

Сравнительный анализ эффективности применения нейросетевых архитектур и традиционных алгоритмов в задаче сегментации движений глаз

Е. А. Пекло

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

lizapeklo02042002@gmail.com

Аннотация. Задача сегментации данных о движении глаз на саккады, фиксации, прослеживающие движения, моргания и др. является ключевой задачей в отслеживании глазодвигательной активности для исследований и создания инновационных методологий в медицине. Основная цель данной работы заключается в анализе эффективности классических алгоритмов сегментации движений глаз в сравнении с нейросетевыми подходами на базе CNN (Convolutional Neural Network), LSTM (Long Short-Term Memory), гибридных архитектур. В ходе работы проведен обзор алгоритмической базы (традиционных пороговых и статистических алгоритмов), даны описания архитектур используемых моделей, приведена сравнительная таблица результатов валидации решений (Accurasy, F1-score, точность детекции отдельных типов событий и др.). Полученные данные позволяют сформулировать рекомендации по выбору оптимального подхода для решения практических задач в области нейротехнологий и искусственного интеллекта в медицине.

Ключевые слова: нейронные сети; айтрекинг в медицине; сегментация глазодвигательных событий; классификация движений глаз; глазодвигательная активность; фиксации; саккады

I. ОБЩАЯ ИНФОРМАЦИЯ

A. Актуальность и постановка задачи

Движения глаз остаются предметом научного интереса на протяжении нескольких столетий. Первые исследования, посвященные изучению типов глазных движений [1–3] и рассмотрению глазодвигательного поведения в рамках описания когнитивных процессов человека, упоминаются в литературе конца XVIII – XX веков.

Согласно современным исследованиям, отслеживание взгляда для выявления паттернов движения глаз – перспективный метод оценки работы нервной системы и психики человека.

Одним из медицинских направлений, где айтрекинг находит свое применение, является выявление нейрогенеративных патологий, таких как болезнь Паркинсона [4], рассеянный склероз [5], болезнь Альцгеймера [6], постковидный синдром [7] и др.

Кроме того, выявление соответствующих биомаркеров в движении глаз используется в исследованиях, посвященных отклонениям психики и

нервной системы. Например, айтрекинг может быть применен для дополнительной оценки наличия расстройства аутистического спектра (РАС) [8], синдрома дефицита внимания и гиперактивности (СДВГ) [9], депрессии [10] и пр.

Эти данные подчеркивают актуальность применения айтрекинга в области медицины, где своевременно поставленный диагноз – основа сохранения здоровья и качества жизни человека.

В табл. 1 [11] представлены исследования по применению айтрекинга в медицинских задачах и рассматриваемые в них данные.

ТАБЛИЦА I. ДАННЫЕ О ДВИЖЕНИИ ГЛАЗ В МЕДИЦИНСКИХ ЗАДАЧАХ

Тема исследования	Используемыми признаки
Анализ когнитивной нагрузки в физиологии	диаметр зрачка
Когнитивные реакции во время локомоции (движения)	моргание, фиксации, саккады
Классификация стресса	траектория движения взгляда, диаметр зрачка, моргание
Когнитивные аспекты восприятия опасности	диаметр зрачка
Понимание РАС (расстройств аутистического спектра)	траектория движения взгляда, диаметр зрачка, моргание, фиксации
Адаптивное вмешательство при РАС	траектория движения взгляда, диаметр зрачка, моргание, фиксации, саккады
Выявление болезни Паркинсона	траектория движения взгляда, диаметр зрачка

Как видим, во многих исследованиях анализируются данные о саккадах и фиксациях, поскольку именно из этих движений глаз строятся характерные для указанных патологий/состояний паттерны.

Паттерном движений глаз является характерный набор движений. Так, например, ярким паттерном при чтении является чередование прослеживающих движений и саккад.

Саккады [12] – это быстрые, скачкообразные движения глаз между точками фиксации. Макрасаккады длятся от 20 до 200 мс. и могут иметь угловую скорость до 500 %/с.

