации. Для экспериментов с ансамблем нейронных сетей была использована готовая реализация нейронной сети прямого распространения, выполненная на языке R в среде R 3.1.2. Из основной программы на языке C# производился вызов команд в среде R. Репозитории данных для решаемых тестовых задач классификации также подгружались в среде R.

На рисунках 3.2 и 3.3 представлены главные окна программ "Эволюционные алгоритмы многокритериальной оптимизации" и "Коэволюционный алгоритм однокритериальной оптимизации".



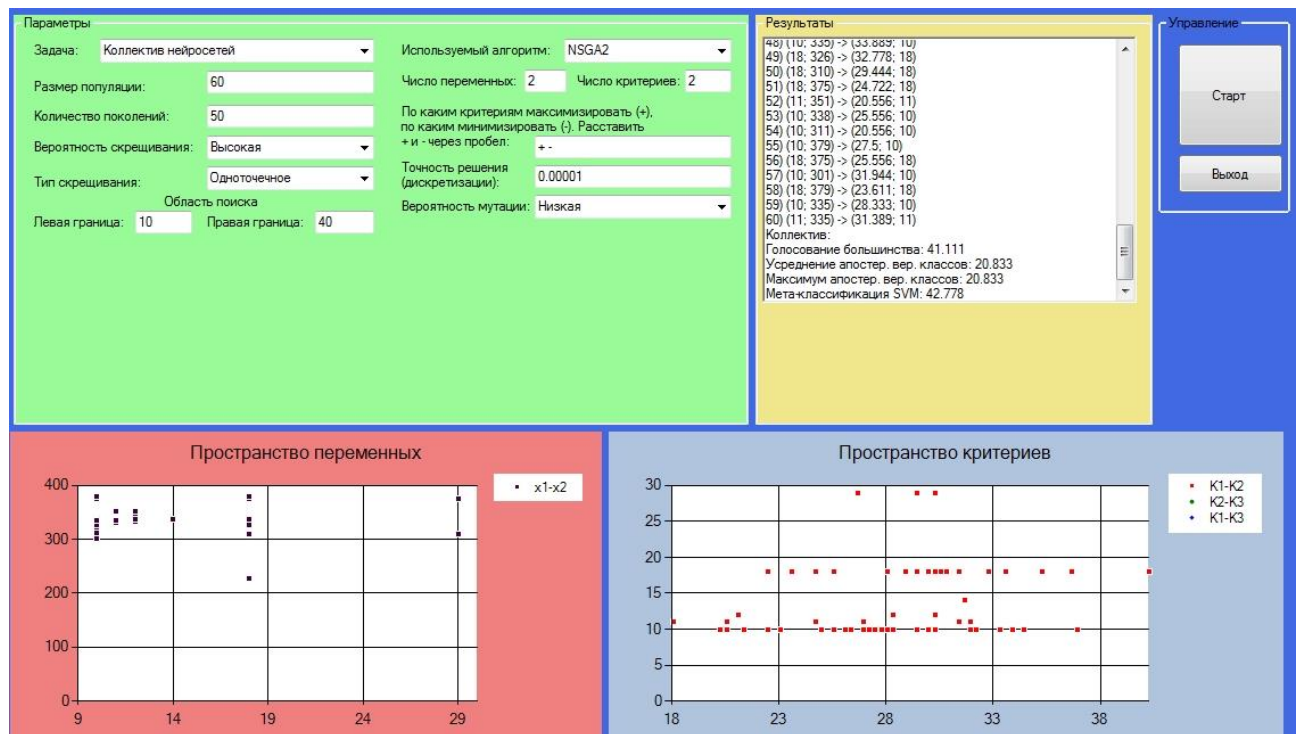


Рисунок 3.2. Главное окно программной системы "Эволюционные алгоритмы многокритериальной оптимизации"

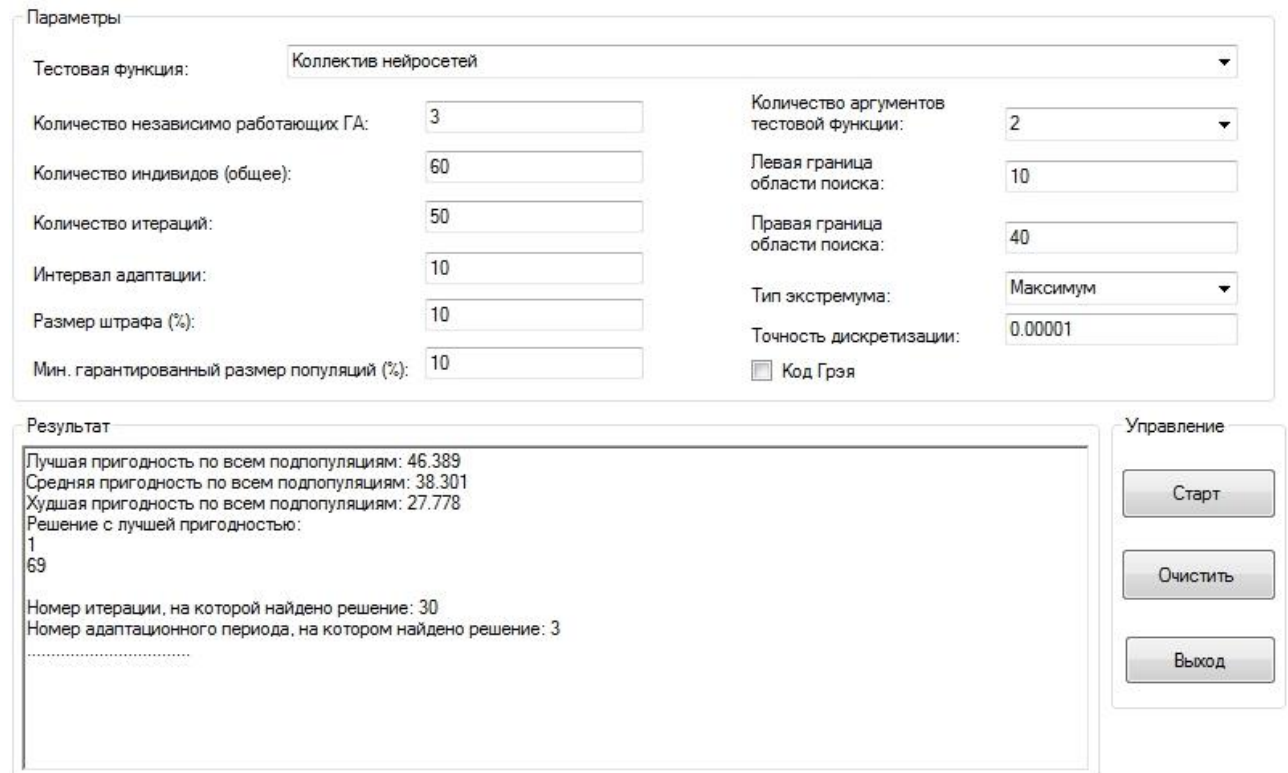


Рисунок 3.3. Главное окно программной системы "Коэволюционный алгоритм однокритериальной оптимизации"



Параметры

Тестовая функция: Feature selection

|                                            |    |                                         |          |
|--------------------------------------------|----|-----------------------------------------|----------|
| Количество независимо работающих ГА:       | 3  | Количество аргументов тестовой функции: | 36       |
| Количество индивидов (общее):              | 60 | Левая граница области поиска:           | 0        |
| Количество итераций:                       | 50 | Правая граница области поиска:          | 1        |
| Интервал адаптации:                        | 10 | Тип экстремума:                         | Максимум |
| Размер штрафа (%):                         | 10 | Точность дискретизации:                 | 0.00001  |
| Мин. гарантированный размер популяций (%): | 10 | <input type="checkbox"/> Код Грэя       |          |

Рисунок 3.6. Окно параметров коэволюционного алгоритма однокритериальной оптимизации, задача отбора информативных признаков

Результат

```

Лучшая пригодность по всем подпопуляциям: 45.625
Средняя пригодность по всем подпопуляциям: 34.663
Худшая пригодность по всем подпопуляциям: 8.125
Решение с лучшей пригодностью:
00011000111101110111000100010000010
Число переменных: 16
Номер итерации, на которой найдено решение: 10
Номер адаптационного периода, на котором найдено решение: 5
.....

```

Рисунок 3.7. Окно вывода результатов работы алгоритма однокритериальной оптимизации, задача отбора информативных признаков

### § 3.3 Разработка и реализация многокритериального подхода к проектированию ансамбля нейросетевых классификаторов

Как было отмечено в предыдущем параграфе, эволюционные алгоритмы часто используются для оптимизации параметров алгоритмов машинного обучения. Традиционно данная задача формулируется в однокритериальной постановке, то есть для каждого набора значений параметров алгоритма рассчитывается некий критерий, который необходимо оптимизировать. Чаще всего в качестве такого критерия выступает точность классификации, если решается задача классификации, либо аналогичный критерий.

Многокритериальный подход к оптимизации алгоритмов машинного обучения отличается тем, что в нем одновременно оптимизируются несколько, часто конфликтующих, критериев.

В данной главе описывается многокритериальный подход, примененный для оптимизации параметров нейронной сети (рисунок 3.8). Нейронная сеть прямого распространения была выбрана в качестве алгоритма обучения по нескольким причинам. Во-первых, нейронные сети ранее успешно применялись для решения практических задач анализа изображений. Во-вторых, эффективность нейронных сетей в значительной мере зависит от настройки их параметров. Была использована однослойная нейронная сеть с переменным числом нейронов и сигмоидной активационной функцией.

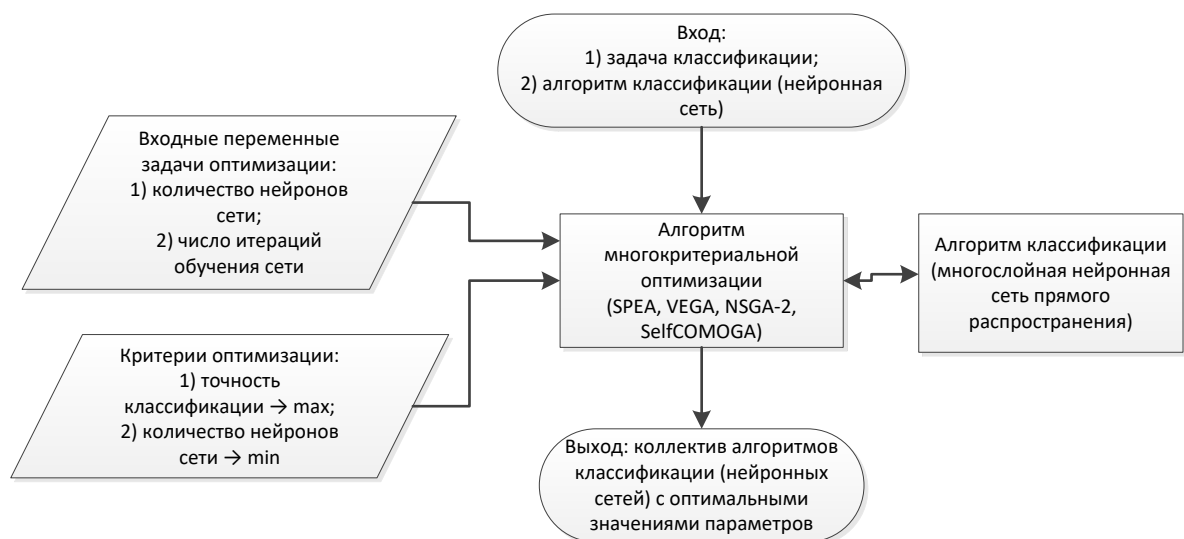


Рисунок 3.8. Схема многокритериального подхода к проектированию ансамбля нейросетевых классификаторов

В однокритериальной постановке входные переменные алгоритма оптимизации включают в себя общее число нейронов сети  $N_n$  (границы изменения  $N_n = \overline{2, 50}$ ) и число итераций обучения сети  $N_t$  (границы изменения  $N_t = \overline{2, 200}$ ). Оптимизируемый критерий - точность классификации.

В многокритериальной постановке отличие заключается в том, что добавляется второй (минимизируемый) критерий - число нейронов сети. Данный критерий одновременно является первой входной переменной.

Для решения сформулированных задач оптимизации были использованы эволюционные алгоритмы оптимизации: в однокритериальной постановке - коэволюционный ГА (КГА), в многокритериальной постановке - Strength Pareto Evolutionary Algorithm (SPEA), Non-dominated Sorting Genetic Algorithm (NSGA-2), Vector Evaluated Genetic Algorithm (VEGA) и Self-configuring Coevolutionary Multiobjective Genetic Algorithm (SelfCOMOGA). В алгоритме SelfCOMOGA алгоритмы SPEA, NSGA-2 и VEGA работают параллельно, и через фиксированное число итераций производят процедуру миграции индивидов и элитизма подобно тому, как это делается в КГА. Для ранжирования индивидов производится их недоминируемая сортировка, которая является обобщением обычной сортировки на случай нескольких критериев.

В однокритериальной постановке решением задачи является нейронная сеть с найденными оптимальными значениями параметров. В многокритериальной постановке мы получаем множество Парето оптимальных нейронных сетей. Для того чтобы можно было сравнивать эти 2 подхода, получаемое Парето-оптимальное множество нейросетей объединяется в коллектив. Были опробованы 3 метода объединения классификаторов в коллектив:

1. Метод голосования - выбирается класс, предсказанный наибольшим количеством базовых алгоритмов ансамбля.

2. Усреднение вероятностей классов - получаемые на выходе классификатора вероятности классов усредняются по всем классификаторам коллектива.

3. Мета-классификация - обучающая выборка делится на 2 части, первая часть используется для обучения классификаторов коллектива. Получаемые на выходе классификаторов вероятности классов используются в качестве входных переменных для дополнительного мета-классификатора, который обучается на

второй части выборки. Мета-классификатор на выходе выдает класс, предсказанный для данного объекта выборки (рисунок 3.9).

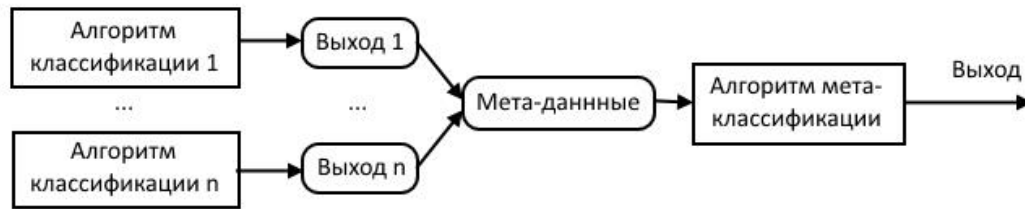


Рисунок 3.9. Схема метода мета-классификации

В программных системах "Эволюционные алгоритмы многокритериальной оптимизации" и "Козволюционный алгоритм однокритериальной оптимизации", описанных в предыдущем параграфе, был также реализован описанный оптимизационный подход к проектированию ансамбля нейросетевых классификаторов.

На рисунке 3.10 представлено окно ввода параметров эволюционного алгоритма многокритериальной оптимизации для решения задачи проектирования ансамбля нейронных сетей, на рисунке 3.11 - окно вывода результатов решения данной задачи. На рисунке 3.12 показано аналогичное окно ввода параметров коэволюционного алгоритма однокритериальной оптимизации для решения задачи оптимизации параметров одиночной нейронной сети, на рисунке 3.13 - окно вывода результатов работы коэволюционного алгоритма.

| Параметры                                                                                           |                      |
|-----------------------------------------------------------------------------------------------------|----------------------|
| Задача:                                                                                             | Коллектив нейросетей |
| Используемый алгоритм:                                                                              | NSGA2                |
| Размер популяции:                                                                                   | 60                   |
| Число переменных:                                                                                   | 2                    |
| Число критериев:                                                                                    | 2                    |
| Количество поколений:                                                                               | 50                   |
| По каким критериям максимизировать (+), по каким минимизировать (-). Расставить + и - через пробел: | + -                  |
| Вероятность скрещивания:                                                                            | Высокая              |
| Точность решения (дискретизации):                                                                   | 0.00001              |
| Тип скрещивания:                                                                                    | Одноточечное         |
| Вероятность мутации:                                                                                | Низкая               |
| Область поиска                                                                                      |                      |
| Левая граница:                                                                                      | 10                   |
| Правая граница:                                                                                     | 40                   |

Рисунок 3.10. Окно параметров алгоритма многокритериальной оптимизации, задача проектирования ансамбля нейронных сетей



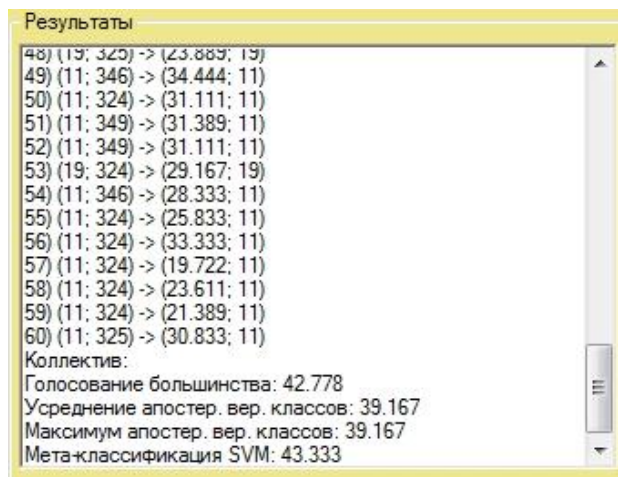


Рисунок 3.11. Окно вывода результатов работы алгоритма многокритериальной оптимизации, задача проектирования ансамбля нейронных сетей

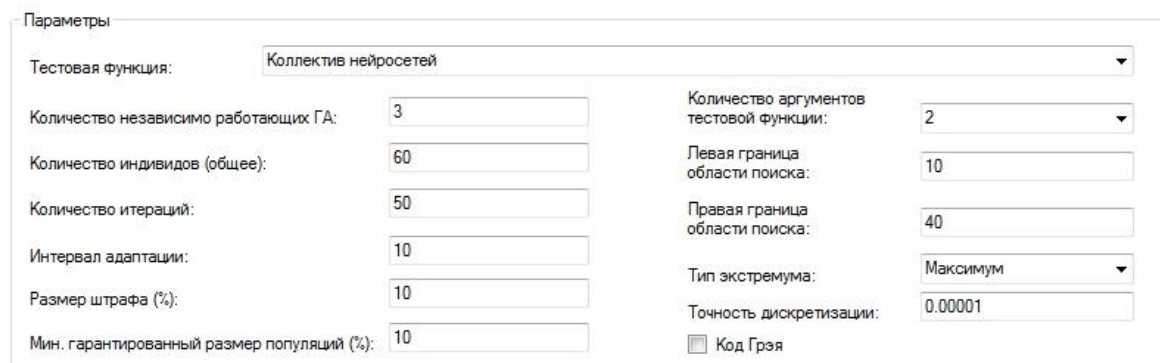


Рисунок 3.12. Окно параметров алгоритма однокритериальной оптимизации, задача оптимизации параметров нейронной сети

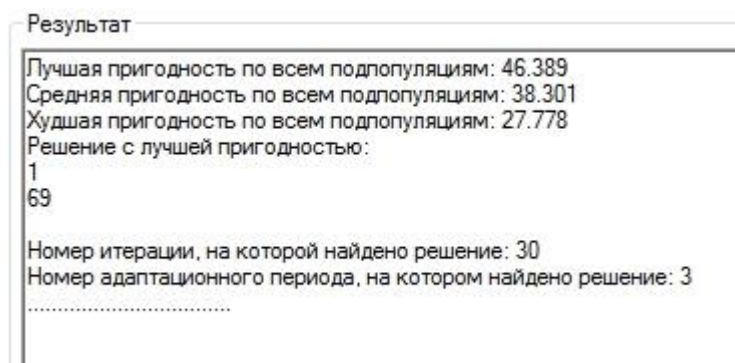


Рисунок 3.13. Окно вывода результатов работы алгоритма однокритериальной оптимизации, задача оптимизации параметров нейронной сети

### **§ 3.4 Исследование эффективности многокритериального подхода к отбору информативных признаков и проектированию ансамбля нейросетевых классификаторов**

В данном параграфе приведены результаты численных экспериментов на множестве практических задач, являющихся тестовыми для оценки и сравнения однокритериального и многокритериального подходов к решению задач отбора информативных признаков и проектирования ансамбля нейросетевых классификаторов, включая задачу определения типа раковой опухоли груди, задачу распознавания рукописных цифр, задачу определения наличия заболевания диабетом, задачу классификации типов почв на спутниковых снимках. Более детальные эксперименты и анализ представлены для поставленной в диссертации задачи распознавания эмоций.

Значения параметров эволюционных алгоритмов оптимизации, использованные в экспериментах, представлены в таблице 3.1.

Так как используемые эволюционные алгоритмы являются рандомизированными, в ходе экспериментов метрики эффективности усреднялись по 20 независимым запускам.

Кроме того, было проведено тестирование эффективности решения каждой из задач набором базовых алгоритмов классификации. Данный набор включает следующие алгоритмы:

1. Метод опорных векторов (Support Vector Machine, SVM).
2. Многослойный перцептрон (Multilayer perceptron, MLP).
3. Логистическая регрессия (Logistic regression, Logit) [85].
4. Аддитивная логистическая регрессия (Additive logistic regression, LogitBoost) [73].
5. Радиально-базисная нейронная сеть (Radial basis function network, RBF) [46].
6. Наивный байесовский классификатор (Naive Bayes).



7. Дерево решений (Classification ad regression tree, CART).
8. Случайный лес (Random forest) [51].
9. Алгоритм генерирования правил 1R (One rule).

Таблица 3.1. Значения параметров генетических алгоритмов

| <b>Параметр генетических алгоритмов</b>                                                      | <b>Значение</b> |
|----------------------------------------------------------------------------------------------|-----------------|
| Размер популяции                                                                             | 60              |
| Число поколений                                                                              | 50              |
| Вероятность скрещивания                                                                      | Высокая         |
| Тип скрещивания                                                                              | Одноточечное    |
| Вероятность мутации                                                                          | Низкая          |
| Максимальный размер внешнего множества (SPEA)                                                | 60              |
| Интервал адаптации (SelfCOMOGA)                                                              | 10              |
| Размер штрафа, % от размера популяции (SelfCOMOGA)                                           | 10%             |
| Минимальный гарантированный размер популяции, % от начального размера популяции (SelfCOMOGA) | 10%             |

Ниже будет проведена апробация описанных в § 3.2-3.3 подходов на тестовых задачах машинного обучения:

1. Определение типа раковой опухоли груди.
2. Распознавание рукописных цифр (репозиторий penbased).
3. Определение наличия заболевания диабетом.
4. Классификация типов почв на спутниковых снимках.

После этого предложенные подходы будут апробированы на более сложной задаче распознавания эмоций (репозиторий SAVEE), сформулированной в главе 1.

### Задача №1 - определение типа раковой опухоли груди

Сперва было проведено тестирование эффективности базовых алгоритмов классификации на исходной выборке, а также на выборке со сниженной размерностью данных алгоритмом PCA [91]. Результаты тестирования представлены на рисунке 3.14.

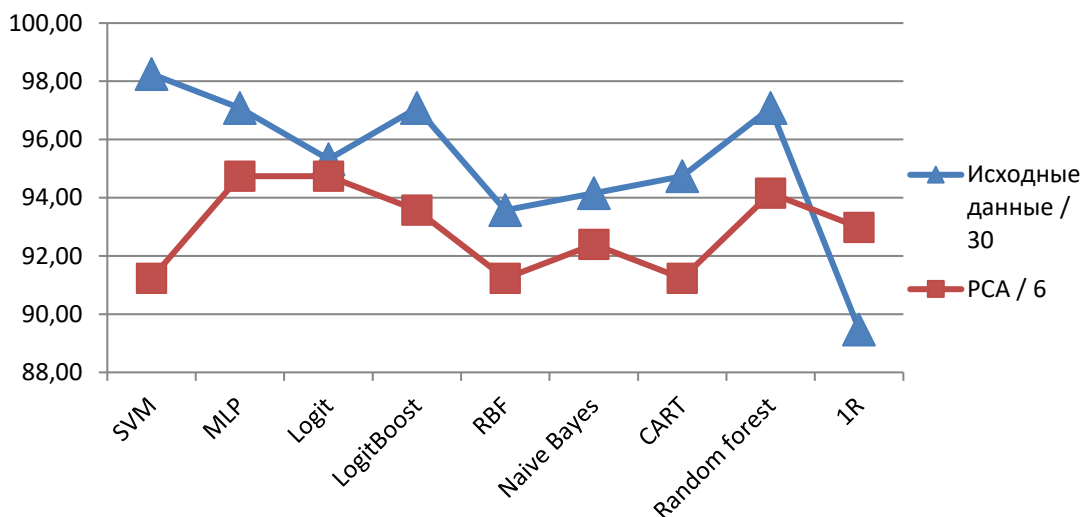


Рисунок 3.14. Точность определения типа раковой опухоли (%) базовыми алгоритмами классификации; данные / число признаков

**Отбор информативных признаков.** Было проведено сравнение различных подходов к снижению размерности признакового пространства, а именно: метод главных компонент, отбор признаков однокритериальным алгоритмом оптимизации, отбор признаков многокритериальным алгоритмом оптимизации. Для однокритериальной оптимизации был применен коэволюционный генетический алгоритм [22], для многокритериальной оптимизации - алгоритмы SPEA, NSGA-2, VEGA, SelfCOMOGA. Все эти алгоритмы были подробно описаны в §2.1-2.2.

Эффективность того или иного подхода определялась по точности классификации эмоций нейронной сетью с использованием выбранного набора признаков. Такой способ проведения экспериментов, когда метод снижения размерности используется в тандеме с алгоритмом классификации, называется

оберточным (wrapper method). Параметры нейронных сетей в данном случае были фиксированы и составляли 20 нейронов на скрытом слое, 200 итераций обучения.

Результаты экспериментов по снижению размерности и отбору информативных признаков для задачи определения типа раковой опухоли груди представлены в таблице Б1 приложения Б.

**Проектирование ансамбля классификаторов.** В качестве алгоритмов оптимизации для проектирования ансамбля нейросетевых классификаторов были использованы те же алгоритмы, что и для отбора признаков - коэволюционный однокритериальный ГА, а также многокритериальные SPEA, NSGA-2, VEGA, SelfCOMOGA. Кроме того, были реализованы различные методы объединения коллектива классификаторов, получаемых на выходе алгоритмов оптимизации.

Выборки, подаваемые на вход нейронных сетей, включали в себя признаки, отобранные соответствующим алгоритмом оптимизации на предыдущем этапе - этапе отбора информативных признаков.

В таблице 3.2 приведен пример Парето-оптимального множества нейронных сетей для задачи определения типа раковой опухоли груди, найденного с помощью алгоритма оптимизации NSGA-2.

Полные результаты экспериментов по проектированию ансамбля нейронных сетей для задачи определения типа раковой опухоли груди представлены в таблице Б2 приложения Б.

Таблица 3.2. Парето-оптимальное множество нейронных сетей для задачи определения типа раковой опухоли груди, алгоритм оптимизации NSGA-2

| № | Число нейронов | Число итераций обучения | Точность классификации |
|---|----------------|-------------------------|------------------------|
| 1 | 10             | 356                     | 62.57                  |
| 2 | 35             | 350                     | 87.72                  |
| 3 | 10             | 303                     | 62.57                  |
| 4 | 18             | 350                     | 63.16                  |
| 5 | 35             | 358                     | <b>92.98</b>           |

Продолжение таблицы 3.2.

| <b>№</b>                        | <b>Число нейронов</b> | <b>Число итераций<br/>обучения</b> | <b>Точность<br/>классификации</b> |
|---------------------------------|-----------------------|------------------------------------|-----------------------------------|
| 6                               | 15                    | 303                                | 88.30                             |
| 7                               | 10                    | 356                                | 88.89                             |
| 8                               | 10                    | 456                                | 90.06                             |
| 9                               | 17                    | 250                                | 88.30                             |
| 10                              | 18                    | 350                                | 91.23                             |
| 11                              | 20                    | 350                                | 88.30                             |
| 12                              | 18                    | 351                                | 62.57                             |
| 13                              | 20                    | 150                                | 90.06                             |
| 14                              | 35                    | 163                                | 62.57                             |
| 15                              | 35                    | 363                                | 92.40                             |
| 16                              | 23                    | 302                                | 87.43                             |
| 17                              | 28                    | 348                                | 89.45                             |
| 18                              | 16                    | 290                                | 90.23                             |
| 19                              | 13                    | 380                                | 78.33                             |
| 20                              | 30                    | 345                                | 86.79                             |
| Голосование                     |                       |                                    | 92.06                             |
| Усреднение вероятностей классов |                       |                                    | 62.57                             |
| Мета-классификация              |                       |                                    | <b>94.15</b>                      |

### Задача №2 - распознавание рукописных цифр (penbased)

См. полные результаты экспериментов в таблицах Б3, Б4 приложения Б.

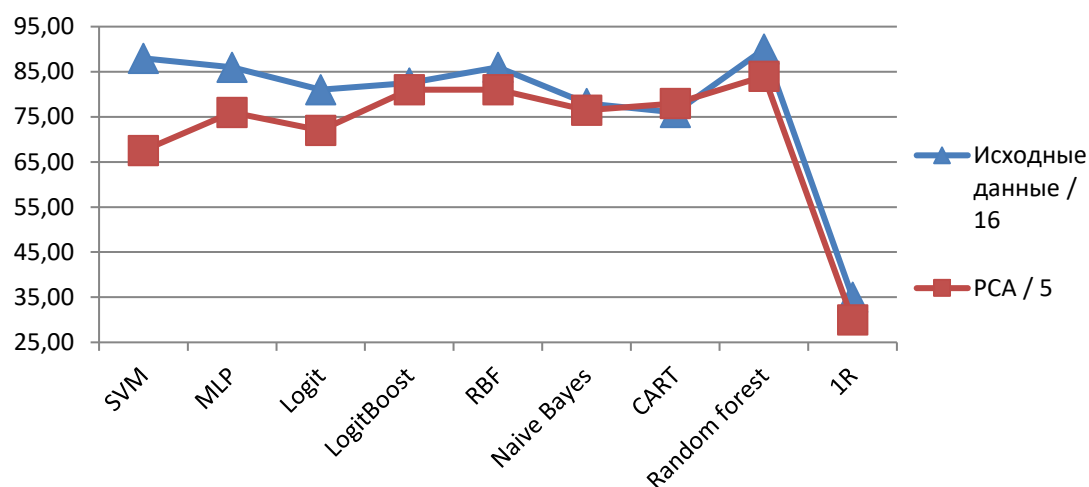


Рисунок 3.15. Точность распознавания рукописных цифр для задачи penbased (%) базовыми алгоритмами классификации; данные / число признаков

### Задача №3 - определение наличия заболевания диабетом

См. полные результаты экспериментов в таблицах Б5, Б6 приложения Б.

