

Г. К. Букалов, д-р техн. наук, проф., e-mail: gk.bukalov44@yandex.ru,
А. О. Бурыгин, аспирант, e-mail: g.t.m.p@yandex.ru,
И. Г. Панин, д-р техн. наук, проф., e-mail: igpanin@list.ru,
А. В. Торцев, аспирант, e-mail: fullfulk47458@gmail.com,
Костромской государственной университет

Модификация метода полносвязных сверточных сетей (FCN) для поиска редко встречающихся дефектов на больших площадях

Рассматривается задача нахождения дефектов на изображениях большой площади, для чего изображение проходит несколько этапов: создание сверточной сети U-Net, извлечение признаков U-Net, классификация алгоритмом Random Forest и выявление дефектных областей. С помощью классификатора на основе Random Forest проводится сегментация частей входного изображения. Проведены вычислительные эксперименты, направленные на исследование эффективности предложенного метода в сравнении с существующими методиками.

Ключевые слова: стропа, изображение, CNN, U-Net, MSER, Random Forest

Введение

У объектов большой площади могут возникать дефекты различного рода. В качестве примера таких объектов можно взять текстильные стропы, дефекты которых могут быть получены как при их производстве, так и во время эксплуатации. Дефекты любого типа значительно сокращают срок службы строп, потенциально приводя к отказам и авариям, поэтому задача их постоянной проверки является актуальной для любого производства. Она заключается, в основном, во внешнем осмотре строп, который хотя и проводится квалифицированными специалистами, но по-прежнему, является непоследовательным, субъективным, утомительным и трудоемким. Для того чтобы сократить рабочую нагрузку инспекторов и обеспечить объективный и эффективный осмотр строп, можно использовать автоматизированные системы инспектирования строп. Но задача нахождения дефектов на изображениях большой площади, в том числе и на изображениях строп, сложна из-за своей специфики, поскольку, во-первых, дефекты (особенно новые) достаточно редки и, во-вторых, не всегда легко получить большое число примеров для обучения классификатора/детектора объектов.

Первоначальная идея с использованием полносвязных сверточных сетей (далее FCN), например U-Net [1] (далее U-сеть), была наиболее

подходящей для поиска дефектов. Однако в процессе исследования было выявлено, что если существует ограничение на число примеров площадей (тканевого полотна) с повреждениями, такие методы работают плохо из-за специфики больших площадей. Тем не менее, в данной работе показано, что алгоритм типа "случайный лес" (далее RF) [2], обученный классифицировать каждый пиксель на основе выборки изображений, достигает гораздо лучшего результата, чем обычный конечный слой Softmax в U-сети.

Для проверки больших областей изображений необходимо их сначала нормализовать, а затем, так как нельзя пропустить через нейронную сеть весь датасет (все множество входных фотографий) разом, необходимо разделить датасет на батчи (пакеты, сетки или партии) участков изображений, которые затем и подаются в U-сеть. U-сеть обучается только на батчах, отобранных вокруг дефектов в обучающем наборе, чтобы избежать переполнения системы нормальными (бездефектными) пикселями, которых подавляющее большинство.

В предлагаемой статье представлены разработки в данном направлении:

1. Рассмотрено описание системы для идентификации небольших областей дефектов на больших площадях изображений.

2. Показано, что использование "случайного леса" значительно улучшает производительность U-сети (в сравнении с Softmax).

3. Показано, что ложные срабатывания могут быть уменьшены с помощью создания максимально устойчивых экстремальных областей MSER [3] для идентификации регионов-кандидатов, а не простого порогового значения.

1. Методология идентификации областей дефектов на изображении

Классические подходы к определению дефектов реализуются с использованием либо базовых методов обработки изображений, либо традиционных классификаторов. Методы сверточной нейронной сети CNN [4], в отличие от классических, показали хорошую производительность для многих современных задач компьютерного зрения. Теоретически обучение CNN требует большого числа маркированных примеров, но если использовать полносвязанный слой FC (слой, на котором каждая характеристика входного вектора имеет влияние на каждую характеристику выходного вектора) предварительно обученной CNN, то он может быть выбран в качестве универсального экстрактора признаков [5]. В своей работе Р. Рэн [6] использовал признаки FC, извлеченные из набора батчей изображений, для обучения классификатора, который может использоваться для прогнозирования содержания батчами дефектов. Необходимо отметить, что извлечение признаков FC из батчей занимает достаточно продолжительное время, что ограничивает целесообразность применения метода в реальных задачах контроля.

