

# РАЗРАБОТКА СИСТЕМЫ ОБНАРУЖЕНИЯ ПОВРЕЖДЕНИЙ АВТОМОБИЛЯ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

## DEVELOPMENT OF A VEHICLE DAMAGE DETECTION SYSTEM USING CONVOLUTIONAL NEURAL NETWORKS

**A. Arysbek  
N. Saparkhodzhaev  
R. Abdrakhmanov**

*Summary.* Images and their processing are an important part of understanding the world. Therefore, incomplete images do not allow us to determine and analyze the state of the image. This robs us of a lot of opportunities, but modern machine learning platforms can recover lost or damaged parts of such images, making it easier to understand the behind-the-scenes context and effectively analyze the created images. Extrapolation in drawing can be done by splitting local structures into unknown parts to create a single pixel (or part) of the missing part, while maintaining harmony with neighboring pixels. This deep learning app is used to identify vehicle damage and obtain initial characteristics before the event for timely insurance payments.

Currently, the development of the automotive industry is directly related to the increase in the number of car accidents. Thus, insurance companies are faced with the spread of many complaints and claims at the same time. Using Mask R-CNN, owned by CNN neural networks based on machine learning and deep learning algorithms, can help solve such problems for insurance companies.

*Keywords:* CNN, Mask R-CNN, machine learning, Deep Learning, computer vision, neural network, Object detection, RoIAlign, RPN.

**Арысбек Алымхан Болатханович**

Международный казахско-турецкий университет  
имени Ходжи Ахмеда Ясави  
arysbek.alymkhan@mail.ru

**Сапарходжаев Нурбек Пажарбекович**

PhD, доцент, Международный казахско-турецкий  
университет имени Ходжи Ахмеда Ясави  
nurbek.saparkhojayev@ayu.edu.kz

**Абдрахманов Рустам Бахтиерович**

К.т.н., и.о. доцента, Международный казахско-  
турецкий университет имени Ходжи Ахмеда Ясави  
a.rustam.a@mail.ru

*Аннотация.* Изображения и их обработка являются важной частью понимания мира. Поэтому неполные изображения не позволяют нам определять и анализировать состояние изображения. Это лишает нас больших возможностей, но современные платформы машинного обучения могут восстанавливать потерянные или поврежденные части таких изображений, что облегчает понимание закулисного контекста и эффективный анализ созданных изображений. Экстраполяция в рисовании может быть выполнена путем разделения локальных структур на неизвестные части для создания одного пикселя (или части) недостающей части, сохраняя гармонию с соседними пикселями. Это приложение глубокого обучения используется для выявления повреждений автомобиля и получения первоначальных характеристик перед событием для своевременных страховых выплат.

В настоящее время развитие автомобильной промышленности напрямую связано с ростом числа автомобильных аварий. Таким образом, страховые компании сталкиваются с распространением множества жалоб и претензий одновременно. Использование Mask R-CNN, принадлежащего нейронным сетям CNN, основанным на алгоритмах машинного обучения и глубокого обучения, может помочь в решении таких проблем для страховых компаний.

*Ключевые слова:* CNN, Mask R-CNN, машинное обучение, Deep Learning, компьютерное зрение, нейронная сеть, обнаружение объекта, RoIAlign, RPN.

**О**бнаружение объектов — одно из основных направлений исследования компьютерного зрения. Он предназначен для определения степени и информации о местонахождении интересующего объекта на рисунке на уровне экземпляра. В настоящее время наиболее популярными алгоритмами обнаружения объектов являются R — CNN [1], Fast R-CNN [2], Faster

R-CNN [3] и SSD [4]. Однако эти структуры требуют больших объемов обучающих данных. Возможность определения окружности положения ограничена, и при извлечении данных часто происходит потеря градиента или взрыв градиента из-за увеличения числа сверточных слоев. Из-за этих недостатков Хе Каминг и т.д. предложил остаточную сеть (ResNet) [5] [8], которая