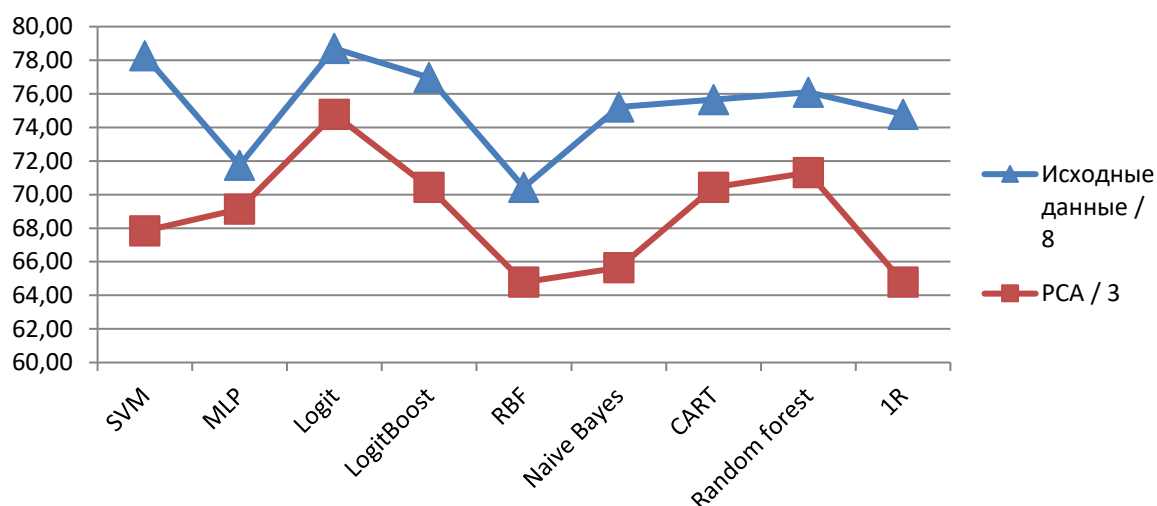


Рисунок 3.16. Точность определения наличия заболевания диабетом (%) базовыми алгоритмами классификации; данные / число признаков

### Задача №4 - классификация типов почв на спутниковых снимках

См. полные результаты экспериментов в таблицах Б7, Б8 приложения Б.

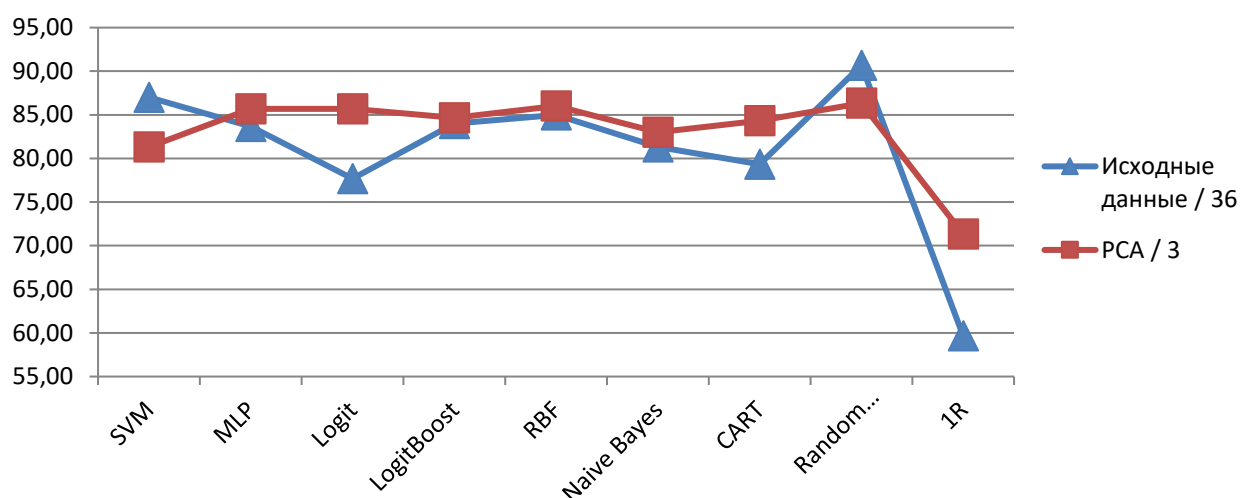


Рисунок 3.17. Точность классификации типов почв (%) базовыми алгоритмами классификации; данные / число признаков

### Задача №5 - распознавание эмоций

На тестовых задачах предлагаемый многокритериальный подход к отбору признаков и проектированию ансамбля нейросетевых классификаторов оказался эффективен. Теперь применим его для решения задачи распознавания эмоций.

В качестве исходных данных для данной задачи был использован репозиторий SAVEE (Surrey Audio-Visual Expressed Emotions), упомянутый в §1.3. Данный репозиторий включает в себя видеозаписи 4 дикторов мужского пола, произносящих набор фраз с определенной эмоциональной окраской - злость, отвращение, страх, счастье, грусть, удивление, а также нейтральная эмоция. Длительность каждой видеозаписи в среднем составляет 3.8 секунд. Репозиторий включает в себя 480 видеозаписей, по 120 на каждого диктора. Суммарная продолжительность всех видеозаписей равна 30.7 минут. На рисунке 3.18 представлено распределение видеозаписей по классам эмоций.

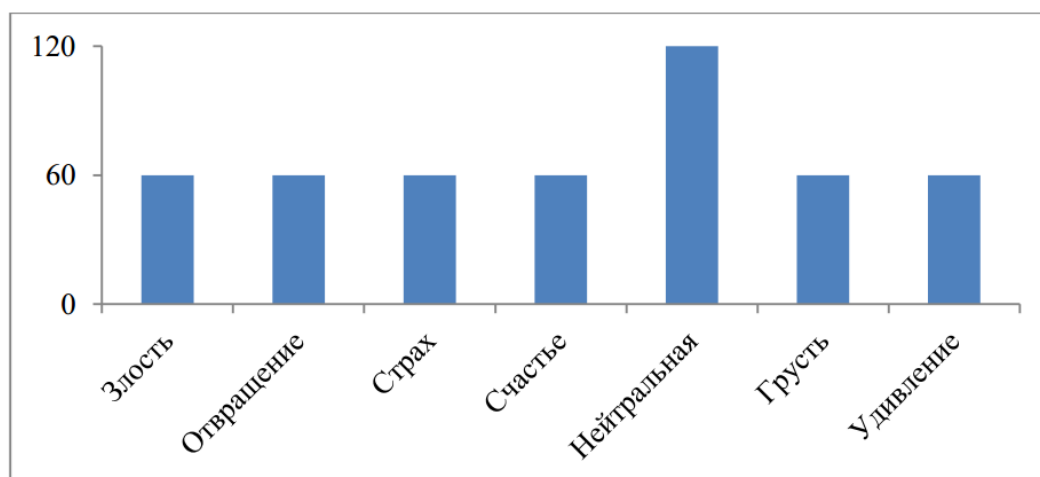


Рисунок 3.18. Распределение видеозаписей репозитория SAVEE по классам эмоций

Репозиторий SAVEE был создан с целью исследования эффективности объединения гетерогенных аудио и видео данных для решения задачи машинного распознавания эмоций. При записи видео на лица дикторов были нанесены специальные точки-маркеры, обозначающие ключевые точки лица. Всего было нанесено 60 маркеров (рисунок 3.19). Данные маркеры с помощью специального прибора отслеживались в ходе записи видео, в результате на выходе были получены координаты каждого маркера на каждом кадре видео. В качестве точки отсчета с нулевыми координатами брались координаты маркера на кончике носа диктора.

На основе полученных данных о координатах маркеров была подготовлена выборка. Координаты каждого маркера были усреднены по всем кадрам видеозаписи, а также вычислены их среднеквадратичные отклонения. Таким образом, была получена выборка, включающая в себя  $60 * 4 = 240$  признаков, и содержащая 480 экземпляров (каждый экземпляр соответствует одной видеозаписи). Соотношение числа признаков и числа экземпляров оказалось слишком высоким, поэтому для предварительного снижения размерности был применен метод главных компонент (principal components analysis, PCA). В итоговую выборку попали 50 наиболее значимых компонент.



Рисунок 3.19. Пример кадра из репозитория SAVEE, на лице диктора нанесены точки-маркеры

Помимо выборки лицевых маркеров, была также подготовлена выборка аудио признаков. Аудио признаки были извлечены из аудиозаписей с помощью программы openSMILE [70] - свободно распространяемого программного обеспечения, предназначенного для извлечения аудио и видео-признаков. Извлеченные аудио признаки включают в себя:

1. Признаки основного тона (pitch features) - извлекаются из контура частоты основной гармоники  $f_0$ .
2. Энергетические признаки (energy features).
3. Признаки звуковой продолжительности (duration features).
4. Спектральные признаки (spectral features) - среднее и среднеквадратичное отклонение мел-частотных кепстральных коэффициентов.

В общем было извлечено 930 различных признаков. С целью предварительного снижения размерности, на аудио выборке также был применен метод PCA, для включения в итоговую выборку были отобраны 50 наиболее значимых компонент.

Выборки лицевых маркеров и аудио признаков были объединены в общую выборку с целью проведения экспериментов по распознаванию эмоций по



гетерогенным аудио-видео данным. Данная выборка включает в себя  $50 + 50 = 100$  признаков.

Все три выборки были разбиты на обучающую и тестовую в дикторонезависимой постановке, т.е. в обучающей и тестовой выборке находятся экземпляры, соответствующие разным дикторам. Данная задача является более сложной, нежели дикторозависимая, т.к. алгоритм тестируется на данных, соответствующих диктору, которого он не видел при обучении.

**Тестирование базовых алгоритмов классификации.** В первую очередь было проведено тестирование эффективности базовых алгоритмов классификации на имеющихся выборках, а также на выборках со сниженной размерностью данных. Результаты тестирования представлены на рисунках 3.20-3.21.

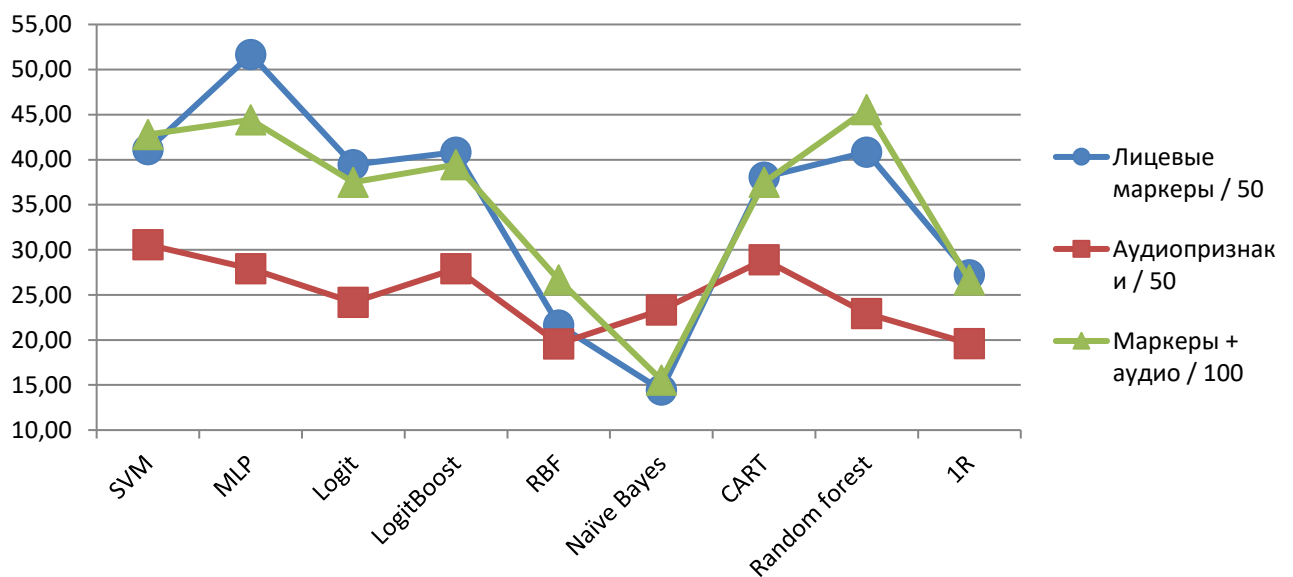


Рисунок 3.20. Точность классификации эмоций (%) базовыми алгоритмами классификации; без снижения размерности методом PCA; данные / число признаков

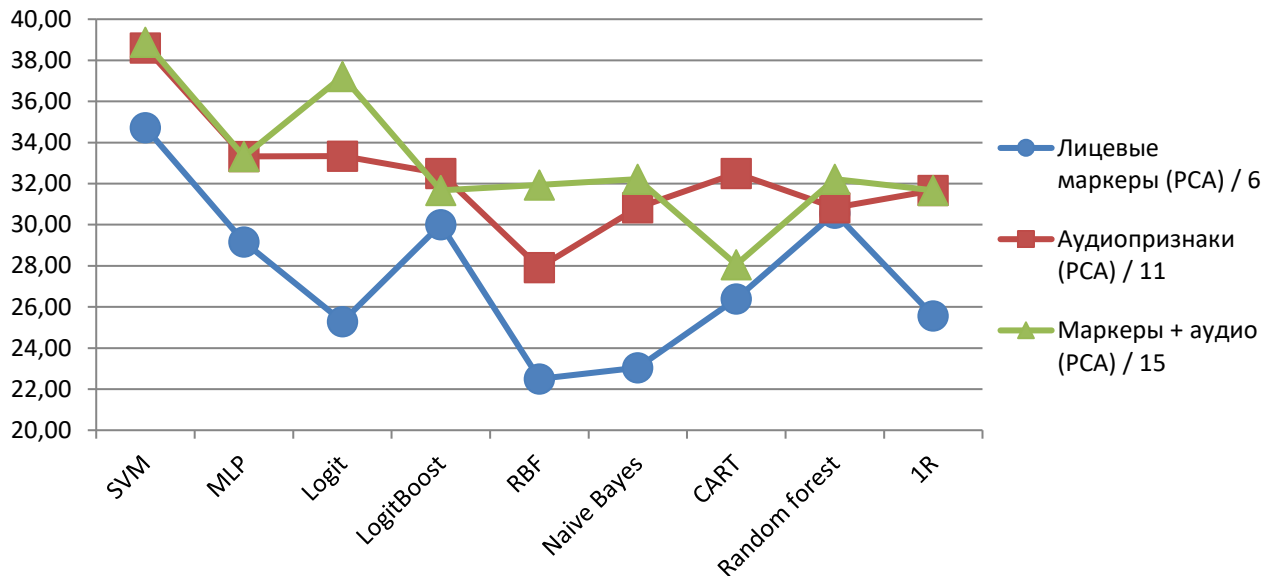


Рисунок 3.21. Точность классификации эмоций (%) базовыми алгоритмами классификации; со снижением размерности методом PCA; данные / число признаков

**Отбор информативных признаков.** Результаты экспериментов по снижению размерности и отбору информативных признаков представлены в таблицах Б9 - Б11 приложения Б.

По результатам видно, что многокритериальный подход к отбору признаков превзошел другие подходы по эффективности на всех трех выборках. Алгоритм SelfCOMOGA оказался наиболее эффективен на выборке аудио признаков и объединенной выборке аудио-видео признаков. Это особенно заметно на последней выборке, преимущество SelfCOMOGA на ней составило ~12%. В целом процедура снижения размерности оправдала себя - на всех трех выборках использование всех признаков без снижения размерности оказалось наименее эффективным.

**Проектирование ансамбля классификаторов.** Эксперименты по оптимизации параметров нейронных сетей проводились для каждой из имеющихся выборок данных (маркеры, аудио, маркеры + аудио). При этом выборки, подаваемые на вход нейронных сетей, включали в себя признаки,

отобранные соответствующим алгоритмом оптимизации на предыдущем этапе - этапе отбора информативных признаков.

Пример полученного множества Парето-оптимальных нейронных сетей с помощью алгоритма многокритериальной оптимизации VEGA, а также точность распознавания эмоций при объединении нейронных сетей в коллектив различными способами, представлены в таблице 3.3.

Заметим, что объединение Парето-оптимальных нейронных сетей, представленных в таблице 3.3 в коллектив увеличило точность классификации эмоций до 40.83%, в то время как наилучшая точность индивидуальной нейронной сети равна 37.50%.

Таблица 3.3. Парето-оптимальное множество нейронных сетей и результат объединения их в коллектив, многокритериальная оптимизация алгоритмом VEGA, данные - аудио признаки, произведен предварительный отбор признаков многокритериальным алгоритмом оптимизации VEGA

| №  | Число нейронов | Число итераций обучения | Точность классификации |
|----|----------------|-------------------------|------------------------|
| 1  | 27             | 259                     | 26.67                  |
| 2  | 11             | 259                     | 25.83                  |
| 3  | 11             | 360                     | 26.67                  |
| 4  | 14             | 368                     | <b>37.50</b>           |
| 5  | 21             | 250                     | 31.67                  |
| 6  | 16             | 255                     | 30.83                  |
| 7  | 11             | 340                     | 28.33                  |
| 8  | 10             | 399                     | 31.67                  |
| 9  | 10             | 293                     | 27.50                  |
| 10 | 10             | 368                     | 30.83                  |
| 11 | 25             | 293                     | 31.67                  |
| 12 | 12             | 255                     | 29.17                  |
| 13 | 11             | 259                     | 27.50                  |

Продолжение таблицы 3.3.

| <b>№</b>                        | <b>Число<br/>нейронов</b> | <b>Число<br/>итераций<br/>обучения</b> | <b>Точность<br/>классификации</b> |
|---------------------------------|---------------------------|----------------------------------------|-----------------------------------|
| 14                              | 12                        | 371                                    | 23.33                             |
| 15                              | 15                        | 302                                    | 26.67                             |
| 16                              | 10                        | 280                                    | 22.50                             |
| 17                              | 12                        | 371                                    | 32.50                             |
| 18                              | 11                        | 360                                    | 25.83                             |
| 19                              | 13                        | 265                                    | 28.33                             |
| 20                              | 14                        | 302                                    | 21.67                             |
| Голосование                     |                           |                                        | <b>39.17</b>                      |
| Усреднение вероятностей классов |                           |                                        | <b>37.50</b>                      |
| Мета-классификация              |                           |                                        | <b>40.83</b>                      |

В таблицах Б12 – Б14 приложения Б представлены результаты экспериментов по оптимизации параметров нейронных сетей и проектированию ансамбля нейронных сетей. Согласно полученным результатам, многокритериальный подход к оптимизации параметров нейронных сетей применительно к задаче распознавания эмоций превосходит однокритериальный подход на всех 3 рассмотренных выборках данных.

Так же как и при отборе информативных признаков, алгоритм SelfCOMOGA обеспечил наивысшую точность на выборке аудио признаков и на объединенной выборке лицевых маркеров и аудио признаков. Это можно объяснить тем, что на этапе отбора признаков алгоритм SelfCOMOGA нашел более удачный набор признаков. Мета-классификация по результатам экспериментов оказалась наиболее эффективным методом объединения коллектива классификаторов, обеспечив лучшую точность классификации на всех 3 выборках. Тот факт, что на разных выборках лучшие результаты были получены разными алгоритмами многокритериальной оптимизации, а именно алгоритмами NSGA-2 и SelfCOMOGA, но с одинаковым методом объединения Парето-

оптимальных классификаторов в коллектив, означает, что метод метаклассификации инвариантен к выбору алгоритма оптимизации. Усреднение вероятностей классов оказалось самым неэффективным методом слияния выходов коллектива нейронных сетей.

### **Выводы**

В данной главе представлен метод многокритериальной оптимизации для отбора информативных признаков и проектирования оптимального коллектива нейронных сетей. Предложенный метод апробирован на задаче распознавания эмоций, а также на ряде других тестовых задач классификации.

Было проведено сравнение эффективности многокритериального и однокритериального подхода к отбору признаков с методом главных компонент, а также с набором наиболее известных и используемых алгоритмов машинного обучения.

Согласно полученным результатам, многокритериальный подход к отбору признаков и проектированию коллектива нейронных сетей обеспечил лучшую точность классификации, чем однокритериальный подход. Было установлено, что многокритериальный и однокритериальный подходы к отбору признаков эффективнее метода главных компонент. В целом, применение процедуры снижения признакового пространства оправдало себя для всех рассмотренных практических задач, так как удалось снизить количество используемых признаков, и при этом улучшить точность классификации.

Среди использованных алгоритмов многокритериальной оптимизации, наиболее эффективным по критерию точности оказался алгоритм SelfCOMOGA. Этот результат в очередной раз подтверждает эффективность разработанного коэволюционного алгоритма, и целесообразность его использования для задач отбора информативных признаков.

Многокритериальный подход к формированию оптимального коллектива нейронных сетей также зарекомендовал себя. Полученные результаты доказывают, что объединение нейронных сетей с Парето-оптимальными настройками в коллектив позволяет получить лучшую точность классификации, чем использование единичной нейронной сети с оптимальными настройками, найденными с помощью однокритериального алгоритма оптимизации. По результатам экспериментов мета-классификация с использованием метода опорных векторов была признана самым эффективным способом объединения Парето-оптимальных классификаторов в коллектив из числа рассмотренных. В дальнейшем, требуется провести апробацию данного подхода для задач проектирования коллективов других алгоритмов интеллектуального анализа данных.

В следующей главе будет описан новый, гибридный алгоритм обучения конволюционной нейронной сети, основанный на последовательной работе эволюционного алгоритма оптимизации, и алгоритма обратного распространения ошибки. Конволюционная нейронная сеть, обученная с помощью этого алгоритма, будет протестирована на все той же задаче распознавания эмоций, однако в этот раз на вход сети будут подаваться непосредственно изображения лиц дикторов.

## Глава 4. Разработка гибридного алгоритма обучения конволюционной нейронной сети с применением эволюционного алгоритма оптимизации

### § 4.1 Конволюционная нейронная сеть и суть методов глубинного обучения

Методы глубинного обучения предназначены для распознавания иерархически организованных признаков (рисунок 4.1). Это означает, что признаки высокого уровня формируются из признаков низкого уровня. Например, в задачах анализа изображений признаками низкого уровня могут быть значения интенсивностей отдельно взятых пикселей, признаками более высокого уровня - горизонтальные и вертикальные линии, признаками еще более высокого уровня - более сложные формы и отдельные объекты, представленные на изображении.

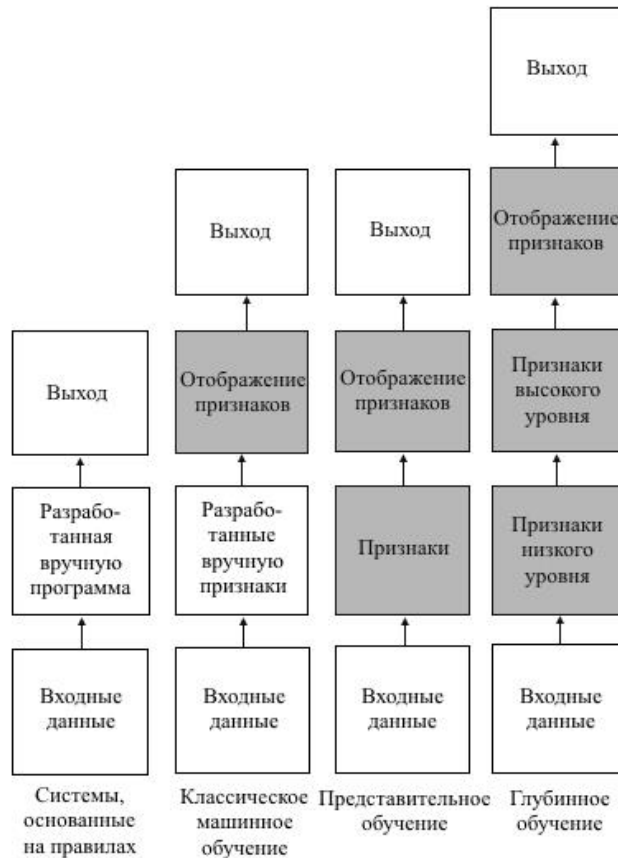


Рисунок 4.1. Типы обучения

Автоматическое распознавание признаков на разных уровнях иерархии позволяет производить классификацию сложных объектов, не прибегая к ручному формированию признаков экспертами предметной области. Это особенно важно при распознавании высокоуровневых объектов, которые зачастую не могут быть явно выражены человеком-экспертом через композицию входных низкоуровневых признаков. Например, в задаче распознавания объекта "дом" по изображению необходимо знать, из каких частей состоит дом. Традиционно в человеческом понимании дом должен иметь окна, крышу, вертикальные стены, однако существует много разновидностей домов, поэтому однозначно выделить признаки, характеризующие дом, не получается. Концепция глубинного обучения состоит в создании систем, способных самообучаться автоматическому распознаванию признаков, пригодных для классификации объектов сложной структуры, таких как "дом", благодаря совместному использованию алгоритмов обучения с учителем и без учителя.

Структурная глубина алгоритма выражается количеством уровней нелинейных преобразований в восстанавливаемой функции. Тогда как большинство современных алгоритмов обучения относятся к неглубоким структурам, состоящим из 1, 2 или 3 таких уровней, мозг млекопитающих имеет глубокую структуру [128], включающую в себя несколько уровней абстракции, причем каждому уровню соответствует своя область коры головного мозга. Мозг обрабатывает информацию в несколько этапов: вначале находит четко выраженные линии, соответствующие краям объектов, затем выделяет примитивные формы, из которых составляет более сложные объекты.

Исследователи в области нейронных сетей на протяжении нескольких десятилетий пытались разработать и обучить глубокие нейронные сети, которые по своему принципу были бы схожи с человеческим мозгом [43, 140], но все эти попытки вплоть до 2006 года были неудачны. Успешными оказывались лишь нейронные сети с 1-2 слоями, более глубокие структуры не давали



удовлетворительных результатов, в основном потому, что не было эффективного алгоритма для их обучения. В 2006 году такой алгоритм был предложен Дж. Хинтоном, разработавшим Сети Глубокого Доверия (Deep Belief Networks, DBN) [83], в которых для обучения каждого слоя по отдельности был использован жадный алгоритм обучения без учителя, Ограниченная Машина Больцмана (Restricted Boltzmann Machine, RBM) [71]. После этого были также предложены алгоритмы-автоассоциаторы [42, 113], в которых использовался тот же принцип: обучение промежуточных слоев нейронной сети по отдельности с применением алгоритма обучения без учителя.

С 2006 года глубокие сети успешно применяются для решения практических задач классификации [34], восстановления регрессии [122], снижения размерности, выделения объектов, в робототехнике, для обработки естественного языка [102]. Несмотря на то, что алгоритмы RBM и DBN могут обучаться на неразмеченных данных (для которых априори неизвестна принадлежность к тому или иному классу), оба эти алгоритма успешно применяются для инициализации весов нейронных сетей, в которых используется алгоритм обучения с учителем.

Так как сети с глубокой структурой представляют собой композицию последовательных этапов обработки данных, важным моментом является то, как эти этапы связываются между собой, иными словами, что является выходом текущего слоя и входом следующего. Множество современных исследований посвящено разработке эффективных алгоритмов обучения, способных выстраивать промежуточные представления данных. К успешным алгоритмам можно отнести RBM [83], автоассоциаторы [42] и разреженные автоассоциаторы (sparse autoencoders) [111]. Эти алгоритмы трансформируют одно представление данных, являющееся выходом предыдущего слоя, в другое, при этом снижая вариативность данных, то есть количество признаков и их взаимозависимость. Поиск хорошей формы представления входных данных на каждом слое глубокой нейронной сети является наиболее сложным этапом ее обучения. Как правило,

если удастся успешно выполнить этот этап, дальше дело остается за малым - инициализировать сеть найденными весами и дообучить ее с помощью градиентного алгоритма оптимизации, например алгоритма обратного распространения ошибки.

Почти все глубокие нейронные сети очень сложно обучать без предварительного обучения без учителя, за исключением конволюционных нейронных сетей (convolutional neural networks, CNN, далее КНС). КНС были созданы по подобию зрительной системы человеческого мозга. Первые расчетные модели, основанные на локальных связях между нейронами и иерархических трансформациях изображений, были представлены в Неокогнитроне Фукушимы [75]. Позднее, Ян Лекун со своими коллегами, развивая эту идею, разработали и обучили КНС, которые показали высокую эффективность при решении некоторых задач распознавания образов [94, 95]. Современные знания о том, как устроена и функционирует зрительная система человека, схожи с принципом работы КНС. Системы распознавания образов, основанные на КНС, на сегодняшний день одни из точнейших, особенно в задачах, связанных с анализом изображений. Ярким примером является ставшая уже классической задача распознавания рукописных цифр [94], которая уже много лет служит тестовой задачей для сравнения алгоритмов распознавания образов.

Что касается проблемы обучения сетей глубокой архитектуры, КНС интересны в данном аспекте тем, что обычно они состоят из 5, 6, либо 7 слоев [131]. Полносвязные нейронные сети с таким количеством слоев почти невозможно обучить, используя случайную начальную инициализацию весов, однако к КНС данное ограничение не относится. Отсюда возникает вопрос, что такого особенного в структуре КНС, что позволяет им достигать хорошей обобщающей способности при анализе и классификации изображений.

КНС, предложенная Лекуном, включает в себя слои двух типов: конволюционные слои (convolutional), или слои свертки, и слои субдискретизации (subsampling). Каждый слой имеет топографическую структуру, то есть каждый

нейрон связан с фиксированным куском входного изображения, который влияет на выход нейрона. На каждом слое находятся нейроны с различным набором входных весов, связанных с нейронами на прямоугольном куске предыдущего слоя.

Существует гипотеза, что малое число входных воздействий нейрона позволяет градиентам проходить через большое количество слоев КНС, при этом не рассеиваясь слишком сильно, что сделало бы их бесполезными. Однако один этот факт не может быть причиной успеха КНС, так как случайной разреженной связности недостаточно для успешного обучения глубоких нейронных сетей. Другая гипотеза успеха данного типа сетей состоит в том, что структура локальной связности, по всей видимости, очень хорошо подходит именно для анализа изображений, устанавливая начальные веса сети в пространстве весов вблизи глобального оптимума, так что градиентная оптимизационная процедура срабатывает. Даже при случайных весах на начальных слоях КНС справляется с тестовыми задачами лучше, чем полностью обученная полносвязная нейронная сеть, но хуже, чем полностью обученная КНС [112].

Не так давно структура КНС была перенесена также на алгоритмы RBM (Restricted Boltzmann Machines) и DBN (Deep Belief Network) [66, 97]. Важным новшеством выступило создание порождающих элементов объединения и субдискретизации, которые прекрасно справились с тестовой задачей классификации рукописных цифр из базы MNIST, и с задачей классификации объектов Caltech-101. Кроме того, визуализация признаков, которые научились распознавать нейроны на разных слоях сети, продемонстрировала состоятельность самой идеи глубинного обучения. Идея эта состоит в последовательном распознавании объектов различного уровня сложности, начиная с простых линий, двигаясь далее к частям объекта, и затем к объекту целиком.

## § 4.2 Достоинства и недостатки алгоритма обратного распространения ошибки и эволюционного алгоритма

Алгоритм обратного распространения ошибки (ОРО) изначально использовался для настройки весов перцептрона. Впоследствии он был перенесен на многослойные нейронные сети прямого распространения сигнала, КНС и многие другие виды сетей [120]. Алгоритм ОРО является градиентным методом оптимизации. В основе его функционирования лежит аппроксимация градиента ошибки относительно изменения весов сети. Градиент задает направление изменения каждого веса сети таким образом, чтобы минимизировать ошибку выхода сети.

Ниже представлена схема алгоритма обратного распространения ошибки применительно к обучению КНС.

### Алгоритм обратного распространения ошибки

#### ***Вход:***

1. КНС заданной структуры;
2. Обучающая выборка  $(x_i, y_i), i = \overline{1, n}$ ;
3. Параметры сети: размер подвыборки  $m$ , число эпох (итераций)  $t$ , шаг обучения  $\alpha$ .

#### ***Инициализация:***

1. Случайно сгенерировать веса ядер свертки на каждом конволюционном слое и веса последнего полносвязного слоя.
2. Рассчитать количество подвыборок  $n_p = n/m$ .

#### ***Основной цикл алгоритма:***

На каждой итерации (эпохе) алгоритма  $t_j$ :

Случайно поделить выборку на  $n_p$  подвыборок.

Для каждой подвыборки  $n_{pk}$ :

1. Произвести прогон подвыборки через сеть:

1.1. В конволюционных слоях произвести операцию свертки выхода предыдущего слоя с каждым из ядер.

1.2. В слоях подвыборки произвести операцию объединения

1.3. Вычислить активационную функцию последнего полносвязного слоя и получить выход сети.

2. Запустить процедуру обратного распространения ошибки, аппроксимирующую градиенты:

2.1. Вычислить ошибку между выходом сети и истинным значением из выборки.

