

Сегментация радужной оболочки глаза на изображении с помощью сверточной нейронной сети архитектуры U-Net

Ю.Х. Гансева¹, Е.В. Мясников^{1,2}

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. Важную роль в работе методов идентификации личности по изображению радужной оболочки глаза играет точность сегментации. В настоящей работе исследуется метод сегментации радужной оболочки глаза, основанный на сверточной нейронной сети архитектуры U-Net. В рамках исследований выполнена ручная сегментация изображений используемого набора данных, определены оптимальные параметры обучения сети и произведена оценка качества сегментации. Все исследования в работе проведены с использованием открытого набора данных MMU Iris Image Database. Полученные результаты показали, что исследуемый подход обеспечивает высокую точность сегментации изображений радужной оболочки глаза.

1. Введение

Идентификация по радужной оболочке глаза является одним из наиболее точных и надёжных способов биометрической идентификации, так как текстура радужки крайне устойчива во времени. Важную роль в достижении высоких показателей качества идентификации по изображению радужной оболочки глаза играет точность сегментации, то есть выделения радужной оболочки глаза на изображении. Так как радужная оболочка представляет собой на изображении кольцевую область между зрачком и склерой, то решение задачи сегментации часто сводится к аппроксимации внутренней и внешней границ радужной оболочки глаза окружностями. Для этой цели в большинстве работ используются метод Даугмана [1] и преобразование Хафа для поиска окружностей [2], а также другие методы, например, основанные на анализе граничных точек [3].

К сожалению, выполненная указанными методами сегментация радужной оболочки глаза часто оказывается недостаточно точной из-за частичного перекрытия радужной оболочки веками и ресницами, а также возникновения бликов от источника света. Кроме того, сама форма радужной оболочки не всегда хорошо аппроксимируется окружностями. Таким образом, актуальной задачей представляется повышение качества сегментации радужной оболочки глаза.

С успехом развития моделей нейронных сетей глубокого обучения, исследователи всё чаще обращают внимание на сверточные нейронные сети для дальнейшего повышения точности существующих методов сегментации радужной оболочки. Для решения задачи сегментации радужной оболочки в настоящей работе предлагается использовать подход, основанный на обучении сверточной нейронной сети (CNN) архитектуры U-Net [4].

U-Net считается одной из стандартных архитектур CNN для задач сегментации изображений, когда нужно не только определить класс изображения целиком, но и произвести сегментацию областей, т. е. создать маску, которая будет разделять изображение на несколько классов. Одним из основных достоинств сети, важным в контексте настоящей работы, является возможность ее обучения на небольшом количестве данных.

Настоящая работа построена следующим образом. Во втором разделе приводится описание исследуемого подхода к сегментации изображения радужной оболочки глаза, описывается архитектура сети, особенности подготовки данных и ее обучения. Третий раздел посвящен описанию экспериментов. В нем описывается использованный в работе набор данных, приводятся результаты экспериментальных исследований. Работа завершается заключением и списком литературы.

2. Метод

2.1. Архитектура U-Net

U-Net представляет собой популярную архитектуру CNN для решения биомедицинских задач (сегментирование различных видов клеток, определение границ между плотными клеточными структурами и др.) и других задач сегментации изображений. Основным преимуществом этой модели является способность обучаться на небольших наборах данных и при этом показывать достаточно высокую точность, что является распространенной проблемой для задач компьютерного зрения.

Архитектура сети приведена на рисунке 1.

