УДК 004.89
DOI 10.58649/1694-9099-2026-3-534-544

**Лэй Синь, Бексултанов Ж.Т.**
Жусуп Баласагын атындагы КУУ
**Лэй Синь, Бексултанов Ж.Т.**
КНУ имени Жусупа Баласагына
**Lei Xin, Beksultanov Zh.T.**
KNU Jusup Balasagyn

# КОШ АГЫМДУУ ОПТИМАЛДАШТЫРУУ: ӨСҮМДҮКТӨРДҮН ПАТОЛОГИЯСЫН ИШЕНИМДҮҮ ДИАГНОЗДОО ҮЧҮН КАТМАР АРАЛЫК КӨҢҮЛ БУРУУНУ ЖАНА ТҮРГӨ МҮНӨЗДҮҮ АПРИОРДУК МААЛЫМАТТАРДЫ ИНТЕГРАЦИЯЛОО

# ДВУХПОТОКОВАЯ ОПТИМИЗАЦИЯ: ИНТЕГРАЦИЯ МЕЖСЛОЙНОГО ВНИМАНИЯ И ВИДОВЫХ АПРИОРНЫХ ДАННЫХ ДЛЯ НАДЕЖНОЙ ДИАГНОСТИКИ ФИТОПАТОЛОГИИ

# DUAL-STREAM OPTIMIZATION: INTEGRATING CROSS-LAYER ATTENTION AND SPECIES PRIORS FOR ROBUST PLANT DISEASE DIAGNOSIS

**Кыскача мүнөздөмө:** Дүйнөлүк айыл чарба экономикасынын туруктуулугу жана азык-түлүк коопсуздугу өсүмдүк ооруларын эрте жана так диагноздоо менен тыгыз байланышта. Терең үйрөтүү айыл чарбасындагы компьютердик көрүү системаларында олуттуу өзгөрүүлөрдү жараткандыгына карабастан, заманбап архитектуралар карама-каршылыкка туш болууда: семантикалык белгилерди бөлүп алуу үчүн терең тармактар талап кылынат, бирок мындай тармактарга мүнөздүү болгон масштабды кичирейтүү оорулардын майда жабыркоолорун аныктоо үчүн зарыл болгон жогорку жыштыктагы текстуралык маалыматтын сапатын төмөндөтөт. Бул карама-каршылыкты чечүү үчүн изилдөөдө белгилерди бөлүп алуунун деталдуулугун жана классификациялоо логикасын жакшыртууга багытталган терең үйрөтүүнүн жаңы алкагы — LANet сунушталат. LANet эки маанилүү инновацияны камтыйт. Биринчиден, мейкиндиктик деталдардын төмөндөн өйдө карай таралышын камсыз кылуу үчүн катмар аралык көңүл буруу (Layer Attention, LA) механизми киргизилип, төмөнкү деңгээлдеги текстураны жогорку деңгээлдеги семантика менен натыйжалуу айкалыштыруу үчүн масштабдар аралык салмактар эсептелет. Экинчиден, таксономиялык жактан башкарылуучу трансфердик үйрөтүү (Transfer Learning, TL) стратегиясы сунушталат. Алгач тармакты өсүмдүктөрдүн түрлөрүн таанууга үйрөтүп, андан кийин ооруга мүнөздүү катмарларды окутуу үчүн бул параметрлерди тоңдуруу аркылуу өсүмдүк түрү менен патогенге сезгичтиктин ортосундагы биологиялык көз карандылык моделдештирилет. PlantVillage маалыматтар топтомунда жүргүзүлгөн эксперименталдык текшерүү LANet 85,62% тактыкка жетишип, тогуз заманбап архитектурадан жогорку натыйжа көрсөткөнүн далилдейт. Бул иш айыл чарбасына автоматташтырылган мониторинг жүргүзүү үчүн масштабдуу жана жогорку тактыктагы чечимди сунуштайт.

**Аннотация:** Стабильность глобальной сельскохозяйственной экономики и продовольственная безопасность неразрывно связаны с ранней и точной диагностикой болезней сельскохозяйственных культур. Хотя глубокое обучение произвело революцию в компьютерном зрении в сельском хозяйстве, современные архитектуры сталкиваются с дихотомией: глубокие сети необходимы для извлечения семантических признаков, однако присущее этим сетям понижение разрешения часто ухудшает высокочастотную текстурную информацию, необходимую для идентификации тонких поражений, вызванных болезнями. Для разрешения этого конфликта в данном исследовании представлена LANet — новая структура глубокого обучения, разработанная для повышения детализации извлечения признаков и логики классификации. LANet вносит два важных нововведения. Во-первых, в нее встроен

механизм внимания слоев (Layer Attention, LA) для облегчения распространения пространственных деталей снизу вверх, вычисляющий межмасштабные веса для эффективного объединения низкоуровневой текстуры с высокоуровневой семантикой. Во-вторых, мы предлагаем таксономически управляемую стратегию трансферного обучения (Transfer Learning, TL). Сначала обучая сеть распознаванию видов растений, а затем замораживая эти параметры для обучения слоев, специфичных для болезней, мы явно моделируем биологическую зависимость между идентичностью хозяина и восприимчивостью к патогену. Экспериментальная проверка на наборе данных PlantVillage демонстрирует, что LANet достигает точности 85,62%, превосходя девять передовых архитектур. Эта работа предлагает масштабируемое высокоточное решение для автоматизированного мониторинга сельского хозяйства.

