

На правах рукописи



Светлаков Михаил Олегович

**МЕТОД И АЛГОРИТМЫ АНАЛИЗА ДАННЫХ
ЭЛЕКТРОЭНЦЕФАЛОГРАФИИ ДЛЯ ВЕРИФИКАЦИИ СУБЪЕКТА**

Специальность 2.3.8 - Информатика и информационные процессы

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Томск – 2023

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Томский государственный университет систем управления и радиоэлектроники» (ТУСУР).

Научный руководитель –

Ходашинский Илья Александрович,
доктор технических наук, профессор

Официальные оппоненты:

Сырямкин Владимир Иванович,
доктор технических наук, профессор,
профессор кафедры управления качеством
факультета инновационных технологий,
ФГАОУ ВО «Национальный исследовательский
Томский государственный университет»
Болодурина Ирина Павловна,
доктор технических наук, профессор,
заведующий кафедрой прикладной математики
института математики и информационных
технологий, ФГБОУ ВО «Оренбургский
государственный университет»

Ведущая организация –

ФГАОУ ВО «Национальный исследовательский
Томский политехнический университет»

Защита состоится «14» декабря 2023 года в 15 час. 15 мин. на заседании диссертационного совета Д 24.2.415.06 ТУСУРа по адресу: 634050, г. Томск, пр. Ленина, 40, ауд. 201.

С диссертацией можно ознакомиться на официальном сайте <https://postgraduate.tusur.ru/urls/ul83wbmi> и в библиотеке ТУСУРа по адресу: 634045, г. Томск, ул. Красноармейская, 146.

Автореферат разослан ___ октября 2023 г.

Ученый секретарь
диссертационного совета



Костюченко Евгений Юрьевич

Общая характеристика работы

Актуальность темы. Системы автоматизированного распознавания субъектов (людей) на основе их уникальных физиологических и поведенческих характеристик более удобны для конечного пользователя, поскольку не требуют запоминания сложных паролей, а используемые физиологические и поведенческие характеристики всегда находятся при нем и, как правило, их сложнее воспроизвести или подобрать, чем пароли. Данные электроэнцефалограммы (ЭЭГ) представляют собой информацию об электрической активности мозга, и алгоритмы машинного обучения могут анализировать эти данные для выявления закономерностей и решения задач классификации. Уникальность данных ЭЭГ и их применимость для задачи верификации многократно подтверждена в работах Zhang S., Lotte F., Maiorana E., Raw-Dowling A., Armstrong B., Ruiz M., Wu Q., Rocca D., Poulos M., Zhi T., Barayeu U., Nishimoto T., Monsy J., Moctezuma L., Kaliramna B., Pham T., Yap H., Suppiah R., Mao Z., Bigdoly A., Jijomon C., Schons T., Das R., Alsumari W.

Под верификацией понимается процесс проверки информации путем сопоставления предоставленной информации с ранее подтвержденной информацией, сравнение «один к одному» - оценка, вычисление или измерение степени схожести или различия между представленным образцом и контрольным шаблоном. В современных системах верификации субъекта применяются подходы на основе машинного обучения. Анализ существующих систем верификации субъекта на основе ЭЭГ выявил, что построенные модели тренируются, в основном, только на предопределенную группу субъектов. Одной из альтернатив является подход на основе машинного обучения и метрического обучения. В ходе обучения модель обучается преобразовывать входные данные в пространство меньшей размерности, где дистанция соответствует мере схожести представленных шаблонов верификации, что позволяет более точно верифицировать субъекта и расширять список субъектов без перестроения системы. Также выявлено несовершенство используемых протоколов обучения и оценки эффективности моделей верификации на основе ЭЭГ, что потенциально приводит к более низкой реальной точности, поэтому задача создания методов и алгоритмов анализа данных электроэнцефалографии для верификации субъекта со скорректированным протоколом обучения и оценки эффективности, а также повышения точности итоговой системы, является актуальной.

Предобработка и преобразование данных является важнейшим этапом обработки сигналов ЭЭГ для повышения точности итоговой модели. Одним из перспективных методов выделения признаков ЭЭГ является спектральный анализ Холо-Гильберта (Holo-Hilbert Spectral Analysis, HHSA), который ранее не применялся для задачи верификации субъекта на основе ЭЭГ. HHSA успешно применялся для обработки данных ЭЭГ в иных задачах, в работах Chu K., Deng

J., Chanh K., Lee P., Gallego-Molina N., Tiwari U., Li Z., Tulay E., поэтому применение данного метода в задаче верификации субъекта на основе ЭЭГ позволит улучшить точность верификации.

Использование методов машинного обучения для систем верификации на основе ЭЭГ также задает актуальную проблему в плане интерпретируемости и объяснимости. Среди существующих подходов к решению проблемы выделяются нечеткие классификаторы, которые имеют в основе принципы человеческого мышления и логики (правила вида ЕСЛИ-ТО), за счет чего более понятны человеку, успешно применяются для решения широкого круга задач машинного обучения. Наиболее значимых результатов в изучении нечетких систем достигли Zadeh L., Sugeno M., Takagi T., Bezdek J., Mamdani E., Herrera F., Angelov P., Yager R., Filev D., Аверкин А., Катасёв А., Леденева Т., Синюк В., Ярушкина Н., Яхьяева Г. В случае сигнала ЭЭГ пространство признаков имеет особенно высокую размерность за счет использования большого количества датчиков (каналов) снятия данных, в отдельных случаях их может быть более 100, число признаков кратно возрастает. В случае нечетких классификаторов для отбора признаков зачастую применяются бинаризованные метаэвристические алгоритмы, которые могут учитывать взаимодействие между признаками и их влияние на целевую переменную. Использование таких алгоритмов приводит к повышению точности и интерпретируемости нечетких классификаторов, за счет снижения избыточности данных.

Цель диссертационной работы - повышение точности и интерпретируемости моделей машинного обучения, предназначенных для верификации субъекта на основе данных ЭЭГ.

Для достижения цели необходимо выполнить следующие задачи:

- 1) провести обзор современных методов верификации субъекта на основе данных ЭЭГ;
- 2) разработать и исследовать метод построения моделей верификации субъекта на основе данных ЭЭГ с помощью нейронных сетей и спектрального анализа Холо-Гильберта;
- 3) разработать и исследовать алгоритм формирования структуры нечеткого классификатора с помощью кластеризации и метаэвристических алгоритмов, а также оптимизации параметров классификатора;
- 4) разработать и исследовать алгоритм отбора признаков с использованием метаэвристических алгоритмов для нечеткого классификатора;
- 5) провести оценку эффективности разработанных метода и алгоритмов на наборах данных ЭЭГ, выполнить сравнение с аналогами.

Объект исследования - процесс построения моделей машинного обучения, предназначенных для верификации субъекта на основе данных ЭЭГ.

Предмет исследования – методы и алгоритмы построения моделей машинного обучения для верификации субъекта на основе данных ЭЭГ.

Методы исследования. В работе применялись методы частотно-временного анализа, классического машинного обучения, глубокого обучения, теория нечетких множеств и нечеткой логики, методы оптимизации.

Достоверность результатов обеспечивается корректностью применения математических методов, результатами проведенных экспериментов и сравнений с базовыми моделями и алгоритмами, проведенных с помощью статистических критериев.

