

Моделирование и реализация модуля определения дорожного покрытия для автоматического управления транспортным средством с использованием нейронной сети *U-NET*

В.А. Евсин, С.Н. Широбокова, С.П. Воробьев, В.А. Евсина

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

Аннотация: В данной статье рассматривается проблематика определения дорожного покрытия для автоматического управления транспортным средством с использованием искусственной нейронной сети. Описано состояние отрасли на текущий момент времени, а также актуальность данных исследований. Представлены входные данные для определения дорожного покрытия. Обосновывается идея применимости метода сегментации изображений для определения дорожного покрытия. Производится формирование структуры искусственной нейронной сети на базе архитектуры U-NET. В частности, визуализируется строение последовательности слоев. Особое внимание уделяется математическому моделированию процесса свертки и максимального пула. Приводится математическая модель процесса обучения искусственной нейронной сети, а так же функций активации: линейной функции и сигмоиды. Предложен алгоритм формирования модели искусственной нейронной сети. Визуализирован на графике процесс обучения данной функции. Представлен результат обучения.

Ключевые слова: искусственные нейронные сети, U-NET, анализ данных, машинное обучение, deep learning, сверточные нейронные сети, свертка, максимальный пул, сегментация изображений, моделирование.

В настоящий момент большое развитие и распространение получают технологии автоматического управления транспортным средством, данной технологией заинтересованы ведущие технологические компании мира, к числу которых относятся: Яндекс, Tesla, *Cognitive Pilot*, *Nvidia*, *Alphabet (Google)* и многие другие. Данная проблема является актуальной, поскольку позволяет повысить качество вождения, что даёт возможность снизить риски аварийных ситуаций. При этом, данная область включает в себя множество различных задач и способов их решения. Основной задачей для реализации автоматического управления транспортным средством является распознавание дорожных объектов, в том числе – дорожного покрытия.

Данная задача относится к классу слабоформализуемых, для решения подобного типа задач наилучшим образом подходит аппарат искусственных

нейронных сетей (далее ИНС). Примеры слабоформализуемых проблематик, в которых уместно решение с использованием ИНС, представлены в [1,2], подробнее об ИНС в [3,4].

Входными данными для определения дорожного покрытия является поток фреймов (изображений) дорожной ситуации перед транспортным средством. Для выделения из полученного изображения дорожного покрытия необходимо сформировать сверточную нейронную сеть (далее СНС), подробнее о данном типе ИНС в [5,6], которая будет выделять маску дорожного покрытия. Для решения поставленной задачи использована архитектура сети *U-NET*, представленная в следующих работах [7,8]. Общая концепция работы данной архитектуры СНС сводится к сегментации изображений, то есть разбиению изображения на отдельные покрывающие его области. Для решения поставленной задачи будет достаточно двух областей: дорожного покрытия и остального изображения. Архитектура сети представляет собой последовательность слоев свертки и максимального пула, подробнее о которых в [9], она пошагово уменьшают размерность изображения, после пытается восстановить маску из сжатого изображения. Архитектура сети *U-NET* представлена на рисунке 1. Данная архитектура состоит из сужающегося и расширяющегося пути. Сужающийся путь представляет собой архитектуру СНС, состоящую из двух слоев свертки с ядром 3×3 , далее следуют активационная функция и операция максимального пула для понижения размерности изображения. Каждый шаг в расширяющемся пути состоит из:

- операции свертки с ядром 2×2 ;
 - конкатенации тензора маски дороги и тензора слоя свертки на соответствующем уровне;
 - операции двух сверток с ядром 3×3 , за которыми следует активационная функция.
-

