

О детектировании трещин в дорожном покрытии с использованием нейронной сети DAUNet

В. С. Половников¹, Д. В. Алексеев², И. В. Виноградов³

В данной работе решается задача сегментации трещин в дорожном покрытии. Предложена сеть Unet-подобной архитектуры с энкодером, представляющим из себя предобученный ResNet18 (без двух классифицирующих слоёв), и стандартным декодером. Показано, что данная архитектура в совокупности с сильной системой аугментации и удачно выбранной функцией потерь способна превзойти существующие решения по используемым в этой области метрикам оценки качества.

Ключевые слова: машинное обучение, нейронные сети, детектирование трещин

1. Введение

Задача попиксельной сегментации трещин дорожного покрытия на изображениях обращала на себя внимание многих исследователей, например [2], [3], [11] и [12]. В последнее время наиболее значимые результаты в этой области были достигнуты с использованием искусственных нейронных сетей. В данной работе показано, что относительно простая архитектура в совокупности с сильной системой аугментации и удачно выбранной функцией потерь способна превзойти существующие решения по точности сегментации.

¹*Половников Владимир Сергеевич* — научный сотрудник каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: polovnikov@intsys.msu.ru

Polovnikov Vladimir Sergeevich — Researcher, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems

²*Алексеев Дмитрий Владимирович* — старший научный сотрудник каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: alekseev@intsys.msu.ru

Alekseev Dmitriy Vladimirovich — Senior researcher, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems

³*Виноградов Иван Владимирович* — аспирант каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: ivan.vinogradoff@gmail.com

Vinogradov Ivan Vladimirovich — PhD student, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems

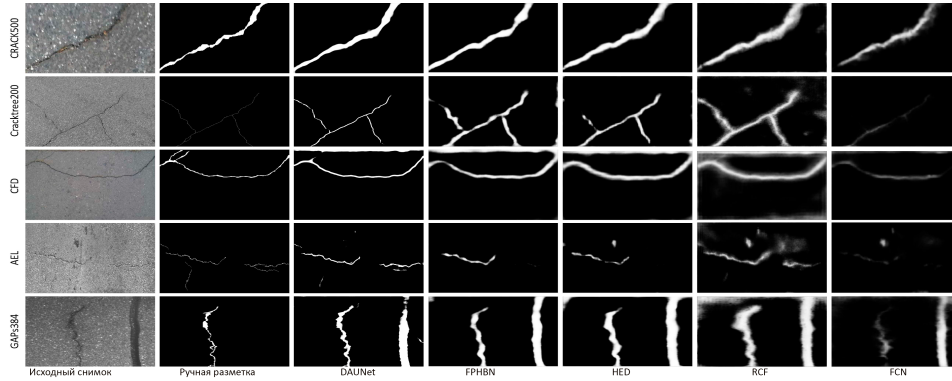


Рис. 1. Сравнение результатов работы различных методов сегментации. Слева направо: исходный снимок, ручная разметка, DAUNet (предложенный), FRHBN([2]), HED([3]), RCF ([11]), FCN([12]).

2. Описание подхода и полученного результата

Была построена сеть Unet-подобной архитектуры с энкодером, представляющим из себя предобученную сеть ResNet18 [1] (без двух классифицирующих слоёв), и стандартным декодером. Обучающим множеством было выбрано CRACK500 [2], состоящее из изображений асфальта и соответствующих бинарных карт трещин.

В качестве функции потерь на первом этапе была взята focal loss [4]. Её значение задаётся формулой

$$FL_{\alpha, \gamma}(Y_{gt}, Y_{pr}) = - \sum (Y_{gt} \cdot \log(Y_{pr})^{\alpha(1-Y_{pr})^{\gamma}} + (1-Y_{gt}) \cdot \log(1-Y_{pr})^{(1-\alpha)Y_{pr}^{\gamma}}),$$

где Y_{gt} – карта трещин, Y_{pr} – предсказание модели, α и γ – параметры, равные в нашем случае 0.25 и 2.

