

Семантическая сегментация ржавчин и пятнистостей пшеницы

И.В. Ариничев¹, С.В. Полянских², И.В. Ариничева³

¹ Кубанский государственный университет,

350040, Россия, г. Краснодар, ул. Ставропольская, д. 149;

² ООО Плариум-ЮГ, 350059, Россия, г. Краснодар, ул. Уральская, д. 75/1;

³ Кубанский государственный аграрный университет им. И.Т. Трубилина,
350044, Россия, г. Краснодар, ул. Калинина, д. 13

Аннотация

В статье исследуется возможность семантической сегментации классификации желтой ржавчины и пятнистости пшеницы с помощью сверточной нейросетевой архитектуры U-Net. На основе собственного набора данных, включающего 268 изображений, собранной в естественных условиях и условиях инфекционных питомников ФНЦ БЗР, показано, что архитектура U-Net с декодерами ResNet способна качественно обнаруживать, классифицировать и локализовывать ржавчины и пятнистости даже в тех случаях, когда болезни присутствуют на растении одновременно. Для отдельных классов болезней основные метрики (ассигасу, micro-/macro precision, recall и F1) колеблются в пределах от 0,92 до 0,96. Это указывает на возможность распознавания даже нескольких болезней на листе с точностью, не уступающей эксперти-фитопатологу. Метрики сегментации IoU и Dice составили соответственно 0,71 и 0,88, что говорит о достаточно высоком качестве попиксельной сегментации и подтверждается при визуальном анализе. Использованная при этом архитектура нейронной сети достаточно легковесна, что делает возможным ее использование на мобильных устройствах без подключения к сети.

Ключевые слова: семантическая сегментация, сверточная нейросеть, U-Net, болезни пшеницы, классификация болезней.

Цитирование: Ариничев, И.В. Семантическая сегментация ржавчин и пятнистостей пшеницы / И.В. Ариничев, С.В. Полянских, И.В. Ариничева // Компьютерная оптика. – 2023. – Т. 47, № 1. – С. 118-125. – DOI: 10.18287/2412-6179-CO-1130.

Citation: Arinichev IV, Polyanskikh SV, Arinicheva IV. Semantic segmentation of rusts and spots of wheat. Computer Optics 2023; 47(1): 118-125. DOI: 10.18287/2412-6179-CO-1130.

Введение

Одним из важнейших биотических факторов, влияющих на урожайность пшеницы, являются возбудители заболеваний. В патокмлексе пшеницы южных регионов России среди листовых болезней преобладают ржавчины: желтая (*Puccinia striiformis* f. sp. *tritici* West), бурая (*Puccinia tritici* cinaf. sp. *tritici* Erikss.), стеблевая (*Puccinia graminis* f. sp. *tritici* Pers.) – и желтая пятнистость (*Pyrenophora tritici-repentis* Drechsler), которые в годы эпифитотийного развития могут уносить до 50 % урожая [1]. При обнаружении конкретного вида ржавчин и пятнистостей нередко возникают трудности, связанные со схожестью в проявлениях заболеваний (например, бурая и стеблевая ржавчина), особенно на ранних стадиях. Решение проблемы качественной, оперативной и доступной диагностики ржавчин и пятнистостей пшеницы является актуальной и практически ценной задачей для фермеров, агрономов и фитопатологов.

Методы глубокого обучения и, в частности, сверточные нейронные сети привели к значительному прогрессу в задачах обнаружения и классификации болезней сельскохозяйственных культур по изображениям. Начиная с 2010 годов опубликовано значительное число работ, посвященных данному вопросу,

отличающихся по источникам происхождения данных, стратегиям обучения, архитектурным решениям моделей, методам регуляризации и т. д. (см., например, обзоры [2–4]).

Несмотря на широту охвата, в большинстве перечисленных работ исследования проводились на общедоступных наборах данных [5–9]. Подход с использованием открытых источников при своей простоте имеет и ряд недостатков. Во-первых, как показывает практика, даже одна и та же культура в разных частях света может выглядеть несколько иначе. То же самое касается и болезней культур. Во-вторых, сами условия съемки могут существенно отличаться от выборки к выборке и совершенно не подходить для текущего исследуемого случая. Таким образом, модели, обученные на открытых данных, зачастую просто неприменимы в качестве базовых решений для каждого конкретного случая. В отличие от имеющихся работ в настоящем исследовании использовалась база изображений болезней пшеницы, специально собранная и размеченная летом/осенью 2021 года в условиях полевого стационара ФГБНУ «Федеральный научный центр биологической защиты растений», г. Краснодар.

Логически данная статья является продолжением серии работ [10–11], в которых последовательно решались задачи строгой классификации грибных бо-

лезней пшеницы и риса в контролируемых условиях и множественной классификации болезней пшеницы, при одновременном проявлении нескольких патогенов на объекте. Однако с практической точки зрения важно не только выявлять наличие тех или иных болезней, но и локализовать их на изображении. Локализация позволяет не только детектировать наличие болезни, но и четко выделить пораженные участки растений, упрощая работу экспертов-фитопатологов и указывая места на изображении, на которые стоит обратить особое внимание. Кроме этого, становится возможной автоматическая количественная оценка степени поражения растения, если дополнительно известны фаза вегетации культуры и некоторые данные о биологии патогена.

