УДК 004.932.2

Д.В. Дементьев, Р.А. Савинов, М.Н. Суетин, А.Г. Подлобошников

СИСТЕМА РАСПОЗНАВАНИЯ ПОВРЕЖДЕНИЙ МЕТАЛЛИЧЕСКИХ КОНСТРУКЦИЙ¹

Дементьев Виталий Евгеньевич, доктор технических наук, окончил экономико-математический факультет Ульяновского государственного технического университета. Заведующий кафедрой «Телекоммуникации» УлГТУ. Имеет монографии, статьи и патенты в областях статистической обработки многомерных сигналов и защиты информации. [e-mail: dve@ulntc.ru].

Савинов Руслан Анатольевич, окончил машиностроительный факультет УлГТУ, ведущий инженер-программист ФНПЦ АО «НПО «Марс». Область научных интересов – разработка автоматизированных систем распознавания и анализа образов с применением нейросетевых технологий. [e-mail: sarus@list.ru].

Суетин Марат Николаевич, аспирант, окончил факультет технологии и предпринимательства Ульяновского государственного педагогического университета. Ведущий инженер-программист ФНПЦ АО «НПО «Марс». Имеет статьи в областях обработки изображений и компьютерного зрения. [e-mail: source81@gmail.com].

Подлобошников Анатолий Геннадиевич, окончил факультет информационных систем и технологий Ульяновского государственного технического университета. Начальник научно-исследовательской лаборатории – главный конструктор ФНПЦ АО «НПО «Марс». Область научных интересов – разработка автоматизированных систем распознавания и анализа образов с применением нейросетевых технологий. [e-mail: agpdonet@mail.ru].

Аннотация

На основе нейросетевого подхода разработана система, позволяющая выявлять внешние дефекты и повреждения объектов инфраструктуры на фото- и видеоизображениях. В качестве обучающего материала для сверточной нейронной сети UNet использовались фото- и видеоизображения конструктивных элементов железнодорожных мостов. Выполнена разметка фото- и видеоизображений, подобрана оптимальная архитектура и гиперпараметры нейронной сети. На наборе тестовых изображений выполнен контроль полученной при обучении нейронной сети. Результаты позволяют сделать вывод, что с помощью полученной системы можно производить поиск дефектов на элементах стальных конструкций на уровне, показываемом в ходе визуального контроля экспертами.

Ключевые слова: неинвазивный контроль, повреждения металлических конструкций, глубокая сверточная нейронная сеть, CNN, мостовые конструкции, трещины, обработка фотоизображений.

doi: 10.35752/1991-2927-2021-2-64-40-45

THE SYSTEM OF DAMAGE RECOGNITION IN METAL STRUCTURES

Vitalii Evgenevich Dementev, Doctor of Sciences in Engineering; graduated from the Faculty of Economics and Mathematics at Ulyanovsk State Technical University; Head of the Department of Telecommunications at UISTU; an author of monographs, articles, and patents in the field of statistical processing of multidimensional signals and information protection. e-mail: dve@ulntc.ru.

Ruslan Anatolevich Savinov, graduated from the Faculty of Machine-Building at Ulyanovsk State Technical University; Leading Software Engineer at Federal Research-and-Production Center Joint Stock Company 'RPA 'Mars'; his research interests are in the field of the development of computer-aided recognition systems and the pattern analysis with neural network technologies. e-mail: sarus@list.ru.

Marat Nikolaevich Suetin, Postgraduate Student; graduated from the Faculty of Technology and Business at Ulyanovsk State Pedagogical University; Leading Software Engineer at FRPC JSC 'RPA 'Mars', an author of articles in the field of image processing and computer vision. e-mail: source81@gmail.com.

¹ Исследование выполнено при финансовой поддержке РФФИ и Правительства Ульяновской области в рамках научного проекта № 19-29-09048.

Anatolii Gennadevich Podloboshnikov, graduated from the Faculty of Information Systems and Technologies of Ulyanovsk State Technical University; Head of a research-and-development laboratory, Chief Designer at FRPC JSC 'RPA 'Mars'; his research interests are in the field of computer-aided recognition and pattern analysis systems development with neural network technologies. e-mail: agpdonet@mail.ru.