Известно, что для обучения нейронных сетей требуется большое количество исходных данных, но для некоторых областей знаний такой набор по различным причинам не может быть получен. Для решения этой проблемы О. Ронненбергом и др. [1] была разработана новая архитектура CNN — U-сеть, содержащая сокращающийся путь, кодирующий контекстную информацию путем повторного применения сверток частей изображения, и симметричный расширяющийся путь, который захватывает локальные данные.

Исходя из этого в данной работе делается предположение, что существует обучающий набор, содержащий примеры изображений площадей (строп) без дефектов и площадей (строп), содержащих типичные дефекты, а также бинарные изображения меток, указывающие пиксели, принадлежащие каждому дефекту.

Поскольку дефекты встречаются редко, то на первом этапе необходимо обучить U-сеть, для чего реализуются следующие шаги:

- каждое изображение дополняется вращениями, отражениями и другими D4-аугментациями;

- изображения вместе с дополнениями нормализуются с помощью линейного преобразования с математическим ожиданием, стремящимся к нулю, и дисперсией, стремящейся к единице;
- вокруг каждого дефекта дополнительно формируются батчи изображений;
- используя полученный набор батчей, U-сеть обучается предсказывать метки двоичных пикселей;
- обучается RF классификатор на использование характеристик, извлеченных в каждой позиции пикселя в U-сеть;
- в вероятностном изображении вычисляются максимально устойчивые экстремальные области (MSERs), и каждая область анализируется для исключения ложных срабатываний.

На втором этапе создается обучающий набор векторов признаков путем извлечения векторов, связанных с каждым положительным пикселем (т.е. пикселем входящего в состав дефекта) и равным числом (случайно отобранных) отрицательных пикселей. Эти выборки используются для обучения классификатора случайных деревьев на основе вектора пространственных характеристик. При появлении нового изображения выходное вероятностное изображение создается путем применения к нему U-сети и дальнейшей оценки вероятности дефекта в каждом пикселе с помощью классификатора RF.

На третьем этапе проводится отбор нужных областей изображений, для чего используется механизм детектора максимально устойчивой экстремальной области (MSER), с помощью которого выбираются "дефектные" пиксели из каждого вероятностного изображения. Этот метод использует характеристики более высокого уровня, закодированные в связанных компонентах, что дает преимущество по сравнению с методом расчета порога изображения. Изображение сначала линейно растягивается до диапазона [0, 255], затем, после прохождения через детектор MSER, выделяется набор "дефектных" кандидатов для каждого тестового изображения. Чтобы оценить, является ли область истинно положительной, для нее вычисляется число пикселей n_p , которые имеют вероятность определения как "дефектных" выше некоторого порогового значения k , а любая область, имеющая вероятность ниже порогового значения k , удаляется.

Чтобы выбрать наилучшие параметры U-сети, необходимо увеличить объем данных, что также увеличивает эффективность методов глубокого обучения [7]. В этом случае применяется конвейер операций аугментаций к изображениям и связанным с ними меткам (операции сдвига, перекоса, переворачивания и упругого искажения), а также Z-нормализация к каждому изо-

бражению, что увеличивает его контрастность и улучшает инвариантность к эффектам освещения. К каждому пикселю применяется преобразование $p' = (p - \mu)/\sigma$, где μ и σ — среднее и стандартное отклонение соответственно, вычисленные в некоторой области размером $W \times W$. Любой блок размером $P \times P$ вокруг каждого дефекта произвольного изображения выбирается с произвольным смещенным положением так, чтобы положение дефекта не было смещено по центру. Для этого используется оптимизатор Adam [8] для оптимизации параметров в U-сети.

2. Проведение экспериментов по определению оптимальных параметров

В качестве обучающих примеров был создан набор из 437 изображений участков строп и связанных с ними изображений этикеток, каждое из которых имело, по крайней мере, один дефект. Для всех экспериментов использовалась 10-кратная схема перекрестной валидации, суть которой состоит в разделении входного батча изображений на 10 частей и дальнейшей попеременной смене их ролей в качестве тестовых и тренировочных батчей. При обучении одной U-сети использовались батчи размером 64×64 , отобранные как из оригинальных, так и из дополнительных изображений, а при тестировании — только из исходных изображений. Все изображения (включая изображения, используемые для выборки батчей) были обработаны с использованием Z-нормализации. Обучение U-сети проводилось с использованием различных уровней и числа фильтров (размер батча брался равным 8, 100 эпох и 32 итерации за эпоху).

