

УДК 004.896

АЛГОРИТМ ОБНАРУЖЕНИЯ ОБЪЕКТА НА ВИДЕОПОТОКЕ НА ОСНОВЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Юдина Д.А.

магистр,

*Московский государственный технический университет им. Н.Э. Баумана,
Москва, Россия*

Аннотация

В данной статье рассматриваются проблемы автоматизированного распознавания объекта на видеопотоке на основе сверточной нейронной сети. В ходе работы были выявлены трудности, встречающиеся при обнаружении заданного объекта на видеопотоке и оказывающие негативное влияние на данный процесс, также проанализированы существующие методы, с помощью которых можно решить данную задачу. Предложен алгоритм, основанный на использовании сверточной нейронной сети, причем для повышения качества работы сети применяются предобработка кадра (изменение экспозиции) и алгоритм отделения движущегося фона от статического. Отделение переднего фона позволило избавиться от некоторых ложных срабатываний, возникающих при отсутствии необходимого объекта на изображении. Как итог было проведено исследование эффективности работы разработанного алгоритма автоматизированного обнаружения объекта на видео. В качестве объекта рассматривался патрубок. На тестовой выборке точность распознавания патрубка составила свыше 92%, что является хорошим результатом.

Ключевые слова: сверточная нейронная сеть, convolution neural network, отделение движущегося фона от статического, background subtractor, предобработка кадров

ALGORITHM FOR DETECTING AN OBJECT ON VIDEO BASED ON CONVOLUTIONAL NEURAL NETWORK

Yudina D.A.

master,

Bauman Moscow State Technical University,

Moscow, Russia

Annotation

This article examined the problems of automated object recognition on video stream based on convolutional neural network. In the course of the work, the difficulties encountered in detecting given object on video stream and which have a negative impact on this process were identified. Also existing methods were analyzed to solve this problem. It is proposed algorithm based on using convolutional neural network. This algorithm also contains frame preprocessing (exposure change) and an algorithm for separating a moving background from a static background that are used to improve the quality of the network. The separation of the foreground made it possible to get rid of some false positives that occur when there is no necessary object on image. As a result, a study was conducted of the effectiveness of the developed algorithm for automated object detection on video. Branch pipe was selected as object. On the test sample, the recognition accuracy of branch pipe was over 92%.

Keywords: convolutional neural network, background subtractor, frame processing

Современные процессы в промышленности требуют повышения качества выпускаемой продукции и постоянного контроля производства. Это сложно реализовать без применения современных методов обработки информации, а именно без методов распознавания образов. Методы распознавания образов позволяют автоматизировать данный процесс и эффективно его решать.

Понятие «объект» является широким, в статье под данным понятием будет пониматься патрубок. Патрубком называется небольшая тонкая трубка, которая является отводом в трубопроводах, резервуарах и тому подобное. По этой специальной трубке поступает аргон в сталеплавильную печь с целью уменьшения поступления различных примесей в выпускаемый продукт. Патрубок изображен на рисунке 1. Под цифрой 1 изображен патрубок, под цифрой 2 - труба, под цифрой 3 - конус. Конус - место крепления патрубка к трубе.