Отметим, что термин «контролируемые условия» связан не с областью машинного обучения, а с контролем условий фотосъемки. Термин «контролируемые условия» в нашей и аналогичных работах других авторов означает фиксацию условия получения изображений: расстояние до объекта, угол наклона объектива, освещение, фон и т.д.

Целью данной работы является исследование возможности семантической сегментации двух болезней пшеницы – желтой ржавчины и желтой пятнистости листьев на основе U-Net модели сверточного нейронного автокодировщика с архитектурой ResNet в качестве базовой в контролируемых условиях сбора данных.

1. Материалы и методы

1.1. Анализ предшествующих работ

Все исследования по сегментации болезней растений можно разделить по типу используемых моделей на три класса – классические, сверточные нейросетевые и гибридные. Модели, основанные на классических подходах машинного обучения, показывают, вообще говоря, невысокое качество работы, их приходится объединять в ансамбли и использовать различные сложные техники пре- и постпроцессинга изображений. Кроме того, классические модели плохо масштабируются и обладают низкой универсальностью, сильно варьируясь даже для однотипных задач. Нейросетевые модели показывают значительно более высокое качество работы, нежели обычные, и достаточно универсальны в том смысле, что дают единый подход к решению целого спектра похожих задач. Гибридные модели обычно включают в себя ансамбли из нескольких моделей с нетривиальным пре- и постпроцессингом изображений.

В исследованиях [12, 13] сегментация и классификация пораженных участков листьев проводилась с помощью классической K-means кластеризации, которая соотносит каждый пиксель изображения с одним из классов заболеваний. Хотя средние значения метрик качества классификации можно считать удовлетворительными (precision, recall и F2-score в [12] соответственно равны 71,35, 60,55 и 60,83 %; accuracy

в [13] ~ 80 %), качество сегментации оставляло желать лучшего. Близкие по значению результаты (precision ~ 70–72 %, recall ~ 70–71 %) демонстрируют методы обучения с учителем Random Forest, Gradient Boosting, которые применялись для классификации пикселей изображений и как результат – сегментации областей заболевания [12]. Как отмечено выше, классические методы работают ожидаемо хуже нейросетевых, имеют низкую для рассматриваемых задач сложность и настраиваемость и, как следствие, обладают более низкой обобщающей способностью.

В работе [14] авторами предложен трехшаговый алгоритм, основанный на обучении нескольких CNN для обнаружения и локализации единственной болезни – северного гельминтоспориоза кукурузы. CNN были обучены сегментировать небольшие области изображений, содержащие болезнь. Их прогнозы объединялись в отдельные тепловые карты и передавались в окончательную CNN, обученную классифицировать все изображения как содержащие больные растения или нет. Авторами использовался набор данных, полученный в неконтролируемых условиях без фокусировки на конкретном органе растения, общим объемом 1796 изображений. Заявленная точность классификации на валидационной выборке составила 97,8 %. Схожий подход обсуждается в статье [15], где процесс предсказания трех болезней пшеницы разбивался на шаги: препроцессинг изображений, извлечение участков листьев – областей кандидатов на наличие болезней (Hot-Spot region); анализ каждого выбранного участка на наличие определенных классов заболеваний; этап интеграции, объединяющий информацию из предыдущих шагов, на котором сохранялись все истинно-положительные образцы и исключались ложноположительные результаты. Благодаря данной реализации, точность классификации на тестовой выборке, включающей 179 изображений, составила 96 %.

В указанных работах сегментация пораженных болезнями участков листьев не является конечной целью, а выступает промежуточным звеном идентификации заболевания. В результате даже при хороших метриках классификации болезней сама сегментация не имеет хорошей точности и требует значительных размеров датасетов. Как показано в настоящей работе, гораздо лучших результатов можно достичь, ставя на первое место именно задачу сегментации и решая задачу классификации уже постфактум.

В статье [12] предложена модель семантической сегментации мучнистой росы по изображениям листьев огурца, основанная на U-Net архитектуре. Объем исходной выборки составил 50 изображений, которая разбивалась на обучение (30 фото) и тест (20 фото). Экспериментальные результаты на 20 тестовых образцах показали, что по сравнению с существующими методами сегментации изображений K-means, Random Forest и GBDT предложенный метод