Научная новизна полученных результатов. В диссертации получены следующие новые научные результаты.

1) Предложен оригинальный метод верификации субъекта на основе данных электроэнцефалограммы и методов глубокого обучения, отличительной особенностью которого является извлечение признаков с помощью спектрального анализа Холо-Гильберта и метрического обучения.

2) Разработан гибридный алгоритм построения базы правил нечеткого классификатора, особенностью которого является совместное использование алгоритма кластеризации k -средних и метаэвристического алгоритма “стая птиц”.

3) Разработан алгоритм отбора признаков, особенностью которого является использование V-образной функции трансформации и метаэвристического алгоритма “стая птиц”.

Теоретическая значимость работы заключается в развитии технологии методов извлечения признаков из сигнала ЭЭГ, построения нечетких систем, интеллектуального анализа данных. Предложенный подход на основе данных ЭЭГ и методов глубокого обучения, отличительной особенностью которого является извлечение признаков с помощью спектрального анализа Холо-Гильберта и использование метрического обучения, также может быть использован для решения иных задач классификации данных ЭЭГ. Предложенные оригинальные алгоритмы отбора признаков и построения базы правил нечеткого классификатора также могут быть использованы для построения нечетких классификаторов в иных задачах.

Практическая значимость работы подтверждается применением предложенных метода и алгоритмов для решения практической задачи классификации – обнаружения эпилептиформ на сегментах ЭЭГ у пациентов Томского НИИ курортологии и физиотерапии ФМБА России, получен акт внедрения. Результаты диссертационной работы также внедрены в учебный процесс ТУСУР. Разработанные метод и алгоритмы использованы при выполнении следующих проектов:

– научный проект при поддержке РФФИ проект № 22–21–00021 «Интерпретируемый нечеткий классификатор рукописных данных для диагностики нейродегенеративных заболеваний» 2022–2023 гг.;

– научный проект при поддержке РФФИ № 16–07–00034 «Методы и инструментальные средства построения самообучающихся систем, основанных на нечетких правилах» 2016–2018 гг.

На защиту выносятся следующие положения:

1) Разработанный оригинальный метод верификации субъекта на основе данных ЭЭГ на общедоступном наборе данных PhysioNet EEG Motor Movement/Imagery Dataset (PEEGMMID) позволяет понизить значение метрики ошибки Equal Error Rate (EER) на 0,97% (с 10,92% до 9,95%; значимые различия по критерию Уилкоксона, $p < 0,01$) по сравнению с использованием признаков на основе преобразования Гильберта-Хуанга, а также снизить значение метрики ошибки EER на 9,38% (значимые различия по критерию Уилкоксона, $p < 0,01$) по сравнению с использованием признаков на основе дискретного вейвлет-преобразования.

Соответствует пункту 4 паспорта специальности: Разработка методов и технологий цифровой обработки аудиовизуальной информации с целью обнаружения закономерностей в данных, включая обработку текстовых и иных изображений, видео контента. Разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения требуемой информации из текстов.

2) Разработанный гибридный алгоритм построения базы правил нечеткого классификатора позволяет повысить точность нечеткого классификатора на рассмотренных наборах данных из различных предметных областей в среднем на 14,38% по сравнению с алгоритмом экстремумов признаков классов¹ и на 6,00% по сравнению с алгоритмом на основе кластеризации k -средних. Для проблемы верификации субъекта по ЭЭГ и набора данных PEEGMMID точность классификатора, построенного гибридным алгоритмом, выше точности классификатора на основе машины опорных векторов с линейным ядром на 6,46%, и на 9,37% выше точности нечеткого классификатора, построенного с помощью алгоритма экстремумов признаков классов.

Соответствует пункту 13 паспорта специальности: Разработка и применение методов распознавания образов, кластерного анализа, нейро-сетевых и нечетких технологий, решающих правил, мягких вычислений при анализе разнородной информации в базах данных.

3) Предложенный алгоритм отбора признаков, извлеченных с помощью дискретного вейвлет-преобразования признаков и нечеткого классификатора, позволил сократить число признаков с 1152 до 98 для набора данных ЭЭГ PEEGMMID.

¹ Мех М.А., Ходашинский И.А. Сравнительный анализ применения методов дифференциальной эволюции для оптимизации параметров нечетких классификаторов // Известия Российской академии наук. Теория и системы управления. 2017. № 4. С. 65-75.

Соответствует пункту 7 паспорта специальности. Разработка методов обработки, группировки и аннотирования информации, в том числе, извлеченной из сети интернет, для систем поддержки принятия решений, интеллектуального поиска, анализа.

Апробация работы. Результаты исследовательской работы использованы для решения практической задачи классификации – обнаружения эпилептиформ на сегментах ЭЭГ пациентов Томского НИИ курортологии и физиотерапии ФМБА России, получен акт внедрения. Предложенные алгоритмы также были использованы при выполнении научного проекта при поддержке РНФ проект № 22-21-00021 «Интерпретируемый нечеткий классификатор рукописных данных для диагностики нейродегенеративных заболеваний» 2022-2023 гг. и научного проекта при поддержке РФФИ № 16-07-00034 «Методы и инструментальные средства построения самообучающихся систем, основанных на нечетких правилах» 2016–2018 гг. По результатам работы проведены выступления на конференциях различного уровня, среди них:

– The Thirteenth International Conference on Swarm Intelligence (ICSI'2022), онлайн, 15-19 July 2022, Xi'an, China.

– 2021 IEEE Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, онлайн, 13-14 мая 2021;

– Международная научно-техническая конференция студентов, аспирантов и молодых учёных Научная сессия ТУСУР – 2020, Томск, 19–21 мая 2020 г., секция 3.4 «Вычислительный интеллект»;

– Международная научно-техническая конференция студентов, аспирантов и молодых учёных Научная сессия ТУСУР – 2021, Томск, 19–21 мая 2021 г., секция 3.4 «Вычислительный интеллект»;

– XVIII Международная конференция студентов, аспирантов и молодых ученых Перспективы развития фундаментальных наук, Томск, 27-30 апреля 2021;

– Международная научно-техническая конференция студентов, аспирантов и молодых учёных Научная сессия ТУСУР – 2022, Томск, 18–20 мая 2022 г., секция 3.4 «Вычислительный интеллект»;

– XXII Всероссийский ежегодный конкурс-конференция студентов и аспирантов по информационной безопасности SIBINFO-2022, Томск, 27 мая 2022 г.;

– XVIII Международная научно-практическая конференция Электронные средства и системы управления, посвящённая 60-летию ТУСУРа, Томск, 16-18 ноября 2022 г., секция 20 «Вычислительный интеллект и машинное обучение»;

– Международная научно-техническая конференция студентов, аспирантов и молодых учёных Научная сессия ТУСУР – 2022, Томск, 18–20 мая 2023 г., секция 3.4 «Вычислительный интеллект».

Публикации по теме диссертации. По теме исследования опубликовано 25 печатных работ, из которых в рекомендованных ВАК периодических изданиях – 5. В международной базе SCOPUS проиндексированы 15 работ, 12 работ – в международной базе Web of Science. Получены 3 свидетельства о государственной регистрации программ для ЭВМ.