2.2. Пройти по слоям сети в обратном порядке и вычислить конечную аппроксимацию градиента ошибки от каждого из весов сети; весами сети являются ядра в конволюционных слоях и веса полносвязного слоя.

3. Обновить веса сети в соответствии с вычисленным градиентом.

**Выход:** Обученная КНС.

Алгоритм обратного распространения ошибки ищет минимум функции ошибки в пространстве весов сети, используя метод градиентного спуска. В качестве нелинейной активационной функции на последнем полносвязном слое чаще всего используется сигмоида, имеющая следующий вид:

$$\text{sigm}(x) = \frac{1}{1+e^{-cx}}, \quad (4.1)$$

где  $c$  - параметр сигмоиды, причем при  $c \rightarrow \infty$   $\text{sigm}(x)$  становится ступенчатой функцией в точке  $x = 0$  (рисунок 4.2).

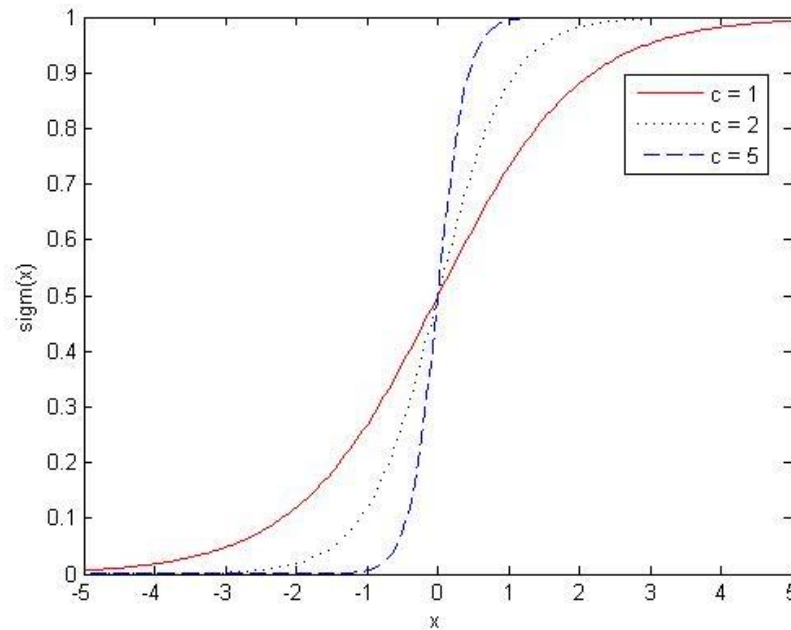


Рисунок 4.2. График сигмоидной функции при различных значениях параметра  $c$

Производная сигмоиды по  $x$  при  $c = 1$  равна:

$$\frac{d}{dx} \text{sigm}(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \text{sigm}(x)(1 - \text{sigm}(x)). \quad (4.2)$$

Данная формула используется в пункте 2.2 алгоритма обратного распространения ошибки для расчета аппроксимации градиента ошибки от весов сети.

Преимущества данного алгоритма состоят в его концептуальной и вычислительной простоте, что особенно важно для КНС, так как количество параметров в такой сети может достигать нескольких сотен тысяч и даже миллионов.

Недостаток алгоритма заключается в использовании градиентного спуска для поиска минимума ошибки в пространстве весов сети. Градиентный спуск не является методом глобальной оптимизации, и зачастую застревает в локальных оптимумах. В этой ситуации исход оптимизации полностью зависит от начальной инициализации весов сети. Если вектор начальных весов в пространстве весов сети окажется недостаточно "близко" к глобальному оптимуму, то процедура градиентного спуска закончится в одном из локальных оптимумов. Учитывая, что начальная инициализация весов КНС проводится случайным образом, а

многомерное пространство весов сети имеет очень сложную структуру, вероятность достижения глобального оптимума очень низкая.

В связи с этим есть смысл заменить алгоритм обратного распространения ошибки алгоритмом глобальной оптимизации. Одним из таких алгоритмов является генетический алгоритм. Класс генетических алгоритмов уже был описан в § 2.1. Здесь будут описаны их достоинства и недостатки.

Эволюционные алгоритмы отличаются от классических методов оптимизации по следующим пунктам [77]:

1. В эволюционных алгоритмах в процессе поиска не используется информация о градиенте оптимизируемой функции. Следовательно, будучи методами прямого поиска, эволюционные алгоритмы могут быть применены к решению широкого круга задач.

2. В эволюционных алгоритмах используется популяционный подход, при котором на каждой итерации алгоритм находит более одного решения. В классических алгоритмах оптимизации одно решение обновляется на каждой итерации. Использование популяционного подхода обладает несколькими преимуществами:

- а. Эволюционный алгоритм находит несколько оптимальных решений, что в свою очередь является плюсом при решении задач мультимодальной и многокритериальной оптимизации.

- б. Эволюционным алгоритмам присуща параллельная вычислительная структура, идеально подходящая для использования параллельных вычислений, значительно ускоряющих процедуру поиска.

- в. В популяции существует возможность сравнивать решения по оптимизируемому критерию, отдавая предпочтение решениям с лучшей пригодностью (суть процесса эволюции), а также применять другие операции, например элитизм.

3. В эволюционных алгоритмах используются стохастические операторы, тогда как в большинстве классических алгоритмов оптимизации

используются детерминированные операторы. Локальная стохастичность в сочетании с глобальной направленностью поиска в сторону наиболее пригодных решений предотвращает алгоритм от застревания в локальных оптимумах при решении сложных задач оптимизации.

Главный недостаток эволюционных алгоритмов состоит в том, что, будучи алгоритмами глобального поиска, они не гарантируют сходимости к глобальному оптимуму. Поэтому при обучении нейронных сетей эволюционные алгоритмы используются для поиска стартовых точек, из которых ведется дальнейшее обучение с помощью алгоритма ОРО. Несмотря на существование некоторых обоснований в пользу эффективности эволюционных алгоритмов [84], до сих пор не предложено строгого математического доказательства их сходимости. Второй недостаток связан с высокой вычислительной сложностью. При решении сложных задач оптимизации может потребоваться несколько сотен итераций поиска при размере популяции также в несколько сотен, то есть на каждой итерации целевая функция вычисляется несколько сотен раз. Все это ведет к значительным вычислительным и временным затратам. Также к недостаткам эволюционных алгоритмов можно отнести проблему стагнации [148], или преждевременной сходимости, при которой в ходе процесса оптимизации популяция оказывается полностью заполнена одинаковыми решениями.

Успех эволюционных алгоритмов во многом зависит от тонкой настройки их параметров [35, 59]. По этой причине была разработана самонастраивающаяся модификация стандартного генетического алгоритма - коэволюционный генетический алгоритм, в котором не требуется ручная настройка параметров [106, 141].



### § 4.3 Разработка и реализация гибридного алгоритма обучения конволюционной нейронной сети

В данном разделе описывается новый, гибридный алгоритм обучения конволюционной нейронной сети (КНС), основанный на алгоритме обратного распространения ошибки и коэволюционном генетическом алгоритме оптимизации.

Для обучения использовалась КНС фиксированной структуры, вариант которой представлен на рисунке 4.3. Макропараметры сети, такие как количество и размерность ядер свертки каждого конволюционного слоя, и масштаб уменьшения входного изображения на каждом слое субдискретизации выбирались вручную в зависимости от размера входных изображений. Размерность полносвязного слоя рассчитывалась автоматически в зависимости от размера изображения на предыдущем слое. Ручной выбор параметров сети связан с тем, что в процессе прохождения через сеть с входным изображением производится ряд операций по его разбиению на фрагменты и масштабированию. В процессе выполнения этих операций может возникнуть ситуация, когда изображение будет иметь нецелый пиксельный размер, например  $23.5 \times 23.5$ , что приведет к сбою работы сети. Поэтому необходимо заранее выбирать такие параметры сети, которые не приводили бы к подобной ситуации.

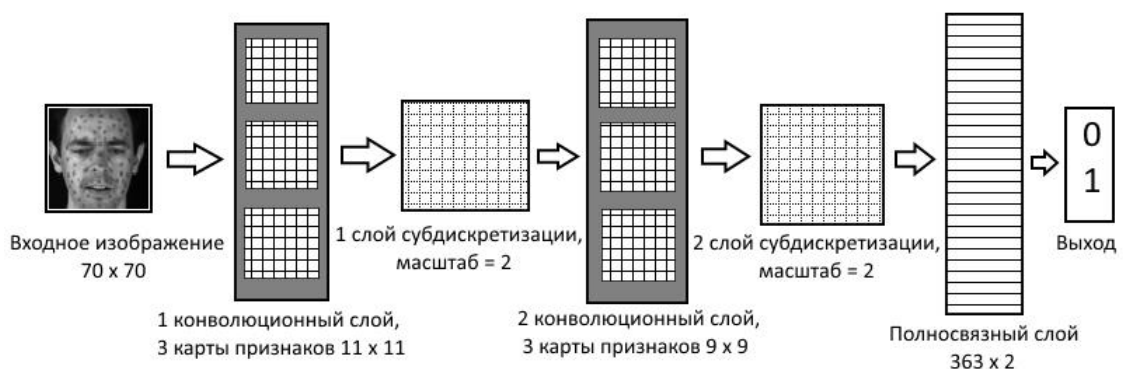


Рисунок 4.3. Вариант структуры конволюционной нейронной сети

Предлагаемый гибридный алгоритм обучения КНС состоит из двух этапов (рисунок 4.4). На первом этапе происходит оптимизация весов сети с помощью коэволюционного генетического алгоритма оптимизации. В качестве индивида в ГА выступает КНС с определенным набором весов. Количество входных переменных ГА (весов сети) в случае задачи распознавания эмоций равно ~500-1000 в зависимости от того, изображения какого размера подаются на вход сети. Применение ГА на 1 этапе обучения сети позволяет более полно охватить пространство поиска и избежать стагнации. Задача ГА - в общем пространстве поиска найти подобласть потенциального оптимума, чтобы на втором этапе с помощью градиентного спуска найти сам оптимум.

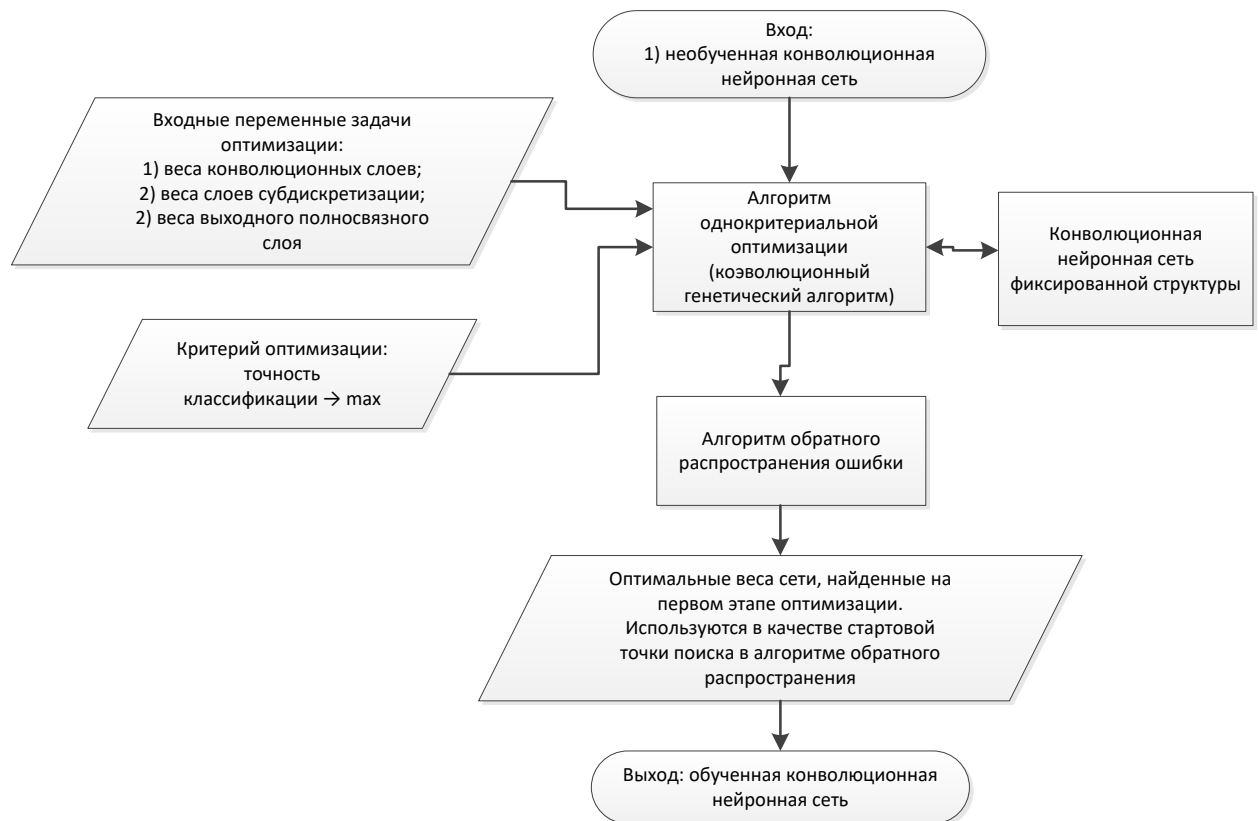


Рисунок 4.4. Схема гибридного алгоритма обучения конволюционной нейронной сети

Важный момент - выбор оптимизируемого критерия для генетического алгоритма. Так как каждый индивид ГА - это КНС, а в качестве целевой функции

выступает задача классификации, то естественно выбрать точность классификации максимизируемым критерием (или ошибку классификации минимизируемым критерием, соответственно). Изначально в ходе разработки данного гибридного алгоритма так и было сделано, однако точность классификации оказалась неудовлетворительным критерием, так как ГА в большинстве экспериментов впадал в стагнацию, то есть все КНС - индивиды популяции, спустя некоторое число итераций оказывались одинаковыми, что прекращало процесс оптимизации. В связи с этим был использован другой критерий оптимизации - усредненная F-мера, которая для случая двух классов вычисляется по формуле (4.3).

$$F = 2 \times \frac{\text{точность} \times \text{полнота}}{\text{точность} + \text{полнота}} \quad (4.3)$$

Точность (precision) и полнота (recall) вычисляются по формулам (4.4) и (4.5), понятия TP (true-positive, истинно-положительное решение), TN (true-negative, истинно-отрицательное решение), FP (false-positive, ложно-положительное решение) и FN (false-negative, ложно-отрицательное решение) проиллюстрированы в таблице 4.1.

$$\text{Точность} = \frac{TP}{TP + FP} \quad (4.4)$$

$$\text{Полнота} = \frac{TP}{TP + FN} \quad (4.5)$$

Таблица 4.1. Таблица сопряженности, иллюстрирующая понятия точности и полноты

| Класс <i>i</i> |   | Истина |    |
|----------------|---|--------|----|
|                |   | 1      | 0  |
| Оценка         | 1 | TP     | FP |
|                | 0 | FN     | TN |

F-мера более адекватно отражает способность системы классификации к обобщению, то есть классификации объектов, не использовавшихся при обучении системы. Формула (4.3) используется для вычисления F-меры каждого класса в задаче с двумя классами. В случае большего числа классов задачу всегда можно свести к двум классам, используя схему "один против всех" (one-vs-all), в которой все классы кроме 1-го заменяются на один класс - не-первый, затем все классы кроме 2-го заменяются на не-второй класс, и т.д.

Для ускорения работы генетического алгоритма при вычислении F-меры каждой сети в популяции на ее вход подавалась не вся обучающая выборка, а лишь часть.

В конце первого этапа из конечной популяции генетического алгоритма выбиралась КНС, F-мера которой была максимальной. Веса выбранной сети, оптимизированные генетическим алгоритмом, брались в качестве начального приближения для второго этапа обучения. На втором этапе обучение выбранной сети продолжалось с помощью традиционного алгоритма ОРО. Так как данный алгоритм является градиентным, идея его использования после начального приближения с помощью ГА состоит в том, что он должен в суженной с помощью ГА области поиска найти глобальный оптимум. Так как алгоритм ОРО в качестве начального приближения использует случайно сгенерированные значения весов КНС, он склонен застревать в локальных оптимумах. Получение начального приближения весов с помощью генетического алгоритма призвано решить эту проблему.

#### **§ 4.4 Исследование эффективности гибридного алгоритма обучения конволюционной нейронной сети на задачах анализа изображений**

В данном параграфе приведены результаты численных экспериментов на множестве практических задач, являющихся тестовыми для сравнения эффективности конволюционных сетей, обученных алгоритмом обратного

распространения ошибки и разработанным гибридным алгоритмом, включая задачу распознавания рукописных цифр, а также задачу распознавания объектов на изображениях. Более детальные эксперименты и анализ представлены для поставленной в диссертации задачи распознавания эмоций.

**Задача распознавания рукописных цифр (MNIST).** Исследование эффективности гибридного алгоритма обучения проводилось на задаче распознавания рукописных цифр, исходные данные для которой были взяты из репозитория MNIST (Mixed National Institute of Standards and Technology). Репозиторий MNIST представляет собой обширный набор изображений образцов рукописных цифр от 0 до 9, размер каждого изображения составляет  $28 \times 28$  пикселей. Данный репозиторий был предложен Национальным институтом стандартов и технологий США для тестирования и сравнения эффективности различных алгоритмов машинного обучения и распознавания образов. Репозиторий MNIST включает в себя данные для 5 различных задач, связанных с распознаванием рукописных цифр:

1. MNIST - рукописные цифры от 0 до 9 на черном фоне;
2. MNIST rotated - повернутые на случайный угол изображения из базы MNIST;
3. MNIST random background - MNIST, наложенные поверх случайно сформированного фона (белого шума);
4. MNIST background image - MNIST, наложенные поверх фрагментов картинок из ограниченного набора;
5. MNIST rotated background image - повернутые на случайный угол изображения MNIST, наложенные поверх фрагментов картинок.

В экспериментах была использована КНС с 2 конволюционными слоями, 2 слоями субдискретизации и 1 полносвязным слоем на выходе. В таблице 4.2 представлены использованные в экспериментах значения параметров КНС. На первом этапе предложенного гибридного алгоритма используется

коэволюционный генетический алгоритм оптимизации (ГА), его параметры представлены в таблице 4.3.

Таблица 4.2. Значения параметров конволюционной нейронной сети для задач распознавания рукописных цифр

| Параметр                                            | Значение     |
|-----------------------------------------------------|--------------|
| Число карт признаков на 1 конволюционном слое       | 6            |
| Число карт признаков на 2 конволюционном слое       | 12           |
| Размерность ядра свертки 1 и 2 конволюционного слоя | $5 \times 5$ |
| Масштаб слоев субдискретизации                      | 2            |
| Число нейронов полносвязного слоя                   | 192          |
| Число эпох обучения                                 | 50           |

Таблица 4.3. Параметры коэволюционного генетического алгоритма

| Параметр                             | Значение  |
|--------------------------------------|-----------|
| Размер популяции                     | 50        |
| Число итераций                       | 50        |
| Интервал поиска по каждой переменной | $[-3; 2]$ |
| Точность разбиения интервала поиска  | 0.0001    |
| Процент элитарных решений            | 4%        |

Сравнение эффективности проводилось между КНС, обученными алгоритмом обратного распространения ошибки (ОРО), и гибридным алгоритмом (ГА + ОРО). КНС, обученная с помощью только генетического алгоритма, ощутимой эффективности не дала по причине слишком большой размерности пространства поиска, поэтому результаты для этого случая были опущены.

Сравниваемые критерии - точность классификации и F-мера. Критерий F-меры был выбран потому, что он более адекватно оценивает качество классификации при классах, неоднородно представленных в выборке (число объектов одного класса существенно больше числа объектов другого класса).

Результаты эффективности подходов для 5 вышеописанных задач семейства MNIST по критерию точности классификации представлены в таблице 4.4, по критерию F-меры - в таблице 4.5.

Таблица 4.4. Сравнение средней точности (%) распознавания рукописных цифр по итерациям обучения в конволюционной нейронной сети: а) обратное распространение ошибки (ОРО); б) гибридный алгоритм (ГА + ОРО)

| Данные                         | Алгоритм обучения КНС |          |       |       |             |          |       |       |
|--------------------------------|-----------------------|----------|-------|-------|-------------|----------|-------|-------|
|                                | а) ОРО                |          |       |       | б) ГА + ОРО |          |       |       |
|                                | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| MNIST                          | 92.05                 | 3.23     | 87.72 | 95.44 | 95.82       | 2.14     | 92.93 | 98.17 |
| MNIST rotated                  | 70.17                 | 4.12     | 62.50 | 76.19 | 74.61       | 2.65     | 69.24 | 81.94 |
| MNIST random background        | 81.29                 | 3.65     | 76.33 | 90.18 | 85.39       | 3.11     | 81.07 | 88.96 |
| MNIST background image         | 75.12                 | 5.33     | 64.94 | 81.74 | 70.07       | 3.56     | 65.48 | 75.68 |
| MNIST rotated background image | 38.76                 | 3.45     | 32.75 | 42.42 | 41.26       | 2.41     | 38.61 | 46.26 |

Таблица 4.5. Сравнение средней F-меры распознавания рукописных цифр по итерациям обучения: а) ОРО; б) ГА + ОРО

| Данные                         | Алгоритм обучения КНС |          |       |       |             |          |       |       |
|--------------------------------|-----------------------|----------|-------|-------|-------------|----------|-------|-------|
|                                | а) ОРО                |          |       |       | б) ГА + ОРО |          |       |       |
|                                | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| MNIST                          | 0.917                 | 0.025    | 0.901 | 0.948 | 0.952       | 0.017    | 0.913 | 0.975 |
| MNIST rotated                  | 0.681                 | 0.043    | 0.657 | 0.717 | 0.742       | 0.035    | 0.675 | 0.808 |
| MNIST random background        | 0.804                 | 0.076    | 0.706 | 0.896 | 0.849       | 0.058    | 0.717 | 0.879 |
| MNIST background image         | 0.742                 | 0.034    | 0.720 | 0.800 | 0.711       | 0.021    | 0.691 | 0.736 |
| MNIST rotated background image | 0.346                 | 0.056    | 0.259 | 0.430 | 0.368       | 0.032    | 0.310 | 0.409 |

Согласно полученным эмпирическим результатам, КНС, обученная с помощью гибридного алгоритма, справилась лучше, чем алгоритм обратного распространения ошибки, на 4 из 5 рассмотренных задач.

Далее будет проведена апробация предложенного гибридного алгоритма обучения на задаче распознавания эмоций SAVEE.

**Задача распознавания объектов (CIFAR-10).** Репозиторий CIFAR-10 (Canadian Institute for Advanced Research) включает в себя цветные изображения размера 32×32 пикселей, на которых изображены объекты, принадлежащие одному из 10 взаимоисключающих классов: самолет, автомобиль, птица, кошка, олень, собака, лягушка, лошадь, корабль, грузовик (рисунок 4.5).



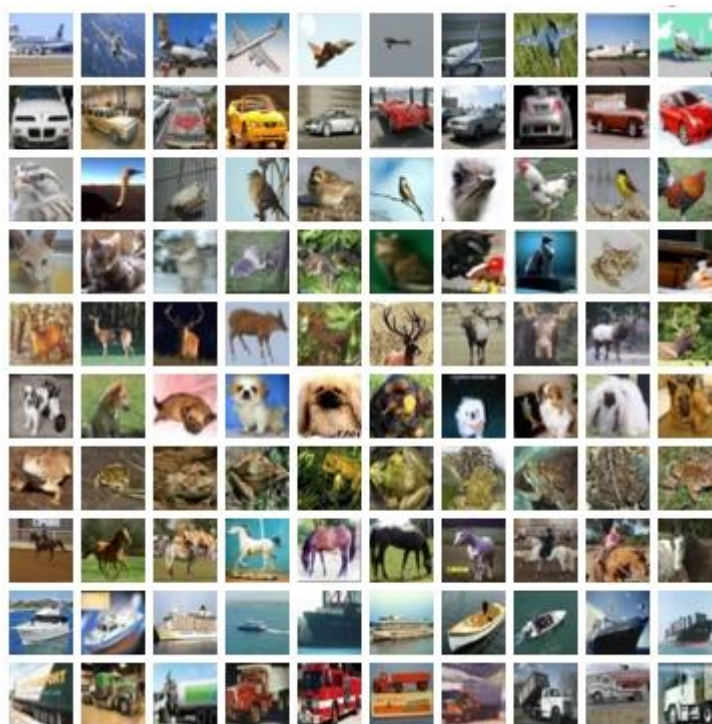


Рисунок 4.5. Изображения из репозитория CIFAR-10

В целях тестирования предложенного гибридного алгоритма обучения конволюционной нейронной сети исходные цветные изображения были преобразованы в изображения с оттенками серого. В обучающую выборку вошли 700 изображений, в тестовую - 300 изображений, каждому из классов соответствует примерно одинаковое количество объектов выборок.

Результаты тестирования алгоритма обратного распространения ошибки и гибридного алгоритма на данной задаче представлены в таблице 4.6.

Таблица 4.6. Сравнение средней точности (%) и F-меры для задачи CIFAR-10 по итерациям обучения в конволюционной нейронной сети: а) обратное распространение ошибки (ОРО); б) гибридный алгоритм (ГА + ОРО)

| Критерий    | Алгоритм обучения КНС |          |       |       |             |          |       |       |
|-------------|-----------------------|----------|-------|-------|-------------|----------|-------|-------|
|             | а) ОРО                |          |       |       | а) ГА + ОРО |          |       |       |
|             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Точность, % | 76.86                 | 4.26     | 71.71 | 80.82 | 83.29       | 3.98     | 78.91 | 89.13 |
| F-мера      | 0.764                 | 0.032    | 0.719 | 0.813 | 0.831       | 0.047    | 0.784 | 0.891 |

**Задача распознавания эмоций (SAVEE).** На вход КНС подавались изображения лиц дикторов, изображающих различные эмоции, размер подаваемых изображений составлял  $40 \times 40$  пикселей,  $50 \times 50$  и  $70 \times 70$ . Для каждого размера входных изображений были использованы различные значения параметров КНС, они представлены в таблице 4.7, значения параметров коэволюционного генетического алгоритма были использованы те же, что и для задач MNIST.

Таблица 4.7. Значения параметров конволюционной нейронной сети для различных размеров входных изображений в задаче распознавания эмоций

| Параметр                                        | Размер входных изображений |                |                |
|-------------------------------------------------|----------------------------|----------------|----------------|
|                                                 | $40 \times 40$             | $50 \times 50$ | $70 \times 70$ |
| Число карт признаков на конволюционных слоях    | 3                          | 3              | 3              |
| Размерность ядра свертки 1 конволюционного слоя | $7 \times 7$               | $9 \times 9$   | $11 \times 11$ |
| Размерность ядра свертки 2 конволюционного слоя | $6 \times 6$               | $6 \times 6$   | $9 \times 9$   |
| Масштаб слоев субдискретизации                  | 2                          | 2              | 2              |
| Число нейронов полносвязного слоя               | 108                        | 192            | 363            |

Сравнение эффективности проводилось между КНС, обученными с помощью алгоритма обратного распространения ошибки (ОРО),

коэволюционного генетического алгоритма (ГА) и гибридного алгоритма (ГА + ОРО).

Задача распознавания эмоций была сформулирована как задача классификации один против всех для каждого типа эмоций - например, класс "счастье" и класс "не счастье", в который входят объекты выборки, принадлежащие ко всем остальным классам, кроме "счастья".

Результаты сравнения средней точности распознавания эмоций по итерациям обучения в дикторозависимой постановке представлены в таблицах В1 – В3 приложения В, средней F-меры - в таблицах В4 – В6 приложения В. Аналогичные результаты для дикторонезависимой постановки представлены в таблицах В7 – В12 приложения В.

Из представленных результатов видно, что ГА в рассмотренных постановках задачи классификации проигрывает по критериям эффективности алгоритму ОРО. Гибридный же алгоритм в среднем превышает по эффективности алгоритм ОРО.

Полученные результаты указывают на то, что гибридный алгоритм обучения КНС, сочетающий в себе эволюционную стратегию и градиентный спуск, эффективнее алгоритма ОРО на задаче распознавания эмоций.

Генетический алгоритм не показал хорошей сходимости к глобальному оптимуму при оптимизации параметров КНС. Вероятная причина этого - слишком большая размерность пространства поиска (~500 - 1000 переменных). Задача ГА в разработанном гибридном алгоритме состоит в поиске подобласти в пространстве входных переменных, в которой находится глобальный оптимум, и со своей задачей ГА справляется. Однако очевидно, что для поиска точного расположения глобального оптимума генетический алгоритм в данной постановке задачи непригоден, так как для этого потребуются неоправданно высокие вычислительные затраты. Поэтому необходимо использовать процедуру градиентного спуска, либо иную процедуру, для поиска глобального оптимума в выделенной подобласти.

## Выводы

Конволюционные нейронные сети с успехом применяются для решения задач распознавания шаблонов в изображениях на протяжении многих лет. Стандартным алгоритмом обучения таких сетей является алгоритм обратного распространения ошибки, являющийся по сути алгоритмом градиентного спуска. Недостатком алгоритмов градиентного спуска является тот факт, что они сходятся к глобальному экстремуму, только если начальная точка поиска находится вблизи этого глобального экстремума.

В данной главе описан гибридный алгоритм обучения КНС, призванный решить вышеописанную проблему. Разработанный алгоритм состоит из двух этапов: поиск подобласти глобального экстремума генетическим алгоритмом, и градиентный спуск к глобальному экстремуму в этой подобласти с помощью алгоритма обратного распространения ошибки.

Было произведено сравнение разработанного гибридного алгоритма с алгоритмом обратного распространения ошибки и с генетическим алгоритмом на задаче распознавания эмоций (SAVEE), на задаче распознавания рукописных цифр (MNIST), а также на задаче распознавания объектов на изображениях (CIFAR-10). Разработанный гибридный алгоритм превзошел по эффективности стандартный алгоритм обратного распространения ошибки как по критерию усредненной точности классификации, так и по критерию усредненной F-меры. Генетический алгоритм, работающий отдельно от алгоритма обратного распространения ошибки, оказался неэффективен ввиду слишком высокой размерности пространства поиска (большое число весов КНС).

В следующей главе будет описан обобщенный метод для решения задач машинного обучения, включающих в себя гетерогенные данные. Данный метод будет основан на алгоритмах, разработанных в данной работе и описанных в главах 2-4, а именно: коэволюционный алгоритм многокритериальной оптимизации SelfCOMOGA, многокритериальный подход к отбору признаков и

проектированию оптимального коллектива нейронных сетей, КНС с гибридным алгоритмом обучения. Разработанный обобщенный метод будет апробирован на задаче распознавания эмоций.

Кроме того, будет описан метод слияния гетерогенной информации на уровне данных и на уровне классификаторов, который также будет апробирован на задаче распознавания эмоций.

## **Глава 5. Разработка обобщенного метода для решения задач анализа гетерогенных данных на основе разработанных алгоритмов**

### **§ 5.1 Разработка и исследование эффективности метода слияния аудио-видео информации на уровне данных и на уровне классификаторов в рамках задачи распознавания эмоций**

Задача распознавания эмоций, являющаяся частью более общей задачи человеко-машинного взаимодействия, ранее уже решалась другими исследователями. Распространенный подход - использовать аудио-информацию в качестве исходных данных. Из сырых данных извлекаются количественные признаки, такие как громкость голоса (амплитуда), тембр и пр. Признаки можно извлекать с различной степенью дискретизации, например каждую секунду, или каждые 5 секунд, и усреднять значения по текущему промежутку времени.

Другой подход заключается в использовании видео-информации. Лицо диктора записывается камерой, и отдельные кадры видеозаписи можно использовать для анализа. Опять же, можно либо брать каждый кадр видеопоследовательности, либо брать кадры с некоторой частотой, например каждый 10 кадр. Важным пунктом является извлечение из видеопоследовательности информативных количественных признаков, то есть признаков, релевантных с точки зрения решаемой задачи.

