

УДК 622.271

**МЕТОД “ИНТЕЛЛЕКТУАЛЬНОГО” ОБНАРУЖЕНИЯ ТРЕЩИН
В БОРТАХ КАРЬЕРОВ
НА ОСНОВЕ УСОВЕРШЕНСТВОВАННОЙ АРХИТЕКТУРЫ MASK R-CNN**

Шуньлин Жуань^{1,2}, Даньян Лю¹, Цинхуа Гу^{1,2}, Ин Цзин¹

¹Сианьский университет архитектуры и технологии,
710055, г. Сиань, Китай

²Главная лаборатория перцептивных вычислений и принятия решений в сфере “умных” технологий,
710055, г. Сиань, Китай

Предложен алгоритм обнаружения трещин на основе усовершенствованной архитектуры Mask R-CNN для предотвращения катастрофических последствий, вызванных нарушением устойчивости бортов. Архитектура направлена на преодоление ограничения традиционного алгоритма обработки изображений и классической модели глубокого обучения применительно к обнаружению трещин в бортах карьеров. Продемонстрированы возможности Mask R-CNN, направленные на их целевое обнаружение, сегментацию и позиционирование. В усовершенствованной архитектуре устранены основные недостатки, связанные с низкой четкостью изображения краев трещин и ложным обнаружением. Показано, что в ней реализован программный “каркас” для обнаружения и сегментации изображений трещин в бортах карьеров. Данный метод представляет собой расширенный вариант сверточной нейронной сети. С его помощью можно выполнять задачи по классификации, сегментации и итерационной передискретизации в ветви маски, что позволяет решить проблему низкой четкости краев трещины. Результаты испытаний показали, что по сравнению традиционным алгоритмом разделения трещин предлагаемый метод обладает более высокой точностью обнаружения и более качественной сегментацией.

Борта карьеров, устойчивость бортов, обнаружение трещины, сегментация, Mask R-CNN, прикладное исследование

DOI: 10.15372/FTPRI20220317

Открытая разработка полезных ископаемых обладает рядом преимуществ по сравнению с подземной и занимает лидирующее положение в общемировой добыче полезных ископаемых, составляя ~86 % [1]. Глубина разработки постепенно увеличивается, борта карьеров становятся выше и круче, а объем добычи ископаемых в наиболее простых горно-геологических условиях уменьшается. В будущем это не только повысит затраты на добычу, но и приведет к возможному обрушению бортов, что является ключевой проблемой, ограничивающей без-

Работа выполнена при финансовой поддержке Национального фонда естественных наук Китая (грант № 51974223), Фонда естественных наук провинции Шаньси (грант № 2019JM-492), Фонда гуманитарных наук провинции Шаньси (грант № 2020R005) и Фонда гуманитарных наук Министерства образования Китая (грант № 18XJC630007).

опасность горных работ. С теоретической и практической точек зрения вопросы геомеханического мониторинга за бортами карьера, обнаружения в них трещин и эффективного управления устойчивостью пород в рамках обеспечения безопасности добычи и повышения экономических показателей крайне значимы.

Обрушение бортов карьера представляет собой движение массива пород вдоль поверхности трещины сдвига. В связи с постепенным увеличением глубины карьеров и угла наклона бортов вопросы их устойчивости становятся все более актуальными из-за угроз безопасности персоналу и инфраструктуре, оказания негативного влияния на окружающую среду и ресурсы, становясь одной из проблем предотвращения чрезвычайных происшествий [2]. Причиной обрушения бортов, как серьезной геологической катастрофы, является вовремя необнаруженный дефект. Под продолжительным воздействием природных и техногенных напряжений, действующих в массиве пород борта, образуются отдельности, щели и трещины, которые нарушают его целостность и приводят к обрушению [3]. Важно выбрать подходящие научно-обоснованные методы точного выявления и обнаружения трещин в бортах карьеров, что позволит получать достоверную и своевременную информацию для предотвращения обрушения массива пород. В связи с быстрым развитием компьютеризации, в частности обработки изображений, глубокого обучения и машинного наблюдения, технологии неразрушающего контроля, в основе которых лежит анализ иллюстраций, стали областью повышенного научного интереса для решения проблемы обнаружения трещин.

С целью точного и эффективного выявления трещин на изображениях проведены обширные исследования и получены определенные научные результаты. Лю Сяожуй и др. [4] предложили быстрый метод обнаружения трещин, сочетающий два разных метода пороговой сегментации для бинаризации изображения; Гао и др. [5] использовали метод быстрой коррекции для изображения с неравным распределением серого цвета, что повысило качество картины и позволило выполнить пороговую сегментацию — обнаружить трещину на исходном эскизе. Ли и др. [6] разработали целевую функцию для максимизации разницы между двумя категориями и предложили алгоритм пороговой сегментации трещины на изображении с помощью метода цветowych гистограмм, который позволяет точно выделять нарушения. Алгоритм обнаружения трещины такого типа, основанный на пороговой сегментации, является простым в вычислениях и несложным в применении, однако он чувствителен к шуму и подходит к определению целей с очевидными признаками. С помощью данного алгоритма трудно обеспечить устойчивое обнаружение трещин в бортах в сложных горно-геологических условиях массива пород.

Ли Цзиньхуэй [7] использовал 8-направленный оператор Собеля для точного обнаружения многих типов трещин, включая поперечные, продольные и нестандартные. Жан Цзюань и др. [8] выявили недостатки традиционных алгоритмов обнаружения трещин, основанные на выделении их краев по шкале серого цвета, и предложили метод фазовой группировки, который подходит для обнаружения незначительных по размеру нарушений со слабым контрастом. Ван Хуа и др. [9] предложили метод фрактальных размерностей для сегментации трещин и сравнили его с оператором Собеля для получения продолжительных и частых нарушений. Данный тип алгоритма, основанный на обнаружении краев трещины и их морфологических особенностей, способен эффективно определять границы нарушений. Амхаз Р. и др. [10] разработали алгоритм автоматического обнаружения трещин на основе двухмерных изображений, которые содержат как фотометрические, так и геометрические характеристики. Нгуен и др. [11] предложили новый метод обнаружения трещин, который в полной мере учитывает яркость и связность изображения. Для своей работы

данный метод не требует обучения текстурам трещин, подобно алгоритму FFA (анизотропия свободной формы). В области обнаружения трещин, помимо основных традиционных алгоритмов, существует множество других методов [12–18].

Указанные методы обладают хорошей эффективностью обнаружения типовых трещин с явными характеристиками, однако изображения их в бортах карьеров имеют более сложные черты, такие как текстура фона, различный шум, нестандартное распределение, крайне изометрическая морфология поверхности, сложность контактного режима и т. д. Эффективность традиционных алгоритмов выявления трещин в бортах карьеров невысока. В последние годы алгоритм обнаружения трещин на основе контролируемого обучения стал областью повышенного научного интереса [19].

Классическая сеть глубокого обучения широко применяется для опознавания изображения, голоса и личности. Сверточные нейронные сети (CNN) обладают заметными преимуществами при обнаружении трещин. Анализ научной литературы показал, что специализированная модель обнаружения трещин в бортах карьеров для сложных условий отсутствует. Методы обнаружения трещин для типовых сооружений (мосты, дороги, бетонные здания) включают в себя алгоритмы обработки изображений, выявления типового рисунка, машинного, а также глубокого обучения. Результаты сравнительного анализа этих алгоритмов (табл. 1) показали, что метод обработки изображений является крайне простым и обладает наименьшей вычислительной сложностью и затратами памяти, тогда как глубокое обучение требует наибольших объемов хранилища информации и эксплуатационных издержек, что сопряжено с трудностями в расчетах.

