

УДК 004.032.26

ПОВЫШЕНИЕ ТОЧНОСТИ НЕЙРОННОЙ СЕТИ ПУТЕМ РАСШИРЕНИЯ ТРЕНИРОВОЧНОГО НАБОРА ДАННЫХ С ПОМОЩЬЮ АУГМЕНТАЦИИ



А.А. Бирилло
студент кафедры
ИПиЭ БГУИР



О.С. Медведев
магистр техники и технологии, инженер-
программист, кафедра ИПиЭ БГУИР



А.Л. Шаршун
студент кафедры
ИПиЭ БГУИР



О.Н. Дудич
кандидат
медицинских наук,
доцент, кафедра
офтальмологии,
БелМАПО



В.Л. Красильникова
доктор
медицинских наук,
профессор,
кафедра
офтальмологии
БелМАПО



В.С. Осипович
кандидат технических
наук, доцент, кафедра
ИПиЭ БГУИР



К.Д. Яшин
Заведующий кафедрой
ИПиЭ БГУИР,
кандидат технических
наук, доцент

Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь, Минск.

Белорусская медицинская академия последипломного образования, отделение офтальмологии, Республика Беларусь, Минск.

E-mail: o.med@bsuir.by.

А. А. Бирилло

Студент БГУИР, специальность «Информационные системы и технологии (в обеспечении промышленной безопасности)». Будущая квалификация инженер-системотехник.

О. С. Медведев

В 2013 году окончил БГУИР, специальность «Инженерно-психологическое обеспечение информационных технологий». Квалификация инженер-системотехник. В 2019 году окончил магистратуру по специальности «Управление безопасностью производственных процессов» с присвоением академической степени магистра техники и технологии. Ведет курсы «CISCO».

А. Л. Шаршун

Студент БГУИР, специальность «Информационные системы и технологии (в обеспечении промышленной безопасности)». Будущая квалификация инженер-системотехник.

О. Н. Дудич

Окончила Саратовский государственный медицинский институт по специальности «Лечебное дело». Защитила диссертацию на соискание ученой степени кандидата медицинских наук на тему «Профилактика послеоперационных воспалительных осложнений в хирургии возрастной катаракты с использованием инъекционной формы кислоты ацетилсалициловой». Является автором 62 печатных работ, 1 патента, 10 учебнометодических пособий, 5 инструкций по применению. Основной исполнитель научного проекта «Разработка и внедрение изделия на основе нерезорбируемых материалов для

устранения посттравматических дефектов и деформаций глазницы и глазничного органакомплекса с использованием технологий 3D моделирования и прототипирования».

В. Л. Красильникова

Окончила Гродненский государственный медицинский институт по специальности «Лечебное дело». Защитила диссертацию на соискание ученой степени кандидата медицинских наук. В 2007 г. защитила диссертацию на соискание ученой степени доктора медицинских наук на тему «Медико-социальная реабилитация пациентов с анофтальмом с помощью композиционного офтальмологического имплантата: кликоэкспериментальное исследование». Является автором более 130 печатных работ, 5 патентов, 10 учебно-Пятая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», Минск, Республика Беларусь, 13-14 марта 2019 года 55 методических пособий, 3 инструкции по применению. Под ее руководством защищена 1 диссертационная работа на соискание ученой степени кандидата медицинских наук. Научный руководитель проекта «Разработка и внедрение изделия на основе нерезорбируемых материалов для устранения посттравматических дефектов и деформаций глазницы и глазничного органакомплекса с использованием технологий 3D моделирования и прототипирования».

К. Д. Яшин

Руководитель научно-исследовательской группы НИГ – 7.1 «Системы и приборы экологического мониторинга в управлении безопасностью жизнедеятельности».

В. С. Осипович

В 2004 году окончил Белорусский государственный университет информатики и радиоэлектроники по специальности «микроэлектроника», а в 2005 году получил степень магистра по той же специальности. В 2010 году защитил диссертацию на соискание ученой степени кандидата технических наук по приборам, системам и медицинской технике. Он является автором 115 публикаций, 2 патентов, 12 учебных пособий. Ежегодно он устно выступает на международных научных форумах. Основные направления научной деятельности: разработка алгоритмов и технологий обработки больших данных, исследования и разработки в области программной обработки медицинских изображений.