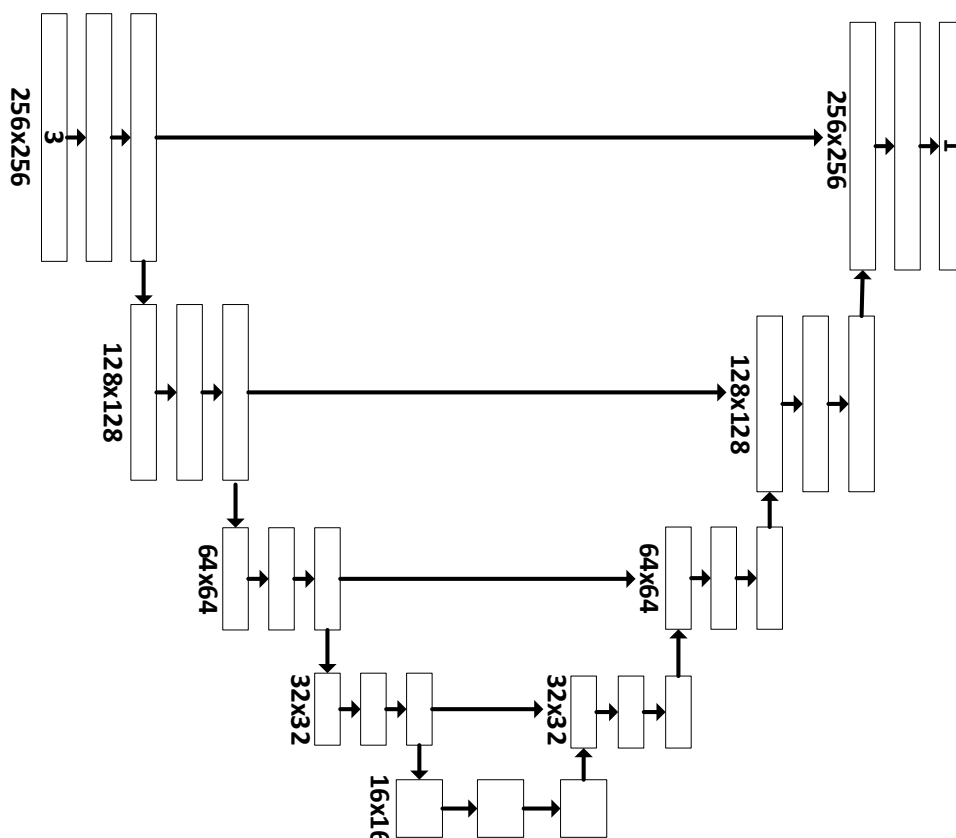


Рис. 1. – Архитектура сети *U-NET*

Формальное определение операции свертки состоит в том, что это процесс, который на входе принимает тензор A размерности $n_A \times n_A$ и ядро свертки K , размерности $n_k \times n_k$ дает на выходе тензор B размерности $n_B \times n_B$, причем:

$$n_B = \left\lfloor \frac{n_A - n_k}{s} + 1 \right\rfloor, \quad (1)$$

где s — шаг сдвига тензора исходного изображения. С учетом того, что основной размерностью ядра для решения данной задачи является ядро 3×3 , каждый элемент рассчитывается следующим образом:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} * \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} = \sum_{i=1}^3 \sum_{j=1}^3 a_{i,j} * k_{i,j}. \quad (2)$$

Размерность тензора входных сигналов для *RGB*-изображения, следующая:

$$PH * PW * 3, \quad (3)$$

где *PH* – количество пикселей изображения в высоту, *PW* – количество пикселей изображения в ширину. Для полученной свертки необходимо провести нелинейное преобразование функцией активации *ReLU*, которая имеет следующую форму:

$$RELU(z) = \begin{cases} z, & \text{если } z \geq 0 \\ 0, & \text{иначе} \end{cases}, \quad (4)$$

где *z* – выходное значение нейрона. Данная функция делает значения каждого нейрона выходного тензора операции свертки неотрицательными. Полученный выходной тензор именуется картой признаков. В ходе проведения операции максимального пула из субтензора размерности ядра выбирается максимальное значение. Для обучения нейронной сети используется метод адаптивной оценки момента (*ADAM*):

$$\begin{aligned} \delta_t &= \nabla_{w_t} J(w_t), \\ m_t &= \beta_1 m_{t-1} + (1 - \beta_1) * \delta_t, \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) \delta_t^2, \\ w_{t+1} &= w_t - \frac{\eta}{\sqrt{v_t + \varepsilon}} * m_t \end{aligned} \quad (5)$$