ТАБЛИЦА 1. Сравнение преимуществ и недостатков основных методов обнаружения трещин

Показатель	Методы обработки изображения			Методы опознавания типового изображения и машинного обучения			Методы глубокого обучения	
	Пороговая сегментация	Морфологический анализ	Метод периферийных вычислений	Метод опорных векторов (SVM)	Многослойный перцептрон в ИНС	Метод Adaboost	Классификация	Сегментация
Затраты	Низкие	Низкие	Низкие	Низкие	Средние	Средние	Высокие	Высокие
Требования к памяти	Низкие	Низкие	Низкие	Низкие	Средние	Средние	Высокие	Высокие
Точность	Низкая	Низкая	Низкая	Высокая	Средняя	Средняя	Средняя	Высокая
Надежность	Низкая	Низкая	Низкая	Средняя	Средняя	Средняя	Высокая	Высокая
Способность к обобщению	Низкая	Средняя	Низкая	Высокая	Высокая	Высокая	Высокая	Высокая

Несмотря на вычислительную эффективность и низкие затраты, метод обработки изображения обеспечивает неточное обнаружение и выявление трещин со сложной фоновой текстурой, различным шумом, нестандартным их распределением, крайне изометрической морфологией поверхности, сложностью контактного режима. Метод глубокого обучения обладает наибольшей точностью опознавания, надежностью алгоритма и способностью к обобщению. Метод опознавания типового изображения или машинного обучения имеет низкие вычислительные затраты, высокую вычислительную эффективность, точность обнаружения и надеж-

ность, равную алгоритмам глубокого обучения. Такое решение не подходит для изображений трещины в борту карьера, так как оно обычно не содержит геометрической и структурной информации для отделения покрытия трещины от фонового слоя. В рамках глубокого обучения сверточные нейронные сети (CNN) обладают высокоэффективной способностью выявления признаков и целевого опознавания (рис. 1*з, д*).

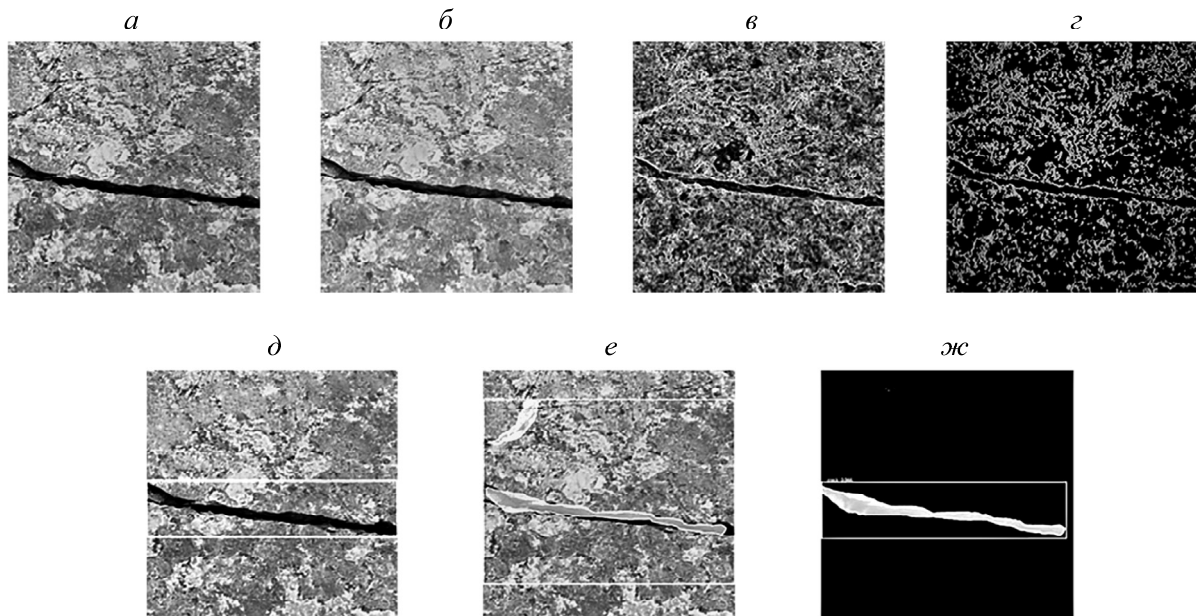


Рис. 1. Идентификация трещин: исходное изображение в борту карьера (*а*); изображение, полученное по шкале серого цвета (*б*); с помощью операторов Собеля (*в*) и Кэнни (*г*); целевого машинного обучения (*д*); обнаружение и сегментация трещины с помощью глубокого обучения (*е, ж*)

Сравнение различных методов обнаружения трещин в борту карьера (рис. 1) показало, что программная архитектура глубокого обучения Mask R-CNN, в которой интегрировано целевое опознавание и сегментация, может быть выбрана в качестве основы построения модели определения нарушений в бортах карьеров, поскольку обладает высокой точностью и эффективностью их нахождения. Тем не менее при использовании модели Mask R-CNN с ResNet50 в качестве стержневой архитектуры для определения трещины в борту карьера прямым способом возникают такие проблемы, как ложное обнаружение и низкая четкость краев. В настоящей работе предложен “интеллектуальный” алгоритм обнаружения трещин в борту карьера, основанный на усовершенствованной архитектуре Mask R-CNN, которая предназначена для обнаружения, сегментации и определения их местоположения.

Первоначально алгоритм выполняет предварительную обработку картин с помощью метода скользящего окна для повышения количества собранных изображений. Затем используется классификационная модель изображений на основе SVM для классификации и распознавания элементов трещины и фона. На заключительном этапе предполагается, что модель с помощью усовершенствованной архитектуры Mask R-CNN реализует обнаружение и сегментацию трещин, где целями являются их элементы в борту карьера. Данная модель способна выполнять три задачи: классификацию, целевое опознавание и сегментацию трещины на изображении в условиях ограниченности данных (рис. 2). Большое количество испытаний подтвердило, что по сравнению с традиционными алгоритмами обнаружения трещин предлагаемый алгоритм обладает лучшей эффективностью опознавания и обобщающей способностью благодаря более точной сегментации их краев.

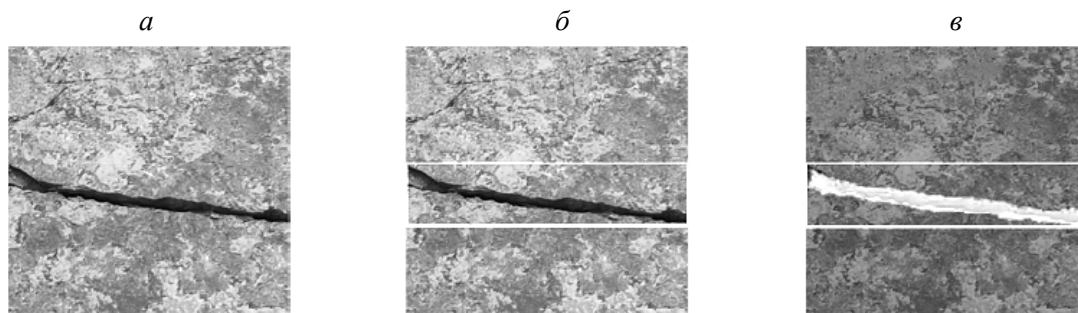


Рис. 2. Трещина в борту карьера: исходное изображение (а); целевое обнаружение (б); сегментация (в)

АРХИТЕКТУРА MASK R-CNN

Mask R-CNN представляет собой гибкую и универсальную архитектуру быстрой сегментации объектов [20]. Она способна эффективно обнаруживать объекты на изображениях и генерировать сегментации высокого качества для каждого из них. Схема работы архитектуры Mask R-CNN показана на рис. 3.

