



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Московский государственный технический университет  
имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

---

ФАКУЛЬТЕТ \_\_\_\_\_ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ \_\_\_\_\_  
КАФЕДРА \_\_\_\_\_ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ \_\_\_\_\_

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**  
***К НИР ПО ОБРАБОТКЕ И АНАЛИЗУ ДАННЫХ***  
***НА ТЕМУ:***

**Классификация и сравнительный анализ методов машинного  
обучения для обработки и моделирования глазодвигательной  
активности**

Студент группы ИУ5-35М  
(код группы)

\_\_\_\_\_  
(подпись, дата)

Д.С. Лобанов  
(инициалы и фамилия)

Научный руководитель

\_\_\_\_\_  
(подпись, дата)

Ю.Е. Гапанюк  
(инициалы и фамилия)

Оценка \_\_\_\_\_

Москва, 2025 г.

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Московский государственный технический университет имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

УТВЕРЖДАЮ  
Заведующий кафедрой ИУ5  
(Индекс)  
В.И. Терехов  
(И.О.Фамилия)  
«   »                      2025 г.

**ЗАДАНИЕ**  
**на выполнение научно-исследовательской работы**

по теме Классификация и сравнительный анализ методов машинного обучения для обработки и моделирования глазодвигательной активности

Студент группы ИУ5-35М

Лобанов Дмитрий Сергеевич  
(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)  
исследовательская

Источник тематики (кафедра, предприятие, НИР) КАФЕДРА

График выполнения НИР: 25% к 8 нед., 50% к 10 нед., 75% к 13 нед., 100% к 16 нед.

**Техническое задание** Рассмотреть подходы к интеллектуальной обработке данных глазодвигательной активности. Подготовить экспериментальный набор данных, включающий реальные координаты взгляда и синтетические тепловые карты. Реализовать и исследовать четыре группы алгоритмов: методы без учителя (DBSCAN, HMM), классификаторы с учителем (SVM, Random Forest), рекуррентные (LSTM, GRU) и сверточные (CNN) нейронные сети. Провести сравнительный анализ эффективности моделей в задачах выделения зон интереса, классификации состояний и прогнозирования траекторий. Оценить устойчивость алгоритмов к шуму и способность выявлять поведенческие паттерны. Сформулировать рекомендации по выбору методов машинного обучения в зависимости от целей анализа визуального внимания

**Оформление научно-исследовательской работы:**

Расчетно-пояснительная записка на 15 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Дата выдачи задания «3» сентября 2025 г.

Научный руководитель

Ю.Е. Гапанюк  
(Подпись, дата) (И.О.Фамилия)

Студент группы ИУ5-35М

Д.С. Лобанов  
(Подпись, дата) (И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	4
Методология .....	5
Алгоритмы машинного обучения.....	6
Эксперимент и визуализация результатов .....	8
Обсуждение .....	12
ЗАКЛЮЧЕНИЕ .....	14
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ .....	15

## ВВЕДЕНИЕ

Айтрекинг – это технология, которая позволяет отслеживать направление и траекторию взгляда человека. Благодаря ей можно понять, какие элементы визуальной сцены привлекают внимание, как человек воспринимает информацию и как он принимает решения. Подобные данные находят применение в самых разных областях: от психологии и педагогики до маркетинга, эргономики и цифрового дизайна. Современные системы регистрации взгляда способны фиксировать движения глаз с высокой точностью и частотой, создавая богатые массивы координат. Однако для того, чтобы извлечь из этих данных полезную информацию, необходимо применить методы анализа, которые позволят выявить в них закономерности и интерпретировать поведение наблюдателя.