Для устранения размытости границ в предсказаниях трещин сеть была дообучена с использованием dice loss [7], формула которой для истинной карты трещин Y_{gt} и предсказанной карты трещин Y_{pr} дана ниже.

$$DL(Y_{gt}, Y_{pr}) = 1 - \frac{(1 + \beta^2) \cdot tp + \epsilon}{(1 + \beta^2) \cdot tp + \beta^2 \cdot fn + fp + \epsilon},$$

где $tp = \sum(Y_{gt} \cdot Y_{pr})$, $fn = \sum(Y_{gt} \cdot (1 - Y_{pr}))$, $fp = \sum((1 - Y_{gt}) \cdot Y_{pr})$, а β и ϵ – параметры, равные в нашем случае 1 и $1e - 5$. Выбор именно таких функций потерь был обусловлен следующими соображениями: focal loss придаёт большее значение тяжёлым для классификации пикселям, а dice loss по сути содержит в себе F1-метрику, увеличение которой приводит к ужесточению границ.

Обучение велось минибатчами, каждый из которых состоял из случайных кропов изображений обучающего множества. Для лучшей аугментации они были подвергнуты растяжениям, поворотам, отражениям и цветовым искажениям.

Оптимизация на первом этапе велась методом Adam [5] с косинусным уменьшением скорости обучения [6], на этапе дообучения – через стохастический градиентный спуск. Для оценки качества полученного результата использовались метрики AIU, ODS, OIS [2], а также предложенные в данной работе sODS и sOIS. Метрика AIU является усреднённым по всем возможным порогам отношением пересечения предсказанных и истинных областей трещин к их объединению, ODS – максимальным значением F-меры для объединения всех изображений, OIS – средним по всем изображениям значением максимальной F-меры, здесь под максимальной понимается максимальность по порогу бинаризации.

Таблица 1. Сравнение методов сегментации на наборе CRACK500

Methods	AIU	ODS	OIS	sODS	sOIS
DAUNet	0.565	0.676	0.706	0.750	0.731
FPHBN[2]	0.489	0.604	0.635	0.647	0.591
HED [3]	0.481	0.575	0.625	N/A	N/A
RCF [11]	0.403	0.490	0.586	N/A	N/A
FCN [12]	0.379	0.513	0.577	N/A	N/A

Таблица 2. Сравнение методов на наборе CrackTree200

Methods	AIU	ODS	OIS	sODS	sOIS
DAUNet	0.128	0.781	0.805	0.234	0.276
FPHBN[2]	0.041	0.517	0.579	0.095	0.125
HED [3]	0.040	0.317	0.449	N/A	N/A
RCF [11]	0.032	0.255	0.487	N/A	N/A
FCN [12]	0.008	0.334	0.333	N/A	N/A

ODS и OIS предполагают расчёт точности и полноты для скелетонизированных трещин, и за ними стоит алгоритм поиска оптимального паросочетания, подробно описанный в [13]. sODS и sOIS являются аналогичными метриками для пикселей трещин, взятых напрямую после бинаризации, их вычисление обходится без поиска паросочетаний.

Было проведено исследование вкладов составляющих получившейся системы в итоговый результат, выяснилось, что наибольшее влияние в смысле АИУ оказывает аугментация. Её наличие играет большую роль, чем выбор функции потерь (для сравнения рассматривались варианты бинарной кроссэнтропии и мера Жаккара), двухэтапность схемы обучения и предобученность используемой сети ResNet18.

В качестве тестовых множеств были взяты CRACK500 (та его часть, что не была задействована в обучении), CrackTree200 [10], CFD [9], AEL [2] и GAPS384 [8]. Поскольку обучение велось на CRACK500, а остальные множества имеют другие распределения толщин, их изображения были растянуты или сжаты, основываясь на одном изображении для каждого множества.

Разработанный метод сравнивался с FPHBN [2], HED [3], RCF [11] и FCN [12]. Для последних трёх использовались данные, предоставленные авторами, среди которых не было значений sODS и sOIS, поэтому по ним сравнение не проводилось. В оставшихся случаях разработанный метод показал лучшее качество. Ниже приводится сравнение с существующими методами на наборах CRACK500 (таб.1) и CrackTree200 (таб.2).