**Abstract:** The stability of global agricultural economics and food security is intrinsically linked to the early and accurate diagnosis of crop diseases. While deep learning has revolutionized computer vision in agriculture, current architectures face a dichotomy: deep networks are required to extract semantic features, yet the downsampling inherent in these networks often degrades the high-frequency textural information necessary to identify subtle disease lesions. To resolve this conflict, this study introduces LANet, a novel deep learning framework designed to enhance the granularity of feature extraction and the logic of classification. LANet introduces two critical innovations. First, a Layer Attention (LA) mechanism is embedded to facilitate a bottom-up propagation of spatial details, computing cross-scale weights to fuse low-level texture with high-level semantics effectively. Second, we propose a taxonomically guided Transfer Learning (TL) strategy. By first training the network to recognize plant species and subsequently freezing these parameters to train disease-specific layers, we explicitly model the biological dependency between host identity and pathogen susceptibility. Experimental validation on the PlantVillage dataset demonstrates that LANet achieves an accuracy of 85.62%, outperforming nine state-of-the-art architectures. This work offers a scalable, high-precision solution for automated agricultural monitoring.

**Негизги сөздөр:** өсүмдүктөрдүн ооруларын таануу; сүрөттөрдү иштетүү; терең үйрөтүү; катмар аралык көңүл буруу; трансфердик үйрөтүү.

**Ключевые слова:** распознавание болезней растений; обработка изображений; глубокое обучение; межслойное внимание; трансферное обучение.

**Keywords:** plant disease recognition; image processing; deep learning; layer attention; transfer learning.

**Введение.** Модернизация сельского хозяйства основывается на интеграции интеллектуальных систем, способных осуществлять высокоточный мониторинг состояния сельскохозяйственных культур. Болезни растений, вызываемые грибами, бактериями и вирусами, выступают основными стрессовыми факторами для глобальных продовольственных цепочек поставок, будучи способными приводить к резкому снижению урожайности и ухудшению качества продукции при отсутствии своевременного контроля [1]. Экономические последствия могут быть крайне серьёзными; например, соевая ржавчина или пшеничная пирикулёриозная инфекция (wheat blast) способны привести к полной гибели урожая в течение нескольких недель после заражения. В этой связи автоматизация диагностики заболеваний является не просто техническим удобством, а объективной необходимостью для обеспечения устойчивого производства продовольствия. Последние достижения в области искусственного интеллекта (ИИ), в частности сверточных нейронных сетей (Convolutional Neural Networks, CNN), предоставили инструменты для масштабного анализа изображений растений. Однако применение универсальных моделей компьютерного зрения к задачам фитопатологии сопряжено с рядом специфических трудностей [2]. В отличие от задач общего распознавания объектов (например, идентификации автомобиля или кошки), распознавание болезней растений в значительной степени опирается на выявление тонких текстурных аномалий — некротических поражений, хлороза или грибных пятен, — которые занимают лишь незначительную долю пиксельного пространства

изображения. Как показано на рис. 1, заболевания различных видов растений могут одновременно проявляться схожими симптомами на листьях. Стандартные архитектуры глубокого обучения используют агрессивное понижение пространственного разрешения (операции пулинга и свёртки со шагом больше единицы) для расширения рецептивного поля и захвата глобальной семантической информации. Однако данный процесс неизбежно приводит к утрате высокочастотных пространственных деталей, необходимых для дифференциации заболеваний, проявляющихся в виде сходных некротических пятен на листьях различных растений [3].



(a)Черная гниль винограда



(b) Ожог листьев земляники



(c) Септориозная пятнистость листьев томата

**Рис. 1. Иллюстрация листьев различных растений и соответствующих им заболеваний**

Кроме того, существующие сквозные (end-to-end) модели классификации часто игнорируют иерархическую биологическую таксономию рассматриваемой задачи. Для решения указанных проблем нами предлагается LANet—специализированная архитектура, разработанная для устранения компромисса между семантической абстракцией и сохранением пространственного разрешения. LANet состоит из двух основных компонентов. В обобщённом виде наше исследование вносит три ключевых вклада в область распознавания болезней растений:

1. Мы предлагаем LANet — лёгкую, но в то же время высокоэффективную нейронную сеть, интегрирующую модифицированную базовую архитектуру MobileNetV3 с новым механизмом внимания, специально оптимизированным с учётом ограничений краевых вычислений в сельском хозяйстве.
2. Мы разрабатываем модуль межмасштабного внимания, который динамически переназначает веса признаков, поступающих из последующих слоёв сети. Это позволяет модели «обращаться» к признакам нижних уровней для восстановления пространственных деталей, утраченных в процессе понижения разрешения, тем самым обеспечивая сохранение мельчайших симптомов заболевания в глубоких слоях сети.