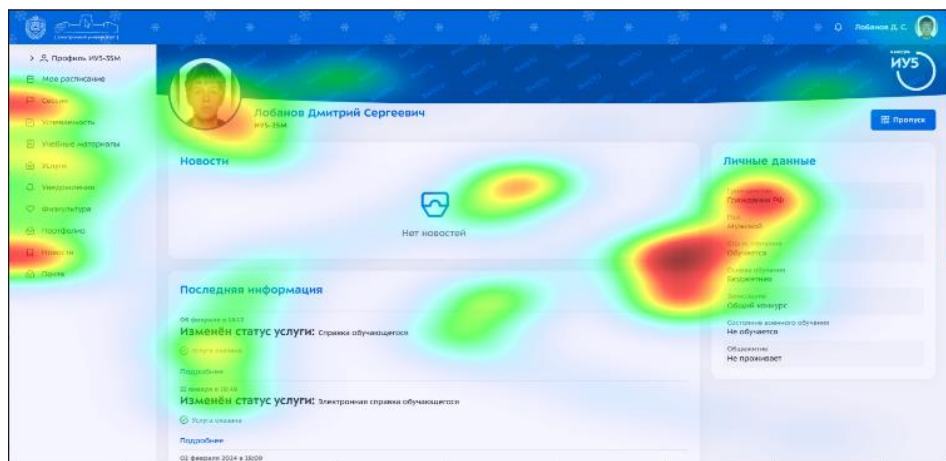
Одним из наиболее перспективных направлений в этой области является использование алгоритмов машинного обучения. Такие методы позволяют решать целый спектр задач: выделять фиксации и саккады, определять зоны интереса (areas of interest, AoI), классифицировать типы визуального поведения, а также строить предсказания дальнейших перемещений взгляда. Выбор конкретного подхода зависит от поставленной задачи, доступного объема данных и требований к результату: где-то важна высокая точность, где-то – интерпретируемость, а где-то – способность обрабатывать большие массивы информации в реальном времени.

В данной работе рассматриваются четыре группы алгоритмов, которые активно применяются в анализе данных айтрекинга: методы без учителя (в том числе кластеризация и скрытые марковские модели), методы с учителем (например, опорные векторы и случайный лес), последовательные модели (LSTM, GRU) и сверточные нейронные сети (CNN). Каждая из этих групп решает свою подзадачу: от выделения устойчивых паттернов до анализа тепловых карт и прогнозирования движения взгляда во времени. В статье описаны особенности применения этих подходов, приведены примеры реализации, а также проведено

их сравнение на общем наборе данных. В качестве исходной информации использовались как реальные записи координат взгляда, так и синтетически сгенерированные визуализации. Цель работы – дать обоснованные рекомендации по выбору алгоритма в зависимости от конкретной задачи анализа визуального поведения.

## Методология

Для анализа применялись координаты взгляда, собранные с помощью самописного айтрекинг-комплекса. Система регистрирует положение глаз пользователя в реальном времени и сохраняет данные в виде последовательности экранных координат с временными метками [1]. В рамках подготовки данных использовались как реальные записи – полученные в процессе взаимодействия с визуальными стимулами, – так и синтетически сгенерированные тепловые карты, имитирующие различные стили визуального внимания. В качестве визуального стимула был выбран интерфейс личного кабинета студента МГТУ им. Н.Э. Баумана.

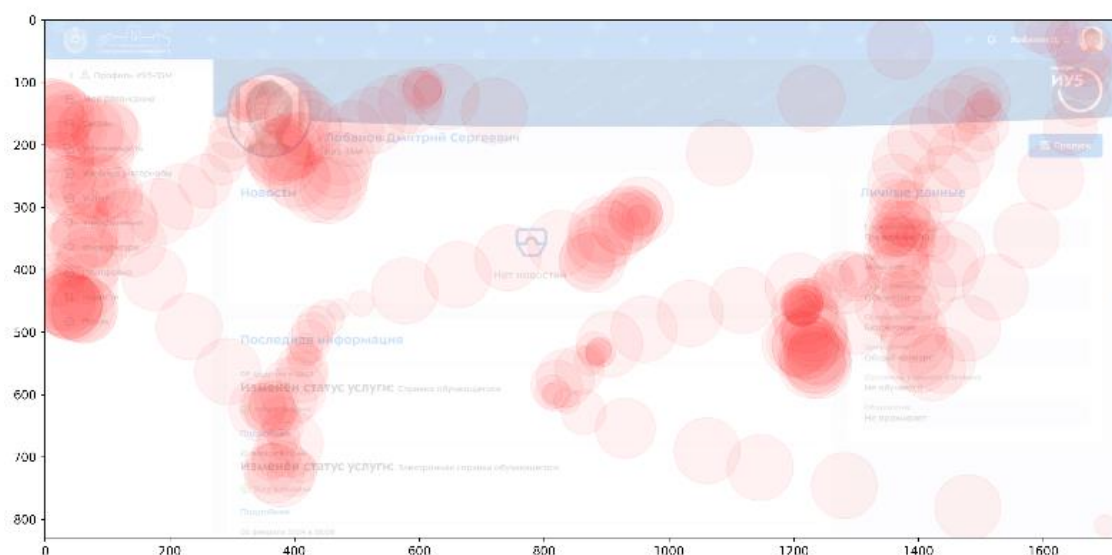


**Рисунок 1.** Тепловая карта, отображающая распределение фиксаций на изображении

Перед подачей в модели координаты проходили базовую фильтрацию: устранялись аномалии, выходящие за границы экрана, применялось сглаживание