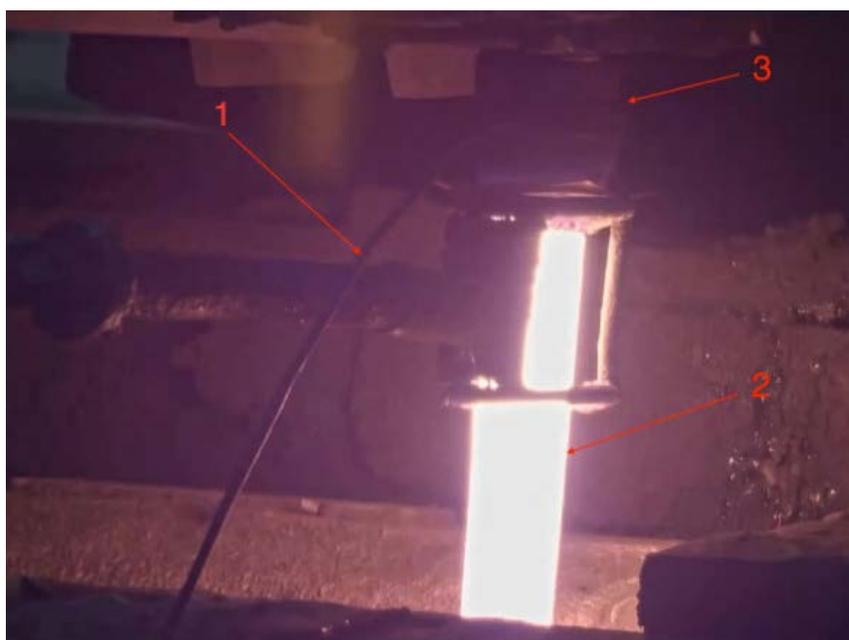


Рис. 1 - Патрубок

Распознавание образов [5, 6, 8] в промышленности осложняется наличием множества специфических факторов и особенностей: освещение, наличие дыма и брызгов, особенно если процесс связан с производством металла. В процессе обнаружения патрубка на изображение/видео встречаются следующие сложности: темное видео, отсутствие стандартного положения патрубка, форма патрубка различна, патрубок внешне схож с другими объектами, наличие дыма и брызгов, возникающих в процессе плавки стали.

Исходя из описанных выше трудностей, возникающих в процессе обнаружения патрубка на изображение, для повышения качества распознавания помимо обнаружения патрубка будет еще распознавать объект «конус», так как он является местом крепления патрубка и данная информация поможет принять более правильное решение в процессе распознавания.

1 Анализ методов распознавания объектов

В настоящее время существует ряд подходов, используемых для обнаружения необходимых объектов на изображение. Принцип действия всех методов разный. Некоторые методы обрабатывают образы целиком, сравнивая исследуемое изображение с заранее подготовленной базой шаблонов или по заранее заданной форме; некоторые - занимаются поиском определенных признаков на изображение или детальным анализом всего изображения. Наиболее интересным является метод, основанный на использование нейронных сетей [9]. Сравнение методов представлено в таблице 1.

Таблица 1 - Сравнение методов распознавания

Метод	Достоинства	Недостатки	Неприменимость к данной задаче
Метод перебора	1. Простота использования	1. Низкая скорость распознавания при большой базе эталонов 2. Распознаваемые образы должны быть «идеальными», то есть приближены к базе эталонов 3. Трудности при составлении базы эталонов 4. Необходимо много ресурсов для хранения данных	В кадре могут находиться объекты, внешне похожие на патрубок (свисающий провод, свисающая цепь).

Классификация по положению	1. Скорость распознавания	1. Не все объекты располагаются в определенном месте/положении	Патрубок не имеет определенного положения и на него невозможно повесить вспомогательные метки, ввиду особенности конструкции плавки стали.
Анализ характеристик образа	1. Точность распознавания	1. Составление векторов признаков 2. Выявление уникальных признаков	Затруднительно выделить особенности для объекта «патрубок», так как он не обладает никакими специфическими особенностями
Нейронные сети	1. Универсальность 2. Самообучение 3. Высокая точность распознавания	1. Неоднозначное определение начальных характеристик 2. Трудности при создании обучающей выборки 3. Результат напрямую зависит от обучающей выборки 4. Скорость обучения	+

2 Предобработка кадров

В работе предобработка кадров заключается в удаление шумов и изменение экспозиции. Изменение экспозиции кадра, а именно ее увеличение, позволяет решить проблему, связанную с плохой освещенностью, в результате чего искомый объект начинает сливаться с фоном. Алгоритм изменения экспозиции основан на применение специального линейного фильтра к

изображению (операция свертки). Фильтр имеет размер 3x3 и содержит следующие значения:

-0,05	-0,05	-0,05
-0,05	2	-0,05
-0,05	-0,05	-0,05

Значения для фильтра были подобраны опытным путем, так, чтобы увеличение экспозиции происходило равномерно, без сильного засвечивания.