В данной главе будет описан метод слияния аудио-видео информации на уровне данных и на уровне классификаторов для решения задачи распознавания эмоций.

На первом этапе извлекаются аудио и видео-признаки из сырых данных видеозаписи. Аудио признаки были извлечены с помощью программы openSMILE. Видео-признаки были извлечены тремя разными алгоритмами:

1. Квантовые локальные моменты Цернике (Quantized Local Zernike Moments, QLZM) [123].
2. Локальные бинарные шаблоны (Local Binary Patterns, LBP).
3. Локальные бинарные шаблоны на трех ортогональных плоскостях (Local Binary Patterns on Three Orthogonal Planes, LBP-TOP).

Алгоритмы QLZM и LBP извлекают признаки из каждого отдельно взятого кадра видеопоследовательности, тогда как LBP-TOP работает в пространственно-временном пространстве, извлекая признаки для нескольких подряд идущих кадров. Количество подряд идущих кадров является настраиваемым параметром алгоритма LBP-TOP. Изображения подавались на вход алгоритма QLZM в оригинальном разрешении. Для алгоритмов LBP и LBP-TOP размер изображений был уменьшен с 1280:1024 до 200 пикселей по ширине с сохранением пропорций изображения. Были использованы следующие параметры алгоритма LBP: равномерный тип отображения с 8 пробными точками, радиус - 1. Параметры алгоритма LBP-TOP: радиусы по осям X, Y, T - (1; 1; 1), где (X; Y) - оси изображения, T - ось времени, количество пробных точек в плоскостях XY, XT и YT - (8; 8; 8).

Элементы выборки были получены путем усреднения извлеченных количественных признаков по всей аудио/видео последовательности. Таким образом, 1 аудио/видео последовательности соответствует 1 вектор признаков выборки.

Все данные были нормализованы, и был применен метод главных компонент (PCA) для снижения размерности данных. Число главных компонент было сокращено по правилу Кайзера: главные компоненты, у которых значения  $\lambda \leq 1$ , удалялись из модели. Здесь  $\lambda$  - собственные значения ковариационной матрицы выборки.

Унимодальные (то есть основанные только на аудио, либо только на видео) выборки аудио/видео признаков подавались на вход двух алгоритмов классификации: метод опорных векторов, обученный алгоритмом

последовательной минимальной оптимизации (SVM-SMO), и нейронная сеть прямого распространения. Были использованы программные реализации данных алгоритмов в средах Rapidminer и R. Алгоритмы классификации были применены для каждой унимодальной выборки (аудио и видео), чтобы определить точность классификации эмоций получающихся унимодальных систем. Кроме того, для каждой комбинации "алгоритм-данные" были определены точности классификации каждого класса эмоций в отдельности. Наиболее перспективные унимодальные выборки данных (по критерию точности) были объединены в единую выборку. Данный процесс называется слиянием гетерогенной информации на уровне данных. В данном контексте под гетерогенностью информации понимается разнотипность информации - аудио/видео. Также была сформирована выборка всех аудио-видео признаков.

Некоторые комбинации "алгоритм-данные" показали лучшую точность классификации на отдельных классах эмоций. Они были объединены в ансамбль, то есть проведено слияние гетерогенной информации на уровне классификаторов, тремя способами, описанными в § 3.3:

1. Метод голосования.
2. Усреднение вероятностей классов.
3. Мета-классификация алгоритмом SVM.

Для решения задачи распознавания эмоций были использованы два алгоритма классификации: SVM-SMO и нейронная сеть. Использовалась схема дикторонезависимой классификации. Точности классификации усреднялись по четырем возможным разбиениям на обучающую и тестовую выборки (4-кратная кросс-валидация). Параметры нейронной сети и алгоритма SVM-SMO представлены в таблице 5.1.

Для проведения экспериментов была использована программная реализация нейронной сети прямого распространения ошибки и алгоритма SVM-SMO в среде R. Упомянутые алгоритмы слияния гетерогенной информации были также реализованы в среде R.



Таблица 5.1. Параметры нейронной сети и алгоритма SVM-SMO

| <b>Параметры нейронной сети</b>       |                                                    |
|---------------------------------------|----------------------------------------------------|
| Число скрытых слоев                   | 2                                                  |
| Число нейронов на каждом скрытом слое | (число входных переменных + число классов) / 2 + 1 |
| число итераций обучения               | 200                                                |
| коэффициент скорости обучения         | 0.3                                                |
| <b>Параметры SVM-SMO</b>              |                                                    |
| Ядро                                  | Полиномиальное                                     |
| Константа сложности C                 | 1                                                  |
| Точность L                            | 0.001                                              |

Результаты экспериментов по классификации эмоций двумя описанными алгоритмами на различных выборках данных представлены в таблице 5.2. Из таблицы 5.2 видно, что признаки, извлеченные из аудио, а также из видео с помощью алгоритма LBP-TOP, обеспечили лучшую точность классификации эмоций, чем признаки, извлеченные из видео алгоритмами QLZM и LBP. Объединение признаков аудио и LBP-TOP в одну выборку, то есть слияние гетерогенной информации на уровне данных, позволило добиться улучшения точности классификации до 40.78%.

В таблице 5.3 представлены точности распознавания отдельных классов эмоций для различных комбинаций "алгоритм-данные". Все эмоции, за исключением грусти, распознаются достаточно хорошо на унимодальных и гетерогенных выборках. Злость лучше всего распознается путем применения алгоритма SVM-SMO на выборке LBP-TOP, счастье лучше всего распознается нейронной сетью, примененной к объединенной аудио-видео выборке, и т.д. В нижней части таблицы 5.3 представлены значения точности классификации отдельных классов для систем, полученных слиянием гетерогенной информации на уровне классификаторов 3 способами - методом голосования, усреднением

вероятностей классов, и мета-классификацией алгоритмом SVM. Оказалось, что наилучшие точности распознавания отдельных классов эмоций достигаются именно при слиянии аудио-видео информации на уровне классификаторов.

Таблица 5.2. Точность классификации эмоций (%), дикторонезависимая классификация, унимодальные системы и слияние на уровне данных

| Выборка                         | SVM-SMO      | Нейронная сеть |
|---------------------------------|--------------|----------------|
| Аудио                           | 36.73        | 35.42          |
| QLZM                            | 14.58        | 14.44          |
| LBP                             | 20.35        | 14.24          |
| LBP-TOP                         | 29.03        | 26.39          |
| Аудио + LBP-TOP                 | <b>40.78</b> | 35.18          |
| Аудио + QLZM + LBP +<br>LBP-TOP | 27.24        | 27.35          |

В таблице 5.4 представлена точность классификации эмоций, достигнутая при построении ансамбля (слияние на уровне классификаторов) тех комбинаций "алгоритм-данные", которые показали наилучшую точность распознавания отдельных классов эмоций (см. таблицу 5.3). Усреднение выходных вероятностей классов у систем, входящих в ансамбль, улучшило общую точность распознавания эмоций до 43.48%, что на 3% превышает лучшую точность, достигнутую с помощью слияния данных.

Таблица 5.3. Точности распознавания отдельных классов эмоций (%), слияние наиболее эффективных комбинаций "алгоритм-данные" на уровне данных и на уровне классификаторов

| Алгоритм<br>м<br>классифи<br>кации   | Выборка                            | Эмоция       |                |              |              |                 |              |               |
|--------------------------------------|------------------------------------|--------------|----------------|--------------|--------------|-----------------|--------------|---------------|
|                                      |                                    | Злость       | Отвраще<br>ние | Страх        | Счастье      | Нейтраль<br>ная | Грусть       | Удивлен<br>ие |
| SVM-<br>SMO                          | Аудио                              | 41.90        | 38.31          | 43.21        | 35.61        | 39.95           | 27.71        | 30.56         |
|                                      | QLZM                               | 13.31        | 8.48           | 4.47         | 7.50         | 27.49           | 2.08         | 0.00          |
|                                      | LBP                                | 16.07        | 5.23           | 0.00         | 23.44        | 0.00            | 0.00         | 1.58          |
|                                      | LBP-TOP                            | <b>63.24</b> | 8.93           | 12.36        | 43.85        | 50.93           | 4.35         | 25.00         |
|                                      | Аудио + LBP-<br>TOP                | 62.22        | <b>41.17</b>   | 51.82        | 58.64        | 69.95           | <b>28.24</b> | <b>64.26</b>  |
|                                      | Аудио +<br>QLZM + LBP<br>+ LBP-TOP | 46.50        | 12.77          | 16.67        | 51.72        | <b>83.33</b>    | 6.67         | 55.56         |
| Нейронна<br>я сеть                   | Аудио                              | 57.81        | 35.62          | 45.64        | 42.44        | 38.99           | 11.74        | 19.38         |
|                                      | QLZM                               | 3.40         | 32.84          | 20.35        | 12.50        | 60.30           | 1.67         | 11.36         |
|                                      | LBP                                | 5.58         | 2.34           | 0.00         | 21.70        | 8.33            | 0.00         | 0.00          |
|                                      | LBP-TOP                            | 43.75        | 7.63           | 20.77        | 25.14        | 19.24           | 19.71        | 23.30         |
|                                      | Аудио + LBP-<br>TOP                | 51.34        | 39.54          | <b>53.95</b> | 58.66        | 66.40           | 21.16        | 56.74         |
|                                      | Аудио +<br>QLZM + LBP<br>+ LBP-TOP | 39.50        | 21.31          | 32.86        | <b>61.31</b> | 51.89           | 3.70         | 52.78         |
| Голосование                          |                                    | <b>68.65</b> | 38.63          | 48.66        | <b>65.00</b> | 71.82           | 19.91        | 76.39         |
| Усреднение вероятностей<br>классов   |                                    | 50.85        | 35.52          | <b>71.76</b> | 46.78        | 66.87           | 33.93        | 68.75         |
| Мета-классификация<br>алгоритмом SVM |                                    | 63.89        | <b>41.91</b>   | 62.44        | 58.16        | <b>86.74</b>    | <b>38.89</b> | <b>90.00</b>  |

Таблица 5.4. Точность классификации эмоций (%), построение коллектива классификаторов, эффективно распознающих отдельные эмоции

| Метод построения коллектива | Точность классификации |
|-----------------------------|------------------------|
| Голосование                 | 41.88                  |
| Усреднение вероятностей     | <b>43.48</b>           |
| Мета-классификация          | 43.12                  |

## § 5.2 Разработка обобщенного метода для решения задач анализа гетерогенных данных на основе слияния данных, многокритериального отбора признаков и оптимизации алгоритмов машинного обучения, и конволюционных нейронных сетей

В § 5.1 были описаны два подхода к слиянию гетерогенных данных при построении систем машинного обучения - слияние на уровне данных (рисунок 5.1) и на уровне классификаторов (рисунок 5.2).

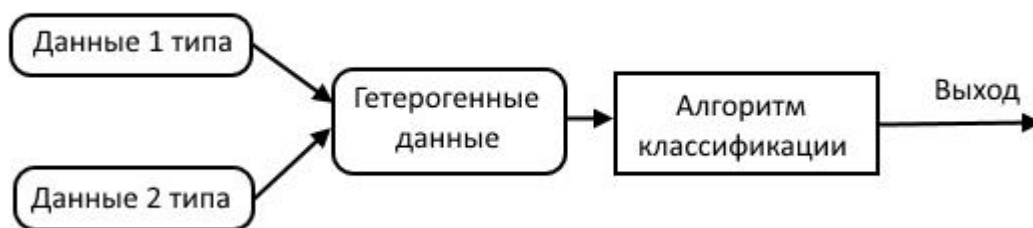


Рисунок 5.1. Слияние гетерогенной информации на уровне данных



Рисунок 5.2. Слияние гетерогенной информации на уровне классификаторов

Слияние на уровне данных в простейшем варианте представляет собой простое объединение нескольких выборок в общую выборку. Помимо этого, может быть применена процедура снижения размерности признакового пространства (до слияния, после слияния, либо и до, и после слияния).

Слияние на уровне классификаторов связано с процедурой объединения алгоритмов классификации, построенных на основе данных различного типа, в коллектив. В коллектив объединяются как правило не сами алгоритмы, а лишь результаты классификации, получаемые на выходе. В связи с этим, в качестве параллельно работающих алгоритмов классификации можно использовать однотипные алгоритмы (например, нейронные сети), либо алгоритмы разного типа (например, нейронная сеть и метод опорных векторов). Для объединения выходов алгоритмов классификации в ансамбль были применены 3 метода (описаны в § 3.3):

1. Метод голосования.
2. Усреднение апостериорных вероятностей классов.
3. Мета-классификация.

На основе методов, описанных в главах 2-4, была разработана обобщенная схема решения задач машинного обучения, в которых приходится иметь дело с гетерогенными данными. К таким задачам прежде всего относятся задачи анализа видеопоследовательностей. Звуковые признаки, отдельные кадры видео, и признаки, извлекаемые из нескольких последовательно идущих кадров, являются гетерогенной (неоднородной) информацией.

На рисунке 5.3 представлена схема разработанного обобщенного метода на примере задачи распознавания эмоций по видеопоследовательности.

Исходные данные включают в себя аудио и видео признаки, а также отдельные кадры видеозаписи. Аудио и видео признаки объединяются в общий пул, и производится отбор информативных признаков с помощью многокритериальной процедуры, описанной в § 3.2. Среди найденного Парето-

оптимального множества наборов признаков выбирается набор, обеспечивший наибольшую точность решения конечной задачи.

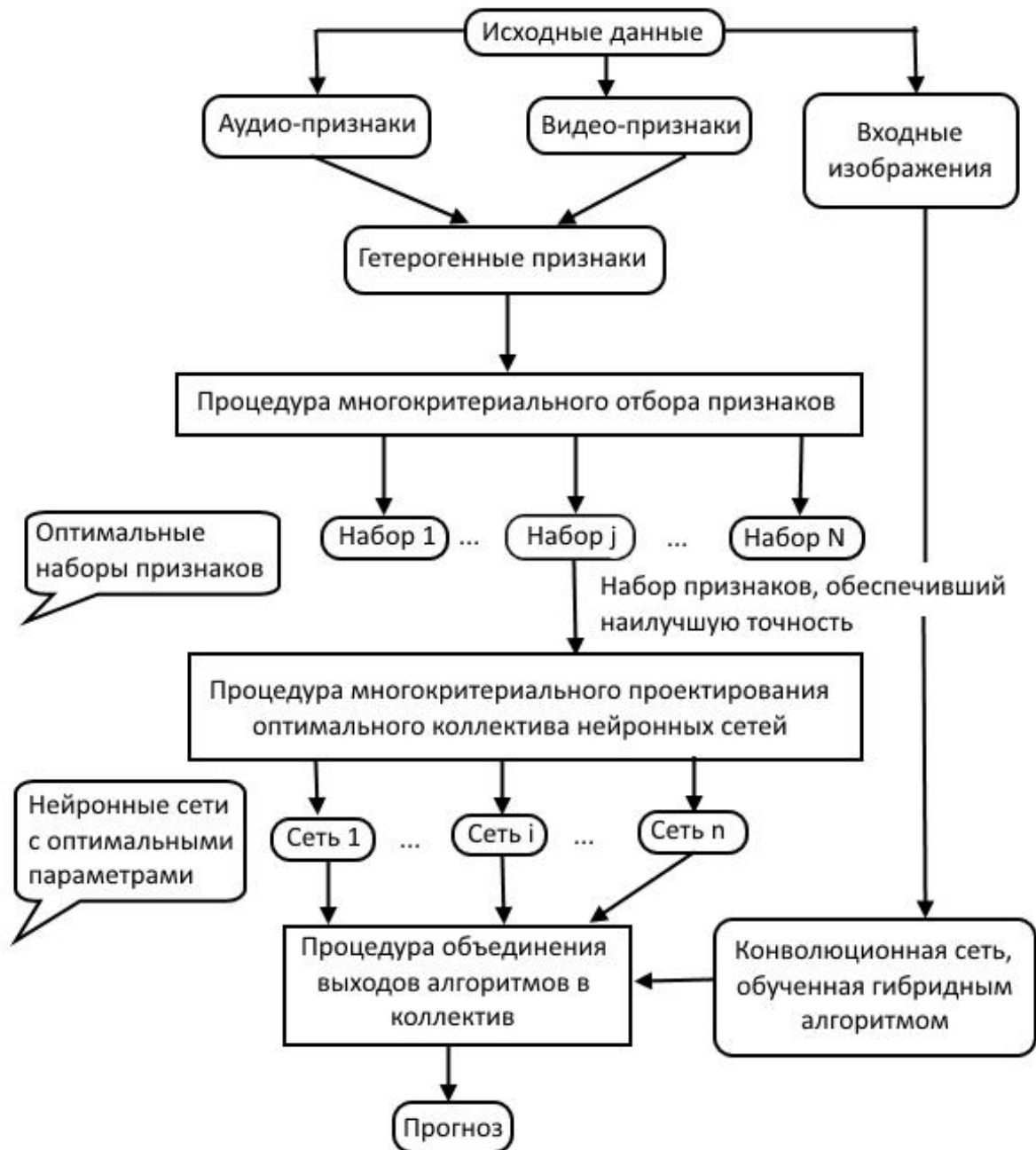


Рисунок 5.3. Схема обобщенного метода для решения задач анализа гетерогенных данных на примере задачи распознавания эмоций

Следующий этап - процедура многокритериального проектирования оптимального коллектива нейронных сетей, описанная в § 3.3. На вход нейронных сетей поступают данные, содержащие оптимальный набор признаков, найденный на предыдущем этапе. Заметим, что в общем случае, на месте нейронных сетей могут находиться любые другие алгоритмы классификации, по своей специфике пригодные для решения конечной задачи классификации. Единственное ограничение - алгоритмы должны быть одного типа. На выходе имеем найденное множество нейронных сетей с оптимальными параметрами.

Параллельно описанной процедуре производится обработка еще одного типа данных - входных изображений. Анализ и классификация входных изображений производятся с помощью КНС, обучаемой гибридным алгоритмом (описан в § 4.3).

Финальный этап - объединение Парето-оптимального набора нейронных сетей, а также КНС, в коллектив одним из вышеописанных способов.

### **§ 5.3 Исследование эффективности разработанного обобщенного метода на задаче распознавания эмоций**

Описанный в § 5.2 обобщенный метод был протестирован на задаче распознавания эмоций. Несмотря на то, что изначально данный метод включал в себя слияние аудио и видео признаков на уровне данных, с целью получения более общей картины были проведены эксперименты с использованием только видео, и только аудио данных. Изображения, подаваемые на вход КНС, были использованы во всех трех случаях.

Помимо этого, эксперименты проводились для различных алгоритмов многокритериальной оптимизации - SPEA, NSGA-2, VEGA, SelfCOMOGA (описаны в § 2.1-2.2), различных схем слияния классификаторов в ансамбль - голосование, усреднение апостериорных вероятностей классов, мета-

классификация SVM (описаны в § 5.2), и различного размера изображений, подаваемых на вход КНС -  $40 \times 40$ ,  $50 \times 50$ ,  $70 \times 70$ ,  $100 \times 100$  пикселей.

При проведении экспериментов значения параметров алгоритмов оптимизации SPEA, NSGA-2, VEGA и SelfCOMOGA были приняты теми же, что и при проведении экспериментов, представленных в 3 главе, значения параметров конволюционной нейронной сети аналогичны использованным в 4 главе.

Значения критериев точности классификации и F-меры для всех серий экспериментов представлены на рисунках Г1 – Г24 приложения Г. Приведенные графики включают в себя следующую информацию:

1. среднее значение критерия;
2. среднее значение критерия +  $\sigma$  (где  $\sigma$  – среднеквадратическое отклонение критерия);
3. среднее значение критерия –  $\sigma$ ;
4. максимум критерия;
5. минимум критерия.

Из представленных результатов видно, что эффективность обобщенного метода в целом сильно зависит от исходных используемых данных, а также от выбранной схемы слияния классификаторов в коллектив на финальном этапе.

На задаче распознавания эмоций наибольшую эффективность обеспечил метод слияния - мета-классификация SVM. Данный вывод оказался справедлив для всех рассмотренных алгоритмов многокритериальной оптимизации.

Что касается эффективности различных алгоритмов многокритериальной оптимизации применительно к отбору признаков и проектированию оптимального коллектива нейронных сетей, наиболее эффективным оказался алгоритм SelfCOMOGA. Данный алгоритм сочетает в себе достоинства алгоритмов SPEA, NSGA-2 и VEGA, и, как было показано, превосходит их по эффективности.



## Выводы

В данной главе был описан обобщенный метод для решения задач анализа гетерогенных данных. Преимущество данного метода заключается в том, что при решении сложных задач машинного обучения обеспечивается использование полного набора имеющихся данных различной природы. Кроме того, данный метод универсален - в него могут включаться различные типы входных данных, различные алгоритмы классификации и оптимизации, могут использоваться различные методы слияния классификаторов.

В рамках данной работы, в разработанный обобщенный метод были включены подходы, разработанные и описанные в данной диссертации. Проведена его апробация на задаче распознавания эмоций.

Экспериментальное исследование показало, что разработанный обобщенный метод достаточно эффективно справился с задачей распознавания эмоций. Алгоритм многокритериальной оптимизации SelfCOMOGA в составе обобщенного метода обеспечил наибольшую эффективность по сравнению с другими рассмотренными алгоритмами оптимизации. Мета-классификация оказалась самым эффективным методом слияния классификаторов в коллектив, обойдя другие рассмотренные методы слияния практически во всех проведенных экспериментах.

## ЗАКЛЮЧЕНИЕ

В ходе выполнения работы в соответствии с поставленными задачами были достигнуты следующие результаты:

1. Проведен анализ подходов к решению задачи распознавания эмоций, алгоритмов оптимизации, машинного обучения и глубинного обучения.

2. Разработан самоконфигурируемый коэволюционный алгоритм многокритериальной оптимизации SelfCOMOGA, превосходящий по эффективности алгоритмы оптимизации, входящие в него в качестве компонент, на тестовых задачах многокритериальной оптимизации, а также на практических задачах отбора признаков и проектирования ансамбля нейросетевых классификаторов.

3. Разработан и реализован многокритериальный подход к отбору информативных признаков и формированию оптимального коллектива нейросетевых классификаторов на основе эволюционных алгоритмов многокритериальной оптимизации. Многокритериальный подход превзошел по эффективности однокритериальный на исследованных задачах. Также многокритериальный подход к отбору информативных признаков превзошел по эффективности метод главных компонент.

4. Разработан и реализован гибридный алгоритм обучения конволюционной нейронной сети, основанный на последовательном использовании генетического алгоритма (ГА) оптимизации и алгоритма обратного распространения ошибки (ОРО). Предложенный гибридный алгоритм превосходит по эффективности алгоритмы ОРО и ГА, работающие по отдельности, на задаче распознавания эмоций, задаче распознавания рукописных цифр и объектов.

5. Предложен обобщенный метод для решения задач классификации, включающих использование гетерогенных аудио-видеоданных, на основе

разработанных в диссертации методов, проведена его успешная апробация на задаче распознавания эмоций.

6. Синтез эволюционных алгоритмов оптимизации и алгоритмов машинного обучения оказался эффективным по критерию точности классификации на наборе рассмотренных тестовых задач.

В диссертации была показана целесообразность использования методов машинного обучения и оптимизации применительно к задаче человеко-машинного взаимодействия в целом, и к задаче распознавания эмоций в частности. Применение эволюционных алгоритмов однокритериальной и многокритериальной оптимизации для отбора информативных признаков, проектирования ансамблей нейронных сетей и для обучения конволюционных нейронных сетей позволило увеличить точность решения задачи распознавания эмоций. Следовательно, была достигнута цель выполнения диссертационной работы - усовершенствованы методы проектирования нейросетевых систем глубинного и машинного обучения с помощью эволюционных алгоритмов оптимизации.

Теоретические идеи и практические результаты, полученные в данной диссертации, будут развиты и применены в дальнейших исследованиях при работе над диссертацией на соискание степени доктора наук. Среди возможных путей развития можно выделить следующие: применение многокритериального подхода для проектирования ансамблей других алгоритмов классификации; применение других алгоритмов оптимизации для отбора признаков, проектирования ансамблей классификаторов и предобучения конволюционной нейронной сети; построение ансамблей конволюционных нейронных сетей, а также других алгоритмов глубинного обучения. Кроме того, представленные подходы являются достаточно универсальными, а следовательно могут использоваться в других НИР, посвященных решению практических задач машинного обучения.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Адлер, Ю.П. Планирование эксперимента при поиске оптимальных условий. – Рипол Классик. - 1976. - 279 с.
2. Айвазян, С.А., Бухштабер, В.М., Енюков, Е.С. Прикладная статистика: Классификация и снижение размерности. — М.: Финансы и статистика. - 1989. - 607 с.
3. Айзерман, М.А., Браверман, Э.М., Розоноэр, Л.И. Метод потенциальных функций в теории обучения машин. – М.: Наука. - 1970. - 384 с.
4. Аркадьев, А.Г., Браверман, Э.М. Обучение машины классификации объектов. – М.: Наука. - 1971. - 192 с.
5. Базара, М., Шетти, К. Нелинейное программирование. Теория и алгоритмы. – М.: Мир. - 1982. - 583 с.
6. Банди, Б. Методы оптимизации. – Рипол Классик. - 1988. - 128 с.
7. Барсегян, А. Методы и модели анализа данных: OLAP и Data Mining. – БХВ-Петербург. - 2004. - 335 с.
8. Белецкая, С.Ю., Асанов, Ю.А., Поваляев, А.Д., Гаганов, А.В. Исследование эффективности генетических алгоритмов многокритериальной оптимизации //Вестник Воронежского государственного технического университета. – 2015. – Т. 11. – №. 1.
9. Божич, В.И., Лебедев, О.Б., Шницер, Ю.Л. Разработка генетического алгоритма обучения нейронных сетей //Известия Южного федерального университета. Технические науки. – 2001. – Т. 22. – №. 4, С. 170-174.
10. Брестер, К.Ю., Семенкин, Е.С. О решении задач многокритериальной оптимизации самонастраивающимся генетическим алгоритмом //Актуальные проблемы авиации и космонавтики. – 2012. – Т. 1. – №. 8, С. 290-291.
11. Ворожейкин, А.Ю., Гончар, Т.Н., Панфилов, И.А., Сопов, Е.А., Сопов, С.А. Об одной модификации вероятностного генетического алгоритма для решения сложных задач условной оптимизации //Вестник Сибирского

государственного аэрокосмического университета им. академика МФ Решетнева. – 2009. – №. 4, С. 79-84.

12. Гаврилова, Т.А., Хорошевский, В.Ф. Базы знаний интеллектуальных систем. – 2000.

13. Гладков, Л., Курейчик, В., Курейчик, В. Генетические алгоритмы. – Litres. 2017. - 365 с.

14. Джарратано, Д., Райли, Г. Экспертные системы: принципы разработки и программирование, 4-е издание. – Издательский дом Вильямс. - 2007. - 1152 с.

15. Дорофеюк, А.А. Алгоритмы автоматической классификации (обзор) //Автоматика и телемеханика. – 1971. – Т. 12. – С. 78-113.

16. Дюран, Б., Оделл, П. Кластерный анализ //М.: статистика. – 1977. – 128 с.

17. Епанечников, В.А. Непараметрическая оценка многомерной плотности вероятности //Теория вероятностей и ее применения. – 1969. – Т. 14. – №. 1. – С. 156-161.

18. Ершов, А.А. Стабильные методы оценки параметров (обзор) //Автоматика и телемеханика. – 1978. – №. 8. – С. 66-100.

19. Жуков, В.Г., Паротькин, Н.Ю. Исследование дифференцированного адаптивного генетического алгоритма решения задач условной оптимизации //Программные продукты и системы. – 2014. – №. 1 (105). - С. 82-86.

20. Журавлев, С.Ю., Фейгин, В.С. Генетический алгоритм решения многокритериальной задачи оптимизации энергозатрат при использовании машинно-тракторных агрегатов //Вестник Красноярского государственного аграрного университета. – 2013. – №. 10. - С. 182-191.

21. Иванов, В.К., Мескин, П.И.. Реализация генетического алгоритма для эффективного документального тематического поиска //Программные продукты и системы. – 2014. – №4(108). - С. 118-126.

22. Иванов, И.А., Сопов, Е.А. Исследование эффективности коэволюционного генетического алгоритма //Решетневские чтения. – 2011. – Т. 2. – №. 15. – С. 453-454.
23. Иванов, И.А., Сопов, Е.А. Самоконфигурируемый генетический алгоритм решения задач поддержки многокритериального выбора //Вестник Сибирского государственного аэрокосмического университета им. академика МФ Решетнева. – 2013. – №. 1 (47). - С. 30-35.
24. Карманов, В.Г. Математическое программирование. – М.: Наука. - 1980. - 256 с.
25. Курейчик, В.М., Кныш, Д.С. Нечеткий оператор кроссинговера для задачи трассировки коммутационного блока //Национальная конференция по искусственному интеллекту. - 2008. - С. 179-191.
26. Люгер, Д.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е издание. – Издательский дом Вильямс. - 2003. - 864 с.
27. Попов, Э.В. Экспертные системы: Решение неформализованных задач в диалоге с ЭВМ. - М.: Наука. – 1987. - 288 с.
28. Растригин, Л.А. Адаптация сложных систем. – Рига: Зинатне. - 1981. - 375 с.
29. Растригин, Л.А. Случайный поиск. - М.: Знание. – 1979. - 64 с.
30. Рутковская, Д., Рутковский, Л., Пилиньский, М. Нейронные сети, генетические алгоритмы и нечеткие системы. – М.: Горячая линия-Телеком. - 2003. - 384 с.
31. Себер, Д. Линейный регрессионный анализ. - М.: Мир. – 1980. - 456 с.
32. Сергиенко, Р.Б. Исследование эффективности коэволюционного генетического алгоритма условной оптимизации //Вестник Сибирского государственного аэрокосмического университета им. академика МФ Решетнева. – 2009. – №. 3. - С. 31-36.

33. Цой, Ю.Р., Спицын, В.Г. Эволюционный подход к настройке и обучению искусственных нейронных сетей //Электронный журнал «Нейроинформатика. – 2006. – Т. 1. – №. 1. – С. 34-61.
34. Ahmed, A., et al. Training hierarchical feed-forward visual recognition models using transfer learning from pseudo-tasks //European Conference on Computer Vision. – Springer Berlin Heidelberg. - 2008. – pp. 69-82.
35. Angelova, M., Pencheva, T. Tuning genetic algorithm parameters to improve convergence time //International Journal of Chemical Engineering. – 2011.
36. Anguelov, D., Lee, K., Gokturk, S.B., Sumengen, B. Contextual identity recognition in personal photo albums //2007 IEEE Conference on Computer Vision and Pattern Recognition. – IEEE. - 2007. – pp. 1-7.
37. Ani, R., Augustine, A., Akhil, N.C., Deepa, O.S. Random forest ensemble classifier to predict the coronary heart disease using risk factors //Proceedings of the International Conference on Soft Computing Systems. – Springer India. - 2016. - pp. 701-710.
38. Arya, S., et al. An optimal algorithm for approximate nearest neighbor searching fixed dimensions //Journal of the ACM (JACM). – 1998. – Vol. 45(6). – pp. 891-923.
39. Bahlmann, C., Haasdonk, B., Burkhardt, H. Online handwriting recognition with support vector machines-a kernel approach //Frontiers in handwriting recognition. – IEEE. - 2002. – pp. 49-54.
40. Batliner, A., et al. "You Stupid Tin Box" - children interacting with the AIBO robot: A cross-linguistic emotional speech corpus //LREC. – 2004.
41. Bekios-Calfa, J., Buenaposada, J.M., Baumela, L. Robust gender recognition by exploiting facial attributes dependencies //Pattern Recognition Letters. – 2014. – Vol. 36. – pp. 228-234.
42. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H. Greedy layer-wise training of deep networks //Advances in neural information processing systems. – 2007. – Vol. 19. – pp. 153–160.