Abstract

The article deals with a system developed on the basis of a neural network approach that allows the system to detect the visual defects and damage of infrastructure facilities using photo/video data processing. The photo/video images of structural railroad bridge members were used as training data for the convolutional neural network U-Net. The authors carried out ranging and labeling the photo/video data sets as well as they selected an optimal architecture and hyperparameters for the neural network. The image test sets were used for testing the neural network trained. The findings suggest that the resulting system can detect defects in structural steel members at the level which can be achieved if experts performed the visual inspection.

Keywords: noninvasive inspection, damage in metal structures, deep convolutional neural network, CNN, bridge structures, cracks, photo image processing.

ВВЕДЕНИЕ

Важнейшая задача предприятий, имеющих в своих активах искусственные сооружения, – это мониторинг их состояния. Примером таких сооружений являются железнодорожные мосты. В данной статье рассматриваются железнодорожные мостовые сооружения. Мосты подвергаются существенным динамическим воздействиям, которые приводят к появлению тех или иных повреждений (повреждение – неисправность, появившаяся во время эксплуатации сооружения). Развитие повреждений влечет за собой частичное или полное разрушение элементов конструкций, что приводит к аварийным и катастрофическим последствиям. Своевременное обнаружение, мониторинг состояния и локализация (устранение) повреждений – основная задача служб, контролирующих такие сооружения. Контроль и мониторинг мостовых сооружений требует привлечения высококвалифицированного персонала, что сопряжено со значительными финансовыми и организационными проблемами.

Решением в сложившейся ситуации является использование перспективных систем неинвазивного мониторинга, предполагающих минимальное участие человека. В настоящей статье рассматривается система, позволяющая в автоматизированном режиме обрабатывать фото- и видеоматериалы, выявлять и размечать повреждения.

В последние годы значительное число исследований в области поиска и выделения объектов на изображениях было сосредоточено на применении нейросетевого подхода к распознаванию [1–4]. Данный подход показал наибольшую эффективность по сравнению с другими процедурами. Применение при распознавании объектов нейронных сетей требует меньшего числа настроек и хорошо подходит для решения широкого спектра задач.

В то же время решение задачи семантической сегментации изображений с помощью нейросетевого

подхода требует создания обучающих выборок большого объема. Найти подходящий готовый набор изображений для решения специализированной задачи, как правило, не представляется возможным. Разметка обучающих и тестовых выборок, в свою очередь, требует привлечения экспертов [5]. Кроме того, для работы нейросетевых алгоритмов в режиме реального времени требуется прибегать к применению специальных ухищрений [6].

В данной работе представлена специально подобранная архитектура нейронной сети для решения задачи семантической сегментации дефектов на поверхности стальных конструкций. Изложены принципы разметки изображений при подготовке обучающего и тестового наборов изображений. Представлены результаты обучения и тестирования на основной и усиленных выборках. Выполнен анализ полученных результатов. Нарботки, сделанные в ходе исследования, используются как составные элементы системы автоматического мониторинга дефектов и повреждений на поверхности элементов стальных конструкций.

АРХИТЕКТУРА НЕЙРОННОЙ СЕТИ

По ряду соображений в качестве наиболее подходящей архитектуры для выполнения семантической сегментации дефектов на поверхности материала была выбрана нейронная сеть UNet. Модель UNet часто используется для сегментации медицинских изображений, где выполняет поиск аномалий в тканях и внутренних органах человека, и для обработки спутниковых фотоизображений с целью выделения крупных объектов [7–9]. Примечательно, что UNet требует для получения приемлемого результата в процессе обучения сравнительно небольшой объем входных изображений.

Нейронная сеть UNet, созданная для решения задачи семантической сегментации, делится на нисходящую (шифровальщик) и восходящую (дешифровальщик) части (рис. 1). Шифровальщик состоит из повторного при-

менения двух сверток 3×3 , за которыми следует активация ReLU и операция субдискретизации 2×2 степени 2 для понижения разрешения. На каждом этапе понижающей дискретизации каналы свойств удваиваются. Для первичной инициализации сверточных слоев шифровальщика используются параметры, полученные путем применения алгоритма Kaiming. Каждый шаг в дешифровальщике состоит из операции повышающей дискретизации карты свойств, за которой следуют: свертка 2×2 , которая уменьшает количество каналов свойств; объединение с соответствующим образом обрезанной картой свойств из шифровальщика; две свертки 3×3 , за которыми следует ReLU.