Личный вклад автора. Диссертантом лично разработаны алгоритм извлечения признаков для сигнала ЭЭГ, алгоритм формирования структуры нечеткого классификатора, алгоритм настройки параметров термов на основе метаэвристики “стая птиц”, алгоритм отбора признаков термов на основе метаэвристики “ стая птиц”, разработан метод верификации субъекта на основе данных ЭЭГ, отличительной особенностью которого является извлечение признаков с помощью спектрального анализа Холо-Гильберта и использование метрического обучения; лично получены результаты экспериментов, проведена апробация разработанных алгоритмов. Определение целей и задач исследования, анализ результатов экспериментов, подготовка публикаций по полученным результатам проводилась совместно с научным руководителем.

Основное содержание

Во введении описана актуальность работы, определены цель, задачи, объект, предмет, методы исследования, перечислены основные полученные результаты, их научная новизна, теоретическая и практическая значимость, приведены защищаемые положения.

В первой главе приведен обзор сигнала ЭЭГ, его использования в задачах машинного обучения, применяемые при этом методы обработки сигнала, рассмотрены применяемые и планируемые к применению для решения задачи верификации субъекта на основе данных ЭЭГ модели и методы, поставлена задача верификации.

Во второй главе приведено описание разработанных алгоритмов и метода для решения задачи верификации на основе данных ЭЭГ.

Алгоритм построения базы правил нечеткого классификатора с использованием алгоритма кластеризации k -средних и метаэвристического алгоритма “стая птиц”. В данной работе предлагается алгоритм генерации базы правил с использованием метаэвристической кластеризации, на основе метаэвристического алгоритма “стая птиц”. В этом случае генерировалась случайным образом популяция векторов $CS = [CS_1, \dots, CS_i, \dots, CS_{size}]$ ($size$ – размер популяции, $|CS_i| = R * l * pl$, pl – количество параметров выбранной функции

принадлежности, R – число правил, l – число образцов в наборе данных), каждый из которых представляет собой совокупность центроид-правил. Непараметрический алгоритм Autonomous Data Partitioning (ADP) позволяет определить число кластеров-правил для классификаторов. Далее приведен псевдокод алгоритма.

Алгоритм Генерация нечеткой системы с помощью метаэвристической кластеризации:

Вход: таблица наблюдений (\mathbf{X}, \mathbf{Y}) , $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_l]$, $\mathbf{Y} = [y_1, \dots, y_i, \dots, y_l]$, $size$ - размер популяции \mathbf{CS} , R – число правил, l – число образцов в наборе данных, pl – количество параметров выбранной функции принадлежности, **params** - вектор параметров метаэвристического алгоритма, $type$ - тип функции принадлежности, $metaheuristic$ – функция, выполняющая оптимизацию метаэвристическим алгоритмом BSA, $ffclust$ – фитнес-функция.

Выход: \mathbf{CS}_{best} - сгенерированная база правил.

1. $\mathbf{CS} \leftarrow generate(type, size, pl, R, l)$ - генерация популяции баз правил;
 2. $\mathbf{CS} \leftarrow metaheuristic(params, \mathbf{CS}, \mathbf{X}, \mathbf{Y}, ffclust)$
 3. Выбор базы правил \mathbf{CS}_{best} по вычисленному лучшему значению $ffclust(\mathbf{X}, \mathbf{Y}, R, l, \mathbf{CS}_j)$
-

Алгоритм Фитнесс-функция $ffclust$:

Вход: таблица наблюдений (\mathbf{X}, \mathbf{Y}) , $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_l]$, $\mathbf{Y} = [y_1, \dots, y_i, \dots, y_l]$, R , l , $\mathbf{CS}_j = [\mathbf{c}_{j1}, \dots, \mathbf{c}_{ji}, \dots, \mathbf{c}_{jR}]$.

Выход: $ffvalue$.

1. Вычисление матрицы евклидовых расстояний $\mathbf{D}_{R \times l} = (d_{km})$ между центроидами-правилами \mathbf{c}_{ji} и векторами данных \mathbf{x}_i из таблицы \mathbf{X} ;
 2. $Clusters \leftarrow \{Clusters_1, \dots, Clusters_i, \dots, Clusters_R\}$, $Clusters_i \leftarrow \emptyset$
 3. **пока** $m=1 \dots l$
 4. $nearest \leftarrow \arg \min_i d_{km}$
 $k \in 1..R, m \in 1..l$
 5. $Clusters_{nearest} \leftarrow Clusters_{nearest} \cup \mathbf{x}_m$
 6. **конец пока**
 8. $FS \leftarrow \emptyset$
 9. **пока** $m=1 \dots R$
 10. **если** $|Cluster_m| > 0$
 11. Сгенерировать правило $Rule$ на основе мин. и макс. значений признаков в $Cluster_m$
 12. **иначе**
 13. Сгенерировать правило $Rule$ на основе мин. и макс. значений признаков в \mathbf{X}
 - 14: $FS \leftarrow FS \cup Rule$
 - 15: **конец пока**
 - 16: Вычисление точности классификации для набора правил \mathbf{FS} - $ffvalue \leftarrow acc(\mathbf{FS}, \mathbf{X}, \mathbf{Y})$
-

Алгоритм отбора признаков с использованием функции трансформации и метаэвристического алгоритма “стая птиц”. Классический подход к бинаризации метаэвристик заключается в использовании функций трансформации, которые позволяют отображать непрерывные значения на бинарную область. В работе предлагается использовать алгоритм отбора признаков с использованием V-образной функции трансформации и метаэвристического алгоритма “стая птиц”. В ходе работы алгоритма $\theta_{i,j}^t$ - позиция каждой поисковой частицы метаэвристического популяционного алгоритма - используется как бинарная переменная, от которой зависит использование или игнорирование i -го признака на итерации t . Для использования бинарного представления метаэвристики было рассмотрено использование S-образной функции трансформации (Sigm):

$$\text{Если } 0.5 < \frac{1}{1 + e^{-\theta_{i,j}^t}}, \text{ то } \theta_{i,j}^t = 1, \text{ иначе } 0. \quad (1)$$

Помимо S-образной функции, в работе рассмотрена V-образная функция (V):

$$\text{Если } 0.5 < \left| \frac{\theta_{i,j}^t}{\sqrt{1 + \theta_{i,j}^t{}^2}} \right|, \text{ то } \theta_{i,j}^t = 1, \text{ иначе } 0. \quad (2)$$

Также, рассмотрено использование Tanh-образной функции (Tanh):

$$\text{Если } |\tanh(\theta_{i,j}^t)| > 0.5, \text{ то } \theta_{i,j}^t = 1, \text{ иначе } 0. \quad (3)$$

В качестве альтернативы рассмотрено еще одно правило (Naive):

$$\text{Если } \theta_{i,j}^t > 0.5 \text{ то } \theta_{i,j}^t = 1, \text{ иначе } 0. \quad (4)$$

Далее приведен псевдокод алгоритма.

Алгоритм Бинарный алгоритм “стая птиц”

Вход: T – обучающие данные, N – размер популяции, $MaxGen$ – количество итераций, trf - функция трансформации, FQ – параметр алгоритма, P – параметр алгоритма

Выход: бинарный вектор s_{best} , указывающий на информативные признаки.

