

УДК 004.8:004.9

КЛАССИФИКАЦИЯ ГЛАЗНЫХ ЗАБОЛЕВАНИЙ ПО СНИМКАМ ГЛАЗНОГО ДНА С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ EFFICIENTNET



К.И. Свистунова
Студентка кафедры
информационных систем
управления ФПМИ БГУ
kseniyasvistunova@icloud.com



С.В. Абламейко
Профессор БГУ
Доктор технических наук,
Академик НАН Беларуси,
Гл. научный сотрудник ОИПИ
ablameyko@bsu.by

К.И. Свистунова

Является студенткой четвертого курса Белорусского государственного университета. Область научных интересов связана с решением практических задач в области машинного обучения.

С.В. Абламейко

Является профессором БГУ, гл. научным сотрудником ОИПИ. Область научных интересов связана с разработкой методов и алгоритмов для решения задач обработки изображений и распознавания образов.

Аннотация. Предложен алгоритм распознавания заболеваний глаз человека по снимкам глазного дна. В результате проведенного анализа была выбрана сверточная нейронная сеть EfficientNet и на ее основе разработан алгоритм, использующий архитектуры EfficientNetB6 и EfficientNetB0. Проведены экспериментальные исследования, которые показали высокую эффективность данного подхода для диагностики заболеваний глаз.

Ключевые слова: нейронные сети, EfficientNet, аугментация, анализ изображений, распознавание образов, классификация, изображения глазного дна.

Введение. Практически каждый человек в своей жизни сталкивается с проблемой болезни глаз – это любое состояние или расстройство, которое нарушает способность глаза правильно работать или оказывает вредное воздействие на остроту зрения. Болезни глазного дна являются основной причиной слепоты у людей, поэтому своевременная диагностика на начальных стадиях крайне важна для их лечения, так как позволяет предотвратить дальнейшее прогрессирование заболеваний.

Ручная диагностика глазных болезней по снимкам глазного дна на ранних этапах может представлять собой значительную трудность. В связи с этим целесообразно разработать автоматизированную систему для обнаружения заболеваний глаз с помощью компьютерных технологий, которая будет способна выявлять различные нарушения, основываясь на изображениях глазного дна.

Алгоритмы, основанные на глубоком обучении, становятся все более распространенными в анализе медицинских изображений. Исследования показали, что такие модели эффективно справляются с задачами, связанными с обнаружением объектов, классификацией медицинских изображений и выявлением заболеваний.

К настоящему времени опубликовано достаточно много работ по анализу глазных (офтальмологических) изображений. В [2] предложен метод автоматизации анализа

флюоресцентных ангиограмм глазного дна человека, а в [3] – методы для оценки состояния липидного слоя в интермаргинальном пространстве века глаза человека, анализа степени плотности клеточными структурами (целлюлярности) роговичной ткани глаза человека, оценки состояния ретинального кровотока при анализе флюоресцентных ангиограмм глазного дна человека.

В докладе приведен результат анализа эффективности нейронной сети EfficientNet для распознавания различных заболеваний глаз.

Выбор сверточной нейронной сети. EfficientNet - это семейство сверточных нейронных сетей (CNN), целью которого является достижение высокой производительности с меньшими вычислительными ресурсами по сравнению с предыдущими архитектурами. Оно было представлено Mingxing Tan и Quoc V. Le из Google Research в своей статье 2019 года «EfficientNet: переосмысление масштабирования модели для сверточных нейронных сетей». Основной идеей EfficientNet является новый метод масштабирования, называемый составным коэффициентом, для масштабирования моделей простым, но эффективным способом. Вместо случайного масштабирования ширины, глубины или разрешения, составное масштабирование равномерно масштабирует каждое измерение с определенным фиксированным набором коэффициентов масштабирования. Этот подход гарантирует, что модель легко адаптируется к различным вычислительным ограничениям, сохраняя при этом свою производительность в различных масштабах и задачах. Используя данный метод масштабирования, авторы разработали семь моделей различных размеров, которые превосходили самую современную точность большинства сверточных нейронных сетей и с гораздо лучшей эффективностью [8].

Исходный набор данных. Выбранный датасет представляет собой набор данных, доступный на Kaggle, для выявления глазных заболеваний. Он содержит снимки глаз более пяти тысяч пациентов. Набор данных представляет собой цветные снимки глазного дна левого и правого глаза, которые размечены ключевыми диагнозами, поставленными врачами. Врачи классифицируют болезни глаз по восьми категориям: нормальный(здоровый) (Н), диабет (Д), глаукома (Г), катаракта (С), гипертония (Н), патологическая близорукость (М), возрастная макулярная дегенерация (А), другие заболевания/аномалии (О).