по времени. Далее из временного ряда извлекались признаки: скорость и ускорение движения взгляда, расстояние между точками, направление перемещения, радиус разброса и другие. Эти признаки использовались в моделях классификации и предсказания.

Для методов с учителем была выполнена полуавтоматическая разметка: на основе скорости и амплитуды движения точки делились на фиксации и саккады. Полученная разметка служила целевой переменной в задачах бинарной классификации. Для моделей, работающих с изображениями, координаты преобразовывались в тепловые карты заданного разрешения. Эти карты отражали плотность фиксаций в разных областях экрана и служили входом для сверточных нейросетей.



**Рисунок 2.** Визуализация тепловой карты данных взгляда, используемых для обучения и тестирования моделей

## **Алгоритмы машинного обучения**

Для анализа данных движений взгляда в работе были реализованы четыре группы моделей, отличающиеся по способу обучения и типу входных данных. Такой выбор позволил охватить разные подходы – от простых интерпретируемых алгоритмов до нейросетевых архитектур. Каждая модель

применялась в своей задаче: выделение зон внимания, классификация состояний, прогнозирование движения взгляда или анализ тепловых карт. Это обеспечило более полное представление о возможностях машинного обучения в контексте айтрекинга.

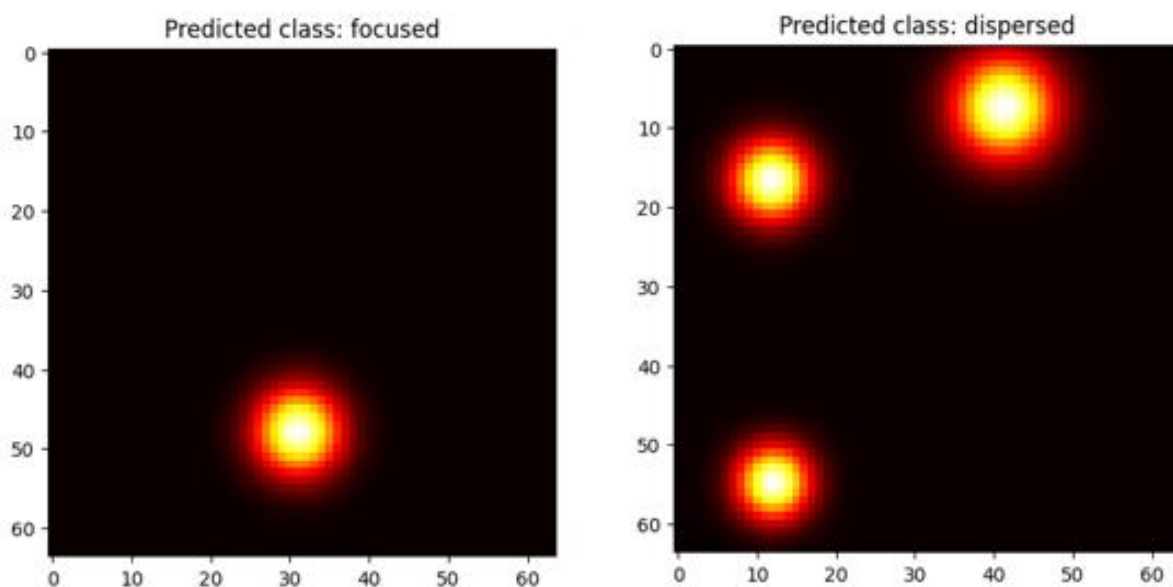
Модели без учителя не требуют заранее размеченных данных и применяются на начальном этапе, когда структура внутри выборки ещё не известна. Кластеризация с помощью DBSCAN позволила выделить области, в которых взгляд задерживался дольше всего, – так называемые зоны интереса. Алгоритм автоматически находит плотные скопления точек, не требуя заранее указывать их количество [2]. Дополнительно использовалась скрытая марковская модель, которая анализирует изменения скорости движения взгляда и делит последовательность на участки, соответствующие фиксациям и саккадам [3]. Эти методы особенно полезны, когда разметка отсутствует или её сложно получить.