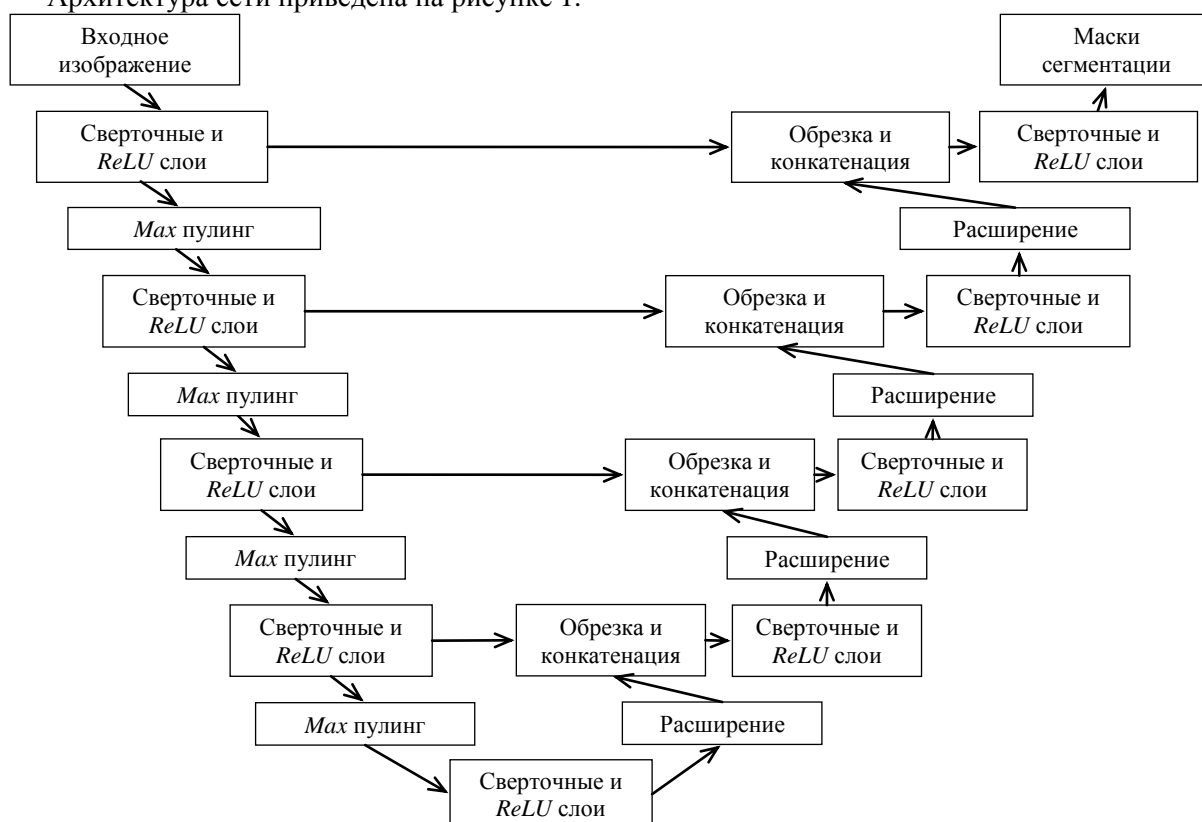


Рисунок 1. Схематическое изображение нейронной сети архитектуры U-Net.

Как видно из рисунка, архитектура состоит из сжимающего пути (левая часть) для захвата контекста и расширяющего пути (правая часть), позволяющего осуществить точную локализацию. В рамках данной работы два класса: радужная оболочка и не радужная оболочка.

Сжимающий путь соответствует типичной структуре сверточной нейронной сети. Он состоит из повторного применения двух сверток размерами (conv 3x3), за которыми следует слой ReLU и пулинг с функцией максимума (max pool 2x2) с шагом 2.

Каждый шаг расширяющего пути содержит расширяющий карту признаков слой (up conv), обратный пулингу, после которого следует свертка (conv 2x2), уменьшающая количество каналов признаков. После идет конкатенация с соответствующим образом обрезанной картой признаков (conv and crop) из сжимающего пути и два сверточных слоя (conv 3x3), после каждого из которых стоит слой ReLU.