1. Инициализировать популяцию $S \leftarrow \{s_1, s_2, \dots, s_N\}$, $s_i \leftarrow \text{rand}\{0, 1\}^n$
 2. Инициализировать счётчик популяции $i \leftarrow 1$
 3. **пока** $i \leq N$
 4. Создать обучающие данные T' из T с признаками s_i
 5. Построить классификатор на T' и вычислить значение целевой функции f_i на основе s_i и функции трансформации trf
 6. $i \leftarrow i + 1$
 7. **конец пока**
-

-
8. Определить лучшее решение \mathbf{s}_{best} в S ($\forall i f_{\text{best}} \leq f_i$)
 9. Инициализировать счётчик итераций $t \leftarrow 1$
 10. **пока** $t \leq \text{MaxGen}$
 11. Инициализировать популяцию вещественных решений $\Theta^t \leftarrow \{\theta_1^t, \dots, \theta_N^t\}$ на основе S
 12. Инициализировать счётчик популяции $i \leftarrow 1$
 13. **если** $(t \bmod FQ \neq 0)$:
 14. **пока** $i \leq N$:
 15. **если** $(\text{rand}(0, 1) < P)$:
 16. Обновить решение θ_i^t по формуле стратегии 1
 17. **иначе:**
 18. Обновить решение θ_i^t по формуле стратегии 2
 19. Скорректировать θ_i^t если вектор выходит за пределы пространства поиска
 20. Обновить бинарное решение \mathbf{s}_i через функцию трансформации $\mathbf{s}_i \leftarrow \text{trf}(\theta_i^t)$
 21. Создать обучающие данные T^{\setminus} из T с признаками \mathbf{s}_i
 22. Построить классификатор на T^{\setminus} , вычислить значение целевой функции f_i на основе \mathbf{s}_i
 23. $i \leftarrow i + 1$
 24. **конец пока**
 25. **иначе:**
 26. Разделить частицы на 1-й и 2-й тип
 27. **пока** $i \leq N$
 28. **если** частица 1-го типа
 29. Обновить решение θ_i^t по формуле стратегии 3
 30. **иначе:**
 31. Обновить решение θ_i^t по формуле стратегии 4
 32. Скорректировать θ_i^t если вектор выходит за пределы пространства поиска;
 33. Обновить бинарное решение \mathbf{s}_i через функцию трансформации $\mathbf{s}_i \leftarrow \text{trf}(\theta_i^t)$
 34. Создать обучающие данные T^{\setminus} из T с признаками \mathbf{s}_i
 35. Построить классификатор на T^{\setminus} , вычислить значение целевой функции f_i на основе \mathbf{s}_i
 36. $i \leftarrow i + 1$
 37. **конец пока**
 38. $t \leftarrow t + 1$
 39. Обновить наилучшее решение \mathbf{s}_{best} в S ($\forall i f_{\text{best}} \leq f_i$)
 40. **конец пока**
-

Метод и архитектура модели верификации субъекта на основе данных электроэнцефалограммы с использованием извлечения признаков с помощью преобразования Холо-Гильберта и метрического обучения. Метод спектрального анализа Холо-Гильберта (Holo-Hilbert Spectral Analysis, HNSA) использует вложенное эмпирическое модовое разложение (EMD) и преобразование Гильберта-Хуанга (Hilbert-Huang Transform, ННТ). Размерность полученного этим методом представления выглядит следующим образом – $[N$ диапазонов частотной модуляции (FM) $\times M$ диапазонов амплитудной модуляции (AM) \times время]. Размерность обычного представления на основе преобразования Гильберта-Хуанга выглядит следующим образом – $[N$ диапазонов FM \times время]. В работе опробованы две архитектуры на основе преобразования Холо-Гильберта: сеть на основе трехмерных сверток и на основе одномерных сверток. В случае одномерных сверток рассматривается такое представление спектрограммы, когда каждый частотный интервал считается отдельной временной серией. Если спектральное представление эпохи ЭЭГ имеет представление размерности [число_каналов_ЭЭГ $\times N$ диапазонов FM \times время], то, в случае модели на основе одномерных сверток, данные преобразовываются к размерности [(число_каналов_ЭЭГ * N диапазонов FM) \times время]. Схема предлагаемого метода приведена на Рисунке 1, а итоговая архитектура сети – в Таблице 1.

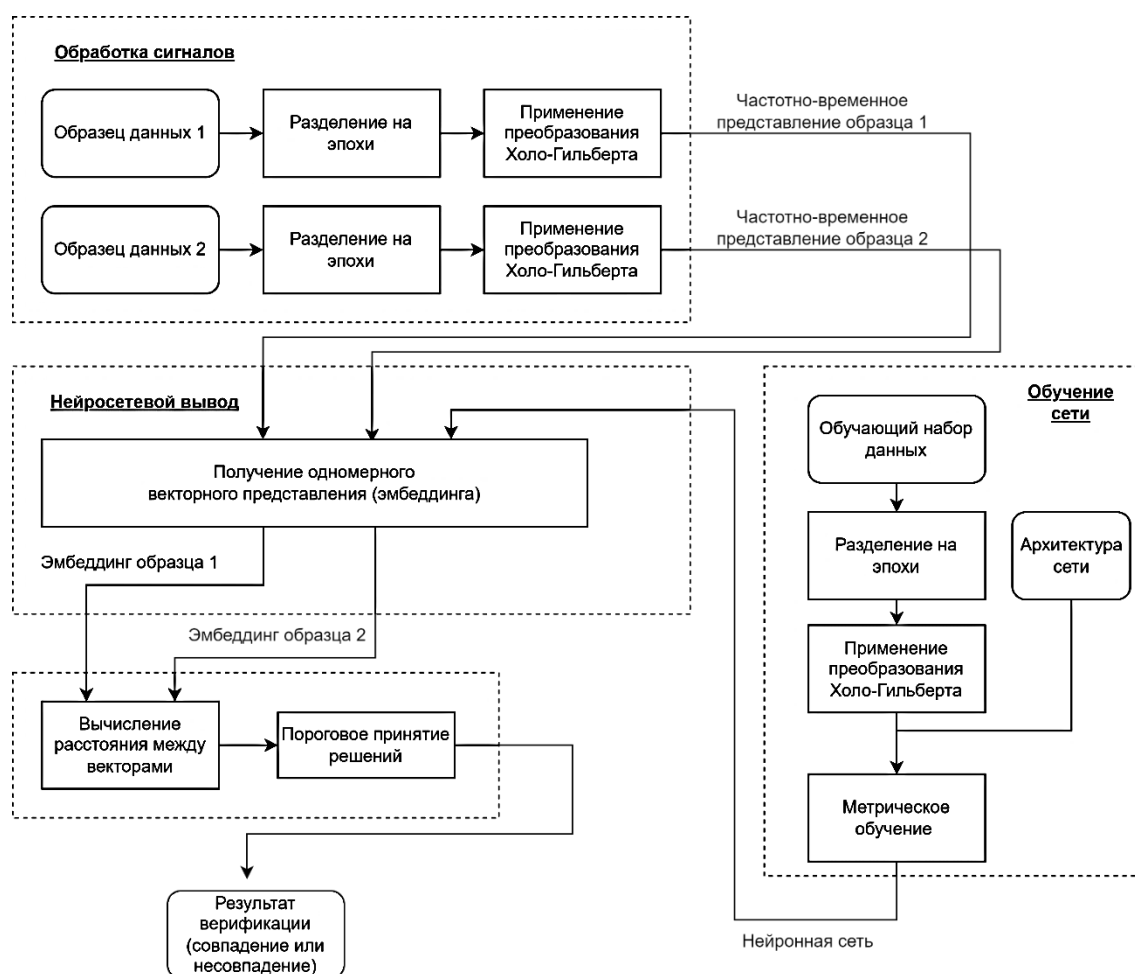


Рисунок 1 – Схема предлагаемого метода

Таблица 1 - Архитектура сети на основе преобразования Холо-Гильберта и одномерных сверток 1DCNN-Holo

