

## Моделирование и реализация процесса определения дорожных объектов с применением аппарата сверточных сетей *RetinaNet*

*В.А. Евсина, С.Н. Широбокова, В.А. Жжонов, В.А. Евсин*

*Южно-Российский государственный политехнический университет (НПИ)  
имени М.И. Платова, Новочеркасск*

**Аннотация:** В данной статье рассматривается проблематика построения сверточных нейронных сетей для определения дорожных объектов. Представлена общая актуальность и постановка проблемы определения дорожных объектов. Сформировано обоснование применения искусственных нейронных сетей для определения дорожных объектов. В качестве основной архитектуры искусственной нейронной сети для определения дорожных объектов использована архитектура сети *Retinanet*. Визуализирована общая концепция данной архитектуры и основные подсети. Описаны функции ошибки для основных подсетей сети *Retinanet*. Дано проектное описание алгоритмов построения аннотации данных для обучения искусственной нейронной сети, а также алгоритмов построения архитектуры нейронной сети классификации, регрессии и пирамиды признаков. Определена динамика изменения общей функции ошибки при определении дорожных объектов. Представлен результат обучения искусственной нейронной сети.

**Ключевые слова:** сверточные нейронные сети, классификация, регрессия, сверточные нейронные сети, deep learning, big data, математическое моделирование, информатика, архитектура *RetinaNet*.

В процессе передвижения беспилотных автотранспортных средств по дорогам общего пользования возникает необходимость в распознавании объектов дорожного движения, к которым относятся пешеходы и автомобили. При определении ситуации на дороге необходимо обеспечивать обмен информацией между автомобилем и транспортной инфраструктурой в виде информации о ситуации на дорогах. Для решения данной задачи лучше всего подходит аппарат искусственных нейронных сетей для определения объектов дорожного движения (далее ИНС), пример в [1,2]. Преимуществом ИНС является способность самообучаться, то есть строить обобщения информации, определяя характеристические показатели обучающей выборки.

Архитектура сети для определения объектов на изображении должна определять объекты на фотографии в определенной области. Для реализации данной задачи использована архитектура сети *Retinanet*. Данная

одноступенчатая сверточная нейронная сеть позволяет определить объекты на фотографии и состоит из следующих основных частей, представляющих собой нейронные подсети:

– *backbone*: базовая сверточная нейронная сеть предназначена для извлечения общих признаков из входного изображения с помощью операции свертки, в качестве базовой нейронной сети используется архитектура *ResNet50*, подробнее о которой в [3];

– *пирамида признаков*: сверточная нейронная сеть, используемая для объединения достоинств карт признаков верхних, которые имеют низкое разрешение, но высокую обобщающую способность, и нижних, которые имеют высокое разрешение, но низкую обобщающую способность, уровней сети;

– *подсеть классификации*: сверточная нейронная сеть, которая извлекает из пирамиды признаков информацию о классе объектов. Данная сеть решает задачу классификации, состоит из пяти сверточных слоев. Первые четыре слоя состоят из 256 нейронов на каждом слое, пятый слой состоит из  $k * \alpha$  нейронов, где  $\alpha$  – максимальное количество объектов, которое возможно определить на изображении,  $k$  – количество классов, подробнее о сверточных нейронных сетях в [4-6];

– *подсеть регрессии*: сверточная нейронная сеть, которая извлекает из пирамиды признаков информацию о координатах изображения. Данная сеть решает задачу регрессии, состоит из пяти сверточных слоев. Первые четыре слоя состоят из 256 нейронов на каждом слое, пятый слой состоит из  $4 * \alpha$  нейронов, где  $\alpha$  – максимальное количество объектов, которое возможно определить на изображении. Архитектура данной сети представлена на рис. 1.



Рис. 1. – Архитектура *retinanet*

Пирамида признаков состоит из трех основных частей: восходящий путь, нисходящий путь и боковые соединения. Восходящий путь представляет собой последовательность сверточных слоев с уменьшающейся размерностью, этот путь имеет уязвимость при извлечении признаков – потеря важной информации об объекте, например, из-за зашумления небольшого, но значимого объекта фоном, так как к концу сети информация сжата и обобщена. Нисходящий путь представляет собой пирамиду, в которой карты признаков верхнего слоя имеют размер карт признаков верхнего слоя восходящего пути и увеличиваются вдвое по направлению вниз. В качестве функции ошибки подсети классификации используется модель *focal loss* ( $FL$ ), подробнее о которой в [7]. Данная модель функции ошибки является развитием концепции мультиклассовой кроссэнтропии и позволяет сбалансировать классы переднего и заднего плана изображения во время обучения. Общая модель  $FL$ :  $FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$ ,

где  $P_t$  – вероятность отнесения объекта к  $t$ -му классу,  $\gamma$  – коэффициент модуляции, который используется для того, чтобы уменьшить потери от

объектов с высокой степенью уверенности в классе,  $\alpha_i$  – взвешенный параметр.