Алгоритмы с учителем, такие как метод опорных векторов и случайный лес, обучались на данных, где каждое наблюдение было подписано вручную или полуавтоматически. Целью таких моделей было определить, к какому типу относится текущее состояние взгляда [4]. В качестве признаков использовались скорость, угол между движениями, амплитуда и плотность точек. Эти модели показывают хорошую точность и не требуют больших вычислительных ресурсов, что делает их удобными для сравнительного анализа.

Последовательные модели – LSTM и GRU – были реализованы для обработки временных последовательностей координат. Такие сети позволяют учитывать не только текущую точку, но и историю движения взгляда, предсказывая следующую фиксацию. Это особенно полезно при анализе устойчивости поведения, поиске закономерностей в перемещении взгляда и построении адаптивных систем, реагирующих на внимание пользователя в реальном времени.

Сверточные нейронные сети применялись к изображениям – тепловым картам, на которых визуализировалась плотность фиксаций. Эти карты

строились по координатам взгляда и подавались на вход CNN в виде матриц [5]. Модель обучалась различать типы распределения внимания (например, сфокусированное или рассеянное) на основе синтетических примеров, а затем тестировалась на реальных данных. Такой подход позволил использовать визуальное представление взглядов и применять методы компьютерного зрения к задачам анализа поведения.



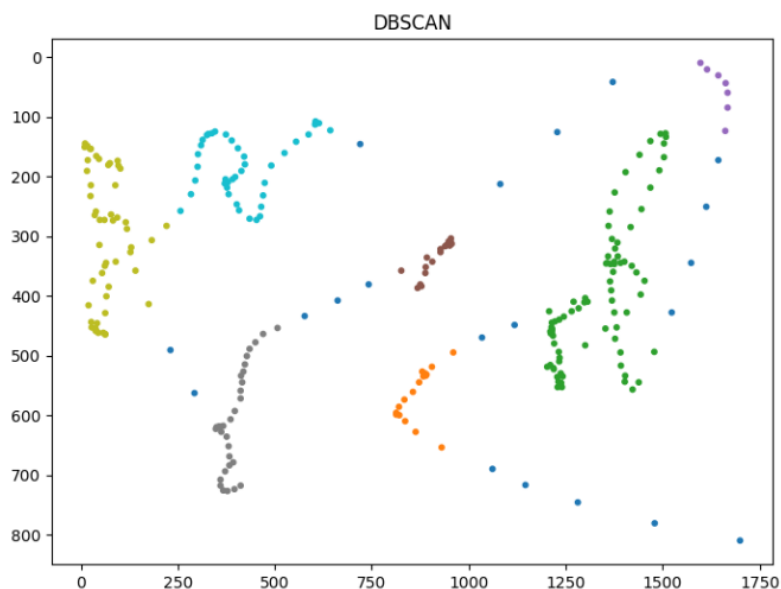
**Рисунок 3.** Примеры сгенерированных тепловых карт, на которых обучалась нейросеть

### **Эксперимент и визуализация результатов**

Для оценки эффективности реализованных моделей был проведён ряд экспериментов на едином наборе данных. В качестве основной информации использовались реальные координаты взгляда, собранные в ходе визуальных заданий. Такой подход позволил протестировать алгоритмы в разных сценариях: от сегментации и классификации точек фиксации до распознавания пространственных шаблонов поведения. Результаты оценивались как количественно, так и визуально – с помощью графиков, карт и сопоставления с базовой разметкой.