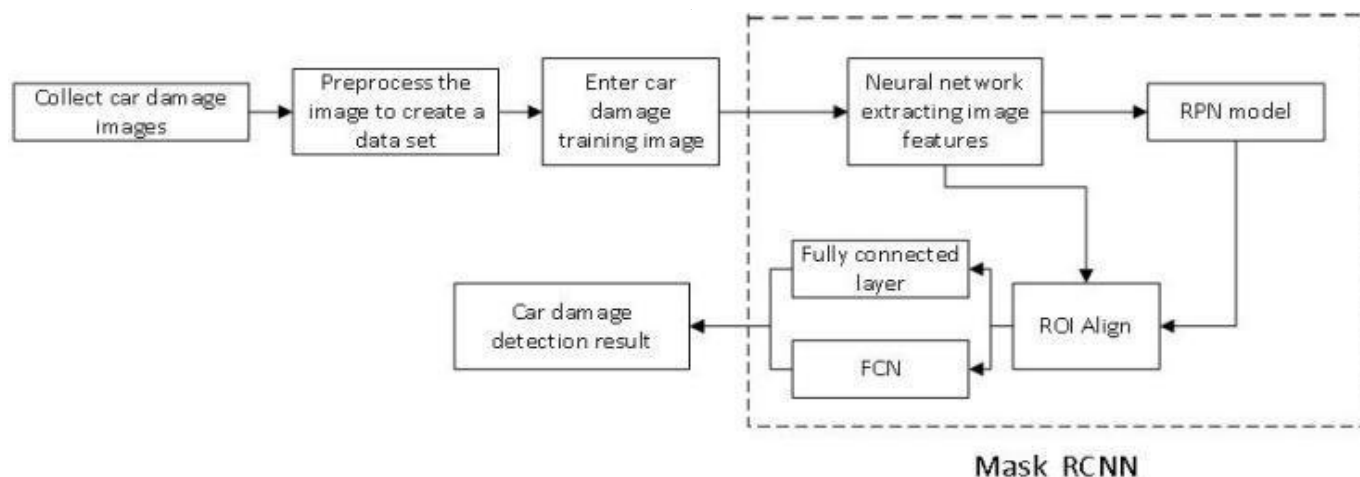


Рис. 1. Фреймворк системы обнаружения повреждений автомобилей

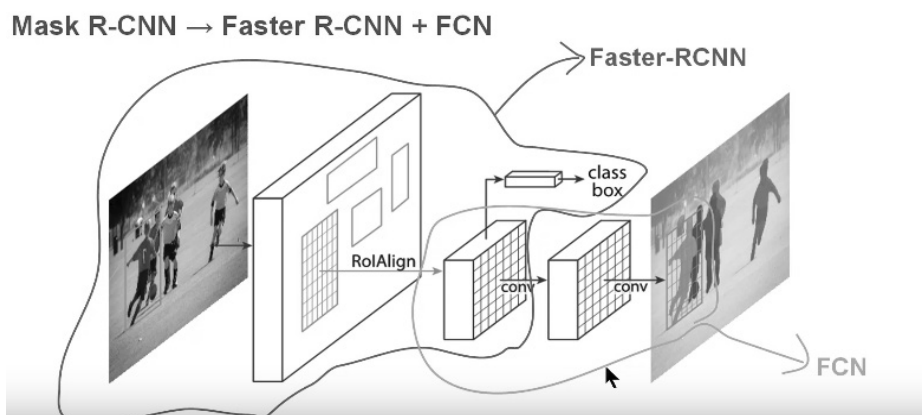


Рис. 2. Архитектура Mask R-CNN

ускоряет обучение нейронной сети с использованием остаточного модуля и интегрируется с моделью обнаружения объекта Mask R-CNN [6] [9] [10] для реализации обнаружения и сегментации объектов, тем самым значительно улучшая точность определения модели. Mask R-CNN-первая модель глубокого обучения, объединяющая определение объекта и сегментацию в единую сеть [7]. По сравнению с традиционным методом обнаружения объекта, модель обнаружения объекта Mask R-CNN не только значительно повышает точность обнаружения, но и обладает большими преимуществами при обнаружении небольших объектов.

В статье используется алгоритм Mask R — CNN для выявления и сегментации зон повреждения автомобилей в дорожно-транспортных происшествиях. Это имеет очень важное исследовательское значение и широко используется в области транспортировки. В этой статье структура линии модели будет улучшена путем кор-

ректировки внутренней структуры, повышения способности к обобщению и коррекции функции потерь для повышения точности повреждения транспортного средства, чтобы уменьшить количество слоев в остаточной сети и усилить регуляризацию модели. В результате проекта будет предложена улучшенная модель Mask R-CNN, применимая к области обнаружения поврежденных автомобилей и на ее основе предназначенная для обнаружения и разделения поврежденной части транспортного средства в результате аварии.

Система обнаружения и сегментации повреждений транспортных средств на основе разработанной в данной статье модели Mask R-CNN представлена на рисунке 1.