где δ_t – градиент функции ошибки, $J(w_t)$ – функция ошибки, m_t – момент расчета, w_t – веса модели, η – темп обучения, $\beta_1 \beta_2$ – параметры обучения. В качестве функции ошибки данной модели используется бинарная кроссэнтропия, которая подробно описана в [10], и имеет форму:

$$Loss = -(y \log(p) + (1 - y) \log(1 - p)), \quad (6)$$

где p – это предсказанное значение на выходе модели, y – это соответствующее истинное значение. В качестве функции активации модели используется функция *sigmoid*, которая используется для предсказания

вероятности того, относится ли данный пиксель к дороге или нет. Общая формула данной функции имеет следующую форму:

$$SIGM(z) = \frac{1}{1 + e^{-z}}, \quad (7)$$

где z – выходное значение нейрона. Итоговый алгоритм формирования нейронной сети представлен на рисунке 2.

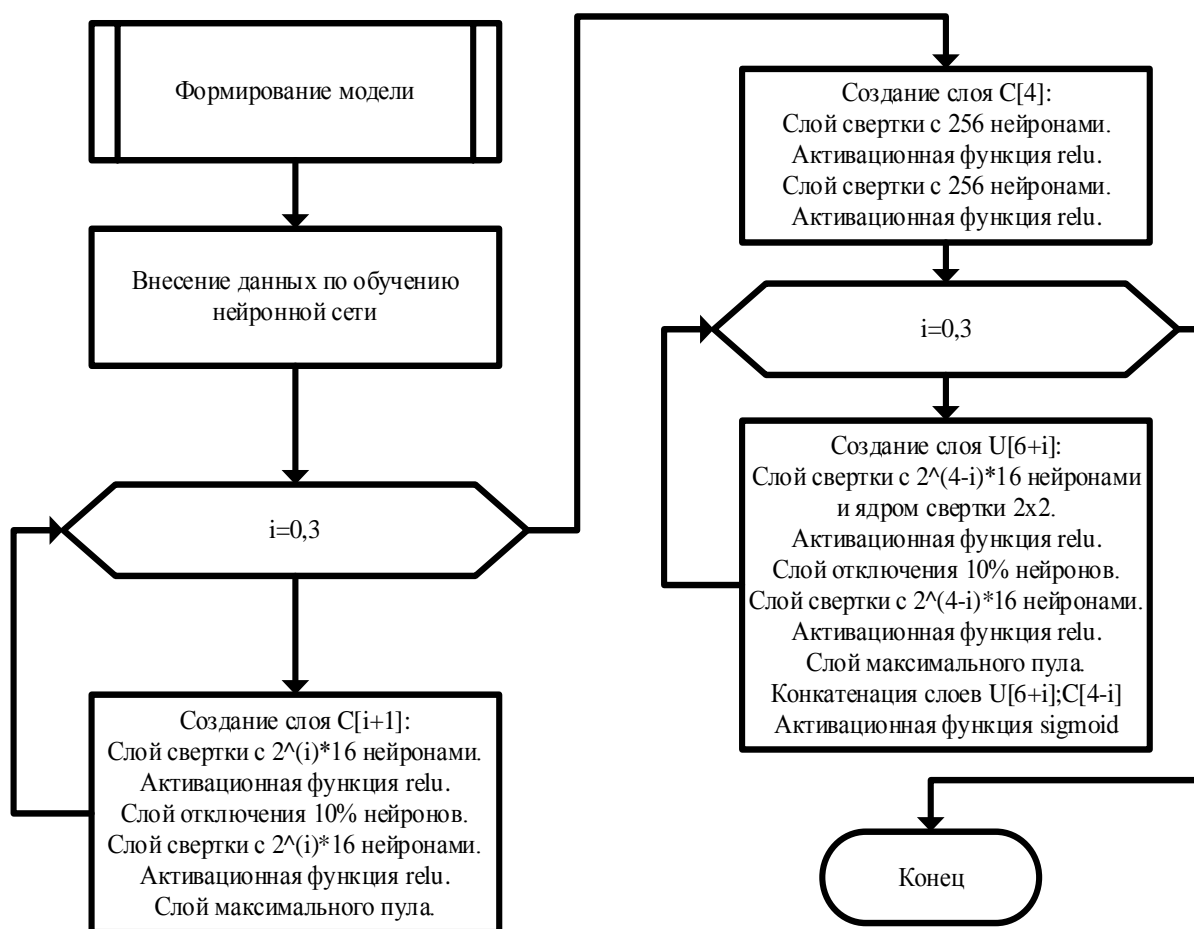


Рис. 2. – Алгоритмизация построения архитектуры сети *U-NET* для определения дорожного покрытия

Для реализации модуля использован фреймворк *Tensorflow* языка программирования *Python*. Модель была обучена на обучающей выборке, включающей различные погодные эффекты, а также различную степень изношенности трассы. Динамика изменения функции ошибки в процессе обучения представлена на рисунке 3.

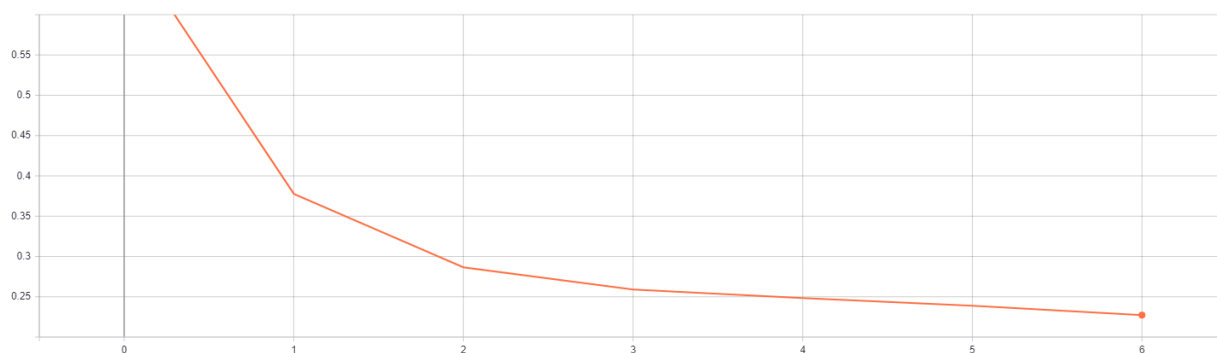


Рис. 3. – Динамика изменения функции ошибки нейронной сети *U-NET*

Тестовое изображение определения маски дорожного покрытия представлено на рисунке 4.