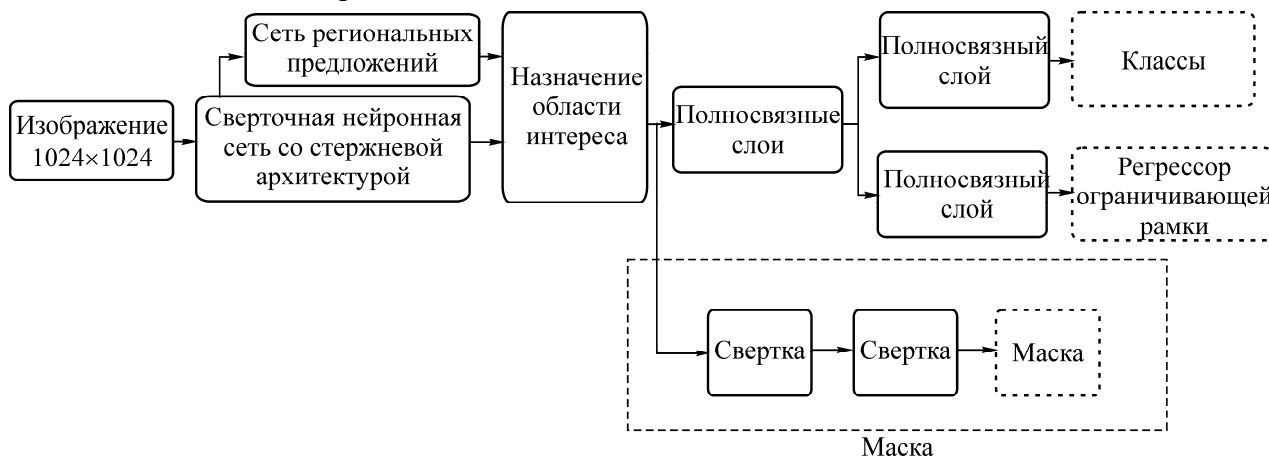


Рис. 3. Сетевая диаграмма архитектуры Mask R-CNN

С концептуальной точки зрения построение Mask R-CNN является крайне простым. На основе класса и выходного значения ограничивающей рамки Faster R-CNN для каждого объекта Mask R-CNN добавляет третью выходную ветвь — маску объекта [21]. Это архитектура нейронной сети глубокого обучения, в которую интегрировано целевое обнаружение и сегментация. На рис. 4 представлена сетевая архитектура Mask R-CNN, которую можно разделить на две части: (I) сверточная стержневая структура для извлечения признаков на всем изображении; (II) головная сеть, используемая для выявления ограничивающей рамки (классификация или регрессия) и ветвь прогнозирования маски. Многие китайские и зарубежные исследователи спроектировали CNN-архитектуры с отличной эффективностью, например Alexnet, VGG, ResNet, DenseNet и т. д. [22–25]. Эффективность этих CNN-архитектур отличается в зависимости от целей обнаружения: стержневая архитектура определяет эффективность последующих построений при решении задач по целевому обнаружению и быстрой сегментации. В качестве стержневой сетевой архитектуры Mask R-CNN обычно используется ResNet. Однако отсутствуют испытания, определяющие наиболее подходящую сверточную стержневую архитектуру для обнаружения трещин в бортах карьера.

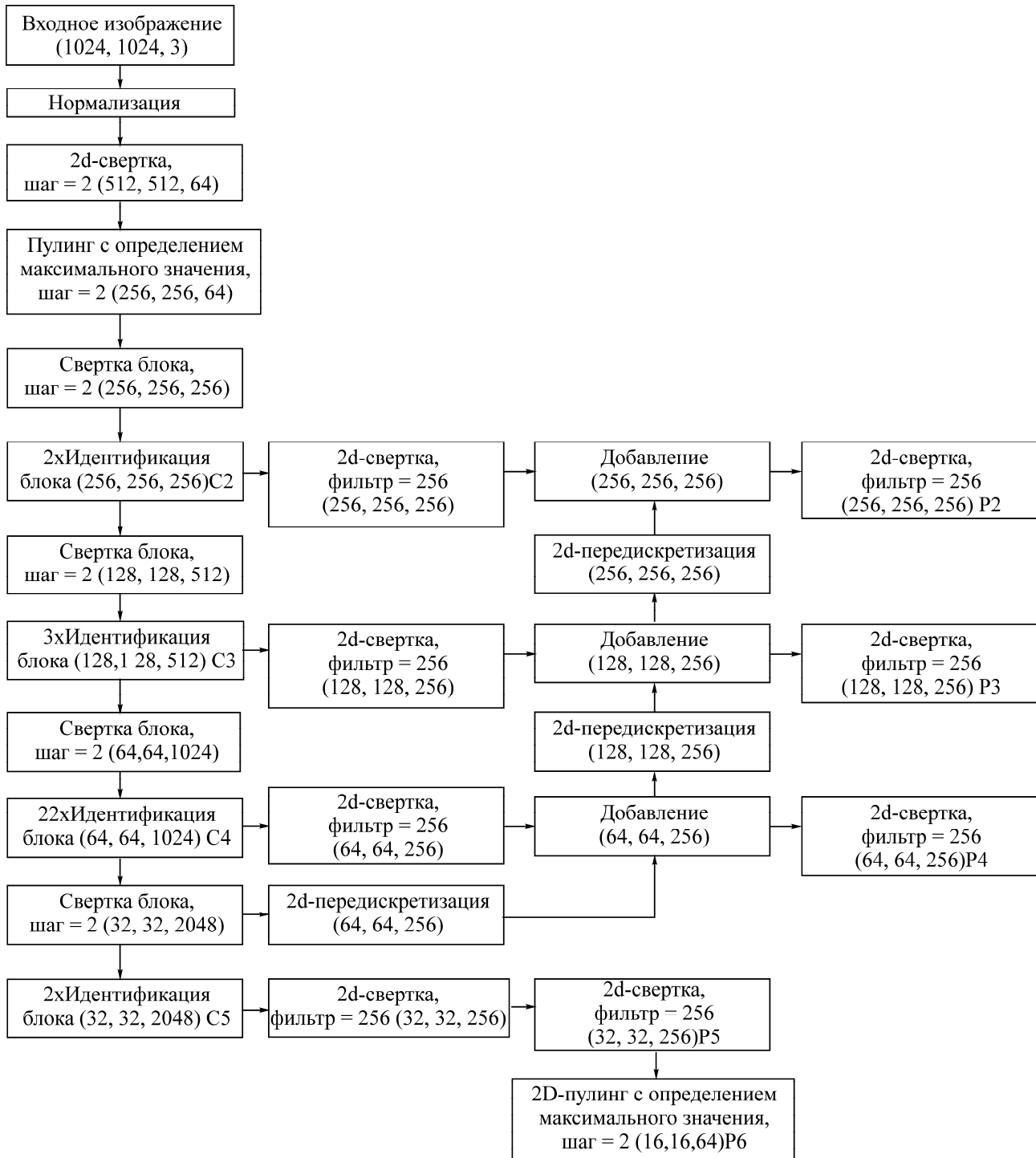


Рис. 4. Сетевая архитектура Mask R-CNN

Mask R-CNN — двухстадийная сетевая модель быстрой сегментации. На первой стадии применяются сети региональных предложений (RPN) для выполнения скользящего сканирования изображения, определяющего области интереса (ROI), где находится цель. На второй стадии проводится параллельно прогнозирование классификации и местоположения каждой ROI, найденной RPN. В качестве выходного параметра выступает соответствующая маска. В процессе обучения модели признаки, извлеченные стержневой сетью, используются в качестве

эффективных слоев признаков в ветви RPN. Сеть RPN применяет классификатор для предварительной проверки содержания целей в ROI, а затем регрессию для корректировки положения приоритета и в конечном итоге генерирует соответствующие предложения для обнаруженных категорий и областей. Mask R-CNN определяет функцию многозадачных потерь для каждой ROI в виде образца:

$$L = L_{cls} + L_{box} + L_{mask},$$

где L_{cls} , L_{box} — потери сети классов и предложений соответственно; L_{mask} — средние бинарные кросс-энтропийные потери ветви маски.