Мы предлагаем новый подход к переносу обучения, заключающийся в предварительном обучении сети классификации категорий растений с последующей фиксацией её параметров и использовании полученной модели в качестве направляющего механизма при классификации заболеваний растений, что приводит к повышению точности и устойчивости модели.

**Связанные исследования.** Ранняя и точная идентификация болезней растений имеет решающее значение для поддержания эффективного управления сельскохозяйственным производством и сохранения урожайности. В данном разделе рассматриваются современные исследования и практические применения как традиционных, так и основанных на глубоком обучении методов распознавания болезней растений по изображениям.

**a. Традиционные методы.** До эпохи глубокого обучения системы распознавания опирались на вручную заданные дескрипторы. Исследователи применяли преобразования цветовых пространств (из RGB в HSV или CIELAB) для выделения очагов поражения на основе хроматических различий. Для количественной оценки текстуры использовались статистические показатели, такие как эксцесс,

асимметрия и матрицы совместной встречаемости уровней серого (Gray Level Co-occurrence Matrix, GLCM) [4]. Несмотря на вычислительную экономичность, данные методы обладали низкой обобщающей способностью: незначительные изменения условий освещения или ориентации листа часто приводили к неактуальности вручную подобранных пороговых значений. Кроме того, дескрипторы формы (эксцентриситет, сплошность), применявшиеся для классификации бактериальных заболеваний риса [5], оказались недостаточно эффективными для болезней, проявляющихся в виде нерегулярных, аморфных пятен, а не чётко очерченных геометрических структур.

**b. Методы, основанные на глубоком обучении.** Появление сверточных нейронных сетей (CNN) ознаменовало парадигмальный сдвиг в данной области. Prathiksha и соавт. продемонстрировали эффективность архитектуры VGG-19 в задачах виноградарства, достигнув точности 98% на наборах данных по винограду [6]. Однако архитектуры семейства VGG характеризуются высокой вычислительной сложностью, что делает их непригодными для развертывания на беспилотных летательных аппаратах или портативных устройствах. Более современные облегчённые архитектуры, такие как MobileNet и EfficientNet, применяются для сокращения числа параметров, однако в задачах тонкой классификации они нередко сталкиваются с проблемой «исчезновения деталей».

Для компенсации потери информации были предложены механизмы внимания. Канальное внимание (определяющее, какие признаки являются значимыми) и пространственное внимание (определяющее, где расположен значимый признак) позволили повысить точность распознавания [7]. Тем не менее большинство существующих модулей внимания функционируют в пределах одного слоя. Они не решают явно проблему утраты информации, возникающую между слоями при переходе от высокоразрешающих поверхностных признаков к низкоразрешающим глубоким представлениям. LANet устраняет данное ограничение за счёт применения межслойного механизма внимания.

**Метод.** В данном разделе подробно рассматриваются проектирование и структура предлагаемой модели LANet.

**Базовая архитектура (Backbone).** Структура сети представлена на рис. 2. В качестве базового модуля извлечения признаков используются первые пять инвертированных остаточных блоков MobileNetV3, включающие глубинную свёртку (depthwise convolution), слой расширения и слой проекции. Слой расширения применяет свёртку 1 × 1 для отображения признаков из пространства меньшей размерности в пространство большей размерности на уровне каналов. В соответствии с архитектурой MobileNetV3 число каналов по умолчанию увеличивается в шесть раз. Глубинно-разделимая свёртка (depthwise separable convolution) заменяет традиционный полный свёрточный оператор факторизованной версией, состоящей из двух последовательных слоёв. Первый слой выполняет облегчённую фильтрацию, используя по одному свёрточному фильтру на каждый входной канал. Второй слой представляет собой поточечную свёртку (pointwise convolution), реализуемую посредством свёртки 1 × 1, которая формирует новые признаки за счёт линейных комбинаций входных каналов. Узкое место (bottleneck) реализуется с использованием свёрток 1 × 1 с линейными преобразованиями, что позволяет эффективно сокращать размерность пространств активаций, минимизируя потери информации и обеспечивая оптимальный баланс между вычислительной эффективностью и точностью модели. Наконец, каждый инвертированный остаточный блок включает структуру Squeeze-and-Excitation для усиления представления признаков.