На рисунке видно, что изображение поврежденной части автомобиля подобрано и собрано в соответствии с требованиями. Данные передаются в Mask R-CNN для

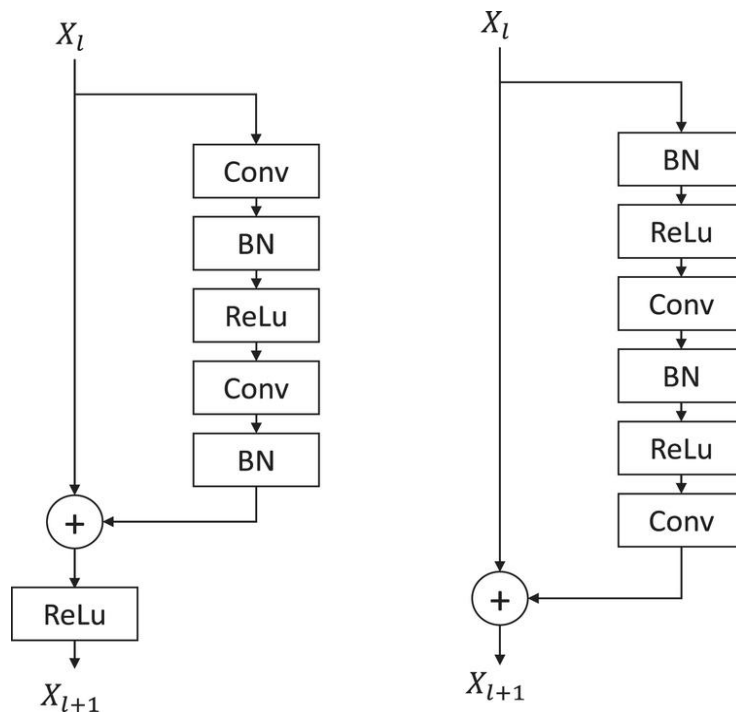


Рис. 3. Структура ResNet и улучшенная структура ResNetV2

получения признаков, прогнозирования классификации и маскировки сегментации, и отображается результат обнаружения повреждения автомобиля.

### Алгоритм Mask R-CNN

Mask R-CNN-это усовершенствованная структура сегментирования объектов с помощью Faster RCNN. Он делится на два этапа: первый этап сканирует изображение и делает предложение, а второй этап классифицирует предложение и создает ограничительную рамку и маску. Блок-схема сетевой структуры алгоритма Mask R-CNN представлена на рисунке 2.

Схема алгоритма Mask R-CNN выглядит следующим образом.

1. Введите редактируемое изображение в заранее подготовленную сетевую модель ResNet50 + FPN, извлеките метки и получите соответствующие карты символов.
2. Эта карта описания использует классификатор softmax для выполнения предыдущей и фоновой двоичной классификации с использованием регрессии кадров для получения более точного кандидата, который получает много кадров-кандидатов (т.е. область интересов или ROI) через RPN.
3. Карта функции и последняя область оставшегося интереса передаются на слой RoIAlign, таким

образом, каждая область интереса создает карту объектов заданного размера.

4. Наконец, поток проходит через две ветви: одна ветвь входит в полностью связанный слой для классификации объектов и регрессии кадра, а другая входит в полную сверточную сеть (FCN) для сегментации пикселей.

### Увеличение структуры магистральной сети

Как правило, сеть поддержки Mask R-CNN использует ResNet101; то есть количество уровней сети равно 101, но слишком много уровней значительно снижает скорость сетевой структуры. Степень повреждения автомобилей, изученных в этой статье, относительно проста, а требования к линейному уровню ниже. Таким образом, мы использовали ResNet50 для дальнейшего увеличения скорости алгоритма.

Поскольку размер транспортного повреждения на изображениях различен, только одна сверточная нейронная сеть не может хорошо воспроизводить все атрибуты изображения. Поэтому в этой статье была использована структура ствола ResNet50 и сеть пирамидальных функций FPN. FPN [21] использует иерархию сверху вниз с боковыми связями от одномасштабных входов до построения пирамиды сетевых функций, что решает задачу получения целевых объектов на изображении.