В Mask R-CNN внедрена пирамидальная сеть признаков (FPN). Слои признаков, сжатые в 2, 3, 4 и 5 раз в стержневой сети, используются для построения FPN, которая реализует мультимасштабное сочетание показателей. Это позволяет сети не только эффективно задействовать нижний слой для точного определения местоположения признаков с высоким разрешением, но и показатели с низким разрешением верхнего слоя. Вертикальная структура FPN с горизонтальными связями нужна для построения пирамиды признаков в сети с одномасштабным входом (рис. 5). Mask R-CNN реализует стержневую сеть ResNet-FPN для извлечения признаков, что повышает точность и скорость обнаружения трещины [20].

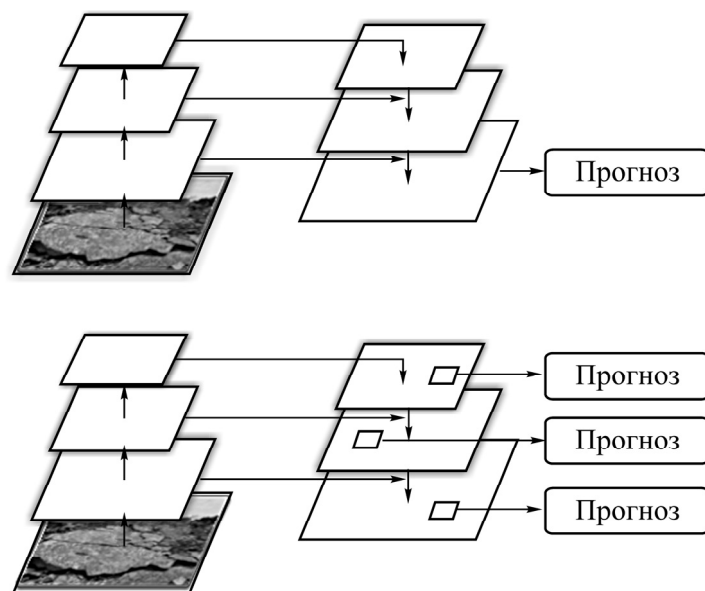


Рис. 5. Пирамидальная сетевая структура признаков FPN

МОДЕЛЬ ОБНАРУЖЕНИЯ ТРЕЩИН В БОРТУ КАРЬЕРА

Модель оптимизации сети. Модель “интеллектуального” обнаружения трещин в бортах карьера, основанная на усовершенствованной архитектуре Mask R-CNN, состоит из трех частей: (1) выполнение операций предварительной обработки — увеличение четкости изображений массива с помощью метода скользящего окна и использование SVM-модели данных картин для предварительной классификации набора показателей; (2) построение мультимасштабной сети извлечения признаков трещин на основе ResNet; (3) разработка модели сегментации трещины на основе Пирамиды пространственного пулинга “Atrous” (ASPP) для решения проблемы ложного обнаружения и низкой четкости ее краев при применении Mask R-CNN для идентификации нарушения в борту карьера.

Процесс обнаружения трещин в бортах карьера с помощью предлагаемой модели включает следующие шаги.

- Первоначально выполняется предварительная обработка исходного изображения и нормализация пикселей с дальнейшим увеличением количества картин с помощью метода скользящего окна.
- Второй шаг — использование SVM-модели классификации изображений для разделения увеличенного набора данных на элементы фона и компоненты, содержащие трещины, с последующей генерацией комплектов обучающих, контрольных и проверочных величин для модели “интеллектуального” обнаружения нарушений в бортах карьера.
- Третий шаг — обучение модели “интеллектуального” обнаружения трещины в бортах карьера, в основе которой лежит усовершенствованная архитектура Mask R-CNN, для получения оптимальных весовых значений каждого слоя.
- Последний этап — выбор оптимальных весовых значений для построения модели “интеллектуального” обнаружения трещины в бортах карьера с высокой точностью.

Мультимасштабное получение признаков трещин в борту карьера. Для решения проблемы потери локальной и детальной информации изображения в процессе субдискретизации адаптирован метод внедрения ASPP в ветвь маски.

В стержневой сверточной нейронной сети модели обнаружения, из-за продолжительного сочетания между признаками нижнего и верхнего уровней в пирамидальной ветви, информация о показателях нижнего слоя теряется в процессе прямого распространения при их получении. Это влияет на последующую классификацию, регрессию ограничивающей рамки и ее генерацию, приводя к ложным обнаружениям трещин и низкой точности сегментации. Для решения данной проблемы информация о признаках на нижних уровнях пирамиды полностью сохраняется при их получении; модуль ASPP внедрен в архитектуру Mask R-CNN для увеличения рецептивного поля информации о признаках [27]. Используется дилатация свертки со скоростями образцов 1, 6, 12 и 18 для параллельной семантической свертки на верхнем уровне признаков, что позволяет увеличить рецептивное поле центра показателей эффективного слоя без потери разрешения. Рецептивное поле — область пикселей на карте признаков каждого эффективного слоя входного изображения, которые рассчитывается следующим образом:

$$\begin{cases} j_l = j_{l-1} s_l; \\ j_l = \prod_{i=1}^l s_i; \\ r_l = r_{l-1} + (k_l - 1) j_{l-1}; \\ \text{start}_l = \text{start}_{l-1} + \left(\frac{k_l - 1}{2} - p_l \right) j_{l-1}, \end{cases}$$

где j_l — пиксельное расстояние между соседними элементами l -й карты слоя признаков; s_l — размер шага l -го слоя; k_l — размер сверточного ядра на l -м слое; r_l — боковая длина рецептивного поля l -го слоя; start_l — координаты центра рецептивного поля карты признаков l -го слоя.

После того как стержневая сверточная сеть выполнит субдискретизацию исходного изображения 4 раза, разрешение карты признаков входящего слоя дилатации свертки станет $32 \times 32 \times 2048$. После трех дилатаций свертки с разрешением сверточного ядра 3×3 последовательные разрешения рецептивных полей центрального признака на карте становятся 13×13 , 25×25 , 37×37 . Пирамида признаков, образованная параллельной дилатацией свертки, может сочетаться с мультимасштабными признаками трещин в бортах, что заставляет последующий слой удерживать больший объем информации о трещине [28].

Сегментация трещины. Полученная маска с разрешением 28×28 прямым способом интерполируется к исходному пиксельному размеру. Данная операция приводит к низкой точности пикселей краев маски. Для решения проблемы низкой четкости краев в Mask R-CNN подобные точки рассчитываются отдельно с помощью метода PointRend [29]. Затем к пикселям в других частях применяется метод прямой интерполяции. Таким образом в настоящей работе выполнено усовершенствование прогнозирования ветви маски в архитектуре Mask R-CNN (рис. 6).

