

# СЕГМЕНТАЦИЯ СОСУДИСТОЙ СЕТИ НА ВЫСОКОДЕТАЛИЗИРОВАННЫХ ИЗОБРАЖЕНИЯХ ГЛАЗНОГО ДНА

**Косик И. И.<sup>1)</sup>, Недзьведь А. М.<sup>2)</sup>**

*<sup>1)</sup> Белорусский государственный университет, Минск, Беларусь,  
e-mail: nedzveda@gmail.com*

*<sup>2)</sup> Белорусский государственный медицинский университет, Минск, Беларусь,  
e-mail: ivankosik91@gmail.com*

Анализ состояния глазного дна позволяет получить важную информацию о наличии и стадии важных заболеваний, таких как: гипертоническая болезнь, сахарный диабет, ревматизм. Возможности автоматизированного получения информации о состоянии сетчатки являются определяющими для выявления патологий глазного дна и лечения заболеваний, ставших их причиной. Сегментация сосудистой сети и анализ ее характеристик является одним из важнейших этапов получения данной информации.

Высокое разрешение позволяет визуализировать множество мелких сосудов. Поэтому в данной работе предлагается использовать нейронную сеть с предсказанием на тайлах для достижения высокой точности распознавания тонких сосудов на высокодетализированных изображениях. Тайлы – небольшие изображения одинаковых размеров, которые и служат фрагментами большой картины.

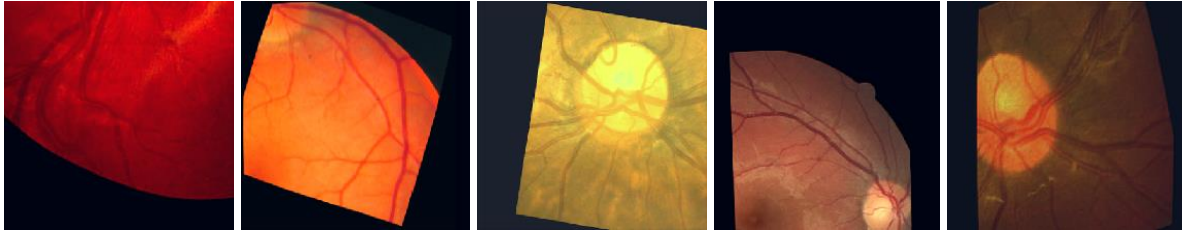
Для обучения нейронной сети использовались четыре открытых набора данных: DRIVE, STARE, CHASE DB1 and HRF. Для обучения изображения из всех наборов смешивались, чтобы получить модель нейронной сети, демонстрирующую хорошее качество сегментации на большем разнообразии снимков. Обучающая выборка составила 87 изображений; валидационная – 26; тестовая – 15.

Для получения высокого качества сегментации на реальных клинических изображениях, отличающихся большой вариативностью и наличием различных артефактов, к обучающим изображениям на каждой эпохе применялись различные фильтры из библиотеки albumentations: ColorJitter (для небольшой смены цветовых характеристик); ShiftScaleRotate, HorizontalFlip, VerticalFlip, RandomRotate90, OpticalDistortion, RandomBrightnessContrast, RandomGamma, MultiplicativeNoise и последующий Blur (для добавления размытых шумов), IAABoss, IAASharp, MotionBlur, GridDistortion. Дополнительно была реализована аугментация сжатия одной из сторон изображения для искажения пропорций.

Использовалась архитектура нейронной сети Unet с бэкбоном inceptionv3, предобученном на наборе данных ImageNet [1].

Основная сложность при обучении модели с данной архитектурой заключается в большом количестве параметров. Поэтому разрешение входного изображения сильно ограничено текущими аппаратными возможностями (количеством памяти видеокарты). На используемой для обучения видеокарте (GeForce GTX 1080 Ti с 11 ГБ видеопамяти) оптимальным является разрешение входного изображения около 352 x 352. При этом набор данных HRF содержит изображения с разрешением 3504 x 2336. Обычное сжатие изображения до такого маленького разрешения приводит к потере мелких сосудов. Поэтому использовался подход, при котором нейронная сеть

обучается на тайлах - частях изображений. Исходные изображения сжимались до разрешения 704 x 704, при котором мелкие сосуды еще не затираются. Каждое изображение валидационной выборки разбивалось на 4 тайла с размерами 352 x 352, таким образом количество валидационных изображений увеличивается в 4 раза. Для обучающей выборки использовался динамический подход: с помощью аугментации RandomSizedCrop на каждой эпохе из изображения 704 x 704 вырезался случайный кусок размером 352 x 352 либо немного больше, и сжимался (при необходимости) до 352 x 352. Пример полученных обучающих изображений приведен на рисунке 1.



