

ВЛИЯНИЕ МОДУЛЯ ИМИТАЦИИ СЕТЧАТКИ ГЛАЗА ЧЕЛОВЕКА НА КАЧЕСТВО РАСПОЗНАВАНИЯ ОБРАЗОВ НЕЙРОННОЙ СЕТЬЮ

М.В. Меркелов

merkelovmv@bmstu.ru

Д.А. Локтев

loktevdan@bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Рассмотрены влияние модуля имитации сетчатки глаза человека на возможность распознавания образов нейронной сетью, существующие программные модели сетчатки и их применение в задаче распознавания образов. Разработан модуль имитации сетчатки глаза человека. Модуль состоит из трех слоев. Первый слой имитирует рецепторы сетчатки глаза человека, предложены варианты расположения на этом слое программных рецепторов. Второй и третий слои модуля имитируют слои биполярных и ганглиозных клеток сетчатки. Для этих слоев разработан вариант организации рецептивных полей нейронов. Выполнена оценка качества работы модуля сетчатки глаза человека на задаче распознавания образов. Тестирование проведено на полносвязной нейронной сети. Создано две модели с одинаковыми параметрами. Одна обучалась на выборке изображений фруктов, сфотографированных с разных ракурсов, в оттенках серого, другая — на той же выборке, но предварительно обработанной с помощью модуля имитации сетчатки глаза человека. Приведены графики обучения нейронной сети на обеих выборках, результаты распознавания образов нейронной сетью и рекомендации по использованию описанного модуля

Ключевые слова

Распознавание образов, классификация, нейронные сети, сетчатка глаза, полносвязные нейронные сети, обучение с учителем

Поступила 07.04.2025

Принята 15.09.2025

© Автор(ы), 2025

Введение. Задача распознавания образов является основной во многих отраслях человеческой деятельности, например, в системах безопасности, медицине, геологии [1, 2]. В связи с этим появляется необходимость в надежной системе распознавания объектов.

Повторение физиологического устройства и работы мозга привело к зарождению искусственных нейронных сетей и является одним из перспективных на данный момент направлений исследований в области искусственного интеллекта.

Модель зрительной коры головного мозга разработана Кунихико Фукушимой в 1979 г. в виде неокогнитрона [3], затем переработана Яном Лекуном в 1988 г. [4]. В результате получена сверточная нейронная сеть, используемая в настоящее время для решения множества задач.

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 зрительную систему, состоящую из таких основных элементов, как сетчатка, зрительный нерв и зрительная кора. Исследование и анализ подобной системы человека, визуализирующей окружающую среду, до сих пор актуальны и могут привести к появлению новых методов и алгоритмов детектирования и распознавания объектов, определения параметров объектов, построения виртуальной/дополненной реальности и т. д. В связи с этим в настоящей работе предложено разработать модуль имитации сетчатки глаза человека.

Модель зрительной системы человека в дальнейшем может быть использована во многих областях человеческой деятельности, например, в компьютерном зрении при проектировании имплантата глаза или для улучшения работы программного обеспечения цифрового фотоаппарата.

Основной задачей настоящей работы является разработка модуля имитации сетчатки глаза человека на основе ее физиологического устройства. *Цель настоящей работы* — оценка влияния разработанного модуля на точность распознавания образов нейронной сетью.

Обзор существующих моделей сетчатки глаза. Проведен поиск существующих моделей сетчатки и работ, в которых описано использование подобных моделей в качестве модулей для предварительной обработки изображений в целях повышения качества распознавания образов.

Модели, реализующие процесс выделения контуров в зрительной системе, рассмотрены в [5], там же приведена бионическая модель зрительной системы человека с неоднородной сетчаткой, разработанная с помощью метода внедрения системы неоднородная сетчатка–однородная кора в системе технического зрения.

Оценка амакриновых клеток с помощью модели, созданной на основе снимка сетчатки макаки, приведена в [6]. Модель предсказывает, что усиление красно-зеленой оппонентности является важной функцией амак-

риновых клеток в пути, обеспечивающем входные данные карликовым ганглиозным клеткам в сетчатке приматов.

Архитектурные аспекты организации сетчатки глазного яблока рассмотрены в [7–9]. Приведены варианты структурной и функциональной организации сетчатки. Показано возможное расположение двух групп рецепторов сетчатки (колбочек и палочек) на плоскости прямоугольной формы для систем компьютерного зрения, но колбочка рассматривается как рецептор только одного вида, т. е. не описано взаимное расположение видов колбочек, реагирующих на разные длины волн.

Клеточная модель сетчатки приматов разработана в [10]. Рассмотрены колбочки только двух типов, они распределялись в сетчатке случайным образом, палочки не моделировались совсем.

Задача компьютерного моделирования методов обработки изображений, основанных на принципах работы зрительной системы человека, рассмотрена в [11]. Созданная компьютерная модель имитирует процессы обработки изображений глазом, но делает это с помощью математических методов, а не биологического устройства глаза.

Симулятор сетчатки Virtual Retina приведен в [12]. Этот симулятор преобразует видеоизображение в кадры, имитирующие изображение реального глаза человека. Его нельзя применить в задаче распознавания образов, так как симулятор, имитируя сферическую форму глаза, удаляет края изображения, на которых мог находиться объект, распознаваемый системой компьютерного зрения.

В приведенных работах 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 не представлено.

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

Первый этап работы над модулем — проектирование слоя фоторецепторов. В реальной сетчатке он содержит два типа клеток: колбочки и палочки. Колбочки можно разделить на три группы по длине волны поглощаемого излучения [13, 14]: 1) коротковолновые *S* (380...500 нм, пик наблюдается на излучении из синей области спектра); 2) средневолновые *M* (450...630 нм, пик — на излучении из зеленой области спектра); 3) длинноволновые *L* (500...700 нм, пик — на излучении из красной области спектра).