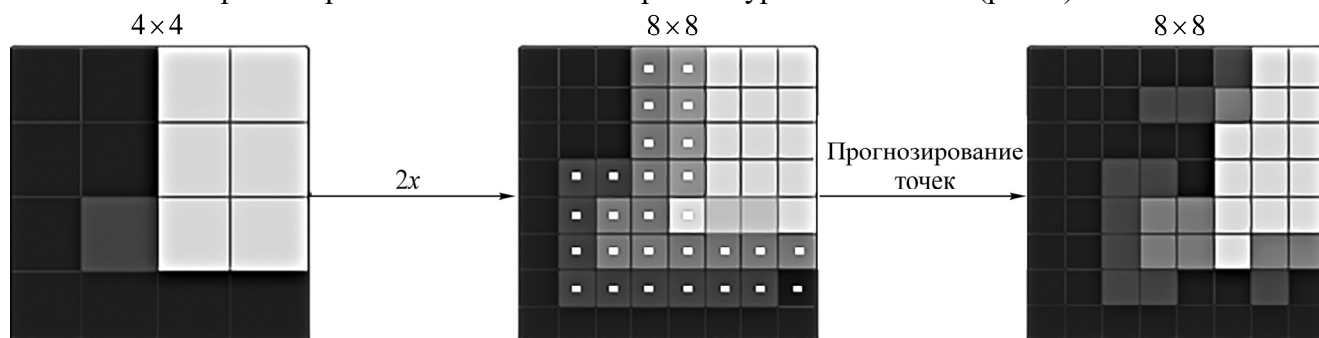


Рис. 6. Модуль PointRend

На стадиях обучения и прогнозирования предлагаемый алгоритм устанавливает гибкую и адаптивную стратегию выбора точки, для которой перестраивает вектор признака, а затем прогнозирует ее класс и повторно сегментирует адресную метку маски. Оптимизирован метод дискретизации сегментации и выполнена итерационная передискретизация классифицированной сегментации. Сегментация маски проводится пошагово до соответствия требованиям пиксельной точности, что позволяет избежать проблему избыточных вычислений при расчете пикселей по одному.

На каждой итерации фазы прогнозирования билинейная интерполяция используется для передискретизации ранее спрогнозированных результатов сегментации трещины. Затем происходит определение точек края трещины на более плотной сети (рис. 6). На фазе обучения применяется неитерационная стратегия, в основе которой лежит случайная дискретизация, для очередной установки неопределенных пиксельных точек маски трещины. Наиболее неопределенные точки на верхнем уровне βN выбираются в качестве первых для дискретизации. Такая гибкая стратегия выбора точек позволяет избежать избыточных вычислений выходных изображений с высоким разрешением, что приводит к снижению вычислительной мощности. После извлечения показателей из выбранных точек выполняется сочетание мелких признаков с грубо спрогнозированной маской для их перестроения одна за другой. Это обеспечивает более детализированную сегментацию и полную информацию о фоне для точек дискретизации. В итоге в качестве прогнозирующей головной сети используется легковесный многослойный перцептрон для получения новой маски, пока ее точность не достигнет требуемой. Каждый раз генерируется новая маска трещины в борту карьера с более высоким разрешением.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Условия проведения испытаний и подготовка данных. Алгоритм “интеллектуального” обнаружения трещин в бортах карьера, в основе которого лежит усовершенствованная архитектура Mask R-CNN, использует сверточную нейронную сеть. Для обучающего, контрольного и прогнозирующего набора данных модели необходимо большое количество изображений трещин в бортах с пояснениями и классификационными метками. Однако

в настоящее время в Китае и за рубежом отсутствует в публичном доступе массив изображений трещин в бортах карьеров с пояснениями и классовыми отметками для сетей глубокого обучения.

В рамках настоящей работы с помощью квадрокоптеров выполнено натурное наблюдение и сбор изображений бортов карьера большого размера. Собрано 500 изображений трещин. Данное количество не может полностью удовлетворить требования для получения знаний, проверки и прогнозирования модели глубокого обучения. Использование натурального наблюдения является трудной задачей. На основе 500 изображений реализован метод скользящего окна для расширения массива изображений трещин в бортах карьера (рис. 7). Метод позволяет расширить массив данных более чем в 4 раза.

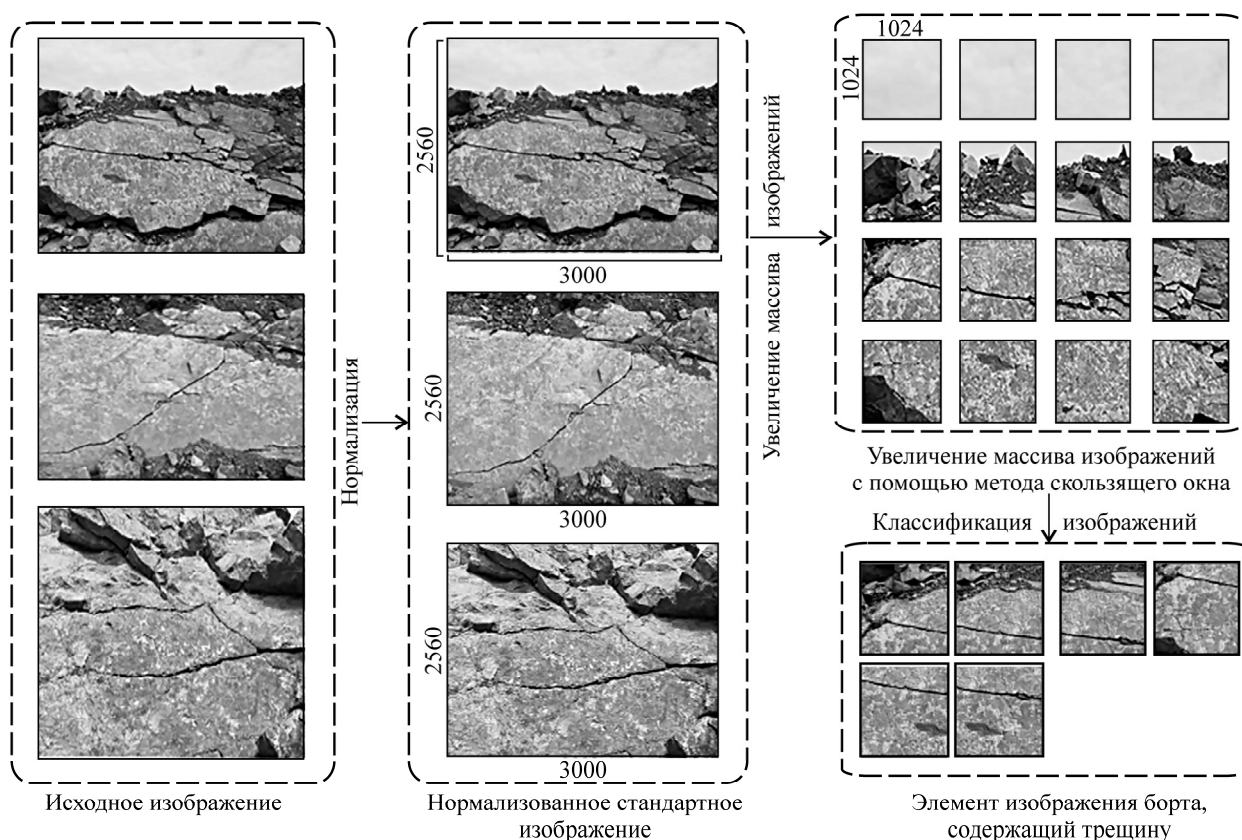


Рис. 7. Процесс увеличения массива изображений с помощью алгоритма скользящего окна

Первоначально проводилась нормализация 500 собранных изображений трещин и унификация до картин с размером пикселей 3000×2560 и разрешением 75×75 . Затем использовалось окно с фиксированным размером 1024×1024 для выполнения накладывающегося скольжения с шагом 512 на нормализованных изображениях. Область картины, покрытая окном после каждого скольжения, принималась как элемент изображения борта. Часть изображения борта, содержащего трещину, считалась элементом с трещиной, а без трещин — фоновым элементом.

