

DOI: 10.15514/ISPRAS-2025-37(6)-25



Применение нейронных сетей семейства YOLO для обнаружения полезных сигналов на вихревоковых дефектограммах рельсов

А.Н. Гладков, ORCID: 0009-0007-0211-5660 <a.gladkov@uniyar.ac.ru>

Л.Ю. Быстров, ORCID: 0000-0002-0610-5466 <l.bystrov@uniyar.ac.ru>

Е.В. Кузьмин, ORCID: 0000-0003-0500-306X <kuzmin@uniyar.ac.ru>

Ярославский государственный университет им. П.Г. Демидова,

Россия, 150003, Ярославль, ул. Советская, д. 14.

Аннотация. Повышение уровня безопасности железнодорожного движения напрямую связано с необходимостью оперативного обнаружения структурных аномалий элементов рельсового пути. Данная задача реализуется посредством регулярных проверок состояния рельсов с применением методов неразрушающего контроля. Среди современных технологий, используемых для этой цели, выделяется вихревоковая дефектоскопия. Дефектоскоп формирует многоканальный дискретный сигнал, который называется дефектограммой. Дефектограммы нуждаются в анализе, то есть в выявлении полезных сигналов, указывающих на дефект или конструктивные элементы рельса. В работе рассматривается применение детектирующих свёрточных нейронных сетей семейства YOLO (You Only Look Once) для автоматического обнаружения полезных сигналов рельсов на вихревоковых дефектограммах рельсов. Цель исследования – оценить эффективность различных способов преобразования многоканального сигнала в двумерные изображения, совместимые с YOLO. Исследованы четыре метода преобразования: пороговое, основанное на сравнении амплитуд с пороговым уровнем шума, оконное преобразование Фурье, непрерывное вейвлет-преобразование и преобразование Гильберта-Хуанга. Набор данных для обучения состоит из фрагментов дефектограмм по 50 тыс. отсчётов с полезными сигналами трёх классов (болтовыестыки, электроконтактные и алюминотермитные сварки). Данные разделены на обучающую, валидационную и тестовую выборки. Обученные на этих данных модели YOLO для всех рассмотренных методов преобразования продемонстрировали высокие показатели сбалансированной средней точности mAP. Наилучшие показатели были достигнуты при использовании непрерывного вейвлет-преобразования, в то время как пороговое преобразование оказалось наименее вычислительно затратным. Окноное преобразование Фурье позволило достичь лучшего баланса между точностью и полнотой обнаружения полезных сигналов. Результаты исследования подтверждают потенциал использования сетей YOLO для анализа вихревоковых дефектограмм и сигналов в целом.

Ключевые слова: свёрточные нейронные сети; сети YOLO; вихревоковая дефектоскопия; оконное преобразование Фурье; непрерывное вейвлет-преобразование; преобразование Гильберта-Хуанга.

Для цитирования: Гладков А.Н., Быстров Л.Ю., Кузьмин Е. В. Применение нейронных сетей семейства YOLO для обнаружения полезных сигналов на вихревоковых дефектограммах рельсов. Труды ИСП РАН, том 37, вып. 6, часть 2, 2025 г., стр. 131–150. DOI: 10.15514/ISPRAS-2025-37(6)-25.

Благодарности: Исследование выполнено в рамках инициативной НИР ЯрГУ им. П.Г. Демидова № VIP-016.

Application of YOLO Family Neural Networks for Useful Signals Detection on Eddy Current Rail Defectograms

A.N. Gladkov, ORCID: 0009-0007-0211-5660 <a.gladkov@uniyar.ac.ru>

L.Y. Bystrov, ORCID: 0000-0002-0610-5466 <l.bystrov@uniyar.ac.ru>

E.V. Kuzmin, ORCID: 0000-0003-0500-306X <kuzmin@uniyar.ac.ru>

P.G. Demidov Yaroslavl State University,
14 Sovetskaya str., Yaroslavl 150003, Russia.

Abstract. Improving the level of safety of railway traffic is directly related to the need for prompt detection of structural anomalies of track elements. This task is implemented through regular inspections using non-destructive testing methods. Among the modern technologies used for this purpose, eddy current flaw detection stands out. The flaw detector generates a multi-channel discrete signal, which is called a defectogram. Defectograms require analysis, that is, the identification of useful signals from a defect or structural elements of the rail. This paper investigates the use of YOLO (You Only Look Once) family convolutional neural networks for automated detection of useful signals in eddy current rail defectograms. The main objective was to evaluate the effectiveness of different transformations of multichannel time-series data into two-dimensional images suitable for YOLO processing, and to explore the trade-off between detection accuracy and computational cost. Four transformation methods are examined: Threshold Transform, based on amplitude comparisons against a twice threshold noise level, Short-Time Fourier Transform, Continuous Wavelet Transform and Hilbert–Huang Transform. The dataset comprises defectogram fragments of 50 thousand counts with annotated useful signals from three classes (flash butt welds, aluminothermic welds, and bolt joints), split into training, validation, and test sets. YOLO models trained on this data achieved high mean Average Precision scores in useful signals detection for all considered transformation methods. Continuous Wavelet Transform yielded the best scores while the Threshold Transform proved to be the least computationally expensive. Short-Time Fourier Transform method offered the best balance between precision and recall. Hilbert–Huang Transform showed slightly lower effectiveness. These results demonstrate the suitability of YOLO networks for eddy current defectogram analysis and useful signals detection in general.

Keywords: convolutional neuron networks; YOLO neuron networks; eddy current testing; short-time Fourier transform; continuous wavelet transform; Hilbert–Huang transform.

For citation: Gladkov A.N., Bystrov L.Y., Kuzmin E.V. Application of YOLO Family Neural Networks for Useful Signals Detection on Eddy Current Rail Defectograms. *Trudy ISP RAN/Proc. ISP RAS*, vol. 37, issue 6, part 2, 2025, pp. 131-150 (in Russian). DOI: 10.15514/ISPRAS-2025-37(6)-25.

Acknowledgements. This work was supported by P.G. Demidov Yaroslavl State University Project № VIP-016.

1. Введение

Вихретоковая дефектоскопия – это один из современных методов неразрушающего контроля железнодорожного полотна, основанный на взаимодействии вихревых токов с материалом рельса. Он применяется для выявления дефектов в электропроводящих материалах без повреждения объекта. Дефектоскоп формирует сигналы для каждого миллиметра рельса с помощью набора физических датчиков. Использование нескольких датчиков придаёт сигналам пространственную структуру. Количество датчиков определяет число каналов для каждого сигнала. Полученный таким образом набор данных называется дефектограммой. Дефектограммы отражают информацию о расположении конструктивных элементов и состоянии поверхности катания рельсов (рис. 1). Анализ дефектограмм заключается в поиске на ней сигналов от дефектов и конструктивных на фоне рельсового шума и помех разного рода. Системы автоматического анализа дефектограмм позволяют своевременно обнаружить и устранить дефекты рельсов, что прямым образом влияет на безопасность движения железнодорожного транспорта.

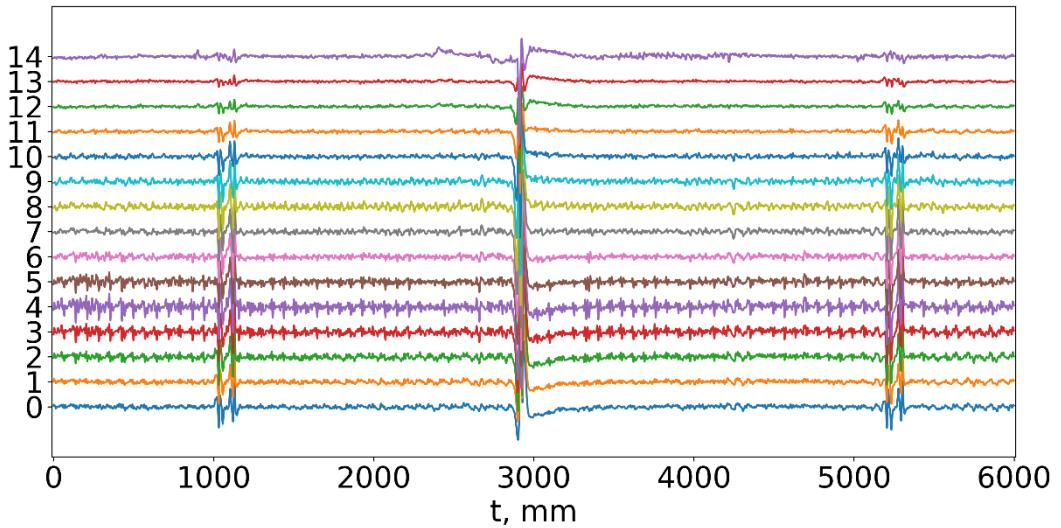


Рис. 1. Пример вихревоковой дефектограммы рельсов.