Палочки поглощают излучение с длиной волны 400...600 нм и более чувствительны к свету, чем колбочки [15, 16].

Для модуля интенсивность пикселей входного изображения является аналогом потоков пучков света, поступающих на рецепторы реального глаза. В связи с этим каждый пиксель должен соответствовать каждой «клетке» первого слоя программной сетчатки.

Программные колбочки должны реагировать только на конкретный цвет, поэтому они должны выбрать 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.

Фоторецепторы реальной сетчатки плотно расположены друг к другу и образуют гексагональную (шестигранную) решетку [13, 17]. Мозаика расположения колбочек и палочек фиксированная, но рецепторы распределены по сетчатке неравномерно [17–19].

Так, в работах [20, 21] выявлено, что соотношение между колбочками *L*- и *M*-типов сильно различается у разных людей с нормальным зрением (например, приведено, что у двух мужчин соотношение между колбочками *L*- и *M*-типов составляет 75,8 % к 20 % и 50,6 % к 44,2 %).

В рассмотренных моделях сетчатки моделировался слой рецепторов по координатам колбочек и палочек с реальных снимков сетчатки (например, в [6] взят участок сетчатки макаки). Наложить участки сетчатки на изображение невозможно, так как они различаются по форме: реальная сетчатка имеет сферическую форму, а изображение — прямоугольную. Брать конкретный участок сетчатки прямоугольной формы тоже нельзя, поскольку, как отмечено ранее, рецепторы распределены неравномерно. Для разрабатываемого модуля необходимо определить закон, который будет позволять моделировать распределение рецепторов на реальной сетчатке, но даст возможность обрабатывать изображения.

Предлагаемая программная модель сетчатки глаза. За основу взяты матрицы распределения пикселей на фотоматрице фотоаппарата, называемой массивом цветных светофильтров, так как они прямоугольной формы.

Самым популярным светофильтром является фильтр Байера, разработанный в компании Kodak Б.Э. Байером в 1976 г. Этот фильтр размером 2×2 пикселя. Его левый верхний пиксель имеет зеленый цвет, правый верхний — красный, нижний левый — синий, а нижний правый — зеленый.

Большинство современных светофильтров являются видоизмененным вариантом оригинального фильтра Байера. Самым подходящим для программной сетчатки ввиду большой концентрации зеленого цвета и его

размера 6×6 пикселей, является фильтр X-Trans, разработанный компанией Fujifilm в 2012 г.

Оба фильтра в оригинальном виде не подходят для программной сетчатки, так как не имеют пикселей, считывающих интенсивность света по всем цветовым каналам сразу. Поэтому их следует видоизменить для возможности имитировать палочки на сетчатке.

На основе рассмотренных фильтров разработано семь вариантов программного слоя рецепторов, моделирующих расположение колбочек и палочек на первом слое сетчатки. Разработанные цветные фильтры для моделирования распределения фоторецепторов на сетчатке приведены в табл. 1.

Таблица 1

Разработанные цветные фильтры для моделирования распределения фоторецепторов на сетчатке

Номер сетчатки	Мозаика расположения программных рецепторов	Описание	Размер элемента, пиксели
1		Видоизмененный фильтр X-Trans, добавлены белые пиксели, выступающие аналогом палочек на сетчатке	6×6
2		Фильтр Байера, окруженный палочками	4×4
3		Видоизмененный фильтр «Сетчатка № 2», увеличено число зеленых пикселей	4×4
4			
5			
6		Видоизмененный фильтр X-Trans, добавлены белые пиксели, выступающие аналогом палочек на сетчатке	6×6
7		Видоизмененный фильтр «Сетчатка № 4»	4×4

Из предложенных фильтров мозаика расположения программных рецепторов после применения фильтра к изображению получается близкой к гексагональной у фильтров «Сетчатка № 1» и «Сетчатка № 6».

Все рецепторы реальной сетчатки поглощают в определенной степени излучение в зеленой области спектра, вследствие чего человеческий глаз максимально чувствителен к нему [22]. В связи с этим в модуле будет использован фильтр «Сетчатка № 6», так как в нем доля зеленого цвета выше, чем у фильтра «Сетчатка № 1». Результат применения фильтра к изображению приведен на рис. 1.

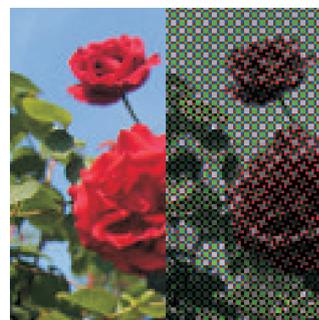


Рис. 1. Результат применения фильтра «Сетчатка № 6» к изображению

Второй этап работы над модулем — разработка слоев биполярных и ганглиозных клеток. Биполярные и ганглиозные клетки бывают с двумя видами организации рецептивных полей: *on*- и *off*-типа. Центр рецептивного поля *on*-типа возбуждает нейрон, а периферия тормозит, у рецептивного поля *off*-типа — центр тормозит нейрон, а периферия возбуждает [23].

Для модуля разработано рецептивное поле клеток, имитирующее работу полей реальных клеток сетчатки. Рецептивные поля биполярных и ганглиозных клеток реальной и реализованной в программном виде сетчатки приведены на рис. 2.

