

## РАЗРАБОТКА КАСКАДНОГО АЛГОРИТМА МОНИТОРИНГА ДВИЖЕНИЯ ДЕТАЛЕЙ В ХОДЕ ИХ ИЗГОТОВЛЕНИЯ

Киселева П.И.  
Печенина Е.Ю.  
Печенин В.А.

Самарский национальный исследова-  
тельский университет им. академика  
С.П. Королева

Московское шоссе, д. 34, г. Самара,  
443084, Российская Федерация

[v.a.pechenin@ssau.ru](mailto:v.a.pechenin@ssau.ru)

*Разработан каскадный алгоритм, позволяющий проводить идентификацию содержимого в производственных тарах. Алгоритм состоит из двух этапов: детектирование ячеек тары и классификация содержимого каждой ячейки. Предложенный алгоритм позволяет добиться точности классификации в 89% при обучении на относительно меньшем объеме выборки, чем потребовалось бы при использовании алгоритма прямого детектирования деталей без этапа детектирования ячеек. Алгоритм таким образом подходит для использования в системах мониторинга состояния производственной среды в аэрокосмическом производстве.*

**Ключевые слова:** рабочая лопатка; детектирование; алгоритм; свёрточная нейронная сеть; тара

### 1 Введение

Для наукоёмкого промышленного производства самолётов, ракет, автомобилей, металлообрабатывающих станков и газотурбинных двигателей характерны длинные циклы изготовления продукции, отдельные детали проходят несколько стадий (операций) обработки в цеху, перемещаются между цехами внутри предприятия или между предприятиями-смежниками. В аэрокосмической отрасли характерно изготовление деталей малыми партиями при широкой номенклатуре их изготовления. Для того чтобы выполнялся главный календарный план и план производства в соответствии с директивными датами и своевременно производились корректирующие мероприятия, необходимо отслеживать движение материальных потоков и оптимизировать загрузку производственных мощностей по цепочке технологического маршрута. Для этой цели служат системы сбора информации и мониторинга (MDC-системы) и системы оперативного учёта выполнения производственных заданий (MES-системы) [1]. Существующие MES-системы направлены на повышение эффективности производства

за счёт снижения избыточных складских запасов, объёмов незавершённого производства, повышения степени использования производственных ресурсов, уровня организации хранения и логистики, пропускной способности производственной системы и т.п. [2].

В настоящее время существует разрыв между используемыми в производстве программами среднесрочного и оперативного планирования (ERP) и используемыми в производстве системами сбора информации (MDC+MES) [3]. На предприятиях аэрокосмического профиля данные собираются либо со станков, далеко не со всех, либо с терминалов, где присутствует ручной ввод. Такой избирательный характер поступления информации и человеческий фактор существенно снижают её достоверность [4]. В этой связи разработка эффективных алгоритмов интеллектуализации видеонаблюдения для проведения автоматического анализа видеопотока при помощи технологий машинного обучения [5] позволит сократить существующий разрыв в сборе и анализе производственных данных, повышая эффективность планирования и исполнения. Целью исследования является разработка каскадного ал-

горитма, позволяющего решать задачу идентификации состояния производственного передела методами машинного обучения с использованием унифицированной тары для межоперационной логистики.

## 2 Разработанный алгоритм

Разработанный алгоритм приведён на рисунке 1. Входное изображение тары с деталями проходит компьютерную обработку (изменение размера) и поступает на вход в нейронную сеть детектирования ячеек. В качестве архитектуры нейронной сети выступает YOLO. Все детектированные фрагменты, содержащие ячейки, проходят в цикле через нейронную сеть для классификации VGG. Перед классификацией происходит также изменение размера и, при необходимости, цветового режима. Опишем используемые нейронные сети подробнее.