значительно повысил точность (метрика Dice = 83,45 %, IoU = 72,11 %). Исследование [16] направлено на решение проблемы обнаружения и сегментации бактериальной пятнистости листьев риса. Для решения данной задачи использовалась модель сверточной нейронной сети BLSNet, основанная на архитектуре U-Net и механизме внимания. BLSNet обучалась на изображениях, полученных в полевых условиях. По сравнению с эталонными моделями DeepLabV3+ и U-Net, для которых значения IoU составили 0,895 и 0,912 соответственно, BLSNet продемонстрировала самую высокую точность (IoU = 0,956). В работе [17] проводится сравнительный анализ трех семейств сверточных архитектур для обнаружения и сегментации болезней томатов в полевых условиях: Faster R-CNN (Faster Region-based CNN), R-FCN (Region-based Fully CNN), SSD (Single Shot Multibox Detector). Каждая из указанных архитектур объединялась с экстракторами признаков на основе VGG и ResNet. В общей сложности в статье представлены значения метрик IoU по восьми моделям для каждого из девяти классов болезней. Лучший средний результат показала R-FCN с экстракторами ResNet-50 (средняя точность составила 0,86). В [18] рассмотрена модель Fully-Convolutional-Network (FrCNN) для сегментации частей листа манго, пораженных антракнозом. Предлагаемая сеть напрямую изучает особенности каждого пикселя входных данных после применения некоторых методов предварительной обработки. Для оценки результатов моделирования качество сегментации сравнивалось с результатами сегментации на основе архитектур Vgg16, Vgg-19 и U-Net на наборе данных, содержащем 7040 изображений. Точность сопоставления сегментированной части с исходным изображением для предложенной авторами модели составила 99,2 % (для U-Net 91,7 %).

В настоящей работе исследуются возможности одной из наиболее простых, но в то же время эффективных архитектур семейства U-Net для семантической сегментации двух болезней пшеницы. Данная архитектура не слишком тяжела и вполне может быть использована в приложениях на мобильных устройствах, что важно для работы в полевых условиях. Также исследуются возможности обучения нейросети на выборке малого размера – всего 268 изображений против тысяч и десятков тысяч в других работах. Особенностью семейства архитектур U-Net является возможность качественного обучения модели в условиях малых наборов данных.

1.2. Подготовка обучающей выборки

Искусственные инфекционные фоны желтой ржавчины и желтой пятнистости листьев пшеницы были созданы в условиях полевого стационара ФГБНУ «Федеральный научный центр биологической защиты растений» с соблюдением пространственной изоляции в весенне-летний период 2021 г. Для иноку-

ляции растений, необходимой для получения инфекционного фона вышеописанных болезней, использовалась смесь урединиоспор с тальком в соотношении 1:100 при нагрузке 5 мг спор/м² для желтой ржавчины, для пиренофороза – водно-конидиальная суспензия с концентрацией $3-5 \times 10^3$ спор/мл (нагрузка 70–100 мл/м²). Учет развития болезней осуществлялся начиная с момента первичного проявления до фазы молочно-восковой спелости зерна с интервалом 10–12 суток.

Фотоснимки пораженных болезнями листьев пшеницы производились при искусственном освещении, на белом фоне, под углом 90°, на расстоянии 30–50 см до объекта съемки. Разрешение каждого фотоснимка – 1024 на 682 пикселя. В общей сложности объем выборки составил 268 объектов: желтая ржавчина – 111, желтая пятнистость – 101, оба класса одновременно – 56.

Чтобы обучить CNN сегментировать области с болезнями на листьях, необходимо разметить исходные изображения. Аннотации всех фотоснимков производились вручную. На рис. 1а представлен исходный образец одного изображения из выборки с одновременным проявлением болезней двух классов. На рис. 1б приведена целевая маска, на которой каждая болезнь аннотирована своим цветом (ржавчина – красным, пятнистость – синим).

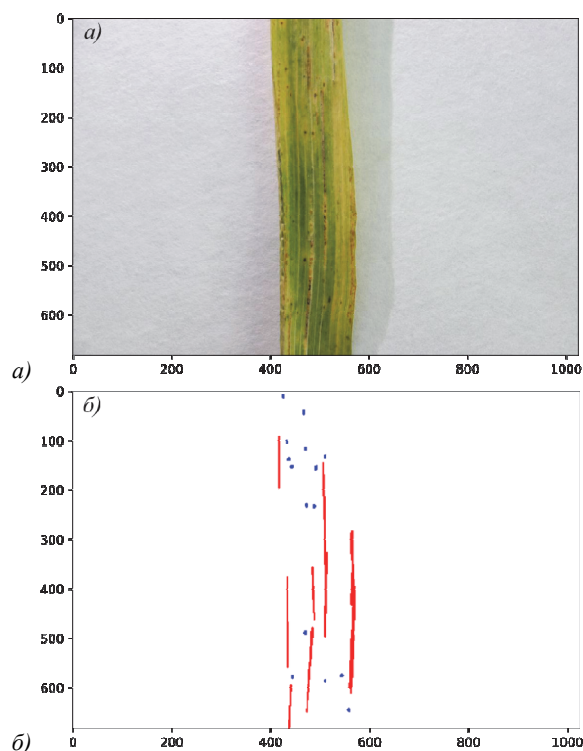


Рис. 1. Исходное изображение из выборки (а) и соответствующая ему размеченная маска болезней (б)

Из 268 изображений случайно отбиралось 214 пар (изображение и маска) в качестве тренировочного набора и 54 пары в качестве тестового набора для оценки производительности модели.

1.3. Препроцессинг данных

В данной работе задача сегментации решается в контролируемых условиях. Но даже следя за качеством и условиями съемки как при сборе данных, так и при использовании обученной модели, может возникнуть ряд проблем принципиального характера, способных существенно ухудшить качество модели. Среди них:

- недостаточный объем выборки;
- естественная инвариантность предсказаний относительно поворотов/отражений изображения;
- неустойчивость предсказаний, когда даже незначительный шум может изменить результат;
- эффект переобучения, когда качество предсказаний на новых изображениях оказывается значительно ниже, чем на обучающих.