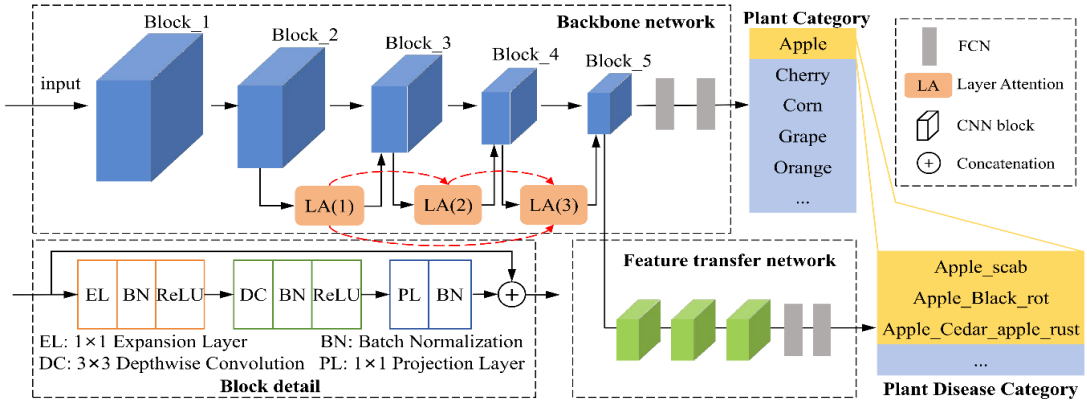



**Рис. 2. Структура сети предлагаемой нами локальной сети (LANet)**

**b. Межслойное внимание (Layer attention, LA).** Традиционное распознавание болезней листьев растений, как правило, основывается на разработке правил на пиксельном уровне для выделения деталей поражённых областей. Эти методы демонстрировали определённую эффективность. Большинство подходов, основанных на глубоком обучении, проектируют глубокие нейронные сети для автоматического извлечения многомасштабных признаков, при этом различные свёрточные ядра фокусируются на разнообразных геометрических и текстурных характеристиках областей поражения на изображениях. Однако по мере увеличения глубины сети утрата пиксельной информации, связанной с заболеванием, приводит к возникновению узкого места в производительности распознавания. Для эффективного выявления заболеваний, требующих сохранения детализированной информации, нами предлагается модуль межслойного внимания, интегрируемый между уровнями восходящей (bottom-up) структуры с целью передачи полезных деталей от низкоуровневых признаков к более высоким уровням. Наш модуль внимания к слоям построен на основе многомасштабных признаков базовой сети. Как показано на рис. 2, каждый масштабный признак генерируется из блока CNN. Всего у нас 5 блоков, и мы строим модуль внимания к слоям на слоях со 2 по 5. Схема модуля внимания к слоям показана на рис. 3.

Выходы блоков, показанных на рис. 2, обозначаются как $b_1$, $b_2$, $b_3$, и $b_4$. Для хранения выходов каждого модуля LA используется множество $s = \{f_i\}$, где $i = \{1, 2, 3\}$.Как показано на рис. 3, признак $b_i$ соответствует выходу предыдущего блока CNN, тогда как признак $b_{i+1}$ (жёлтый фон) является выходом последующего блока CNN. Для пары $b_i$ и $b_{i+1}$ выполняется операция самовнимания (self-attention). Для многоголового самовнимания с H головами вычисление матриц запроса query **Q**, key **K**, и value **V** можно представить в виде: $\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2,..., \mathbf{Q}_H]$, $\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2,..., \mathbf{K}_H]$, и $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2,..., \mathbf{V}_H]$, $\mathbf{Q}_h$ , $\mathbf{K}_h$ , $\mathbf{V}_{h \in R_H^D}$ , $h \in \{1, 2,..., H\}$, а $D$— размерность признакового пространства. Обозначим $\mathbf{Q}_h = w_q^h b_i$, $\mathbf{K}_h = w_k^h b_{i+1}$, $\mathbf{V}_h = w_v^h \cdot b_{i+1}$, где $w_q^h$ , $w_k^h$ , $w_v^h \in R_H^D$. Для позиционного кодирования используются функции синуса и косинуса [19]. Процесс межслойного внимания может быть представлен следующим образом:

$$A_h = \text{Softmax}\left(\frac{\mathbf{Q}_h \mathbf{K}_h^T}{\sqrt{d_k}}\right)\mathbf{V}_h, h = 1, ..., H \quad (1)$$

$$A = \text{Concat}(A_1, A_2, ..., A_h) \quad (2)$$

$$f_i = \text{MLP}(A) \quad (3)$$

где $d_k = \sqrt{\frac{D}{H}}$ а MLP()-многослойный персептрон (MLP) с одним скрытым слоем, $f_i$ обозначает выход $i$+1-го блока CNN с механизмом внимания слоя на начальном этапе обучения. В процессе $b_2$ обрабатывается с помощью LA для получения $f_1$. Затем $f_1$ и $b_2$ объединяются на уровне каналов, после чего выполняется еще одна операция LA для получения $f_2$. Наконец, $f_1$, $f_2$, $b_3$ объединяются, и LA применяется еще раз для генерации $f_3$. Таким образом, каждый последующий слой может учитывать все предыдущие слои, сохраняя более подробную информацию в представлении признаков. Из соображений производительности модуль LA не применяется к Block_1.