На следующем этапе применялся метод опорных векторов (SVM) для классификации расширенного массива изображений и отделения элементов с трещинами [30]. Картины после просмотра подверглись вертикальному и горизонтальному симметричному отражению и случайному вращению для дальнейшего расширения массива изображений. Окончательный массив включал в себя 2052 изображения с трещинами. Далее его разделяли на две группы: для обучения и проверки использовалось 1900 изображений, остальные 152 — для прогнозирования эффективности обучения модели.

Подготовка, проверка и прогнозирование модели глубокого обучения требует большого объема данных с метками. Крайне важно точно отмечать данные. В настоящей работе искусственно расширенный массив изображений трещин получен точным способом. Ввиду трудоемкости процесса форма и структура трещин на изображениях не является типовой. Принято, что на изображении трещины обозначены белым цветом, а остальная часть фона — черным. Пример изображения представлен на рис. 8.

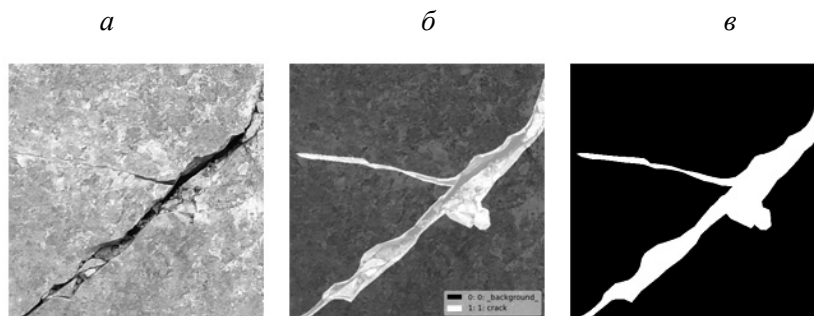


Рис. 8. Исходное изображение трещины (а); метка (б); маска (в)

Программа разработана на языке программирования Python на базе широко распространенных архитектур глубокого обучения TensorFlow Keras и открытой библиотеки алгоритмов компьютерного зрения OpenCV. Для работы алгоритма использован компьютер с процессором Intel(R)Core(TM)i7-9700 с частотой 3.00 ГГц и видеокартой NVIDIA GeForce GTX 1660Super с объемом видеопамати 6 Гб.

Анализ качества обнаружения. Независимо от пригодности используемого массива изображений применение стержневой архитектуры ограничивает качество обнаружения трещин в бортах. В ходе испытаний сначала выявлялась наиболее подходящая стержневая архитектура для идентификации трещин, затем выполнялась проверка влияния увеличения данных на качество обнаружения.

Первый набор испытаний проводился для проверки влияния различных стержневых архитектур на качество обнаружения трещин в бортах. Набор состоял из трех испытаний. В испытании 1 для построения модели использовалась архитектура ResNet50-RPN-FPN. После 50 генераций обучающего набора данных выбирались окончательные аппроксимированные весовые значения для создания прогнозирующего комплекта показателей. В испытании 2 для построения модели применялась архитектура ResNet101-RPN-FPN, а также установочные параметры, обучающие и прогнозирующие изображения, как в испытании 1. В испытании 3 для построения модели использовалась архитектура ResNet152-RPN-FPN. Установочные параметры были аналогичны испытаниям 1 и 2. Изменения значений потерь при обучении и проверке в ходе данных испытаний представлены на рис. 9.



Рис. 9. Диаграммы потерь при следующих стержневых архитектурах: ResNet50 (а); ResNet101 (б); ResNet152 (в): 1 — потери при проверке; 2 — потери при обучении

Из рис. 9а видно, что при использовании ResNet50 в качестве стержневой архитектуры разрабатываемой модели значение потерь исчезает к 17-й генерации обучающих данных. В этом случае модель не может быть полностью обучена и обладать достаточным качеством обнаружения трещин. При использовании ResNet101 и ResNet152 в качестве стержневых архитектур модель постепенно стала конвергентной к 40-й генерации обучающих данных. Изменения значений потерь при обучении и проверке более гладкие, а обучающий эффект выше. Далее выполнялись сравнительные испытания для проверки влияния увеличения данных на основе метода скользящего окна на точность обнаружения трещин моделью. Результаты представлены в табл. 2.

ТАБЛИЦА 2. Влияние увеличения массива изображений на точность модели обнаружения трещин в бортах карьера

Модель	Потери при обучении	Потери при проверке	Количество изображений	РА	МРА	MIoU	FWIoU
Архитектура ResNet101-Mask R-CNN	0.235	0.450	500	0.8576	0.6878	0.6262	0.7807
Архитектура ResNet152-Mask R-CNN	0.369	0.869	500	0.8358	0.6471	0.6098	0.7785
Архитектура ResNet101-Mask R-CNN (увеличенный массив изображений)	0.130	0.424	2000	0.9206	0.7179	0.6753	0.8871
Архитектура ResNet152-Mask R-CNN (увеличенный массив изображений)	0.202	0.458	2000	0.8967	0.6891	0.6383	0.8847

В процессе сегментации изображения обычно применяется множество стандартизованных величин для измерения точности алгоритма. В настоящей работе выбраны четыре оценочных индекса: РА (пиксельная точность), МРА (средняя пиксельная точность), MIoU (среднее пересечение над объединением), FWIoU (взвешенная частота пересечения над объединением). РА — наиболее простая величина, представляющая соотношение точно обнаруженных пикселей к общему их количеству; МРА — усовершенствованная версия РА, которая рассчитывает долю точно обнаруженных пикселей в каждом классе, а затем определяет среднее значение по всем классам; MIoU — стандартная величина для семантической сегментации, которая рассчитывает соотношение пересечений и объединений эталонных и спрогнозированных наборов; FWIoU — усовершенствованная версия MIoU, устанавливающая весовое значение для каждого класса исходя из частоты появления. Ниже представлены формулы для определения данных величин:

$$\begin{aligned}
 \text{РА} &= \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}; \\
 \text{МРА} &= \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}}; \\
 \text{MIoU} &= \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}; \\
 \text{FWIoU} &= \frac{1}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}},
 \end{aligned}$$

где k — количество классов, которые не содержат фон; p_{ii} — количество верно классифицированных пикселей; p_{ij} — количество пикселей, изначально принадлежащих классу i , но спрогнозированных к классу j ; p_{ji} — количество пикселей, изначально принадлежащих классу j , но спрогнозированных к классу i .

Согласно табл. 2, независимо от выбранной стержневой архитектуры для Mask R-CNN (ResNet101 или ResNet152) увеличение объема показателей с помощью метода скользящего окна положительно влияет на точность обнаружения. При сравнении PA, MPA, MIOU и FWIoU установлено, что значение индекса оценки архитектуры ResNet101 для модели обнаружения трещин является достаточно высоким. Увеличение массива изображений позволяет значительно повысить точность обнаружения при использовании архитектуры ResNet101.

Анализ эффективности обнаружения трещин. Результаты испытаний качества обнаружения трещин показали, что архитектуры ResNet101 и ResNet152 имеют схожие идентификации трещин в бортах карьера. Это свидетельствует о необходимости дальнейшей экспериментальной оценки их эффективности обнаружения. Некоторые результаты испытаний представлены на рис. 10.