В качестве функции ошибки подсети регрессии используется модель *smooth*, подробнее о которой в [8]. Данная модель функции ошибки позволяет оценить среднюю абсолютную ошибку между каждым значением в тензоре прогнозируемой и целевой модели. На первом этапе определяется абсолютная разница между данными значениями, после чего вычисляется сумма значений. На втором этапе определяется среднее значение полученной суммы, что является средней абсолютной ошибкой (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n (Y_i - \tilde{Y}_i)^2,$$

где  $Y_i$  – тензор целевых данных объекта (координаты на изображении),  $\tilde{Y}_i$  – тензор целевых данных объекта,  $n$  – количество объектов. В качестве модели оптимизации функции ошибки используется метод стохастического градиентного спуска, подробнее о которой в [9,10].

Основной модуль для определения дорожных объектов включает в себя методы для обучения модели определения и прогнозирования дорожных объектов. Обобщенно, процесс может быть разделен на следующие этапы:

- построение аннотации: в ходе выполнения этапа строится файл с аннотацией, представляющий собой текстовый файл с описанием пути к фотографии, а также координат углов прямоугольника;
- построение модели: в ходе выполнения этапа формируется обучающая модель на основании файла аннотации.
- формирование *backbone*-сети;
- формирование пирамиды признаков;
- формирование подсети классификации;

– формирование подсети регрессии. На рис. 2. представлена блок-схема модели нейронной сети классификации и регрессии.