Fig. 1. The Example of Eddy Current Rail Defectogram.

В данной работе рассматриваются дефектограммы, сформированные 14-разрядным 15-канальным вихревоковым дефектоскопом. Значения амплитуд сигналов – целые числа от –8192 до 8191. Ввиду большой протяжённости дефектограммы нарезают на фрагменты. Длина одного фрагмента дефектограммы – 50 тыс. отсчётов. Далее под дефектограммой будут подразумеваться её фрагменты описанного размера. Таким образом, дефектограммы можно рассматривать как 15-канальный дискретный сигнал $\bar{x}(j)$:

$$\bar{x}(j) = (x_0(j), x_1(j), \dots, x_{14}(j)), \quad j \in \overline{0, 49999}.$$

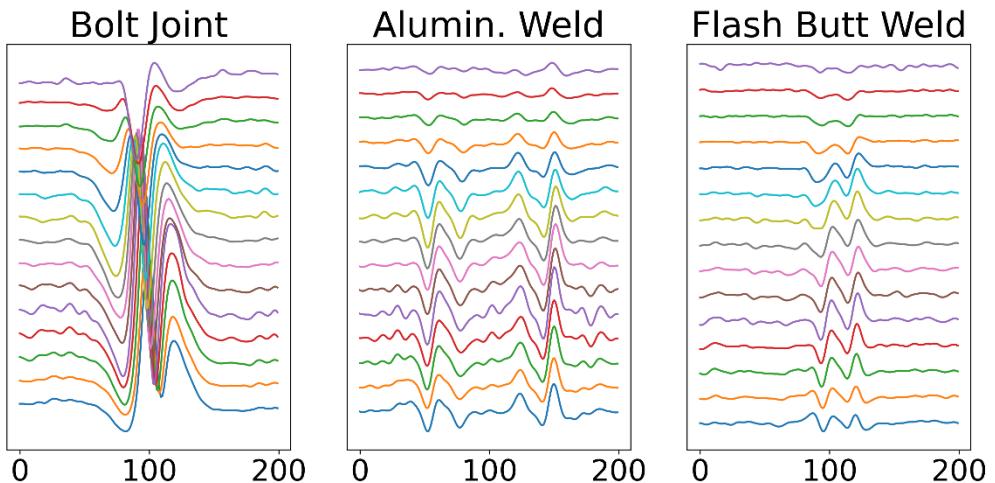


Рис. 2. Примеры сигналов коротких конструктивных элементов на вихревоковых дефектограммах рельсов.

Fig. 2. The Examples of Short Structural Element Signals on Eddy Current Rail Defectogram.

На вихревоковых дефектограммах можно выделить сигналы от коротких и протяжённых конструктивных элементов, помех оборудования и рельсовых дефектов. Среди полезных

сигналов дефектограммы в данной статье рассматриваются сигналы, формирующие образы коротких конструктивов: болтовых стыков (Bolt Joint), электроконтактных (Flash Butt Weld) и алюминотермитных сварок (Aluminothermic Weld). Эти образы представлены на рис. 2. Стандартный подход к обнаружению полезных сигналов в дефектограмме основан на вычислении порогового уровня рельсового шума [1]. Полезные сигналы выделяются на фоне шума высокими значениями амплитуд, поэтому естественной идеей их обнаружения является сравнение амплитуд сигналов со среднеквадратическим отклонением рельсового шума. Перечислим проблемы такого подхода:

- Пороговый уровень шума может быть завышен из-за присутствующих в данных помех [2].
- Не все полезные сигналы могут выходить за пороговый уровень. Возникают трудности с точностью выделения границ полезных сигналов.
- Полезные сигналы могут быть записаны с более низкой чувствительностью по сравнению с рельсовым шумом и не выходить за пределы порогового уровня.

Таким образом, существует потребность в поиске новых методов обнаружения полезных сигналов на вихревых дефектограммах, которые могли бы учитывать форму и сложные свойства сигналов, а не только их высокочастотность.

В последние годы активно развивается подход представления одномерных сигналов в виде двумерных изображений для последующей обработки свёрточными нейронными сетями (Convolutional Neuron Network, CNN) [3]. Такие методы позволяют использовать проверенные архитектуры компьютерного зрения для задач анализа сигналов. В задаче обнаружения объектов нейронные сети могут выступать в качестве основного инструмента, в этом случае вся логика решения реализуется в архитектуре сети напрямую и сеть носит название детектирующей. Наиболее популярным семейством детектирующих свёрточных нейронных сетей являются сети YOLO (You Only Look Once), демонстрирующие впечатляющие результаты в области анализа изображений. Модели YOLO обеспечивают баланс между точностью и скоростью обнаружения объектов в реальном времени [4]. В отличие от двухэтапных детекторов, таких как R-CNN, разделяющими задачу формирования ограничивающих рамок (bounding boxes) в одном модуле и их отбор в другом, YOLO рассматривает задачу обнаружения объектов как проблему регрессии, предсказывая ограничивающие рамки и вероятности классов объектов непосредственно из входного изображения за один проход через нейронную сеть. Изначальная архитектура YOLO состояла из свёрточных слоёв и слоёв макс-пулинга (max-pooling), но с течением времени она эволюционировала в более модульную структуру. Каждая новая версия YOLO вносила улучшения в архитектуру, функцию потерь и методы обучения, что приводило к повышению точности обнаружения, особенно для объектов малого размера, и снижению вычислительных затрат [5]. Высокая скорость работы и эффективность YOLO позволяют использовать эти сети для обработки потенциально больших объёмов данных в реальном времени.

В данной работе исследуются методы применения сетей YOLO в решении задачи обнаружения полезных сигналов на вихревых дефектограммах. Методы отличаются друг от друга способом преобразования 15-канального сигнала $\bar{x}(j)$ в двумерное изображение $F(x; y)$. Рассматриваются 4 вида преобразований: оконное преобразование Фурье, непрерывное вейвлет-преобразование, преобразование Гильберта-Хуанга и пороговое преобразование. Пороговое преобразование является новым методом и вводится впервые. Все виды рассмотренных методов успешно интегрируются с YOLO и показывают высокое качество обнаружения полезных сигналов.

2. Обзор литературы

Существует большое количество техник преобразования сигналов в изображения. Среди них можно выделить две основные группы: преобразования, основанные на частотно-временном анализе, и методы векторного кодирования изображений. К первой категории относятся спектрограммы (Short-time Fourier Transform, STFT), скалограммы (Continuous Wavelet Transform, CWT), спектры Гильберта-Хуанга (Hilbert-Huang Transform, HHT), распределения Вигнера-Вилля (Wigner-Ville Distribution, WVD). Вторая категория представлена грамиановыми угловыми полями (Gramian Angular Field, GAF), рекуррентными изображениями (Recurrence Plot, RP), разностными полями Маркова (Markov Transition Field, MTF) и полями разности мотивов (Motif Difference Field, MDF). Отдельно от этих двух больших групп отстоят изображения гомологической устойчивости (Persistence Image, PI). Все эти преобразования визуально представлены на рис. 3.