Рис. 1. Архитектура UNet

При обучении используется стандартный алгоритм оптимизации Adam. В качестве функции потерь применяется бинарная кросс-энтропия:

$$BCE(p, y) = -(y \ln p + (1 - y) \ln(1 - p)),$$

где $y = 0$ – для пикселей фона и $y = 1$ – для пикселей объекта, p – вероятность того, что пиксель принадлежит объекту.

Для устранения эффекта переобучения, применяются процедуры аугментации [10]: горизонтальное отра-

жение и поворот изображения. С целью нивелирования последствий дисбаланса пикселей фона и объектов на изображениях используется процедура плиточного разбиения изображения (Tiling).

ПОИСК И РАСПОЗНАВАНИЕ ПОВРЕЖДЕНИЙ

Для поиска и распознавания повреждений металлоконструкций (мостов) было предложено, помимо непосредственного выявления трещин (рис. 2), выявлять потенциально опасные места (области внимания) появления трещин, распознавания специальных маркировок и контроля изменения состояния повреждений во времени.

Рассмотрим решение задачи поиска повреждений (трещин) на изображениях пролетных строений металлоконструкций мостовых сооружений. Для решения задачи поиска повреждений предложено использовать глубокую сверточную нейронную сеть (CNN) с архитектурой UNet, которая позволяет определить не только класс изображения целиком, но и сегментировать его области по классу, т. е. создать маску, которая будет разделять изображение на несколько классов. Архитектура состоит из стягивающего пути для захвата контекста и симметричного расширяющегося пути, который позволяет осуществить точную локализацию. Начальной выборкой послужило несколько сотен фотографий, содержащих маркировочные знаки, области внимания и различные дефекты в элементах металлических конструкций.

Для организации процесса обучения нейронной сети, с целью проведения последующей валидации, все имеющиеся в наличии фотографии были поделены на обучающую и тестовую выборки (в соотношении 80%–20%). Была выполнена разметка трещин в виде бинарных изображений, на которых трещины выделены белым цветом, а фон – черным (рис. 3).

Для проведения эксперимента, связанного с исследованием влияния объема обучающей выборки на результат обучения, была произведена нарезка областей, содержащих фрагменты трещин, из исходных фотографий и на их основе подготовлены три набора данных,

Рис. 2. Примеры дефектов

Рис. 3. Формирование обучающей выборки при распознавании дефектов

содержащих соответственно 342/509/842 изображения. Обучение нейронной сети проводилось на исходном и на усиленных наборах данных, кроме того была задействована аугментация для внесения максимального разнообразия в обучающую выборку. Результаты обучения были проверены на тестовом наборе изображений, не участвующем в обучении, и представлены в таблице 1 в виде минимального значения функции ошибки / числа эпох, прошедших до получения минимального значения функции ошибки.

Для получения наилучшего результата была применена технология, связанная с последовательным уменьшением шага градиентного спуска в течение обучения. 75 эпох обучение проводилось с шагом 10-2, 30 эпох – с шагом 10-3 и еще 30 эпох – с шагом 10-4.

На рисунке 4 представлены зависимости функции потерь от числа эпох для случаев использования оригинальных (рис. 4а) и аугментированных выборок (рис. 4б). Заметно, что второй вариант обеспечивает более контролируемый вариант сходимости.

а

б

Рис. 4. Сходимость функции потерь при обучении нейронной сети

Таблица 1

Коэффициент Сёренсена DSC
для различных обучающих выборок

Набор изображений (число изображений)	Без аугментации	С аугментацией
Исходный (175)	0,43	0,37
Усиленный № 1 (342)	0,56	0,49
Усиленный № 2 (509)	0,6	0,55
Усиленный № 3 (842)	0,62	0,6

Для сокращения числа ложных срабатываний и повышения точности сегментации был предложен второй подход к детектированию дефектов, который заключается в предварительном выделении областей внимания нейронной сетью Darknet [11, 12], с последующей их сегментацией, обученной во время реализации первого подхода нейронной сетью UNet. Для обучения нейронной сети Darknet выделению областей внимания для изображений из обучающей выборки была выполнена разметка в стиле открытого набора изображений Pascal VOC.

Тестирование нейронных сетей выполнялось на двух наборах изображений элементов стальных конструкций с дефектами в каждом, но с разных объектов (рис. 5).

Изображения из наборов для тестирования не использовались в процессе обучения нейронной сети. В ходе тестирования оценивались процент фактического успешного обнаружения дефекта, процент ложного срабатывания, а также коэффициент Сёренсена как мера сходства результатов детектирования и сделанной вручную разметки (табл. 2).