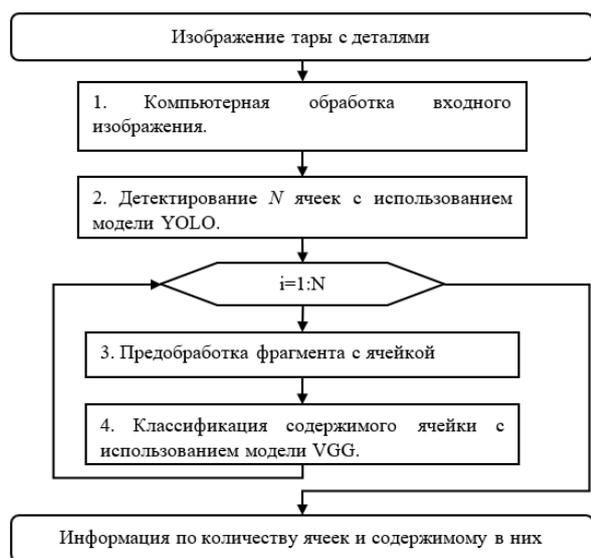


Рисунок 1. Разработанный каскадный алгоритм

### 2.1 Модель детектирования ячеек

Детектирование ячеек в разработанном алгоритме выполняется с использованием нейронной сети архитектуры YOLOv3 [6]. Сеть YOLOv3 содержит 106 слоёв, её особенность состоит в применении основной свёрточной нейронной сети (Darknet-53) один раз ко всему изображению сразу. Это

позволяет работать быстрее и точнее обнаруживать объекты [7].

### 2.2 Модель классификации содержимого ячеек

Задача классификации объекта внутри детектированного фрагмента изображения решалась с использованием свёрточной нейронной сети VGG-19 [8], состоящей из 144 миллионов параметров.

В качестве функции потерь для классификации используется категориальная кросс-энтропия, вычисляемая по формуле:

$$L_{CE} = - \sum_i^C t_i \cdot \log(s_i), \quad (1)$$

где  $C$  – количество классов;  $t_i$  – значение метки для  $i$ -го случая,  $s_i$  – значение на выходе из нейронной сети для  $i$ -го случая.

### 2.3 Оценка качества работы алгоритма

Для оценки точности работы YOLO используется метрика – интерполированная средняя точность (mean average precision,  $mAP$ ). Данная метрика вычисляется на основании средней точности ( $AP$ ). Для расчёта  $AP$  используется в свою очередь метрика  $IoU$  (Intersection over Union), также известная как индекс Жаккара – число от 0 до 1, показывающее, насколько совпадают два объекта (эталонный и детектируемый). Метрика  $IoU$  равна области перекрытия  $S_I$  между прогнозируемой с использованием нейронной сети площадью и размеченной (истинной) областью, разделённой на площадь объединения  $S_U$  между прогнозируемой и истинной областью:

$$IoU = S_I / S_U. \quad (2)$$

При классификации изображений деталей использована метрика доли правильных ответов алгоритма ( $ac$ ):

$$ac = (TP+TN)/(TP+TN+FP+FN), \quad (3)$$

где  $TP$ ,  $FP$ ,  $TN$  и  $FN$  – элементы матрицы ошибок, количество истинно положительных, ложно положительных, истинно отрицательных и ложно отрицательных значений при классификации.

### 3 Эксперименты и результаты

Разработанный алгоритм был реализован на языке Python с использованием

фреймворка Tensorflow и библиотеки OpenCV. Используемая в экспериментах тара и детали (рабочая лопатка турбины) приведены на рисунке 2. Тара, приведённая на рисунке 2, имеет 12 ячеек. Для проведения экспериментов использовались фотографии, сделанные с ip-камер DH-IPC-HFW5442EP-ZE (корпусная цилиндрическая IP-видеокамера 4Мп).