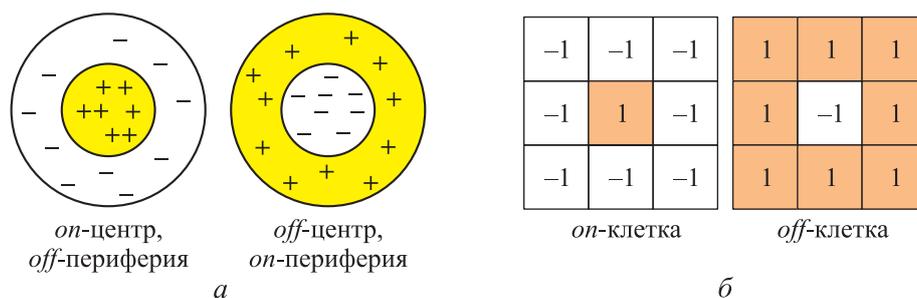


Рис. 2. Рецептивные поля биполярных и ганглиозных клеток реальной (а) и реализованной в программном виде (б) сетчатки

При такой организации выход программного биполярного нейрона *on*-типа определяется по формуле

$$bip_{on} = v(x_{stim}, y_{stim}) + \sum_{k=0}^K (-1)^k v(x_{inhib_k}, y_{inhib_k}), \quad (1)$$

где $v(x_{stim}, y_{stim})$ — значение яркости пикселя изображения, находящегося в центре рецептивной области биполярной клетки (оранжевая область *on*-клетки, см. рис. 2); K — число пикселей изображения, находящихся на периферии рецептивного поля клетки; $v(x_{inhib_k}, y_{inhib_k})$ — значение яркости k -го пикселя изображения, находящегося на краю области биполярной клетки (белая область *on*-клетки, см. рис. 2). Тогда выход программного биполярного нейрона *off*-типа будет вычисляться так:

$$bip_{off} = (-1)^* v(x_{inhib}, y_{inhib}) + \sum_{k=0}^K v(x_{stim_k}, y_{stim_k}), \quad (2)$$

где $v(x_{inhib}, y_{inhib})$ — значение яркости пикселя изображения, находящегося в центре рецептивной области биполярной клетки (белая область *off*-клетки, см. рис. 2); $v(x_{stim_k}, y_{stim_k})$ — значение яркости k -го пикселя изображения, находящегося на краю области биполярной клетки (оранжевая область *off*-клетки, см. рис. 2).

В такой реализации рецептивные поля *on*- и *off*-типов в слое чередуются. По четным индексам расположены *off*-клетки, по нечетным — *on*-клетки.

Для слоя ганглиозных клеток используются такая же организация рецептивных полей *on*- и *off*-типов и такое же расположение *on*- и *off*-клеток в слое. Следовательно, формулы для вычисления их выхода аналогичны формулам (1) и (2).

Разные виды рецептивных полей в сетчатке глаза человека обеспечены амакриновыми и горизонтальными клетками. В программном модуле достаточно реализовать их в качестве обычных синапсов, соединяющих нейрон с его соседями. Имитация сетчатки глаза человека и структура программного модуля сетчатки глаза приведены на рис. 3.

Результаты. Для проверки модуля использовалась выборка, состоящая из фотографий фруктов с разных ракурсов (размер 100×100 пикселей). Представлено 39 классов фруктов. В обучающей выборке содержится 18 505 изображений, в тестовой выборке — 6215. Из оригинальной выборки создано две, на которых в дальнейшем проводилось обучение и тестирование нейронной сети. Первая выборка состояла из фотографий фруктов, переведенных в оттенки серого. Вторая выборка получена путем пропуска изображений из оригинальной выборки через модуль имитации сетчатки. Полученные выборки приведены на рис. 4.

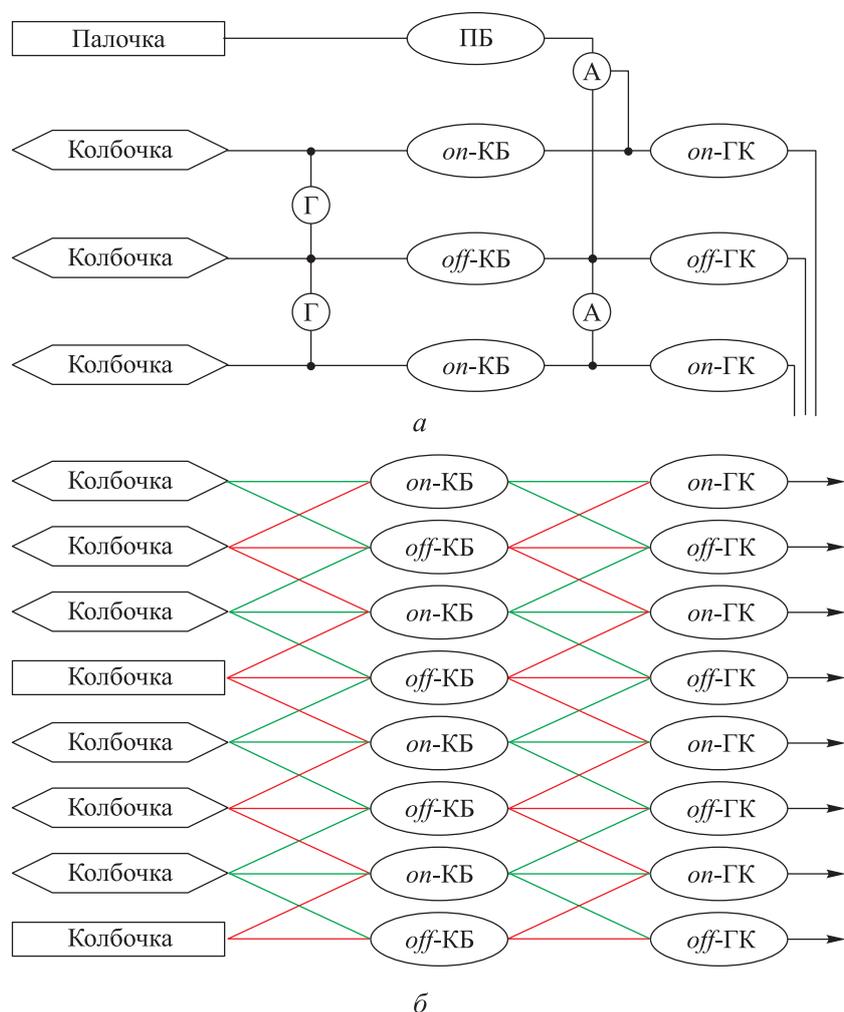


Рис. 3. Имитация сетчатки глаза человека (а), структура программного модуля сетчатки глаза (б):

А, Г, ГК — амакриновая, горизонтальная и ганглиозная клетки;
 КБ, ПБ — колбочковый и палочковый биполярные нейроны

Для распознавания образов создана полносвязная нейронная сеть на языке Python с помощью библиотеки Keras. Полученная нейронная сеть состоит из входного слоя, принимающего изображения размером 100×100 пикселей, и четырех полносвязных слоев. Последний слой имеет 39 нейронов, что соответствует числу классов объектов в выборке. Архитектура полносвязной нейронной сети приведена на рис. 5.