2.2. Обучение модели и подготовительные этапы

На этапе подготовки к обучению была взята предварительно обученная модель [5, 8]. Набор данных был разделен на тренировочное и валидационное множества. Цветные исходные изображения конвертировались в полутоновые, и приводились к размеру 320 x 320 пикселей. Значения пикселей масштабировались до диапазона [0;1].

В качестве функции потерь, определяющей, как оценивается качество сети на обучающих данных, использовалась бинарная кросс-энтропия. В качестве оптимизатора, определяющего механизм обновления весов, был выбран Adam [6].

3. Эксперименты и результаты

Для проведения экспериментальных исследований, была выполнена реализация на языке Python с использованием библиотек Keras, Numpy, OpenCV и Matplotlib.

В качестве набора данных изображений радужной оболочки глаза использовалась база данных MMU Iris Database [7]. База содержала по 5 изображений для левого и правого глаза 45 человек, что дало в общей сложности 450 изображений. Для обучения и оценки качества сегментации была произведена ручная сегментация всех изображений из набора.

Было проведено девять экспериментов для определения лучших параметров, затем, используя их, было проведено обучение нейронной сети со значением количества эпох обучения (epochs) – 50. На этапе поиска лучших параметров при обучении предварительно обученная нейронная сеть [5, 8] настраивалась на указанных выше данных со следующими параметрами: скорость обучения (learning rate) для первого, второго и третьего экспериментов - 10^{-3} , для четвертого, пятого и шестого 10^{-4} , для седьмого, восьмого и девятого - 10^{-5} , количество эпох обучения (epochs) - 10, размер случайной подвыборки (batch size), используемой для оценки градиента, для первого, четвертого и седьмого экспериментов – 10, второго, пятого и восьмого – 20 и третьего, шестого, девятого - 40.

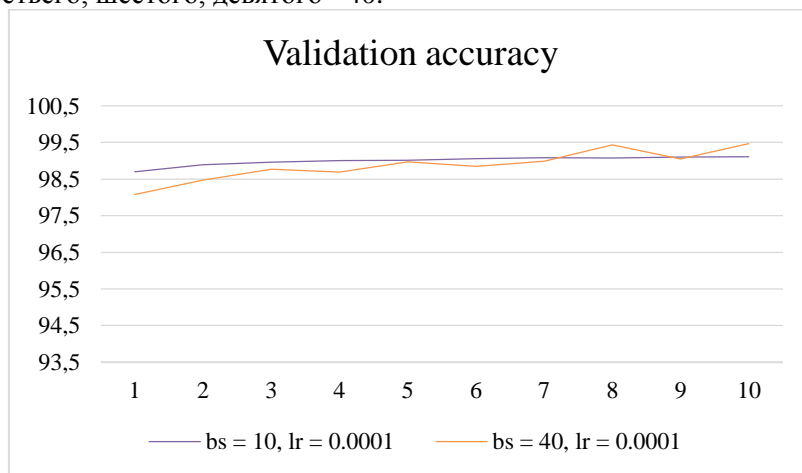


Рисунок 2. Обучение нейронной сети: график зависимости точности от эпохи обучения нейронной сети.

Для мониторинга процесса обучения по завершении каждой эпохи обучения по валидационной выборке оценивалось значение кросс-энтропии и точности сегментации.

Точность сегментации оценивалась как доля корректно классифицированных пикселей изображений валидационной выборки.

Значения функции потерь и точности обученной модели на тренировочных и валидационных данных при различных значениях гиперпараметров batch size и learning rate показаны в таблице 1.

Таблица 1. Результаты обучения для определения лучших параметров.