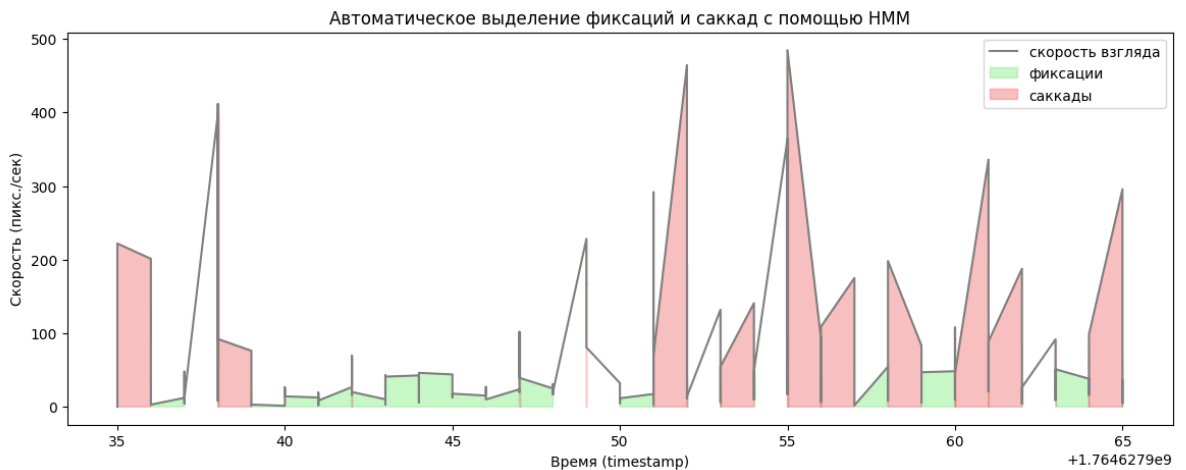


Кластеризация DBSCAN применялась к пространственным данным и визуализировалась в виде цветных облаков точек. Кластеры хорошо соответствовали зонам, в которых взгляд задерживался дольше всего. Алгоритм также продемонстрировал устойчивость к шуму: одиночные случайные точки не попадали в кластеры, что позволило отсеивать нерелевантные движения.



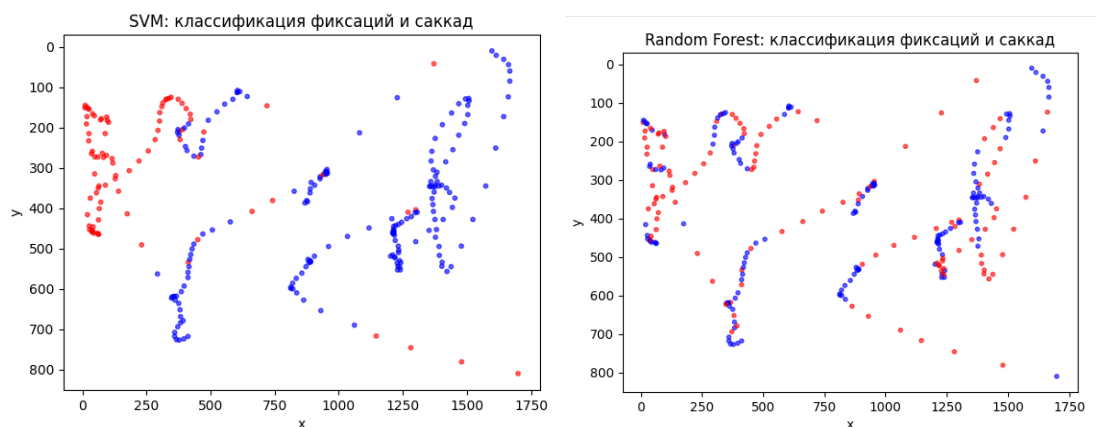
**Рисунок 4.** Кластеры, полученные с помощью DBSCAN на выборке фиксаций

Скрытые марковские модели позволяли автоматически сегментировать временные ряды без ручной разметки. При визуализации на графике координат модель выделяла участки, соответствующие фиксациям и саккадам. Это дало возможность сравнивать такие сегменты с разметкой из других моделей и убедиться в совпадении. Кроме того, НММ оказалась полезной как предварительный фильтр для отбора устойчивых фрагментов движения взгляда.