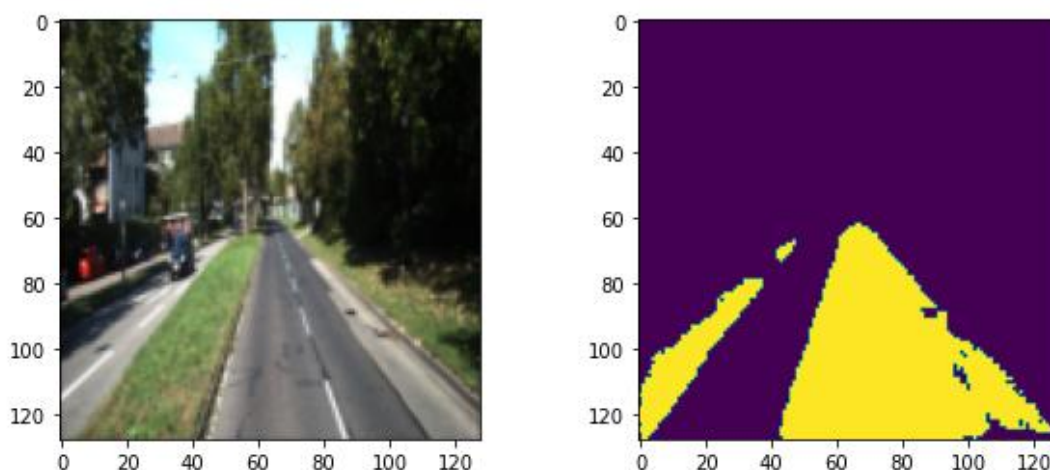


Рис. 4. – Тестовое определение маски дорожного покрытия

Таким образом, получено изображение маски дорожного покрытия. Дальнейшие исследования будут направлены на уточнение данной модели, а также на разработку модуля определения дорожных объектов.

Литература

1. Параскевопуло К.М., Наркевич А.Н. Применение классических нейронных сетей для распознавания злокачественных новообразований на цифровых изображениях кожи // Инженерный вестник Дона, 2021, №5. URL: ivdon.ru/uploads/article/pdf/IVD_31_5_Paraskevopulo.pdf_70f8c97ba7.pdf.

2. Евсин В.А., Евсина В.А., Широбокова С.Н., Продан Е.А. Математическое моделирование процесса оценки стоимости аренды



недвижимости с использованием искусственных нейронных сетей // Инженерный вестник Дона, 2019, №1. URL: ivdon.ru/ru/magazine/archive/n1y2019/5456

3. Хайкин С. Нейронные сети. Полный курс. 2 изд. М.: Издательский дом "Вильямс", 2006. 1104 с.

4. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия. Телеком, 2002. 382 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. Аггарвал, Ч. Нейронные сети и глубокое обучение. М.: Диалектика, 2020. 752 с.

7. Albishri A.A., Shah S.J.H., Kang S.S. AM-UNet: automated mini 3D end-to-end U-net based network for brain claustrum segmentation // Multimed Tools Appl, 2022. DOI: doi.org/10.1007/s11042-021-11568-7.

8. Liu S., Ye H., Jin K. CT-UNet: Context-Transfer-UNet for Building Segmentation in Remote Sensing Images // Neural Process Lett, 2021, № 53. DOI: doi.org/10.1007/s11063-021-10592-w.

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

10. Kline D.M., Berardi V.L. Revisiting squared-error and cross-entropy functions for training neural network classifiers // Neural Comput and Applic, 2005, № 14. DOI: doi.org/10.1007/s00521-005-0467-y.

References

1. Paraskevopulo K.M., Narkevich A.N. Inzhenernyj vestnik Dona, 2021, №5 URL: ivdon.ru/uploads/article/pdf/IVD_31__5_Paraskevopulo.pdf_70f8c97ba7.pdf



2. Evsin V.A., Evsina V.A., SHirobokova S.N., Prodan E.A. Inzhenernyj vestnik Dona, 2019, №1. URL: ivdon.ru/ru/magazine/archive/n1y2019/5456.
3. Hajkin S. Nejronnye seti. Polnyj kurs. [Neural networks. A Comprehensive Foundation]. 2 vol. M.: Williams Publishing House, 2006. 1104p.
4. Kruglov V.V., Borisov V.V. Iskusstvennye nejronnye seti. Teoriya i praktika. [Artificial neural network. Theory and practice.] M.: Hotline. Telekom, 2002. 382p.
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. Aggarval, CH. Nejronnye seti i glubokoe obuchenie [Neural networks and deep learning]. M.: Dialektika, 2020. 752p.
7. Albishri A.A., Shah S.J.H., Kang S.S. Multimed Tools Appl, 2022. DOI: doi.org/10.1007/s11042-021-11568-7.
8. Liu S., Ye H., Jin K. Neural Process Lett, 2021, № 53. DOI: doi.org/10.1007/s11063-021-10592-w.
9. Mikelluchi, U. Prikladnoe glubokoe obuchenie. Podhod k ponimaniyu glubokih nejronnyh 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.
10. Kline D.M., Berardi V.L. Neural Comput and Applic, 2005, № 14. DOI: doi.org/10.1007/s00521-005-0467-y.