43. Bengio, Y., LeCun, Y. Scaling learning algorithms towards AI //Large-scale kernel machines. – 2007. – Vol. 34(5). – pp. 1-41.
44. Bengio, Y., LeCun, Y., Nohl, C., Burges, C. LeRec: A NN/HMM hybrid for on-line handwriting recognition //Neural Computation. – 1995. – Vol. 7(6). – pp. 1289-1303.
45. Bergstra, J., Bengio, Y. Random search for hyper-parameter optimization //Journal of Machine Learning Research. – 2012. – Vol. 13(Feb). – pp. 281-305.
46. Billings, S.A., Zheng, G.L. Radial basis function network configuration using genetic algorithms //Neural Networks. – 1995. – Vol. 8(6). – pp. 877-890.
47. Bocklet, T., et al. Age and gender recognition for telephone applications based on gmm supervectors and support vector machines //IEEE International Conference on Acoustics, Speech and Signal Processing. – 2008. – pp. 1605-1608.
48. Borchert, M., Dusterhoft, A. Emotions in speech-experiments with prosody and quality features in speech for use in categorical and dimensional emotion recognition environments //International Conference on Natural Language Processing and Knowledge Engineering. – 2005. – pp. 147-151.
49. Boser, B.E., Guyon, I.M., Vapnik, V.N. A training algorithm for optimal margin classifiers //Proceedings of the fifth annual workshop on Computational learning theory. – 1992. – pp. 144-152.
50. Breiman, L. Bagging predictors //Machine learning. – 1996. – Vol. 24(2). – pp. 123-140.
51. Breiman, L. Random forests //Machine learning. – 2001. – Vol. 45(1). – pp. 5-32.
52. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A. Classification and regression trees. Wadsworth & Brooks //Monterey, CA. – 1984.
53. Burkhardt, F., et al. A database of German emotional speech //Interspeech. – 2005. – Vol. 5. – pp. 1517-1520.



54. Busso, C., et al. Analysis of emotion recognition using facial expressions, speech and multimodal information //Proceedings of the 6th international conference on Multimodal interfaces. – 2004. – pp. 205-211.
55. Busso, C., Narayanan, S.S. Interrelation between speech and facial gestures in emotional utterances: a single subject study //IEEE Transactions on Audio, Speech, and Language Processing. – 2007. – Vol. 15(8). – pp. 2331-2347.
56. Caridakis, G., et al. Multimodal emotion recognition from expressive faces, body gestures and speech //IFIP International Conference on Artificial Intelligence Applications and Innovations. – 2007. – pp. 375-388.
57. Chapelle, O., Scholkopf, B., Zien, A. Semi-Supervised Learning //MIT Press. - 2006. - 508 p.
58. Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S. Choosing multiple parameters for support vector machines //Machine learning. – 2002. – Vol. 46(1-3). – pp. 131-159.
59. Chebbi, O., Chaouachi, J. Effective parameter tuning for genetic algorithm to solve a real world transportation problem //Methods and Models in Automation and Robotics (MMAR). – 2015. – pp. 370-375.
60. Chen, L.S.H. Joint processing of audio-visual information for the recognition of emotional expressions in human-computer interaction. Ph.D. Thesis, University of Illinois at Urbana-Champaign, 2000.
61. Dahl, G.E., Yu, D., Deng, L., Acero, A. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition //IEEE Transactions on Audio, Speech, and Language Processing. – 2012. – Vol. 20(1). – pp. 30-42.
62. De Silva, L.C., Ng, P.C. Bimodal emotion recognition //Automatic Face and Gesture Recognition. – IEEE, 2000. – pp. 332-335.
63. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II //IEEE transactions on evolutionary computation. – 2002. – Vol. 6(2). – pp. 182-197.

64. Dehshibi, M.M., Bastanfard, A. A new algorithm for age recognition from facial images //Signal Processing. – 2010. – Vol. 90(8). – pp. 2431-2444.
65. Dempster, A.P., Laird, N.M., Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm //Journal of the royal statistical society. Series B (methodological). – 1977. – pp. 1-38.
66. Desjardins, G., Bengio, Y. Empirical evaluation of convolutional RBMs for vision //DIRO, Université de Montréal. – 2008. – pp. 1-13.
67. Elfenbein, H.A., Ambady, N. On the universality and cultural specificity of emotion recognition: a meta-analysis //Psychological bulletin. – 2002. – Vol. 128(2). – pp. 203-235.
68. El-Hajj, R., Likforman-Sulem, L., Mokbel, C. Arabic handwriting recognition using baseline dependant features and hidden Markov modeling //Eighth International Conference on Document Analysis and Recognition (ICDAR'05). – IEEE, 2005. – pp. 893-897.
69. Engberg, I.S., Hansen, A.V. Documentation of the danish emotional speech database des //Internal AAU report, Center for Person Kommunikation, Denmark. – 1996.
70. Eyben, F., Wöllmer, M., Schuller, B. Opensmile: the munich versatile and fast open-source audio feature extractor //Proceedings of the 18th ACM international conference on Multimedia. – ACM, 2010. – pp. 1459-1462.
71. Freund, Y., Haussler, D. Unsupervised learning of distributions on binary vectors using two layer networks //Technical Report UCSC-CRL-94-25, University of California. - 1994. - pp. 912-919.
72. Freund, Y., Schapire, R.E. A desicion-theoretic generalization of on-line learning and an application to boosting //European conference on computational learning theory. – Springer Berlin Heidelberg, 1995. – pp. 23-37.
73. Friedman, J., Hastie, T., Tibshirani, R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors) //The annals of statistics. – 2000. – Vol. 28(2). – pp. 337-407.

74. Friedman, J.H. Greedy function approximation: a gradient boosting machine //Annals of statistics. - 2001. - Vol. 29(5). - pp. 1189-1232.
75. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position //Biological Cybernetics. - 1980. - Vol. 36. - pp. 193–202.
76. Funahashi, K., Nakamura, Y. Approximation of dynamical systems by continuous time recurrent neural networks //Neural networks. – 1993. – Vol. 6(6). – pp. 801-806.
77. Goldberg, D.E. Genetic Algorithms for Search, Optimization, and Machine Learning. Addison-Wesley Longman Publishing Co. - 1989. - 432 p.
78. Graves, A., Schmidhuber, J. Offline handwriting recognition with multidimensional recurrent neural networks //Advances in neural information processing systems. – 2009. – pp. 545-552.
79. Gunes, H., Piccardi, M. A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior //18th International Conference on Pattern Recognition. – IEEE, 2006. – Vol. 1. – pp. 1148-1153.
80. Hansen, L.K., Salamon, P. Neural network ensembles //IEEE transactions on pattern analysis and machine intelligence. - 1990, - Vol. 12. - pp. 993-1001.
81. Haq, S., Jackson, P.J.B., Edge, J. Audio-visual feature selection and reduction for emotion classification //Proc. Int. Conf. on Auditory-Visual Speech Processing, Australia. – 2008.
82. Hinton, G., et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups //IEEE Signal Processing Magazine. – 2012. – Vol. 29(6). – pp. 82-97.
83. Hinton, G.E., Osindero, S., Teh, Y. A fast learning algorithm for deep belief nets //Neural Computation. – 2006. – Vol. 18. — pp. 1527–1554.
84. Holland, J.H. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press. - 1975. - 183 p.

85. Hosmer Jr, D.W., Lemeshow, S., Sturdivant, R.X. Applied logistic regression. John Wiley & Sons. - 2013. – 268 p.
86. Hsu, C.W., Chang, C.C., Lin, C.J. A practical guide to support vector classification. National Taiwan University. - 2003. - 16 p.
87. Hsu, W.N., Zhang, Y., Lee, A., Glass, J. Exploiting depth and highway connections in convolutional recurrent deep neural networks for speech recognition //Interspeech. – 2016. – Vol. 50. – pp. 395-399.
88. Huang, X.D., Ariki, Y., Jack, M.A. Hidden Markov models for speech recognition. Edinburgh : Edinburgh university press. - 1990.
89. Ioannou, S.V. et al. Emotion recognition through facial expression analysis based on a neurofuzzy network //Neural Networks. – 2005. – Vol. 18(4). – pp. 423-435.
90. Ishibuchi, H., Nozaki, K., Yamamoto, N., Tanaka, H. Selecting fuzzy if-then rules for classification problems using genetic algorithms //IEEE Transactions on fuzzy systems. – 1995. – Vol. 3(3). – pp. 260-270.
91. Jolliffe, I. Principal component analysis. John Wiley & Sons, Ltd. - 2002.
92. Khritonenko, D.I., Semenkin, E.S. Distributed self-configuring evolutionary algorithms for artificial neural networks design //Вестник Сибирского государственного аэрокосмического университета им. академика М. Ф. Решетнева. - 2013. - № 4(50). С. 112-116.
93. Kuo, C.H., Nevatia, R. How does person identity recognition help multi-person tracking? //IEEE Conference on computer vision and pattern recognition (CVPR). – IEEE, 2011. – pp. 1217-1224.
94. LeCun, Y., et al. Backpropagation applied to handwritten zip code recognition //Neural Computation. - 1989. - Vol. 1(4). - pp. 541–551.
95. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. Gradient-based learning applied to document recognition //Proceedings of the IEEE. -1998. - Vol. 86(11). - pp. 2278–2324.

96. Lee, H.K., Kim, J.H. An HMM-based threshold model approach for gesture recognition //IEEE Transactions on pattern analysis and machine intelligence. – 1999. – Vol. 21(10). – pp. 961-973.
97. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations //Proceedings of the twenty-sixth international conference on machine learning. - 2009. - pp. 609-616.
98. Margner, V., El Abed, H. Arabic handwriting recognition competition //Ninth international conference on document analysis and recognition (ICDAR). – IEEE, 2007. – Vol. 2. – pp. 1274-1278.
99. Marti, U.V., Bunke, H. The IAM-database: an English sentence database for offline handwriting recognition //International Journal on Document Analysis and Recognition. – 2002. – Vol. 5(1). – pp. 39-46.
100. Mitchell, M. An introduction to genetic algorithms. MIT press. - 1998. - 158 p.
101. Mitra, S., Acharya, T. Gesture recognition: A survey //IEEE Transactions on systems, man, and cybernetics, part C (applications and reviews). – 2007. – Vol. 37(3). – pp. 311-324.
102. Mnih, A., Hinton, G.E. A scalable hierarchical distributed language model //Advances in neural information processing systems. - 2009. - Vol. 21. - pp. 1081–1088.
103. Morency, L.P., Quattoni, A., Darrell, T. Latent-dynamic discriminative models for continuous gesture recognition //IEEE conference on computer vision and pattern recognition. – IEEE, 2007. – pp. 1-8.
104. Moulines, E., Charpentier, F. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones //Speech communication. – 1990. – Vol. 9(5-6). – pp. 453-467.

105. Murakami, K., Taguchi, H. Gesture recognition using recurrent neural networks //Proceedings of the SIGCHI conference on human factors in computing systems. – ACM, 1991. – pp. 237-242.
106. Nepomuceno, L., Baptista, E.C., Balbo, A.R., Soler, E.M. Coevolutionary genetic algorithm based on the augmented lagrangian function for solving the economic dispatch problem //IEEE Latin America Transactions. - 2015. - Vol. 13(10). - pp. 3277-3286.
107. O'Toole, A.J., et al. A video database of moving faces and people //IEEE Transactions on pattern analysis and machine intelligence. – 2005. – Vol. 27(5). – pp. 812-816.
108. Quinlan, J.R. C4.5: programs for machine learning. – Elsevier, 2014. - 301 p.
109. Quinlan, J.R. Induction of decision trees //Machine learning. – 1986. – Vol. 1(1). – pp. 81-106.
110. Rabiner, L.R. A tutorial on hidden Markov models and selected applications in speech recognition //Proceedings of the IEEE. – 1989. – Vol. 77(2). – pp. 257-286.
111. Ranzato, M., Boureau, Y.-L., LeCun, Y. Sparse feature learning for deep belief networks //Advances in Neural Information Processing Systems. - 2008. - Vol. 20. - pp. 1185–1192.
112. Ranzato, M., Huang, F., Boureau, Y., LeCun, Y. Unsupervised learning of invariant feature hierarchies with applications to object recognition //Proceedings of the computer vision and pattern recognition conference. - 2007. - pp. 1-8.
113. Ranzato, M., Poultney, C., Chopra, S., LeCun, Y. Efficient learning of sparse representations with an energy-based model //Advances in neural information processing systems. -2007. - Vol. 19. - pp. 1137–1144.
114. Reiter, E., Dale, R., Feng, Z. Building natural language generation systems. – Cambridge : Cambridge university press, 2000.

115. Ringeval, F., Sonderegger, A., Sauer, J., Lalanne, D. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions //10th IEEE International Conference and Workshops on automatic face and gesture recognition. - IEEE, 2013. – pp. 1-8.
116. Rish, I. An empirical study of the naive Bayes classifier //IJCAI 2001 workshop on empirical methods in artificial intelligence. – IBM New York, 2001. – Vol. 3(22). – pp. 41-46.
117. Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J. Rotation forest: A new classifier ensemble method //IEEE transactions on pattern analysis and machine intelligence. - 2006. - Vol. 28(10). - pp. 1619-1630.
118. Roisman, G.I., Tsai, J.L., Chiang, K.H.S. The emotional integration of childhood experience: physiological, facial expressive, and self-reported emotional response during the adult attachment interview //Developmental psychology. – 2004. – Vol. 40(5). – pp. 776-789.
119. Royston, P. et al. Multiple imputation of missing values //Stata journal. – 2004. – Vol. 4(3). – pp. 227-241.
120. Rumelhart, D.E., Hinton, G.E., Williams, R.J. Learning representations by back-propagating errors //Nature. -1986. - Vol. 323(6088). - pp. 533–536.
121. Sahu, A., Runger, G., Apley, D. Image denoising with a multi-phase kernel principal component approach and an ensemble version //IEEE applied imagery pattern recognition workshop. - 2011. - pp. 1-7.
122. Salakhutdinov, R., Hinton, G.E. Using deep belief nets to learn covariance kernels for Gaussian processes //Advances in neural information processing systems. - 2008. - Vol. 20. - pp. 1249–1256.
123. Sariyanidi, E., Gunes, H., Gokmen, M., Cavallaro, A. Local Zernike Moment Representation for Facial Affect Recognition //BMVC. – 2013. - 13 p.
124. Schaffer, J.D. Multiple objective optimization with vector evaluated genetic algorithms //Proceedings of the 1st international conference on genetic algorithms and their applications. - 1985. - pp. 93–100.

125. Scholkopf, B., Mullert, K.R. Fisher discriminant analysis with kernels //Neural networks for signal processing IX. – 1999. – Vol. 1(1). – pp. 41-48.
126. Schuller, B., Rigoll, G., Lang, M. Hidden Markov model-based speech emotion recognition // IEEE International conference on acoustics, speech, and signal processing. - IEEE, 2003. – Vol. 2. – pp. I-401-404.
127. Sebe, N., Cohen, I., Gevers, T., Huang, T.S. Emotion recognition based on joint visual and audio cues //18th International Conference on Pattern Recognition (ICPR'06). – IEEE, 2006. – Vol. 1. – pp. 1136-1139.
128. Serre, T., et al. A quantitative theory of immediate visual recognition //Progress in brain research, computational neuroscience: Theoretical insights into brain function. - 2007. - Vol. 165. - pp. 33–56.
129. Shan, C. Learning local binary patterns for gender classification on real-world face images //Pattern Recognition Letters. – 2012. – Vol. 33(4). – pp. 431-437.
130. Shinde, A., Sahu, A., Apley, D., Runger, G. Preimages for variation patterns from kernel PCA and bagging //IIE Transactions. - 2014. - Vol.46(5). - pp. 429-456.
131. Simard, D., Steinkraus, P.Y., Platt, J.C. Best practices for convolutional neural networks //International conference on document analysis and recognition. - 2003. - Vol. 3. - pp. 958-962.
132. Sims, C.A., Stock, J.H., Watson, M.W. Inference in linear time series models with some unit roots //Econometrica: Journal of the econometric society. – 1990. – pp. 113-144.
133. Snoek, J., Larochelle, H., Adams, R.P. Practical bayesian optimization of machine learning algorithms //Advances in neural information processing systems. - 2012. - pp. 2951-2959.
134. Soleymani, M., Pantic, M., Pun, T. Multimodal emotion recognition in response to videos //IEEE transactions on affective computing. – 2012. – Vol. 3(2). – pp. 211-223.



135. Srinivas, N., Deb, K. Multiobjective optimization using nondominated sorting in genetic algorithms //Evolutionary computation. - 1994. - Vol. 2(3). - pp. 221–248.
136. Sutton, R.S., Barto, A.G. Reinforcement learning: An introduction. – Cambridge : MIT press, 1998. – Vol. 1(1).
137. Tax, D.M.J., Duin, R.P.W. Using two-class classifiers for multiclass classification //16th International conference on pattern recognition. – IEEE, 2002. – Vol. 2. – pp. 124-127.
138. Thornton, C., Hutter, F., Hoos, H., Leyton-Brown, K. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms //Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining. - 2013. - pp. 847-855.
139. Tian, J., Li, M., Chen, F., Kou, J. Coevolutionary learning of neural network ensemble for complex classification tasks //Pattern recognition. - 2012. - Vol. 45(4). - pp. 1373-1385.
140. Utgoff, P.E., Stracuzzi, D.J. Many-layered learning //Neural computation. - 2002. - Vol. 14. - pp. 2497–2539.
141. Weigand, R.P., Liles, W.C., De Jong, K.A. An empirical analysis of collaboration methods in cooperative coevolutionary algorithms //Proceedings of the genetic and evolutionary computation conference. - 2001. - Vol. 2611. - pp. 1235-1245.
142. Yin, L., et al. A 3D facial expression database for facial behavior research //7th international conference on automatic face and gesture recognition (FGR06). – IEEE, 2006. – pp. 211-216.
143. Zadeh, L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic //Fuzzy sets and systems. – 1997. – Vol. 90(2). – pp. 111-127.
144. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S. A survey of affect recognition methods: Audio, visual, and spontaneous expressions //IEEE transactions on pattern analysis and machine intelligence. – 2009. – Vol. 31(1). – pp. 39-58.

145. Zhang, Z., et al. Enhanced semi-supervised learning for multimodal emotion recognition //2016 IEEE International conference on acoustics, speech and signal processing (ICASSP). – IEEE, 2016. – pp. 5185-5189.
146. Zhang, Q., et al. Multiobjective optimization test instances for the CEC 2009 special session and competition //University of Essex and Nanyang technological university, Tech. Rep. CES-487. - 2008.
147. Zhao, Z.S., et al. Evolved neural network ensemble by multiple heterogeneous swarm intelligence //Neurocomputing. - 2015. - Vol. 149. - pp. 29-38.
148. Zhou, Z., Harris, K.D. Counteracting stagnation in genetic algorithm calculations by implementation of a micro genetic algorithm strategy //Physical chemistry chemical physics. - 2008. - Vol. 10. - pp. 7262-7269.
149. Zhou, Z.H., Wu, J., Tang, W. Ensembling neural networks: many could be better than all //Artificial intelligence. - 2002. - Vol. 137(1). - pp. 239-263.
150. Zitzler, E., Laumanns, M., Thiele, L. SPEA2: Improving the strength Pareto evolutionary algorithm //Eurogen. - 2001. - Vol. 3242(103). - pp. 95-100.
151. Zitzler, E., Thiele, L. An evolutionary algorithm for multiobjective optimization: The strength pareto approach //Technical report №43, Computer engineering and communication networks lab. - 1998. - 40 p.

## СПИСОК ПУБЛИКАЦИЙ АВТОРА

### **Статьи в ведущих рецензируемых научных журналах и изданиях**

1. Ivanov I. Multi-objective based feature selection and neural networks ensemble method for solving emotion recognition problem // Вестник СибГАУ. - 2016. - Т. 17, № 1. - С. 45-49.
2. Иванов И.А. Гибридный алгоритм обучения конволюционной нейронной сети / Иванов И.А., Сопов Е.А. // Вестник СибГАУ. - 2016. - Т. 17, № 4.
3. Иванов И.А. Многокритериальный подход к проектированию ансамбля нейросетевых классификаторов с отбором информативных признаков для решения задачи распознавания эмоций / Иванов И.А., Сопов Е.А., Панфилов И.А. // Вестник СибГАУ. - 2015. - Т. 16, № 4. - С. 819-827.
4. Иванов И.А. Исследование эффективности самоконфигурируемого коэволюционного алгоритма решения сложных задач многокритериальной оптимизации / Иванов И.А., Сопов Е.А. // Системы управления и информационные технологии. - 2013. - Т. 51, № 1.1. - С. 141-145.
5. Иванов И.А. Самоконфигурируемый генетический алгоритм решения задач поддержки многокритериального выбора / Иванов И.А., Сопов Е.А. // Вестник СибГАУ. - 2013. - № 1 (47). - С. 30-35.

### **Свидетельства на программу для электронных вычислительных машин**

6. Программа для решения задачи распознавания эмоций человека по видеозаписи методом оптимизации коллектива нейронных сетей и отбора информативных признаков с помощью генетического алгоритма оптимизации / Иванов И.А., Сопов Е.А. // Свидетельство №2017610772 о государственной регистрации в Реестре программ для ЭВМ от 18.01.2017.
7. Программа автоматической сегментации левого желудочка сердца на снимках магнитно-резонансной томографии на основе кластерного подхода /

Иванов И.А., Брестер К.Ю. // Свидетельство №2017611091 о государственной регистрации в Реестре программ для ЭВМ от 19.01.2017.

8. Генетический алгоритм для решения задачи совмещения маршрутов с предварительной кластеризацией точек отправления и назначения / Иванов И.А., Сопов Е.А. // Свидетельство №2016611721 о государственной регистрации в Реестре программ для ЭВМ от 09.02.2016.

9. Программная система решения сложных задач многокритериальной оптимизации самоконфигурируемым коэволюционным алгоритмом SelfCOMOGA / Иванов И.А., Сопов Е.А. // Свидетельство №2013613773 о государственной регистрации в Реестре программ для ЭВМ от 16.04.2013.

#### **Публикации в изданиях, индексируемых в международных базах**

10. Ivanov I. Self-configuring ensemble of neural network classifiers for emotion recognition in the intelligent HMI / Ivanov I., Sopov E. // 2015 IEEE Symposium Series on Computational Intelligence, 8-10 December 2015, Cape Town, South Africa.

11. Ivanov I. Feature and decision level audio-visual data fusion in emotion recognition problem / Ivanov I., Sopov E., Sidorov M., Minker W. // 12th International Conference on Informatics in Control, Automation and Robotics, 21-23 July 2015, Colmar, France.

12. Ivanov I. Design efficient technologies for context image analysis in dialog HCI using self-configuring novelty search genetic algorithm / Sopov E., Ivanov I. // 11th International Conference on Informatics in Control, Automation and Robotics, 1-3 September 2014, Vienna, Austria.

#### **Прочие публикации, отражающие основное содержание диссертации**

13. Иванов И.А. Система автоматического выделения левого желудочка сердца на снимках магнитно-резонансной томографии // Материалы V Всероссийской научно-методической конференции с международным участием "Информационные технологии в математике и математическом образовании" (Красноярск, 16-17 ноября 2016 г.). - С. 25-29. - 2016.

14. Ivanov I. Artificial intelligence application in medicine: automatic calculation of the left ventricle ejection fraction // Материалы XV Международной научной конференции бакалавров, магистрантов, аспирантов и молодых ученых "Молодежь. Общество. Современная наука, техника и инновации" (Красноярск, 12 мая 2016 г.). - С. 249-252. - 2016.

15. Иванов И.А. Проектирование нейросетевого классификатора для решения задачи распознавания эмоций // Материалы XIX Международной научной конференции "Решетневские чтения" (Красноярск, 10-13 ноября 2015 г.). - Т. 2. - С. 42-44. - 2015.

16. Иванов И.А. Самоорганизующийся алгоритм решения многокритериальных задач оптимизации на базе коэволюционного генетического алгоритма / Иванов И.А., Сопов Е.А. // Материалы V Международной конференции "Системный анализ и информационные технологии" (САИТ-2013, Красноярск, 19-25 сентября 2013 г.). - 2013.

17. Иванов И.А. Система поддержки принятия решений многокритериального выбора на базе коэволюционного генетического алгоритма // Материалы IX Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых "Молодежь и наука" с международным участием (Красноярск, 15-25 апреля 2013 г.). - 2013.

**Приложение А. Сравнение разработанного алгоритма SelfCOMOGA с  
алгоритмами-победителями соревнования CEC по метрике IGD**

Таблица А1. Тестовая задача 1

| <b>Ранг</b> | <b>Название алгоритма</b> | <b>IGD</b>     |
|-------------|---------------------------|----------------|
| 1           | MOEAD                     | 0.00435        |
| 2           | GDE3                      | 0.00534        |
| 3           | MOEADGM                   | 0.0062         |
| 4           | MTS                       | 0.00646        |
| 5           | LiuLiAlgorithm            | 0.00785        |
| 6           | DMOEADD                   | 0.01038        |
| 7           | NSGAIIS                   | 0.01153        |
| 8           | OWMOSaDE                  | 0.0122         |
| 9           | ClusteringMOEA            | 0.0299         |
| 10          | AMGA                      | 0.03588        |
| 11          | MOEP                      | 0.0596         |
|             | <b>SelfCOMOGA</b>         | <b>0.06632</b> |
| 12          | DECMOSA-SQP               | 0.07702        |
| 13          | OMOEAI                    | 0.08564        |

Таблица А2. Тестовая задача 2

| <b>Ранг</b> | <b>Название алгоритма</b> | <b>IGD</b> |
|-------------|---------------------------|------------|
| 1           | MTS                       | 0.00615    |
| 2           | MOEADGM                   | 0.0064     |
| 3           | DMOEADD                   | 0.00679    |
| 4           | MOEAD                     | 0.00679    |
| 5           | OWMOSaDE                  | 0.0081     |
| 6           | GDE3                      | 0.01195    |
| 7           | LiuLiAlgorithm            | 0.0123     |

Продолжение таблицы A2.

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 8    | NSGAIILS           | 0.01237        |
| 9    | AMGA               | 0.01623        |
| 10   | MOEP               | 0.0189         |
| 11   | ClusteringMOEA     | 0.0228         |
|      | <b>SelfCOMOGA</b>  | <b>0.02689</b> |
| 12   | DECMOSA-SQP        | 0.02834        |
| 13   | OMOEAI             | 0.03057        |

Таблица A3. Тестовая задача 3

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | MOEAD              | 0.00742        |
| 2    | LiuLiAlgorithm     | 0.01497        |
| 3    | DMOEADD            | 0.03337        |
| 4    | MOEADGM            | 0.049          |
| 5    | MTS                | 0.0531         |
| 6    | ClusteringMOEA     | 0.0549         |
| 7    | AMGA               | 0.06998        |
| 8    | DECMOSA-SQP        | 0.0935         |
| 9    | MOEP               | 0.099          |
| 10   | OWMOSaDE           | 0.103          |
| 11   | NSGAIILS           | 0.10603        |
| 12   | GDE3               | 0.10639        |
|      | <b>SelfCOMOGA</b>  | <b>0.18667</b> |
| 13   | OMOEAI             | 0.27141        |

Таблица А4. Тестовая задача 4

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | MTS                | 0.02356        |
| 2    | GDE3               | 0.0265         |
| 3    | DECMOSA-SQP        | 0.03392        |
|      | <b>SelfCOMOGA</b>  | <b>0.04025</b> |
| 4    | AMGA               | 0.04062        |
| 5    | DMOEADD            | 0.04268        |
| 6    | MOEP               | 0.0427         |
| 7    | LiuLiAlgorithm     | 0.0435         |
| 8    | OMOEAI             | 0.04624        |
| 9    | MOEADGM            | 0.0476         |
| 10   | OWMOSaDE           | 0.0513         |
| 11   | NSGAIILS           | 0.0584         |
| 12   | ClusteringMOEA     | 0.0585         |
| 13   | MOEAD              | 0.06385        |

Таблица А5. Тестовая задача 5

| Ранг | Название алгоритма | IGD     |
|------|--------------------|---------|
| 1    | MTS                | 0.01489 |
| 2    | GDE3               | 0.03928 |
| 3    | AMGA               | 0.09405 |
| 4    | LiuLiAlgorithm     | 0.16186 |
| 5    | DECMOSA-SQP        | 0.16713 |
| 6    | OMOEAI             | 0.1692  |
| 7    | MOEAD              | 0.18071 |
| 8    | MOEP               | 0.2245  |
| 9    | ClusteringMOEA     | 0.2473  |
| 10   | DMOEADD            | 0.31454 |



Продолжение таблицы А5.

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 11   | OWMOSaDE           | 0.4303         |
| 12   | NSGAIILS           | 0.5657         |
|      | <b>SelfCOMOGA</b>  | <b>0.67732</b> |
| 13   | MOEADGM            | 1.7919         |

Таблица А6. Тестовая задача 6

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | MOEAD              | 0.00587        |
| 2    | MTS                | 0.05917        |
| 3    | DMOEADD            | 0.06673        |
| 4    | OMOEAI             | 0.07338        |
| 5    | ClusteringMOEA     | 0.0871         |
|      | <b>SelfCOMOGA</b>  | <b>0.10073</b> |
| 6    | MOEP               | 0.1031         |
| 7    | DECMOSA-SQP        | 0.12604        |
| 8    | AMGA               | 0.12942        |
| 9    | LiuLiAlgorithm     | 0.17555        |
| 10   | OWMOSaDE           | 0.1918         |
| 11   | GDE3               | 0.25091        |
| 12   | NSGAIILS           | 0.31032        |
| 13   | MOEADGM            | 0.5563         |

Таблица А7. Тестовая задача 7

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | MOEAD              | 0.00444        |
| 2    | LiuLiAlgorithm     | 0.0073         |
| 3    | MOEADGM            | 0.0076         |
| 4    | DMOEADD            | 0.01032        |
| 5    | MOEP               | 0.0197         |
| 6    | NSGAIILS           | 0.02132        |
| 7    | ClusteringMOEA     | 0.0223         |
| 8    | DECMOSA-SQP        | 0.02416        |
| 9    | GDE3               | 0.02522        |
| 10   | OMOEAI             | 0.03354        |
| 11   | MTS                | 0.04079        |
|      | <b>SelfCOMOGA</b>  | <b>0.05321</b> |
| 12   | AMGA               | 0.05707        |
| 13   | OWMOSaDE           | 0.0585         |

Таблица А8. Тестовая задача 8

| Ранг | Название алгоритма | IGD     |
|------|--------------------|---------|
| 1    | MOEAD              | 0.0584  |
| 2    | DMOEADD            | 0.06841 |
| 3    | LiuLiAlgorithm     | 0.08235 |
| 4    | NSGAIILS           | 0.0863  |
| 5    | OWMOSaDE           | 0.0945  |
| 6    | MTS                | 0.11251 |
| 7    | AMGA               | 0.17125 |
| 8    | OMOEAI             | 0.192   |
| 9    | DECMOSA-SQP        | 0.21583 |
| 10   | ClusteringMOEA     | 0.2383  |

Продолжение таблицы А8.