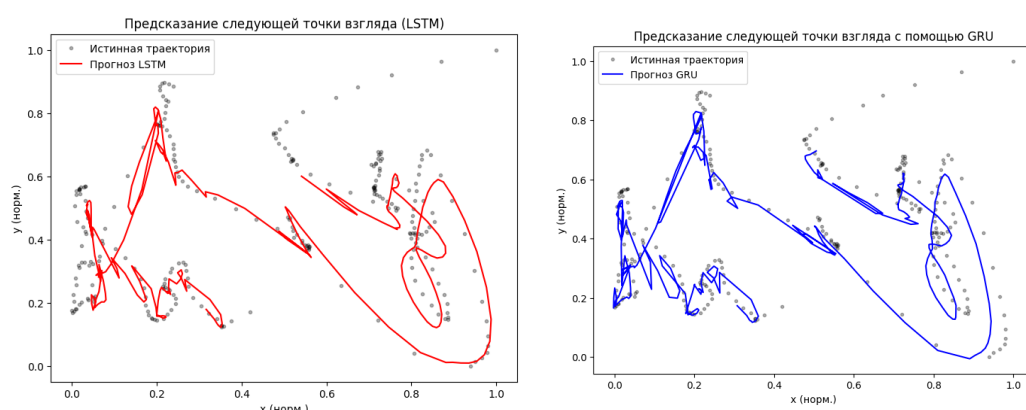
**Рисунок 5.** Сегментация траектории взгляда на фиксации и саккады (НММ)

Для классификации состояний взгляда (фиксация/саккада) использовались SVM и Random Forest. Их производительность оценивалась по точности, полноте и F1-мере на размеченной выборке. Оба алгоритма показали сопоставимо высокие значения метрик, однако визуализация результатов выявила различия в поведении моделей. На рисунке видно, что SVM формирует более чёткие и компактные области классов, с меньшим количеством точек на границе между фиксациями и саккадами. В то же время Random Forest склонен к более рассеянной классификации: встречаются одиночные ошибки, и граница между классами менее выражена. Это указывает на то, что SVM более чувствителен к структуре признаков, тогда как случайный лес показывает лучшую устойчивость к шуму, но может давать менее уверенные границы.



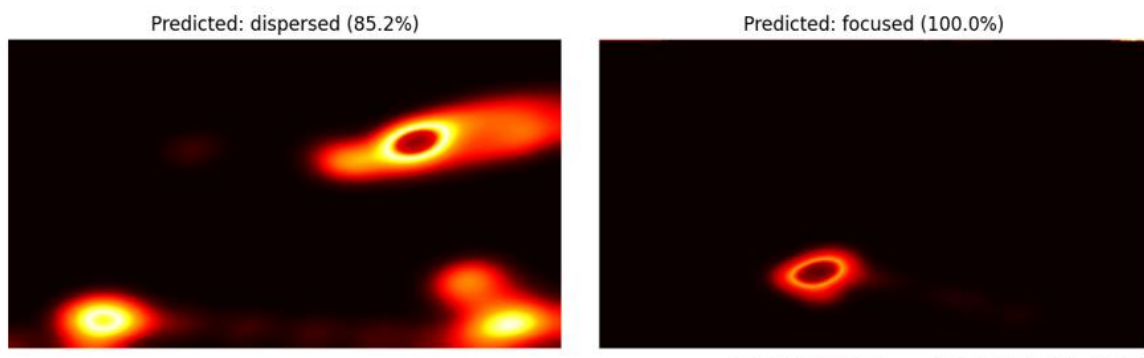
**Рисунок 6.** Результаты классификации SVM/Random Forest на изображении

Модели LSTM и GRU применялись для предсказания координат следующей точки взгляда по данным за предыдущие моменты времени. На предоставленных визуализациях видно, что обе модели в целом следуют за истинной траекторией, воспроизводя форму движения взгляда. При этом GRU даёт более плавный прогноз и чуть лучше сохраняет направление перемещений. LSTM же местами демонстрирует резкие отклонения, особенно при переходах между участками траектории. Это может быть связано с особенностями её внутренней памяти, которая чувствительна к резким скачкам. В целом обе модели способны воспроизводить общую структуру движения, но точность предсказания зависит от типа фрагмента: на прямолинейных участках результаты точнее, чем на резких поворотах или паузах.



**Рисунок 7.** Сравнение предсказанных и реальных точек взгляда (LSTM/GRU)

Сверточная нейросеть обучалась на синтетических тепловых картах, представляющих разные стили визуального внимания. Классы карт задавались заранее и имитировали разные сценарии – например, сосредоточенное внимание в центре или рассеянное по краям. После обучения модель тестировалась на тепловых картах, построенных по реальным координатам взгляда. Визуальное сравнение показало, что модель способна обобщать признаки и корректно определять тип внимания даже на ранее не встречавшихся примерах.