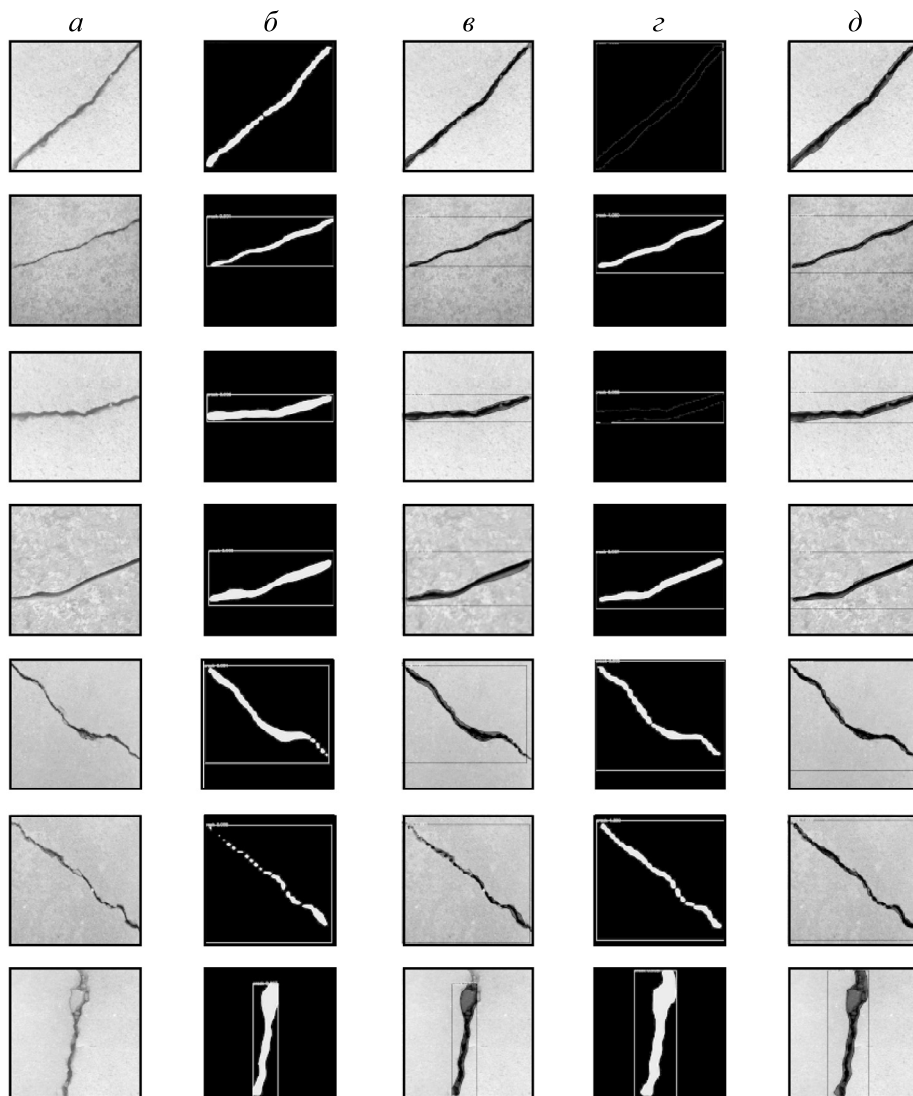


Рис. 10. Трещины в бортах карьера: исходные изображения из прогнозирующего набора (а); изображение трещины при архитектуре ResNet152 (б, в); изображение трещины при архитектуре ResNet101 (г, д)

При сравнении результатов идентификации трещин в бортах карьера с помощью ResNet101 и ResNet152 в качестве стержневых архитектур установлено, что при прямом использовании Mask R-CNN для поиска трещин ветвь обнаружения целей опознает их с большей эффективностью. Тем не менее наблюдается низкая точность сегментации маски, слабая четкость краев трещин, а также сложность разделения головных и хвостовых концов. Сравнение рис. 10б и г на четвертой и шестой линиях показало, что картина маски трещины, полученной ResNet101, более точная, чем ResNet152. При этом сегментация краев более детальна. Однако существуют некоторые недостатки — увеличение ширины трещины при сегментации и низкий эффект соответствия с фактической трещиной. Основываясь на более высоком качестве обнаружения при увеличении массива, можно сделать вывод, что архитектура ResNet101 является наиболее подходящей для модели поиска трещин в бортах карьера.

Для подтверждения увеличения точности обнаружения предлагаемого алгоритма выполнено сравнение изображений второго и четвертого столбцов (рис. 11б, г). Сопоставление выявило, что архитектура Aspp-PointRend-Mask R-CNN позволяет более детально обнаруживать и сегментировать трещины. Разработанная модель предлагает более точные результаты обнаружения трещин и сегментации их краев, высокую степень соответствия фактическим параметрам, а также более точный эффект позиционирования трещин из исходного изображения.

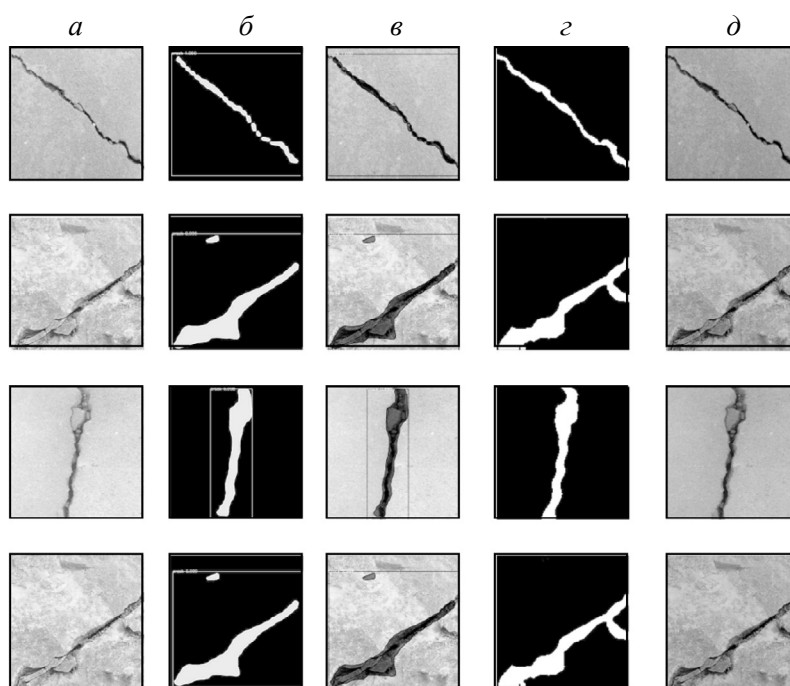


Рис. 11. Сравнение предлагаемого алгоритма обнаружения трещин в бортах карьера с неусовершенствованной архитектурой Mask R-CNN

Сопоставление экспериментальных результатов с численным анализом показало, что модель “интеллектуального” обнаружения трещин в бортах карьера обладает наилучшей применимостью и более детальной сегментацией по сравнению с широко распространенными алгоритмами идентификации нарушений. В результате создан прогнозирующий набор, содержащий 152 изображения трещин в бортах карьера, и рассчитаны значения четырех показателей оценки. В табл. 3 представлено сравнение результатов обнаружения между предлагаемым алгоритмом и неусовершенствованной архитектурой Mask R-CNN.