Для измерения производительности классификации строился график кривой ROC соотношения истинно-положительной доли пикселей (для которых прогноз совпал с разметкой изображения, предсказанной U-сетью) к ложно-положительной доле "дефектных" пикселей (которые U-сеть посчитала ошибочно как "дефектные") для каждого из выбранных пороговых значений вероятности, а в качестве единой меры была выбрана площадь AUC (характеристика качества классификации) под кривой ROC. Чем больше значение AUC, тем "лучше" модель классификации.

Для понимания сути кривой ROC введем понятия элементов таблицы сопряженности, строящейся на основе классификации модели и реальной (фактической) принадлежности к следующим четырем классам:

- истинно положительные (True Positives (TP), или правильно классифицированные положительные случаи);

- истинно отрицательные значения (True Negative (TN), или правильно классифицированные отрицательные случаи);
- ложноотрицательные значения (False Negatives (FN), или неверно классифицированные отрицательные случаи, т. е. те случаи, когда объект ошибочно не обнаруживается);
- ложноположительные значения (False Positives (FP), или неверно классифицированные положительные случаи, т. е. ложные обнаружения объекта).

При анализе результатов принято оперировать не абсолютными показателями, а их долями, выраженными в процентах:

- доля истинно положительных примеров (True Positives Rate)

$$TPR = \frac{TP}{TP + FN} \cdot 100 \%;$$

- доля ложноположительных примеров (False Positives Rate)

$$FPR = \frac{FP}{TN + FP} \cdot 100 \%.$$

Кривая ROC представляет собой линию в координатах истинно положительных (True Positive) и ложноположительных (False Positive) значений. Для оценки ROC кривой используется параметр AUC:

$$AUC = \sum_i \left[\frac{X_{i+1} + X_i}{2} \right] (Y_{i+1} + Y_i),$$

где X_i , Y_i — соответствующие значения на осях кривой ROC.

Поскольку работа направлена на обнаружение дефектов, а не отдельных пикселей, была принята мера уровня объекта, т.е. рабочая характеристика свободного отклика (FROC), показывающая долю обнаруженных дефектов к числу ложных срабатываний на изображение. Дефект считается обнаруженным, если более половины его площади входит в положительную область обнаружения. Далее приводятся результаты нескольких экспериментов.

1. На рис. 1 показаны результаты, полученные при обучении U-сети с помощью **различного числа уровней** (от 1 до 4), построены кривые ROC и вычислены значения AUC для всех уровней: 0,478; 0,649; 0,872; 0,817 соответственно. Видно, что наилучший результат получен при использовании трех уровней.

2. На рис. 2 показаны результаты, полученные при обучении U-сети с числом уровней, равным 3, **для разного числа фильтров сегментации** (32, 64, 128). Как видно, наилучшее значение точности получено с использованием 64 фильтров.

3. Графики на рис. 3 дают представление о том, что **обучение U-сети с нормализацией** на

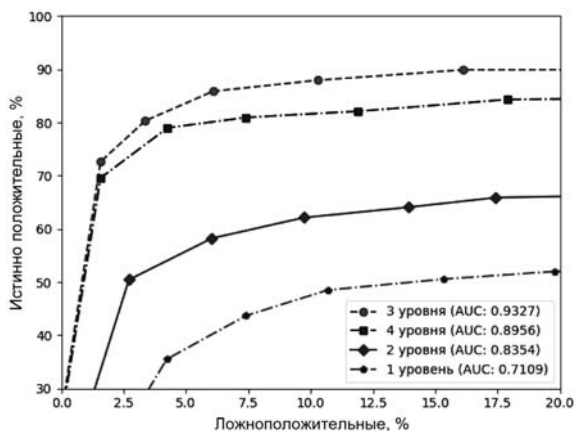


Рис. 1. Влияние числа уровней U-сети

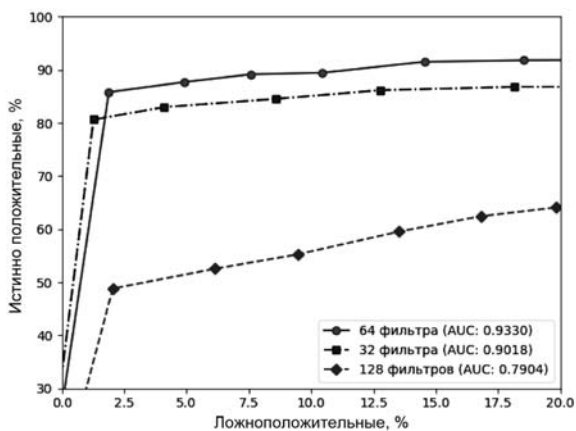


Рис. 2. Влияние числа фильтров

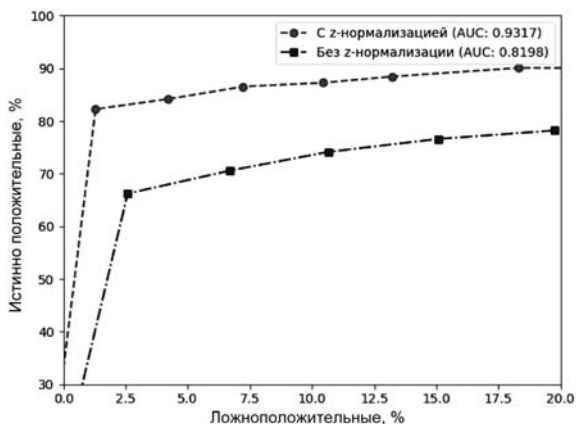


Рис. 3. Влияние Z-нормализации

входных батчах значительно превышает производительность сети без нормализации, при этом значение AUC увеличивается с 0,83 до 0,91.

4. На рис. 4 показаны результаты работы U-сети для трех случаев: с помощью слоя *Softmax*, с помощью классификатора "случайных лесов" с 20 и с 150 деревьями. Результаты работы классификатора со 150 деревьями имеют превосходство по производительности, хотя и не очень значительное.

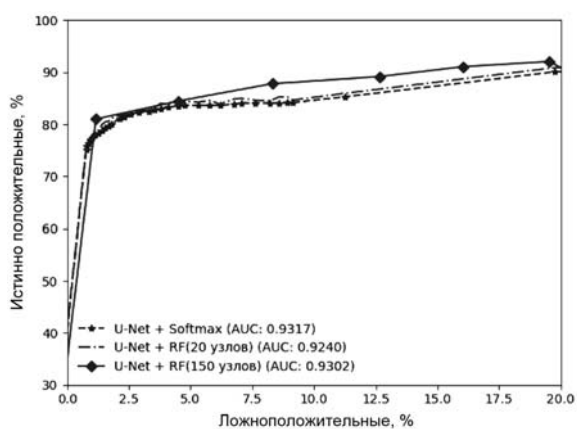


Рис. 4. Сравнение работы *Softmax*-слоя и RF-классификатора

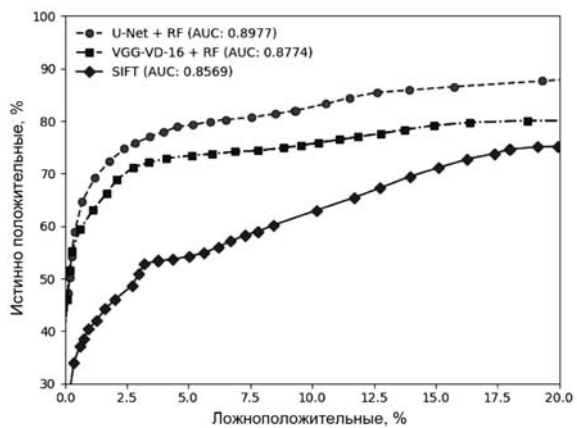


Рис. 5. Сравнения подходов обнаружения дефектов

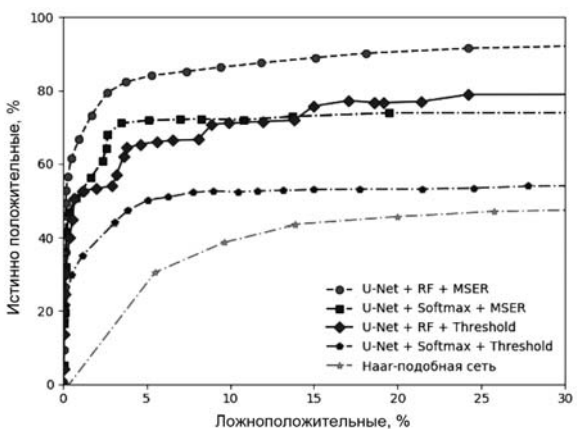


Рис. 6. Сравнение с существующими предобученными сетями

5. В данном эксперименте *сравнивалась производительность рассматриваемой сверточной U-сети с результатами, полученными с помощью многослойных аналогов: SIFT и VGG-VD-16.* На рис. 5 показано, что результаты предлагаемой сети значительно превосходят аналоги.