Фиксации [12] — это периоды, когда глаза остаются относительно неподвижными, позволяя человеку сосредоточиться на конкретном объекте или области.

Длительность фиксации варьируется от 100 до 600 мс, в зависимости от задачи и сложности обрабатываемой информации.

Задача сегментации движений глаз сводится к анализу траектории движения взгляда и выделению на ней сегментов, относящихся к тому или иному виду движений глаз.

Классические алгоритмы сегментации, основанные на пороговых значениях и статистических методах, имеют ряд ограничений, связанных с необходимостью подбора параметров, неустойчивостью к шуму в данных и др. Поэтому актуальной задачей в рамках поиска лучших решения является проведение сравнительного анализа классических алгоритмов с современными нейросетевыми подходами. Эта задача была решена в рамках данной статьи.

В. Исходные данные

Для валидации алгоритмов и обучения моделей нейронных сетей был использован набор данных «Eye Movement Detector Evaluation» [13].

Данный набор содержит записи движений глаз, полученные на профессиональном айтрекере с высокой частотой (500 Гц). Записи имеют «эталонную разметку» саккад и фиксаций, выполненную с помощью алгоритмических и ручных методов. Набор данных включает данные, полученные в различных сценариях: свободный просмотр, чтение, выполнение визуальных задач.

Структура набора данных приведена на рис. 1.

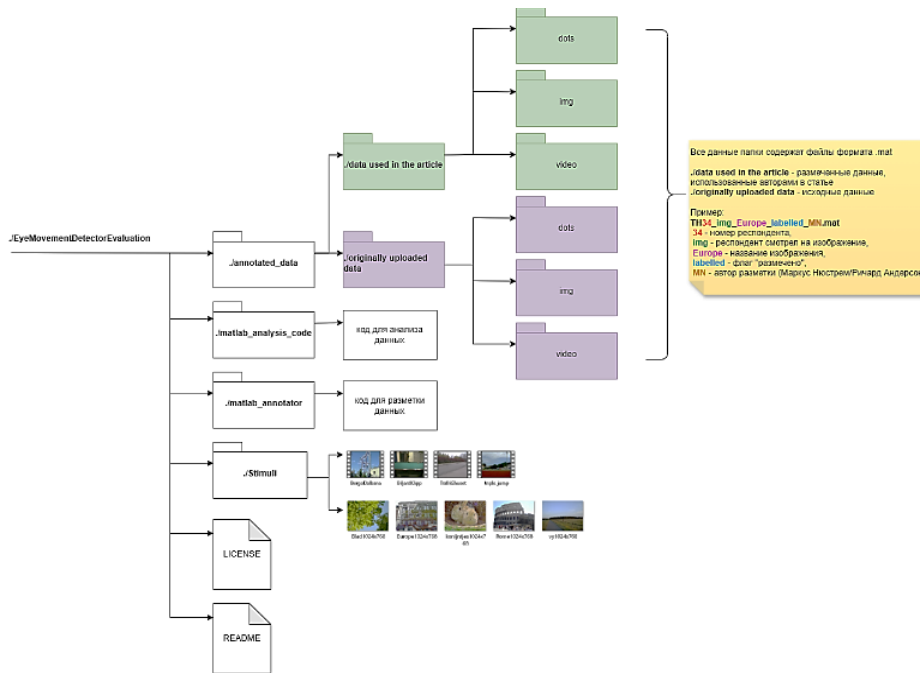


Рис. 1. Структура набора данных «Eye Movement Detector Evaluation»

На рис. 2 на примере записи «TH34_img_Europe_labelled_MN.mat» показаны графики изменения координат взгляда по вертикали и по горизонтали, а также график профиля скорости.