Аннотация. Представлены результаты исследований по улучшению точности распознавания поиска глазных яблок и костных структур глазниц. Для достижения цели было использовано глубокое обучение нейронной сети с использованием аугментации и без неё. Показано, как с расширением количества тренировочного набора увеличивается точность нейронной сети.

Ключевые слова: глазное яблоко, костные структуры глазницы, аугментация, разметка, смещение, деформация, повороты, шумы.

Введение.

Машинное обучение (ML) и особенно глубокое обучение (DL), способны выявлять, локализовать и количественно определять патологические особенности, фактически при любом заболевании, в том числе, и при патологии глазницы.

В то же время, анализ изображений, полученных с применением микроспирального компьютерного томографа MSCT [5-7] в формате DICOM [8] и трехмерная реконструкция костей лицевого черепа, позволяет хирургу более надежно оценить анатомические особенности отдельного пациента, локализацию, границы и распространенность патологического процесса, и планировать объем операции [9, 10].

Для работы нейронной сети с полученными результатами работы компьютерного томографа за основной элемент программной обработки были взяты биомаркеры.

Определение биомаркеров – основные элементы глазницы и глазного яблока, которые должна распознать нейросеть после ее обучения.

В настоящий момент тренировочный набор данных для последующего обучения сети ограничен, что не позволяет улучшить точность нейронной сети. Для того, чтобы расширить исходный набор будет применена аугментация. То есть построение дополнительных данных из исходных при решении задач машинного обучения путём внесения искажений, шумов и применения других приёмов.

Цель работы разработка программного средства, позволяющего однозначно идентифицировать биомаркеры, ограничивающие костную глазницу, а также анализ результатов

использования аугментации при обучении нейронной сети.

Методы и материалы

Исходные данные. В качестве исходных данных для обучения нейронной сети были использованы результаты микроспиральной компьютерной томографии пациентов с переломами костей глазницы различной тяжести. По каждому пациенту анализировался набор изображений в формате DICOM, полученных с использованием результате Микроспирального компьютерного томографа Siemens Emotion 6 (Германия).

Подготовка изображений для обучения нейронной сети осуществлялась путём послойной разметки этих изображений. До процесса разметки была осуществлена конвертация DICOM изображений в RGB изображения. В качестве инструмента для разметки данных использовали приложение VGG Image Annotator – это приложение для ручного аннотирования изображений с возможностью осуществлять множественную разметку. Результатом разметки являются файлы в формате csv и json, которые хранят информацию о координатах точек, ограничивающих костную глазницу в привязке к названию файла. Рисунок 1 отражает пример разметки исходных файлов.



Рисунок 1. Результат разметки исходных данных в одном слое

Все исходные данные были разделены на тренировочные и тестовые в процентном соотношении 90 и 10 % соответственно. Тестовые исходные данные использовались для проверки работы нейронной сети после обучения. Кроме того, тестовые исходные данные использовались в качестве контрольного эксперимента для сравнения результатов расчёта объёма глазниц размеченных нейронной сетью с объёмом глазниц, размеченных вручную.

Обучение нейронной сети. Решение поставленной задачи было осуществлено средствами языка программирования Python с использованием платформы Anaconda. В качестве основных фреймворков для работы с нейронными сетями мы использовали Tensorflow и Keras. Количество изображений, использованных для обучения нейронной сети постоянно увеличивалось, начиная с одной сотни в начале и 8000 в конце. Также была разработана собственная архитектура нейронной сети, которая представлена на рисунке 2.

Улучшение точности работы программного средства. Поскольку изначальное количество изображений не удовлетворяло минимально допустимому, было принято решения об использовании методов, увеличивающих число изображений для обучения. Ответом на данный запрос является аугментация.

Также при выборе методов было учтено с каким их подобием может столкнуться модель в действительности. К примеру, это может быть разная яркость изображения, различные шумы, растяжение/сжатие, смещение относительно центра, повороты изображения и совокупность всех вышеуказанных методов в разных пропорциях.