ТАБЛИЦА 3. Результаты сравнения предлагаемого алгоритма и неусовершенствованной архитектуры Mask R-CNN

Модель	Потери при обучении	Потери при проверке	Количество изображений	PA	MPA	MIoU	FWIoU
Предлагаемый алгоритм	0.1160	0.2985	2000	0.9825	0.8173	0.7835	0.9669
Aspp-ResNet101 – Mask RC-NN (усовершенствованная архитектура)	0.1300	0.4240	200	0.9606	0.7579	0.7153	0.9279
ResNet101 – Mask RC-NN (неусовершенствованная архитектура)	0.2024	0.4510	2000	0.9206	0.7179	0.6753	0.8871

Предлагаемая модель демонстрирует более высокое качество и точность. Например, пиксельная точность (PA) достигает 0.9825, что на 0.0619 больше, чем у архитектуры Mask R-CNN (0.9206). Среднее пересечение над объединением (MIoU) достигает 0,7835, что на 0.1082 выше, чем у неусовершенствованной архитектуры Mask R-CNN (0.6753). Отметим, что предлагаемая модель и неусовершенствованная архитектура обладают высокими результатами обнаружения в рамках рассматриваемой ветви и классовой сети. Точность ограничивающей рамки, сгенерированной для цели в виде трещины в борту карьера, превышает 98 %.

ВЫВОДЫ

Предложен алгоритм “интеллектуального” обнаружения трещин в бортах карьера на основе усовершенствованной архитектуры Mask R-CNN. Использовано увеличение объема массива изображений с помощью метода скользящего окна. В архитектуре Mask R-CNN выделены точки улучшения. Опытным путем определена наиболее подходящая стержневая архитектура для Mask R-CNN, а также выполнена экспериментальная проверка усовершенствованного алгоритма. Результаты испытаний показали, что по сравнению с традиционным алгоритмом обнаружения трещин архитектурой Mask R-CNN модель “интеллектуального” поиска обладает большей эффективностью опознавания и точностью сегментации.

Предложенная стержневая архитектура извлечения признаков в модели “интеллектуального” обнаружения трещин ResNet101-RPN позволяет получить высокие результаты обучения и проверки. Разработанный алгоритм показал высокую эффективность в процессе практического использования при мониторинге устойчивости бортов карьера.

СПИСОК ЛИТЕРАТУРЫ

1. **Yang T. H. et al.** Research situation of open-pit mining high and steep slope stability and its developing trend, *Rock & Soil Mechanics*, 2011, Vol. 32, No. 5. — P. 1437–1451.
2. **Jianping Chen, Ke Li, Kuo-Jen Chang, Giulia Sofia, Paolo Tarolli.** Open-pit mining geomorphic feature characterisation, *Int. J. App. Earth Observations & Geoinformation*, 2015, Vol. 42. — P. 76–86.
3. **Chen Q.** Influencing factors and preventive measures of slope stability in open-pit mines, *Opencast Min. Technol.*, 2019, Vol. 034, No. 1. — P. 92–94.
4. **Liu X. and Xie X.** Rapid crack inspection of tunnel surface based on image processing, *Chinese J. Underground Space and Eng.*, 2009, Vol. 5(a02). — P. 1624–1628.

5. **Jianzhen Gao et al.** Automatic road crack detection and identification, *Comput. Eng.*, 2003, Vol. 02. — P. 154–155.
6. **Li Q. and Liu X.** Novel approach to pavement image segmentation based on neighboring difference histogram method, *IEEE Comput. Soc.*, 2008.
7. **Jinhui L.** Image processing algorithm for detecting the pavement crack diseases, *Comput. Eng. and Appl.*, 2003, Vol. 39(035). — P. 212–213.
8. **Zhang J. et al.** Pavement crack automatic recognition based on phase-grouping method, *China J. Highway and Transport*, 2008(2). — P. 43–46.
9. **Hua Wang, Ning Zhu, and Wang Qi.** Segmentation of pavement cracks using differential box-counting approach, *J. Harbin Institute Technol.*, 2007, Vol. 1. — P. 142–144.
10. **Amhaz R. et al.** Automatic crack detection on 2D pavement images : An algorithm based on minimal path selection, *IEEE Transactions on Intelligent Transportation Systems*, 2016, Vol. 17, No. 10. — P. 2718–2729.
11. **Nguyen T. S. et al.** Free-form anisotropy: A new method for crack detection on pavement surface images. *IEEE Int. Conf. Image Proc.*, 2011.
12. **Zhibiao S. and Yanqing G.** Algorithm on contourlet domain in detection of road cracks for pavement images, *Int. Symp. on Distributed Comput. & Appl. to Business*, 2010.
13. **Chen F. C. and Jahanshahi R. M. R.** NB-CNN: Deep learning-based crack detection using convolutional neural network and naïve bayes data fusion, *IEEE Transactions on Industrial Electronics*, 2018, Vol. 65(99). — P. 4392–4400.
14. **Mustafa R. and Mohamed E. A.** Concrete crack detection based multi-block CLBP features and SVM classifier, *J. Theor. & Appl. Information Technol.*, 2015, Vol. 70, No. 6.
15. **Qin Zou et al.** Deep crack: learning hierarchical convolutional features for crack detection, *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2018.
16. **Quintana M., Torres J., and Menéndez J. M.** A Simplified computer vision system for road surface inspection and maintenance, *IEEE Transactions on Intelligent Transportation Systems*, 2016, Vol. 17(3). — P. 608–619.
17. **Tsai Y. C., Kaul V., and Mersereau R. M.** Critical assessment of pavement distress segmentation methods, *J. Transportation Eng.*, 2010, Vol. 136, No. 1. — P. 11–19.
18. **Yamaguchi T. and Hashimoto S.** Fast crack detection method for large-size concrete surface images using percolation-based image processing, *Machine Vision & Applications*, 2010, Vol. 21, No. 5. — P. 797–809.
19. **Yu Kai J. L. and Chen Yuqiang Wei Xu.** Deep learning: Yesterday, Today and Tomorrow, *J. Comput. Res. Development*, 2013.
20. **He K. et al.** Mask R-CNN, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, Vol. 42. — P. 1–1.
21. **Ren S. et al.** Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, Vol. 39, No. 6. — P. 1137–1149.
22. **Krizhevsky A., Sutskever I., and Hinton G.** Image net classification with deep convolutional neural networks, *Advances in Neural Information Proc. Systems*, 2012, Vol. 25, No. 2.
23. **Simonyan K. and Zisserman A.** Very deep convolutional networks for large-scale image recognition, *Int. Conf. on Learning Representations (ICLR)*, 2015.
24. **He K. et al.** Deep residual learning for image recognition, *IEEE Conf. Comput. Vision and Pattern Recognition*, 2016.

25. **Huang G. et al.** Densely connected convolutional networks, IEEE Conf. Comput. Vision and Pattern Recognition, 2017.
26. **Lin T. Y. et al.** Feature pyramid networks for object detection, IEEE Conf. Comput. Vision and Pattern Recognition, 2017.
27. **Liang-Chieh Chen G. P., Florian Schroff, and Hartwig Adam.** Rethinking atrous convolution for semantic image segmentation, arXiv:1706.05587v3 [cs.CV] 5 Dec 2017, 2017.
28. **Yu F. and Koltun V.** Multi-scale context aggregation by dilated convolutions, Int. Conf. Learning Representations (ICLR), 2016.
29. **Kirillov A. et al.** Point Rend: image segmentation as rendering, arXiv:1912.08193v2 [cs.CV] 16 Feb 2020, 2019.
30. **Platt J. C.** A fast algorithm for training support vector machines, J. Information Technol., 1998, Vol. 2, No. 5. — P. 1–28.

Поступила в редакцию 05/IV 2022
После доработки 05/V 2022
Принята к публикации 06/V 2022