Методом оптимизации при обучении выбран стохастический градиентный спуск с моментами Нестерова со скоростью обучения 0,01. В качестве функции потерь установлена категориальная кросс-энтропия. Размер

пакета составил 32 изображения, число эпох 10. Из обучающей выборки 20 % выделено на валидационную выборку, нейронная сеть в результате обучалась на 14 804 изображениях. Размер валидационной выборки составил 3701 изображение.

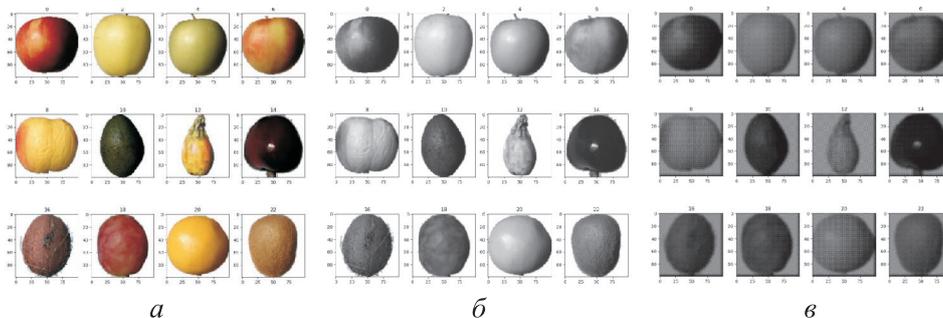


Рис. 4. Оригинальная (а), в оттенках серого (б), после преобразования модулем (в) выборки

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 10000)	0
dense (Dense)	(None, 512)	5120512
dense_1 (Dense)	(None, 256)	131328
dense_2 (Dense)	(None, 128)	32896
dense_3 (Dense)	(None, 39)	5031
Total params: 5,289,767		
Trainable params: 5,289,767		
Non-trainable params: 0		

None

Рис. 5. Скриншот архитектуры используемой полносвязной нейронной сети

Создано два варианта сети с одинаковыми настройками. Одна сеть обучалась исключительно на изображениях в оттенках серого, другая — на изображениях, пропущенных через модуль. Графики обучения нейронных сетей на 10 эпохах приведены на рис. 6.

Точность распознавания фруктов нейронной сетью, обученной на фотографиях в оттенках серого, на тестовой выборке составила 79 %, обученной на обработанных модулем изображениях — 89 %, т. е. модуль в этом случае позволил нейронной сети распознать на 10 % больше фруктов. Поскольку начальные веса связей между нейронами в нейронной сети выбираются случайно, каждое обучение приводит к различным результатам.

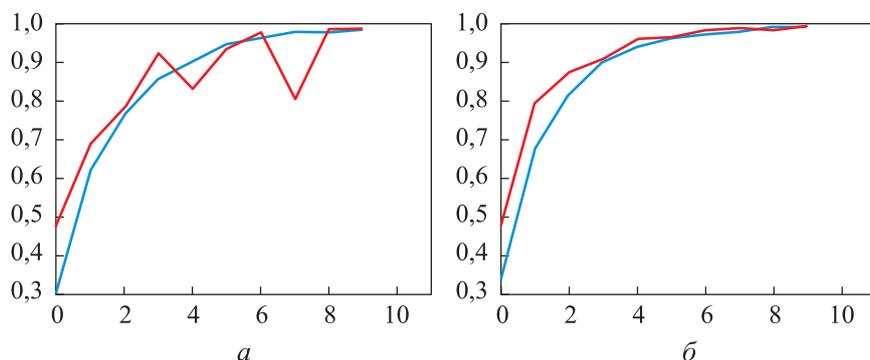


Рис. 6. Графики обучения (на 10 эпохах) нейронных сетей для обучающей (—) и валидационной (—) выборок без предварительной (а) и с предварительной (б) обработками

Для каждой сети выполнено 100 запусков. При каждом запуске происходила очистка сетей, чтобы предыдущие результаты обучения не влияли на последующие запуски. Перед обучением на вход сетям подавалась каждый раз заново перемешанная случайным образом выборка. На каждом запуске первая и вторая выборки различались по расположению изображений, т. е. если в первой выборке на первой позиции лежало яблоко, то во второй выборке на этой же позиции во время этого запуска мог лежать другой фрукт.

Результаты распознавания образов обученными сетями двух типов приведены на рис. 7.

