

УДК 004.93'12

## Предварительная обработка изображений с помощью модуля имитации сетчатки

Меркелов Михаил Валерьевич<sup>(\*)</sup>

merkelovmv@bmstu.ru

SPIN-код: 1900-0253

Локтев Даниил Алексеевич

loktevdan@bmstu.ru

SPIN-код: 4883-9877

МГТУ им. Н.Э. Баумана, Москва, Россия

**Аннотация.** Оценивается влияние предварительной обработки изображений с помощью модуля имитации сетчатки на распознавание образов нейронной сетью. Тестирование работы модуля имитации сетчатки проводилось на полносвязной нейронной сети. Было создано две модели с одинаковыми параметрами. Одна обучалась на выборке изображений фруктов, сфотографированных с разных ракурсов, в оттенках серого. Другая обучалась на той же выборке, но предварительно обработанной модулем имитации сетчатки. Приведены графики обучения нейронной сети на обеих выборках. Представлены графики точности распознавания образов нейронной сетью. Приведены рекомендации по использованию описанного модуля.

**Ключевые слова:** распознавание образов, классификация, нейронные сети, сетчатка глаза, обучение с учителем

## Image preprocessing using the retina simulation module

Merkelov Mikhail Valerievich<sup>(\*)</sup>

merkelovmv@bmstu.ru

SPIN-code: 1900-0253

Loktev Daniil Alekseevich

loktevdan@bmstu.r

SPIN-code: 4883-9877

BMSTU, Moscow, Russia

**Abstract.** The paper evaluates the effect of image preprocessing using the retinal simulation module on image recognition by a neural network. The retinal simulation module was tested on a fully connected neural network. Two models with the same parameters were created. One was trained on a sample of fruit images photographed from different angles, in shades of gray. The other was trained on the same sample, but pre-processed by the retinal simulation module. The graphs of neural network training on both samples are given. The results of image recognition by a neural network are presented. Recommendations for using the described module are given.

**Keywords:** pattern recognition, classification, neural networks, retina, supervised learning

**Введение.** Во многих отраслях человеческой деятельности, таких как медицина, строительство, транспорт и т. д. [1, 2], актуальной задачей является получение важной для процессов информации и автоматизация деятельности

в целом, что возможно, в том числе, путем внедрения визуального распознавания объектов в системы их контроля и мониторинга и дальнейшее отображение информации с помощью средств дополненной реальности, что в итоге повлияет на своевременное реагирование сотрудников и позволит повысить безопасность их работы.

Существующие методы и алгоритмы распознавания образов часто основаны на существующих в природе механизмах восприятия информации, в том числе на перцептивной системе человека, включающей в себя и зрительную систему, состоящую из таких основных элементов, как сетчатка, зрительный нерв, зрительная кора. Исследование и анализ подобной системы человека, визуализирующей окружающую среду, до сих пор актуально и может привести к появлению новых методов и алгоритмов детектирования и распознавания объектов [3, 4], определения параметров объектов [5, 6], построения виртуальной/дополненной реальности и т. д. Для этого в данной работе предлагается разработать модуль имитации сетчатки глаза и оценить его влияние на точность распознавания образов нейронной сетью.

**Материалы и методы.** В первую очередь был проведен поиск существующих моделей сетчатки и работ, в которых описано использование подобных моделей в качестве модулей для предварительной обработки изображений для повышения качества распознавания образов.

В рассмотренных работах [7–10] модели создавались для исследования физиологических особенностей зрительной системы человека и систем в области робототехники. Использование модели сетчатки в качестве средства предварительной обработки изображения для улучшения распознавания образов в настоящий момент описано не было.

Разрабатываемый модуль состоит из слоя фоторецепторов, биполярных клеток и ганглиозных клеток.

Первым этапом разработки является проектирование слоя фоторецепторов. Для модуля интенсивность пикселей входного изображения является аналогом потоков пучков света, поступающих на рецепторы реального глаза. Поэтому каждый пиксель должен соответствовать каждой «клетке» первого слоя программной сетчатки. Программные колбочки должны реагировать только на конкретный цвет, поэтому они должны брать значение интенсивности с конкретного цветового канала изображения. Палочки просто возбуждаются при наличии какого-то уровня света, поэтому в модуле они должны переводить пиксель в оттенки серого и уже полученное значение передавать на другие слои программной сетчатки.

В качестве закона, моделирующего расположение рецепторов в реальной сетчатке, была взята матрица распределения пикселей на фотоматрице фотоаппарата X-Trans [11].

Фильтр в оригинальном виде не подходит для программной сетчатки, так как не имеет пикселей, считывающих интенсивность света по всем цветовым каналам сразу. Поэтому он был видоизменен для возможности имитировать палочки на сетчатке. Полученный фильтр показан на рис. 1.

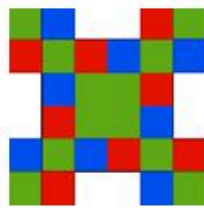


Рис. 1. Видоизмененный фильтр X-Trans

Следующим этапом является разработка слоев биполярных и ганглиозных клеток.

Для модуля был разработан вариант рецептивного поля клеток, имитирующего работу полей реальных клеток сетчатки. На рис. 2 показана организация рецептивных полей клеток реализованной в программном виде сетчатки.

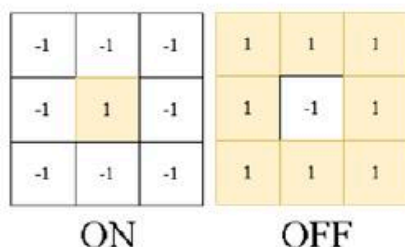


Рис. 2. Организация рецептивных полей биполярных и ганглиозных клеток сетчатки, реализованной в программном виде

В данной реализации рецептивные поля ON и OFF клеток в слое чередуются. По четным индексам расположены OFF клетки, по нечетным — ON.

Для слоя ганглиозных клеток используются такие же варианты организации рецептивных полей клеток ON-типа и OFF-типа и такие же варианты расположения клеток ON-типа и OFF-типа в слое.

**Результаты.** Для проверки модуля использовалась выборка [12], состоящая из фотографий фруктов, сфотографированных с разных ракурсов, размером 100 на 100 пикселей, которые проклассифицированы по 39 классам. Данный набор был разделен на две части — обучающую (18505 изображений), из которой 20 % было выделено на валидационную выборку, и тестовую (6215 изображений).

Для распознавания образов была создана полносвязная нейронная сеть на языке Python с помощью библиотеки keras. Полученная нейронная сеть состоит из входного слоя, принимающего изображения размером 100 на 100 пикселей, и четырех полносвязных слоев. Последний слой имеет 39 нейронов, что соответствует количеству классов объектов в выборке [12].

Было создано два варианта обучения описанной выше нейронной сети: в первом случае оно происходило на изображениях, представленных в градациях серого цвета (рис. 3, а), во втором — на изображениях, пропущенных через модуль имитации сетчатки глаза (рис. 3, б).

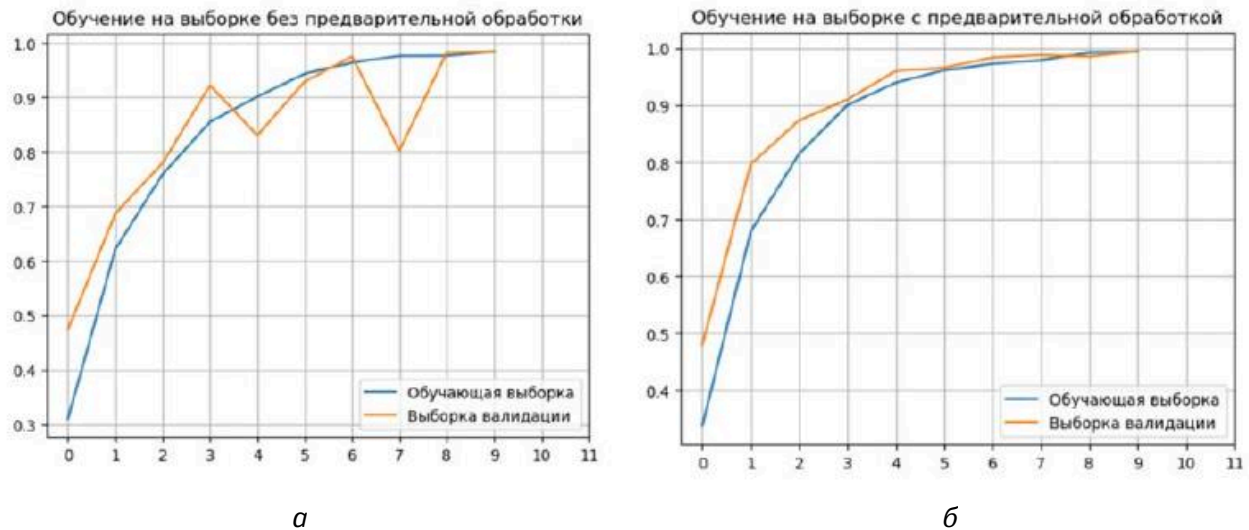


**Рис. 3.** Графики обучения нейронной сети на выборках без (а) и с (б) предварительной обработкой

Первая сеть, обученная на фотографиях в оттенках серого, на тестовой выборке показала точность распознавания в 79 % (рис. 3, а). Обученная на обработанных модулем изображениях сеть показала точность распознавания фруктов в 89 % (рис. 3, б). То есть модуль в данном случае позволил нейронной сети распознать на 10 % больше объектов.