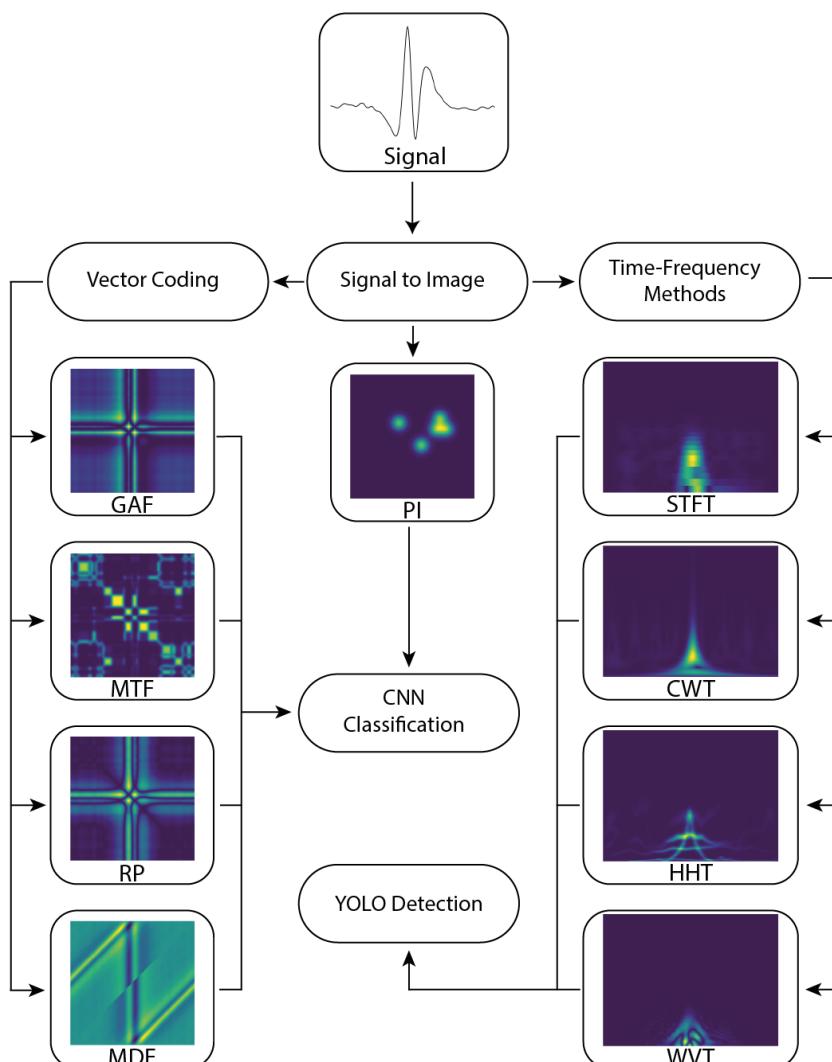


Рис. 3. Основные способы преобразования сигналов в изображения.
Fig. 3. The Main Ways to Convert Signals into Images.

Оконное преобразование Фурье (STFT) позволяет анализировать гармонический спектр в зависимости от времени. В дискретном случае оно описывается функцией

$$F(t, \omega) = \frac{1}{N} \sum_{j=0}^{N-1} x(j) \cdot W(j - t) \cdot e^{-\frac{i\omega 2\pi}{N}},$$

где $x(j)$ – дискретный сигнал, $j \in \overline{0, N-1}$, $W(t)$ – оконная функция, $\omega \in \overline{0, N-1}$, $i = \sqrt{-1}$. Этот метод широко используется в распознавании аудио- и вибросигналов [6]. Получающиеся с помощью него изображения называются спектрограммами. Спектрограмма отражает как временную, так и частотную составляющую сигнала, что может быть полезно при поиске полезных сигналов. Спектрограммы являются наиболее популярными двумерными представлениями сигналов, активно применяющимися на практике. Так в работах [7] и [8] для вибрационных сигналов, полученных от электродвигателей и вращающихся машин, были построены спектрограммы, на которых обучена свёрточная нейронная сеть для классификации. Такой подход позволил значительно усовершенствовать системы диагностики неисправностей. Аналогичный подход применялся в статьях [9] и [10] и позволил достичь высокого качества обнаружения дефектов в свёрлах и дорожном покрытии. Авторы [11] применили сеть YOLOv10 вместе с STFT в задаче обнаружения и классификации радиосигналов и показали важность правильного выбора оконной функции для качества обнаружения. Успех применения спектрограмм в данных задачах обусловлен тем, что вибрации вращающихся систем, звуковые колебания и радиосигналы характеризуются выраженным гармоническим составом. Если же сигналы являются нестационарными и обладают сложной гармонической структурой, спектрограмма может не отражать их особенностей. В этом случае применяют непрерывное вейвлет-преобразование [12].

Главным отличием вейвлет-преобразования от преобразования Фурье является использование в разложении сигналов в качестве базиса не бесконечных функций, а ограниченных, локализованных по времени и по частоте. Это позволяет выявлять в сигналах резкие и быстро затухающие скачки амплитуд. Непрерывное вейвлет-преобразование $F(a, b)$ – это свёртка исходного ряда $x(j)$ с материнской вейвлет-функцией $\psi(j)$:

$$F(a, b) = \frac{1}{\sqrt{a}} \sum_{j=0}^{N-1} x(j) \cdot \psi\left(\frac{j-b}{a}\right),$$

где a – параметр масштаба, b – смещение, $a > 0$, $b \in \overline{0, N-1}$.

Получаемые на основе CWT изображения называются скалограммами (от англ. *scale* – масштаб). Скалограммы использовались в [13] для обработки вибрационных сигналов с целью повышения производительности моделей на основе CNN при обнаружении неисправностей в винтовых редукторах, благодаря чему была достигнута высокая точность классификации (более 99%). Авторы [14] применяли CWT в качестве промежуточного слоя в архитектурах LSTM и CNN для классификации аномалий в вентиляторных системах. В работе [15] проводится сравнительный анализ непрерывных вейвлет-преобразований с разными материнскими вейвлет-функциями (WT-Amor, WT-Bump, WT-Morse) в задаче диагностики неисправностей электродвигателей с использованием свёрточных нейронных сетей. Самая высокая доля правильных ответов (accuracy) была достигнута при использовании вейвлетов семейства Морзе. Авторы [16] решали задачу обнаружения утечек в трубопроводе. Они подавали на вход свёрточной сети двухканальное изображение, где один канал представлял собой спектрограмму, а второй – скалограмму сигналов акустической эмиссии. Такое простое сочетание методов позволило значительно улучшить качество обнаружения утечек.

Итак, непрерывное вейвлет-преобразование является универсальным инструментом преобразования различных типов сигналов в информативные для CNN изображения. Главное

ограничение CWT – высокая трудоёмкость по времени и по памяти при генерации матриц [17] (прогнозируется квадратичная сложность по размеру сигнала).

Отдельно остановимся на статье [18], авторы которой решали задачу обнаружения полезных сигналов в вихревоковых дефектограммах рельсов. Изображения в статье формируются с использованием непрерывного вейвлет-преобразования Морле. Авторы заявляют, что им удалось для сформированных изображений обучить свёрточную нейронную сеть и классифицировать с 98% точностью полезные сигналы трёх видов: скваты (вид дефекта), сварки и болтовые стыки. Однако результаты исследования нуждаются в тщательной проверке. Сомнения в полученных результатах вызывает использование авторами одноканального дефектоскопа, который, очевидно, даёт ограниченное представление о состоянии поверхности катания рельсов. Более того, наличие всего лишь одного канала затрудняет формирование образов конструктивных элементов рельсов (сварок и стыков рельсов). Полезные сигналы обнаруживаются с помощью сравнения амплитуд с постоянным пороговым уровнем шума при том, что уровень шума рельсов, как правило, является динамической величиной (зависит от настроек оборудования и типов рельсов) и для разных участков рельсового пути может сильно отличаться. Таким образом, достоверность результатов исследования [18] вызывает сомнения, результаты работы нуждаются в апробации.