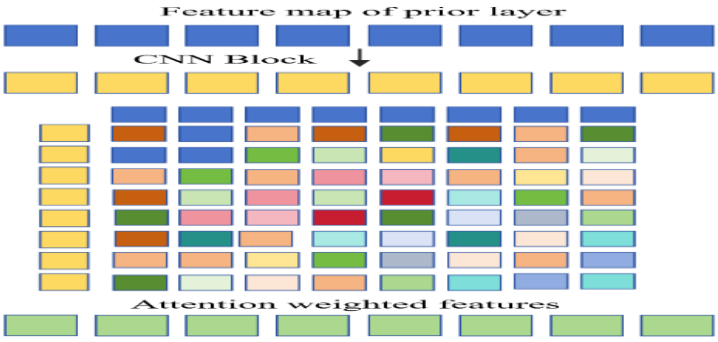



**Рис. 3. Схема модуля внимания слоя**

**c. Перенос обучения для распознавания болезней растений.** В существующих исследованиях по распознаванию болезней растений преимущественно используются сквозные (end-to-end) модели, в которых метки заболеваний напрямую применяются в качестве целевых переменных при обучении. Однако детальный анализ показывает, что такие модели игнорируют взаимосвязь между категориями растений и соответствующими им заболеваниями, тогда как тип растения во многом определяет спектр возможных болезней. Например, Grape Black Rot и Strawberry Leaf Scorch проявляют сходные симптомы, но поражают разные виды растений. В то время как большинство алгоритмов машинного обучения сосредоточено на решении отдельных задач прогнозирования, перенос обучения позволяет интегрировать данные из различных источников для решения целевой задачи, что потенциально приводит к более эффективным результатам.

Исходя из этого, нами разработана новая модель, основанная на концепции переноса обучения. Предлагаемый метод включает два основных этапа. На первом этапе компактная сеть, содержащая пять блоков CNN с межслойным вниманием, обучается для распознавания категорий растений. Затем параметры данной сети фиксируются, после чего добавляются ещё три блока CNN для наделения модели способностью классифицировать болезни листьев растений. В частности, как показано на рис. 2, для распознавания заболеваний используется сеть переноса признаков. Блоки CNN в этой сети имеют ту же структуру, что и Block_5. Процесс обучения для распознавания категорий растений можно представить следующим образом:

$$f_{cate}=\mathrm{CNN}_{\mathrm{LA}}(I) \qquad (4)$$

где $f_{cate}$ -выход сети $\mathrm{CNN}_{\mathrm{LA}}$ , представляющей базовую архитектуру, а $I$-входное изображение. Параметры $\theta$ данных блоков оптимизируются путём минимизации функции потерь $L_{cat}$:

$$\theta^{*}=\mathrm{argmin}_{\theta}L_{cat}(y_{cat},\hat{y}_{cat}) \qquad (5)$$

где $y_{cat}$ и $\hat{y}_{cat}$ — истинные и предсказанные метки категорий растений соответственно. Функция $L_{cat}$ определяется как кросс-энтропийная потеря. После завершения обучения первой части параметры $\theta^{*}$ фиксируются. Далее добавляются три новых блока CNN $G^{(j)}$ для распознавания заболеваний. Блок $G^{(1)}$ получает в качестве входа выход Block_5. Параметры $\phi$ новых блоков оптимизируются посредством минимизации функции потерь $L_{disease}$:

$$\phi^{*}=\mathrm{argmin}_{\phi}L_{disease}(y_{disease},\hat{y}_{disease}) \qquad (6)$$

где $L_{disease}$ также представляет собой кросс-энтропийную функцию потерь.

Такая организация обеспечивает чёткое определение каждого этапа и переходов между слоями, подчёркивая стратегию переноса признаков, применяемую в сети для оптимизации распознавания заболеваний на основе выявленных признаков, связанных с категорией растения.

**Эксперименты.** В данном разделе сначала описываются экспериментальная постановка и детали обучения предложенного метода, затем проводится его валидация и сравнение с современными моделями уровня SOTA на общедоступном наборе данных PlantVillage [8], и, наконец, выполняются абляционные исследования для дополнительной проверки эффективности предложенных модулей.

**a. Экспериментальная постановка.** 1) Набор данных: В работе использован набор данных PlantVillage — эталонный бенчмарк в области сельскохозяйственного компьютерного зрения. Полная версия содержит 54 305 изображений по 38 категориям. Для обеспечения строгого тестирования и сбалансированности классов случайным образом было отобрано 10 000 изображений, сформировав подмножество из 13 видов растений и 26 классов заболеваний. Для обучающей и валидационной выборок поддерживалось строгое соотношение 1:1 с целью предотвращения переобучения.