**Рисунок 8.** Примеры классификации тепловых карт моделью CNN

## Обсуждение

Полученные результаты позволяют сопоставить возможности разных подходов к анализу данных айтрекинга. Алгоритмы без учителя показали себя как удобный инструмент на этапе первичного изучения траекторий. DBSCAN оказался особенно полезен для выделения устойчивых зон интереса без необходимости в предварительной разметке, а HMM – для автоматической сегментации по типу активности. Однако эти методы ограничены в возможностях точной настройки и плохо подходят для задач, требующих индивидуального обучения под конкретного пользователя или сценарий.

Методы с учителем обеспечили хорошее качество классификации состояний взгляда. SVM продемонстрировал чёткое разделение классов, но требовал тщательной подготовки признаков. Случайный лес оказался более устойчив к шуму, но при этом менее уверенно разделял пограничные области. Оба метода показали, что даже на ограниченной выборке можно добиться высокой точности, если данные предварительно структурированы и размечены. Однако сам процесс разметки требует значительных временных затрат и не всегда легко масштабируется.

Рекуррентные модели (LSTM и GRU) дали возможность не только предсказывать направление взгляда, но и выявлять поведенческие шаблоны. GRU показал более стабильное поведение на длинных отрезках, тогда как LSTM оказался чувствителен к резким изменениям траектории. Такие модели особенно

полезны для задач, где важно учитывать контекст движения взгляда, но для их обучения требуется значительно больше данных, чем для простых классификаторов. Кроме того, их поведение менее прозрачно по сравнению с классическими методами.

Сверточная нейросеть, обученная на тепловых картах, показала хорошие результаты при переносе на реальные данные, несмотря на обучение на синтетических. Это демонстрирует применимость визуального представления взглядов для задач классификации паттернов внимания. Однако модели такого типа требуют больших вычислительных ресурсов и заранее сформированных изображений, что ограничивает их применение в режиме реального времени.

## ЗАКЛЮЧЕНИЕ

В работе был проведён сравнительный анализ различных подходов к обработке и интерпретации данных движений взгляда с применением методов машинного обучения. Были реализованы и протестированы четыре группы моделей: алгоритмы без учителя, методы с учителем, рекуррентные нейросети и сверточные сети. Эксперименты показали, что выбор модели зависит от конкретной задачи: кластеризация зон интереса, классификация состояний, предсказание траектории или анализ плотности внимания.

Модели без учителя оказались полезными на этапе первичного анализа и при отсутствии разметки, методы с учителем продемонстрировали высокую точность при условии подготовки обучающей выборки. Последовательные сети позволили учитывать временную структуру данных, а сверточные архитектуры – анализировать пространственное распределение внимания. Обобщая результаты, можно рекомендовать использовать комбинацию подходов: простые модели – для базовой фильтрации и классификации, более сложные – для выявления поведенческих паттернов и визуального анализа.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Горячкин Б.С., Якубов А.Р., Аникин Ф.А. Разработка инструментария трекинга глаз // Нейрокомпьютеры: разработка, применение. 2024. Т. 26. № 2. С. 5-12.
2. Naqshbandi K., Gedeon T. Automatic clustering of eye gaze data for machine learning // Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. — 2016. — P. 4089–4094. URL: <https://www.researchgate.net/publication/313586647> (дата обращения: 04.12.2025).
3. Li X. S., Fan Z. Z., Ren Y. Y., Zheng X. L., Yang R. Classification of eye movement and its application in driving based on refined pre-processing and machine learning algorithms // IEEE Access. — 2021. — Vol. 9. — P. 9550773. DOI: 10.1109/ACCESS.2021.9550773.
4. Franceschiello B., Di Noto T., Bourgeois A. Machine learning algorithms on eye tracking trajectories to classify patients with spatial neglect // Computer Methods and Programs in Biomedicine. — 2022. — Vol. 221. — Article 106881. DOI: 10.1016/j.cmpb.2022.106881.
5. Selim A. M., Barz M., Bhatti O. S. A review of machine learning in scanpath analysis for passive gaze-based interaction // Frontiers in Artificial Intelligence. — 2024. — Vol. 7. — Article 1391745. DOI: 10.3389/frai.2024.1391745.