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 11   | MOEADGM            | 0.2446         |
| 12   | GDE3               | 0.24855        |
| 13   | MOEP               | 0.423          |
|      | <b>SelfCOMOGA</b>  | <b>0.85356</b> |

Таблица А9. Тестовая задача 9

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | DMOEADD            | 0.04896        |
| 2    | NSGAIILS           | 0.0719         |
| 3    | MOEAD              | 0.07896        |
| 4    | GDE3               | 0.08248        |
| 5    | LiuLiAlgorithm     | 0.09391        |
| 6    | OWMOSaDE           | 0.0983         |
| 7    | MTS                | 0.11442        |
| 8    | DECMOSA-SQP        | 0.14127        |
|      | <b>SelfCOMOGA</b>  | <b>0.15932</b> |
| 9    | MOEADGM            | 0.1878         |
| 10   | AMGA               | 0.18861        |
| 11   | OMOEAI             | 0.23179        |
| 12   | ClusteringMOEA     | 0.2934         |
| 13   | MOEP               | 0.342          |

Таблица A10. Тестовая задача 10

| Ранг | Название алгоритма | IGD            |
|------|--------------------|----------------|
| 1    | MTS                | 0.15306        |
| 2    | DMOEADD            | 0.32211        |
| 3    | AMGA               | 0.32418        |
| 4    | MOEP               | 0.3621         |
| 5    | DECMOSA-SQP        | 0.36985        |
| 6    | ClusteringMOEA     | 0.4111         |
| 7    | GDE3               | 0.43326        |
| 8    | LiuLiAlgorithm     | 0.44691        |
| 9    | MOEAD              | 0.47415        |
|      | <b>SelfCOMOGA</b>  | <b>0.50667</b> |
| 10   | MOEADGM            | 0.5646         |
| 11   | OMOEAI             | 0.62754        |
| 12   | OWMOSaDE           | 0.743          |
| 13   | NSGAIILS           | 0.84468        |

**Приложение Б. Результаты экспериментов по исследованию эффективности  
многокритериального подхода к отбору информативных признаков и  
проектированию ансамбля нейросетевых классификаторов**

Таблица Б1. Точность определения типа раковой опухоли (%) при снижении  
размерности ( $n$  - усредненное по экспериментам число признаков,  
соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |       |       | n  |
|-------------------------|--------------------------------|-------------------|-----------|----------|-------|-------|----|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min   | Max   |    |
| Все признаки            |                                |                   | 90.35     | 1.81     | 88.30 | 94.15 | 30 |
| Метод главных компонент |                                |                   | 84.50     | 2.23     | 82.46 | 88.30 | 6  |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 93.27     | 2.87     | 92.40 | 98.25 | 17 |
|                         | Многокритериальная оптимизация | SPEA              | 89.18     | 5.83     | 83.04 | 95.91 | 6  |
|                         |                                | NSGA-2            | 92.11     | 4.10     | 81.29 | 96.49 | 8  |
|                         |                                | VEGA              | 87.43     | 3.61     | 81.87 | 90.06 | 7  |
|                         |                                | SelfCOMOGA        | 97.37     | 4.75     | 90.64 | 98.83 | 5  |

Таблица Б2. Точность определения типа раковой опухоли (%) при проектировании ансамбля нейронных сетей эволюционными алгоритмами  
оптимизации

| Алгоритм<br>оптимизации | Схема слияния выходов<br>классификаторов ансамбля | Точность     |          |       |       |
|-------------------------|---------------------------------------------------|--------------|----------|-------|-------|
|                         |                                                   | $\bar{x}$    | $\sigma$ | min   | max   |
| Коэволюцион<br>ный ГА   | -                                                 | 94.44        | 4.23     | 85.38 | 98.83 |
| SPEA                    | Голосование                                       | 86.26        | 3.27     | 84.80 | 90.64 |
|                         | Усреднение вероятностей классов                   | 62.28        | 7.12     | 58.48 | 74.27 |
|                         | Мета-классификация                                | 90.35        | 3.78     | 81.87 | 95.32 |
| NSGA-2                  | Голосование                                       | 91.52        | 5.21     | 88.30 | 100   |
|                         | Усреднение вероятностей классов                   | 62.28        | 8.45     | 50.29 | 72.51 |
|                         | Мета-классификация                                | 94.44        | 2.43     | 89.47 | 95.91 |
| VEGA                    | Голосование                                       | 87.43        | 5.21     | 77.19 | 94.74 |
|                         | Усреднение вероятностей классов                   | 62.28        | 6.33     | 51.46 | 70.76 |
|                         | Мета-классификация                                | 95.61        | 3.11     | 88.89 | 99.42 |
| SelfCOMOGA              | Голосование                                       | 96.78        | 2.45     | 93.57 | 98.83 |
|                         | Усреднение вероятностей классов                   | 72.81        | 7.90     | 64.91 | 89.47 |
|                         | Мета-классификация                                | <b>97.95</b> | 1.13     | 97.08 | 100   |

Таблица Б3. Точность распознавания рукописных цифр для задачи penbased (%) при снижении размерности ( $n$  - усредненное по экспериментам число признаков, соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |      |       | n  |
|-------------------------|--------------------------------|-------------------|-----------|----------|------|-------|----|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min  | max   |    |
| Все признаки            |                                |                   | 86.75     | 2.18     | 83.5 | 91.00 | 16 |
| Метод главных компонент |                                |                   | 81.25     | 3.62     | 73.5 | 83.00 | 5  |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 87.25     | 3.11     | 81   | 88.00 | 15 |
|                         | Многокритериальная оптимизация | SPEA              | 86.50     | 4.75     | 80.5 | 94.50 | 6  |
|                         |                                | NSGA-2            | 89.00     | 5.65     | 81   | 93.50 | 8  |
|                         |                                | VEGA              | 85.25     | 2.44     | 80   | 86.00 | 10 |
|                         |                                | SelfCOMOGA        | 91.25     | 3.16     | 88   | 94.50 | 13 |

Таблица Б4. Точность распознавания рукописных цифр для задачи penbased (%) при проектировании ансамбля нейронных сетей эволюционными алгоритмами оптимизации

| Алгоритм оптимизации | Схема слияния выходов классификаторов ансамбля | Точность     |          |       |       |
|----------------------|------------------------------------------------|--------------|----------|-------|-------|
|                      |                                                | $\bar{x}$    | $\sigma$ | min   | max   |
| Коэволюционный ГА    | -                                              | 90.75        | 4.89     | 88.00 | 97    |
| SPEA                 | Голосование                                    | 87.25        | 5.76     | 78.00 | 95.50 |
|                      | Усреднение вероятностей классов                | 68.25        | 7.98     | 57.50 | 82.50 |
|                      | Мета-классификация                             | 92.75        | 4.87     | 86.50 | 99.50 |
| NSGA-2               | Голосование                                    | 90.25        | 3.32     | 83.00 | 97.50 |
|                      | Усреднение вероятностей классов                | 79.25        | 6.42     | 71.50 | 85.50 |
|                      | Мета-классификация                             | 92.75        | 2.56     | 89.50 | 96.50 |
| VEGA                 | Голосование                                    | 91.25        | 3.15     | 87.50 | 97.50 |
|                      | Усреднение вероятностей классов                | 83.75        | 5.32     | 69.50 | 92.00 |
|                      | Мета-классификация                             | 91.00        | 3.76     | 84.00 | 98.00 |
| SelfCOMOGA           | Голосование                                    | 89.25        | 2.43     | 85.50 | 94.50 |
|                      | Усреднение вероятностей классов                | 79.25        | 4.83     | 74.50 | 90.00 |
|                      | Мета-классификация                             | <b>96.25</b> | 2.67     | 90.50 | 100   |



Таблица Б5. Точность определения наличия заболевания диабетом (%) при снижении размерности ( $n$  - усредненное по экспериментам число признаков, соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |       |       | n |
|-------------------------|--------------------------------|-------------------|-----------|----------|-------|-------|---|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min   | max   |   |
| Все признаки            |                                |                   | 67.61     | 2.52     | 64.35 | 73.04 | 8 |
| Метод главных компонент |                                |                   | 74.57     | 3.11     | 68.26 | 79.13 | 3 |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 79.78     | 3.54     | 74.35 | 84.78 | 6 |
|                         | Многокритериальная оптимизация | SPEA              | 78.48     | 4.32     | 74.78 | 87.83 | 6 |
|                         |                                | NSGA-2            | 78.91     | 5.11     | 73.04 | 89.57 | 2 |
|                         |                                | VEGA              | 79.35     | 2.12     | 75.22 | 81.74 | 5 |
|                         |                                | SelfCOMOGA        | 83.70     | 2.65     | 80.87 | 88.70 | 3 |

Таблица Б6. Точность определения наличия заболевания диабетом (%) при проектировании ансамбля нейронных сетей эволюционными алгоритмами  
оптимизации

| Алгоритм<br>оптимизации | Схема слияния выходов<br>классификаторов ансамбля | Точность     |          |       |       |
|-------------------------|---------------------------------------------------|--------------|----------|-------|-------|
|                         |                                                   | $\bar{x}$    | $\sigma$ | min   | max   |
| Козволюционн<br>ый ГА   | -                                                 | 80.65        | 4.87     | 72.61 | 89.57 |
| SPEA                    | Голосование                                       | 74.57        | 9.32     | 58.70 | 85.22 |
|                         | Усреднение вероятностей классов                   | 77.61        | 5.21     | 69.13 | 89.13 |
|                         | Мета-классификация                                | 84.57        | 5.32     | 70.87 | 93.04 |
| NSGA-2                  | Голосование                                       | 79.78        | 5.90     | 73.04 | 89.13 |
|                         | Усреднение вероятностей классов                   | 76.30        | 5.56     | 70.87 | 82.17 |
|                         | Мета-классификация                                | 83.26        | 3.78     | 78.26 | 88.26 |
| VEGA                    | Голосование                                       | 76.74        | 2.45     | 71.30 | 81.30 |
|                         | Усреднение вероятностей классов                   | 75.00        | 2.87     | 72.61 | 81.74 |
|                         | Мета-классификация                                | 75.43        | 3.55     | 68.70 | 83.04 |
| SelfCOMOGA              | Голосование                                       | 78.91        | 5.65     | 69.13 | 85.22 |
|                         | Усреднение вероятностей классов                   | 81.96        | 6.87     | 73.04 | 94.78 |
|                         | Мета-классификация                                | <b>85.43</b> | 3.23     | 80.43 | 88.26 |

Таблица Б7. Точность классификации типов почв (%) при снижении размерности  
( $n$  - усредненное по экспериментам число признаков, соответствующее лучшей  
точности)

| Метод отбора признаков  |                                |                   | Точность     |          |       |       | n        |
|-------------------------|--------------------------------|-------------------|--------------|----------|-------|-------|----------|
|                         |                                |                   | $\bar{x}$    | $\sigma$ | min   | max   |          |
| Все признаки            |                                |                   | 75.17        | 1.31     | 73.33 | 76.00 | 36       |
| Метод главных компонент |                                |                   | 83.50        | 5.16     | 80.00 | 97.33 | 3        |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 93.83        | 3.67     | 85.33 | 98.00 | 19       |
|                         | Многокритериальная оптимизация | SPEA              | 91.83        | 5.89     | 80.67 | 99.00 | 8        |
|                         |                                | NSGA-2            | 91.50        | 2.35     | 88.33 | 98.33 | 10       |
|                         |                                | VEGA              | 91.17        | 3.26     | 86.33 | 95.00 | 20       |
|                         |                                | SelfCOMOGA        | <b>94.17</b> | 4.83     | 90.67 | 100   | <b>6</b> |

Таблица Б8. Точность классификации типов почв (%) при проектировании ансамбля нейронных сетей эволюционными алгоритмами оптимизации

| Алгоритм оптимизации | Схема слияния выходов классификаторов ансамбля | Точность     |          |       |       |
|----------------------|------------------------------------------------|--------------|----------|-------|-------|
|                      |                                                | $\bar{x}$    | $\sigma$ | min   | max   |
| Коеволюционный ГА    | -                                              | 95.17        | 2.43     | 92.00 | 97.67 |
| SPEA                 | Голосование                                    | 86.83        | 3.65     | 81.67 | 90.33 |
|                      | Усреднение вероятностей классов                | 90.50        | 3.23     | 85.67 | 96.67 |
|                      | Мета-классификация                             | 91.83        | 3.88     | 84.00 | 99.00 |
| NSGA-2               | Голосование                                    | 89.17        | 4.12     | 82.00 | 97.33 |
|                      | Усреднение вероятностей классов                | 90.50        | 2.45     | 85.00 | 95.00 |
|                      | Мета-классификация                             | 92.83        | 2.11     | 87.67 | 100   |
| VEGA                 | Голосование                                    | 86.50        | 4.71     | 78.67 | 94.67 |
|                      | Усреднение вероятностей классов                | 90.17        | 1.54     | 89.00 | 90.67 |
|                      | Мета-классификация                             | 88.83        | 5.76     | 82.33 | 97.67 |
| SelfCOMOGA           | Голосование                                    | 88.17        | 1.76     | 85.33 | 91.67 |
|                      | Усреднение вероятностей классов                | 90.17        | 2.89     | 88.33 | 94.00 |
|                      | Мета-классификация                             | <b>96.50</b> | 1.23     | 94.00 | 97.67 |

Таблица Б9. Точность классификации эмоций (%) для различных подходов к снижению размерности, входные данные – лицевые маркеры ( $n$  - усредненное по экспериментам число признаков, соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |       |       | n  |
|-------------------------|--------------------------------|-------------------|-----------|----------|-------|-------|----|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min   | max   |    |
| Все признаки            |                                |                   | 38.49     | 3.70     | 36.67 | 42.50 | 50 |
| Метод главных компонент |                                |                   | 40.00     | 2.45     | 37.50 | 43.33 | 6  |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 50.20     | 5.12     | 45.00 | 59.17 | 24 |
|                         | Многокритериальная оптимизация | SPEA              | 51.67     | 5.43     | 45.83 | 57.50 | 3  |
|                         |                                | NSGA-2            | 50.83     | 6.98     | 43.33 | 60.00 | 4  |
|                         |                                | VEGA              | 55.28     | 1.68     | 53.33 | 58.33 | 23 |
|                         |                                | SelfCOMO GA       | 54.67     | 3.12     | 48.33 | 63.33 | 5  |

Таблица Б10. Точность классификации эмоций (%) для различных подходов к снижению размерности, входные данные – аудиопризнаки ( $n$  – усредненное по экспериментам число признаков, соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |       |       | n  |
|-------------------------|--------------------------------|-------------------|-----------|----------|-------|-------|----|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min   | max   |    |
| Все признаки            |                                |                   | 30.57     | 4.11     | 29.17 | 35.00 | 50 |
| Метод главных компонент |                                |                   | 33.14     | 2.78     | 27.50 | 35.83 | 11 |
| Отбор признаков         | Однокритериальная оптимизация  | Козволюционный ГА | 39.17     | 6.11     | 29.17 | 50.83 | 29 |
|                         | Многокритериальная оптимизация | SPEA              | 35.56     | 4.38     | 29.17 | 40.00 | 1  |
|                         |                                | NSGA-2            | 40.00     | 4.67     | 35.00 | 45.83 | 2  |
|                         |                                | VEGA              | 39.17     | 4.16     | 35.83 | 47.50 | 3  |
|                         |                                | SelfCOMO GA       | 44.56     | 3.83     | 40.00 | 54.17 | 4  |

Таблица Б11. Точность классификации эмоций (%) для различных подходов к снижению размерности, входные данные – лицевые маркеры + аудиопризнаки ( $n$  - усредненное по экспериментам число признаков, соответствующее лучшей точности)

| Метод отбора признаков  |                                |                   | Точность  |          |       |       | n   |
|-------------------------|--------------------------------|-------------------|-----------|----------|-------|-------|-----|
|                         |                                |                   | $\bar{x}$ | $\sigma$ | min   | max   |     |
| Все признаки            |                                |                   | 32.57     | 3.59     | 28.33 | 38.33 | 100 |
| Метод главных компонент |                                |                   | 36.67     | 4.65     | 32.50 | 42.50 | 15  |
| Отбор признаков         | Однокритериальная оптимизация  | Коэволюционный ГА | 42.5      | 4.32     | 38.33 | 46.67 | 49  |
|                         | Многокритериальная оптимизация | SPEA              | 47.5      | 4.90     | 45.00 | 56.67 | 3   |
|                         |                                | NSGA-2            | 51.11     | 4.31     | 43.33 | 56.67 | 5   |
|                         |                                | VEGA              | 52.22     | 1.90     | 50.00 | 55.00 | 22  |
|                         |                                | SelfCOMO GA       | 64.31     | 3.06     | 60.83 | 68.33 | 24  |

Таблица Б12. Точность классификации эмоций (%) при оптимизации коллектива нейронных сетей различными алгоритмами оптимизации, дикторонезависимая классификация, входные данные – лицевые маркеры

| Алгоритм<br>оптимизации<br>(число критериев) | Схема слияния выходов<br>ансамбля  | Точность     |          |       |       |
|----------------------------------------------|------------------------------------|--------------|----------|-------|-------|
|                                              |                                    | $\bar{x}$    | $\sigma$ | min   | max   |
| Коэволюционный<br>ГА (1)                     | -                                  | 46.39        | 2.45     | 42.50 | 49.17 |
| SPEA (2)                                     | Голосование                        | 50.00        | 3.12     | 47.50 | 55.00 |
|                                              | Усреднение вероятностей<br>классов | 45.00        | 1.23     | 43.33 | 47.50 |
|                                              | Мета-классификация                 | 53.89        | 2.35     | 51.67 | 58.33 |
| NSGA-2 (2)                                   | Голосование                        | 51.11        | 2.76     | 44.17 | 55.83 |
|                                              | Усреднение вероятностей<br>классов | 45.06        | 3.11     | 40.83 | 48.33 |
|                                              | Мета-классификация                 | <b>55.56</b> | 2.09     | 52.50 | 58.33 |
| VEGA (2)                                     | Голосование                        | 48.33        | 1.87     | 45.00 | 50.83 |
|                                              | Усреднение вероятностей<br>классов | 36.72        | 2.45     | 32.50 | 39.17 |
|                                              | Мета-классификация                 | 51.11        | 1.53     | 48.33 | 52.50 |
| SelfCOMOGA (2)                               | Голосование                        | 49.44        | 2.78     | 42.50 | 50.00 |
|                                              | Усреднение вероятностей<br>классов | 47.78        | 2.54     | 45.00 | 52.50 |
|                                              | Мета-классификация                 | 53.36        | 2.13     | 48.33 | 55.00 |



Таблица Б13. Точность классификации эмоций (%) при оптимизации коллектива нейронных сетей различными алгоритмами оптимизации, дикторонезависимая классификация, входные данные – аудиопризнаки

| Алгоритм оптимизации<br>(число критериев) | Схема слияния выходов ансамбля  | Точность     |          |       |       |
|-------------------------------------------|---------------------------------|--------------|----------|-------|-------|
|                                           |                                 | $\bar{x}$    | $\sigma$ | min   | max   |
| Коэволюционный<br>ГА (1)                  | -                               | 41.05        | 2.65     | 35.83 | 45.83 |
| SPEA (2)                                  | Голосование                     | 37.33        | 3.90     | 34.17 | 43.33 |
|                                           | Усреднение вероятностей классов | 30.28        | 2.87     | 28.33 | 35.00 |
|                                           | Мета-классификация              | 43.5         | 1.54     | 40.00 | 45.83 |
| NSGA-2 (2)                                | Голосование                     | 35.78        | 1.86     | 33.33 | 40.83 |
|                                           | Усреднение вероятностей классов | 33.06        | 2.76     | 30.00 | 37.50 |
|                                           | Мета-классификация              | 39.78        | 1.32     | 37.50 | 43.33 |
| VEGA (2)                                  | Голосование                     | 31.39        | 1.06     | 29.17 | 32.50 |
|                                           | Усреднение вероятностей классов | 29.22        | 1.21     | 28.33 | 31.67 |
|                                           | Мета-классификация              | 38.72        | 1.41     | 37.50 | 40.00 |
| SelfCOMOGA (2)                            | Голосование                     | 42.43        | 2.76     | 40.00 | 48.33 |
|                                           | Усреднение вероятностей классов | 38.83        | 2.06     | 33.33 | 42.50 |
|                                           | Мета-классификация              | <b>47.34</b> | 1.54     | 45.83 | 50.00 |

Таблица Б14. Точность классификации эмоций (%) при оптимизации коллектива нейронных сетей различными алгоритмами оптимизации, дикторонезависимая классификация, входные данные – лицевые маркеры + аудиопризнаки

| Алгоритм<br>оптимизации<br>(число критериев) | Схема слияния выходов<br>ансамбля  | Точность     |          |       |       |
|----------------------------------------------|------------------------------------|--------------|----------|-------|-------|
|                                              |                                    | $\bar{x}$    | $\sigma$ | min   | max   |
| Коэволюционный ГА<br>(1)                     | -                                  | 55.00        | 2.14     | 51.67 | 60.00 |
| SPEA (2)                                     | Голосование                        | 62.67        | 1.35     | 60.83 | 64.17 |
|                                              | Усреднение вероятностей<br>классов | 49.67        | 1.48     | 47.50 | 52.50 |
|                                              | Мета-классификация                 | 59.44        | 2.89     | 55.00 | 64.17 |
| NSGA-2 (2)                                   | Голосование                        | 41.11        | 1.78     | 37.50 | 42.50 |
|                                              | Усреднение вероятностей<br>классов | 49.83        | 1.56     | 46.67 | 53.33 |
|                                              | Мета-классификация                 | 58.78        | 1.06     | 57.50 | 60.00 |
| VEGA (2)                                     | Голосование                        | 42.78        | 0.76     | 42.50 | 44.17 |
|                                              | Усреднение вероятностей<br>классов | 39.17        | 1.34     | 38.33 | 40.83 |
|                                              | Мета-классификация                 | 52.33        | 0.38     | 51.67 | 53.33 |
| SelfCOMOGA (2)                               | Голосование                        | 54.33        | 1.87     | 52.50 | 58.33 |
|                                              | Усреднение вероятностей<br>классов | 48.17        | 2.54     | 45.83 | 49.17 |
|                                              | Мета-классификация                 | <b>69.20</b> | 3.12     | 63.33 | 70.83 |

**Приложение В. Результаты экспериментов по исследованию эффективности гибридного алгоритма обучения конволюционной нейронной сети на задаче распознавания эмоций**

Таблица В1. Сравнение средней точности (%) распознавания эмоций с помощью конволюционной нейронной сети (КНС): а) обратное распространение ошибки (ОРО); б) генетический алгоритм (ГА); в) гибридный алгоритм (ГА + ОРО); дикторозависимая классификация; размер входных изображений 40×40 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 92.27                 | 1.12     | 90.97 | 94.51 | 85.48     | 2.35     | 80.67 | 87.52 | 95.31       | 1.22     | 93.63 | 96.56 |
|        | Отвращение  | 87.91                 | 1.56     | 85.89 | 90.69 | 79.77     | 2.68     | 73.24 | 86.57 | 91.45       | 1.57     | 88.71 | 94.43 |
|        | Страх       | 87.28                 | 1.71     | 84.35 | 90.22 | 82.59     | 2.77     | 77.62 | 85.09 | 90.15       | 1.47     | 86.92 | 91.57 |
|        | Счастье     | 88.30                 | 1.56     | 85.36 | 90.76 | 87.53     | 2.21     | 86.45 | 91.24 | 92.98       | 2.11     | 88.96 | 96.56 |
|        | Нейтральная | 87.43                 | 2.46     | 81.85 | 89.87 | 87.53     | 2.57     | 82.58 | 90.28 | 92.98       | 2.42     | 89.64 | 95.98 |
|        | Грусть      | 87.31                 | 1.45     | 85.72 | 89.98 | 88.43     | 2.89     | 83.09 | 93.84 | 90.79       | 1.67     | 87.59 | 92.13 |
|        | Удивление   | 88.97                 | 2.45     | 84.59 | 92.28 | 83.44     | 3.08     | 77.42 | 87.44 | 91.82       | 1.23     | 89.49 | 93.20 |

Таблица В2. Сравнение средней точности (%) распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторозависимая классификация; размер входных изображений 50×50 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 93.39                 | 1.45     | 91.86 | 94.80 | 79.62     | 3.12     | 77.05 | 85.25 | 95.87       | 1.43     | 93.43 | 98.95 |
|        | Отвращение  | 89.64                 | 1.18     | 87.05 | 90.67 | 80.94     | 1.90     | 77.63 | 84.07 | 94.36       | 1.08     | 91.81 | 95.78 |
|        | Страх       | 88.24                 | 1.36     | 86.77 | 90.55 | 78.10     | 2.29     | 73.99 | 83.35 | 91.61       | 0.87     | 89.99 | 92.90 |
|        | Счастье     | 90.40                 | 1.67     | 87.32 | 90.77 | 84.22     | 2.55     | 80.63 | 87.19 | 94.06       | 1.76     | 91.01 | 95.18 |
|        | Нейтральная | 88.31                 | 2.78     | 82.79 | 92.88 | 71.32     | 2.89     | 68.88 | 78.31 | 97.29       | 1.09     | 95.55 | 98.17 |

Продолжение таблицы В2.

| Данные |           | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-----------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |           | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |           | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
|        | Грусть    | 88.67                 | 1.67     | 86.01 | 90.81 | 85.66     | 2.14     | 82.78 | 89.47 | 92.09       | 1.23     | 90.11 | 94.23 |
|        | Удивление | 90.51                 | 1.69     | 86.92 | 93.29 | 85.59     | 2.49     | 82.80 | 90.57 | 94.29       | 1.88     | 92.45 | 96.17 |

Таблица В3. Сравнение средней точности (%) распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторозависимая классификация; размер входных изображений 70×70 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 96.45                 | 0.89     | 95.53 | 98.27 | 88.89     | 2.67     | 87.34 | 91.87 | 97.31       | 1.24     | 95.25 | 98.29 |
|        | Отвращение  | 93.73                 | 0.67     | 92.92 | 94.95 | 82.64     | 3.76     | 76.33 | 88.47 | 94.67       | 1.87     | 91.45 | 96.07 |
|        | Страх       | 89.04                 | 0.76     | 87.66 | 90.82 | 84.64     | 2.76     | 82.72 | 89.57 | 91.09       | 1.65     | 88.72 | 93.59 |
|        | Счастье     | 90.30                 | 1.48     | 88.29 | 93.28 | 74.72     | 3.12     | 71.55 | 78.41 | 92.02       | 1.41     | 88.72 | 94.53 |
|        | Нейтральная | 93.59                 | 1.34     | 91.52 | 95.12 | 61.65     | 3.18     | 57.21 | 66.59 | 90.82       | 0.56     | 89.99 | 93.68 |
|        | Грусть      | 91.40                 | 1.12     | 90.42 | 92.40 | 85.94     | 2.96     | 80.80 | 91.83 | 93.52       | 1.56     | 91.12 | 95.52 |
|        | Удивление   | 93.96                 | 1.78     | 90.99 | 95.40 | 73.77     | 2.78     | 68.73 | 79.57 | 94.56       | 0.75     | 93.19 | 95.54 |

Таблица В4. Сравнение средней F-меры распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторозависимая классификация; размер входных изображений 40×40 пикселей

| Данные |            | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |            | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |            | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость     | 0.759                 | 0.044    | 0.665 | 0.834 | 0.773     | 0.049    | 0.712 | 0.854 | 0.885       | 0.066    | 0.788 | 0.962 |
|        | Отвращение | 0.628                 | 0.048    | 0.595 | 0.708 | 0.663     | 0.024    | 0.603 | 0.696 | 0.760       | 0.036    | 0.730 | 0.792 |
|        | Страх      | 0.514                 | 0.026    | 0.472 | 0.569 | 0.702     | 0.063    | 0.600 | 0.796 | 0.674       | 0.040    | 0.592 | 0.722 |
|        | Счастье    | 0.517                 | 0.060    | 0.372 | 0.562 | 0.759     | 0.020    | 0.724 | 0.798 | 0.782       | 0.029    | 0.722 | 0.827 |

Продолжение таблицы В4.

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
|        | Нейтральная | 0.698                 | 0.026    | 0.669 | 0.722 | 0.759     | 0.022    | 0.727 | 0.789 | 0.782       | 0.036    | 0.721 | 0.824 |
|        | Грусть      | 0.479                 | 0.051    | 0.443 | 0.568 | 0.737     | 0.042    | 0.716 | 0.822 | 0.689       | 0.036    | 0.619 | 0.723 |
|        | Удивление   | 0.504                 | 0.020    | 0.458 | 0.539 | 0.685     | 0.022    | 0.675 | 0.714 | 0.689       | 0.023    | 0.659 | 0.705 |

Таблица В5. Сравнение средней F-меры распознавания эмоций с помощью КНС:

а) ОРО; б) ГА; в) ГА + ОРО; дикторозависимая классификация; размер входных изображений 50×50 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 0.796                 | 0.018    | 0.770 | 0.838 | 0.683     | 0.043    | 0.625 | 0.740 | 0.894       | 0.048    | 0.822 | 0.970 |
|        | Отвращение  | 0.705                 | 0.034    | 0.638 | 0.751 | 0.679     | 0.062    | 0.558 | 0.787 | 0.861       | 0.022    | 0.820 | 0.896 |
|        | Страх       | 0.568                 | 0.047    | 0.491 | 0.659 | 0.622     | 0.055    | 0.587 | 0.776 | 0.750       | 0.036    | 0.681 | 0.797 |
|        | Счастье     | 0.632                 | 0.041    | 0.579 | 0.690 | 0.683     | 0.046    | 0.546 | 0.751 | 0.824       | 0.054    | 0.744 | 0.906 |
|        | Нейтральная | 0.721                 | 0.028    | 0.647 | 0.788 | 0.642     | 0.025    | 0.632 | 0.667 | 0.946       | 0.043    | 0.900 | 0.983 |
|        | Грусть      | 0.573                 | 0.015    | 0.545 | 0.589 | 0.673     | 0.062    | 0.578 | 0.743 | 0.738       | 0.040    | 0.704 | 0.781 |
|        | Удивление   | 0.602                 | 0.021    | 0.547 | 0.623 | 0.670     | 0.026    | 0.625 | 0.696 | 0.797       | 0.043    | 0.745 | 0.868 |

Таблица В6. Сравнение средней F-меры распознавания эмоций с помощью КНС:

а) ОРО; б) ГА; в) ГА + ОРО; дикторозависимая классификация; размер входных изображений 70×70 пикселей

| Данные |            | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |            | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |            | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость     | 0.898                 | 0.035    | 0.854 | 0.992 | 0.804     | 0.028    | 0.753 | 0.869 | 0.939       | 0.051    | 0.852 | 0.992 |
|        | Отвращение | 0.694                 | 0.047    | 0.645 | 0.746 | 0.664     | 0.038    | 0.640 | 0.763 | 0.777       | 0.050    | 0.692 | 0.821 |

Продолжение таблицы В6.