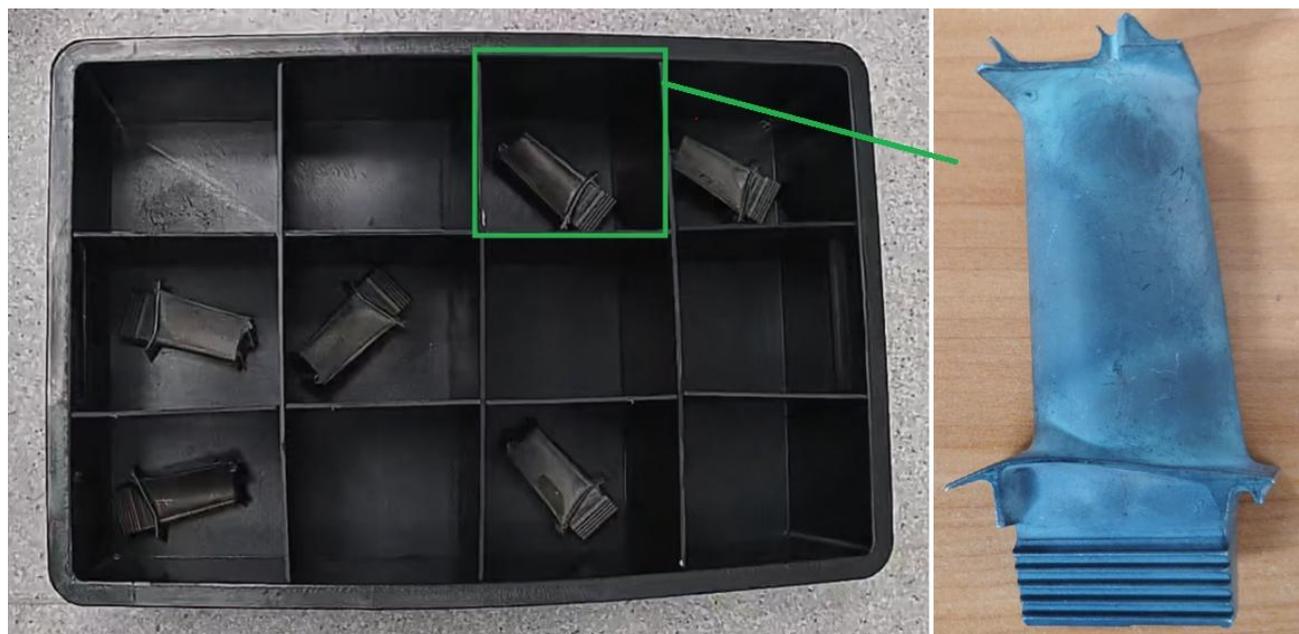


Рисунок 2. Тара с рабочими лопатками турбины

#### 3.1 Обучение сети по детектированию ячеек на таре

Для обучения нейронной сети на детектирование ячеек в таре было использовано 22 изображения тары, выполненных при разном освещении, на разном фоне и с небольшими изменениями ракурса, из которых 16 составили обучающую выборку, 6 – тестовую.

В приложении labelling была выполнена разметка обучающей выборки (пример на рисунке 3), детектируется один класс «ячейка». Кроме того, были использованы 6 тестовых изображений, на которых проверялась точность работы и вычислялась метрика  $mAP$ .

Оптимальный результат при качественном осмотре достигается при 2000 эпохах

обучения. Величина функции потерь  $L$  (1) на 2000 эпохе составила 0,23, средняя величина  $IoU$  достигла значения 0,879. Метрика  $mAP$  на 2000 эпохе составила 100%.