Работа с данными. В поставленной задаче было принято решение рассматривать пары классов (здоровые и заболевание) для выявления классификации заболевания.

Для начала было рассмотрено количество снимков, соответствующие каждому классу, для анализа того, какую подготовку данных следует совершить для решения задачи. В таблице 1 представлено количество снимков, содержащих определенное заболевание.

В таблице 1 показано, что при рассмотрении любого класса заболевания в сравнении со здоровым, у нас будет несбалансированный набор данных. Для решения этой проблемы было решено применять аугментацию на лету. Данный метод был использован для того, чтобы каждый раз на обучение поступали изображения, которые модель еще не видела. Это помогает улучшить обобщающую способность модели, снизить вероятность переобучения и делает её более устойчивой к изменениям в данных. Изучив изображения, на которых модель ошибается, было принято решение использовать способы аугментации, такие как поворот изображений на случайный угол, смещение изображения по ширине, смещение изображения по высоте. Для заполнения новых пикселей, которые могут возникнуть в результате трансформаций (например, при повороте или изменении размера) используется "nearest" (ближайший сосед). Это означает, что новые пиксели будут заполнены значениями ближайших пикселей.

Таблица 1. Количество снимков согласно заболеваниям

№ п/п	Заболевание	Количество
1	Здоровый	5501
2	Катаракта	594
3	Диабетическая ретинопатия	165
4	Миопия	457
5	Гипертензия	382
6	Возрастная макулярная дегенерация	551
7	Глаукома	616

Обучение нейронной сети. Для решения задачи были выбраны EfficientNetB6 и EfficientNetB0 для выявления того, как изменение размера изображения повлияет на точность определения глаза с заболеванием. Для изображений размера 224x224 применим EfficientNetB0, для изображений 512x512 (исходный размер изображений) - EfficientNetB6.

Данные из набора преобразуем в массивы изображений и меток. Разделяем их на обучающую и тестовую выборки, где 20 % данных выделяем для тестирования. На следующем этапе создаем генератор данных для выполнения аугментации, что помогает улучшить обобщающую способность модели за счет разнообразия обучающих данных.

Создаем экземпляры моделей, использующих предварительно обученные веса на наборе данных ImageNet. Указание параметра `include_top=False` исключает верхние полностью связанные слои, `input_shape=(image_size, image_size, 3)` задает форму входных изображений. Создаем последовательную модель с помощью Sequential из TensorFlow, которая позволяет добавлять слои последовательно друг за другом. Добавление предварительно обученной модели к последовательной модели позволяет использовать архитектуру и веса модели как часть нашей конечной модели. К модели EfficientNetB6 добавляем слой GlobalAveragePooling2D, который выполняет глобальное усреднение по каждому из пространственных размеров входных данных. Он уменьшает пространство признаков, создавая вектор признаков фиксированной длины из выходных данных EfficientNetB6, и устраняет избыточность, что может помочь предотвратить переобучение. Также добавляем слой Dense - полносвязный слой с 256 нейронами. Активация ReLU (Rectified Linear Unit) применяется для введения нелинейности в модель, что позволяет ей обучаться более сложным функциям. Добавляем полносвязный слой Dense с одним нейроном и функцией активации sigmoid. В данном случае используется один нейрон с активацией sigmoid, так как задача является бинарной классификацией. Модель компилируется с использованием оптимизатора Adam, функции потерь binary_crossentropy и метрики точности.

Модель обучается с использованием `data_gen.flow(X_train, y_train, batch_size=16)`, который создает генератор данных для подачи изображений и меток в модель по партиям (batch). Это полезно для обработки больших наборов данных, так как позволяет загружать данные по частям в память. Обучение происходит в течение 15 эпох. Для предотвращения переобучения используются механизмы ранней остановки и сохранения лучшей модели.

Результаты классификации. Для оценки эффективности предложенного алгоритма были проведены экспериментальные исследования на наборе данных изображений глазного дна. Нейронная сеть EfficientNet была обучена и протестирована. В таблице 2 представлены результаты классификации для разных типов заболеваний.