Список литературы

- [1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition”, *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, 770–778.
- [2] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, “Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection”, *IEEE Trans. Intell. Transp.*, **21**:4 (2020), 1525–1535.
- [3] S. Xie and Z. Tu, “Holistically-Nested Edge Detection,” pp. 1395–1403, ICCV, 2015
- [4] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**:2 (2020), 318–327.
- [5] D. Kingma and J. Lei Ba, “Adam: A Method for Stochastic Optimization”, *arXiv:1412.6980v9*, 2014.
- [6] I. Loshchilov and F. Hutter, “SGDR: Stochastic Gradient Descent with Restarts”, *arXiv:1608.03983*, 2016.
- [7] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, *arXiv:1606.04797*, 2016.
- [8] M. Eisenbach, R. Stricker, D. Seichter, K. Amende, K. Debes, M. Sesselmann, D. Ebersbach, U. Stoeckert and H.-M. Gross, “How to Get Pavement Distress Detection Ready for Deep Learning? A Systematic Approach”, *International Joint Conference on Neural Networks*, 2017, 2039–2047.
- [9] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, “Automatic Road Crack Detection Using Random Structured Forests”, *IEEE Transactions on Intelligent Transportation Systems*, **17**:12 (2016), 3434–3445.

- [10] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, “Cracktree: Automatic Crack Detection from Pavement Images”, *Pattern Recognition Letters*, **33**:3 (2012), 227–238.
- [11] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, and X. Bai, “Richer Convolutional Features for Edge Detection”, *CVPR*, 2017, 5872–5881.
- [12] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation”, *CVPR*, 2015, 3431–3440.
- [13] D. R. Martin, C. C. Fowlkes and J. Malik, “Learning to detect natural image boundaries using local brightness, color, and texture cues”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**:5 (2004), 530–549.

**On pavement crack detection using DAUNet neural network
Polovnikov V.S., Alekseev D.V., Vinogradov I.V.**

This work is devoted to road pavement cracks segmentation problem. Proposed a Unet-like network with a pre-trained ResNet18 as an encoder (without two last classification layers) and a standard decoder is proposed. It is shown that this architecture, together with a strong augmentation system and a well-chosen loss function, is capable of surpassing state-of-the-art solutions in terms of various quality metrics. *Keywords:* machine learning, neural networks, crack detection

References

- [1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition”, *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, 770–778.
- [2] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, “Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection”, *IEEE Trans. Intell. Transp.*, **21**:4 (2020), 1525–1535.
- [3] S. Xie and Z. Tu, “Holistically-Nested Edge Detection,” pp. 1395–1403, ICCV, 2015
- [4] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**:2 (2020), 318–327.
- [5] D. Kingma and J. Lei Ba, “Adam: A Method for Stochastic Optimization”, *arXiv:1412.6980v9*, 2014.
- [6] I. Loshchilov and F. Hutter, “SGDR: Stochastic Gradient Descent with Restarts”, *arXiv:1608.03983*, 2016.
- [7] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, *arXiv:1606.04797*, 2016.
- [8] M. Eisenbach, R. Stricker, D. Seichter, K. Amende, K. Debes, M. Sesselmann, D. Ebersbach, U. Stoeckert and H.-M. Gross, “How to Get Pavement Distress Detection Ready for Deep Learning? A Systematic Approach”, *International Joint Conference on Neural Networks*, 2017, 2039–2047.

- [9] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, “Automatic Road Crack Detection Using Random Structured Forests”, *IEEE Transactions on Intelligent Transportation Systems*, **17**:12 (2016), 3434–3445.
- [10] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, “Cracktree: Automatic Crack Detection from Pavement Images”, *Pattern Recognition Letters*, **33**:3 (2012), 227–238.
- [11] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, and X. Bai, “Richer Convolutional Features for Edge Detection”, *CVPR*, 2017, 5872–5881.
- [12] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation”, *CVPR*, 2015, 3431–3440.
- [13] D. R. Martin, C. C. Fowlkes and J. Malik, “Learning to detect natural image boundaries using local brightness, color, and texture cues”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**:5 (2004), 530–549.