Слой	Число фильтров / нейронов	Размер фильтров / ядер	Размерность входа	Число параметров	Функция активации выхода
Conv1D (padding=1)	1024	2	[64*15*2, 115]	7865344	-
Maxpool1D	-	2	[1024, 116]	-	-
BatchNorm1D	-	-	[1024, 59]	2048	SiLU
Conv1D	512	2	[1024, 59]	1049088	-
Maxpool1D	-	2	[512, 58]	-	-
BatchNorm1D	-	-	[512, 29]	1024	SiLU
Conv1D	256	2	[1024, 29]	262400	-
Maxpool1D	-	2	[256, 28]	-	-
BatchNorm1D	-	-	[256, 14]	512	SiLU
Conv1D	128	2	[256, 14]	65664	-
Maxpool1D	-	2	[128, 13]	-	-
BatchNorm1D	-	-	[128, 6]	256	SiLU
Conv1D	64	2	[128, 6]	16448	-
Maxpool1D	-	2	[64, 5]	-	-
BatchNorm1D	-	-	[64, 2]	128	SiLU
Flatten	-	-	[128]	-	-
Linear	128	-	[128]	16512	-
Output	-	-	[128]	-	-

В данной архитектуре и остальных архитектурах в работе в качестве функции активации используется функция SiLU (Sigmoid Linear Unit). Linear – полносвязный слой, Flatten – слой преобразования размерности в одномерный вектор, Output – выходной слой сети, не изменяющий и не преобразующий поступающие данные. В первом слое используется заполнение (padding, дополнение нулями), чтобы сохранить размерность входа (кроме размеров каналов) для корректной обработки значений пограничных значений. Также, для учета влияния краевых эффектов, искажающих спектрограммы, были исключены около 5% данных результатов преобразования Гильберта-Хуанга и Холо-Гильберта по оси времени, в начале и в конце каждого сегмента. Для метрического обучения использовалась функция потерь Multi-Similarity Loss. Число итераций обучения для архитектуры в Таблице 1 равно 20, размер пакета образцов (батча) равен 512.

Третья глава посвящена проверке эффективности предложенных метода и алгоритмов на данных KEEL и на данных ЭЭГ PEEGMMID. Набор данных PEEGMMID после обработки имеет следующую размерность – [8715 образцов, 64 канала, 126 временных отсчетов]. Для применения метода Холо-Гильберта и Гильберта-Хуанга, согласно теореме Найквиста-Котельникова, был выбран диапазон частотной модуляции от 0 Гц до 60 Гц. Для итогового процессинга данных был выделен диапазон частотной модуляции 0–60 Гц с 15 частотными субдиапазонами по 4 Гц. Для преобразования Холо-Гильберта при исследовании наличия амплитудной модуляции, обнаружено, что большая часть спектральной мощности сосредоточена в области низких частот (0–8 Гц). Для итогового процессинга данных в случае преобразования Холо-Гильберта был выделен диапазон амплитудной модуляции 0–8 Гц с 2 частотными субдиапазонами по 4 Гц. Для учета влияния краевых эффектов, искажающих спектрограммы, были исключены около 5% данных спектрограммы в начале и в конце рассматриваемого временного сегмента ЭЭГ. После этого набор данных имеет следующую размерность в случае преобразования Гильберта-Хуанга – [8715 образцов, 64 канала, 15 FM диапазонов, 115 временных отсчетов]. Для преобразования Холо-Гильберта размерность также скорректирована – [8715 образцов, 64 канала, 15 FM диапазонов, 2 AM диапазона, 115 временных отсчетов].

В данной работе для оценки нейросетевых методов верификации используется метрика EER и доля (процент) правильной классификации. Для нечетких систем используется только доля (процент) правильной классификации. Для оценки эффективности на данных ЭЭГ в режиме верификации генерировались пары “положительных” векторов субъектов (идентификаторы субъекта для обеих векторов пары совпадают) и пары “негативных” (идентификаторы субъекта для обеих векторов пары не совпадают). Всего генерировалось по 5000 “положительных” векторов и по 5000 “негативных” векторов. Процесс генерации векторов различается в зависимости от выбранной схемы кросс-валидации. В работе используется два протокола

валидации: обычной кросс-валидации (Kfold) и с исключением групп (LeavePGroupsOut). В данной работе для данных ЭЭГ используется протокол LeavePGroupsOut с генерацией 20 разбиений и 20-кратный Kfold, для данных репозитория Keel – 10-кратный Kfold. В первом случае пары “положительных” векторов генерируются только для субъектов в тестовой выборке, в парах сгенерированных “негативных” всегда присутствует субъект из тестовой выборки, а второй субъект может быть как из тестовой, так и обучающей выборки.

Исследование эффективности построения базы правил нечеткого классификатора с использованием алгоритма кластеризации k -средних и метаэвристического алгоритма “стая птиц”. Для применения указанного алгоритма сначала необходимо задать необходимое число кластеров (правил нечеткой системы). Для этого на первом этапе предлагается использовать непараметрический алгоритм кластеризации ADP. На основе полученных алгоритмом ADP или иным алгоритмом кластеризации k кластеров формируются правила нечеткой системы ($R=k$), полученная совокупность правил оценивается критерием. Использование алгоритма k -средних в предложенном алгоритме заключается в использовании идеи передвижения центроид-правил и оптимизации их положения в многомерном пространстве поиска, а после этого кластеризации данных по полученным центроидам на каждой итерации для вычисления критерия оптимизации.

Алгоритмы-аналоги k -средних ($k=ADP_k$) и алгоритм построения по экстремумам признаков классов (АЭПК) были сравнены с предложенным оригинальным гибридным алгоритмом генерации базы правил с использованием кластеризации на основе метаэвристического алгоритма “стая птиц”. Предложенный гибридный алгоритм обозначен как BSA-GEN-ADPK.

По результатам проверки на наличие значимых различий между результатами с помощью критерия Уилкоксона алгоритмов генерации (Таблица 2) было выявлено, что BSA-GEN-ADPK превосходит оба алгоритма-аналога (превосходит АЭПК в среднем на 14,38% и k -средних ($k=ADP_k$) на 6,00%).

Таблица 2 – Проверка на наличие значимых различий между результатами алгоритмов генерации с помощью критерия Уилкоксона

Нулевая гипотеза (НГ)	Между результатами сопоставляемых алгоритмов имеются лишь случайные различия		
Сопоставляемые алгоритмы	BSA-GEN-ADPK и АЭПК	BSA-GEN-ADPK и k -средних ($k=ADP_k$)	k -средних ($k=ADP_k$) и АЭПК
Значение критерия	170,00	20,00	63,87
Асимптотическая значимость	<0,01	<0,01	<0,01
Результат	НГ отклоняется	НГ отклоняется	НГ отклоняется

Исследование эффективности отбора признаков с использованием функции трансформации и метаэвристического алгоритма “стая птиц”. Для данных KEEL результаты эксперимента показали, что по числу признаков результаты статистически различимы, по проценту правильной классификации результаты неразличимы. Наилучший результат показали статистически неразличимые между собой Tanh- и V-образная функции трансформации. V-образная функции трансформации выбрана для дальнейшего использования.

