

## ПРОГРАММНОЕ СРЕДСТВО ДЛЯ НАХОЖДЕНИЯ КООРДИНАТ ЗРАЧКА И УГОЛКОВ ГЛАЗ С ВЕБ-КАМЕРЫ

*Самаль И.Д., студент гр. 050501*

*Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Перцев Д.Ю. – канд. техн. наук*

В данной работе представлены особенности проектирования и обучения нейронной сети, предназначенной для определения координат центра зрачка и уголков глаза.

В настоящее время задача распознавания и отслеживания направления взгляда имеет достаточно широкое применение в различных сферах жизни, таких как психология, медицина, маркетинг, разработка программных продуктов и так далее. Решение данной задачи сводится к нахождению области глаз на изображении, а также трёх ключевых точек: центра зрачка и уголков глаза.

Определение области глаз на первом кадре видеопотока с веб-камеры выполняется при помощи Face Landmark Detection [1], находящего на изображении 478 ключевых точек лица, в том числе уголки глаз и центр зрачка. Задачей этой библиотеки является не отслеживание направления взгляда, а мониторинг мимики лица и эмоций человека, поэтому точность нахождения необходимых координат глаза весьма низкая, чтобы использовать её на каждом кадре. Так как в видеопотоке изменения между кадрами являются небольшими, для ускорения решаемой задачи было решено на каждом кадре вырезать область глаз по угловым координатам с прошлого кадра с некоторым запасом, после чего проводить её обработку.

На сегодняшний день существует множество различных алгоритмов цифровой обработки изображения для поиска зрачка, например, преобразование Хафа или морфологические методы, опирающиеся на предположение о том, что зрачок является наиболее тёмной областью глаза.

В отличие от предыдущей задачи, получившей большое развитие за счёт применения в медицине, распознавание уголков глаза не получило такого распространения. Также сложность данного вопроса заключается в том, что в зависимости от освещения края глаза могут являть как самой тёмной точкой, так и самой светлой.

Оптимальным вариантом решения нахождения ключевых точек глаза является свёрточная нейронная сеть, которая позволяет учитывать не только различное освещение, но и индивидуальные черты строения глаза, благодаря своей способности извлекать и анализировать визуальные признаки на разных уровнях абстракции.

За основу была взята нейронная сеть из проекта [2], определяющего центр зрачка в режиме реального времени, что делает её подходящей для обработки видеопотока. В качестве датасета использовался MPIIGaze [3], представляющий собой 37,667 изображений области глаз и координат трёх ключевых точек. Процесс предобработки изображений представляет собой преобразование к чёрно-белой модели, вырезание области глаза по известным координатам с запасом по оси Y пропорциональным расстоянию между краями глаза, приведение обрезанных изображений к единому размеру 32x16, пересчёт координат с учётом нового размера. Также проводится усреднение положения левого и правого зрачка с целью ускорения работы сети, что значительно снижает точность распознавания.

Исходя из вышесказанного, первоначальным вариантом было убрать усреднение, а также изменить структуру нейронной сети, чтобы результатом её работы были три точки, а не одна. Такая трансформация помогла немного увеличить точность нахождения центра зрачка, однако из-за физического различия левого и правого глаза, данная сеть не справилась с поиском уголков глаза. В результате работы нейронной сети было получено изображение (см. рисунок 1), на котором координаты одного из углов на обоих глазах симметрично подняты вверх от реального края глаза.



Рисунок 1 – Результаты работы первоначального варианта нейронной

На основе вышесказанного, было решено разделить обучение для левого и правого глаза как обучение двух нейронных сетей с одинаковой структурой. Также был изменён процесс предобработки изображения: при вырезании области глаза запас брался не только по оси Y, но также и по X, что

позволило при работе с видеопотоком уменьшить влияние предыдущих недочётов в предсказании, а также предусмотреть возможное изменение расстояния глаза от веб-камеры. В результате данных изменений координаты зрачков сместились ближе к его центру, уголки глаз также стали определяться намного точнее (см. рисунок 2). На наборе из 40 последовательно идущих кадров с незначительным изменением положения зрачка колебания точек практически отсутствуют, что говорит о стабильности работы сети.

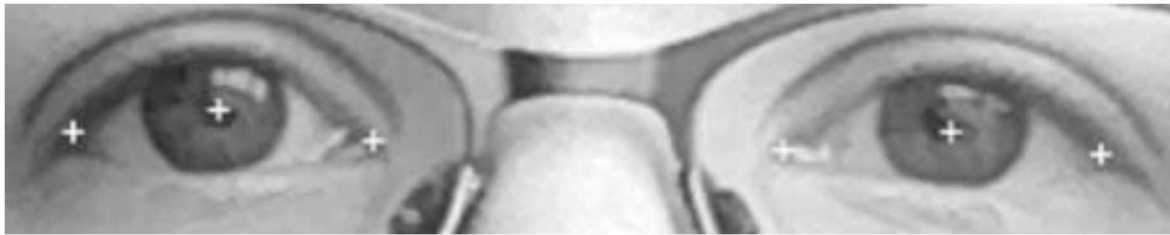


Рисунок 2 – Результаты работы двух нейронных

Из-за того, что угол направления взгляда на изображениях из датасета MPIIGaze варьируется от  $-1.5$  до  $20$  градусов по вертикали и от  $-18$  до  $+18$  градусов по горизонтали, обученная на нём сеть не справляется с обнаружением зрачка в случаях сильного отклонения взгляда, как, например, взгляд в левую сторону на  $45$  градусов, когда на изображении часть зрачка скрывается за веком.

На текущий момент в открытом доступе имеется не так много датасетов с изображениями области глаза и координатами трёх точек, а из подходящих под данное условие оптимальным является MPIIGaze, так как он содержит фотографии, сделанные с различного ракурса с учетом разного освещения.

Исходя из вышесказанного, было решено дополнить датасет собственными изображениями, содержащими направления взгляда, отсутствующие в MPIIGaze. Для этого было сделано около двух тысяч изображений пяти людей при различной освещённости, а также в полуавтоматическом режиме при помощи Face Landmark Detection проведена разметка изображений. Итоговый размер тренировочного датасета для каждого глаза составил  $7,000$  изображений, из которых  $1,500$  являются новыми. Это позволило улучшить результаты работы нейронной сети при сильном отклонении взгляда (см. рисунок 3), однако при обработке видеопотока с быстрыми и значительными изменениями направления взгляда возможны скачки координат.



Рисунок 3 – Результаты работы двух нейронных сетей после обучения на расширенном датасете

Таким образом, было разработано программное средство на основе свёрточных нейронных сетей для нахождения координат центров зрачка и уголков глаз в видеопотоке с веб-камеры, а также датасет размером в  $2,000$  изображений, состоящий из изображений лица и аннотаций к ним. Данное средство может применяться для решения задачи определения направления взгляда при небольшом повороте или наклоне головы.

**Список использованных источников:**

1. Face landmark detection guide [Электронный ресурс]. – Электронные данные. – Режим доступа: [https://developers.google.com/mediapipe/solutions/vision/face\\_landmarker#get\\_started](https://developers.google.com/mediapipe/solutions/vision/face_landmarker#get_started) – Дата доступа: 04.04.2024.
2. Мобильный eye-tracking на PyTorch [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://habr.com/ru/articles/501412/> – Дата доступа: 04.04.2024.
3. MPIIGaze: Real-World Dataset and Deep Appearance-Based Gaze Estimation [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://arxiv.org/pdf/1711.09017.pdf> – Дата доступа: 04.04.2024.