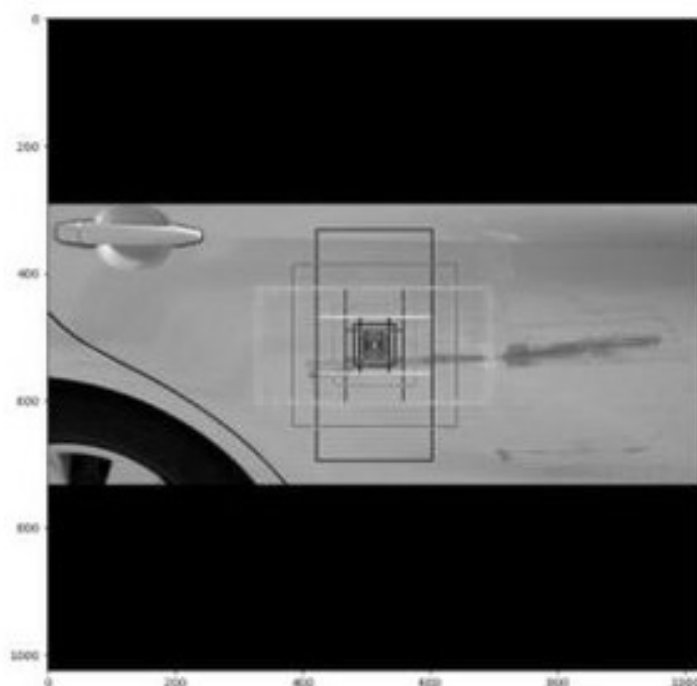


Рис. 4. Изображение, обработанное в модели RPN

жениях. Эта структура очень надежна и гибка и требует меньше настроек.

### Совершенствование модели RPN

Этот раздел принимает структуру сети пирамид знаков и дает изображения различных размеров для создания объектов, которые соответствуют различным критериям. Мелкие элементы могут различать простые большие цели, в то время как более глубокие элементы могут различать маленькие цели.

RPN эквивалентен бесклассовому детектору объектов на основе раздвижного окна. Он основан на структуре сверточной нейронной сети. Сканирование подвижной рамы производит якоря из якорной рамы. Предлагаемая область может создавать множество якорей различных размеров и соотношений, и они перекрываются, чтобы охватить как можно больше изображений; размер предлагаемой области и предпочтительная область перекрытия (IOU) напрямую влияют на эффект классификации. Чтобы приспособиться к поврежденным участкам автомобиля, алгоритм регулирует масштаб» точки привязки « $\{32 \times 32, 64 \times 64, 128 \times 128, 256 \times 256, 512 \times 512\}$ , а соотношение точек привязки изменяется, как показано на рисунке 4{1: 2,1: 1,3: 1}. Под IoU мы подразумеваем заранее предсказанную крышку коробки и фактическую коробку, значение которой равно пересечению двух коробок,

разделенных на объединение двух коробок. В этой статье значение IoU устанавливается на 0,8; то есть, когда соответствующая рамка якоря и фактическая целевая область больше 0,8, на переднем плане; когда коэффициент перекрытия меньше 0,2, это фон; между двумя значениями он удаляется. Усовершенствованный RPN дает меньший ROI, что, в свою очередь, повышает эффективность модели, уменьшает объем вычислений, экономит время.

### RoIAlign модели

В сетевой структуре Mask RCNN ветка маски должна определить, является ли данный пиксель частью цели, и точность должна быть на уровне пикселей. После того, как исходное изображение было сильно упаковано и объединено, размер изображения изменился. Когда сегментация на уровне пикселей выполняется непосредственно, невозможно точно расположить целевой объект изображения, поэтому Mask RCNN Faster улучшается на основе RCNN, а слой RoI Pooling заменяется слоем выравнивания интересующей области (RoIAlign). Метод билинейной интерполяции [23] сохраняет пространственную информацию на карте функций, что часто решает ошибку, вызванную двумя квантованиями карты функций на слое интеграции RoI, и решает проблему регионального несоответствия объекта изображения. Таким образом, можно добиться сегментации детектирования на уровне пикселей.