Преобразование Гильберта-Хуанга (ННТ) – это ещё один метод анализа данных, предназначенный для нелинейных и нестационарных сигналов, то есть сигналов, частотные характеристики которых меняются со временем. ННТ состоит из двух этапов: эмпирическая модовая декомпозиция и преобразование Гильберта. На первом этапе сигнал $x(t)$ разлагается на набор простых колебательных компонент, называемых внутренними модовыми функциями (Intrinsic Mode Functions, IMF). IMF являются монохроматическими функциями, то есть локально выглядят как синусоиды с медленно меняющимися параметрами. В результате эмпирической модовой декомпозиции ряд раскладывается на n модовых функций c_j и остаток r :

$$x(t) = \sum_{j=1}^n c_j(t) + r(t).$$

Далее, к каждой IMF применяется преобразование Гильберта

$$H[c_j(t)] = \frac{1}{\pi} P.V. \int_{-\infty}^{+\infty} \frac{c_j(\tau)}{t - \tau} d\tau,$$

где $P.V.$ означает главное значение интеграла по Коши (Cauchy Principal Value).

Сигнал вида $h_j(t) = c_j(t) + iH[c_j(t)]$ называется аналитическим. Представление аналитического сигнала в полярной форме позволяет перейти к спектру Гильберта.

$$h_j(t) = A_j(t)e^{i\Theta_j(t)}, A_j(t) = \sqrt{(c_j(t))^2 + (H[c_j(t)])^2}, \Theta_j(t) = \arctan \frac{H[c_j(t)]}{c_j(t)}.$$

Мгновенная частота в спектре находится как производная фазы по времени

$$\omega_j(t) = \frac{1}{2\pi} \frac{d\Theta_j(t)}{dt}.$$

Наконец, набор троек $(t, \omega_j(t), A_j(t))$ образует спектр Гильберта для всех IMF j и моментов времени t . Такое представление даёт богатую информацию о динамике нестационарных процессов. При этом преобразование Гильберта-Хуанга, как и вейвлет-преобразование, может быть вычислительно трудоёмким. Другой недостаток ННТ связан с тем, что эмпирическая модовая декомпозиция может иногда приводить к смешению мод, когда различные частотные компоненты смешиваются в одной и той же IMF. В этом случае спектр Гильберта становится тяжело интерпретируемым.

В статье [19] решается задача диагностики неисправностей подшипников в реактивных двигателях с переключаемым сопротивлением. Для этого по вибрационным сигналам строится спектр Гильберта, на котором с помощью двухэтапного детектора R-CNN выполняется поиск ограничивающих рамок. Двухэтапные детекторы значительно уступают в качестве одноэтапным, таким как YOLO. Авторы заявляют о высокой доле правильных ответов, но считают её на обучающей, а не тестовой выборке. В связи с тем, что сам метод R-CNN не доказал свою эффективность [20], возникают сомнения насчёт того, что модель сможет показать такие же впечатляющие результаты обнаружения объектов на новых данных.

Авторы [21] представили автоматическую систему для оценки качества бетона методом ударно-эхового контроля. В ней сигналы преобразовывались в изображения с помощью распределения Вигнера-Вилля (WVD). Далее, в отличие от работ, рассмотренных ранее, свёрточные сети использовались напрямую для обнаружения объектов. Авторы обучили разные версии YOLO и с помощью них обнаружили трещины, поверхностные волны и вибрации в бетоне. При этом вместо классических метрик точности (Precision), полноты (Recall) и средней точности (mAP) они использовали свою метрику Detection rate, равную отношению числа найденных отметок к числу правильных отметок. Интерпретация данной метрики затруднительна. Если сеть найдёт слишком много лишних отметок, то Detection rate превысит единицу. Таким образом, используемая авторами метрика не позволяет адекватно оценить качество работы сети и, соответственно, подвергает сомнению результаты исследования.

Вторая группа преобразований сигналов в изображения основана на векторном кодировании. Ввиду того, что получаемые в данной группе изображения сложно использовать вместе с YOLO, им удалено меньше внимания.

Грамиановы угловые поля (GAF) и разностные поля Маркова (MTF) преобразуют сигнал в матрицу попарных отношений. Это позволяет получать информацию о корреляции всех пар амплитуд во времени [22]. Как правило, GAF/MTF применяются там, где важна общая структура и глобальные временные закономерности. Кратковременные сигналы могут быть размыты или менее заметны. С некоторыми проблемами GAF и MTF справляются поля разностей мотивов (MDF), которые обладают простотой реализации и отлавливают больше значимых признаков сигнала [23]. Известно, что эти методы требуют значительных вычислительных ресурсов (для сигнала размера N создаётся матрица $N \times N$) и порождают изображения большого размера, не пригодного для YOLO. Если исходный сигнал разбить на множество небольших фрагментов и для них выполнить GAF/MTF/MDF, то полученные изображения будут отлично сочетаться с моделями классификации (одно изображение — один класс), но не подойдут для задач, где нужно обнаруживать несколько объектов [24].

Рекуррентные изображения (RP) визуализируют моменты времени, когда траектория в фазовом пространстве возвращается в ранее посещённое состояние. Они создаются путём вычисления матрицы расстояний между всеми парами точек в реконструированном фазовом пространстве и последующей пороговой обработкой этой матрицы для выявления рекуррентности [25]. Рекуррентные изображения имеют все те же проблемы, что и GAF/MTF/MDF: высокая трудоёмкость и большой размер изображений. Кроме того, они выделяют только повторяющиеся закономерности и не кодирует пространственное и временное расположение сигналов, что имеет принципиальное значение для обнаружения полезных сигналов вихревых дефектограмм.

Ещё один способ создания изображений из сигналов основан на вычислении устойчивых гомологий. Устойчивые гомологии — это объект из топологического анализа данных, который извлекает информацию о «форме» данных, идентифицируя и количественно определяя топологические признаки, такие как компоненты связности и дыры, которые сохраняются при различных масштабах (разрешениях) данных — фильтрациях. Изображения гомологической устойчивости (PI) — это способ представления выходных данных

топологического анализа (диаграмм устойчивости) в виде изображений с помощью свёртки диаграмм с гауссовским ядром [26]. Они часто используются в машинном обучении [27]. PI отображают время жизни топологических характеристик сигналов, но по ним нельзя восстановить временную локализацию сигналов. Таким образом, изображения устойчивых гомологий можно использовать для классификации сигналов, но не для их обнаружения, то есть сети YOLO к PI не применимы.

Итак, лучше всего для интеграции с YOLO подходят преобразования сигналов в изображения, основанные на частотно-временном анализе. С одной стороны, они сохраняют временную локализацию сигналов, с другой – кроме пространственных признаков позволяют анализировать частотные характеристики. Изображения на основе GAF, MTF, RP и PI могут использоваться вместе с нейросетевыми классификаторами, но не пригодны для детектирующих моделей.

3. Методология

В данном разделе рассматриваются методы преобразования изображений, которые впоследствии интегрируются с YOLO. Три из четырёх методов основаны на частотно-временном анализе и уже были частично рассмотрены в предыдущем разделе: оконное преобразование Фурье, непрерывное вейвлет-преобразование и преобразование Гильберта-Хуанга. Предложен также авторский вариант преобразования, названный пороговым преобразованием (Threshold Transform, TT).

Требуется не просто преобразовать дефектограмму в изображение, но в изображение приемлемого для YOLO размера. Сети YOLO на вход принимают изображения размером 640×640 в формате RGB. Поэтому полученные из дефектограммы изображения нарязаются на несколько фрагментов длиной 640 пикселей каждый. После данной операции фрагментированные изображения выстраиваются вертикально друг под другом и дополнительно разделяются зелёными полосами для наглядного представления границ фрагментации. На основе каждого из преобразований строятся 2 матрицы, соответствующие красному и синему цвету, а зелёный цвет используется для отображения полос. Таким образом достигается размер и формат изображения, воспринимаемые YOLO.

3.1. Пороговое преобразование

Данное преобразование не основано на частотно-временном анализе и вводится впервые. Будем непосредственно сопоставлять каждой амплитуде сигнала пиксель с определённой цветовой интенсивностью. Такой простой подход позволяет значительно сократить вычислительные ресурсы, которые требуются для создания изображений преобразованию Фурье и более трудоёмким алгоритмам, таким как непрерывное вейвлет-преобразование и преобразование Гильберта-Хуанга. На основе дефектограммы $\bar{x}(j)$ построим 2 матрицы $F_r(i, j)$ и $F_b(i, j)$. Матрица $F_r(i, j)$ хранит информацию об интенсивностях красного цвета каждого пикселя и соответствует положительным амплитудам сигналов. Матрица $F_b(i, j)$ хранит информацию об интенсивностях синего цвета и соответствует отрицательным амплитудам.