Со всеми этими проблемами в определенной степени можно справиться, организовав грамотный препроцессинг, то есть предобработку исходных изображений. В данном исследовании мы используем следующие этапы препроцессинга исходного датасета:

- повороты на углы, кратные 45°;
- отражения относительно главных осей;
- случайные повороты на малые углы;
- стандартная нормализация RGB-каналов изображения, масштабирование изображений до размера 512×640×3.

В результате размер обучающей выборки увеличивается, повышая устойчивость предсказаний и обеспечивая их инвариантность к поворотам изображения.

1.4. Архитектура нейронной сети

Сверточная нейронная сеть, построенная в статье, базируется на архитектуре U-Net, показавшей отличные результаты в задаче сегментации биомедицинских изображений [19]. В отличие от полносверточных сетей, в которых теряется часть пространственной информации из-за наличия узких мест между слоями и резкого увеличения размера тензора в процессе увеличения разрешения, в моделях, базирующихся на U-Net, апсемплинг происходит более сглаженно. Архитектура U-Net состоит из двух частей – энкодера, сжимающего изображение, извлекая из него семантически богатые признаки, и декодера, действующего в обратном направлении, постепенно увеличивая размерность тензора, приводя его к исходным размерам. Чтобы обеспечить точность по픽сельно и сохранить пространственную информацию, между симметричными блоками энкодеров и декодеров в модели присутствуют горизонтальные связи. Конкатенируя соответствующие блоки в U-Net, производится агрегация признаков, которые были получены в энкодере с признаками, извлеченными в процессе декодинга, сохраняя таким образом часть информации, которая неизбежно терялась в полносверточных сетях и при апсемплинге.

ResNet как отдельная архитектура отлично показала себя в задачах компьютерного зрения. Универсальность данной архитектуры позволяет ее использовать не только саму по себе, но и встраивать в архитектуры более высокого уровня, как U-Net. В этом случае ResNet блоки могут быть использованы как в энкодере, так и в декодере. В настоящей работе мы использовали ResNet-34 вариант этой архитектуры, обладающий хорошей обобщающей способностью при оптимальном количестве параметров в окончательной модели. Структура сети ResNet-U-Net, которая использовалась в статье, приведена на рис. 2.

Важной особенностью U-Net является возможность тренировать «end-to-end» модели из выборок небольшого объема, что очень ценно для отрасли сельского хозяйства, где не всегда можно получить большие наборы данных для исследования. Сегментация болезней по существу является задачей классификации, которая выполняется для каждого пикселя. Функция потерь рассчитывалась по формуле (1), которая сочетает в себе стандартные методы оценки производительности моделей сегментации IoU (2) и Dice (3) с бинарной кросс-энтропией (4). Комбинация трех функций допускает некоторое разнообразие потерь, при этом выигрывая от стабильности энтропии. Преимущества комбинированной функции потерь в задачах сегментации описаны, например, в работе [20].

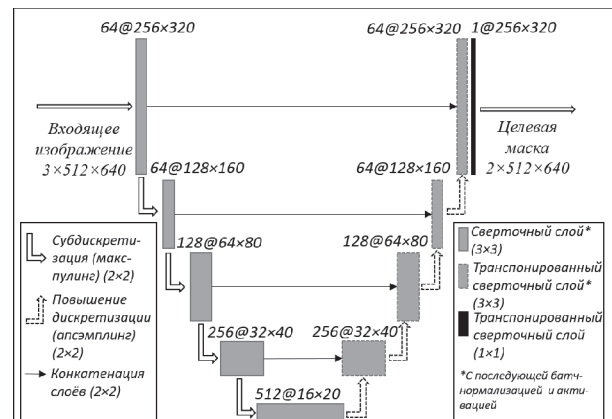


Рис. 2. Архитектура сверточной сети ResNet-U-Net

$$loss(y; \hat{y}) = \alpha \cdot BCE(y; \hat{y}) + \beta \cdot (1 - Dice(y; \hat{y})) + \gamma \cdot (1 - IoU(y; \hat{y})), \quad (1)$$

$$IoU(y; \hat{y}) = \left(\sum_{n=1}^m y_n \hat{y}_n \right) / \sum_{n=1}^m (y_n + \hat{y}_n - y_n \hat{y}_n), \quad (2)$$

$$Dice(y; \hat{y}) = 2 \cdot \left(\sum_{n=1}^m y_n \hat{y}_n \right) / \sum_{n=1}^m (y_n + \hat{y}_n), \quad (3)$$

$$BCE(y; \hat{y}) = - (1/m) \sum_{n=1}^m [y_n \log \hat{y}_n + (1 + y_n) \log(1 + \hat{y}_n)], \quad (4)$$

где m – число пикселей на изображении, занумерованных каким-либо образом; y_i – истинное значение тензора (наличие болезни) на i -м пикселе, равное 0

или 1; \hat{y}_i – прогнозируемое значение болезни на i -м пикселе в диапазоне от 0 до 1; α , β , γ – веса целевой функции, подбираются по отложенной выборке.