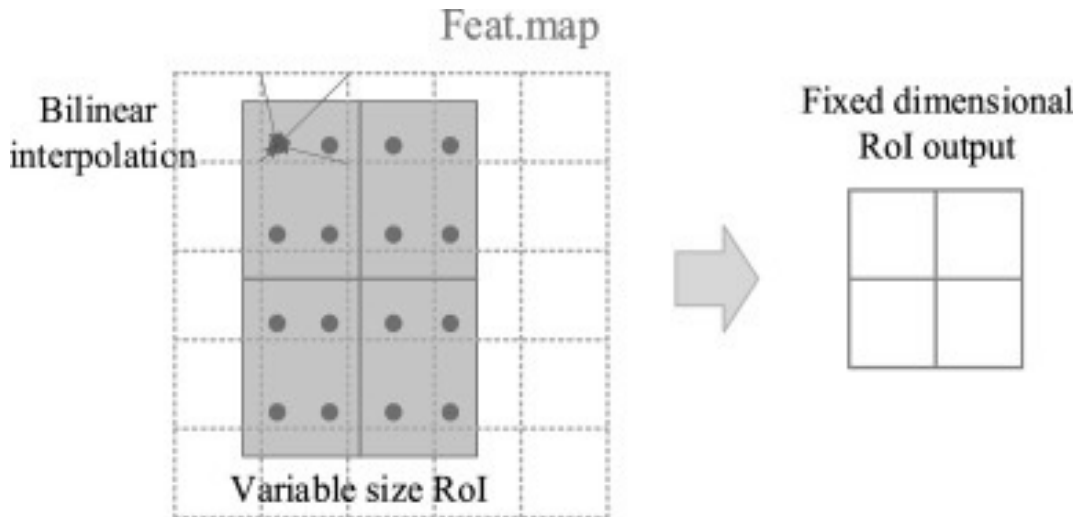


Рис. 5. Схема RoIAlign

Слой выравнивания интересующей области RoIAlign отличается от объединения ROI тем, что он удаляет операцию квантования и не пересчитывает область интереса и границы единицы, а использует нелинейную интерполяцию для вычисления фактического местоположения точек выборки в каждой единице с сохранением десятичного числа, а затем использует максимальное или среднее объединяющее действие для получения последнего фиксированного значения ROI. Как показано на рисунке 5, синяя пунктирная линия представляет собой карту знаков 5 × 5 после сборки, сплошная линия представляет собой небольшой блок знаков, соответствующий интересующей области на карте знаков, в то время как RoIAlign сохраняет границы числа с плавающей точкой без квантования. Сначала блок меньшего знака делится на блоки 2 × 2 (каждая граница блока не считается), а затем делится на четыре меньших блока в каждом блоке; центральная точка воспринимается как синяя точка на рисунке как четыре координаты. Затем значения четырех позиций вычисляются путем нелинейной интерполяции и, наконец, создается максимальная или средняя единица операций для получения карты характеристик 2 × 2.

Улучшение затрат

Многофункциональная функция потерь Mask RCNN:

$$L = L_{cls} + L_{box} + L_{Mask} \tag{1}$$

Приведенное выше уравнение аналогично функции потерь в модели Faster R-CNN, которая представляет соответственно ошибку классификации и ошибку определения. Ветка маски и ветка прогнозирования класса являются отдельными, тогда как ветка классификации

бинарной маски прогнозируется индивидуально для каждой категории, не полагаясь на результаты прогнозирования. Faster R-функция потерь CNN:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \tag{2}$$

В приведенной выше формуле  $i$  — индекс бокса якоря в небольшой партии;  $N_{cls}$  и  $N_{reg}$  указывают количество уровней классификации и уровней регрессии соответственно;  $p_i$  — прогнозируемое значение того, является ли  $i$  связывание объектом;  $p_i^*$  равен 0, если блок привязки отрицательный, а блок привязки положительный, равен 1;  $t_i$  указывает 4 параметризованные координаты окна кандидата на предсказание;  $t_i^*$  относится к 4 параметризованным координатам области действительного значения;  $L_{cls}$  и  $L_{reg}$  означают соответственно потерю классификации и потерю регрессии.  $\lambda$  — это коэффициент равновесия, используемый для управления соотношением двух функций потерь.

Faster RCNN вводит контрольный баланс гиперпараметра  $\lambda = 10$  между потерей классификации и потерей регрессии, в то время как крупномасштабная цель и меньшая цель разделяют этот единственный параметр.

В Mask RCNN функция ошибки пункта прогнозирования класса может быть вычислена по формуле:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda[u \geq 1]L_{loc}(t^u, v) \tag{3}$$



Car Damage Images

Рис. 6. Набор данных

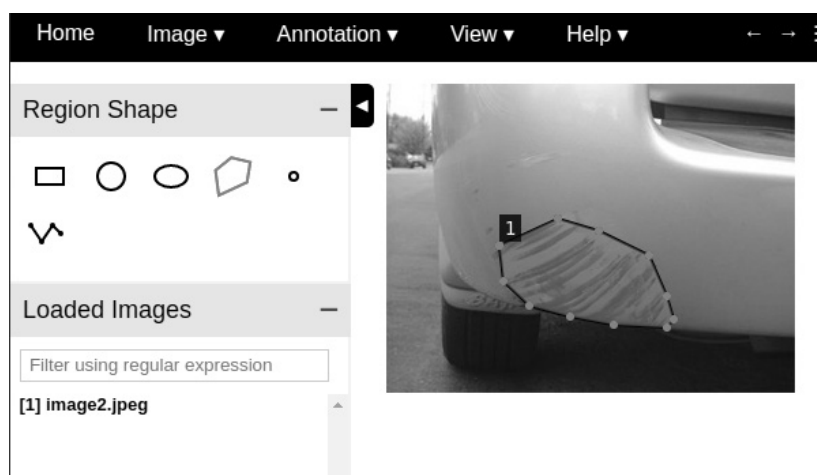


Рис. 7. VGG image Annotator — инструмент аннотатора V 1.0.6

где  $p$  — предполагаемый класс,  $u$  — класс GT,  $t^u$  — предполагаемая ограничительная рамка для класса  $u$ ,  $v$  — ограничительная рамка GT.

