

## ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ ПАРАУНИТАРНЫХ БАНКОВ ФИЛЬТРОВ В НЕЙРОННЫХ СЕТЯХ

*Кузьма В.В., Леченко А.В*

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

В работе представлены результаты применимости параунитарных банков фильтров для уменьшения вычислительной сложности сверточных нейронных сетей в задаче сегментации изображения. В данной работе рассматривалась нейронная сеть U-Net, которая была обучена распознавать пешеходные переходы на 3083 размеченных дорожных изображениях.

Каждый год на дорогах гибнет около 300000 пешеходов, что составляет до 26% всех смертей в дорожно-транспортных происшествиях [1]. Это делает обнаружение пешеходных переходов важной задачей для обеспечения безопасности автомобильных средств. Для решения данной задачи можно использовать сверточные нейронные сети [2]. Они используются для сегментирования изображений, в том числе для выделения дорожных объектов, таких как пешеходные переходы [3]. В данной ситуации наиболее важным является время распознавания объекта. Одним из способов уменьшения времени распознавания является понижение размерности распознаваемого изображения методами пространственной фильтрации сигналов.

В данной работе представлено решение задачи сегментации пешеходных переходов на изображениях с использованием параунитарного банка фильтров на основе алгебры кватернионов [4], далее Q-ПУБФ и нейронной сети U-net [5]. Q-ПУБФ используется для понижения размерности входного изображения нейронной сети с целью понижения вычислительной сложности сегментации изображений.

Банк фильтров – цифровая система, состоящая из секций анализа и синтеза. Входной сигнал разбивается при помощи фильтров блока анализа на субполосные составляющие. На вход подается изображение размером  $N \times N$ . Разделение на субполосы позволяет наиболее декоррелировать входные данные. При этом, Q-ПУБФ концентрирует перцептуальную значимую информацию в низкочастотном канале. Для четырехполосного фильтра, размер канала равен  $N/4 \times N/4$ , то есть размер изображения уменьшен в 16 раз. Этот низкочастотный канал передается на вход U-net, которая на выходе выдает маску определяющую вероятность того, что данный пиксель принадлежит пешеходному переходу. Но эта маска также размера  $N/4 \times N/4$ . Для ее восстановления до размеров исходного изображения можно использовать блок синтеза Q-ПУБФ. Операции в блоке синтеза, являются обратными операциями блока анализа, т. е. позволяют восстановить исходный сигнал из его субполосных компонент.

Для проверки предложенного решения был выбран набор данных из 3083 изображений, на части из них пешеходного перехода не было. Снимки были сделаны как днем, так и ночью, а также при плохих погодных условиях. Все изображения трехцветные, размером  $256 \times 256 \times 3$ . Система анализа Q-ПУБФ преобразует входное изображение в тензор  $16 \times 64 \times 64 \times 3$ . Из этого тензора выделяется низкочастотная компонента исходного изображения размером  $64 \times 64 \times 3$  и передается на вход нейронной сети. На выходе получаем маску, которая представляет собой двумерный тензор размером  $64 \times 64$ . Каждое значение маски представляет собой вероятность того, что соответствующий пиксель принадлежит пешеходному переходу на входном изображении. Это решение сравнивалось с нейросетью той же архитектуры, которая распознавала изображения исходного размера [6]. Для корректности сравнения, из маски  $64 \times 64$  системой синтеза Q-ПУБФ восстанавливалась маска размера  $256 \times 256$ . Маска  $64 \times 64$  использовалась как низкочастотный канал, значение остальных каналов задавалось нулями.

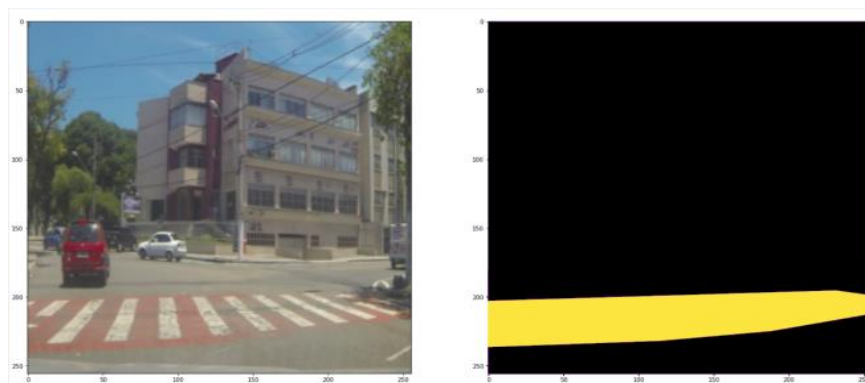


Рисунок 1 – Пример изображения (слева) и его маски (справа)

В качестве метрики качества сегментации использовалась Intersection over Union (IoU) [7]. Она измеряет сходство между предсказанной и размеченной областями. Чтобы её вычислить нужно найти отношение пересечения областей к их объединению. IoU вычислялось на масках размером 256x256.