Пороговое преобразование основано на сравнении амплитуд сигналов с удвоенным пороговым уровнем шума $level$, который находится с помощью алгоритма из [1]. Если амплитуда сигнала $x_i(j)$ в канале i по модулю превысила $level$, то пикселию (i, j) в зависимости от знака амплитуды сопоставляется интенсивность красного или синего цвета по логарифмическому правилу $(\log_{10} \frac{|x_i(j)|}{level}) \cdot 191 + 64$; в противном случае используется линейное правило $\frac{|x_i(j)|}{level} \cdot 64$. Каждая интенсивность округляется до целого числа. При превышении допустимого значения интенсивности полученное число понижается до 255. Гиперпараметр 64, соответствующий моменту перехода от линейной функции к

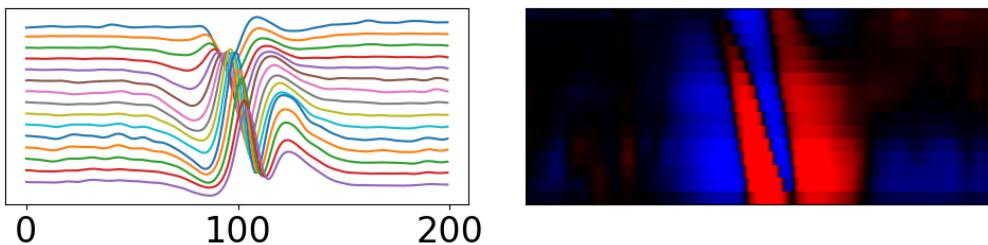
логарифмической, подобран эмпирически. В виде формулы пороговое преобразование можно описать следующим образом:

$$F(i, j) = \begin{cases} \left\lfloor \frac{|x_i(j)|}{level} \cdot 64 \right\rfloor, & |x_i(j)| < level, \\ \min \left(\left\lfloor \left(\log_{10} \frac{|x_i(j)|}{level} \right) \cdot 191 + 64 \right\rfloor; 255 \right), & |x_i(j)| \geq level. \end{cases}$$

Элементы матриц $F_r(i, j)$ и $F_b(i, j)$ формируются с помощью порогового преобразования $F(i, j)$ в соответствии со знаком амплитуды $x_i(j)$. Например, для $F_r(i, j)$ преобразование выглядит как

$$F_r(i, j) = \begin{cases} F(i, j), & \text{если } x_i(j) \geq 0, \\ 0, & \text{иначе.} \end{cases}$$

Пример изображения, полученного применением порогового преобразования к сигналу болтового стыка, приведён на рис. 4.



*Рис. 4. Применение порогового преобразования к сигналу болтового стыка.
Fig. 4. Application of Threshold Transform to a Bolt Joint Signal.*

Когда матрицы сформированы, они фрагментируются на матрицы меньшего размера с сохранением исходной высоты, но длиной 640 пикселей. Фрагменты пересекаются друг с другом. Длина пересечения равна 200. Пересечение призвано предотвратить разрез сигналов конструктивных элементов, длина которых не превышает 200 отсчётов. Для 50 тыс. отсчётов дефектограммы по каждому каналу описанная процедура позволяет получить 114 фрагментированных изображений. Из них после вертикального выстраивания с разделяющими полосами высотой 1 пиксель получаются 3 изображения размером 640×640 .

3.2 Оконное преобразование Фурье

Преобразование Фурье выполняется отдельно для каждого канала дефектограммы с окном Ханна размера 64 и шагом 5. Спектр Фурье симметричен относительно нуля, то есть отрицательные и положительные частоты в спектре дублируют друг друга. В связи с этим рассматривается только положительная часть спектра. В результате по каждому каналу получается матрица размером 32×10000 .

Основная часть информации о частотах полезных сигналов в вихревоковых дефектограммах представлена на отрезке угловой частоты от 0 до 1 [2]. Если рассматривать только этот диапазон частот, то размер матрицы для каждого канала сократится до 10×10000 . Для того чтобы из спектра Фурье получить две матрицы (интенсивностей красного и синего цвета), красному цвету сопоставляются действительные значения гармоник, а синему – мнимые. Далее, матрицы объединяются по вертикали отдельно для каждого цвета, образуя 2 матрицы размером 150×10000 . Необходимые матрицы размера 640×640 получаются с помощью описанной ранее операции фрагментации с длиной 140

пересечения фрагментов, равной 40. Высота разделяющих полос равна 10. Каналы из одного фрагмента соединяются друг с другом без разделения полосами. По дефектограмме длиной 50 тыс. отчётов получается 5 изображений 640×640 .

3.3 Непрерывное вейвлет-преобразование

Скалограмма строится отдельно для каждого канала дефектограммы. В качестве материнской вейвлет-функции используется вейвлет Морле с параметрами $B = 2,5$ и $C = 0,5$:

$$\psi(j) = \frac{1}{\sqrt{\pi B}} e^{-\frac{j^2}{B}} e^{2i\pi C j}, i = \sqrt{-1}.$$

В качестве параметров масштаба a равномерно выбраны 8 значений из отрезка $[0; 1]$ по логарифмической шкале. Вейвлет-преобразование имеет квадратичную сложность. Для того чтобы сократить количество вычислений, амплитуды берутся с шагом 5 (сжатие в 5 раз). Если скалограммы каждого из каналов выстроить вертикально, то получится матрица размером 120×10000 . Как и в случае преобразования Фурье, красный цвет формируется действительными значениями вейвлет-преобразования, а синий – мнимыми. Далее, матрицы фрагментируются с длиной пересечения фрагментов 40. Высота разделительной полосы равна 8 пикселей. Из 50 тыс. амплитуд получается 5 изображений размера 640×640 .

3.4 Преобразование Гильберта-Хуанга

Для каждого канала независимо считаются внутренние модовые функции (IMF) по методу маскирующих сигналов [28], что позволяет улучшить качество эмпирической модовой декомпозиции. Далее, применяется преобразование Гильbertа. Полученные спектры на каждом канале нормализуются по стандартному отклонению и суммируются друг с другом. Размер полученной матрицы сокращается в 2 раза по длине за счёт отбrosa каждого второго столбца матрицы (сжатие в 2 раза). Высота матрицы может быть сколь угодно большой ввиду того, что функция мгновенной частоты $\omega(t)$ является непрерывной. В нашем случае функция дискретизирована так, чтобы высота матрицы Гильбертова спектра составляла 63 пикселя. Полученная матрица 63×25000 фрагментируется с длиной пересечения фрагментов 140. Используется только красный канал изображения, синий канал содержит нулевые значения. Зелёные разделяющие полосы имеют высоту 1 пиксель. Из дефектограммы размера 50 тыс. отчётов получается 5 изображений.

4. Обучение YOLO

В этом разделе даётся описание того, каким образом производилось обучение модели YOLO, приводятся результаты обучения для разных видов изображений, построение которых было описано в предыдущем разделе.

4.1 Метрики оценки качества обнаружения объектов

При оценке качества обнаружения объектов необходимо контролировать как точность локализации объектов, так и надёжность их классификации. Для этого используется стандартная метрика – сбалансированная средняя точность (mean Average Precision, mAP). Она основана на понятии истинноположительных (True Positive, TP), ложноположительных (False Positive, FP) и ложноотрицательных (False Negative, FN) предсказаний модели. TP означает количество истинных предсказаний модели, FP – количество ложных предсказаний модели, а FN – количество не найденных моделью объектов. Истинность или ложность предсказания модели обычно определяется на основании вычисления пересечения над объединением (Intersection over Union, IoU), которое позволяет оценить степень

геометрической схожести отметок. Обычно для сопоставления отметок выбирается порог IoU равный 0,5.