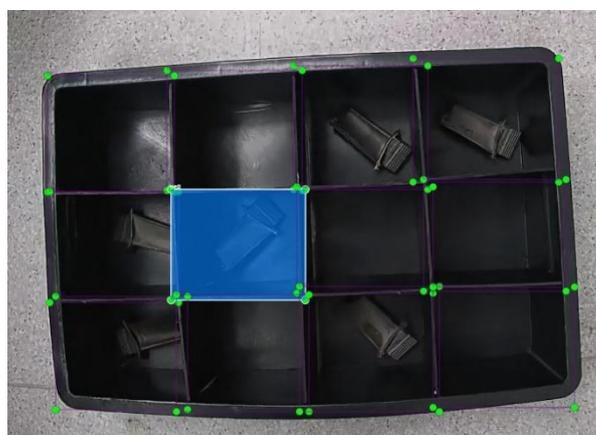


Рисунок 3. Разметка изображений для обучения Yolo

### 3.2 Обучение нейронных сетей для классификации деталей в таре

За этапом детектирования ячеек в таре следует этап классификации содержимого в них (блок-схема на рисунке 1). С использованием обученной на предыдущем этапе сети для детектирования ячеек YOLOv3 с изображений тар с деталями, например как на рисунке 2, выделялись фрагменты с ячейками. Для проведения экспериментов, как отмечалось ранее, использовался 1 вид деталей – рабочая лопатка турбины. Вторым классом являлась пустая ячейка. Для классификации была использована нейронная сеть VGG19. Для оценки точности работы сети VGG19 использовались следующие метрики: доля правильных ответов алгоритма  $ac$  (3) и категориальная кросс-энтропия  $L_{CE}$  (1).

При проведении экспериментов ставилась задача подбора нескольких параметров. Прежде всего выполнялся поиск минимально необходимого размера изображений для обучения и тестирования. Все изображения ячеек делились в пропорции 9:1, то есть размер обучающей выборки составлял 90%, тестовой – 10%. Кроме того, во всех экспериментах использовалась ещё проверочная выборка из 125 изображений, не участвующая в обучении.

Вторым параметром поиска являлся размер батча [9]. Батч – небольшая коллекция данных, которая обычно выбирается из всего обучающего множества случайно. То есть на вход в сеть для обновления весов при обучении одновременно подаётся не одно, а несколько изображений. Это ускоряет обучение и позволяет достигнуть большей обобщающей способности при обновлении весов.

Третьим параметром являлось количество эпох обучения. Необходимо для конкретных условий найти такое число эпох, при которых функция потерь или перестанет уменьшаться на тестовой выборке, или же начинает расти.

Четвёртым параметром являлся цветовой режим изображения: можно либо ис-

пользовать исходное цветное изображение RGB, либо преобразовать его в оттенки серого.

Пятым параметром являлся размер изображения – размерность входной матрицы в нейронной сети. Использовались две размерности: 224x224 и 112x112. Меньшая размерность позволяет экономить место на диске при хранении структуры нейронной сети.

На рисунке 4 приведены графики метрик для тестовой и обучающей выборки в одном из экспериментов, где обучающая и тестовая выборка в сумме составляют 1332 цветных изображения

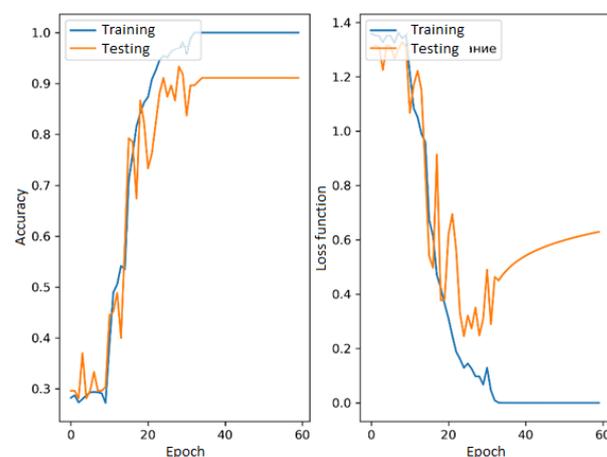


Рисунок 4. Функции потерь и точность при обучении VGG19

Как следует из графиков на рисунке 3.13, начиная с 40 эпохи функция потерь обучающей выборки перестаёт уменьшаться, а для тестовой начинает немного расти, что свидетельствует о переобучении и необходимости остановиться на 38-40 эпохе.

В проводимых экспериментах количество эпох обучения варьировалось от 40 до 60. Ключевыми критериями успешности обучения являлись метрики именно для тестовой и проверочной выборок.

В таблице 1 приведён план экспериментов по подбору параметров сети и получаемая точность для классификации содержимого ячеек.