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
|        | Страх       | 0.644                 | 0.025    | 0.625 | 0.690 | 0.684     | 0.053    | 0.565 | 0.722 | 0.710       | 0.046    | 0.652 | 0.763 |
|        | Счастье     | 0.763                 | 0.037    | 0.701 | 0.811 | 0.582     | 0.041    | 0.508 | 0.625 | 0.679       | 0.051    | 0.633 | 0.748 |
|        | Нейтральная | 0.749                 | 0.039    | 0.724 | 0.800 | 0.533     | 0.058    | 0.465 | 0.657 | 0.813       | 0.058    | 0.706 | 0.871 |
|        | Грусть      | 0.691                 | 0.032    | 0.626 | 0.732 | 0.700     | 0.008    | 0.689 | 0.723 | 0.732       | 0.046    | 0.679 | 0.799 |
|        | Удивление   | 0.772                 | 0.062    | 0.648 | 0.802 | 0.571     | 0.039    | 0.508 | 0.653 | 0.790       | 0.016    | 0.760 | 0.816 |

Таблица В7. Сравнение средней точности (%) распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных изображений 40×40 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 93.31                 | 1.51     | 90.67 | 96.23 | 87.22     | 2.80     | 83.31 | 92.52 | 95.26       | 2.14     | 90.12 | 97.29 |
|        | Отвращение  | 92.80                 | 1.63     | 90.11 | 96.11 | 84.52     | 2.17     | 82.28 | 87.56 | 94.95       | 3.02     | 91.10 | 99.18 |
|        | Страх       | 90.15                 | 1.64     | 88.45 | 92.26 | 81.96     | 2.33     | 79.53 | 85.42 | 90.61       | 1.88     | 89.02 | 93.07 |
|        | Счастье     | 90.27                 | 0.92     | 87.40 | 91.91 | 86.64     | 1.33     | 84.99 | 90.93 | 95.81       | 2.15     | 90.10 | 98.22 |
|        | Нейтральная | 92.93                 | 1.97     | 90.39 | 96.74 | 85.66     | 1.91     | 80.90 | 87.49 | 95.81       | 1.71     | 92.27 | 96.75 |
|        | Грусть      | 90.58                 | 0.39     | 89.75 | 91.19 | 90.50     | 1.25     | 88.98 | 92.29 | 94.18       | 2.17     | 89.86 | 95.77 |
|        | Удивление   | 90.66                 | 2.68     | 88.23 | 94.34 | 91.26     | 0.53     | 90.39 | 91.91 | 94.04       | 2.75     | 89.61 | 97.47 |

Таблица В8. Сравнение средней точности (%) распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных изображений 50×50 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 94.75                 | 1.53     | 92.34 | 98.99 | 81.41     | 1.93     | 79.44 | 86.72 | 94.26       | 1.75     | 90.92 | 96.49 |
|        | Отвращение  | 94.37                 | 2.74     | 88.52 | 99.86 | 77.88     | 1.02     | 76.81 | 79.39 | 94.22       | 1.79     | 90.98 | 96.66 |
|        | Страх       | 90.83                 | 2.56     | 85.15 | 90.85 | 87.88     | 0.74     | 87.12 | 89.35 | 93.13       | 1.42     | 91.86 | 95.87 |
|        | Счастье     | 92.71                 | 1.47     | 91.32 | 95.03 | 42.42     | 1.71     | 40.29 | 45.72 | 89.12       | 2.33     | 85.33 | 91.35 |
|        | Нейтральная | 94.15                 | 1.39     | 92.13 | 97.18 | 77.12     | 2.22     | 71.49 | 81.51 | 96.35       | 1.28     | 93.92 | 98.61 |
|        | Грусть      | 91.78                 | 2.57     | 85.98 | 94.39 | 88.27     | 1.30     | 86.88 | 91.46 | 94.31       | 0.51     | 93.93 | 95.27 |
|        | Удивление   | 92.51                 | 1.49     | 89.69 | 94.20 | 70.36     | 2.49     | 64.40 | 76.65 | 93.05       | 1.42     | 91.56 | 95.48 |

Таблица В9. Сравнение средней точности (%) распознавания эмоций с помощью КНС: а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных изображений 70×70 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 93.71                 | 1.18     | 91.68 | 96.90 | 85.96     | 1.55     | 82.53 | 89.15 | 93.94       | 2.89     | 89.18 | 96.48 |
|        | Отвращение  | 92.47                 | 0.92     | 90.69 | 93.54 | 81.39     | 1.92     | 80.90 | 88.21 | 93.91       | 2.25     | 90.49 | 95.64 |
|        | Страх       | 94.11                 | 1.42     | 92.42 | 95.74 | 80.08     | 1.61     | 77.34 | 82.38 | 95.10       | 2.41     | 91.65 | 99.93 |
|        | Счастье     | 92.74                 | 1.46     | 89.83 | 95.98 | 75.03     | 1.53     | 72.71 | 76.57 | 94.29       | 0.82     | 91.72 | 96.55 |
|        | Нейтральная | 93.05                 | 1.94     | 91.93 | 94.43 | 70.22     | 1.41     | 68.48 | 71.47 | 94.73       | 2.18     | 89.27 | 97.63 |
|        | Грусть      | 92.71                 | 0.72     | 91.65 | 94.31 | 85.63     | 1.49     | 83.05 | 88.37 | 93.72       | 1.73     | 90.66 | 95.53 |
|        | Удивление   | 95.71                 | 1.21     | 94.68 | 97.47 | 79.76     | 2.73     | 78.84 | 82.79 | 96.34       | 3.22     | 90.85 | 99.64 |

Таблица В10. Сравнение средней F-меры распознавания эмоций с помощью КНС:  
а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных  
изображений 40×40 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 0.796                 | 0.038    | 0.701 | 0.876 | 0.761     | 0.045    | 0.725 | 0.860 | 0.874       | 0.056    | 0.796 | 0.948 |
|        | Отвращение  | 0.798                 | 0.058    | 0.698 | 0.868 | 0.711     | 0.030    | 0.656 | 0.766 | 0.877       | 0.052    | 0.819 | 0.951 |
|        | Страх       | 0.662                 | 0.051    | 0.656 | 0.746 | 0.690     | 0.045    | 0.580 | 0.718 | 0.725       | 0.025    | 0.664 | 0.786 |
|        | Счастье     | 0.639                 | 0.034    | 0.591 | 0.684 | 0.730     | 0.021    | 0.714 | 0.756 | 0.867       | 0.046    | 0.814 | 0.983 |
|        | Нейтральная | 0.835                 | 0.033    | 0.762 | 0.879 | 0.740     | 0.046    | 0.638 | 0.776 | 0.918       | 0.047    | 0.839 | 1.007 |
|        | Грусть      | 0.719                 | 0.050    | 0.647 | 0.799 | 0.760     | 0.021    | 0.710 | 0.792 | 0.851       | 0.028    | 0.791 | 0.875 |
|        | Удивление   | 0.595                 | 0.028    | 0.542 | 0.644 | 0.795     | 0.047    | 0.733 | 0.891 | 0.775       | 0.053    | 0.687 | 0.822 |

Таблица В11. Сравнение средней F-меры распознавания эмоций с помощью КНС:  
а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных  
изображений 50×50 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 0.846                 | 0.025    | 0.793 | 0.909 | 0.693     | 0.031    | 0.653 | 0.739 | 0.853       | 0.049    | 0.775 | 0.889 |
|        | Отвращение  | 0.845                 | 0.065    | 0.750 | 0.927 | 0.661     | 0.041    | 0.627 | 0.692 | 0.843       | 0.066    | 0.695 | 0.969 |
|        | Страх       | 0.687                 | 0.037    | 0.589 | 0.730 | 0.685     | 0.063    | 0.573 | 0.774 | 0.822       | 0.037    | 0.755 | 0.891 |
|        | Счастье     | 0.740                 | 0.054    | 0.692 | 0.875 | 0.389     | 0.034    | 0.333 | 0.455 | 0.647       | 0.060    | 0.535 | 0.723 |
|        | Нейтральная | 0.862                 | 0.034    | 0.797 | 0.923 | 0.649     | 0.041    | 0.599 | 0.731 | 0.924       | 0.027    | 0.865 | 0.937 |
|        | Грусть      | 0.762                 | 0.042    | 0.713 | 0.798 | 0.756     | 0.047    | 0.638 | 0.858 | 0.855       | 0.027    | 0.796 | 0.916 |
|        | Удивление   | 0.690                 | 0.032    | 0.643 | 0.737 | 0.522     | 0.013    | 0.500 | 0.539 | 0.775       | 0.031    | 0.717 | 0.810 |



Таблица В12. Сравнение средней F-меры распознавания эмоций с помощью КНС:  
а) ОРО; б) ГА; в) ГА + ОРО; дикторонезависимая классификация; размер входных  
изображений 70×70 пикселей

| Данные |             | Алгоритм обучения КНС |          |       |       |           |          |       |       |             |          |       |       |
|--------|-------------|-----------------------|----------|-------|-------|-----------|----------|-------|-------|-------------|----------|-------|-------|
|        |             | а) ОРО                |          |       |       | б) ГА     |          |       |       | в) ГА + ОРО |          |       |       |
|        |             | $\bar{x}$             | $\sigma$ | min   | max   | $\bar{x}$ | $\sigma$ | min   | max   | $\bar{x}$   | $\sigma$ | min   | max   |
| Эмоция | Злость      | 0.811                 | 0.041    | 0.720 | 0.851 | 0.771     | 0.031    | 0.757 | 0.803 | 0.839       | 0.061    | 0.746 | 0.879 |
|        | Отвращение  | 0.825                 | 0.078    | 0.736 | 0.909 | 0.674     | 0.031    | 0.645 | 0.715 | 0.863       | 0.037    | 0.765 | 0.890 |
|        | Страх       | 0.817                 | 0.028    | 0.758 | 0.839 | 0.663     | 0.012    | 0.643 | 0.666 | 0.841       | 0.050    | 0.736 | 0.946 |
|        | Счастье     | 0.737                 | 0.032    | 0.689 | 0.771 | 0.585     | 0.040    | 0.509 | 0.627 | 0.797       | 0.063    | 0.695 | 0.887 |
|        | Нейтральная | 0.820                 | 0.049    | 0.766 | 0.897 | 0.598     | 0.070    | 0.536 | 0.769 | 0.884       | 0.052    | 0.704 | 0.935 |
|        | Грусть      | 0.830                 | 0.045    | 0.742 | 0.887 | 0.717     | 0.037    | 0.650 | 0.771 | 0.795       | 0.019    | 0.752 | 0.833 |
|        | Удивление   | 0.837                 | 0.030    | 0.798 | 0.891 | 0.616     | 0.036    | 0.546 | 0.682 | 0.864       | 0.058    | 0.793 | 0.923 |

**Приложение Г. Результаты экспериментов по исследованию эффективности  
обобщенного метода анализа гетерогенных данных на задаче распознавания  
эмоций**

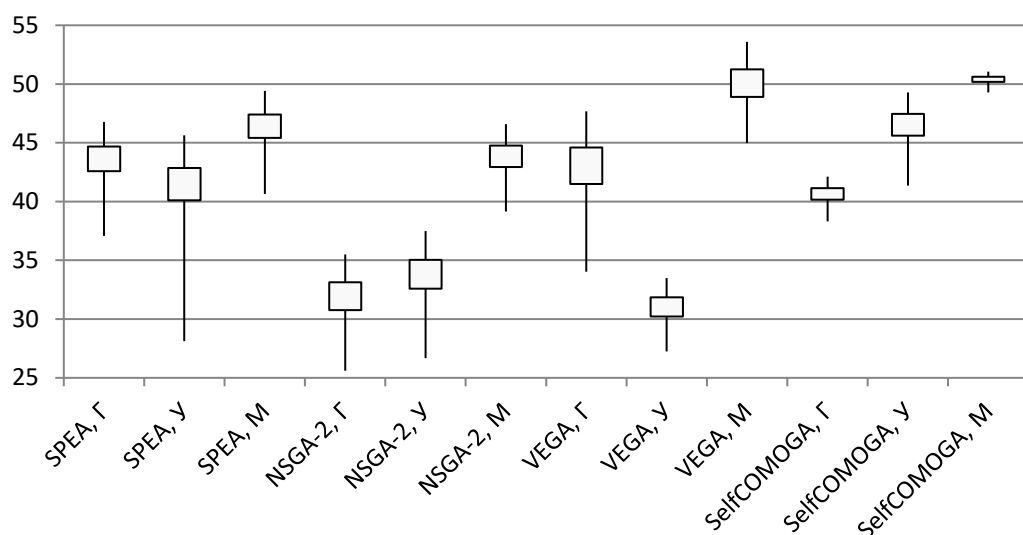


Рисунок Г1. Точность распознавания эмоций (%), входные данные - лицевые маркеры, обобщенный подход; Г - голосование, У - усреднение апостериорных вероятностей классов, М - мета-классификация; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

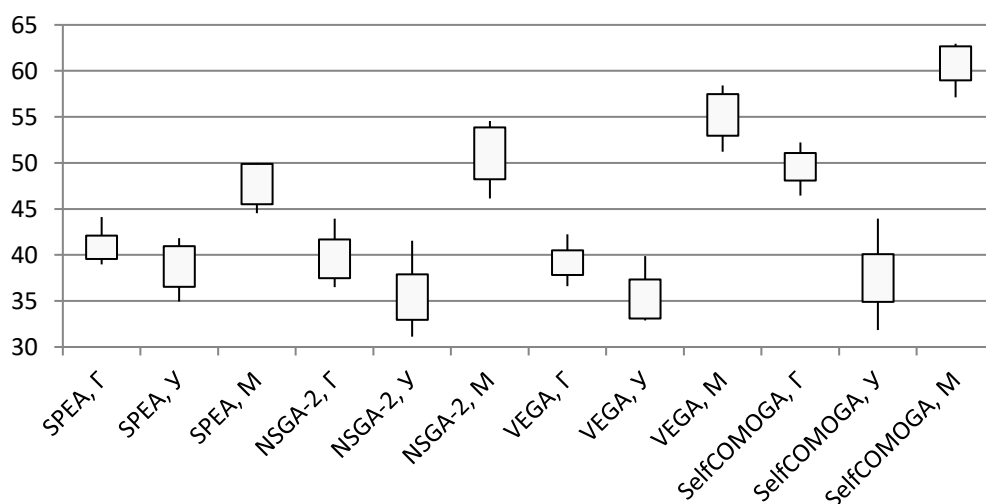


Рисунок Г2. Точность распознавания эмоций (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей

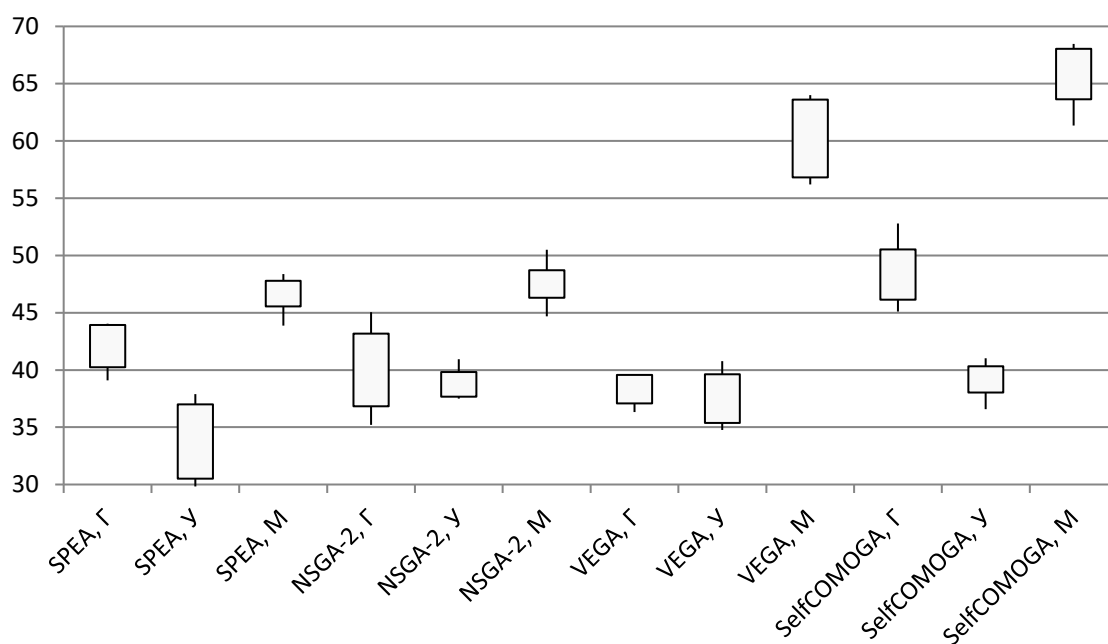


Рисунок Г3. Точность распознавания эмоций (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 70×70 пикселей

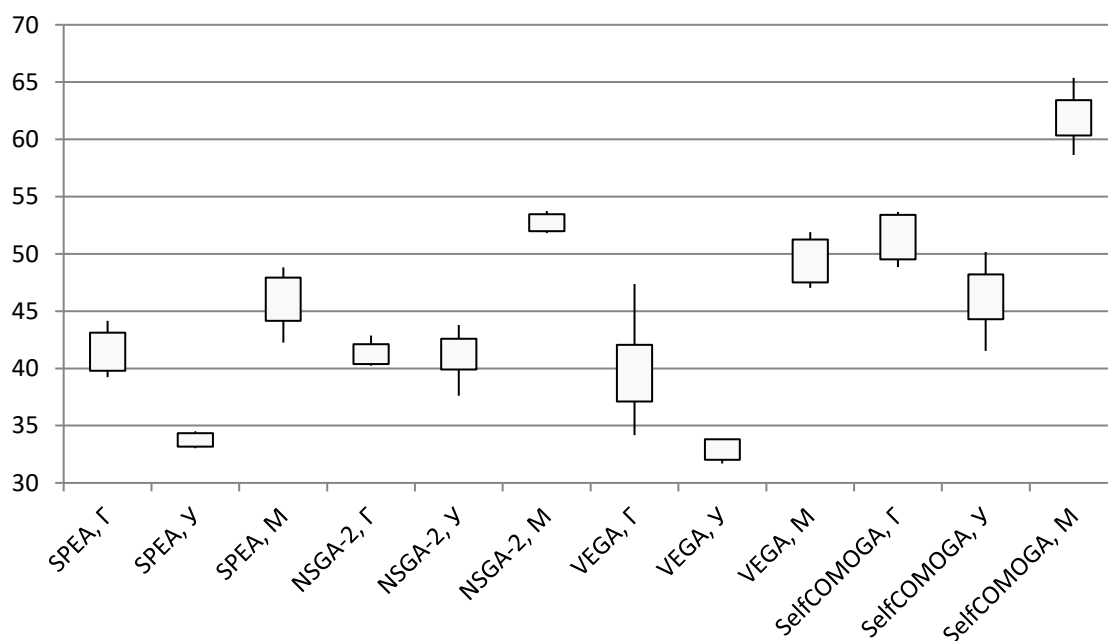


Рисунок Г4. Точность распознавания эмоций (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 100×100 пикселей

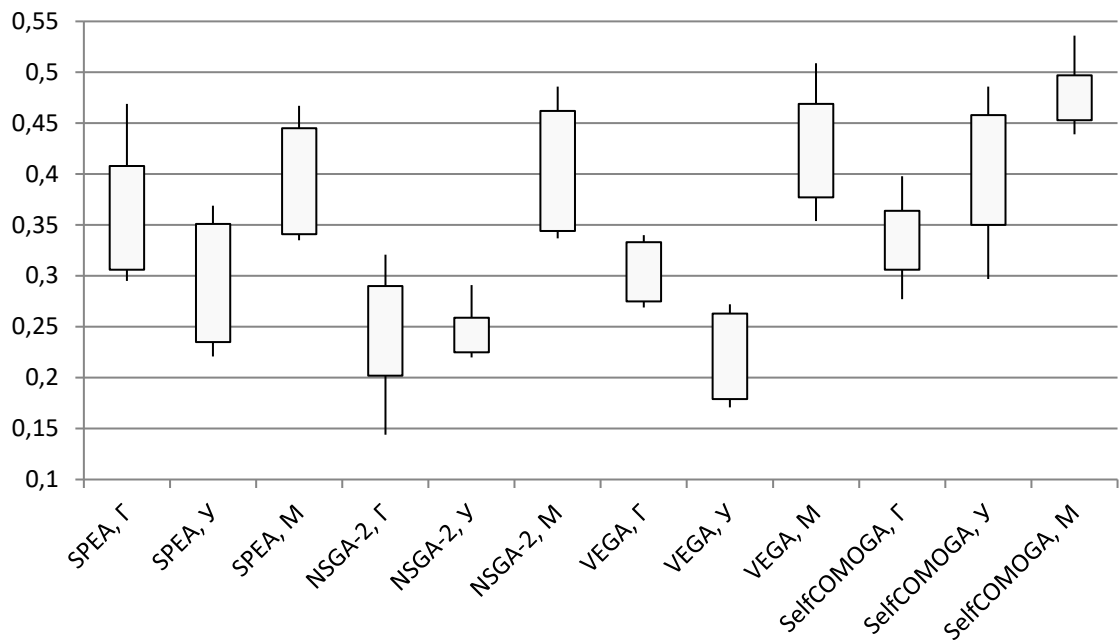


Рисунок Г5. F-мера распознавания эмоций, (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

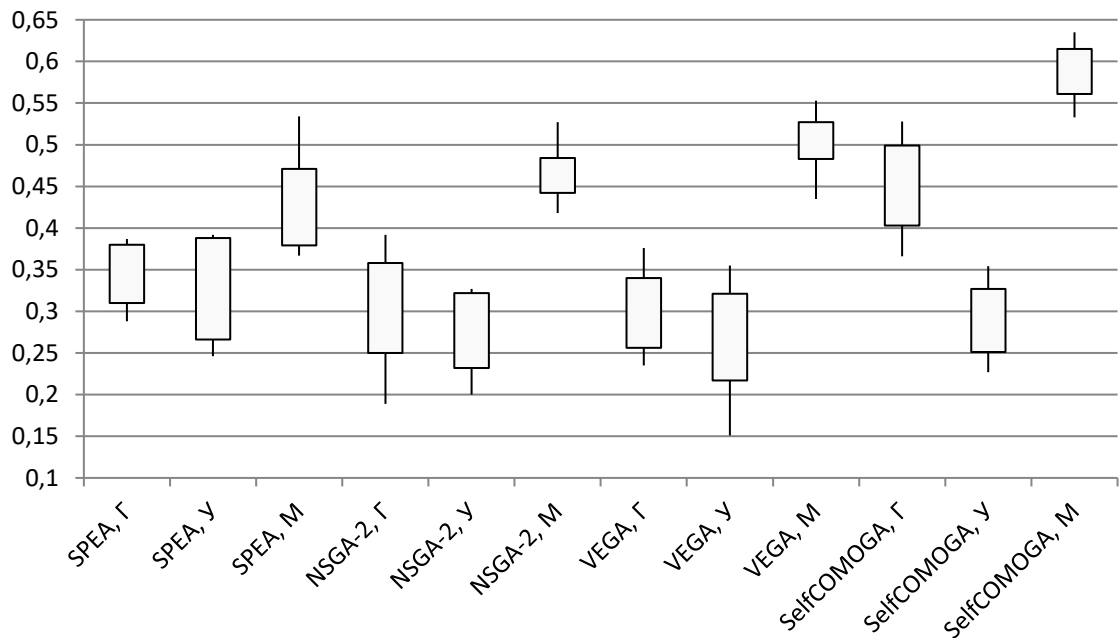


Рисунок Г6. F-мера распознавания эмоций, (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей

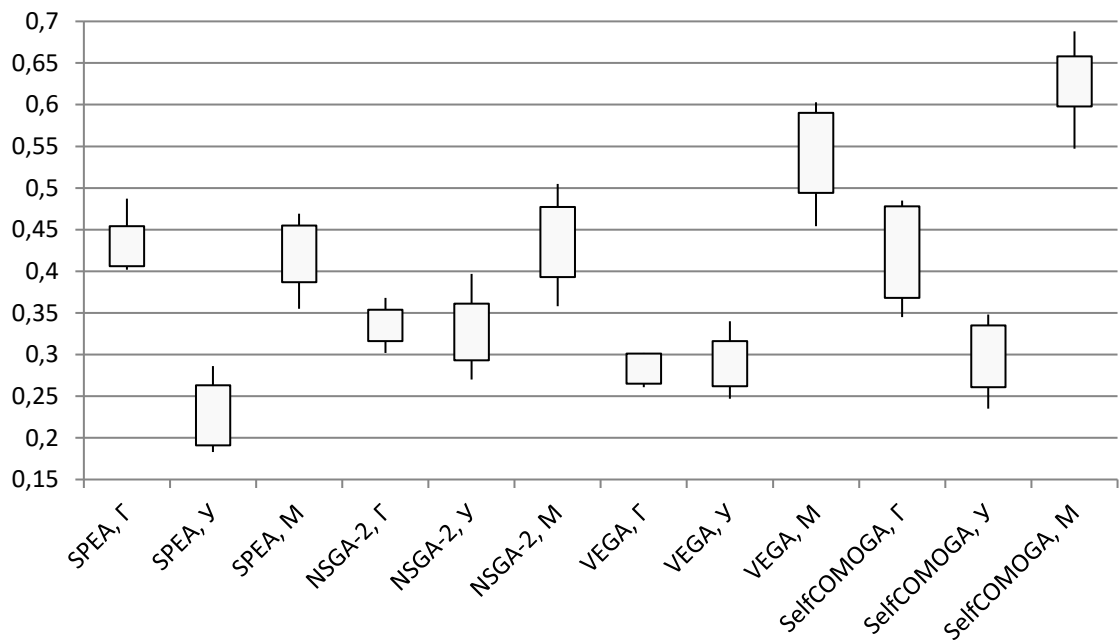


Рисунок Г7. F-мера распознавания эмоций, (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 70×70 пикселей

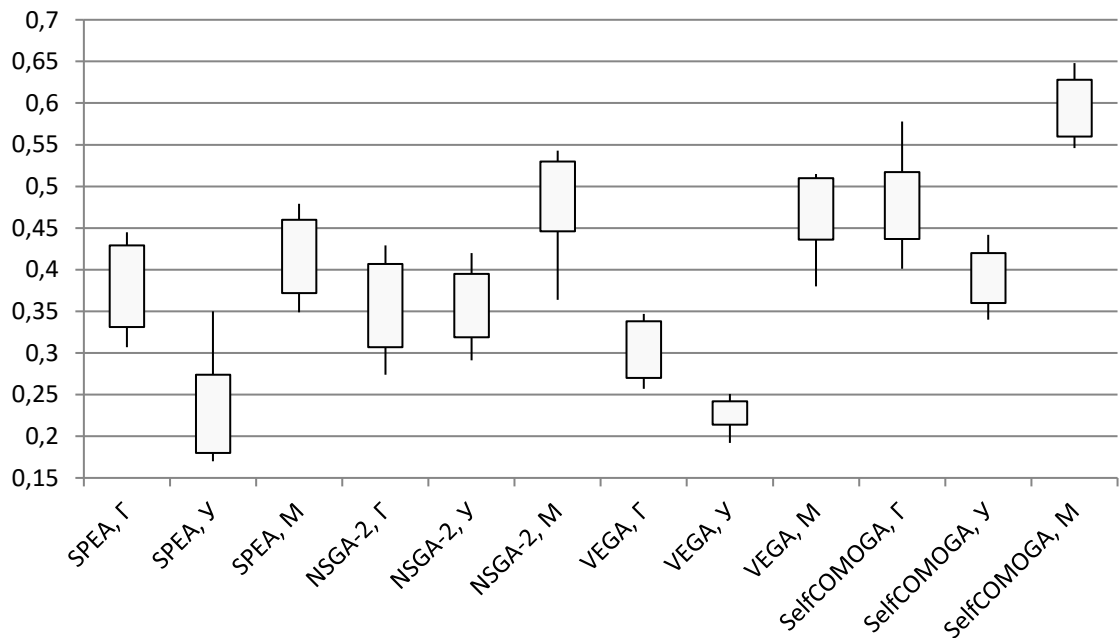


Рисунок Г8. F-мера распознавания эмоций, (%), входные данные - лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 100×100 пикселей

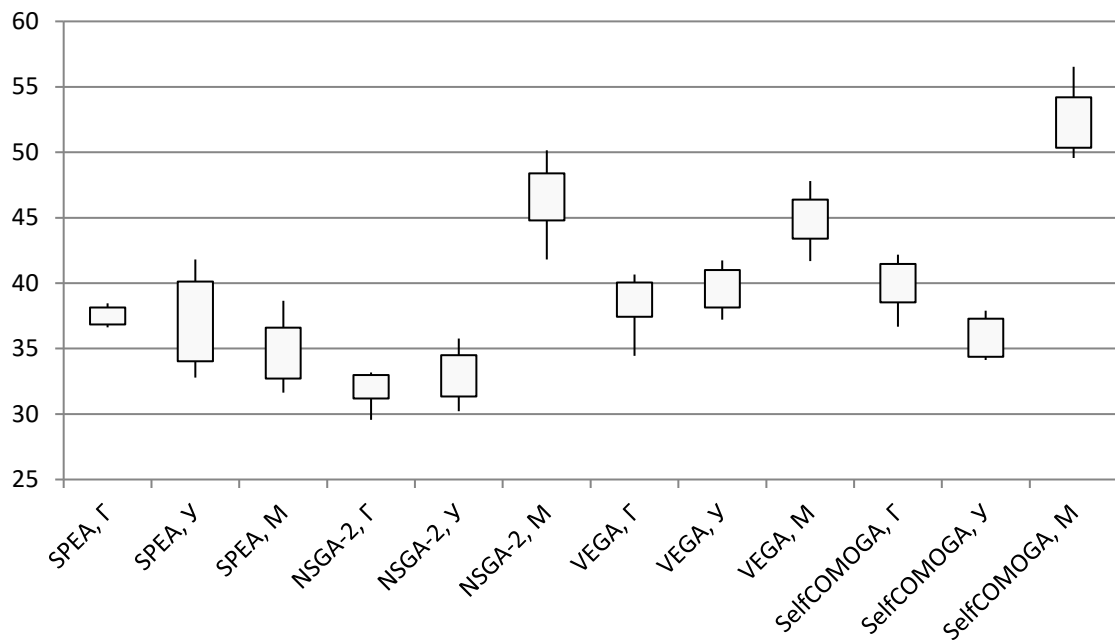


Рисунок Г9. Точность распознавания эмоций (%), входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

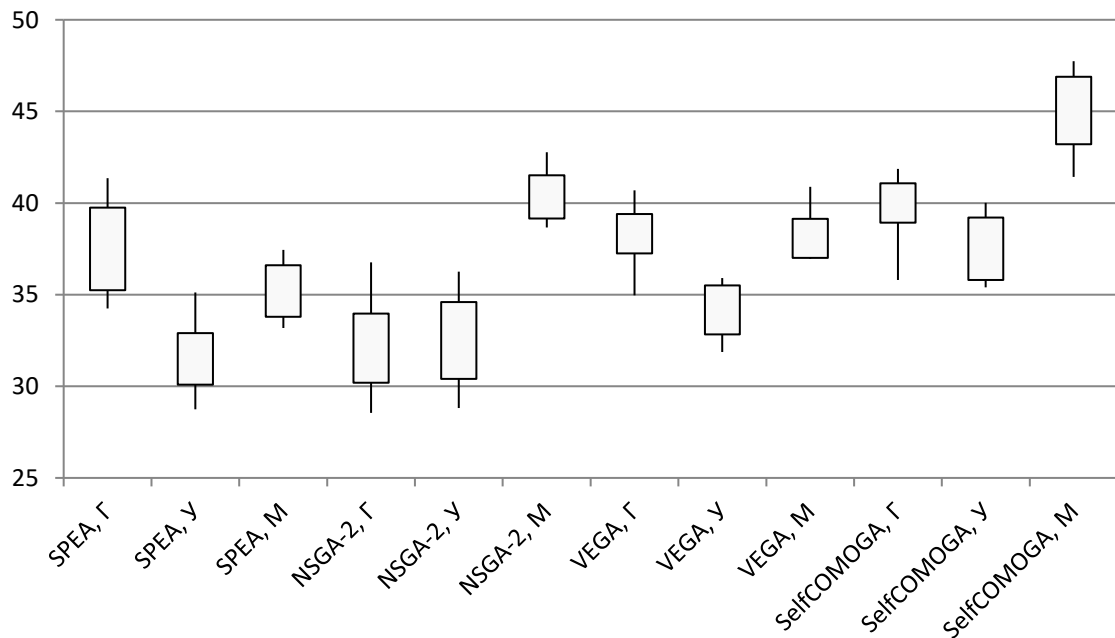


Рисунок Г10. Точность распознавания эмоций (%), входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей

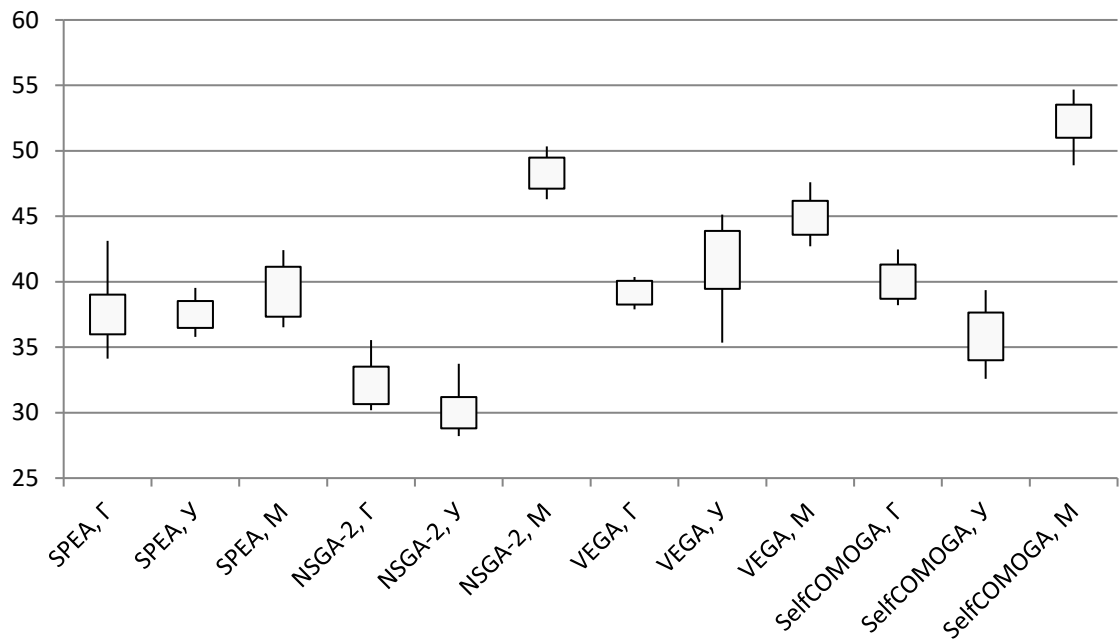


Рисунок Г11. Точность распознавания эмоций (%), входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети  $70 \times 70$  пикселей

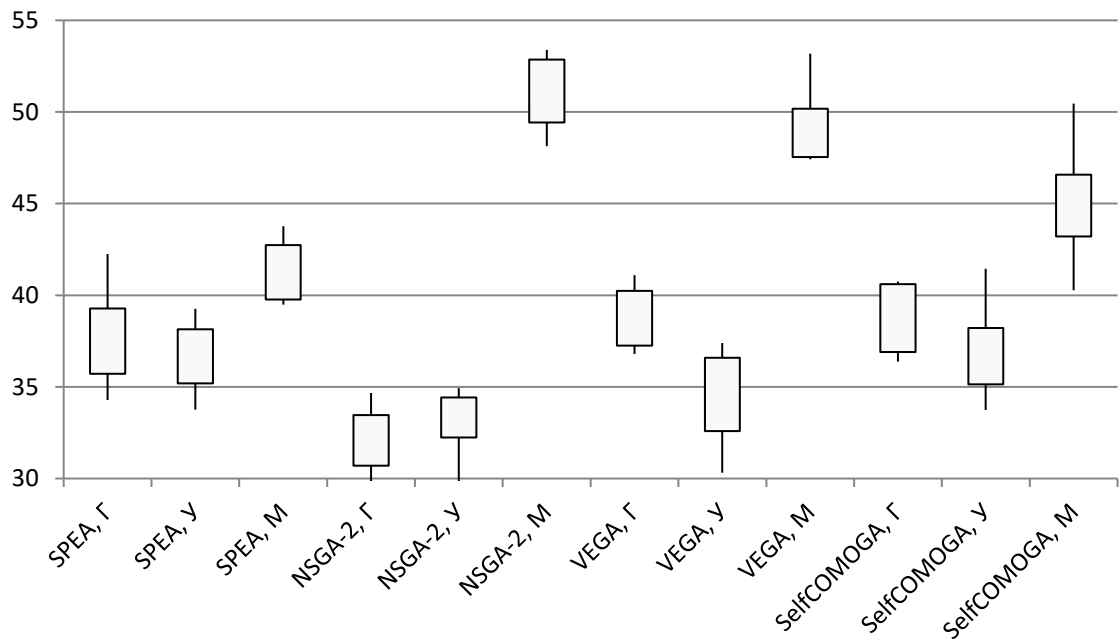


Рисунок Г12. Точность распознавания эмоций (%), входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети  $100 \times 100$  пикселей

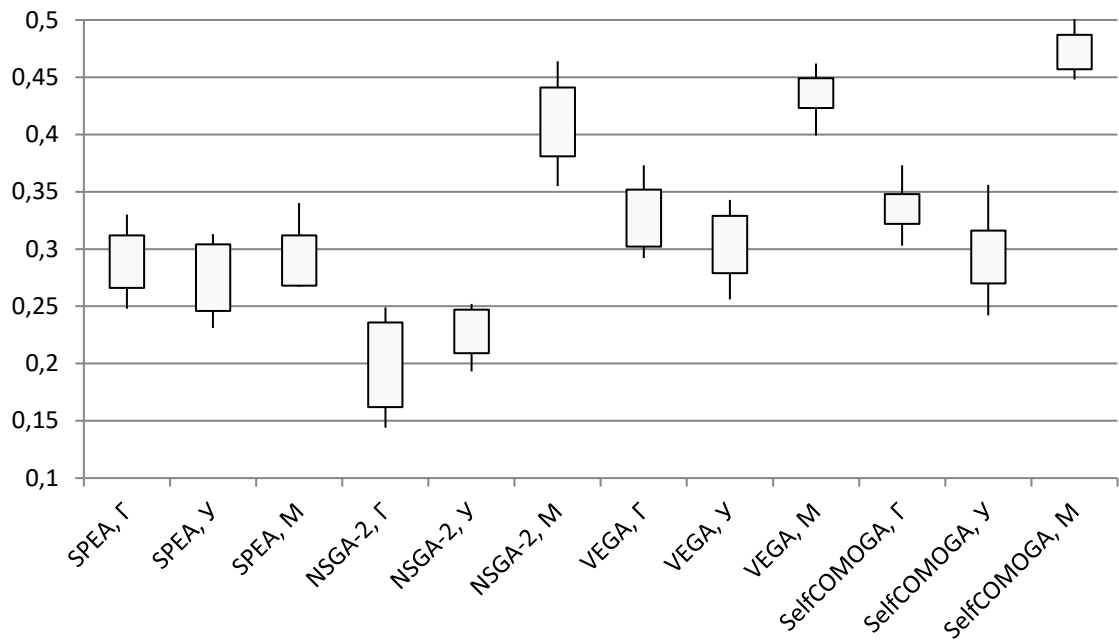


Рисунок Г13. F-мера распознавания эмоций, входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

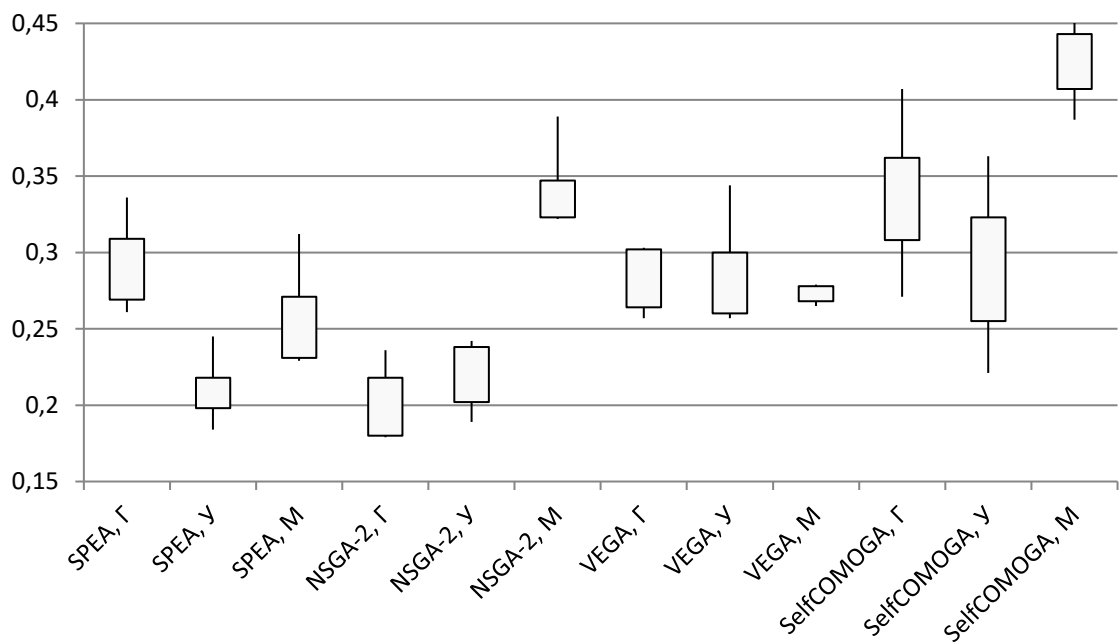


Рисунок Г14. F-мера распознавания эмоций, входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей



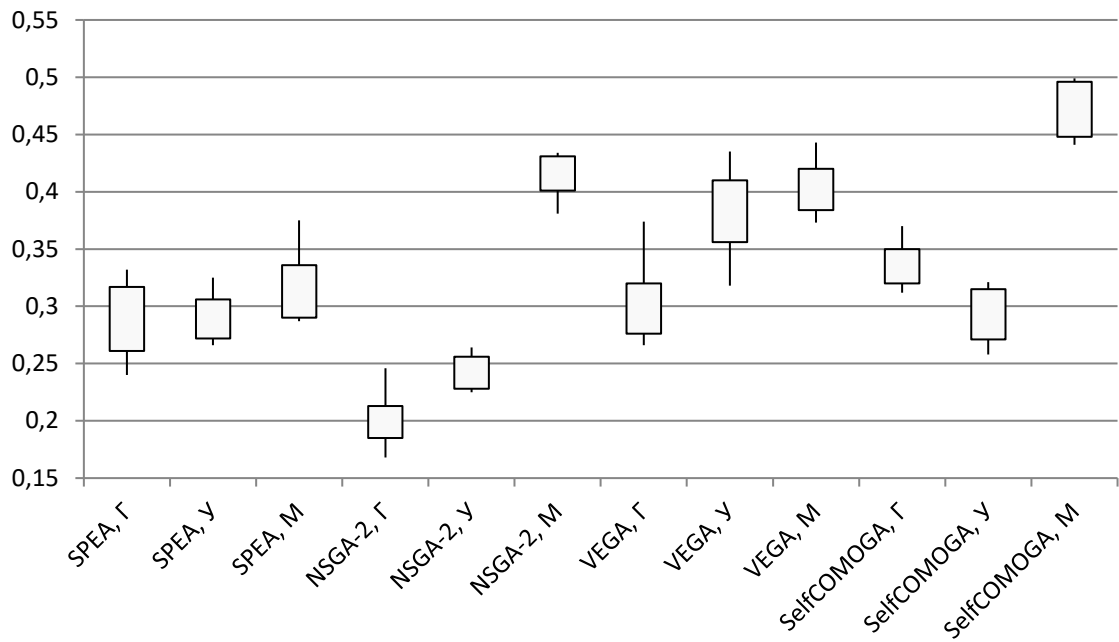


Рисунок Г15. F-мера распознавания эмоций, входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 70×70 пикселей

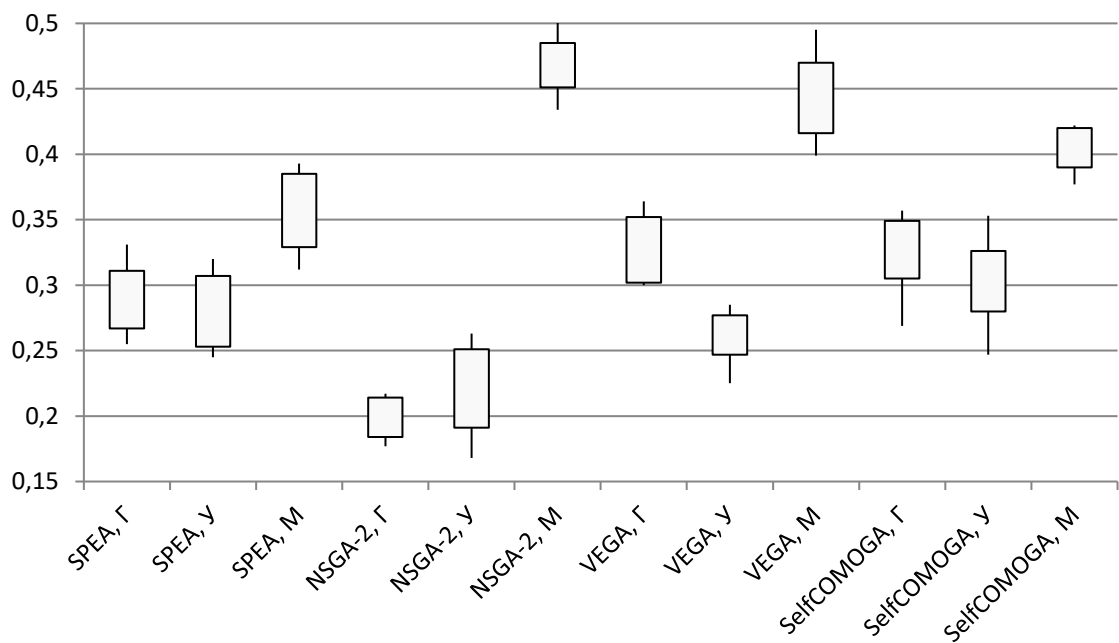


Рисунок Г16. F-мера распознавания эмоций, входные данные - аудио признаки, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 100×100 пикселей

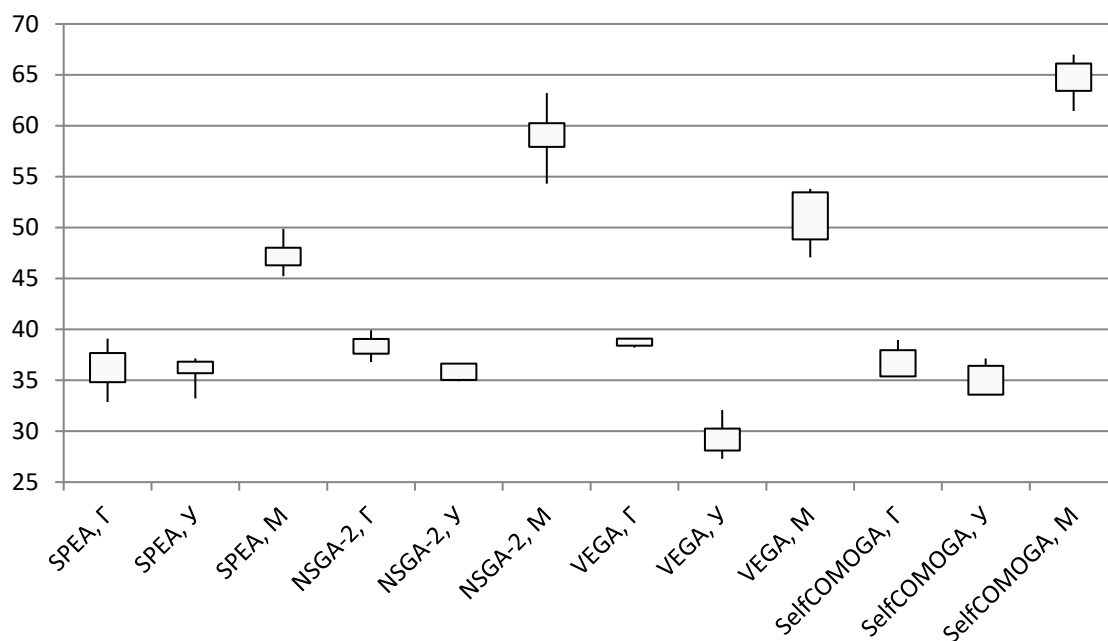


Рисунок Г17. Точность распознавания эмоций (%), входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

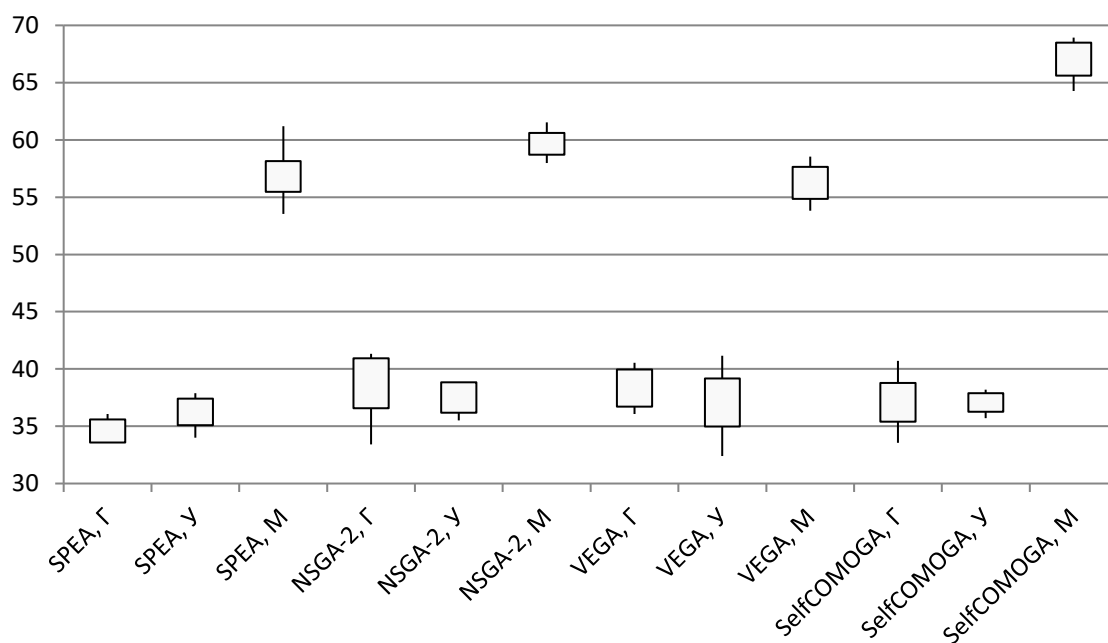


Рисунок Г18. Точность распознавания эмоций (%), входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей

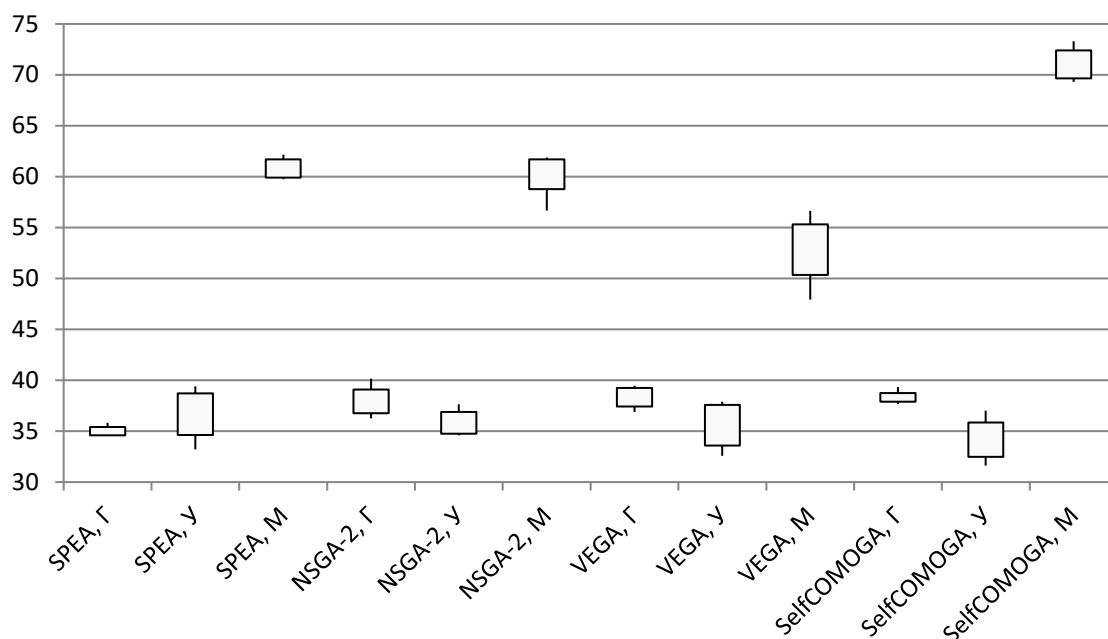


Рисунок Г19. Точность распознавания эмоций (%), входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 70×70 пикселей

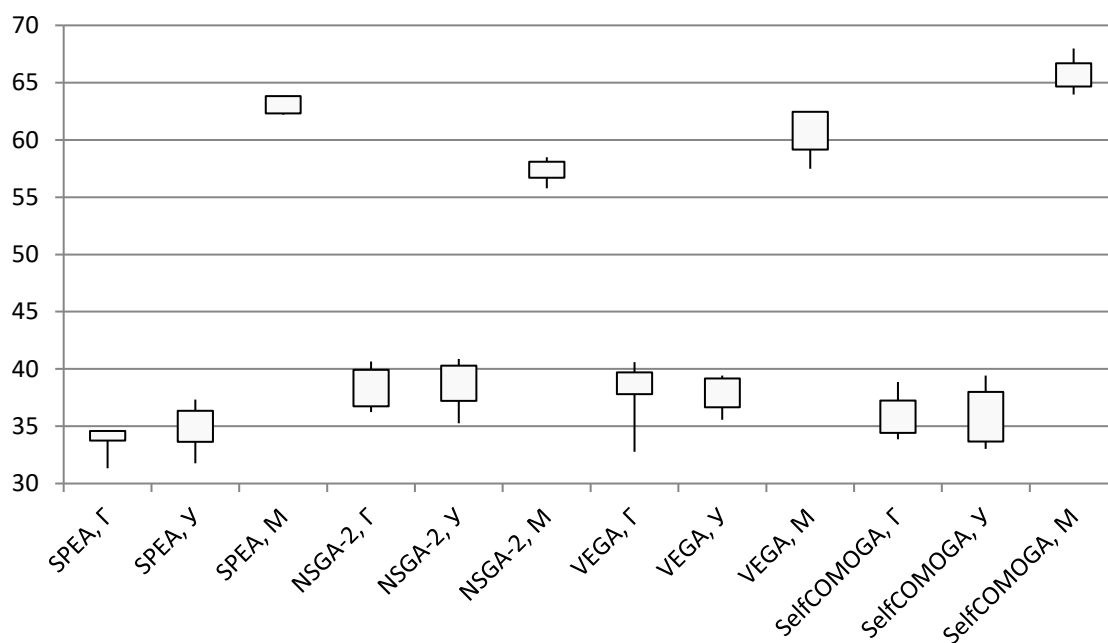


Рисунок Г20. Точность распознавания эмоций (%), входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 100×100 пикселей

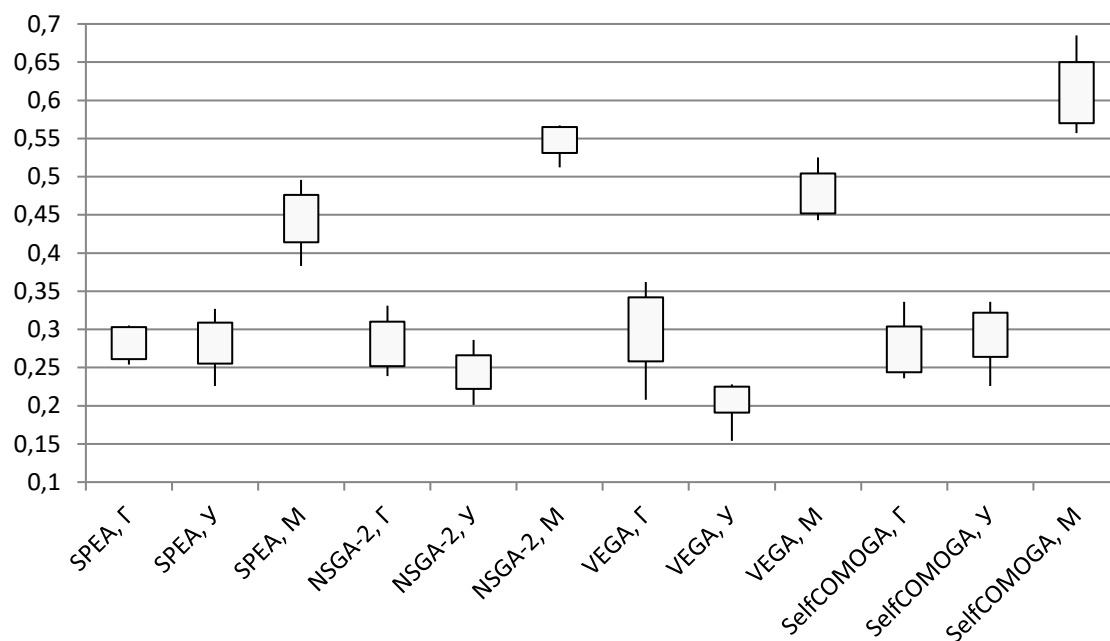


Рисунок Г21. F-мера распознавания эмоций, входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 40×40 пикселей

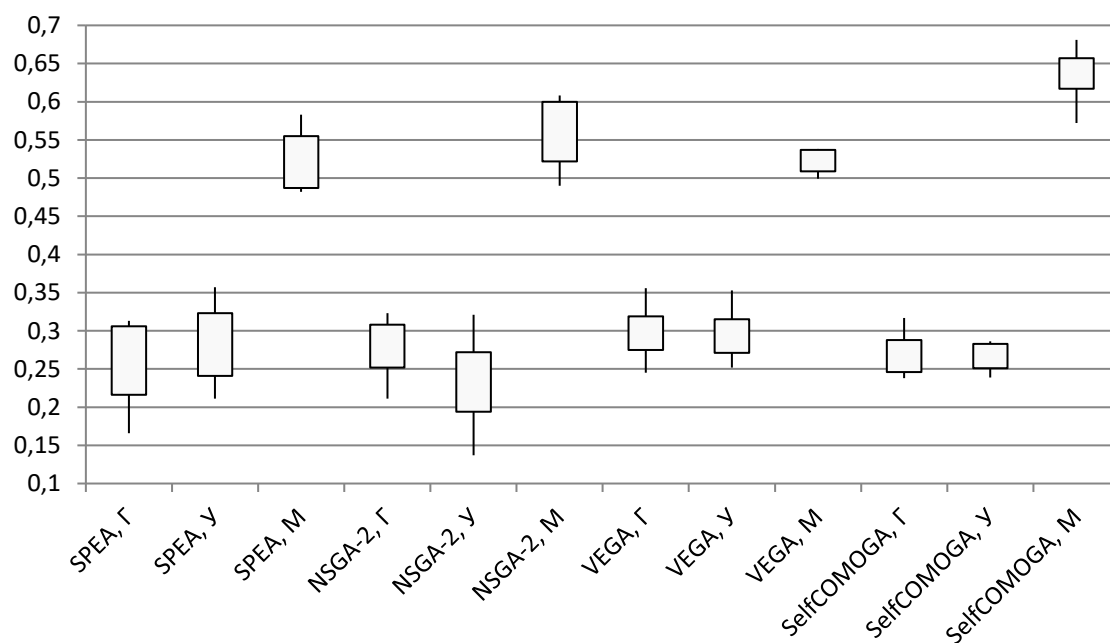


Рисунок Г22. F-мера распознавания эмоций, входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 50×50 пикселей

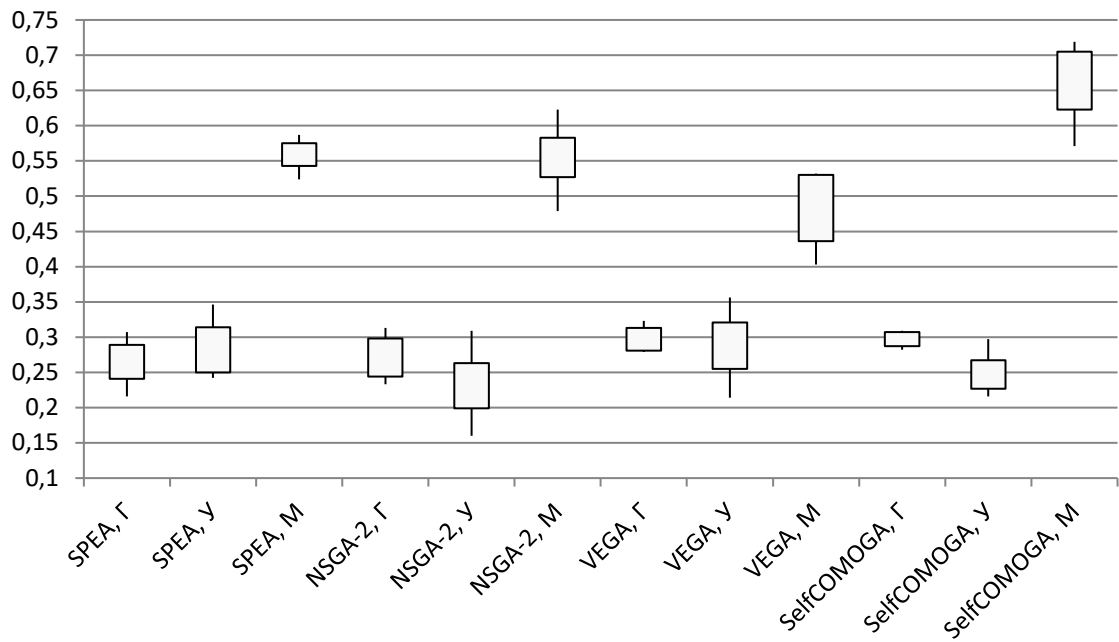


Рисунок Г23. F-мера распознавания эмоций, входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 70×70 пикселей

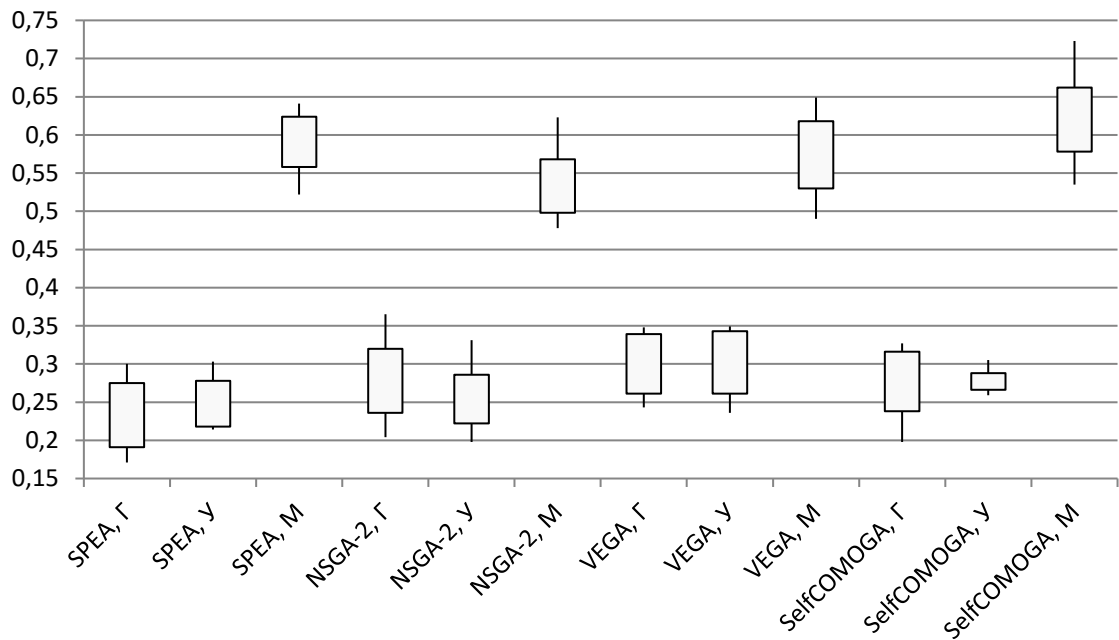


Рисунок Г24. F-мера распознавания эмоций, входные данные - аудио признаки + лицевые маркеры, обобщенный подход; размер изображений, подаваемых на вход конволюционной нейронной сети 100×100 пикселей