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|}.$$

Точность (P) вычисляется как отношение количества правильных предсказаний к количеству всех предсказаний модели:

$$P = \frac{TP}{TP + FP}.$$

Полнота (R) находится как отношение количества правильных предсказаний модели к количеству действительно правильных объектов:

$$R = \frac{TP}{TP + FN}.$$

Средняя точность (AP) считается как площадь под кривой точность-полнота (PR-curve). Для формирования PR-curve предсказания модели сортируются по убыванию уверенности (confidence) предсказаний модели, затем вычисляется точность и полнота для каждого порога уверенности. Средняя точность вычисляется для каждого класса отдельно, сбалансированная средняя точность (mAP) находится как среднее арифметическое средних точностей для каждого класса. Формально это записывается следующим образом:

$$AP = \int_0^1 P(R)dR,$$
$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i,$$

где $P(R)$ – значение на RP-curve в точке R, AP_i – средняя точность для i -го класса, n – количество классов.

Число mAP50 – это та же метрика mAP, вычисленная при пороге IoU = 0,5. Число mAP50-95 – это среднее арифметическое значений метрики mAP, вычисленных при пороге IoU от 0,5 до 0,95 с шагом 0,05.

4.2 Описание набора данных

Набор данных состоит из 1540 фрагментов дефектограммы длиной 50000 отсчётов каждый, на которых отмечены полезные сигналы от болтовых стыков (Bolt Joint), электроконтактных (Flash Butt Weld) и алюминотермитных (Aluminothermic Weld) сварок. Это сигналы коротких конструктивов вихревоковых данных. В зависимости от метода преобразования дефектограмм в изображения для набора данных можно получить от 4620 до 7700 изображений. В табл. 1 представлена информация о количестве изображений в обучающей, валидационной и тестовой выборках для каждого преобразования. Из набора данных исключены изображения, не содержащие отметок. Всего в наборе присутствует 4123 полезных сигнала. Информация о количестве объектов каждого класса в выборках представлена в табл. 2.

Для разных способов формирования изображений сохранялось единообразное разбиение на выборки. Вихревоковые данные изначально формируются для целых проездов – продолжительных записей одного участка железнодорожного пути. Далее, проезды нарезаются на дефектограммы по 50 тыс. отсчётов. Выборки составлялись таким образом, чтобы все участки данных, полученные из одного проезда, попадали в общую выборку. Такой подход позволяет уменьшить корреляцию между дефектограммами обучающей, валидационной и тестовой выборок и контролировать переобучение (overfitting).

Табл. 1. Размеры обучающей, валидационной и тестовой выборок.

Table 1. Sizes of Training, Validation and Test Sets.

Способ формирования изображений	Общее количество	Обучающая выборка	Валидационная выборка	Тестовая выборка
Пороговое преобразование	3360	1961	974	425
Оконное преобразование Фурье	3744	2259	1027	458
Непрерывное вейвлет-преобразование	3540	2098	1006	436
Преобразование Гильберта-Хуанга	3853	2336	1055	462

Табл. 2 Число полезных сигналов в наборе данных.

Table 2. The Number of Useful Signals in the Data Set.

Выборка	Электроконтактные сварки	Алюминотермитные сварки	Болтовыестыки	Всего
Общая	2831	1232	60	4123
Обучающая	1609	858	28	2495
Валидационная	920	184	17	1121
Тестовая	302	190	15	507

4.3 Результаты обучения

Среди сетей YOLO была выбрана одна из последних версий модели YOLOv11n. Параметры обучения перечислены в табл. 3. На рис. 5 представлены графики функции ошибки во время обучения на обучающей и валидационной выборках.

Табл. 3. Гиперпараметры обучения YOLO.

Table 3. Hyperparameters of YOLO Training.

Параметр	Значение
Количество эпох	300
Размер пакета	30
Оптимизатор	AdamW
Начальный темп обучения	0,00149
Инерция	0,9

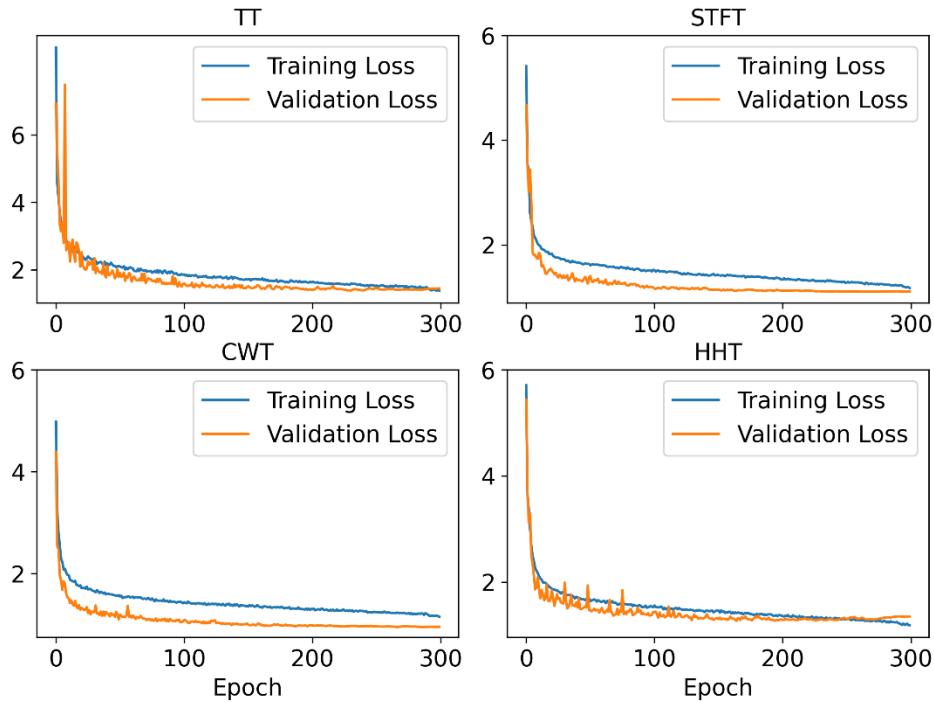


Рис. 5. Графики функции ошибки на обучающей и валидационной выборках.

Fig. 5. Training and Validation Loss Curves.

В табл. 4 отражены значения сбалансированной средней точности mAP50 и mAP50-95 для обученных моделей YOLO на разных видах изображений. В целом, все виды преобразований сигналов в изображения в интеграции с YOLO позволяют получить высокие показатели точности обнаружения полезных сигналов вихревых дефектограмм.

Табл. 4 Сбалансированная средняя точность mAP обученных моделей YOLO.

Table 4. Mean Average Precision (mAP) of Trained YOLO Models.

Способ формирования изображений	mAP50	mAP50-95
Пороговое преобразование	0,9792	0,9694
Окноное преобразование Фурье	0,9790	0,9656
Непрерывное вейвлет-преобразование	0,9823	0,9742
Преобразование Гильберта-Хуанга	0,9292	0,9203

Лучшие результаты показывает вейвлет-преобразование. Пороговое преобразование выигрывает у других методов простотой вычислений и линейной трудоёмкостью. На создание одного изображения ушло 372 мс для STFT, 97 мс – CWT, 9716 мс – ННТ и всего 61 мс – для порогового преобразования ТТ. Среди частотно-временных методов STFT имеет близкие к CWT показатели mAP, тогда как ННТ им немного уступает. Это можно объяснить используемым подходом к формированию гильбертова спектра многоканального сигнала,

когда каналы объединяются простым усреднением спектра и не используется синий канал RGB формата изображения. Видимо, следует рассмотреть более надёжные методы ННТ, например, ННТ с Multidimensional Ensemble Empirical Mode Decomposition (MEEMD) [29]. Также за счёт RGB формата можно соединить несколько частотных представлений сигнала в одном изображении, например, как в работе [16].