Таблица 1. Сводная таблица параметров и результатов экспериментов

№	Размер батча	Кол-во эпох	Выборка	Цвет	Размер, пикс.	Обучение		Тестирование		Проверка	
						$L_{CE}$	$ac$	$L_{CE}$	$ac$	$L_{CE}$	$ac$
1	4	50	119	серый	224x224	1,39	27,62	1,38	33,33	1,38	33,33
2	4	50	189	серый	224x224	0,68	70,00	0,86	63,16	0,86	63,16
3	4	50	297	серый	224x224	0,21	94,01	0,72	66,67	1,47	51,85
4	4	50	297	серый	224x224	0,00	100,00	1,35	73,33	3,37	57,41
5	4	50	297	серый	224x224	0,00	100,00	1,12	73,33	0,56	86,67
6	4	50	297	цвет	224x224	0,01	100,00	1,49	63,33	2,80	62,96
7	1	50	297	серый	224x224	1,38	28,46	1,40	23,33	1,42	20,37
8	2	50	297	серый	224x224	0,49	85,39	1,22	66,67	2,40	53,70
9	8	50	297	серый	224x224	6,12	21,35	6,30	26,67	4,70	33,33
10	4	50	369	серый	224x224	0,53	80,12	1,12	56,76	1,21	42,59
11	4	50	369	цвет	224x224	0,94	64,16	1,13	48,65	0,99	50,00
12	4	60	369	серый	224x224	1,38	23,19	1,38	35,14	1,37	37,04
13	4	60	369	цвет	224x224	0,09	96,69	1,66	72,97	2,46	66,67
14	4	60	369	цвет	224x224	0,00	100,00	2,10	67,57	3,50	62,96
15	4	60	1332	цвет	224x224	0,00	100,00	0,63	91,11	1,22	88,89
16	4	40	1332	цвет	112x112	0,00	100,00	0,24	94,81	0,46	88,89
17	4	40	1332	серый	112x112	0,04	98,84	0,30	89,63	0,64	74,07

Оптимальными параметрами обучения для рассматриваемого тестового примера классификации двух случаев содержимого ячеек являются: размер батча 4, количество эпох обучения 40, входной размер цветных изображений в нейронную сеть 112x112; размер обучающей выборки 1406. При таких параметрах метрика  $ac$  на тестовой выборке в 156 картинок достигает 94,81%; на проверочной выборке в 54 изображения 88,89 %.

На рисунке 5 приведён результат работы разработанного каскадного алгоритма на одном из проверочных изображений тары.

#### 4 Заключение

В работе описан разработанный алгоритм идентификации состояния производственного передела, адаптированный под условия производства: детали перемещаются в тарах, необходимо определять их тип и количество для дальнейшего исполь-

зования этой информации в управленческих целях. В алгоритме последовательно производится детектирование ячеек тары и последующая классификация содержимого в них. Результаты экспериментов показали 100% точность при детектировании и 89% точности при классификации.



Рисунок 5. Результат работы алгоритма

Разработанный алгоритм и проведённые исследования в будущем предполагается использовать при разработке программных

средств, позволяющих выполнять предиктивную аналитику потенциальных несоответствий с использованием технологий компьютерного зрения и машинного обучения. При этом необходимо будет предусмотреть идентификатор для каждой тары в виде RFID-метки или QR-кода, позволяющий загружать веса нейронных сетей, обученных на конкретный вид тары и типы деталей в них. Внедрение подобных систем продиктовано, кроме того, мировым трендом и объективной необходимостью создания «цифровых двойников» производственной системы, которые являются методологической основой и инструментарием для решения задач управления производством.

#### **Список использованных источников**

- [1] Introductory overview: Systems and control methods for operational management support in agricultural production systems / S. van Mourik, R. van der Tol, R. Linker [et al.] // *Environmental Modelling & Software*. – 2021. – Vol. 139. – P. 105031.
- [2] Challenges for the cyber-physical manufacturing enterprises of the future / H. Panetto, B. Lung, D.

Ivanov [et al.] // *Annual Reviews in Control*. – 2019. – Vol. 47. – PP. 200-213.