Таким образом, решено создать следующие виды аугментации и их смешение:

1. Смещение (сдвиг) изображения (bias augmentation).

2. Изменение яркости изображения (вариация метода главных компонент, brightness augmentation).
3. Шумы (noise augmentation) – создание шумов различной интенсивности по всему изображению.
4. Повороты изображения (rotate_augmentation).
5. Деформация изображения.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 497, 497, 16)	12304
max_pooling2d_1 (MaxPooling2D)	(None, 248, 248, 16)	0
conv2d_2 (Conv2D)	(None, 241, 241, 32)	32800
max_pooling2d_2 (MaxPooling2D)	(None, 120, 120, 32)	0
conv2d_3 (Conv2D)	(None, 116, 116, 32)	25632
max_pooling2d_3 (MaxPooling2D)	(None, 58, 58, 32)	0
conv2d_4 (Conv2D)	(None, 57, 57, 64)	8256
max_pooling2d_4 (MaxPooling2D)	(None, 28, 28, 64)	0
conv2d_5 (Conv2D)	(None, 26, 26, 128)	73856
max_pooling2d_5 (MaxPooling2D)	(None, 13, 13, 128)	0
conv2d_6 (Conv2D)	(None, 11, 11, 256)	295168
flatten_1 (Flatten)	(None, 30976)	0
dropout_1 (Dropout)	(None, 30976)	0
dense_1 (Dense)	(None, 512)	15860224
dense_2 (Dense)	(None, 256)	131328
dense_3 (Dense)	(None, 8)	2056
Total params: 16,441,624		
Trainable params: 16,441,624		
Non-trainable params: 0		

Рисунок 2. Архитектура работы нейронной сети

Алгоритм подготовки дата сета и подачи его в нейронную сеть для обучения построен таким образом, что аугментация осуществляется перед подачей изображений в нейронную сеть для момента обучения. Это гарантирует, что нейросеть каждый раз будет получать на вход псевдоновое, изменённое изображение, что позволит увеличить её точность.

Рассмотрим в качестве примера на одном слое (рисунок 3) варианты преобразования изображения (рисунок 3.).



Рисунок 3. Исходное изображение до аугментации и смещений

Смещение (сдвиг) изображения можно получить путём нахождения размерности матрицы изображения (`num_rows`, `num_cols`), составления матрицы для аффинных преобразований. Поскольку изображение перемещается лишь по оси x и y , то угол альфа равен 90° . Далее применяется оптимизированная функция (`cv2.warpAffine`), которая автоматически создаёт удовлетворяющий размер изображения. Выходное изображение представлено на рисунке 4 (А).

Изменение яркости изображения получается созданием исходной матрицы изображения такой же размерности, чтобы избежать изменения исходной. Далее идёт прохождение по каждому RGB пикселю и умножение значения каждого канала на введённый коэффициент яркости, проверяя при этом, чтобы он не вышел за пределы диапазона $(0, 255)$. После прохождения всего изображения получается «осветлённое» изображение. Результат обработки представлен на рисунке 4 (Б).