Очень важным результатом является относительно низкий результат ложных срабатываний разработанного алгоритма. Также в ходе работы были получены

Таблица 2

Параметры эффективности сегментации дефектов

	Аугментация	Контрольная выборка №1			Контрольная выборка №2		
		% фактического успешного обнаружения дефекта	% ложных срабатываний	Коэффициент Сёренсена	% фактического успешного обнаружения дефекта	% ложных срабатываний	Коэффициент Сёренсена
Только U-Net	Нет	0,76	0,06	0,16	0,7	0,06	0,11
	Да	0,59	0,07	0,09	0,5	0,07	0,06
Darknet + U-Net	Нет	0,81	0,01	0,71	0,75	0,01	0,5
	Да	0,76	0,01	0,57	0,77	0,02	0,4

результаты, свидетельствующие о хорошей восприимчивости полученных нейронных сетей к дообучению на блоках, новых данных. Это дает обоснованную надежду на востребованность алгоритма на практике.

В случае, если на фотографии присутствуют изображения и маркера, и дефекта (трещины), то появляется возможность оценки геометрических размеров этой трещины. Действительно, имея результат выделения маркера на изображении в виде некоторой пространственной области (четырехугольника), а также зная истинные значения этого маркера, несложно с помощью простых геометрических преобразований (а именно пространственных пропорций) рассчитать под каким углом ведется съемка и истинные размеры дефекта. При этом для простоты в работе считается, что дефект располагается на той же плоскости, что и маркер. На рисунке 6 представлен финальный результат работы программы. На изображении найден маркер, идентифицирована цифра на этом маркере («Sign:3»), детектирована трещина и определен ее тип (RoI: T6), выделена область, которая соответствует этой трещине, и рассчитаны ее размеры («1,2x3,2»).

ЗАКЛЮЧЕНИЕ

В данной статье рассмотрено решение задачи поиска повреждений (трещин) на изображениях пролетных строений металлоконструкций мостовых сооружений.

Рис. 5. Примеры изображений из контрольной выборки

Для решения задачи использовалась глубокая сверточная нейронная сеть (CNN) с архитектурой UNet. Нейронная сеть UNet предварительно обучалась и тестировалась на заранее подготовленных наборах изображений с применением процедур аугментации.

Для повышения качества семантической сегментации дефектов и уменьшения числа ложных срабатываний применялась предварительная обработка входных изображений другой нейронной сетью Darknet с целью предварительного выделения областей интереса на изображениях. Области интереса представляют собой, как правило, места стыковых соединений металлоконструкций, в которых с высокой вероятностью может появиться трещина. Нейронная сеть Darknet обучалась на предварительно подготовленном наборе изображений типа Pascal VOC. После проведения такой предобработки, семантическая сегментация дефектов нейронной сетью UNet производится уже не на целом изображении в выявленной области внимания, что значительно улучшает качество сегментации трещин.

В ходе тестирования системы мониторинга были выявлены некоторые проблемы, связанные с недостаточным объемом обучающей выборки, которые проявляются в пропусках трещин других типов, не входящих в текущий набор изображений для обучения нейрон-

Рис. 6. Финальный результат работы программы

ной сети. Кроме того, обнаружено значительное число ложных срабатываний в том случае, если поиск дефектов ведется без предварительного выделения областей внимания, то есть на целом изображении. Выявленные проблемы требуют решения в ходе последующих исследований путем увеличения объема и качества обучающих выборок, а также применения эффективных процедур оптимизации работы нейронной сети UNet с целым изображением.

Тем не менее, полученные результаты свидетельствуют о практической возможности применения разработанных процедур в задачах автоматического или автоматизированного мониторинга состояния металлоконструкций на основе обработки их фотоизображений.