2) Архитектура сети и гиперпараметры: Базовая сеть включает 5 блоков с глубинно-разделимыми свёртками и остаточными структурами. Выходы блоков имеют размерности [16,112,112], [16,56,56], [24,28,28], [40,14,14] и [40,14,14], где значения соответствуют числу каналов и размеру карты признаков. Для снижения вычислительных затрат модули LA добавлены к 3-му, 4-му и 5-му слоям. Эти модули обеспечивают распространение признаков между слоями, компенсируя потерю детализированной информации. Модуль LA включает два энкодера: трансформер по каналам для передачи признаков и свёрточный слой для согласования размерностей. Далее применяется функция активации h-swish. Сеть переноса признаков содержит три блока CNN для классификации заболеваний, каждый идентичен по структуре Block_5, что повышает стабильность и устойчивость обучения.

Блок «neck» для распознавания категорий включает два слоя FCN. Первый слой агрегирует признаки из Block_5. Размер выходного тензора Block_5 составляет [1,576,7,7] (batch_size, channel_num, width, height). После агрегации размерность уменьшается до [1,576], затем признаки подаются во второй FCN-слой. Входная размерность равна 576, скрытая — 1024, выходная — 11 (число категорий листьев). В сети переноса используется аналогичная структура FCN, за исключением изменения числа выходных категорий с 11 до 26.

3) Детали реализации: Применялись методы аугментации, аналогичные MobileNetV3: 1) rслучайный поворот в диапазоне [-π/4, +π/4]; 2) масштабирование в пределах [0.95,1.05]; and 3) случайное горизонтальное и вертикальное отражение с вероятностями ph=0.5 и pv=0.5.

Оптимизация выполнялась с использованием *adam_onecycle* с шагом обучения 0.001, импульсом 0.9 и weight decay 0.01. Обучение проводилось на двух GPU GTX3090 с размером батча 64 в течение 100 эпох. Применялась стратегия Dropout с вероятностью 0.2. Планировщик StepLR использовался с шагом 2 и коэффициентом уменьшения 0.973. Подробные конфигурации приведены в open-mmlab.

**b. Результаты.** Предложенный метод сравнивался с моделями SOTA в одинаковых условиях. Метрикой оценки служила точность (accuracy). Результаты представлены в Таблице 1.

Как видно из Таблицы I, MobileViT и EfficientNet достигли около 80%, что отражает сложность задачи для стандартных облегчённых моделей. Примечательно, что чистые трансформеры, такие как ViT (82.05%), не показали доминирующих результатов. Трансформеры обычно требуют больших объёмов данных для формирования индуктивных смещений (например, локальности), которыми CNN обладают изначально. При использовании подмножества из 10,000 изображений архитектура LANet, объединяющая индуктивные свойства CNN и контекстную чувствительность механизма внимания, продемонстрировала наилучшие результаты. LANet достиг точности 85.62%, превзойдя ближайшего конкурента Riformer на 1.6%. Это подтверждает, что использование априорной информации о виде растения и организация межслойного взаимодействия эффективнее, чем простое увеличение сложности модели.

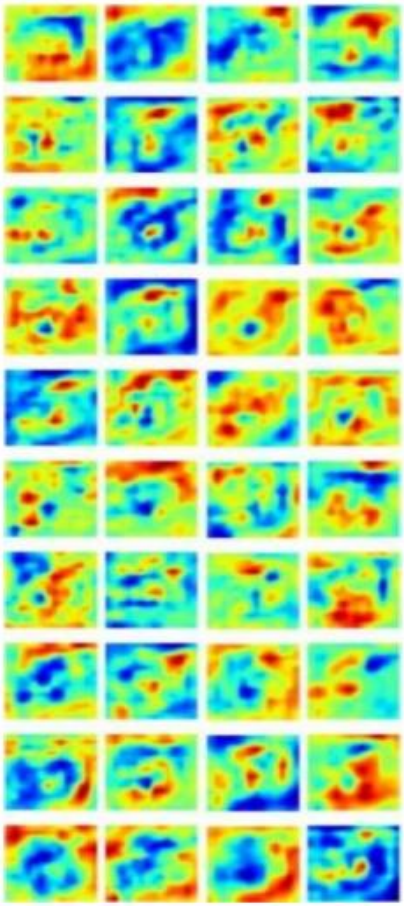
**Таблица 1. Сравнение точности распознавания заболеваний листьев с методами SOTA**

| Архитектура модели | Тип модели | Точность(%) |
|---|---|---|
| MobileViT | Облегчённая | 79.75 |
| EfficientNet | Облегчённая CNN | 80.40 |
| VIG | Графовая | 80.63 |
| Levit | Transformer | 80.77 |
| RevViT | Transformer | 81.92 |
| ViT | Transformer | 82.05 |