3 Отделение динамического фона от статического

Вычитание фона технически основано на извлечение движущегося переднего плана из статического фона [3]. Данные действия могут позволить обнаружить момент присоединения патрубка к конусу, что позволит избавиться от некоторого количества ложных срабатываний, возникающих при работе нейронной сети.

Алгоритм определения переднего фона заключается в сегментации фона, основанной на гауссовском распределении. Важной особенностью данного алгоритма является выбор числа гауссовского распределения для каждого пикселя. Описанная особенность обеспечивает лучшую адаптацию к различным сценам из-за изменений освещения. В алгоритме учитываются такие параметры как история (через какое целое количество кадров произойдет «замена» исходного заднего фона) и порог (допустимая дельта разности между фоновым изображением и текущим).

4 Архитектура разработанной сверточной нейронной сети

Структура сверточной нейронной сети была разработана согласно современным тенденциям развития в данной области. В работе сеть на вход принимает поток изображений размером (416x416) и состоит из 32-х слоев. Архитектура разработанной нейронной сети представляет собой последовательное чередование сверточных слоев со слоями субдискретизации. Дневник науки | www.dnevniknauki.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

Периодически сверточные слои следовали друг за другом. Все сверточные слои разработанной нейронной сети имеют функцию активации Leaky ReLu, так как она позволяет решить проблему, связанную с «умиранием» нейронов. Все слои субдискретизации содержат фильтр, представляемый в виде матрицы размером 2x2. В процессе субдискретизации происходит поиск наибольшего элемента в фильтре и запоминание этого элемента. Последний слой нейронной сети является полносвязным, у которого функцией активации является функция softmax.

5 Алгоритм обнаружения патрубка на видео

5.1 Вспомогательные параметры

Так как нейронная сеть не может всегда показывать сто процентный результат, были введены некоторые вспомогательные параметры, позволяющие приблизиться к лучшему результату:

- Переменная, отвечающая за время, по истечению которого будет приниматься решение об отсутствии патрубка на изображении. В работе значение данной переменной составляет 60 секунд.
- Переменная, отвечающее за время, в течение которого можно с высокой степенью уверенности сказать, что необходимый объект присутствует на видео. В работе значение данной переменной составляет 5 секунд.
- Переменная, отвечающая за допустимое время, в течение которого может отсутствовать распознаваемый объект. В работе значение данной переменной составляет 2 секунды.

5.2 Блок схема алгоритма обнаружения патрубка на видео

На рисунке 2 представлена блок-схема алгоритма обнаружения патрубка

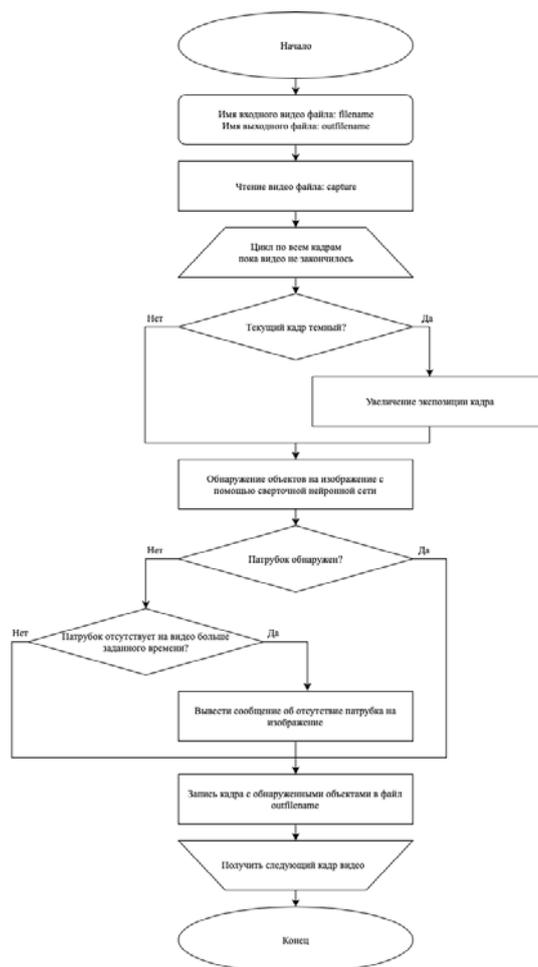


Рис. 2 - Общий алгоритм обнаружения патрубка на изображение

на видео.