[3] Mörth, O. Cyber-physical systems for performance monitoring in production intralogistics / O. Mörth, C. Emmanouilidis, M. Schadler // *Computers & Industrial Engineering*. – 2020. – Vol. 142. – P. 106333.

[4] Deep learning for visual understanding: A review / Y. Guo, Y. Liu, A. Oerlemans [et al.] // *Neurocomputing*. – 2016. – Vol. 187. – PP. 27-48.

[5] Wang, C-Y Scaled-YOLOv4: Scaling cross stage partial network / C-Y Wang, A. Bochkovskiy, H-Y M. Liao // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2021. – PP. 13024–13033.

[6] Redmon, J. YOLOv3: An Incremental Improvement / J. Redmon, A. Farhadi // *arXiv*. – 2018. – Vol. 1804. – P. 02767.

[7] You only look once: Unified, real-time object detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2016. – Vol. 7780460. – PP. 779-788.

[8] Simonyan, K. Very deep convolutional networks for large-scale image recognition / K. Simonyan, A. Zisserman // *3rd International Conference on Learning Representations*. – 2015. – Vol. 1409. – P. 1556.

[9] Ioffe, S.Y. Batch normalization: accelerating deep network training by reducing internal covariate shift / S.Y. Ioffe, C. Szegedy // *Proc. 32nd ICML*. – 2015. – PP. 448-456.

**DEVELOPMENT OF A CASCADE ALGORITHM FOR MONITORING THE MOVEMENT OF PARTS DURING THEIR MANUFACTURE**

**Kiseleva P.I.**  
**Pechenina E.Yu.**  
**Pechenin V.A.**

Samara National Research University

34, Moskovskoe shosse,  
Samara, 443086,  
Russian Federation

[v.a.pechenin@ssau.ru](mailto:v.a.pechenin@ssau.ru)

*A cascade algorithm has been developed that allows identification of contents in production containers. The algorithm consists of two stages: detection of container cells and classification of the contents of each cell. The proposed algorithm makes it possible to achieve a classification accuracy of 89% when trained on a relatively small sample size than would be required when using a direct part detection algorithm, without the cell detection stage. The algorithm is thus suitable for use in environmental monitoring systems in aerospace manufacturing.*

**Keywords:** blade; detection; algorithm; convolutional neural network; container

## References

- [1] van Mourik, S., van der Tol, R., Linker, R., Reyes-Lastiri, D., Kootstra, G., Koerkamp, P. G. and van Henten, E. J. (2021), "Introductory overview: Systems and control methods for operational management support in agricultural production systems", *Environmental Modelling & Software*, vol. 139, p. 105031.
- [2] Panetto, H., Jung, B., Ivanov, D., Weichhart, G. and Wang, X. (2019), "Challenges for the cyber-physical manufacturing enterprises of the future", *Annual Reviews in Control*, vol. 47, pp. 200-213.
- [3] Mörth, O., Emmanouilidis, C. and Schadler, M. (2020), "Cyber-physical systems for performance monitoring in production intralogistics", *Computers & Industrial Engineering*, vol. 142, pp. 106333.
- [4] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. and Lew, M.S. (2016), "Deep learning for visual understanding: A review", *Neurocomputing*, vol. 187, pp. 27-48.
- [5] Wang, C-Y, Bochkovskiy, A., Liao, H-Y M. (2021), "Scaled-YOLOv4: Scaling cross stage partial network", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 13024-13033.
- [6] Redmon, J. and Farhadi, A. (2018), "YOLOv3: An Incremental Improvement", arXiv, vol. 1804, p. 02767.
- [7] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016), "You only look once: Unified, real-time object detection", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 7780460, pp. 779-788.
- [8] Simonyan, K. and Zisserman, A. (2015), "Very deep convolutional networks for large-scale image recognition", *3rd International Conference on Learning Representations*, vol. 1409, p. 1556.
- [9] Ioffe S.Y., Szegedy C. (2015), "Batch normalization: accelerating deep network training by reducing internal covariate shift", *Proc. 32nd ICML*, pp. 448-456.