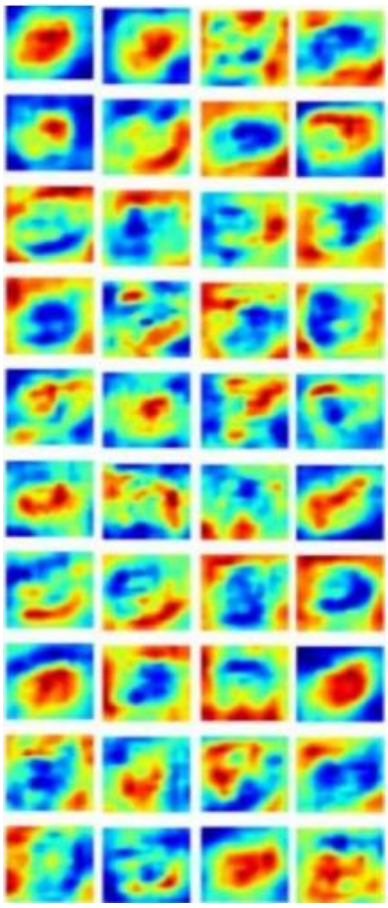
| | | |
|---|---|---|
| ResNest | CNN с вниманием | 82.11 |
| EfficientFormer | Гибридная | 83.89 |
| Riformer | Гибридная | 84.02 |
| **Ours** | **Attentional CNN** | **85.62** |

**c. Визуализация.** Для интерпретации процесса принятия решений использовался метод CAM++ (Class Activation Mapping). В исходном MobileNet (рис. 4a) карты активации были размытыми и часто фокусировались на здоровых участках листа или фоновых элементах. После внедрения Layer Attention (рис. 4b) тепловые карты стали значительно более локализованными в областях некротических поражений.

В сравнении с EfficientFormer и Riformer (рис. 5) LANet продемонстрировал более высокую способность выявлять несколько разрозненных очагов заболевания на одном листе. В то время как другие модели обычно концентрировались на наиболее крупном пятне, механизм внимания LANet позволял агрегировать признаки из распределённых поражений, обеспечивая более целостную диагностику.