Учитывая, что оба тензора y и \hat{y} имеют размер $2 \times \text{Ш} \times \text{В}$, формулы (2–4) для каждого изображения вычисляются дважды, для каждой из болезней, после чего значения метрик суммируются и подставляются в функцию потерь (1).

Модель тренировалась с помощью фреймворка PyTorch. В процессе оптимизации применялся метод Адама со стандартными настройками и подобранным постоянным значением коэффициента скорости обучения, равным 0,0001. Модель обучалась с использованием 16 Гб оперативной памяти на базе графического процессора Tesla V100 в системной среде Ubuntu 18.04 LTS.

2. Результаты и обсуждение

На рис. 3 приведены графики функции потерь (1) на обучающей и тестовой выборках.

За сравнительно небольшое (~30) число эпох функционал вышел на плато со значениями, близкими к нулю. Метрики (2) и (3), усредненные по тестовой выборке, составили соответственно 0,71 и 0,88, откуда следует достаточно высокая точность сегментации.

На рис. 4 приведены результаты семантической сегментации ResNet-UNet для трех случаев проявления болезней на листьях пшеницы: желтая ржавчина (рис. 4а), желтая пятнистость (рис. 4б), желтая ржавчина и желтая пятнистость (рис. 4в). Первое изображение на каждом из трех рисунков – исходный фотоснимок из тестовой выборки; второе – размеченная маска болезней; третье – результаты сегментации с помощью ResNet-UNet. Визуальный анализ изображений на рис. 4 подтверждает численные результаты моделирования, выраженные сходимостью функционала потерь к нулю и высокими значениями метрик (2) и (3).