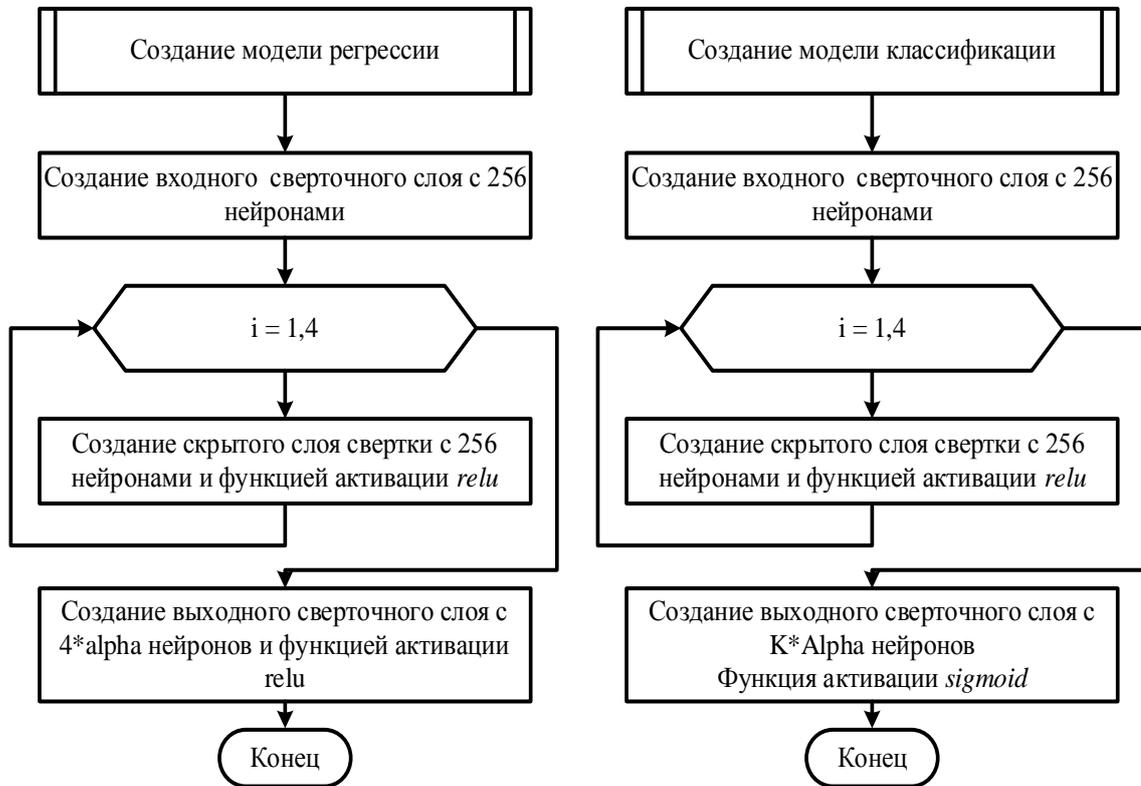


Рис. 2. – Блок-схема архитектуры нейронной сети регрессии и классификации

В ходе построения архитектуры нейронной сети регрессии и классификации были сформированы четыре последовательных слоя свертки изображения с количеством нейронов 256, для каждого слоя сформирован слой максимального пула, а также определена функция активации. В качестве функции активации использована *relu*. В подсети классификации и регрессии в выходном слое использована сверточная сеть с функцией активацией *relu* и *sigmoid* соответственно. Блок-схема построения общей архитектуры пирамиды признаков представлена на рис. 3.

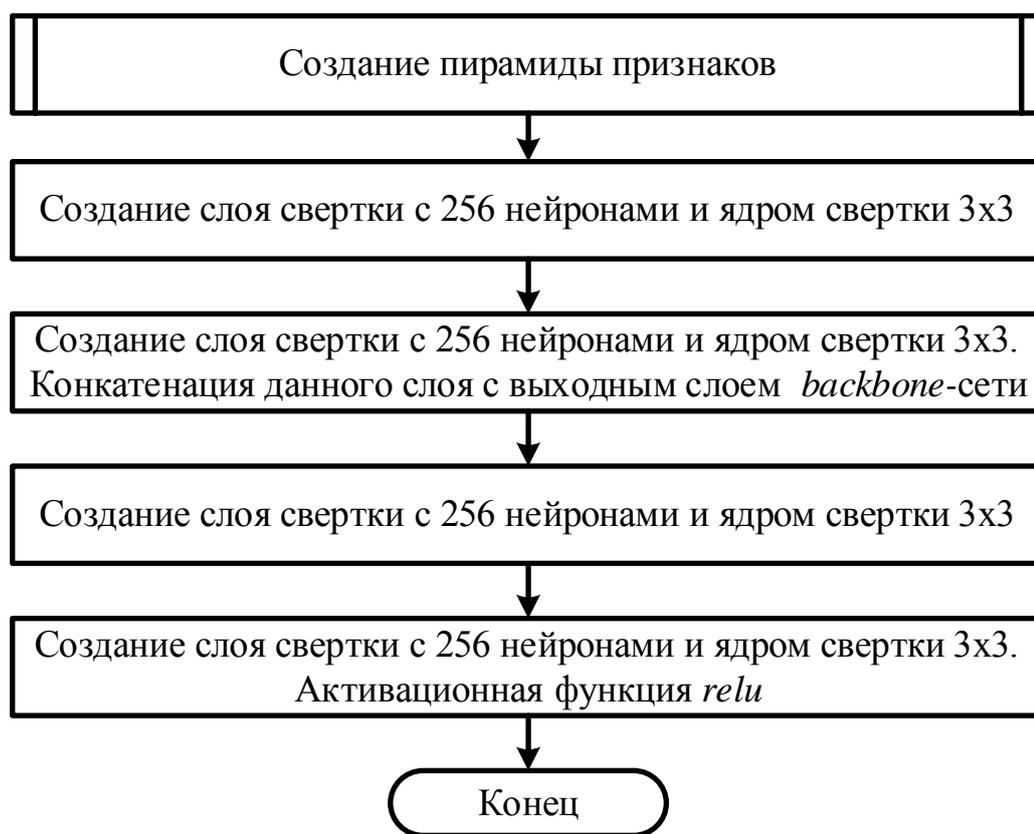


Рис. 3. – Блок-схема архитектуры сети пирамиды признаков

Входные данные модуля тестирования нейронной сети определения дорожных объектов представляют собой изображение и обученную ранее модель. Выходные данные представляют собой изображение с обозначенными объектами.

Динамика изменения функции ошибки представлена на рис. 4.



Рис. 4. – Динамика изменения функции ошибки при определении дорожных объектов

В результате обучения сформирована модель, которая определяет основные объекты дорожного движения. Пример тестирования данной сети представлен на рис. 5.

Дорожные объекты



Рис. 5. – Пример определения дорожных объектов с использованием нейронной сети

Таким образом, проведено проектное моделирование и реализован модуль определения дорожных объектов, который может быть интегрирован в системы автоматического пилотирования транспортным средством.

### Литература

1. Евсин В.А., Широбокова С.Н., Воробьев С.П., Евсина В.А. Моделирование и реализация модуля определения дорожного покрытия для автоматического управления транспортным средством с использованием нейронной сети *U-NET* // Инженерный вестник Дона, 2022, №1. URL: [ivdon.ru/ru/magazine/archive/n1y2022/7424](http://ivdon.ru/ru/magazine/archive/n1y2022/7424).