Матрицы ошибок (рис. 6) показывают, что при использовании порогового преобразования сеть, несмотря на высокие показатели mAP, делает наибольшее число ложноположительных (FP) предсказаний по сравнению с остальными преобразованиями. Наименьшее число ложноположительных предсказаний происходит при интеграции сети YOLO с STFT. При этом преобразование Фурье практически не уступает другим методам по числу истинноположительных предсказаний (TP), а в случае электроконтактных сварок показывает наилучшие результаты. Преобразование Гильберта-Хуанга хуже остальных находит полезные сигналы и многие из них пропускает. Пороговое преобразование лучше реагирует на алюминиотермитные сварки. Таким образом, с точки зрения матрицы ошибок использование STFT по сравнению с другими преобразованиями обладает рядом преимуществ: происходит наименьшее число ложных предсказаний, находятся оказавшиеся наиболее сложными для обнаружения сетью электроконтактные сварки и практически не пропускаются другие полезные сигналы.

		TT				STFT			
Ground Truth	Predicted	Flash Butt Weld	Alumin. Weld	Bolt Joint	Background	Flash Butt Weld	Alumin. Weld	Bolt Joint	Background
	Flash Butt Weld	274	1	0	27	286	4	0	12
	Alumin. Weld	1	189	0	0	6	182	0	2
	Bolt Joint	0	0	15	0	0	0	15	0
	Background	13	17	0	0	11	1	1	0
		CWT				HHT			
Ground Truth	Predicted	Flash Butt Weld	Alumin. Weld	Bolt Joint	Background	Flash Butt Weld	Alumin. Weld	Bolt Joint	Background
	Flash Butt Weld	285	3	0	14	259	1	0	42
	Alumin. Weld	3	186	0	1	3	180	0	7
	Bolt Joint	0	0	15	0	0	0	12	3
	Background	19	7	3	0	6	6	1	0

Рис. 6. Матрицы ошибок обученных моделей YOLO.
Fig. 6. The Confusion Matrixes of Trained YOLO Models.

На рис. 7 представлены примеры работы сетей YOLO, интегрированных с каждым из рассмотренных преобразований, на фрагменте дефектограммы с рис. 1.

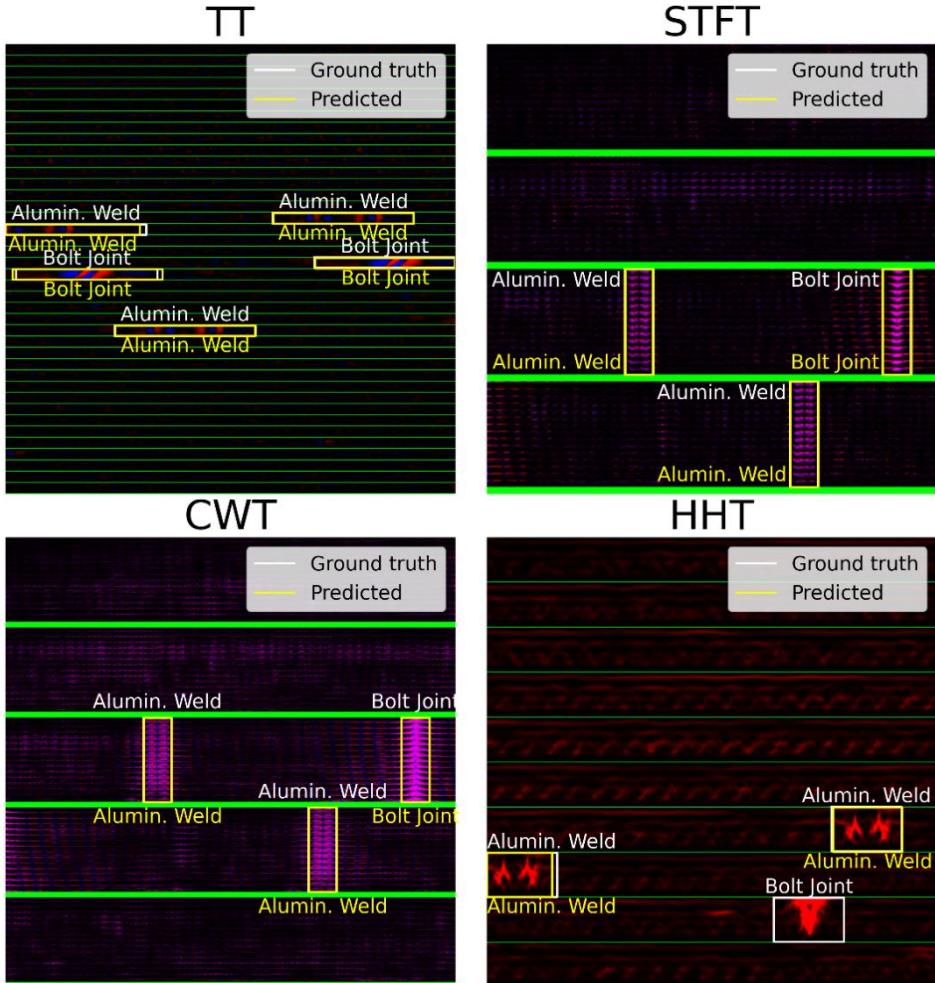


Рис. 7. Примеры работы обученных моделей YOLO.
Fig. 7. The Example of Inference of Trained YOLO Models.

5. Заключение

Задача обнаружения полезных сигналов может решаться разными методами. Один из подходов основан на свёрточных нейронных сетях, которые в последнее время стали основным инструментом в решении задач компьютерного зрения. Для применения этого подхода сигналы необходимо преобразовать в изображения. Существует множество способов выполнить такое преобразование. В данной статье применительно к вихревоковым дефектограммам рельсов, которые представляют собой 15-канальный дискретный сигнал, рассмотрены оконное преобразование Фурье (STFT), непрерывное вейвлет-преобразование (CWT), преобразование Гильберта-Хуанга (HHT) и пороговое преобразование (TT). Все методы кроме TT строят частотно-временной спектр сигналов. Предложенное авторами пороговое преобразование основано на простом сопоставлении каждой амплитуде сигнала значения интенсивности цвета пикселя.

Рассмотренные преобразования можно интегрировать вместе с детектирующей свёрточной нейронной сетью YOLO. В результате было обучено 4 сети YOLO для изображений каждого вида. Наименьшими временными затратами на формирование изображений обладает пороговое преобразование. Наибольшие значения mAP показала сеть на основе CWT. Лучший баланс между точностью и полнотой нахождения полезных сигналов показала сеть на основе STFT. В целом, каждое из преобразований позволило получить высокие показатели метрик.

Таким образом, для вихревоковых дефектограмм могут использоваться разные способы формирования изображений в зависимости от требований на время выполнения, количество пропусков и ложных срабатываний, значимости полезных сигналов определённого класса и т. д. Результаты исследования показывают перспективность применения сетей семейства YOLO при решении задачи обнаружения полезных сигналов вихревоковых дефектограмм и одномерных данных в целом. В дальнейшем планируется рассмотреть подходы к обнаружению полезных сигналов на основе других видов изображений и с использованием одномерной детектирующей свёрточной сети, построенной на основе архитектуры YOLO, для которой не будет требоваться трансляция сигналов в изображения.