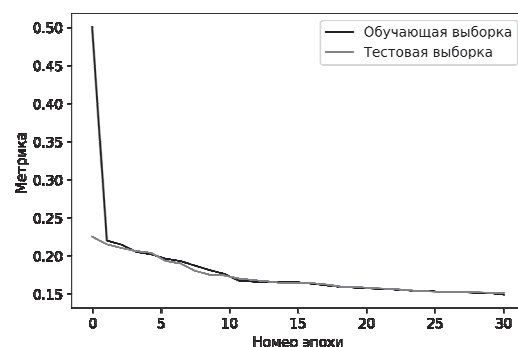


Рис. 3. Функционал потерь на обучающей и тестовой выборках

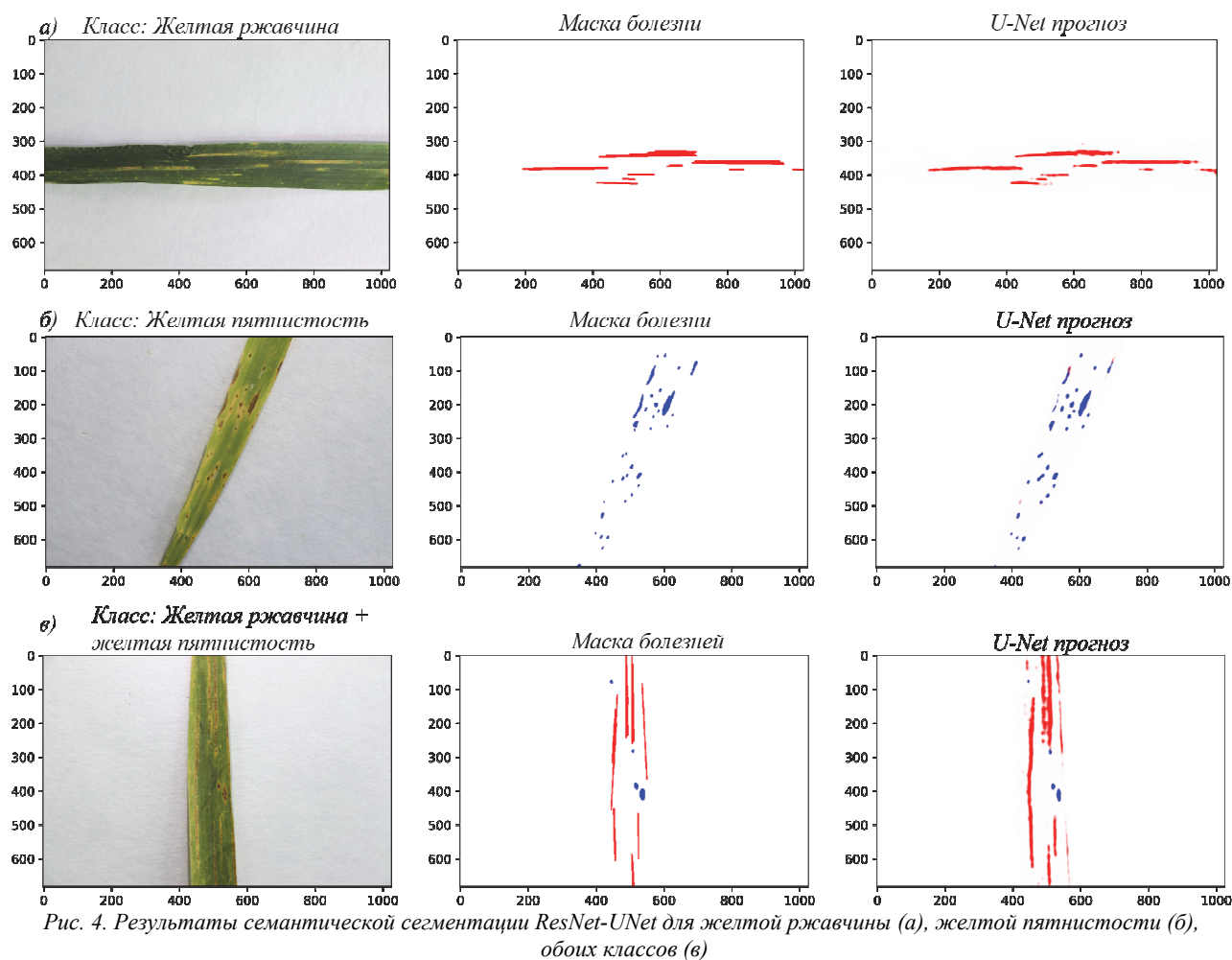


Рис. 4. Результаты семантической сегментации ResNet-UNet для желтой ржавчины (а), желтой пятнистости (б), обоих классов (в)

Как отмечалось выше, на выходе модели ResNet - UNet находится тензор \hat{y} , элементы которого характеризуют вероятность наличия в каждом пикселе одного из двух классов болезней.

Чтобы оценить, как алгоритм предсказывает каждый из классов, перейдем от задачи сегментации к задаче классификации. С этой целью необходимо было установить пороги – степень уверенности модели в том, что болезнь точно присутствует в данном пикселе, а также минимальное количество таких пикселей, когда болезнь может быть соотнесена с целым фотоснимком (а не только пикселем). Множественный классификатор f , устанавливающий соответствие между исходным изображением и классами заболеваний, представим в виде:

$$f_j(x) = [\sum_{i=1}^m \hat{y}_{ij}(x) > t] > p, j = 0, 1, \quad (5)$$

где x – объект выборки (изображение); m – количество пикселей на изображении, занумерованных каким-либо образом; y_{ij} – прогноз модели ResNet-UNet для болезни j -го класса по i -му пикселю; $[\]$ – скобки Айверсона; t, p – пороги (настраиваются по отложенной выборке).

Заметим, что одному и тому же объекту может быть поставлено в соответствие одновременно две единичные метки, что соответствует проявлению сразу двух заболеваний на листе.

Оценка результатов прогнозирования модели у множественных классификаторов несколько сложнее, чем у строгих (один объект – одна болезнь). В случае строгой классификации все основные метрики основываются на матрице ошибок, которая дает ясное представление о качестве модели. В нашем случае на одной фотографии может быть обнаружено более одной болезни, поэтому метрики качества не могут быть полностью сведены к операциям над матрицей ошибок, как это имеет место в задачах классификации (в т.ч. мультиклассовой). По этой причине одновременно с матрицей ошибок (табл. 1) используется набор метрик *accuarcy*, *precision*, *recall* и *F-мера*, усредненных как по классам (*macro-averaging*), так и по числу объектов в классах (*micro-averaging*). Детальное описание этих метрик можно найти в [21].

Табл. 1. Матрица ошибок модели (*P* – желтая ржавчина, *П* – желтая пятнистость, *П+P* – желтая пятнистость и желтая ржавчина)

	<i>P</i>	<i>П</i>	<i>П+P</i>
<i>P</i>	19	0	4
<i>П</i>	0	19	0
<i>П+P</i>	3	2	7

В табл. 2 приведена первая группа наиболее легко интерпретируемых метрик, основанных на доле правильных ответов классификации. Использование доли правильных ответов в качестве показательной

метрики обосновывается не слишком большой стратификацией классов в исследуемом в работе случае.

Первые два столбца – доля изображений с указанной болезнью, которая была верно распознана, независимо от того, встретила ли она отдельно на листе либо пересеклась с другим классом заболеваний на одном объекте. *Exact assuarcy* – это доля фотоснимков, где модель в точности распознала обе болезни (или их отсутствие). Последняя метрика является довольно жесткой, так как небольшие отклонения от идеальной классификации неизбежны и допустимы, но даже в этом случае 83 % верно распознанных болезней – внушительный результат. При этом доля верно классифицированных болезней по отдельности еще выше – от 88 до 96 %, что не хуже по качеству, чем результат классификации для эксперта-фитопатолога.

Табл. 2. Метрика *accuarcy* для итоговой модели

Ассурасу, желтая ржавчина	Ассурасу, желтая пятнистость	Exact assuarcy
0,96	0,88	0,83

В табл. 3. приведены глобальные метрики модели – точность (6), полнота (7) и мера *F1* (8) в их микро- и макровариантах. Данная группа метрик отражает качество модели на всех данных в целом, а не на каждом классе в отдельности. В микро-случае мы усредняем значения, взвешивая их с частотами классов, а в макро- – просто вычисляем их среднее арифметическое.

$$precision = \frac{TP}{TP + FP}, \quad (6)$$

$$recall = \frac{TP}{TP + FN}, \quad (7)$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}, \quad (8)$$

где *TP* – число верно классифицированных примеров некоторой болезни, *FP* – число примеров, неверно классифицированных моделью зараженных некоторой болезнью, *FN* – число примеров, зараженных данной болезнью, которые модель пропустила.