Предложенные алгоритмы построения нечеткого классификатора впоследствии применены для набора данных ЭЭГ PEEGMMID. Перечень используемых комбинаций алгоритмов для построения нечеткого классификатора приведен в Таблице 3. После извлечения признаков получаем следующую размерность набора данных – [8715 образцов, 1152 признаков]. Использовался протокол валидации LeavePGroupsOut ($P=20$, 20 итераций валидации).

Согласно Таблице 4 лучшим по точности классификации является предложенный алгоритм на основе метаэвристической кластеризации (BSA-GEN). Он не только превосходит классификатор на основе машины опорных векторов с линейным ядром (LinearSVC) по точности, но и более интерпретируем. У нечеткого классификатора значительно меньше правил (при более высокой точности и схожем числе признаков) по сравнению с деревьями решений, что свидетельствует о лучшей интерпретируемости. Предложенный алгоритм показывает статистически значимые различия с двумя ближайшими по точности аналогами при проверке с помощью критерия Уилкоксона (Таблица 5).

Таблица 3 – Перечень используемых комбинаций алгоритмов для построения нечеткого классификатора

Обозначение	Генерация	R	Отбор признаков	Оптимизация термов
EX-BSA	АЭПК	2	бинарный алгоритм “стая птиц” с функцией трансформации V	Непрерывный алгоритм “стая птиц”
ADP + k -средних	k -средних ($k=ADP_k$)	20	бинарный алгоритм “стая птиц” с функцией трансформации V	Непрерывный алгоритм “стая птиц”
BSA-GEN	BSA-GEN-ADPK	20	бинарный алгоритм “стая птиц” с функцией трансформации V	Непрерывный алгоритм “стая птиц”

Таблица 4 – Результаты построения классификаторов на наборе данных ЭЭГ

Обозначение	R	Число признаков	Процент правильной классификации
EX-BSA	2	93	75,94
ADP + k -средних	20	90	83,70
BSA-GEN	20	98	85,31
LinearSVC	-	75	78,85
Дерево решений	70 правил	81	74,32
Ансамбль деревьев	100 ансамблей по 60 правил	79	78,59

Таблица 5 – Проверка наличия статистически значимых различий между моделями по критерию Уилкоксона

Нулевая гипотеза (НГ)	Между результатами сопоставляемых алгоритмов имеются лишь случайные различия	
Алгоритмы	BSA-GEN и ADP + k -средних	BSA-GEN и LinearSVC
Значение критерия	202,00	209,00
Асимптот. значимость	<0,01	<0,01
Результат	НГ отклоняется	НГ отклоняется

Проверка эффективности верификации субъекта на основе данных электроэнцефалограммы с использованием извлечения признаков с помощью преобразования Холо-Гильберта и метрического обучения. В Таблице 6 приведены обозначения проверяемых вариантов моделей на основе одномерных сверток, используемых признаков и протоколов валидации, результаты точности. Для двух протоколов валидации, LeavePGroupsOut и Kfold, выявлены статистически значимые различия для архитектуры на основе одномерных сверток и подтверждено, что полное исключение субъектов тестовой выборки из обучающей значительно повышает значение метрики ошибки. Протокол LeavePGroupsOut выбран для применения, как более полно отражающий условия работы системы верификации в реальном мире. В Таблице 7 приведены обозначения проверяемых вариантов моделей без протокола кросс-валидации Kfold. Для указанных в Таблице 6 и Таблице 7 архитектур были проведены попарные сравнения по критерию Уилкоксона (Таблица 8).

Таблица 6 – Описание архитектур и результаты метрики ошибки для двух протоколов валидации в случае одномерных сверток

Архитектура	Используемые признаки или преобразования	Протокол валидации	EER, % (чем меньше, тем лучше)
1DCNN-Hilbert-KF	Преобразование Гильберта-Хуанга	Кросс-валидация Kfold	2,42
1DCNN-Hilbert	Преобразование Гильберта-Хуанга	LeavePGroupsOut	10,92

Таблица 7 – Обозначения проверяемых вариантов моделей, используемых признаков и протоколов валидации, значения точности с исключенным протоколом кросс-валидации Kfold

Архитектура	Используемые признаки или преобразования	Протокол валидации	EER, % (чем меньше, тем лучше)
1DCNN-Hilbert	Преобразование Гильберта-Хуанга	LeavePGroupsOut	10,92
2DCNN-Hilbert	Преобразование Гильберта-Хуанга	LeavePGroupsOut	14,62
1DCNN-Holo	Преобразование Холо-Гильберта	LeavePGroupsOut	9,95
3DCNN-Holo	Преобразование Холо-Гильберта	LeavePGroupsOut	12,98
DWT	Коэффициенты дискретного вейвлет преобразования	LeavePGroupsOut	19,33

Таблица 8 – Проверка наличия статистически значимых различий между архитектурами по критерию Уилкоксона

Нулевая гипотеза (НГ)	Между результатами сопоставляемых методов имеются лишь случайные различия				
Методы	1DCNN- Hilbert и 1DCNN- Hilbert-KF	1DCNN- Hilbert и 2DCNN- Hilbert	1DCNN- Holo и 3DCNN-Holo	1DCNN- Holo и 1DCNN-Hilbert	1DCNN- Holo и DWT
Значение критерия	210,00	209,00	210,00	182,50	210,00
Асимптот. значимость	<0,01	<0,01	<0,01	<0,01	<0,01
Результат	НГ отклоняется	НГ отклоняется	НГ отклоняется	НГ отклоняется	НГ отклоняется

По результатам эксперимента оригинальный метод верификации субъекта на основе данных, отличительной особенностью которого является использование извлечение признаков с помощью преобразования Холо-Гильберта и метрического обучения (1DCNN-Holo), показал статистически значимое снижение значения метрики ошибки по сравнению с архитектурой на основе дискретного вейвлет-преобразования (DWT) (снижение на 9,38% метрики ошибки EER) и преобразования Гильберта-Хуанга (1DCNN-Hilbert) (снижение на 0,97% метрики ошибки EER). Также показано статистически значимое снижение ошибки при использовании архитектур на

основе одномерных сверток, как в случае преобразования Гильберта-Хуанга, так и преобразования Холо-Гильберта.

В четвертой главе описано применение разработанных алгоритмов для решения задачи обнаружения наличия эпилептиформ на сегментах ЭЭГ. Для решения задачи определения эпилептиформ на данных ЭЭГ в НИИКиФ был собран набор данных для построения нечеткого классификатора предложенными алгоритмами. Данный нечеткий классификатор превосходит модели-аналоги, достигая точности правильной классификации 84,40%. На основе построенного нечеткого классификатора было создано веб-приложение. Также предложена альтернативная модель на основе преобразования Холо-Гильберта, используя сторонний набор данных Temple University EEG Events Corpus (TUEV). Данная модель превосходит по точности аналоги (достигая точности правильной классификации 85,29%) и может автоматизировать процесс анализа данных ЭЭГ при использовании вместе с другим аппаратом снятия ЭЭГ в НИИКиФ, который поддерживает открытый формат файла “.edf”.