(a) Карта признаков 5-го слоя в MobileNet



(b) карта признаков 5-го слоя в нашей модели

**Рис. 4. Карты признаков, извлеченные с использованием различных методов на слое 5.**

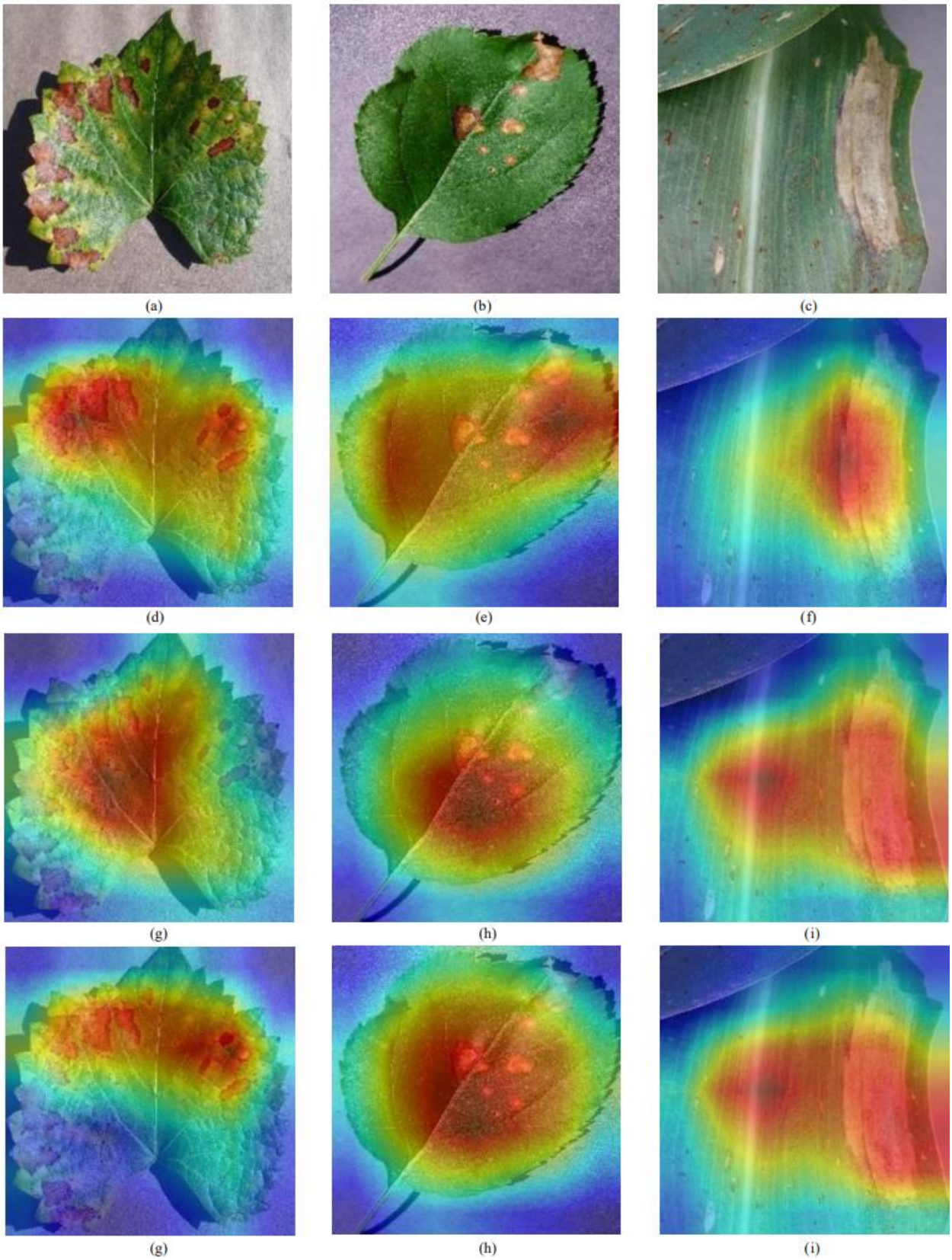



**Рис. 5. Тепловые карты, сгенерированные с использованием различных методов: Верхний ряд: исходные изображения; Второй ряд: предложенный LANet; Третий ряд: EfficientFormer; последний ряд: Riformer**

**d. Абляционное исследование. Эффективность модулей.** Результаты, представленные в Таблице 2, демонстрируют, что как модуль межслойного внимания (Layer Attention), так и модуль Category-Disease Loss способствуют повышению производительности модели, при этом Layer Attention

обеспечивает более выраженный прирост точности. Это указывает на то, что снижение потерь детализированной признаковой информации существенно улучшает качество распознавания болезней растений. Совместное использование обоих модулей позволило достичь наилучших результатов.

**Эффективность модуля переноса признаков.** Для дополнительной валидации процесса переноса обучения проведено абляционное исследование с двумя различными конфигурациями, направленное на анализ влияния переноса признаков на точность распознавания заболеваний.

В модели A блок G(1) инициализируется случайными весами без использования предварительно обученных признаков и служит контрольной группой для оценки влияния переноса обучения. В модели B блок G(1) получает признаки из Block_5 с фиксированными параметрами, реализуя стандартную схему переноса обучения. Обе модели обучаются на одном и том же наборе данных с одинаковыми гиперпараметрами и числом эпох, что обеспечивает корректность сравнения. Результаты представлены в таблице III. Модель B превосходит модель A, что демонстрирует эффективность использования предварительно обученных признаков для повышения точности распознавания болезней растений.

**Таблица 2. Влияние различных модулей в предложенном методе**

| **Внимание к слою** | **Категория – Потеря от болезней** | **Точность** |
|---|---|---|
| | ✓ | 70.31 |
| ✓ | ✓ | 75.88 |
| ✓ | | 82.30 |
| | | **85.62** |

**Таблица 3. Оценка влияния переноса обучения на распознавание болезней растений**

| | **Точность**(%) |
|---|---|
| МодельA | 81.20 |
| МодельB | **85.62** |

**Заключение.** В данной работе предложен метод распознавания болезней листьев растений, включающий два новых модуля. Введён механизм межслойного внимания для компенсации потерь пространственной информации, вызванных понижением разрешения в глубоких слоях сети, что позволило расширить архитектуру по горизонтали и увеличить её представительную способность. Кроме того, разработан подход переноса обучения, усиливающий способность модели к распознаванию заболеваний за счёт использования предварительно обученных признаков, полученных при классификации категорий растений. Проведённые обширные эксперименты подтвердили эффективность предложенных решений.

**Список использованной литературы**

1. Demilie W. B. Plant disease detection and classification techniques: a comparative study of the performances // Journal of Big Data. – 2024. – Vol. 11, № 1. – P. 5.
2. Kumari N., Saini B. Fully automatic wheat disease detection system by using different CNN models // Sentiment Analysis and Deep Learning: Proceedings of ICSADL 2022. – Springer, 2023. – P. 351–365.
3. Sunil C., Jaidhar C., Patil N. Systematic study on deep learning-based plant disease detection or classification // Artificial Intelligence Review. – 2023. – Vol. 56, № 12. – P. 14955–15052.
4. Haridasan A., Thomas J., Raj E. D. Deep learning system for paddy plant disease detection and classification // Environmental Monitoring and Assessment. – 2023. – Vol. 195, № 1. – P. 120.
5. Kaur S., Pandey S., Goel S. Semi-automatic leaf disease detection and classification system for soybean culture // IET Image Processing. – 2018. – Vol. 12, № 6. – P. 1038–1048.

6. Kai S., Zhikun L., Hang S., Chunhong G. A research of maize disease image recognition of corn based on BP networks // 2011 Third International Conference on Measuring Technology and Mechatronics Automation. – IEEE, 2011. – Vol. 1. – P. 246–249.
7. Joshi A. A., Jadhav B. Monitoring and controlling rice diseases using image processing techniques // 2016 International Conference on Computing, Analytics and Security Trends (CAST). – IEEE, 2016. – P. 471–476.
8. Hughes D., Salathé M. et al. An open access repository of images on plant health to enable the development of mobile disease diagnostics // arXiv preprint. – 2015. – arXiv:1511.08060.

**Рецензент: кандидат физико-математических наук, доцент Эшаров Э.А.**