	Learning rate	Batch size	Точность	Значения функции потерь
Обучающий набор	10^{-3}	10	97,71	0,0262
		20	97,84	0,0293
		40	95,61	0,0771
	10^{-4}	10	98,04	0,0165
		20	97,98	0,0178
		40	98,04	0,0182
	10^{-5}	10	98,64	0,0135
		20	98,85	0,0142
		40	97,76	0,0384
Валидационный набор	10^{-3}	10	98,55	0,0245
		20	96,23	0,0322
		40	95,61	0,0812
	10^{-4}	10	99,11	0,015
		20	99,05	0,016
		40	99,47	0,009
	10^{-5}	10	99,19	0,0127
		20	96,89	0,0153
		40	99,03	0,0255

Графики изменения значений точности и функции потерь на первых 10 эпохах обучения нейронной сети для двух лучших моделей можно увидеть ниже на рисунках 2 и 3.



Рисунок 3. Обучение нейронной сети: график зависимости функции потерь от эпохи обучения нейронной сети.

В результате экспериментов были выбраны следующие значения параметров обучения: скорости обучения (learning rate) - 10^{-4} , размер случайной подвыборки (batch size) - 40. Обучение нейронной сети с указанными параметрами в течение 50 эпох позволило достичь на валидационной выборке значений точности и функции потерь равных 99,53 и 0,0071, соответственно.

Результаты работы исследуемого подхода для полученной модели показаны ниже на рисунке 4.