В заключении сделаны выводы о полученных в процессе работы результатах.

В работе представлено решение задачи создания системы анализа электроэнцефалографических данных на основе нейро-сетевых и нечетких методов, имеющей значение для развития научных основ современных информационных технологий.

В процессе выполнения диссертационной работы решалась задача повышения точности и интерпретируемости моделей машинного обучения, предназначенных для верификации субъекта на основе данных ЭЭГ.

Для достижения цели были выполнены следующие задачи:

1) проведен обзор существующих методов верификации субъекта на основе данных электроэнцефалограммы, выявлены недостатки используемых протоколов валидации. Выявлены перспективные методы процессинга данных ЭЭГ, методики обучения и модели, среди которых предложены нечеткие классификаторы, как более интерпретируемые по сравнению с другими моделями;

2) разработан и исследован метод построения моделей верификации субъекта на основе данных электроэнцефалограммы с помощью нейронных сетей и спектрального анализа Холо-Гильберта;

3) разработан и исследован алгоритм формирования структуры нечеткого классификатора с помощью кластеризации и метаэвристического алгоритма “стая птиц”, а также оптимизации параметров нечеткого классификатора;

4) разработан и исследован алгоритм отбора признаков с помощью метаэвристического алгоритма “стая птиц” для нечеткого классификатора;

5) выполнена проверка разработанных алгоритмов на наборах данных ЭЭГ и проведено сравнение с аналогами.

Для подтверждения эффективности предложенного гибридного алгоритма построения базы правил нечеткого классификатора, особенностью которого является использование кластеризации на основе метаэвристического алгоритма “стая птиц” был проведен эксперимент построения соответствующих нечетких классификаторов на 18 наборах данных KEEL. Предложенный алгоритм позволяет повысить точность нечеткого классификатора на рассмотренных наборах данных из различных предметных областей в среднем на 14,38% по сравнению с АЭПК и на 6,00% по сравнению с алгоритмом на основе кластеризации k -средних.

Предложенный гибридный алгоритм построения базы правил нечеткого классификатора, особенностью которого является использование кластеризации на основе метаэвристического алгоритма “стая птиц”, и алгоритм отбора признаков, особенностью которого является использование V-образной функции трансформации и метаэвристического алгоритма “стая птиц”, применены для решения задачи верификации на основе данных ЭЭГ PhysioNet. Нечеткий классификатор, построенный на основе предлагаемых алгоритмов, превосходит классификатор на основе машины опорных векторов с линейным ядром (прирост 6,46% правильной классификации) и нечеткий классификатор на основе АЭПК (прирост 9,37% правильной классификации) по точности. У нечеткого классификатора значительно меньше правил (при более высокой точности и схожем числе признаков) по сравнению с деревьями решений, что свидетельствует о лучшей интерпретируемости.

По результатам эксперимента оригинальный метод верификации субъекта на основе данных, отличительной особенностью которого является использование извлечение признаков с помощью преобразования Холо-Гильберта и метрического обучения, показал статистически значимое снижение ошибки по сравнению с архитектурой на основе дискретного вейвлет-преобразования (снижение на 9,38% метрики ошибки EER) и преобразования Гильберта-Хуанга (снижение на 0,97% метрики ошибки EER, с 10,92% до 9,95%). Также показано статистически значимое снижение значения метрики ошибки при использовании архитектур на основе одномерных сверток. Для двух протоколов валидации, LeavePGroupsOut и Kfold, выявлены статистически значимые различия между протоколами валидации и подтверждено, что полное исключение субъектов тестовой выборки из обучающей значительно повышает значение ошибки верификации.

Полученные метод и алгоритмы могут применяться как для задач верификации на основе ЭЭГ, мультиклассовой классификации сегментов ЭЭГ, так и для других задач обработки сигнала и в сфере машинного обучения в научно-исследовательских и практических целях, что

подтверждено при решении практической задачи определения эпилептиформной активности в данных ЭЭГ.

Автор благодарит за помощь и поддержку в работе над диссертацией научного руководителя, д.т.н., профессора Илью Александровича Ходашинского, а также к.т.н., доцента Константина Сергеевича Сарина и к.т.н., доцента Марину Борисовну Бардамову.

Публикации автора по теме диссертации

Работы, опубликованные в журналах, рекомендованных ВАК:

1) Svetlakov, M. Representation Learning for Electroencephalogram-Based Biometrics Using Holo-Hilbert Spectral Analysis / M. Svetlakov, I. Hodashinsky, K. Sarin // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. – 2022. – Vol. 32. – № 3. – P. 682-688. – DOI: 10.1134/S1054661822030415. (Scopus, WoS).

2) Sarin, K. Extracting Knowledge from Images of Meanders and Spirals in the Diagnosis of Patients with Parkinson's Disease / K. Sarin, I. Hodashinsky, M. Svetlakov // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. – 2022. – Vol. 32. – № 3. – P. 658-664. – DOI: 10.1134/S1054661822030385. (Scopus, WoS).

3) Ходашинский И. А. Биометрические данные и методы машинного обучения в диагностике и мониторинге нейродегенеративных заболеваний: обзор / И. А. Ходашинский, К. С. Сарин, М. Б. Бардамова, М.О. Светлаков [и др.] // Компьютерная оптика. – 2022. – Т. 46, № 6. – С. 988–1020. – DOI: 10.18287/2412–6179-CO-1134. (Scopus, WoS).

4) Ходашинский И. А. Нечёткие классификаторы для диагностики болезни Паркинсона на основе статических рукописных данных / И. А. Ходашинский, Ю. А. Шурыгин, К. С. Сарин, М. Б. Бардамова, А. О. Слёзкин, М. О. Светлаков, Н. П. Коришев // Автометрия. – 2023 – Т. 59, – № 3. – С. 72–85 – DOI: 10.15372/AUT20230200. (Scopus, WoS).

5) Генерация базы правил нечеткого классификатора для диагностики болезни Паркинсона по рукописным данным / М. Б. Бардамова, И. А. Ходашинский, Ю. А. Шурыгин, К. С. Сарин, М. О. Светлаков // Искусственный интеллект и принятие решений. – 2023. – № 2. – С. 31–44. – DOI 10.14357/20718594230203. (Scopus, WoS).

Другие работы, опубликованные по теме диссертации:

6) Svetlakov, M. Gender, Age and Number of Participants Effects on Identification Ability of EEG-based Shallow Classifiers / M. Svetlakov, I. Hodashinsky, A. Slezkin // Proceedings - 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2021, Yekaterinburg, 13–14 мая 2021 года. – Yekaterinburg, 2021. – P. 350-353. – DOI: 10.1109/USBEREIT51232.2021.9455114. (Scopus).

7) Svetlakov, M. O. Clustering-based rule generation methods for fuzzy classifier using Autonomous Data Partitioning algorithm / M. O. Svetlakov, I. A. Hodashinsky // *Journal of Physics: Conference Series*: 18, Tomsk, 27–30 апреля 2021 года. – Tomsk, 2021. – P. 1-7. – DOI: 10.1088/1742-6596/1989/1/012032. (Scopus, WoS).