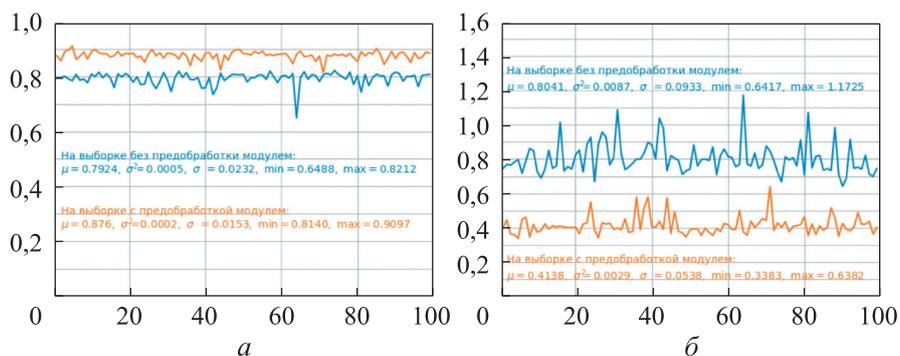


Рис. 7. Скриншот результатов распознавания образов обученными нейронными сетями двух типов на выборках без предварительной (—) и с предварительной (—) обработкой модулем

Выполнено сравнение качества распознавания модулем имитации сетчатки с другими методами предварительной обработки изображения. Результаты сравнения качества распознавания изображений приведены ниже.

Средняя точность распознавания на 10 запусках, %

Метод предварительной обработки изображений:	
оттенки серого	78
модуль имитации сетчатки	87
Фильтры:	
LoG	80
DoG	80
Робертса	71
MSRCR	76

Проведено несколько экспериментов на выборке CIFAR-10, которая представлена 10 классами объектов и содержит 60 000 изображений животных и транспортных средств. В выборке CIFAR-10 находится 50 000 изображений в обучающей выборке и 10 000 изображений в тестовой выборке.

В пакете `tonic` языка Python доступен вариант этой выборки, полученный с помощью DVS-камеры. Этот вариант выборки описывает тот же набор классов объектов, но содержит только 10 000 изображений. Изображения, полученные с таких камер, тоже можно считать вариантом предварительной обработки, поэтому в настоящей работе проводилось два варианта экспериментов на выборке CIFAR-10.

В первом варианте экспериментов рассматривались все методы предварительной обработки. Нейронная сеть обучалась на полном наборе изображений из выборки CIFAR-10. Из обучающей выборки 20 % выделено на валидационную выборку, поэтому нейронная сеть в результате обучалась на 40 000 изображений. Размер валидационной выборки составил 10 000 изображений.

Во втором варианте экспериментов, кроме предварительной обработки, рассматривалась предварительная обработка с помощью DVS-камеры. В этом варианте экспериментов все методы предварительной обработки использовали уменьшенную по размеру до 10 000 изображений выборку CIFAR-10. Из этого набора 20 % выделено на тестовую выборку и 20 % на валидационную выборку, в результате нейронная сеть обучалась на 6000 изображений. Размер валидационной и тестовой выборок составил 2000 изображений. Для случая с DVS-камерой взята выборка из пакета `tonic` и преобразована в изображения методом `tonic.transforms.ToImage`. Для остальных методов обработки из оригинальной выборки в качестве обучающей взята тестовая выборка.

В обоих экспериментах изображения выборки CIFAR-10 интерполировались до размера 100×100 пикселей методом `transform.resize` пакета

skimage языка Python, так как на оригинальных размерах изображений 32×32 пикселя модуль имитации сетчатки сильно искажает изображения.

В обоих экспериментах в архитектуре нейронной сети (см. рис. 5) изменено число нейронов в последнем выходном слое, с 39 на 10, так как CIFAR-10 содержит 10 классов, а не 39, как было ранее в выборке с фруктами. Остальные настройки сети не менялись.

Результаты распознавания образов обученными нейронными сетями на выборке CIFAR-10 с разными вариантами предварительной обработки изображений приведены в табл. 2.

Таблица 2

Результаты распознавания образов обученными нейронными сетями на выборке CIFAR-10 с разными вариантами предварительной обработки изображений

Размер выборки (число изображений)	Метод предварительной обработки изображений						
	Оттенки серого	Модуль имитации сетчатки	Фильтр				DVS-камера
			LoG	DoG	Робертса	MSRCR	
60 000	38	44	42	42	37	39	–
10 000	30	35	33	32	26	30	9

Проведены такие же варианты экспериментов на сверточной архитектуре нейронной сети. Созданная архитектура имеет семь сверточных слоев, на вход нейронная сеть принимает изображения размером 100×100 пикселей. Последний полносвязный слой имеет 10 нейронов, что соответствует числу классов объектов в выборке CIFAR-10. Полная архитектура используемой полносвязной нейронной сети приведена на рис. 8.

Методом оптимизации при обучении выбран стохастический градиентный спуск с моментами Нестерова 0,9 со скоростью обучения 0,01. В качестве функции потерь установлена категориальная кросс-энтропия. Размер пакета составил 32 изображения, число эпох 15.

Результаты распознавания образов обученными сверточными нейронными сетями на выборке CIFAR-10 с разными вариантами предварительной обработки изображений приведены в табл. 3.

Обсуждение полученных результатов. За счет разработанного модуля каждый цвет на изображении выделяется собственным узором, что позволяет нейронной сети запоминать больше информации об объекте на изображении (рис. 9).

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)             (None, 96, 96, 96)         2496
conv2d_1 (Conv2D)           (None, 96, 96, 96)         9312
max_pooling2d (MaxPooling2D) (None, 47, 47, 96)         0
conv2d_2 (Conv2D)           (None, 43, 43, 192)        460992
conv2d_3 (Conv2D)           (None, 43, 43, 192)        37056
max_pooling2d_1 (MaxPooling2D) (None, 21, 21, 192)        0
conv2d_4 (Conv2D)           (None, 19, 19, 192)        331968
conv2d_5 (Conv2D)           (None, 19, 19, 192)        37056
conv2d_6 (Conv2D)           (None, 19, 19, 10)         1930
global_average_pooling2d (GlobalAveragePooling2D) (None, 10)                 0
dense (Dense)                (None, 10)                 110
-----
Total params: 880,920
Trainable params: 880,920
Non-trainable params: 0
    