6 Обучение разработанной сети

Обучение нейронной сети происходит с помощью алгоритма обратного распространения ошибки [1, 2, 7]. В процессе обучения может возникнуть серьезная проблема, называемая переобучением сети. Для решения этой проблемы в работе происходит разделение обучающей выборки на два подмножества: обучающий и тестовый наборы. С помощью первого набора происходит непосредственное обучение сети, с помощью второго - проверка

точности распознавания построенной модели. Эти наборы должны содержать уникальные непересекающиеся данные.

Обучающая выборка состоит из множества изображений, полученных с 5-ти видеокамер, расположенных в цехах плавки стали. Полученная выборка была размечена на 3 класса: патрубок, конус с патрубком и конус без патрубка. Количество изображений для каждого класса было примерно одинаковым. В обучающей выборке содержались также и изображения, на которых отсутствовали выше перечисленные классы (примерно 10% от общего количества изображений). Исходные данные распределялись в приведенные множества случайным образом в отношении 85 : 15. Всего исходная выборка составила более 10 500 изображений. Пример разметки представлен на рисунке



Рис. 3 - Пример разметки

3.

7 Исследование разработанного метода

Тестирование разработанного алгоритма производилось на наборе видео, полученных с 5-ти камер видеонаблюдения, установленных в цехах плавки

Дневник науки | www.dnevniknauki.ru | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

стали. Всего для тестирования использовалось 91-го видео продолжительностью от 30 секунд и до полутора минут. Разрешение видео составляло 1280x720 с частотой кадров 15 fps. Данные видео содержали различные ситуации по положению патрубка: на некоторых видео патрубка не было, где-то происходило отсоединение/присоединение патрубка от/к конуса/конусу, где-то патрубок присутствовал на видео, однако он не был прикреплен к конусу, также присутствовали затемненные видео.

Всего было проведено 11 экспериментов (таблица 2), каждый из которых заключался в разном обучении нейронной сети: различные множества обучающей выборки и различная скорость обучения.

Таблица 2 - Характеристика проведенных экспериментов

Характеристика	Характеристика
Тест 1	Небольшая обучающая выборка (до 1000 изображений). Причем положение патрубка было разделено на 3 класса: «нормальное положение», когда патрубок полностью виден на кадре и ничем не перекрыт, либо перекрыт незначительно; «положение вдоль трубы» и «плохое положение», когда патрубок подведен сбоку и перекрыт сторонними конструкциями (чаще всего такого рода видео оказывались темными).
Тест 2	Аналогичен тесту 1, однако был изменен размер входного видео (уменьшен вдвое).
Тест 3	Выборка из теста 1, однако обучение модели происходило сразу на все ракурсы, также был изменен размер входного видео (уменьшен вдвое).
Тест 4	Дообучение модели, полученной в тесте 3. Результирующее распознавание использовалось без изменения размера видео.
Тест 5	Новая расширенная обучающая выборка (чуть более 4500 изображений).
Тест 6	Дообучение модели, полученной в тесте 5.
Тест 7	Использовалась модуль, размеченная на три класса: патрубок, конус с парубком и конус без патрубка. Выборка использовалась такая же как и в предыдущем тесте, но она была увеличена вдвое за счет размытия всех изображений.
Тест 8	Дообучение модели, полученной в тесте 7.

Таблица 2 - Характеристика проведенных экспериментов

Тест 9	Выборка из предыдущего теста была дополнена новыми ракурсами и из нее были удалены искусственно размытые кадры (около 6500 изображений).
Тест 10	В данном тесте была изменена скорость обучения (уменьшена в 2 раза) и изменены множества для обучающей и тестовой выборки (случайное перераспределение изображений).
Тест 11	Увеличенная обучающая выборка (около 10500 изображений).