8) Hodashinsky, I. Identity Authentication Based on Handwritten Signature Using Fuzzy Classifiers Ensemble / I. Hodashinsky, A. Slezkin, M. Svetlakov, E. Kostyuchenko, I. Sidorov // *International Journal of Advanced Research in Engineering and Technology*, – 2021, – Vol.12, - Iss. 1, – P. 539-568. – DOI: 10.34218/IJARET.12.1.2021.050. (Scopus).

9) Tolosana, R. ICDAR 2021 Competition on On-Line Signature Verification / R. Tolosana, R. Vera-Rodriguez, C. Gonzalez-Garcia, M. Svetlakov [и др.] // *Lecture Notes in Computer Science*. – 2021. – Vol. 12824 LNCS. – P. 723-737. – DOI: 10.1007/978-3-030-86337-1_48. (Scopus, WoS).

10) Tolosana, R. SVC-onGoing: Signature verification competition / R. Tolosana, R. Vera-Rodriguez, C. Gonzalez-Garcia, M. Svetlakov [и др.] // *Pattern Recognition*. – 2022. – Vol. 127. – P. 1-14. – DOI: 10.1016/j.patcog.2022.108609. (Scopus, WoS).

11) Hancer, E. Binary PSO Variants for Feature Selection in Handwritten Signature Authentication / E. Hancer, M. Bardamova, I. Hodashinsky, M. Svetlakov [и др.] // *Informatica*. – 2022. – Vol. 33. – No 3. – P. 523-543. – DOI: 10.15388/21-INFOR472. (Scopus, WoS).

12) Svetlakov, M. Representation Learning for EEG-Based Biometrics Using Hilbert–Huang Transform / M. Svetlakov, I. Kovalev, A. Konev [и др.] // *Computers*. – 2022. – Vol. 11. – P. 1-15. – No 3. – DOI: 10.3390/computers11030047. (Scopus, WoS).

13) Tsybenov, B. Feature Selection Methods Comparison for EEG-based Classifier Constructed Using Discrete Wavelet Transform Features / B. Tsybenov, M. Svetlakov, I. Hodashinsky // *Journal of Physics: Conference Series*. – 2022. – Vol. 2291. – № 1. – P. 1-6. – DOI: 10.34218/IJARET.12.1.2021.050. (Scopus, WoS).

14) Bardamova, M. Construction of Fuzzy Classifiers by a Brain Storm Optimization Algorithm / M. Bardamova, I. Hodashinsky, M. Svetlakov // *Lecture Notes in Computer Science*. – 2022. – P. 391–403. – DOI: 10.1007/978-3-031-09677-8_33. (Scopus, WoS).

15) Sarin K. A three-stage fuzzy classifier method for Parkinson's disease diagnosis using dynamic handwriting analysis / K. Sarin, M. Bardamova, M. Svetlakov, N. Koryshev, R. Ostapenko, A. Hodashinskaya, I. Hodashinsky // *Decision Analytics Journal*. – 2023. – Vol. 8. – P. 1-23. – DOI: 10.1016/j.dajour.2023.100274. (Scopus).

16) Светлаков, М. О. Отбор признаков для нечеткого классификатора с использованием метаэвристического алгоритма "bird swarm" / М. О. Светлаков // *Сборник избранных статей научной сессии ТУСУР*. – 2021. – № 1–2. – С. 230–233.

17) Светлаков, М. О. Алгоритм оптимизации параметров нечеткого классификатора, использующий метаэвристику "стая птиц" и метод "островов" / М. О. Светлаков // Передовые инновационные разработки. Перспективы и опыт использования, проблемы внедрения в производство: Сборник научных статей по итогам шестой международной научной конференции, Казань, 31 июля 2019 года. – Казань: ООО "КОНВЕРТ", 2019. – 2019. – С. 224–226.

18) Светлаков, М. О. Генерация правил нечеткого классификатора с помощью алгоритмов ADP и k-means / М. О. Светлаков // Перспективы развития фундаментальных наук: Сборник научных трудов XVIII Международной конференции студентов, аспирантов и молодых ученых, Томск, 27–30 апреля 2021 года. Том 7. – Томск: Национальный исследовательский Томский политехнический университет, 2021. – 2021. – С. 107–109.

19) Sarin K. Machine Learning Methods for Predicting Tumor Volume in Rats after Termination of Complex Treatment with a Varying Dose of Cyclophosphamide / K. Sarin, M. Bardamova, M. Svetlakov [и др.] // Systematic Reviews in Pharmacy. – 2021. – Vol. 12. – No 3. – P. 983-995.

20) Светлаков, М. О. Применение алгоритма кластеризации Autonomous Data Partitioning для построения начальной базы правил нечетких классификаторов / М. О. Светлаков // Сборник избранных статей научной сессии ТУСУР. – 2020. – № 1–2. – С. 79–82.

21) Цыбенков, Б. Д. Отбор и выделение признаков сигнала ЭЭГ с помощью вейвлет-преобразования для аутентификации пользователя / Б. Д. Цыбенков, М. О. Светлаков // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2021. – № 1–2. – С. 317–319.

22) Светлаков, М. О. Влияние выбора функции потерь на точность при обучении представлений с использованием данных электроэнцефалограммы для решения задачи идентификации субъекта / М. О. Светлаков // Сборник избранных статей научной сессии ТУСУР. – 2022. – № 1–2. – С. 92–94.

23) Светлаков, М. О. Применение нечеткого классификатора для аутентификации пользователя на основе данных электроэнцефалограммы / М. О. Светлаков // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2022. – № 1–2. – С. 214–217.

24) Светлаков, М. О. Метод построения нечетких систем на основе модели островов / М. О. Светлаков // Фундаментальные и прикладные научные исследования: актуальные вопросы, достижения и инновации: сборник статей XXV Международной научно-практической конференции: в 2 ч., Пенза, 30 июля 2019 года. Том Часть 1. – Пенза: "Наука и Просвещение", 2019. – 2019. – С. 143–146.

25) Светлаков, М. О. Построение модифицированных систем типа Ангелова-Ягера с использованием метода отбора признаков / М. О. Светлаков // Сборник материалов международных научно-практических конференций: Материалы конференций, Москва, 19 июля 2019 года / Ред. Я. А. Коротких. – г. Москва: Центр научного развития "Большая книга", 2019. – 2019. – С. 27–34.

Свидетельства о государственной регистрации программы для ЭВМ:

1) Программа построения нечетких классификаторов диагностики онкологических заболеваний на основе дискретных метаэвристических алгоритмов / И. А. Ходашинский, М. О. Светлаков [и др.] // Свидетельство о государственной регистрации программы для ЭВМ № 2022610117. Дата регистрации в реестре: 10.01.2022.

2) Программа генерации правил нечеткого классификатора на основе энтропийного алгоритма кластеризации k -means / И. А. Ходашинский, М. О. Светлаков [и др.] // Свидетельство о государственной регистрации программы для ЭВМ № 2022660268. Дата регистрации в реестре: 01.06.2022.

3) Программа извлечения пространственно-временных признаков из биометрических сигналов динамической рукописной подписи для аутентификации пользователя / И. А. Ходашинский, М. О. Светлаков [и др.] // Свидетельство о государственной регистрации программы для ЭВМ № 2021611177. Дата регистрации в реестре: 25.01.2021.