```

Рис. 8. Скриншот полной архитектуры используемой полносвязной нейронной сети

Таблица 3

Результаты распознавания образов обученными сверточными нейронными сетями на выборке CIFAR-10 с разными вариантами предварительной обработки изображений

Размер выборки (число изображений)	Метод предварительной обработки изображений						
	Оттенки серого	Модуль имитации сетчатки	Фильтр				DVS-камера
			LoG	DoG	Робертса	MSRCR	
60 000	69	66	63	62	10	10	–
10 000	56	56	52	53	9	9	9

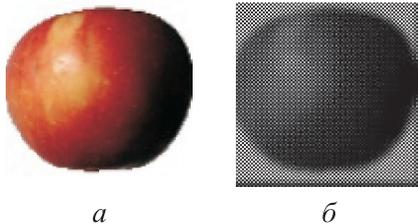


Рис. 9. Изображение яблока (а), результат применения к нему модуля (б)

При изменении фильтра, имитирующего расположение колбочек, палочек, рецептивных полей биполярных и ганглиозных клеток, узоры будут получаться иными, что может привести к другим результатам обучения и тестирования.

В процессе исследования установлено, что изображения с низким разрешением сильно искажаются модулем, и из-за этого качество распознавания сильно падает. В связи с этим рекомендуется использовать изображения с разрешением 50×50 пикселей и более. Ввиду таких размеров изображений нейронной сети потребуется больше времени на обучение. Искажение изображения после использования модуля для различных разрешений приведено на рис. 10.

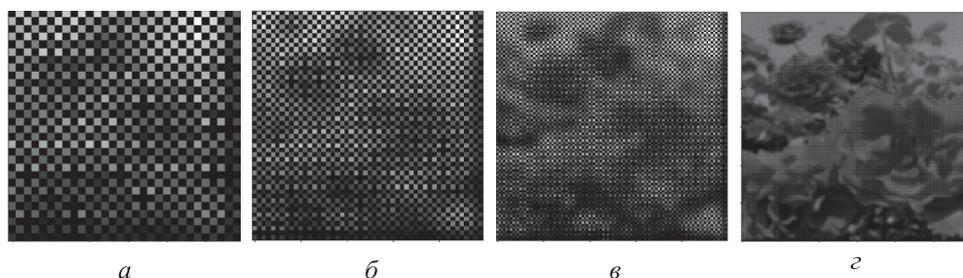


Рис. 10. Искажение изображения после использования модуля для разрешений 30×30 (а), 50×50 (б), 100×100 (в), 300×300 (г) пикселей

В результате 100 запусков выборок с фруктами выявлено, 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 рис. 6: модель, обучающаяся на обработанных изображениях, имеет менее ломаный график (валидационная выборка). Отметим, что на выборках с фруктами с обработанными изображениями дисперсия и среднеквадратическое отклонение ниже, чем на выборках без обработки изображения, что подтверждает тезис о более плавном и стабильном обучении.

Значение функции потерь также меньше на выборке с фруктами, обработанной модулем, на 100 запусках.

Согласно результатам сравнения (см. табл. 2), на полносвязной нейронной сети за счет предварительной обработки модулем точность распознавания образов выше среди других методов предварительной обработки изображений на разных вариантах выборок.

По результатам сравнения (см. табл. 3) на сверточной нейронной сети, модуль имитации сетчатки при выбранных вариантах рецептивных полей

и мозаики расположения программных рецепторов не дает эффекта на данной архитектуре, но выше по точности других методов предварительной обработки. Необходимо протестировать модуль со сверточными нейронными сетями при больших размерах изображений в выборке и при других вариантах рецептивных полей и мозаик расположения программных рецепторов.

Заключение. Проанализировано влияние разработанного модуля на результат распознавания образов нейронной сетью. Показано, что предварительная обработка с помощью модуля имитации сетчатки позволяет повысить качество распознавания образов полносвязной нейронной сетью и в целом стабилизировать ее обучение. На выборке с предварительной обработкой модулем имитации сетчатки полносвязная нейронная сеть лучше адаптируется под новые образы.

Установлено, что модуль сильно искажает изображения с низким разрешением. В связи с этим разработанную сетчатку рекомендуется использовать на разрешениях 50×50 пикселей и более.

Определение фильтров, рецептивных полей клеток и подбор архитектур нейронных сетей, с которыми модуль будет позволять достигать наилучшего результата, является предметом дальнейших исследований.

ЛИТЕРАТУРА

- [1] Шульга Т.Э. Актуальные задачи компьютерного зрения. *Инновационные исследования: теоретические основы и практическое применение. Сб. статей по итогам Всерос. науч.-практ. конф.* Уфа, АМИ, 2021, с. 88–92. EDN: GUFFZW
- [2] Адамова А.А. Методы и технологии машинного обучения и нейросетевых технологий в задачах компьютерного зрения. *Нейрокомпьютеры: разработка, применение*, 2021, т. 23, № 4, с. 25–39. DOI: <https://doi.org/10.18127/j19998554-202104-03>
- [3] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybernetics*, 1980, vol. 36, no. 4, pp. 193–202. DOI: <https://doi.org/10.1007/bf00344251>
- [4] LeCun Y., Boser B., Denker J.S., et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, 1989, vol. 1, no. 4, pp. 541–551. DOI: <https://doi.org/10.1162/neco.1989.1.4.541>
- [5] Гладилин С.А. Математические модели для систем технического зрения. Дис. ... канд. физ.-мат. наук. М., ИППИ РАН, 2005.
- [6] Lebedev D.S., Marshak D.W. Amacrine cell contributions to red-green color opponency in central primate retina: a model study. *Vis. Neurosc.*, 2007, vol. 24, no. 4, pp. 535–547. DOI: <https://doi.org/10.1017/s0952523807070502>

- [7] Утробин В.А. Модели организации поля рецепторов сетчатки глазного яблока. *Труды НГТУ им. П.Е. Алексеева*, 2010, № 4, с. 63–69. EDN: NUXQZH
- [8] Утробин В.А. Информационные модели организации сетчатки глазного яблока. *Труды НГТУ им. П.Е. Алексеева*, 2010, № 3, с. 66–74. EDN: NVAXJN