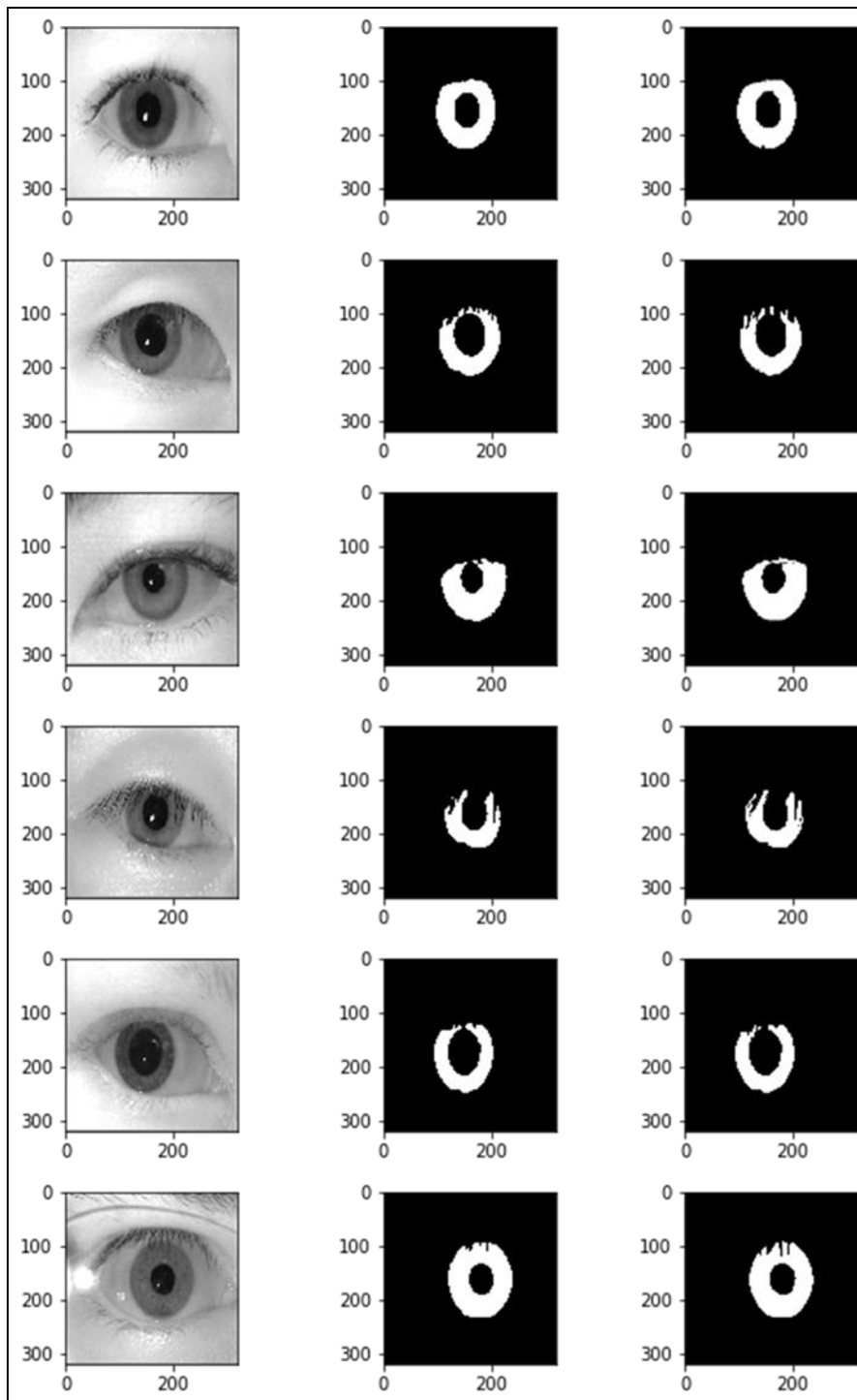


Рисунок 4. Примеры сегментации изображений радужной оболочки глаза: исходное изображение (левый столбец), маска сегментации, сгенерированная с использованием исследуемого подхода (средний столбец) и истинная маска сегментации (правый столбец).

4. Заключение

В настоящей работе исследуется метод сегментации радужной оболочки глаза, основанный на сверточной нейронной сети архитектуры U-Net. Для настройки и оценки метода выполнена ручная сегментация изображений из открытого набора данных MMU Iris Image Database.

С использованием указанного набора данных определены наилучшие значения параметров обучения сети и произведена оценка качества сегментации. Исследования показали, что исследуемый подход позволяет достигнуть 99,53% точности сегментами на валидационном множестве данных.

5. Благодарности

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 18-01-00748 в частях «2. Метод» - «3. Эксперименты и результаты» и Министерства науки и высшего образования РФ в рамках госзадания ФНИЦ «Кристаллография и фотоника» РАН в частях «1. Введение» и «4. Заключение».

6. Литература

- [1] Daugman, J.G. How iris recognition works // Proc. IEEE Trans. Circ. Syst. Video Technol. – 2004. – Vol. 14(1). – P. 21-30.
- [2] Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс – М.: Техносфера, 2005. – 1072 с.
- [3] Semenov, M.S. A comparison of iris image segmentation techniques / M.S. Semenov, E.V. Myasnikov // CEUR Workshop Proceedings. – 2018. – Vol. 2210. – P. 163-169.
- [4] Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // MICCAI, 2015. – P. 234-241.
- [5] Предварительно обученная модель U-Net [Электронный ресурс]. – Режим доступа: <https://github.com/jus390/U-net-Iris-segmentation>. (23.12.2019)
- [6] Kingma, D. Adam: A Method for Stochastic Optimization / D. Kingma, J. Ba // arXiv: 1412.6980v8, 2017.
- [7] MMU Iris Image Database: Multimedia University [Электронный ресурс]. – Режим доступа: <http://pesonna.mmu.edu.my/ccteo/>.
- [8] Lozej, J. End-to-End Iris Segmentation Using U-Net / J. Lozej, B. Meden, V. Struc, P. Peer // IEEE International Work Conference on Bioinspired Intelligence (IWOB), 2018. – P. 1-6.

Iris segmentation in image using U-Net convolutional neural network architecture

Y.K. Ganeeva¹, E.V. Myasnikov^{1,2}

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. The accuracy of segmentation plays an important role in the methods for person identification by iris images. In this paper, we study the iris segmentation method based on the convolutional neural network of the U-Net architecture. As part of the research, manual segmentation of the images of the used data set was performed, the optimal network training parameters were determined and the segmentation quality was evaluated. All studies in the paper were conducted using the open MMU Iris Image Database. The results showed that the studied approach provides high precision segmentation of the iris images.