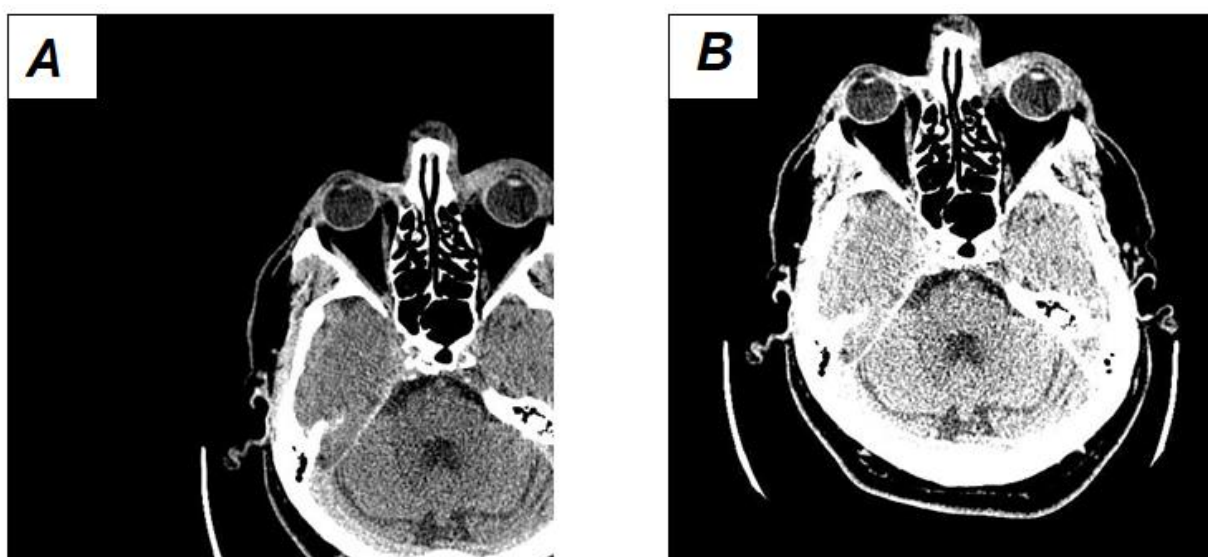


Рисунок 4. Смещение (сдвиг) изображения (А), изменение яркости изображения (В)

Шумы можно получить произвольным обходом по каждому пикселю и, если хотя бы один из трёх каналов не равен нулю, прибавляется к нему число, полученное генератором случайных

чисел, не выходящее из диапазона (-noise_level, noise_level) После окончания прохода проводится проверка, что значения каналов RGB не будут выходить из диапазона (0, 255) и, если данное условие не выполняется, производится обрезание значений до их приемлемого уровня Результат обработки представлен на рисунке 5 (А).

Поворот изображения основан на функции аффинных преобразований Для матрицы определяется модуль синуса и косинуса угла Далее находится значение, которое нужно добавить к разрешению по высоте и ширине, потому как изображение размерность не совпадает с исходной матрицей После применения аффинного преобразования и получается новое изображение, представленное на рисунке 5 (В).

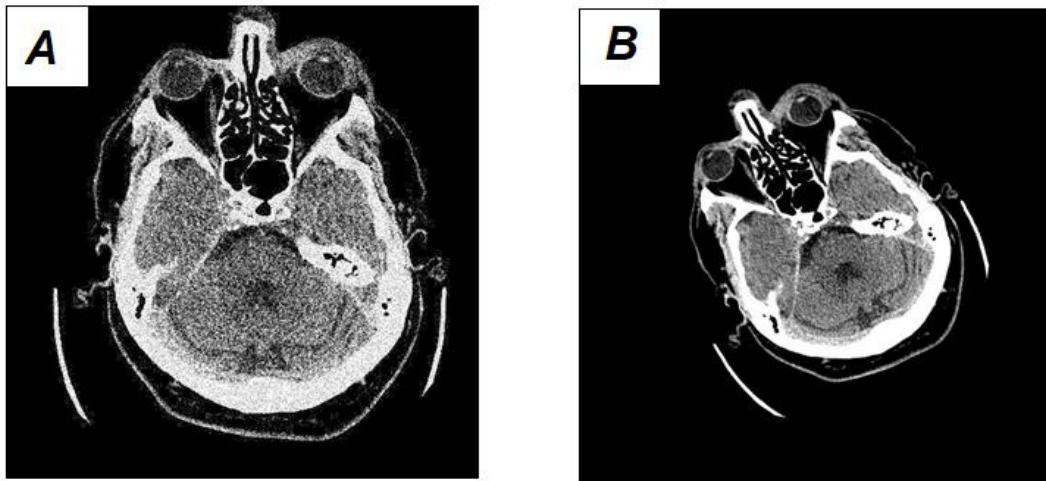


Рисунок 5. Добавление шумов изображению (А), поворот изображения (В)

Деформация изображения получается также путем использования аффинного преобразования. Используется достаточно простая матрица М, ненулевые элементы которой являются коэффициентами растяжения/сжатия по оси Х и Y. Далее применяется само преобразование, и получается новое изображение в виде матрицы, представленное на рисунке 8.