Если гиперпараметр  $\lambda = 10$  все еще вводится в Mask R-CNN, это вызывает явление. В основной элемент вводится семантическая информация высокого уровня. Маленькая цель имеет четкие точки подъема, тогда как более широкая цель не ясна. Дополнительная информация об основной функции вводится или поддерживается в функции более высокого уровня. Цель в больших масштабах имеет четкие точки подъема, в то время как

меньшая цель не ясна. Рамка большой цели на самом деле более точна, но отклонение позиции более важно, поэтому для позиционирования требуется хорошая информация, которая помогает улучшить большую цель на индикаторе карты. Возможное место для маленькой цели более точное. Тем не менее, оценка семантической информации относительно слаба, поэтому семантическая информация более высокого уровня необходима, чтобы помочь различию, которая может помочь улучшить меньшую цель в индексе карты. Таким образом, основное внимание уделяется оптимизации информации о местоположении для больших целей. Для

Таблица 1. Информация об экспериментальной среде

Имя атрибута	Значение атрибута
TensorFlow version	1.14.0
Keras version	2.2.5
Processor	Intell CoreI5-8300H CPU
Graphics	GeForceGTX 1080/Pcle/SSE2
Operating system version	Windows 10

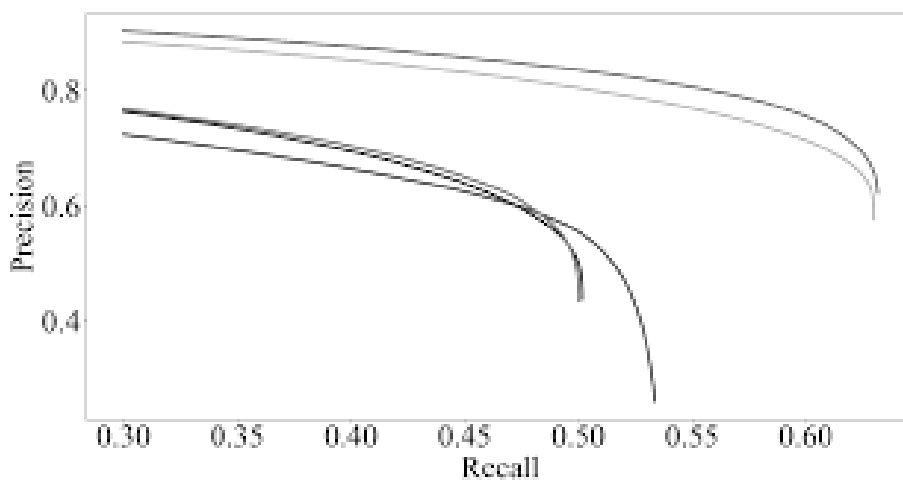


Рис. 8. Кривая P-R

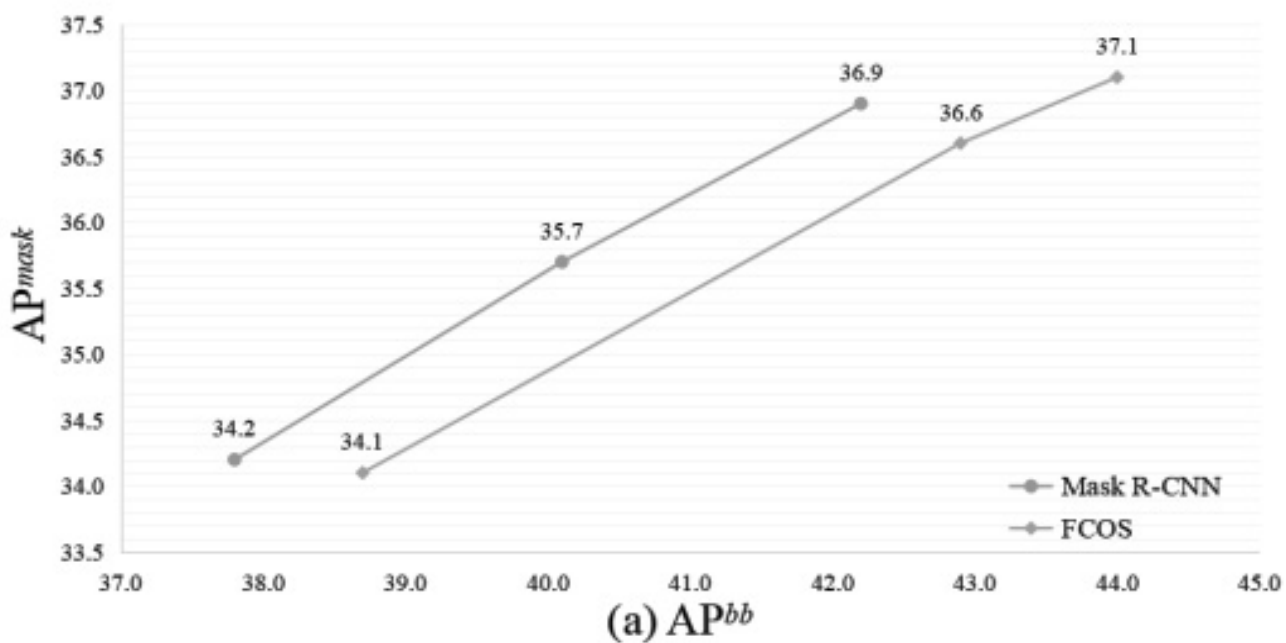


Рис. 9: значения AP двух алгоритмов

Таблица 2. Сравнение точности результатов испытаний

Алгоритм	Точность определения (%)	Точность Mask (MioU) (%)
Mask R-CNN	94.53	81.25
Улучшенный Mask R-CNN	96.68	83.14



Рис. 10. Результат обнаружения повреждений автомобиля на основе алгоритма Mask RCNN.



Рис. 11. Результат обнаружения повреждений автомобиля на основе улучшенного алгоритма Mask R-CNN.

небольших целей основное внимание уделяется оптимизации прогнозирования категорий. То есть для целей разного масштаба необходимо вводить в функцию потерь разные веса для повышения точности пунктов определения.

#### Сбор данных

Для этого проекта я собрал фотографию поврежденных автомобилей из Google.