Для предлагаемых нейронных сетей было проведено по 100 запусков (рис. 4), для каждого из которых сначала производилась очистка, чтобы предыдущие результаты обучения не влияли на последующие запуски, и используемая выборка заново перемешивалась случайным образом.



**Рис. 4.** Точность распознавания образов обученными нейронными сетями на тестовой выборке

**Обсуждение полученных результатов.** В процессе исследования было выявлено, что изображения с низким разрешением сильно искажаются модулем имитации сетчатки, в связи с чем требуется больше времени на обучение нейронной сети. Из результата 100 запусков (рис. 4) видно, что на предварительно обработанной модулем выборке точность распознавания повышается в среднем на 8–9 %. При этом нейронная сеть лучше адаптируется под новые изображения, что видно на графиках обучения, где модель, обучающаяся на обработанных изображениях, имеет менее ломаный график на выборке валидации.

На 100 запусках на выборке с обработанными изображениями получилась дисперсия и среднее квадратичное отклонение ниже, чем на выборке без обработки, что подтверждает, сделанный ранее, тезис о более плавном и стабильном обучении.

**Заключение.** В работе проанализировано влияние разработанного модуля на результат распознавания образов нейронной сетью. Было показано, что предварительная обработка с помощью программной сетчатки позволяет повысить качество распознавания образов на 8–9 % и в целом стабилизировать обучение. На выборке с предварительной обработкой нейронная сеть лучше адаптируется под новые образы. Было выявлено, что модуль сильно искажает изображения с низким разрешением, поэтому разработанную сетчатку рекомендуется использовать на разрешениях 50 на 50 пикселей и более. Нахождение фильтров, рецептивных полей клеток и подбор архитектур нейронных сетей, с которыми модуль будет позволять достигать наилучшего результата, а также их реализация в системах виртуальной и дополненной реальности являются предметом дальнейших исследований.

## Список источников

- [1] Тупилейкина П.Ю. Исследование современного состояния и возможных перспектив в области искусственного интеллекта. *Научные исследования XXI века*, 2023, № 6 (26), с. 41–44.
- [2] Локтева О.С., Локтев Д.А. Безопасность труда в транспортной отрасли. *Внедрение современных конструкций и передовых технологий в путевое хозяйство*, 2017, т. 11, № 11 (11), с. 80–89.
- [3] Адамова А.А., Зайкин В.А., Гордеев Д.В. Методы и технологии машинного обучения и нейросетевых технологий в задачах компьютерного зрения. *Нейрокомпьютеры: разработка, применение*, 2021, т. 23, № 4, с. 25–39. <https://doi.org/10.18127/j19998554-202104-03>
- [4] Локтев Д.А., Алфимцев А.Н. Измерение расстояния до движущегося объекта с помощью комплексной системы видеомониторинга. *Инженерный журнал: наука и инновации*, 2013, № 11(23), с. 4. <https://doi.org/10.18698/2308-6033-2013-11-996>
- [5] Локтев А.А., Локтев Д.А. Определение параметров состояния объектов путем анализа их изображений с помощью нейронной сети. *Транспортное строительство. Всерос. науч.-техн. конф.: сб. ст.* Москва, Перо, 2020, с. 4–16.
- [6] Локтев А.А., Локтев Д.А. Оценка измерений расстояния до объекта при исследовании его графического образа. *Вестник МГСУ*, 2015, № 10, с. 54–65.
- [7] Lebedev D.S., Marshak D.W. Amacrine cell contributions to red-green color opponency in central primate retina: a model study. *Visual Neuroscience*, 2007, vol. 24, no. 4, pp. 535–47. <https://doi.org/10.1017/S0952523807070502>

- [8] Утробин В.А. Информационные модели организации сетчатки глазного яблока. *Труды НГТУ им. Р.Е. Алексеева*, 2010, № 3, с. 66–74
- [9] Маштаков А.П., Путинцева К.А. Программный комплекс обработки изображений на основе модели зрения человека. *Программные системы: теория и приложения*, 2019, т. 10, № 4 (43), с. 111–139. <https://doi.org/10.25209/2079-3316-2019-10-4-111-139>
- [10] Wohrer A., Kornprobst P. Virtual Retina: A biological retina model and simulator, with contrast gain control. *Journal of Computational Neuroscience*, 2009, vol. 26, pp. 219–249. <https://doi.org/10.1007/s10827-008-0108-4>
- [11] Fujifilm X-Trans sensor technology. URL: <https://www.fujifilm.eu/uk/products/digital-cameras/model/x-pro1/features-4483/aps-c-16m-x-trans-cmos> (дата обращения 18.02.2025).
- [12] Oltean M. Fruits-360 dataset. URL: <https://www.kaggle.com/datasets/moltean/fruits> (дата обращения 18.02.2025).