Список литературы / References

- [1]. Kuzmin E. V., Gorbunov O. E., Plotnikov P. O., Tyukin V. A. Finding the Level of Useful Signals on Interpretation of Magnetic and Eddy-Current Defectograms. Automatic Control and Computer Sciences, vol. 52, 2018, pp. 658-666. DOI: 10.3103/S0146411618070179.
- [2]. Быстров Л. Ю., Гладков А. Н., Кузьмин Е. В. Подавление аддитивных периодических низкочастотных помех на вихревоковых дефектограммах. Моделирование и анализ информационных систем, 31(2), с. 164-181. DOI: 10.18255/1818-1015-2024-2-164-181. / Bystrov L.Y., Gladkov A.N., Kuzmin E.V. Suppression of additive periodic low-frequency interference on eddy current defectograms. Modeling and Analysis of Information Systems, 31(2), 2024, pp. 164-181 (In Russian).
- [3]. Chao-Lung Y., Zhi-Xuan C., Chen-Yi Y. Sensor Classification Using Convolutional Neural Network by Encoding Multivariate Time Series as Two-Dimensional Colored Images. Sensors, 20(1), 2019, article no. 168. DOI: 10.3390/s20010168.
- [4]. Terven J., Cordova-Esparza, D. M., Romero-González J. A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Machine Learning and Knowledge Extraction, 5(4), 2023, pp. 1680-1716. DOI: 10.3390/make5040083.
- [5]. Hussain M. YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. Machines, 11(7), 2023, article no. 677. DOI: 10.3390/machines11070677.
- [6]. Wyse L. Audio Spectrogram Representations for Processing with Convolutional Neural Networks. 2017. DOI: 10.48550/arXiv.1706.09559. – Available at: <https://arxiv.org/abs/1706.09559>.
- [7]. Junior R. F. R. et al. Fault Detection and Diagnosis in Electric Motors Using Convolution Neural Network and Short-Time Fourier Transform. Journal of Vibration Engineering & Technologies, vol. 10, 2022, pp. 1-12. DOI: 10.1007/s42417-022-00501-3.
- [8]. Modupalli V. P. CNN-Based Machine Tool Monitoring with STFT Image Analysis. International Journal for Research in Applied Science and Engineering Technology, 12(5), 2024, pp. 3986-3991. DOI: 10.22214/ijraset.2024.62493.
- [9]. Yu Y., Liu Q., Han B. S., Zhou W. Application of Convolutional Neural Network to Defect Diagnosis of Drill Bits. Applied Sciences, 12(21), 2022, article no. 10799, 13 p. DOI: 10.3390/app122110799.
- [10]. Cheng C., Seo H., Zhao Y. A Novel Pavement Transverse Cracks Detection Model Using WT-CNN and STFT-CNN for Smartphone Data Analysis. International Journal of Pavement Engineering, 23(12), 2021, pp. 1-13. DOI: 10.1080/10298436.2021.1945056.
- [11]. Kang X. et al. Optimal Preprocessing for Joint Detection and Classification of Wireless Communication Signals in Congested Spectrum Using Computer Vision Methods. 2024. DOI: 10.48550/arXiv.2408.06545. – Available at: <https://arxiv.org/abs/2408.06545>.

- [12]. Liu C., Wang D., Lin Y., Song, S. Research on Online Monitoring of Chatter Based on Continuous Wavelet Transform and Convolutional Neural Network – Vision Transformer (CNN-ViT). *Mechanical Sciences*, 16(1), 2025, pp. 167-180. DOI: 10.5194/ms-16-167-2025.
- [13]. Lupea I., Lupea M. Continuous Wavelet Transform and CNN for Fault Detection in a Helical Gearbox. *Applied Sciences*, 15(2), 2025, article no. 950, 23 p. DOI: 10.3390/app15020950.
- [14]. Pałczyński C., Olejnik P. Anomalies Classification in Fan Systems Using Dual-Branch Neural Networks with Continuous Wavelet Transform Layers: An Experimental Study. *Information*, 16(2), 2025, article no. 71, 13 p. DOI: 10.3390/info16020071.
- [15]. Piedad E. J., Rosario C., Prieto-Araujo E., Bellmunt O. Exploring Wavelet Transformations for Deep Learning-Based Machine Condition Diagnosis. In Proc. of the International Conference on Diagnostics in Electrical Engineering (Diagnostika), 2024, pp. 1-4. DOI: 10.1109/Diagnostika61830.2024.10693895.
- [16]. Siddique M. F., Ahmad Z., Ullah N., Kim J. A Hybrid Deep Learning Approach: Integrating Short-Time Fourier Transform and Continuous Wavelet Transform for Improved Pipeline Leak Detection. *Sensors (Basel)*, 23(19), 2023, article no. 8079. DOI: 10.3390/s23198079.
- [17]. Ch Vidyasagar K. E. et al. Signal to Image Conversion and Convolutional Neural Networks for Physiological Signal Processing: A Review. *IEEE Access*, vol. 12, 2024, pp. 66726-66764. DOI: 10.1109/ACCESS.2024.3399114.
- [18]. Alvarenga T. A. et al. Detection and Classification System for Rail Surface Defects Based on Eddy Current. *Sensors*, 21(23), 2021, article no. 7937. DOI: 10.3390/s21237937.
- [19]. Sagar B., Kane P. Intelligent Bearing Fault Diagnosis of Switched Reluctance Motor Using Hilbert-Huang Transformation and Region-Based Convolutional Neural Network. *The International Journal of Acoustics and Vibration*, 29(4), 2024, pp. 431-441. DOI: 10.20855/ijav.2024.29.42072.
- [20]. Srivastava S. et al. Comparative analysis of deep learning image detection algorithms. *Journal of Big Data*, 8(66), 2021. DOI: 10.1186/s40537-021-00434-w. Available at: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00434-w>.
- [21]. Po-Liang Y., Pei-Ling L. Automatic Detection of Crack Echo Signals in Spectrogram Using Yolo. Available at: <https://ssrn.com/abstract=4885492>. DOI: 10.2139/ssrn.4885492.
- [22]. Vargas V. M. et al. Gramian Angular and Markov Transition Fields Applied to Time Series Ordinal Classification. *Advances in Computational Intelligence*, 2023, pp. 505-516. DOI: 10.1007/978-3-031-43078-7_41.
- [23]. Zhang Y., Chen X. Motif Difference Field: A Simple and Effective Image Representation of Time Series for Classification. 2020. DOI: 10.48550/arXiv.2001.07582. – Available at: <https://arxiv.org/abs/2001.07582>.
- [24]. Jun-Hao C., Yun-Cheng T. Dynamic Deep Convolutional Candlestick Learner. 2022. DOI: 10.48550/arXiv.2201.08669. – Available at: <https://arxiv.org/abs/2201.08669>.
- [25]. Dan-Feng W. et al. Planetary-Gearbox Fault Classification by Convolutional Neural Network and Recurrence Plot. *Applied Sciences*, 10(3), 2020, article no. 932. DOI: 10.3390/app10030932.
- [26]. Som A. et al. PI-Net: A Deep Learning Approach to Extract Topological Persistence Images. In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 3639-3648. DOI: 10.1109/CVPRW50498.2020.00425.
- [27]. Samani E. U., Banerjee A. G. Persistent Homology Meets Object Unity: Object Recognition in Clutter. In Proc. of the IEEE Transactions on Robotics, vol. 40, 2024, pp. 886-902. DOI: 10.1109/TRO.2023.3343994.
- [28]. Yang Y., Deng J., Kang D. An Improved Empirical Mode Decomposition by Using Dyadic Masking Signals. *Signal, Image and Video Processing*, vol. 9, 2015, pp. 1259-1263. DOI: 10.1007/s11760-013-0566-7.
- [29]. Zhaohua W., Jiaxin F., Fang-Li Q., Zhe-Min T. Fast Multidimensional Ensemble Empirical Mode Decomposition for the Analysis of Big Spatio-Temporal Datasets. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 2016, article no. 20150197. DOI: 10.1098/rsta.2015.0197.

Информация об авторах / Information about authors

Артемий Николаевич ГЛАДКОВ – ассистент кафедры теоретической информатики. Научные интересы: цифровая обработка сигналов, машинное обучение, прикладная статистика.

Artemy Nikolaevich GLADKOV – assistant of the Department of Theoretical Computer Sciences.
Research interests: digital signal processing, machine learning, applied statistics.

Леонид Юрьевич БЫСТРОВ – ассистент кафедры теоретической информатики. Научные интересы: цифровая обработка сигналов, машинное обучение, топологические методы анализа динамических систем, математическая статистика.

Leonid Yurievich BYSTROV – assistant of the Department of Theoretical Computer Sciences.
Research interests: digital signal processing, machine learning, topological methods of dynamic systems analysis, mathematical statistics.

Егор Владимирович КУЗЬМИН – доктор физико-математических наук, доцент, заведующий кафедрой теоретической информатики. Научные интересы: цифровая обработка сигналов, методы верификации программ и теоретические основы информатики.

Egor Vladimirovich KUZMIN – Dr. Sci. (Phys.-Math.), Associate Professor, Head of the Department of Theoretical Computer Sciences. Research interests: digital signal processing, program verification methods and theoretical foundations of computer sciences.