Масштабирование основано на функции деформации, только коэффициент масштабирования по оси Х и Y одинаковый. Всё те же аффинные преобразования и последующая обрезка изображения. Результат преобразования представлен на рисунке 9.

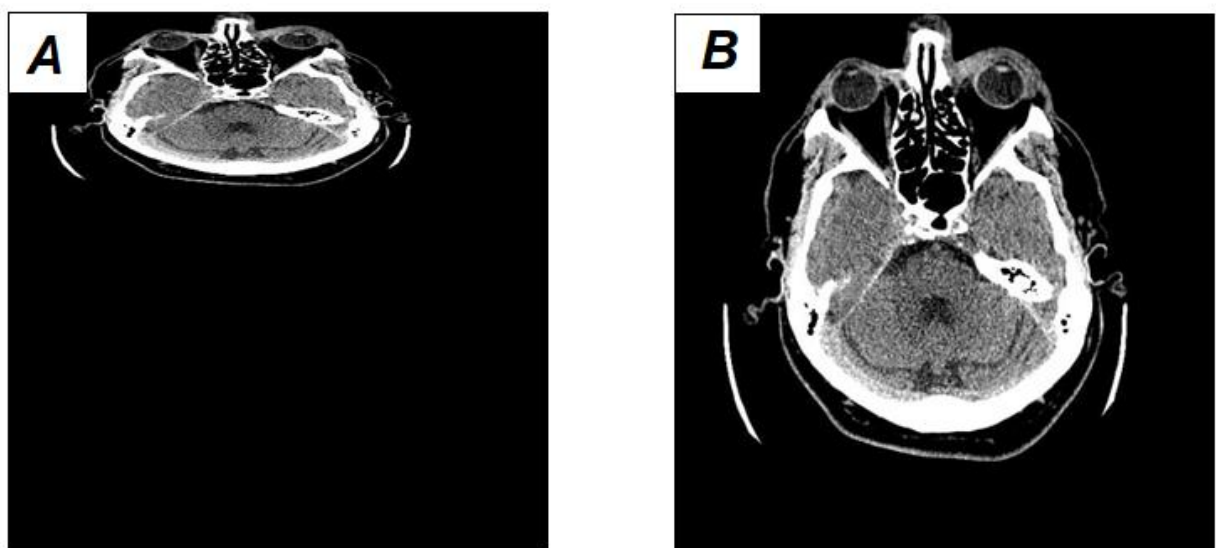


Рисунок 6. Деформация изображения (А), масштабирование изображения (В)

Результаты и обсуждение.

В результате обучения была получена матрица весовых коэффициентов для каждого слоя нейронной сети. Полученные весовые коэффициенты были использованы для поиска координат входа в глазицу на послойных результатах сканирования костей глазиц, тестовых наборах данных.

В качестве показателей точности использовалась такая величина, как MAE – Mean Absolute Error или средний модуль отклонения. В данном случае проще представлять данную метрику ошибки в качестве среднего отклонения разметки нейронной сети от эталонной, сделанной человеком (как разницу между координатами эталонных биомаркеров и сделанных нейронной сетью). Чем меньше ошибка сети при проверке, тем больше и лучше её точность, так как она тем самым наиболее приближена к разметке, выполненной человеком.

Поскольку нейронная сеть обучалась итеративно в течении пятидесяти эпох, то существует пятьдесят значений, характеризующих точность сети за эпоху (MAE). Однако, после первой эпохи существует высокая вероятность «запоминания» нейронной сетью тренировочных изображений и выдаче правильных результатов только на изображения, состоящие в обучающей выборке. Чтобы избежать этого, существует валидационная выборка изображений и разметки, которую нейронная сеть никогда не видела. В конце каждой эпохи сеть проверяется на таком наборе изображений. Таким образом, появляется величина, более точно характеризующая нейронную сеть, а именно Val_MAE – Validation Mean Absolute Error или проверочный средний модуль отклонения.

В итоге обработки изображений тестовых наборов данных с использованием обученной нейронной сети с аугментацией и без её использования были получены зависимости ошибки нейронной сети в зависимости от конкретной эпохи (рис. 11, 12). Как видно, в результате прохождения каждой эпохи, ошибка сети MAE плавно уменьшается, что свидетельствует о работе функции-оптимизатора нейронной сети (рис. 7.).

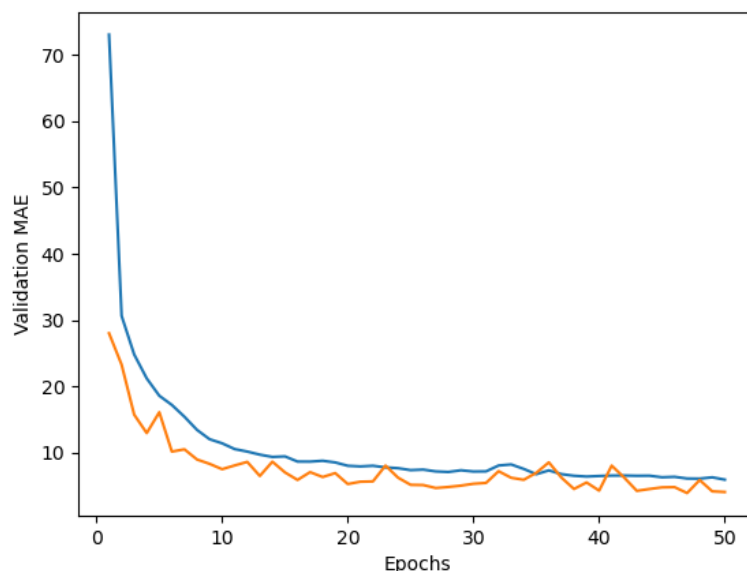


Рисунок 7. Результат обучения без аугментации: синяя линия – ошибка сети при обучении на тренировочном наборе (MAE); Оранжевая линия –ошибка при тестировании на проверочном наборе (Val_MAE)

Как видно из изображений, точность сети в конце обучения с применением аугментации удвоилась, при этом по результатам тестирования сеть, обученная с применением аугментации, оказалась более приспособлена к нестандартным и нетипичным входным данным. Так, ошибка сети, обученной без использования аугментации находилась в диапазоне 6-5 условных единиц (Val_MAE), а в случае использования аугментации конечная точность составляла 3-2 условные единицы (Val_MAE).

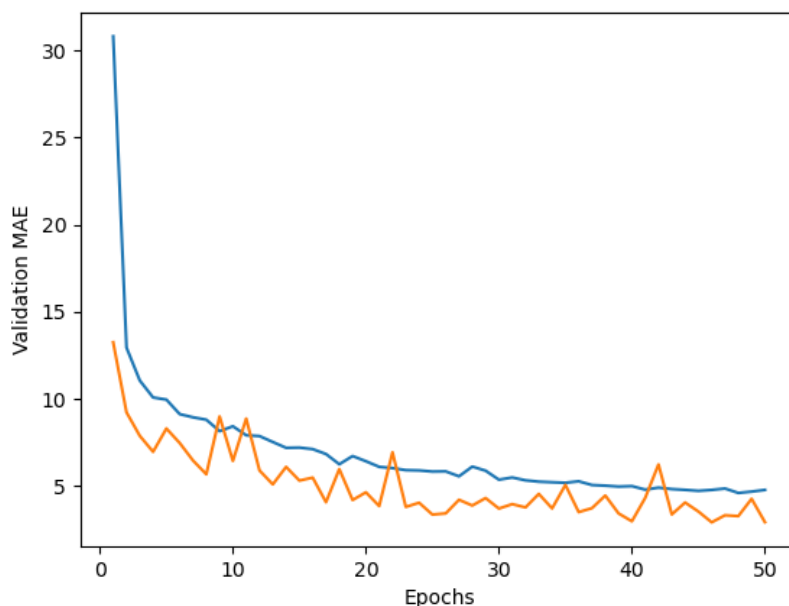


Рисунок 8. Результат обучения с аугментацией: синяя линия – ошибка сети при обучении на тренировочном наборе (MAE); Оранжевая линия – ошибка при тестировании на проверочном наборе (Val_MAE)

Это свидетельствует о высокой точности разметки глазниц нейронной сетью обученной с применением аугментации. Ошибку в пределах 2-3 единиц можно игнорировать, поскольку она представляет из себя всего несколько пикселей, что крайне мало. Предложенное программное средство с увеличенной точностью целесообразно использовать для автоматизации процесса расчёта объёма глазниц на этапе подготовки к операции и при оценке результатов операции по замещению тонких костей глазницы.