- [9] Утробин В.А. Модели организации поля рецепторов сетчатки глазного яблока. *Датчики и системы*, 2013, № 11, с. 37–43. EDN: RIUCVB
- [10] Momiji H., Bharath A.A., Hankins M.W., et al. Numerical study of short-term after-images and associate properties in foveal vision. *Vis. Res.*, 2005, vol. 46, no. 3, pp. 365–381. DOI: <https://doi.org/10.1016/j.visres.2005.08.031>
- [11] Маштаков А.П. Программный комплекс обработки изображений на основе модели зрения человека. *Программные системы: теория и приложения*, 2019, т. 10, № 4, с. 111–139. DOI: <https://doi.org/10.25209/2079-3316-2019-10-4-111-139>
- [12] Wohrer A., Kornprobst P. Virtual Retina: a biological retina model and simulator, with contrast gain control. *J. Comput. Neurosc.*, 2009, vol. 26, no. 2, pp. 219–249. DOI: <https://doi.org/10.1007/s10827-008-0108-4>
- [13] Чупров А.Д., Синькова В.И., Кузнецов И.В. Теории цветовосприятия. Фоторецепторный аппарат сетчатки глаза. *Современные проблемы науки и образования*, 2021, № 6. DOI: <https://doi.org/10.17513/spno.31287>
- [14] Школьник-Яррос Е.Г., Калинина А.В. Нейроны сетчатки. М., Наука, 1986.
- [15] Говардовский В.И. Специфика физиологических и биохимических механизмов возбуждения и адаптации колбочек сетчатки. *Сенсорные системы*, 2015, т. 29, № 4, с. 296–308. EDN: VLAMEB
- [16] Pirenne M.H. *Vision and the eye*. London, Pilot Press, 1948.
- [17] Ahnelt P.K. The photoreceptor mosaic. *Eye*, 1998, vol. 12, no. 3, pp. 531–540. DOI: <https://doi.org/10.1038/eye.1998.142>
- [18] Рожкова Г.И., Белокопытов А.В., Грачева М.А. Загадки слепой зоны и кольца повышенной плотности колбочек на крайней периферии сетчатки. *Сенсорные системы*, 2016, т. 30, № 4, с. 263–281. EDN: XDNLRH
- [19] Jones L.A., Higgins G.C. Photographic granularity and graininess III. Some characteristics of the visual system of importance in the evaluation of graininess and granularity. *J. Opt. Soc. Am.*, 1947, vol. 37, no. 4, pp. 217–263. DOI: <https://doi.org/10.1364/josa.37.000217>
- [20] Roorda A., Williams D.R. The arrangement of the three cone classes in the living human eye. *Nature*, 1999, vol. 397, no. 6719, pp. 520–522. DOI: <https://doi.org/10.1038/17383>
- [21] Алексеенко С.В. Структура сетчатки глаза человека: классические и современные данные. *Сенсорные системы*, 2019, т. 33, № 4, с. 269–286. DOI: <https://doi.org/10.1134/S0235009219040024>
- [22] Judd D., Wyszecki G. *Color in business, science, and industry*. New York, Wiley, 1975.

[23] Измайлов И.А., Соколов Е.Н., Черноризов А.М. Психофизиология цветового зрения. М., Изд-во МГУ, 1989.

Меркелов Михаил Валерьевич — ассистент, аспирант кафедры «Информационные системы и телекоммуникации» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Локтев Даниил Алексеевич — д-р техн. наук, доцент, профессор кафедры «Информационные системы и телекоммуникации» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Просьба ссылаться на эту статью следующим образом:

Меркелов М.В., Локтев Д.А. Влияние модуля имитации сетчатки глаза человека на качество распознавания образов нейронной сетью. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2025, № 4 (153), с. 102–120. EDN: HJIVUH

**THE EFFECT OF THE HUMAN RETINA SIMULATION MODULE
ON THE QUALITY OF IMAGE RECOGNITION
BY A NEURAL NETWORK**

M.V. Merkelov

merkelovmv@bmstu.ru

D.A. Loktev

loktevdan@bmstu.ru

BMSTU, Moscow, Russian Federation

Abstract

This article examines the impact of a human retina simulation module on the neural network's ability to recognize patterns, as well as existing software models of the retina and their application to pattern recognition. A human retina simulation module is being developed. The module consists of three layers. The first layer simulates the receptors of the human retina, and options for arranging software receptors on this layer are proposed. The second and third layers of the module simulate the layers of bipolar and ganglion cells of the retina. An option for organizing the receptive fields of neurons is being developed for these layers. The retina module's performance in pattern recognition is assessed. Testing is conducted on a fully connected neural network. Two models with identical parameters are created. One was trained on a sample of fruit images photographed from different angles in grayscale, and the other on the same sample, but pre-processed by the retina simulation module. Graphs

Keywords

Pattern recognition, classification, neural networks, retina, fully connected neural networks, supervised learning

of neural network training on both samples, the results of pattern recognition by the neural network, and recommendations for using the described module are presented

Received 07.04.2024

Accepted 15.09.2025

© Author(s), 2025

REFERENCES

- [1] Shulga T.E. [Actual tasks of computer vision]. *Innovatsionnye issledovaniya: teoreticheskie osnovy i prakticheskoe primeneniye. Sb. statey po itogam Vseros. nauch.-prakt. konf.* [Innovative Research: Theoretical Foundations and Practical Application. Proc. Russ. Sc.-Pract. Conf.]. Ufa, AMI Publ., 2021, pp. 88–92 (in Russ.). EDN: GUFFZW