2. Евсин В.А., Широбокова С.Н., Жжонов В.А., Евсина В.А. Моделирование и реализация процесса распознавания знаков дорожного

движения при определении ситуации на дороге с помощью искусственных нейронных сетей // Инженерный вестник Дона, 2022, №4. URL: ivdon.ru/ru/magazine/archive/n4y2022/7587.

3. Koonce B. ResNet 50. In: Convolutional Neural Networks with Swift for Tensorflow. Apress, Berkeley, 2021. DOI: doi.org/10.1007/978-1-4842-6168-2\_6

4. Николенко С., Кадурын А., Архангельская Е. Глубокое обучение. СПб. Питер, 2018. 480 с.

5. Dean J. Big Data, Data Mining, and Machine Learning. Value Creation for Business Leaders and Practitioners. New Jersey: John Wiley & Sons Limited, 2014. 208 p.

6. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия. Телеком, 2002. 382 с.

7. Kerkhof M., Wu L., Perin G., Picek S. Focus is Key to Success: A Focal Loss Function for Deep Learning-Based Side-Channel Analysis. In: Balasch, J., O’Flynn, C. (eds) Constructive Side-Channel Analysis and Secure Design. COSADE 2022. Lecture Notes in Computer Science, vol 13211. Springer, Cham, 2022. DOI: doi.org/10.1007/978-3-030-99766-3\_2

8. Аггарвал, Ч. Нейронные сети и глубокое обучение. М.: Диалектика, 2020. 752 с.

9. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Горячая линия – Телеком, 2006. 383 с

10. Микеллучи, У. Прикладное глубокое обучение. Подход к пониманию глубоких нейронных сетей на основе метода кейсов. СПб.: БХВ-Петербург, 2020. 368 с.

### References

1. Evsin V.A., SHirobokova S.N., Vorobjov S.P., Evsina V.A. Inzhenernyj vestnik Dona, 2022, №1. URL: ivdon.ru/ru/magazine/archive/n1y2022/7424.



2. Evsina V.A., Evsin V.A., SHirobokova S.N., Zhzhonov V.A. Inzhenernyj vestnik Dona, 2022, №4. URL: [ivdon.ru/ru/magazine/archive/n4y2022/7587](http://ivdon.ru/ru/magazine/archive/n4y2022/7587).
  3. Koonce B. ResNet 50. In: Convolutional Neural Networks with Swift for Tensorflow. Apress, Berkeley, 2021. DOI: [doi.org/10.1007/978-1-4842-6168-2\\_6](https://doi.org/10.1007/978-1-4842-6168-2_6)
  4. Nikolenko S., Kadurin A., Arhangel'skaya E. Glubokoe obuchenie [Deep learning]. SPb.: Piter, 2018. 480p.
  5. Dean J. Big Data, Data Mining, and Machine Learning. Value Creation for Business Leaders and Practitioners. New Jersey: John Wiley & Sons Limited, 2014. 208 p.
  6. Kruglov V.V., Borisov V.V. Iskusstvennye neyronnye seti. Teoriya i praktika. [Artificial neural network. Theory and practice]. M.: Hotline. Telekom, 2002. 382p.
  7. Kerkhof M., Wu L., Perin G., Picek S. Focus is Key to Success: A Focal Loss Function for Deep Learning-Based Side-Channel Analysis. In: Balasch, J., O'Flynn, C. (eds) Constructive Side-Channel Analysis and Secure Design. COSADE 2022. Lecture Notes in Computer Science, vol 13211. Springer, Cham, 2022. DOI: [doi.org/10.1007/978-3-030-99766-3\\_2](https://doi.org/10.1007/978-3-030-99766-3_2)
  8. Aggarval, CH. Neyronnye seti i glubokoe obuchenie [Neural networks and deep learning]. M.: Dialektika, 2020. 752p.
  9. Rutkovskaja D., Pilin'skij M., Rutkovskij L. Neyronnye seti, geneticheskie algoritmy i nechetkie sistemy. [Neural networks, genetic algorithms and fuzzy systems]. M.: Gorjachaja linija. Telekom, 2006. 383 p.
  10. Mikelluchi, U. Prikladnoe glubokoe obuchenie. Podhod k ponimaniyu glubokih neyronnyh setej na osnove metoda kejsov [Applied deep learning. An approach to understanding deep neural networks based on the case method]. SPb.: BHV-Peterburg, 2020. 368p.
-