Средняя точность обученной модели для каждого из проведенных экспериментов представлена в таблице 3.

Таблица 3 - Сравнение уверенности предсказания патрубка

Название теста	Точность распознавания патрубка, %	Средняя уверенность определения патрубка	Средняя уверенность определения конуса
Тест 1	78,1	0,84	-
Тест 2	79,2	0,84	-
Тест 3	75,6	0,84	-
Тест 4	84,6	0,88	-
Тест 5	87,9	0,88	-
Тест 6	90,3	0,87	-
Тест 7	85,6	0,85	0,83
Тест 8	93,4	0,88	0,87
Тест 9	86,8	0,86	0,85
Тест 10	89,1	0,87	0,86
Тест 11	93,4	0,86	0,82

Как видно из таблицы 3, наивысшую точность и уверенность предсказания показал тест 8, как для патрубка, так и для конуса.

Наилучшую суммарную точность распознавания патрубка на тестовых видео также показала модель, обученная в 8-м тесте. Данная модель позволила обнаружить сложную ситуацию, когда патрубок расположен почти перпендикулярно трубе и перекрыт сторонним сооружением. Причем модель определила необходимый объект с довольно высокой уверенностью (0.86). Также эта обученная модель позволяет распознать патрубок на сильно затемненном видео.

Таким образом, в результате проведенных экспериментов была выбрана модель из 8-го эксперимента. То есть модель, обученная на выборке примерно 9000 изображений: 4500 изображений, полученных с камер видеонаблюдения, и тех же изображений, но подвергнувшихся небольшому размытию.

Заключение

Таким образом, в работе был разработан и реализован алгоритм автоматизированного обнаружения патрубка на видеопотоке. В его основе лежит метод, использующий сверточную нейронную сеть. Для повышения точности обнаружения патрубка использовались предобработка кадров (изменение экспозиции) и отделение движущегося фона от статического. Это позволило решить некоторые проблемы, связанные с обнаружением патрубка на изображении/видеопотоке, и повысить качество распознавания на сильно затемненных кадрах. В результате обучения модели на тестовых данных была достигнута точность более 92%.

Библиографический список:

1. Алгоритм обратного распространения ошибок. [Электронный ресурс]. - Режим доступа - URL: <http://www.aiportal.ru/articles/neural-networks/back-propagation.html> (Дата обращения: 15.05.19)

2. Нейронная сеть. [Электронный ресурс]. - Режим доступа - URL: <http://statsoft.ru/home/textbook/modules/stneunet.html#multilayer> (Дата обращения: 15.05.19)
3. Обработка потоков изображений с целью выделения объекта. [Электронный ресурс]. - Режим доступа - URL: https://storage.tusur.ru/files/11309/АСУ-1402_Ивашко_Мусяенко.Pdf (Дата обращения: 15.05.19)
4. Продукционная модель знаний. [Электронный ресурс]. - Режим доступа - URL: <http://www.aiportal.ru/articles/knowledge-models/production-model.html> (Дата обращения: 15.05.19)
5. Распознавание образов. [Электронный ресурс]. - Режим доступа - URL: https://ru.bmstu.wiki/Распознавание_образов (Дата обращения: 15.05.19)
6. Распознавание образов. Обзорная лекция. [Электронный ресурс]. - Режим доступа - URL: http://it-claim.ru/Persons/Zelencov/Lecture_text.pdf (Дата обращения: 15.05.19)
7. Стохастические методы обучения нейронных сетей. [Электронный ресурс]. - Режим доступа - URL: <http://www.intuit.ru/studies/courses/88/88/lecture/20539?page=1> (Дата обращения: 15.05.19)
8. Б. Яне Цифровая обработка изображений / Яне Б. - М.: Техносфера, 2007.
9. Neural Networks and Deep Learning. [Электронный ресурс]. - Режим доступа - URL: <http://neuralnetworksanddeeplearning.com/> (Дата обращения: 15.05.19)

Оригинальность 96%