#### Аннотирование данных

Модель Mask R-CNN требует аннотирования изображений и определения области повреждения. Поэтому я использовал VGG Image Annotator — инструмент аннотатора V 1.0.6.

В таблице 1 также приведены значения технологий, использованных в ходе проведения эксперимента.



## Результаты научно-исследовательской работы

Для изучения эффективности определения усовершенствованного алгоритма в наборе данных о повреждениях автомобиля я сравнил его с расширенным алгоритмом обнаружения Mask R-CNN. На рисунке 8 показана кривая P-R, полученная с помощью двух алгоритмов. Затем область под кривой P-R получается интегрированием и получается средняя точность двух алгоритмов определения повреждений автомобиля, т.е. значение AP и результат показаны на рис. 9.

На рисунках 9 и 10 видно, что улучшенный алгоритм Mask R-CNN значительно улучшил характеристики обнаружения по сравнению с алгоритмом Mask R-CNN. Как видно из рисунка 10, значение маски Mask R-CNN равно 0,75, а значение AP улучшенного алгоритма обнаружения равно 0,83, что на 0,08 выше значения Mask R-CNN расширенного алгоритма обнаружения цели.

Как видно из таблицы 2, улучшенный по сравнению с Mask R-CNN Mask R — CNN увеличил точность определения на 2,15%, точность маски-на 1,89%. Усовершенствованный алгоритм не только улучшает точность, но и ускоряет скорость обнаружения, обладает хорошими эксплуатационными преимуществами и позволяет увидеть, что на поврежденном участке автомобиля могут быть применены более высокие результаты.

Сравнивая рисунок 10, рисунок 11 и таблицу 2, улучшенный Mask R-CNN показал улучшение детализации и низкой точности в недостающих местах. Таким образом, усовершенствованный алгоритм показал высокую надежность и адаптивность к обнаружению повреждений транспортных средств. Но для того, чтобы усовершенствованный Mask R-CNN обладал более высокой производительностью, необходимо: увеличить

объем набора данных, собрать больше изображений поврежденных автомобилей в различных погодных условиях и на разных уровнях освещенности, улучшить данные, улучшить изображение контуров кромок и более точное отображение поврежденных участков автомобиля.