6. В этом эксперименте сравнивалась производительность рассмотренного подхода для выявления дефектов (U-Net + RF + MSER) со следующими существующими подходами:

- end-to-end предобученная U-сеть с Softmax-слоем и MSER-детектором (U-Net + Softmax + MSER);
- end-to-end предобученная U-сеть с Softmax-слоем и детектором, основанном на применении порога (U-Net + Softmax + Treshhold);
- U-сеть с конечным слоем из классификатора RF и детектором, основанном на применении порога (U-Net + RF + Treshhold);
- использование каскада Хаара с дальнейшей классификацией из RF.

Из рис. 6 видно, что предлагаемый подход дал значительно лучшие результаты *по отношению к другим рассматриваемым методам*, в частности, 84,5 % дефектов были обнаружены при 2,4 ложных срабатываний на одно изображение.

Заключение

Предложен подход, дополняющий U-сеть конечным слоем из классификатора случайных лесов (RF), при котором классификатор идентифицирует каждый пиксель с использованием объектов, извлеченных из сети. Вместо порогового значения массива вероятностей, создаваемого классификатором, предложенный подход U-Net + RF + MSER идентифицирует дефектные кандидаты с помощью детектора максимально устойчивой экстремальной

области (MSER), при котором используются характеристики более высокого уровня, основанные на регионах в изображении, а не простое пиксельное пороговое значение. Наилучшая производительность, полученная с использованием подхода U-Net + RF + MSER, превосходит множество проверенных альтернатив. Хотя данный подход и был ограничен малым числом учебных данных, он все же дал многообещающие результаты.

Список литературы

1. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional networks for biomedical image segmentation // LNCS. 2015. 9351. P. 234–241.
2. **Breiman L.** Random forests // Mach Learn. 2001. 45(1). P. 5–32.
3. **Matas J., Chum O., Urban M., Pajdla T.** Robust wide baseline stereo from maximally stable extremal regions // Proc. British Machine Vision Conference. 2002. P. 384–396.
4. **LeCun Y., Bengio Y., Hinton G.** Deep learning // Nature. 2015. 521. P. 436–444.
5. **Razavian A., Azizpour H., Sullivan J., Carlsson S.** CNN features off-the-shelf: an astounding baseline for recognition // Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2014. P. 512–519.
6. **Ren R., Hung T., Tan K.** A generic deeplearning based approach for automated surface inspection // IEEE Transactions on Cybernetics. 2018. 48(3). P. 929–940.
7. **Taylor L., Nitschke G.** Improving deep learning using generic data augmentation. 2017. arXiv:1708.06020.
8. **Kingma D., Adam J.** A method for stochastic optimization. 2017. arXiv:1412.6980.

G. K. Bukalov, Professor, e-mail: gk.bukalov44@yandex.ru,
A. O. Burygin, PhD student, e-mail: g.t.m.p@yandex.ru, **I. G. Panin**, Professor, e-mail: igpanin@list.ru,
A. B. Tortsev, student, e-mail: fullfulk47458@gmail.com,
 Kostroma State University

Defect Detection Using FCN Modification for Finding Rare Defects on Large Areas

There is problem of finding defects on a textile sling on large areas. For this purpose, the image goes through several stages: creation of a convolutional U-Net network, extraction of U-Net features, classification by the Random Forest algorithm, and identification of defective areas via MSER. The Random Forest classifier is used to segment parts of the input image obtained from U-Net. Computational experiments were conducted to study the effectiveness of the proposed method in comparison with existing methods.

Keywords: sling, image, CNN, U-Net, MSER, Random Forest

DOI: 10.17587/it.26.683-687

Refences

1. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional networks for biomedical image segmentation, LNCS, 2015, 9351, p. 234–241.
2. **Breiman L.** Random forests, Mach Learn, 2001, 45(1), p. 5–32.
3. **Matas J., Chum O., Urban M., Pajdla T.** Robust wide baseline stereo from maximally stable extremal regions, *Proc. British Machine Vision Conference*, 2002, pp. 384–396.

4. **LeCun Y., Bengio Y., Hinton G.** Deep learning, *Nature*, 2015, 521, p. 436–444.
5. **Razavian A., Azizpour H., Sullivan J., Carlsson S.** CNN features off-the-shelf: an astounding baseline for recognition, *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, p. 512–519.
6. **Ren R., Hung T., Tan K.** A generic deeplearning based approach for automated surface inspection, *IEEE Transactions on Cybernetics*, 2018, 48(3), p. 929–940.
7. **Taylor L., Nitschke G.** Improving deep learning using generic data augmentation, 2017, arXiv:1708.06020.
8. **Kingma D., Adam J.** A method for stochastic optimization, 2017, arXiv:1412.6980.