Из табл. 3 видно, что данная группа метрик также указывает на высокое качество модели.

Табл. 3. Глобальные метрики для итоговой модели

Micro			Macro		
Precision	Recall	F1	Precision	Recall	F1
0,924	0,938	0,931	0,923	0,937	0,929

Заключение

В работе рассмотрен «end-to-end» подход к задаче семантической сегментации и классификации болезней желтой ржавчины и желтой пятнистости пшеницы. Собрана и размечена собственная обучающая выборка, включающая 268 изображений листьев пше-

ницы, пораженных грибными болезнями (111 объектов – желтая ржавчина, 101 объект – желтая пятнистость, 56 объектов – пятнистость и ржавчина одновременно). Показано, что сверточная нейросетевая архитектура U-Net с ResNet-основой решает поставленную задачу на уровне точности, сравнимом с уровнем экспертов-фитопатологов. Нейросеть хорошо обучается уже на собственном датасете, содержащем всего 268 изображений. Выбранная архитектура качественно решает сразу три задачи: обнаруживает болезнь на листе, классифицирует и попиксельно локализует ржавчины и пятнистости даже в тех случаях, когда болезни присутствуют на растении одновременно. Стандартные для такого рода задач метрики классификации колеблются в пределах от 0,83 до 0,96, что свидетельствует о высоком качестве модели. Сама архитектура при этом может использоваться как онлайн, так и полностью офлайн благодаря небольшому объему потребляемой памяти.

Благодарности

Исследование выполнено при финансовой поддержке Кубанского научного фонда в рамках научно-го проекта № МФИ-20.1/121.

References

- [1] Matveeva IP, Volkova GV. Yellow rust of wheat. Expansion, harm, control measures (review). Vestnik of Ulyanovsk State Agricultural Academy 2019; 46(2): 102-116. DOI: 10.18286/1816-4501-2019-2-102-116.
- [2] Boulent J, Foucher S, Theau J, St-Charles PL. Convolutional neural networks for the automatic identification of plant diseases. Front Plant Sci 2019; 10: 941. DOI: 10.3389/fpls.2019.00941.
- [3] Ngugi LC, Abelwahab M, Abo-Zahhad M. Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. Inf Process Agric 2021; 8(1): 27-51. DOI: 10.1016/j.inpa.2020.04.004.
- [4] Saleem MH, Potgieter J, Arif KM. Plant disease detection and classification by deep learning. Plants 2019; 8(11): 468. DOI: 10.3390/plants8110468.
- [5] Atabay H. Deep residual learning for tomato plant leaf disease identification. J Theor Appl Inf Technol 2017; 95(24): 6800-6808.
- [6] Brahimi M, Boukhalifa K, Moussaoui A. Deep learning for tomato diseases: classification and symptoms visualization. Appl Artif Intell 2017; 31: 299-315. DOI: 10.1080/08839514.2017.1315516.
- [7] Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Comput Electron Agric 2018; 145: 311-318. DOI: 10.1016/j.compag.2018.01.009.
- [8] Mohanty SP, Hughes DP, Salathe M. Using deep learning for image-based plant disease detection. Front Plant Sci 2016; 7: 1419. DOI: 10.3389/fpls.2016.01419.
- [9] Wang G, Sun Y., Wang J. Automatic image-based plant disease severity estimation using deep learning. Comput Intell Neurosci 2017; 2017: 2917536. DOI: 10.1155/2017/2917536.
- [10] Arinichev IV, Polyanskikh SV, Volkova GV, Arinicheva IV. Rice fungal diseases recognition using modern computer vision techniques. Int J Fuzzy Log Intell Syst 2021; 21(1): 1-11. DOI: 10.5391/IJFIS.2021.21.1.1.
- [11] Polyanskikh SV, Arinicheva IV, Arinichev IV, Volkova GV. Autoencoders for semantic segmentation of rice fungal diseases. Agron Res 2021, 19(2): 574-585. DOI: 10.15159/AR.21.019.
- [12] Lin K, Gong L, Huang Y, Liu C, Pan J. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front Plant Sci 2019; 10: 155. DOI: 10.3389/fpls.2019.00155.
- [13] Zhang S, Wuc X, You Z, Zhang L. Leaf image based cucumber disease recognition using sparse representation classification. Comput Electron Agric 2017; 2017(134): 135-141. DOI: 10.1016/j.compag.2017.01.014.
- [14] DeChant C, Wiesner-Hanks T, Stewart E, Gore M. Automated Identification of Northern leaf blight-infected maize plants from field imagery using deep learning. Phytopathology 2017; 107: 1426-1432. DOI: 10.1094/PHYTO-11-16-0417-R.
- [15] Picon A, Alvarez-Gila A, Seitz M, Ortiz-Barredo A, Echazarra J, Johannes A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Comput Electron Agric 2018; 2018(161): 280-290. DOI: 10.1016/j.compag.2018.04.002.
- [16] Chen S, Zhang K, Zhao Y, Sun Y, Ban W, Chen Y, Zhuang H, Zhang X, Liu J, Yang T. An approach for rice bacterial leaf streak disease segmentation and disease severity estimation. Agriculture 2021, 11: 420. DOI: 10.3390/agriculture11050420.
- [17] Fuentes AF, Yoon S, Lee J, Park DS. High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank. Front Plant Sci 2018; 29(9): 1162. DOI: 10.3389/fpls.2018.01162.
- [18] Saleem R, Shah JH, Sharif M, Ansari GJ. Mango leaf disease identification using fully resolution convolutional network. Comput Mater Contin 2021; 69(3): 3581-3601. DOI: 10.32604/cmc.2021.017700.
- [19] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. arXiv Preprint. 2015. Source: <https://arxiv.org/abs/1505.04597v1>.
- [20] Berman M, Triki AR, Blaschko MB. The Lovász-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. arXiv Preprint. 2017. Source: <https://arxiv.org/abs/1705.08790>.
- [21] Bishop CM. Pattern recognition and machine learning. Cambridge: Springer; 2006. ISBN: 978-0-387-31073-2.