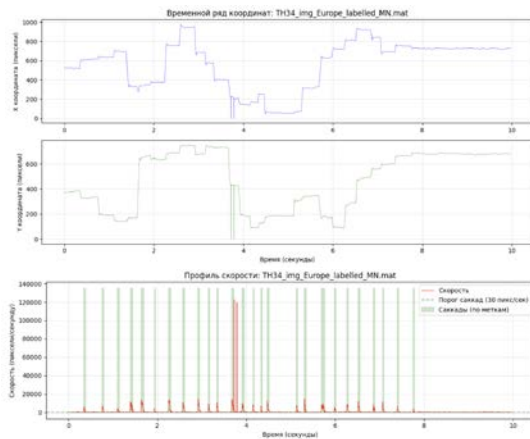


Рис. 2. Пример графиков данных

В рамках предобработки данных набор был подвергнут вычислению дополнительных признаков (скорость, ускорение, дисперсия для каждого сэмпла), сглаживанию траектории взгляда, нормализации признаков для приведения к общему масштабу. После этого набор был разделен на отдельные окна по 100 сэмплов и сбалансирован.

С использованием рассмотренного набора данных была проведена оценка и сравнение реализованных классических и нейросетевых методов с результатами. Каждый из методов был подвергнут валидации на рассмотренном наборе данных, для всех методов построена общая сравнительная табл. 5.

II. КЛАССИЧЕСКИЕ ПОДХОДЫ

А. Краткое описание алгоритмов

В табл. 2 приведено краткое описание каждого из рассмотренных алгоритмов, а также список их входных параметров, каждый из которых требует дополнительного подбора и зависит от данных.

ТАБЛИЦА II. ОПИСАНИЕ КЛАССИЧЕСКИХ АЛГОРИТМОВ

Алгоритм	Краткое описание	Параметры
I-VT [14]	Алгоритм классифицирует точки взгляда на основе порогового значения скорости.	порог скорости, минимальная длительность саккады, частота дискретизации
I-VDT [14]	Комбинированный алгоритм, который использует оба порога – и скорости, и дисперсии.	порог скорости, порог дисперсии, минимальная длительность фиксации, частота дискретизации

Каждый из описанных алгоритмов был описан своим классом с использованием общего родительского интерфейса.

В. Подбор параметров

Для каждого из классических алгоритмов был произведен подбор параметров, при которых алгоритм достигает наилучшего результата по F1-score. Лучшие параметры были использованы для получения финальных результатов работы алгоритма, включенных в итоговую сравнительную таблицу.

В табл. 3 приведен список лучших подобранных параметров для каждого из алгоритмов.

ТАБЛИЦА III. ЛУЧШИЕ ПАРАМЕТРЫ АЛГОРИТМОВ

Алгоритм	Параметры	Лучшие значения
I-VT	порог скорости саккады (пиксель/сек)	1000
	минимальная длительность саккады (мс)	50
	частота дискретизации (Гц)	500
I-VDT	порог скорости (пикс/сек)	100
	порог дисперсии (пиксели)	50
	минимальная длительность фиксации (мс)	20
	частота дискретизации (Гц)	500

С. Результаты

В табл. 4 приведены результаты тестирования классических алгоритмов I-VT и I-VDT на валидационной выборке. Данные алгоритмы выбраны как одни из лучших алгоритмов сегментации саккад [14], поскольку именно саккады являются характерными и наименее численными движениями глаз, представленными в наборе данных.

Для сравнения были выбраны следующие метрики: Accuracy – общая доля правильных ответов модели; F1-score (Saccade) – гармоническое среднее точности и полноты определения саккад; precision (Saccade) – доля верно найденных саккад среди всех предсказанных; recall (Saccade) – доля найденных саккад среди всех реальных; specificity (Fixation) – доля верно классифицированных фиксаций.

ТАБЛИЦА IV. РЕЗУЛЬТАТЫ ВАЛИДАЦИИ АЛГОРИТМОВ

Алгоритм	Accuracy	F1-score (Sac.)	Precision (Sac.)	Recall (Sac.)	Specificity (Fix.)
I-VT	0.9202	0.6366	0.8709	0.5017	0.9880
I-VDT	0.9213	0.7486	0.6747	0.8407	0.9344

III. НЕЙРОННЫЕ СЕТИ

А. Описание архитектур моделей

В рамках данной части работы представлено описание архитектур моделей нейронных сетей, в числе которых CNN, LSTM и гибридная архитектура, сочетающая преимущества которых CNN и LSTM. Все модели применялись для бинарной классификации основных глазоподвижных событий – саккад и фиксаций.

Сверточная нейронная сеть (CNN), архитектура которой приведена на рис. 3, позволяет выделять локальные паттерны во временном ряду.

Layer (type)	Output Shape	Param #
input_layer_32 (InputLayer)	(None, 100, 5)	0
conv1d_174 (Conv1D)	(None, 100, 32)	512
batch_normalization_125 (BatchNormalization)	(None, 100, 32)	128
conv1d_175 (Conv1D)	(None, 100, 32)	3,104
batch_normalization_126 (BatchNormalization)	(None, 100, 32)	128
conv1d_176 (Conv1D)	(None, 100, 64)	6,208
batch_normalization_127 (BatchNormalization)	(None, 100, 64)	256
conv1d_177 (Conv1D)	(None, 100, 64)	12,352
batch_normalization_128 (BatchNormalization)	(None, 100, 64)	256
conv1d_178 (Conv1D)	(None, 100, 128)	24,704
batch_normalization_129 (BatchNormalization)	(None, 100, 128)	512
conv1d_179 (Conv1D)	(None, 100, 128)	49,280
batch_normalization_130 (BatchNormalization)	(None, 100, 128)	512
conv1d_180 (Conv1D)	(None, 100, 2)	258

Total params: 98,210 (383.63 KB)
Trainable params: 97,314 (380.13 KB)
Non-trainable params: 896 (3.50 KB)

Рис. 3. Архитектура CNN

Таковыми локальными паттернами в задаче сегментации саккад и фиксаций могут являться характерные пики скорости саккад и низкоамплитудный шум при фиксациях.

LSTM (рис. 4), с другой стороны, позволяет моделировать долгосрочные зависимости в последовательности, что позволяет учитывать контекст. Такой зависимостью, например, может быть выделено следование саккад за фиксациями.

Layer (type)	Output Shape	Param #
input_layer_33 (InputLayer)	(None, 100, 5)	0
lstm_31 (LSTM)	(None, 100, 64)	17,920
lstm_32 (LSTM)	(None, 100, 32)	12,416
time_distributed_5 (TimeDistributed)	(None, 100, 2)	66

Total params: 30,402 (118.76 KB)
Trainable params: 30,402 (118.76 KB)
Non-trainable params: 0 (0.00 B)

Рис. 4. Архитектура LSTM

Перспективным подходом выглядит также использование гибридной модели, объединяющей описанные преимущества CNN и LSTM. Архитектура гибридной модели представлена на рис. 5.

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 100, 5)	0
conv1d (Conv1D)	(None, 100, 32)	512
batch_normalization (BatchNormalization)	(None, 100, 32)	128
conv1d_1 (Conv1D)	(None, 100, 32)	3,104
batch_normalization_1 (BatchNormalization)	(None, 100, 32)	128
lstm (LSTM)	(None, 100, 64)	24,832
conv1d_2 (Conv1D)	(None, 100, 64)	12,352
batch_normalization_2 (BatchNormalization)	(None, 100, 64)	256
conv1d_3 (Conv1D)	(None, 100, 64)	12,352
batch_normalization_3 (BatchNormalization)	(None, 100, 64)	256
conv1d_4 (Conv1D)	(None, 100, 2)	130

Total params: 54,950 (211.13 KB)
 Trainable params: 53,966 (209.63 KB)
 Non-trainable params: 384 (1.50 KB)

Рис. 5. Архитектура гибридной модели

В. Обучение

Все нейросетевые модели обучались на размеченных данных, описанных в главе 1. Окна длительностью 100 временных точек (200 мс при частоте дискретизации, что соответствует верхнему пределу длительности саккад) подавались на вход моделей. На выходе моделей для каждого отсчета формировался бинарный ответ, указывающий предсказанный класс, – саккада или фиксация. Пример окна с разметкой глазодвигательных событий приведен на рис. 6. Саккады обозначены красным цветом, фиксации – зеленым.

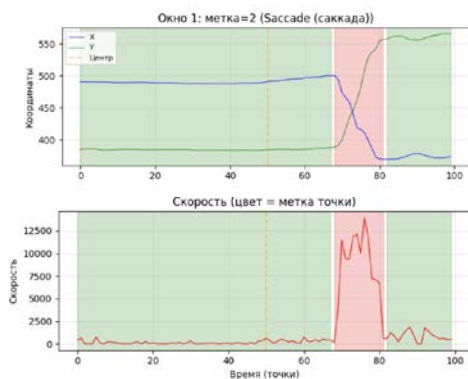


Рис. 6. Визуализация окна данных

В рамках обучения применяли классические способы оптимизации и регуляризации. Применялся оптимизатор Adam с начальной скоростью обучения 0.001 и размером пакета 32 окна. В качестве функции потерь использовалась категориальная кросс-энтропия. Для предотвращения переобучения использовалась ранняя остановка, а также слои регуляризации (Dropout и BatchNormalization). Максимальное количество эпох было ограничено 50 единицами.

С. Результаты

В табл. 5 приведены результаты тестирования моделей нейронных сетей на валидационной выборке.

ТАБЛИЦА V. РЕЗУЛЬТАТЫ ВАЛИДАЦИИ МОДЕЛЕЙ

Модель	Accuracy	F1-score (Sac.)	Precision (Sac.)	Recall (Sac.)	Specificity (Fix.)
CNN	0.9568	0.8355	0.8889	0.7881	0.9841
Hybrid	0.9565	0.8233	0.9494	0.7268	0.9937
LSTM	0.9374	0.7468	0.8548	0.6631	0.9818

IV. СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕЗУЛЬТАТОВ

Наилучшие результаты достигнуты с использованием сверточной нейронной сети (Accuracy = 0.9568, F1 = 0.8355). Гибридная архитектура показала сопоставимую точность (0.9565) и наивысшую прецизионность (0.9494). Среди классических алгоритмов лучшим оказался I-VDT (Accuracy = 0.9213, F1 = 0.7486), однако он уступает нейросетевым моделям.

Итоговая сравнительная табл. 6 содержит обобщенные данные о результатах валидации описанных методов.

ТАБЛИЦА VI. СРАВНИТЕЛЬНАЯ ТАБЛИЦА РЕЗУЛЬТАТОВ

Модель	Accuracy	F1-score (Sac.)	Precision (Sac.)	Recall (Sac.)	Specificity (Fix.)
CNN	0.9568	0.8355	0.8889	0.7881	0.9841
Hybrid	0.9565	0.8233	0.9494	0.7268	0.9937
LSTM	0.9374	0.7468	0.8548	0.6631	0.9818
I-VT	0.9202	0.6366	0.8709	0.5017	0.9880
I-VDT	0.9213	0.7486	0.6747	0.8407	0.9344

V. ЗАКЛЮЧЕНИЕ

В данной работе было проведено экспериментальное исследование классических алгоритмов сегментации саккад и фиксаций в сравнении с современными подходами на базе нейронных сетей.

Полученные результаты позволяют сделать вывод о том, что применение нейронных сетей в задаче сегментации движений глаз позволяет получать результаты превосходящие результаты классических алгоритмов, что делает их использование перспективным подходом, который не зависит от пороговых значений.

БЛАГОДАРНОСТЬ

Выражаю благодарность своему научному руководителю Юлии Александровне Шичкиной и консультанту по реализации исследовательской работы Семену Евгеньевичу Ильину за поддержку в создании настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

- [1] Darwin R.W. Sect. XL. On the ocular spectra of light and colours // Zoonomia; or, The laws of organic life: in 2 vols. Vol. 1. London: Printed for J. Johnson, 1794. P. 534–566.
- [2] Helmholtz H. von. Handbuch der physiologischen Optik : mit 213 in den Text eingedruckten Holzschitten und 12 Tafeln. Leipzig: Leopold Voss, 1866. 874 s.
- [3] Javal É. Essai sur la physiologie de la lecture / É. Javal // Annales d'Oculistique. 1879. Vol. 82. P. 242–253.
- [4] Waldthaler J., Stock L., Student J., Sommerkorn J., Dowiasch S., Timmermann L. Antisaccades in Parkinson's Disease: A Meta-Analysis [Electronic resource] // Neuropsychology Review. 2021. — DOI: 10.1007/s11065-021-09489-1.
- [5] Polet K., Hesse S., Cohen M., Morisot A., Joly H., Kullmann B., Mondot L., Pesce A., Lebrun-Frenay C. Video-oculography in multiple sclerosis: Links between oculomotor disorders and brain magnetic resonance imaging (MRI) [Electronic resource] // Multiple Sclerosis and Related Disorders. 2020. Vol. 40. P. 101969. — DOI: 10.1016/j.msard.2020.101969.
- [6] Wilcockson T. D. W., Mardanbegi D., Xia B., Taylor S., Sawyer P., Gellersen H. W., Leroi I., Killick R., Crawford T. J. Abnormalities of saccadic eye movements in dementia due to Alzheimer's disease and mild cognitive impairment [Electronic resource] // Aging. 2019. Vol. 11, No. 15. P. 5389–5398. DOI: 10.18632/aging.102118.

- [7] Зайцев А.А., Сергеева К.В., Петрова М.Ю., Смирнов И.Н. Изучение когнитивных нарушений у пациентов с постковидным синдромом: аномалии движения глаз и влияние лобно-подкорковых связей с помощью айтрекинга и машинного обучения // Журнал неврологии и психиатрии им. С.С. Корсакова. 2024. Т. 124, № 2. С. 45–52.
- [8] Ван Вэньхао (Wang W. H.). Классификация и оценка аутизма на основе технологии отслеживания движения глаз и глубокого обучения [Electronic resource]: дис. ... магистра психологии / Цзянсийский педагогический университет. Наньчан, 2023. 98 с. — DOI: 10.27178/d.cnki.gjxsu.2023.000617. — URL: <https://oversea.cnki.net/> (дата обращения: 13.04.2026). На кит. яз.
- [9] Lee S., Lee S., Jeong I., Jeong J., Park H., Kwon M.-K., Zanto T., Sul S., Jung D. Exploring early-stage orienting behavior using an eye tracker for attention deficit hyperactivity disorder classification [Electronic resource] // Scientific Reports. 2026. Vol. 16. Article number 8671. — DOI: 10.1038/s41598-026-41419-0. — URL: <https://www.nature.com/articles/s41598-026-41419-0> (дата обращения: 10.03.2026).
- [10] Kiss G., Viktor P. Artificial Intelligence-Based Depression Detection [Electronic resource] // Sensors. 2026. Vol. 26, No. 2. P. 748. — DOI: 10.3390/s26020748. — URL: <https://www.mdpi.com/1424-8220/26/2/748> (дата обращения: 13.04.2026).
- [11] Bozkir E., Özdel S., Wang M., David-John B., Gao H., Butler K., Jain E., Kasneci E. Eye-Tracked Virtual Reality: A Comprehensive Survey on Methods and Privacy Challenges [Electronic resource] // IEEE Transactions on Visualization and Computer Graphics. 2026. Vol. 32, No. 5. P. 1–20. — DOI: 10.1109/TVCG.2026.11366239. — URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=11366239> (дата обращения: 10.03.2026).
- [12] Барабанщиков В.А., Жегалло А.В. Айтрекинг: методы регистрации движений глаз в психологических исследованиях и практике. М.: Когито-Центр, 2014. 128 с. (Университетское образование).
- [13] Andersson R. Eye Movement Detector Evaluation [Electronic resource] // GitHub. — URL: <https://github.com/richardandersson/EyeMovementDetectorEvaluation> (дата обращения: 13.04.2026).
- [14] M3stark. Eye_tracking_proj [Electronic resource] // GitHub. — URL: https://github.com/M3stark/Eye_tracking_proj (дата обращения: 13.04.2026).