Поведение оранжевой линии (рис 12), а именно резкий рост Val_MAE вызваны тем, что в результате обучения и аугментации с последующей оптимизацией своих весовых коэффициентов может попадать в локальные минимумы, повышая свою ошибку. Чем более разнообразные данные сеть получает на входе, а аугментация, как известно, делает входные данные более разнообразными, тем сильнее сеть должна оптимизироваться для достижения лучших результатов.

Заключение.

Установлено, что погрешность расчётов координат биомаркеров на основе биомедицинских изображений (результатов компьютерной томографии) с использованием нейронной с применением аугментации сократилась вдвое по сравнению с обычным обучением.

Список литературы

- [1] Глубокое обучение на Python Автор: Шолле Франсуа Год: 2018 Издательство: Питер ISBN 978-5-4461-0770-4 Страниц: 400
- [2] Сверточная нейронная сеть [Электронный ресурс]. – Режим доступа: <https://intellect.ml/svertochnaya-nejronnaya-set-convolutional-neural-network-cnn-6013>(дата обращения 01.05.2018)
- [3] Применение нейронных сетей для задач классификации [Электронный ресурс]. – Режим доступа: <https://basegroup.ru/community/articles/classification>(дата обращения 08.05.2018)
- [4] Что такое TensorFlow [Электронный ресурс]. – Режим доступа: <https://www.computerworld.ru/articles/Chto-takoe-TensorFlow-i-kak-eto-ispolzuetsya>(дата обращения 15.05.2018)
- [5] Keras: Deep Learning for humans [Электронный ресурс]. – Режим доступа: <https://github.com/keras-team/keras>
- [6] Deep Learning: Сравнение фреймворков для символического глубокого обучения URL:

<https://habr.com/company/microsoft/blog/313318/> (дата обращения 27.05.2018)

[7] Digital imaging and communications in medicine [Электронный ресурс]. – Режим доступа: <https://www.dicomstandard.org/current/> (дата обращения: 25.01.2020).

[8] Aggregation Network for Instance Segmentation [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1803.01534> (дата обращения: 25.01.2020).

[9] Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow [Электронный ресурс]. – Режим доступа: https://github.com/matterport/Mask_RCNN/ (дата обращения: 25.01.2020).

USE OF THE NEURAL NETWORK IN PROCESSING OF MEDICAL IMAGES AND CALCULATION OF VOLUME OF THE EYE

A. BIRILLO

*Student of Human
Engineering and Ergonomics
and Radioelectronics*

O. MEDVEDEV

*Engineer of Human
Engineering and Ergonomics,
Master of Technical Science*

A. SHARSHUN

*Student of Human
Engineering and Ergonomics
and Radioelectronics*

A. DUDZICH

*M.D., associate professor,
Department of Ophthalmology,
Belarusian medical academy of
postgraduate education.*

V. KRASILNIKOVA

*M.D., Professor, Department of
Ophthalmology, Belarusian
medical academy of postgraduate
education*

V. OSIPOVICH,

*PhD, Associate Professor,
the Chair of Engineering
Psychology and
Ergonomics*

K. YASHIN

*PhD, Head of the Department of
Human Engineering and
Ergonomics, BSUIR*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus, Minsk
Belarus Medical Academy of Postgraduate Education, Republic of Belarus, Minsk
E-mail: o.med@bsuir.by*

Abstract. These are the results of research to improve the recognition accuracy for eyeballs and bone eye sockets structures. Deep neural network has applied with and without augmentation. It has shown how neural network accuracy increases with numbers expansions of training sets.

Keywords: eyeball, bone structures of the eye socket, augmentation, marking, displacement, deformation, rotations, noises.