- [2] Adamova A.A. Methods and technologies of machine learning in neural network for computer vision purposes. *Neyrokompyutery: razrabotka, primeneniye* [Neurocomputers], 2021, vol. 23, no. 4, pp. 25–39 (in Russ.).
DOI: <https://doi.org/10.18127/j19998554-202104-03>
- [3] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybernetics*, 1980, vol. 36, no. 4, pp. 193–202. DOI: <https://doi.org/10.1007/bf00344251>
- [4] LeCun Y., Boser B., Denker J.S., et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, 1989, vol. 1, no. 4, pp. 541–551.
DOI: <https://doi.org/10.1162/neco.1989.1.4.541>
- [5] Gladilin S.A. Matematicheskie modeli dlya sistem tekhnicheskogo zreniya. Dis. kand. fiz.-mat. nauk [Mathematical models for technical vision systems. Cand. Sc. (Phys.-Math.). Diss.]. Moscow, IPPI RAN, 2005 (in Russ.).
- [6] Lebedev D.S., Marshak D.W. Amacrine cell contributions to red-green color opponency in central primate retina: a model study. *Vis. Neurosc.*, 2007, vol. 24, no. 4, pp. 535–547. DOI: <https://doi.org/10.1017/s0952523807070502>
- [7] Utrobin V.A. Organization models of the eyeball retina field of receptors. *Trudy NGTU im. R.E. Alekseeva*, 2010, no. 4, pp. 63–69 (in Russ.). EDN: NUXQZH
- [8] Utrobin V.A. Information models of the eyeball retina organization. *Trudy NGTU im. R.E. Alekseeva*, 2010, no. 3, pp. 66–74 (in Russ.). EDN: NVAXJN
- [9] Utrobin V.A. Organization models of the eyeball retina field of receptors. *Datchiki i sistemy* [Sensors & Systems], 2013, no. 11, pp. 37–43 (in Russ.). EDN: RIUCVB
- [10] Momiji H., Bharath A.A., Hankins M.W., et al. Numerical study of short-term after-images and associate properties in foveal vision. *Vis. Res.*, 2005, vol. 46, no. 3, pp. 365–381. DOI: <https://doi.org/10.1016/j.visres.2005.08.031>
- [11] Mashtakov A.P. Image processing toolkit inspired by mechanisms of human visual perception. *Programmnye sistemy: teoriya i prilozheniya* [Program Systems: Theory and Applications], 2019, vol. 10, no. 4, pp. 111–139 (in Russ.).
DOI: <https://doi.org/10.25209/2079-3316-2019-10-4-111-139>
- [12] Wohrer A., Kornprobst P. Virtual Retina: a biological retina model and simulator, with contrast gain control. *J. Comput. Neurosc.*, 2009, vol. 26, no. 2, pp. 219–249.
DOI: <https://doi.org/10.1007/s10827-008-0108-4>

- [13] Chuprov A.D., Sinkova V.I., Kuznetsov I.V. Color perception theories. photoreceptors structure of eye retina. *Sovremennyye problemy nauki i obrazovaniya* [Modern Problems of Science and Education], 2021, no. 6 (in Russ.).
DOI: <https://doi.org/10.17513/spno.31287>
- [14] Shkolnik-Yarros E.G., Kalinina A.V. Neyrony setchatki [Retinal neurons]. Moscow, Nauka Publ., 1986.
- [15] Govardovskiy V.I. Specifics of physiological and biochemical mechanisms of excitation and adaptation in retinal cones. *Sensornyye sistemy* [Sensory Systems], 2015, vol. 29, no. 4, pp. 296–308 (in Russ.). EDN: VLAMEB
- [16] Pirenne M.H. Vision and the eye. London, Pilot Press, 1948.
- [17] Ahnelt P.K. The photoreceptor mosaic. *Eye*, 1998, vol. 12, no. 3, pp. 531–540.
DOI: <https://doi.org/10.1038/eye.1998.142>
- [18] Rozhkova G.I., Belokopytov A.V., Gracheva M.A. Mysteries of the blind zone and cone-enriched rim at the extreme periphery of the human retina. *Sensornyye sistemy* [Sensory Systems], 2016, vol. 30, no. 4, pp. 263–281 (in Russ.). EDN: XDNLRH
- [19] Jones L.A., Higgins G.C. Photographic granularity and graininess III. Some characteristics of the visual system of importance in the evaluation of graininess and granularity. *J. Opt. Soc. Am.*, 1947, vol. 37, no. 4, pp. 217–263.
DOI: <https://doi.org/10.1364/josa.37.000217>
- [20] Roorda A., Williams D.R. The arrangement of the three cone classes in the living human eye. *Nature*, 1999, vol. 397, no. 6719, pp. 520–522.
DOI: <https://doi.org/10.1038/17383>
- [21] Alekseenko S.V. The structure of the human retina: classical and modern data. *Sensornyye sistemy* [Sensory Systems], 2019, vol. 33, no. 4, pp. 269–286 (in Russ.).
DOI: <https://doi.org/10.1134/S0235009219040024>
- [22] Judd D., Wyszecki G. Color in business, science, and industry. New York, Wiley, 1975.
- [23] Izmaylov I.A., Sokolov E.N., Chernorizov A.M. Psikhofiziologiya tsvetovogo zreniya [Psychophysiology of color vision]. Moscow, MSU Publ., 1989.

Merkelov M.V. — Assistant, Post-Graduate Student, Department of Information Systems and Telecommunications, BMSTU (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Loktev D.A. — Dr. Sc. (Eng.), Assoc. Professor, Department of Information Systems and Telecommunications, BMSTU (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Please cite this article in English as:

Merkelov M.V., Loktev D.A. The effect of the human retina simulation module on the quality of image recognition by a neural network. *Herald of the Bauman Moscow State Technical University, Series Instrument Engineering*, 2025, no. 4 (153), pp. 102–120 (in Russ.). EDN: HJIVUH