Таблица 2. Результаты классификации с помощью EfficientNet для разных типов заболеваний

Заболевание	Точность EfficientNetB0 (224x224)	Точность EfficientNetB6 (512x512, size of images)
Миопия	0.9790576100349426	0.9947643876075745
Катаракта	0.9663865566253662	0.9729729890823364
Глаукома	0.9211618304252625	0.9300000071525574
Возрастная макулярная дегенерация	0.9497717022895813	0.9674418568611145
Диабетическая ретинопатия	0.9305555820465088	0.9166666865348816
Гипертензия	0.9044585824012756	0.9426751732826233

Лучшие результаты показывает сеть EfficientNetB6 на изображениях исходного размера. Это показывает, что сохранение размера изображений в решении данной задачи важно и необходимо, так как происходит меньший процент потери данных. Например, в диагностике возрастной макулярной дегенерации точность EfficientNet значительно увеличилась с 0.9498 (224x224) до 0.9674 (512x512), показывая, что размер изображения важен для повышения точности.

Лучшие результаты наблюдаются в выявлении миопии и катаракты, что может быть связано с большим отличием в изображениях здорового глаза и глаза, имеющего данные заболевания.

Для оценки работы были использованы следующие метрики:

Precision (точность) — это доля объектов, названных классификатором положительными и при этом действительно являющимися положительными. Высокие показатели Precision говорят:

- о высокой чувствительности модели – она хорошо обнаруживает случаи с положительным диагнозом;
- о низком уровне ложноотрицательных результатов.

Recall (полнота) — это то, какую долю объектов положительного класса из всех объектов положительного класса нашёл алгоритм. Высокие показатели Recall говорят:

- о высокой чувствительности модели - она хорошо обнаруживает случаи с положительным диагнозом;
- о низком уровне ложноотрицательных результатов.

F1-score — одна из метрик, используемых для оценки качества бинарных классификаторов, которая учитывает, как точность (Precision), так и полноту (Recall). Эта метрика особенно полезна в случаях, когда важно находить баланс между этими двумя показателями, и, соответственно, когда классы имеют неравное количество примеров или, когда неравномерные ошибки имеют разные последствия. Например, в диагностике заболеваний, пропуск положительного случая (ложный отрицательный) может иметь серьезные последствия. В этом случае высокая полнота более критична, и F1-score позволяет учитывать это.

В Таблице 3 представлены показатели recall, precision и F1-score для каждого класса.

Таблица 3. Оценки работы EfficientNetB6 для различных классов

	Класс	Precision	Recall	F1 score
Здоровый и миопия	0	1.00	0,99	0,99
	1	0,99	1.00	0,99
Здоровый и катаракта	0	0.98	0.96	0.97
	1	0.97	0.98	0.97

Здоровый и диабетическая ретинопатия	0	0.92	0.92	0.92
	1	0.91	0.91	0.91
Здоровый и глаукома	0	0.92	0.94	0.93
	1	0.94	0.92	0.93
Здоровый и ВМД	0	0.97	0.96	0.96
	1	0.97	0.97	0.97
Здоровый и гипертензия	0	0.95	0.94	0.95
	1	0.93	0.94	0.94

Для анализа того, какие изображения были распознаны неверно, была прописана функция вывода изображений, на которых метка классов не совпадает с предсказанной меткой. Результат работы такой функции для класса Диабетической ретинопатии представлен на рисунке 1.

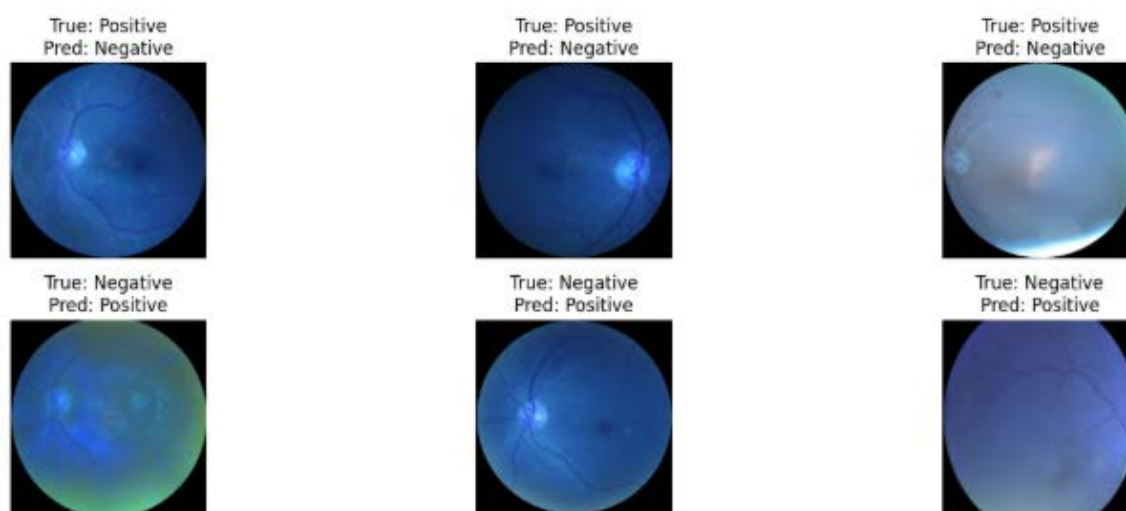


Рисунок 1. Изображения для класса Диабетическая ретинопатия, для которых истинные метки не совпали с предсказанными