Сведения об авторах

Ариничев Игорь Владимирович, 1985 года рождения, в 2008 году с отличием окончил Кубанский государственный университет по специальности «Математика». Работает доцентом кафедры теоретической экономики ФГБОУ ВО КубГУ. Область научных интересов: компьютерное зрение, программирование, автоматизация производственных процессов в растениеводстве. E-mail: iarinichev@gmail.com.

Полянских Сергей Валерьевич, с отличием окончил математический факультет Кубанского государственного университета. Работает ведущим специалистом по машинному обучению в компании ООО «Плариум». Область научных интересов: машинное обучение, программирование, электрогидродинамическая устойчивость микротечений. E-mail: spmathf@gmail.com.

Ариничева Ирина Владимировна, окончила Санкт-Петербургский государственный морской технический университет, доктор биологических наук. Работает профессором кафедры высшей математики Кубанского государственного аграрного университета. Область научных интересов: математические методы и модели в растениеводстве. E-mail: loukianova7@mail.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 22 марта 2022 г. Окончательный вариант – 16 июля 2022 г.

Semantic segmentation of rusts and spots of wheat

I.V. Arinichev¹, S.V. Polyanskikh², I.V. Arinicheva³

¹ Kuban State University, 350040, Krasnodar, Russia, Stavropolskaya 149;

² Plarium, 350059, Krasnodar, Russia, Uralskaya 75/1;

³ Kuban State Agrarian University named after I.T. Trubilin, 350044, Krasnodar, Russia, Kalinina 13

Abstract

The paper explores the possibility of semantic segmentation of the yellow rust and wheat blotch classification using the U-Net convolutional neural network architecture. Based on an own dataset of 268 images, collected in natural conditions and in infectious nurseries of the Federal Research Center for Biological Plant Protection (VNII BZR), it is shown that the U-Net architecture with ResNet decoders is able to qualitatively detect, classify and localize rust and spotting even in cases where diseases are present on the plant at the same time. For individual classes of diseases, the main metrics (accuracy, micro-/macro precision, recall, and F1) range from 0.92 to 0.96. This indicates the possibility of recognizing even a few diseases on a leaf with an accuracy that is not inferior to that of a plant pathology expert. The IoU and Dice segmentation metrics are 0.71 and 0.88, respectively, which indicates a fairly high quality of pixel-by-pixel segmentation and is confirmed by visual analysis. The architecture of the neural network used in this case is quite lightweight, which makes it possible to use it on mobile devices without connecting to the network.

Keywords: semantic segmentation, convolutional neural network, U-Net, wheat diseases, classification of diseases.

Citation: Arinichev IV, Polyanskikh SV, Arinicheva IV. Semantic segmentation of rusts and spots of wheat. *Computer Optics* 2023; 47(1): 118-125. DOI: 10.18287/2412-6179-CO-1130.

Acknowledgements: This work was supported by the Kuban science Foundation (Project No. IFR-20.1/121).

Authors' information

Igor Vladimirovich Arinichev, (b. 1985) graduated from Kuban State University in 2008, majoring in Mathematics. Currently he works as the assistant professor at the Assistant Professor, Theoretical Economics department of Kuban State University. Research interests are computer vision, programming, automation of production processes in crop production. E-mail: iarinichev@gmail.com.

Sergey Valerievich Polyanskikh, graduated with honors from Mathematics faculty of Kuban State University. Works as a senior in machine learning at Plarium LLC. Research interests: machine learning, programming, electrohydrodynamic stability of microflows. E-mail: spmathf@gmail.com.

Irina Vladimirovna Arinicheva, graduated from St. Petersburg State Marine Technical University, Doctor of Biological Sciences. Works as a professor of Higher Mathematics department of Kuban State Agrarian University. Research interests: mathematical methods and models in crop production. E-mail: loukianova7@mail.ru.

Received March 22, 2022. The final version – July 16, 2022.