СПИСОК ЛИТЕРАТУРЫ

1. OverFeat: Integrated recognition, localization and detection using convolutional networks / P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun. URL: <https://arxiv.org/pdf/1312.6229.pdf> (дата обращения: 21.05.2021).
2. Handwritten digit recognition with a backpropagation neural network / Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R. Hovard, W. Hubbard, L. Jackel // *Advances in Neural Information Processing Systems*. 1990. Vol. 2. pp. 396–404.
3. LeCun Y., Huang F.-J., Bottou L. Learning methods for generic object recognition with invariance to pose and lighting // *Proceedings of Computer Vision and Pattern Recognition*. 2004. Vol. 2. pp. 97–104.
4. Tang Y. Deep learning using linear support vector machines. URL: <https://arxiv.org/pdf/1306.0239.pdf> (дата обращения: 21.05.2021).
5. Biau G., Cadre B., Rouviere L. Accelerated gradient boosting. URL: <https://arxiv.org/pdf/1803.02042.pdf> (дата обращения: 21.05.2021).
6. Ye L., Liu Z., Wang Y. Learning semantic segmentation with diverse supervision. URL: <https://arxiv.org/pdf/1802.00509.pdf> (дата обращения: 21.05.2021).
7. Ronneberger O., Fischer F., Brox T. U-Net: Convolutional networks for biomedical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2015. Vol. 9351. pp. 1097–1105.
8. Unet 3+: A full-scale connected Unet for medical image segmentation / H. Huang, L. Lin, H. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, J. Wu. URL: <https://arxiv.org/pdf/2004.08790.pdf> (дата обращения: 21.05.2021).
9. DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation / R. Li, W. Liu, L. Yang, S. Sun, W. Hu, F. Zhang, W. Li. URL: <https://arxiv.org/pdf/1709.00201.pdf> (дата обращения: 21.05.2021).
10. Random erasing data augmentation / Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang. URL: <https://arxiv.org/pdf/1708.04896.pdf> (дата обращения: 21.05.2021).
11. You Only Look Once: unified, real-time object detection / J. Redmon, S. Divvala, R. Girshick, F. Farhadi. URL: <https://arxiv.org/pdf/1506.02640.pdf> (дата обращения: 21.05.2021).
12. Redmon J., Farhadi A. YOLO9000: better, faster, stronger. URL: <https://arxiv.org/pdf/1612.08242.pdf> (дата обращения: 21.05.2021).

REFERENCES

1. Sermanet P., Eigen D., Zhang X., Mathieu M., Fergus R., LeCun Y. *OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks*. Available at: <https://arxiv.org/pdf/1312.6229.pdf> (accessed 21.05.2021).
2. LeCun Y., Boser B., Denker J.S., Henderson D., Hovard R., Hubbard W., Jackel L. Handwritten Digit Recognition with a Backpropagation Neural Network. *Advances in Neural Information Processing Systems*, 1990, vol. 2, pp. 396–404.
3. LeCun Y., Huang F.-J., Bottou L. Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting. *Proceedings of Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 97–104.
4. Tang Y. Deep Learning Using Linear Support Vector Machines. Available at: <https://arxiv.org/pdf/1306.0239.pdf> (accessed 21.05.2021).
5. Biau G., Cadre B., Rouviere L. *Accelerated Gradient Boosting*. Available at: <https://arxiv.org/pdf/1803.02042.pdf> (accessed 21.05.2021).
6. Ye L., Liu Z., Wang Y. *Learning Semantic Segmentation with Diverse Supervision*. Available at: <https://arxiv.org/pdf/1802.00509.pdf> (accessed 21.05.2021).
7. Ronneberger O., Fischer F., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, vol. 9351, pp. 1097–1105.
8. Huang H., Lin L., Tong H., Hu H., Zhang Q., Iwamoto Y., Han X., Chen Y.-W., Wu J. *Unet 3+: A Full-Scale Connected U-Net for Medical Image Segmentation*. Available at: <https://arxiv.org/pdf/2004.08790.pdf> (accessed 21.05.2021).
9. Li R., Liu W., Yang L., Sun S., Hu W., Zhang F., Li W. *DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation*. Available at: <https://arxiv.org/pdf/1709.00201.pdf> (accessed 21.05.2021).
10. Zhong Z., Zheng L., Kang G., Li S., Yang Y. *Random Erasing Data Augmentation*. Available at: <https://arxiv.org/pdf/1708.04896.pdf> (accessed 21.05.2021).
11. Redmon J., Divvala S., Girshick R., Farhadi F. *You Only Look Once: Unified, Real-Time Object Detection*. Available at: <https://arxiv.org/pdf/1506.02640.pdf> (accessed 21.05.2021).
12. Redmon J., Farhadi A. *YOLO9000: Better, Faster, Stronger*. Available at: <https://arxiv.org/pdf/1612.08242.pdf> (accessed 21.05.2021).