## Заключение

В работе, описанной в этой статье, я думаю, что алгоритм обнаружения Mask R-CNN, основанный на глубоком обучении обнаружению повреждений транспортных средств, станет незаменимой системой для решения проблемы страховых компаний в дорожно-транспортных происшествиях. После тестирования и усовершенствования, транспортировка является рекомендуемым методом обнаружения повреждений транспортных средств, основанным на методе обучения, и представляет собой усовершенствованную систему на основе Mask R-CNN. Модель показала хорошие результаты обнаружения в различных сценариях. Независимо от силы света, поврежденного участка нескольких автомобилей или слишком высоких воздействий, эффект соответствия лучше и более надежен. Несмотря на то, что в этой статье принят надежный алгоритм Mask R-CNN, который улучшает исходный алгоритм и дает отличные экспериментальные результаты, некоторые аспекты еще не изучены. Например, точность обнаружения очень высока, но сегментация части маски не может быть полностью правильной, и некоторые области, где повреждение не является очевидным, не могут быть сегментированы. В будущем он может быть предназначен для сбора большего количества изображений поврежденных автомобилей в различных погодных условиях и при различных уровнях освещения, улучшения данных, расширения данных для улучшения краев изображений и маскировки поврежденных участков автомобиля.

## ЛИТЕРАТУРА

1. R. Girshick, J. Donahue, T. Darrell, and J. Malik Rich feature hierarchies for accurate object detection and semantic segmentation // in Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2014. — № 1 pp. 580–587.
2. R. Girshick Fast R-CNN // in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2015. pp. 1440–1448.
3. S. Ren, K. He, R. Girshick, and J. Sun Faster R-CNN: Towards realtime object detection with region proposal networks // IEEE Trans. Pattern Anal. Mach. Intell, 2017. — № 6 pp. 1137–1149. doi: 10.1109/tpami.2016.2577031.
4. W. Liu, D. Anguelov, and D. Erhan SSD: Single shot multibox detector // in Proc. IEEE Eur. Conf. Comput. Vision, 2016. pp. 21–37.
5. K. He, X. Zhang, S. Ren, and J. Sun Deep residual learning for image Recognition // in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016. pp. 770–778.
6. K. He, G. Gkioxari, P. Dollar, and R. Girshick Mask RCNN // in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017. pp. 2980–2988.
7. N. Kumar and R. Verma A multi-organ nucleus segmentation challenge // IEEE Trans. Med. Imag, 2019. — № 1 pp. 34–39. doi: 10.1109/TMI.2019.2947628.
8. G. Han, J. Su, and C. Zhang A method based on multi-convolution layers joint and generative adversarial networks for vehicle detection // in Proc. KSII Trans. Internet Inf. Syst, 2019. pp. 1795–1811.
9. Y. Yu, K. Zhang, L. Yang, and D. Zhang Fruit detection for strawberry harvesting robot in non-structural environment based on mask-RCNN // Comput. Electron. Agricult, 2019. № 104846.

10. Y. Liu, P. Zhang, Q. Song, A. Li, P. Zhang, and Z. Gui Automatic segmentation of cervical nuclei based on deep learning and a conditional random field // IEEE Access, 2018. pp. 53709–53721.
11. I.F. Suvorov, B.B. Utegulov, A.B. Utegulov, A.B. Uakhitova Development of methods for determining the capacitive insulation conduction in three-phase networks with isolated midpoint neutral conductor with voltage up to 1000v // Вестник науки Казахского агротехнического университета им. С. Сейфуллина, 2015. — № 4 (87). С. 98–101
12. Мусалимов Т.К., Ахметов Е.С., Анисимова Л.Н. Использование современных инновационных технологий обучения в профессионально-графической подготовке бакалавра // Вестник науки Казахского агротехнического университета им. С. Сейфуллина (междисциплинарный), 2017. — № 4 (95) С. 101–108
13. S.V. Fedorova B.B. Utegulov, A.B. Utegulov, I.V. Koshkin Development of the arc-suppression coil control method for capacitive current compensation in the network of voltage 6–10 kv // Вестник науки Казахского агротехнического университета им. С. Сейфуллина (междисциплинарный), 2018. — № 4 (99) P. 144–154
14. Nukeshev S., Lichman G., Lichman A. Information technologies in precision agriculture // Вестник науки Казахского агротехнического университета им.С.Сейфуллина (междисциплинарный), 2015. — № 1 (84) С. 216–220
15. Кожажулов Е.Т. Фазово-управляемые колебания нейронных сетей: Сборник тезисов Международной конференции студентов и молодых ученых, Мир науки. — Алматы, 2012. С. 191.
16. B. Amirgaliyev, P. Kupagulova, C. Kenshimov Pedestrian detection algorithm for overlapping and non-overlapping conditions: в 12-th international conference on electronics computer and computation, ICECCO, Almaty, 2015.

© Арысбек Алымхан Болатханович ( arysbek.alymkhan@mail.ru ),

Сапарходжаев Нурбек Пажарбекович ( nurbek.saparkhojayev@ayu.edu.kz ), Абдрахманов Рустам Бахтиерович ( a.rustam.a@mail.ru ).

Журнал «Современная наука: актуальные проблемы теории и практики»



Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави