

Повышение качества обработка видеоизображений на основе использования искусственной нейронной сети Yolo

В. Е. Дементьев
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
dve@ulntc.ru

А.Г. Ташлинский
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
tag54@mail.ru

М. Н. Суетин
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
source81@yandex.ru

Аннотация—В работе рассмотрены вопросы обработки видеоизображений металлических конструкций, потенциально содержащих дефекты. Такую обработку предлагается выполнять на основе сверточных нейронных сетей. Для повышения качества обработки предлагаются варианты использования результатов обработки предшествующих кадров видеопоследовательности. Делается вывод о состоятельности и эффективности таких подходов.

Ключевые слова— Обнаружение аномалий, нейронные сети, обработка видеопотока, траекторное слежение.

1. ВВЕДЕНИЕ

В последнее время поиск объектов интереса на видеоизображениях находит широкое применение в различных областях и становится ключевой темой многих исследований [1-7]. В этих исследованиях обработка видеоизображений связана с решением таких задач как сопровождение и распознавание объектов интереса на последовательности кадров, которые основываются на детектировании объектов и предъявляют высокие требования к локализации объектов интереса на каждом кадре последовательности. Впечатляющие результаты по скорости обработки и качеству локализации объектов интереса демонстрируют подходы к детектированию связанные с применением нейросетевых архитектур (R-CNN, YOLO и т.д.) [1-4]. В работах [5-6] было показано, как на базе применения этих архитектур возможно получить приемлемое качество решения важной задачи обнаружения и оценки параметров дефектов на отдельных изображениях металлических конструкций. Привлекательной при этом выглядит идея совместной обработки нескольких изображений в видеопоследовательности для повышения качества детектирования и оценки параметров дефектов. Однако проведенные эксперименты показали, что перемещение регистрирующей камеры в ходе движения беспилотного аппарата накладывает свои ограничения на процесс обработки видеоизображений, связанные с искажением характеристик отдельных кадров, и ведет к потере качества детектирования. Искажение видеокadres приводит к пропаданию объекта интереса на очередном кадре и существенному ухудшению качества распознавания в целом.

2. ОПИСАНИЕ РЕШЕНИЯ

Для решения указанных проблем предлагается использовать информацию, полученную из предшествующих кадров видеопотока и содержащую результаты детектирования объектов интереса. Для этого выполняется последовательное совмещение кадров в последовательности на основе быстродействующих псевдоградиентных алгоритмов [7]. Оценки смещения могут быть использованы для прогнозирования области размещения объекта на последующих кадрах. Особенно удобным подобная процедура является в случае применения нейросетевых конструкций семейства Yola, поскольку выходной информацией сетей такой архитектуры являются прямоугольные окна, обрамляющие объекты интереса. Тогда, зная параметры деформаций, связывающих предшествующие и текущий кадр, возможно просто получить потенциальное место размещения дефекта на последующем кадре. Для этого достаточно выполнить расчет мест размещения четырех точек, являющихся углами выделенного нейронной сетью прямоугольника на предшествующих кадрах. В случае, если эти точки оказываются в пределах следующего кадра, соответствующий четырехугольник и будет местом потенциального размещения дефекта. Данный подход привлекателен в том числе и тем, что видеоизображения, получаемые с беспилотных платформ, обычно не содержит временных разрывов, и качество их совмещения обычно [5] очень хорошее.

Наличие оценок возможного размещения объекта на последующем кадре вкупе с дополнительной информацией о качестве выполненных псевдоградиентных релаксаций [7], позволяют варьировать коэффициент уверенности (confidence coefficient) нейронной сети в районе возможной локализации объекта. При этом в случае потери объекта на очередном кадре видеопоследовательности алгоритм повышает коэффициент уверенности нейронной сети при предсказании места объекта и тем самым стабилизирует результат работы детектора. Поиск необходимых зависимостей коэффициента уверенности от характеристик обрабатываемой видеопоследовательности и свойств детектируемого объекта также может быть выполнен на основе релейных псевдоградиентных процедур. Объект считается потерянным и исключается из алгоритма прогнозирования только в том случае, если он стабильно

не детектируется на заданном числе последовательных кадров.

3. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

В ходе экспериментов, проведенных на тестовых видеоизображениях железнодорожных конструкций, была выполнена оценка стабильности детектирования областей внимания и трещин с помощью нейросетевой архитектуры YOLOv3 при различных коэффициентах уверенности нейросети в предсказаниях. В качестве гиперпараметров, влияющих на уверенность нейросети YOLOv3 в предсказаниях, были использованы пороговый коэффициент при отсеивании выявленных объектов интереса нейросетью (threshold) и пороговый коэффициент при отсеивании объектов интереса алгоритмом Non-Maximum Suppression (threshold NMS). При этом сравнивались результаты детектора, выполняющего независимую обработку кадров в видеопоследовательности (алгоритм 1), процедуры, предполагающей снижение коэффициентов threshold и threshold NMS в области прогнозируемого размещения дефекта на k процентов (алгоритм 2), процедуры, выбирающей коэффициенты threshold и threshold NMS в области прогнозируемого размещения дефекта в соответствии с предварительно проведенной процедурой псевдоградиентного поиска (алгоритм 3).



Рис. 1. Пример обработки кадра видеопоследовательности

В таблице 1 представлены зависимости средних вероятностей правильного обнаружения дефектов и вероятностей ложных тревог от вида используемого алгоритма.

Таблица 1. АНАЛИЗ РАБОТЫ АЛГОРИТМОВ

Алгоритм	Оценка вероятности правильного обнаружения	Оценка вероятности ложной тревоги
Алгоритм 1	0,56	0,01
Алгоритм 2, $k=10$	0,62	0,01
Алгоритм 2, $k=20$	0,66	0,02
Алгоритм 2, $k=40$	0,73	0,04
Алгоритм 2, $k=60$	0,79	0,13
Алгоритм 3	0,78	0,03

Предложенный в работе подход показал высокие результаты по стабилизации результатов детектирования видеоизображений. При этом заметно, что наилучшие результаты показывает процедура, предполагающая предварительный поиск зависимостей коэффициентов от свойств обрабатываемых изображений и типов детектируемых объектов. В то же время такая процедура сама по себе не является тривиальной и, по сути, также предполагает проведение некоторого обучения относительно свойств обрабатываемого материала. Поэтому на практике возможно ограничиться пропорциональным снижением соответствующих порогов уверенности, обеспечивая до 15% повышения вероятности верного обнаружения дефекта при незначительном увеличении вероятности ложной тревоги. Отметим особо, что рассматриваемые в работе процедуры сами по себе почти никак не сказываются на скорости обработки видеоматериала. Так на видеокарте Nvidia Quadro RTX 3000 обработка видеоизображения алгоритмами 1, 2 и 3 выполняется со скоростью примерно 30 кадров в секунду, что позволяет вести обработку видеопотока в режиме реального времени.

4. ЗАКЛЮЧЕНИЕ

Таким образом, в работе предложено семейство процедур, позволяющих поднять качество нейросетевых детекторов при обработке видеопоследовательностей. Выполнена оценка эффективности этих процедур на реальном материале, по результатам которой можно сделать вывод о повышении качества обработки без принципиального снижения производительности.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-29-09048.

ЛИТЕРАТУРА

- [1] Kalaivani, R. Object Detection in Video Frames Using Various Approaches / R. Kalaivani, C.R. Manicha // International Journal of Advanced Research in Computer and Communication Engineering. – 2013. – Vol. 2(9). – P. 157-160.
- [2] Redmon, J. You Only Look Once: unified, real-time object detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi [Electronical Resource]. – Access mode: <http://itnt-conf.org> (01.02.2022).
- [3] Padmanabula, S.S. Object Detection Using Stacked YOLOv3 / S.S. Padmanabula, V. Sistla, V.K.K. Kolli // International Information and Engineering Technology Association. – 2020. – Vol. 25(5). – P. 691-697.
- [4] Gai, W. An improved Tiny YOLOv3 for real-time object detection / W. Gai, Y. Liu, J. Zhang, G. Jing // System Science & Control Engineering: an open access journal. – 2021. – Vol. 9(1). – P. 314-321.
- [5] Dementiev, V.E. Detecting anomalies in temporal image sequences based on object identification by the stochastic gradient adaptation / V.E. Dementiev, R.G. Magdeev, A.G. Tashlinskii // IEEE Xplore. International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-5.
- [6] Dementiev, V. Improving The Quality of Detecting and Recognizing Defects in Images of Steel Structures / V. Dementiev, M. Suetin, M. Gaponova // IEEE Xplore. International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-5.
- [7] Ташлинский, А.Г. Оценивание параметров пространственных деформаций последовательностей / А.Г. Ташлинский. – Ульяновск : Издательство УлГТУ, 2000. – 139 с.