Визуальный анализ ошибочно классифицированных изображений может помочь выявить общие паттерны среди неверно классифицированных изображений, такие как нечеткость снимков, присутствие шумов и артефактов. Анализируя результаты, можно понять, какие виды аугментации данных наиболее эффективны для улучшения производительности модели и уменьшения количества ошибок. Результаты визуального анализа могут быть использованы для разработки стратегии дальнейшего улучшения модели, включая создание новых наборов данных, улучшение предобработки и аугментации.

Заключение. В докладе рассмотрена нейронная сеть EfficientNet для анализа изображений глазного дна. Предложен алгоритм распознавания заболеваний глаза человека по снимкам глазного дна. Также приведены результаты экспериментальных исследований.

Список литературы

- [1] Волков, Е. Н. Gradient-based методы объяснительного искусственного интеллекта для классификации заболеваний глаза / Е. Н. Волков, А. Н. Аверкин // IV Международная конференция по нейронным сетям и нейротехнологиям (NeuroNT'2023) : сб. докладов, СПб., 16 июня 2023 г. – СПб. : СПбГЭТУ «ЛЭТИ», 2023. – С. 7–10.
- [2] Gurevich I. B., Yashina V. V., Ablameyko S. V., Nedzved A. M., Ospanov A. M., Tleubaev A. T., Fedorov A. A., Fedoruk N. A. Development and experimental investigation of mathematical methods for automating the

diagnostics and analysis of ophthalmological images [Electronic resource]. – Pattern Recognit Image Anal. – 2018. – № 28(4). – P. 612–636. – Mode of access: <https://doi.org/10.1134/S1054661818040120>. – Date of access: 01.08.2024.

[3] Гуревич И. Б., Яшина В. В., Абламейко С. В., Недзведь А. М., Оспанов А. М., Тлеубаев А. Т., Федорук Н. А. Новый математический метод автоматизации анализа флюоресцентных ангиограмм глазного дна человека [Электронный ресурс]. – Вестник офтальмологии. – 2021. – № 137(3). – P. 49–57. – Режим доступа: <https://doi.org/10.17116/oftalma202113703149>. – Дата доступа: 01.08.2024.

[4] Sarki, R. Automatic detection of diabetic eye disease through deep learning using fundus images: a survey / R. Sarki [et al.]. – IEEE Access. – 2020. – Vol. 8. – P. 151133–151149. – DOI: 10.1109/ACCESS.2020.3015258.

[5] Ting, D. Artificial intelligence and deep learning in ophthalmology / D. Ting [et al.]. – Br J Ophthalmol. – 2019. – Vol. 103. – P. 67–75. – DOI: 10.1136/bjophthalmol-2018-313173.

[6] Rahimy, E. Deep learning applications in ophthalmology / E. Rahimy. – Curr Opin Ophthalmol. – 2018. – Vol. 29. – P. 54–60. – DOI: 10.1097/ICU.0000000000000470.

[7] Deep learning in ophthalmology: a review / P. S. Grewal. – Can J Ophthalmol. – 2018. – Vol. 53(4). – P. 309–313. – DOI: 10.1016/j.jcjo.2018.04.019.

[8] EfficientNet paper — Tan, M., & Le, Q. v. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. 36th International Conference on Machine Learning, ICML 2019, 2019-June.

CLASSIFICATION OF EYE DISEASES ON FUNDUS IMAGES USING EFFICIENTNET NEURAL NETWORK

K.I. Svistunova

*Student of the Department of
Information Management Systems of
BSU*

S.V. Ablameyko

*Professor of BSU,
DSc of Technical sciences,
Academician of the National Academy
of Sciences of Belarus,
Senior Researcher of the UIIP*

Abstract. The paper presents an algorithm for recognizing human eye diseases from fundus images. Based on the analysis, the convolutional neural network EfficientNet is selected and the algorithm based on EfficientNetB6 and EfficientNetB0 is proposed. Experimental studies have shown high efficiency for diagnosing eye diseases.

Keywords: neural networks, EfficientNet, augmentation, image analysis, pattern recognition, classification, fundus images.