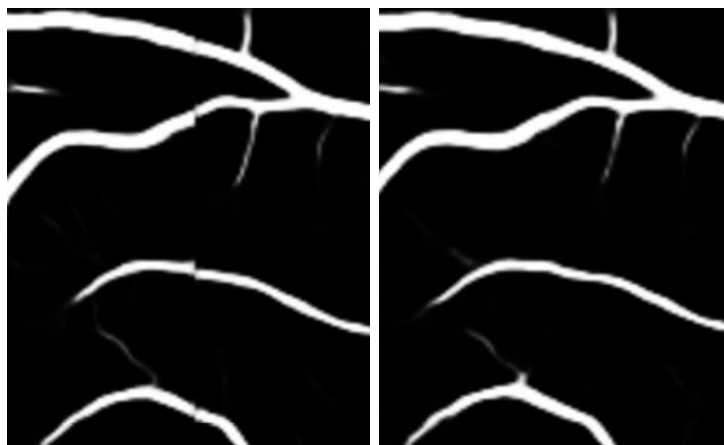
*Рис. 1. Примеры обучающих тайлов после аугментаций*

В качестве функции потерь использовалась комбинация бинарной кросс-энтропии и коэффициента Жаккарда. Данная функция потерь работает лучше с бинарными масками, а не полутоновыми. Поэтому при обучении маски сосудов после аугментации приводились к бинарным. Для последующего улучшения работы сети планируется проверить функции потерь, способные лучше работать именно с полутоновыми масками, например, Adaptive Wing Loss [2].

Первоначально сеть обучалась с замороженными весами архитектуры нейронной сети (которая уже предобучена на ImageNet) [3]. Использовался оптимизатор Adam, начальная скорость обучения (learning rate):  $5e-3$ . После этого веса бэкабона размораживались и использовались различные скорости обучения для начальных (низкая скорость обучения) и конечных слоев (высокая скорость обучения), т.к. начальные слои отвечают за общие фильтры и практически не требуют модификации, а конечные слои отвечают за специфические признаки, поэтому требуют больших изменений, чтобы подстроить их под наши данные. Используемый размер батча: 8.

В обучении использовалось деление валидационных изображений по сетке 2 x 2, т.к. разбиение на большее количество тайлов увеличивало время обучения, не улучшая результата. Для предсказания используется разбиение 3 x 3, т.к. позволяет лучше выделять тонкие сосуды.

Первоначально анализируемое трехканальное изображение сжимается до размера 996 x 996. После этого разбивается на 9 тайлов с разрешением 352 x 352 таким образом, чтобы каждый кусок захватывал часть соседнего тайла (10 пикселей). Это необходимо для устранения искажений на границе между двумя тайлами (рис. 2).



*Рис. 2. Пример вертикальной границы между двумя тайлами. а) Простое соединение соседних тайлов с видимым искажением на границе. б) Удалении части границы и перекрытие с соседним тайлом, используя альфа-смешение*

Для каждого тайла выполняется сегментация сосудов нейронной сетью. После чего они объединяются в одно изображение. На границах тайлов нейронная сеть немного искажает маску, поэтому у тайлов обрезаются граничные 5 пикселей. Затем они объединяются в ряды. Чтобы сгладить переход между двумя тайлами, их границы в 5 пикселей перекрываются друг с другом, и результирующая яркость рассчитывается с помощью альфа-смешения, создавая плавный переход. Таким же образом полученные 3 ряда тайлов объединяются в сосудистую маску целого изображения. После этого полученная маска растягивается до размера исходного изображения.

Работа выполнялась при поддержке проектов БРФФИ-РФФИ Ф20Р-134 и «Разработать прецизионный метод обнаружения и оценки степени выраженности дегенеративного повреждения диска зрительного нерва, как индикатора ранней гибели ганглионарных клеток сетчатки при оптиконеуропатиях различного генеза» в рамках ГПНИ «Трансляционная медицина».

#### **Литература**

1. U-Net: Convolutional Networks for Biomedical Image Segmentation // Olaf Ronneberger, Philipp Fischer, Thomas Brox // Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, Vol.9351: 234--241, 2015, available at arXiv:1505.04597 [cs.CV]
2. C. Gros. SoftSeg: Advantages of soft versus binary training for image segmentation // C Gros, A Lemay, J Cohen-Adad // Medical Image Analysis, 102038, 2020.
3. Chen, C. Robust Multimodal Brain Tumor Segmentation via Feature Disentanglement And Gated Fusion/ C, Chen // Lecture notes in Computer Science. 2019.-Vol. 11766.№4.P.447-456.