Набор данных был случайным образом разделен на обучающий и проверочный наборы. Обучающий набор состоит из 2159 изображений (70% от общего набора данных), а набор для проверки составляет 924 изображения. Нейронная сеть тренируется на обучающем наборе, а затем предсказывает область пешеходного перехода на наборе проверки.

Результаты работы натренированных приведены в таблице 1. Использование Q-ПУБФ в сочетании с U-net ведет к незначительному снижению результата качества сегментации, по сравнению с работой U-net на исходных изображениях, однако значительно сокращает время распознавания. Таким образом, параунитарные банки фильтров могут использоваться в качестве инструмента для повышения производительности сверточных нейронных сетей. Более того, согласно проведенному эксперименту, использование параунитарных банков фильтров оказывает незначительное влияние на результат.

Таблица 1 – Результаты работы.

Решение	Длительность тренировки, мин	Длительность распознавания 924 изображений, с	IoU
U-Net	67	202	0.9496
Q-ПУБФ + U-Net	11	21	0.9165

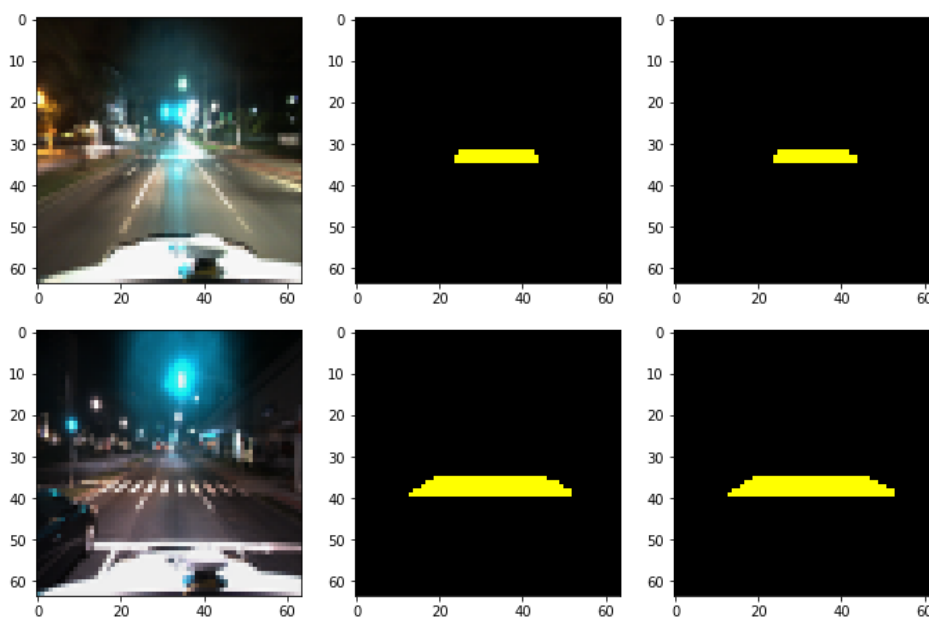


Рисунок 2 – НЧ канал исходного изображения, предсказанная область и область истинности

**Список использованных источников:**

1. W. H. Organization et al., "Global status report on road safety 2018: summary," World Health Organization, Tech. Rep., 2018.
2. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: *Backpropagation Applied to Handwritten Zip Code Recognition*, *Neural Computation*
3. M. A. Malbog, "Mask r-cnn for pedestrian crosswalk detection and instance segmentation," in 2019 IEEE 6th International Conference on Engineering
4. Парфенюк М., Петровский А.А. Параунитарные банки фильтров на основе алгебры кватернионов: теория и применение / М. Парфенюк, А. А. Петровский // «Цифровая обработка сигналов» - 2008 – 15 с.
5. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*.
6. E. Adaska, A. Lechanka, "UNetX: Real-time Pedestrian Crosswalk Segmentation on Mobile Device", in 2021 PRIP 15th International Conference on Pattern Recognition and Information Processing.
7. M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *International symposium on visual computing*. Springer, 2016, pp. 234–244.