

Научная статья | Original paper

УДК 004.85:004.932

Режимы адаптации заранее обученных нейросетей в задаче классификации визуально близких классов: эксперимент на наборе PlantVillage

А.С.А. Мутхана¹ ✉, Е.В. Ляпунцова²

¹ Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСИС»» (НИТУ МИСИС)
Москва, Российская Федерация

² Московский государственный технический университет имени Н.Э. Баумана (ФГАОУ ВО «МГТУ им. Н.Э. Баумана»)
Москва, Российская Федерация

✉ adammadam265@gmail.com

Резюме

Классификация визуально близких категорий остается одной из наиболее трудных задач анализа изображений, поскольку межклассовые различия часто малы, а внутриклассовая вариативность, напротив, велика. Цель исследования состояла в оценке того, как режим адаптации заранее обученной нейросетевой модели влияет на качество распознавания в задаче высокой визуальной сложности при фиксированной архитектуре и единых условиях эксперимента. Рабочая гипотеза предполагала, что частичное дообучение верхних слоев заранее обученной сверточной нейронной сети даст более высокий результат, чем режим, при котором сверточная основа остается замороженной, а обучение затрагивает только завершающий классификационный блок. В качестве стандартизированного испытательного полигона использован открытый набор PlantVillage, включающий 54 303 изображения и 38 классов; предметная область набора рассматривалась как удобная модель сложной классификации визуально близких состояний. Базовой архитектурой служила MobileNetV3Small. Сопоставлялись два режима адаптации: замороженная сверточная основа и частичное дообучение верхней части признакового ядра. Основным результатом получен при частичном дообучении: точность на проверочной части выборки возросла с 0,9707 до 0,9816, а значение функции потерь снизилось с 0,0929 до 0,0576. Поклассовый анализ



лучшей модели на итоговой части выборки показал, что у 33 из 38 классов мера F1 превысила 0,95; наибольшие трудности возникли в группах с высокой визуальной близостью, где минимальные значения меры F1 составили 0,8889 и 0,9078. Научная новизна исследования заключается не в самой идее частичного дообучения заранее обученной сверточной сети, поскольку данный подход широко применяется в задачах компьютерного зрения, а в экспериментальном сопоставлении двух режимов адаптации MobileNetV3Small в строго одинаковых условиях на наборе PlantVillage. Полученные результаты показывают, что для выбранной архитектуры, фиксированного разбиения данных и задачи классификации визуально близких классов частичное дообучение верхней части признакового ядра обеспечивает более высокое качество распознавания и формирует содержательно интерпретируемую структуру ошибок.

Ключевые слова: обучение с переносом, режимы адаптации нейросетей, сверточные нейронные сети, частичное дообучение, классификация изображений, визуально близкие классы, анализ ошибок, межклассовые смещения, стандартизированный набор данных, PlantVillage

Для цитирования: Мутхана, А.С.А., Ляпунцова, Е.В. (2026). Режимы адаптации заранее обученных нейросетей в задаче классификации визуально близких классов: эксперимент на наборе PlantVillage. *Моделирование и анализ данных*, 16(2), 127–145. <https://doi.org/10.17759/mda.2026160207>

Adaptation regimes of pretrained neural networks in visually proximate class classification: evidence from the PlantVillage benchmark

A.S. A.Muthana¹ ✉, E.V. Lyapunтова²

¹ National University of Science and Technology MISIS (NUST MISIS)
Moscow, Russian Federation

² Bauman Moscow State Technical University
(FSAEI HE Bauman Moscow State Technical University)
Moscow, Russian Federation

✉ adamdam265@gmail.com

Abstract

Classifying visually proximate categories remains a difficult problem in image analysis because inter-class differences are often small, whereas intra-class variability is substantial. The aim of the study was to evaluate how the adaptation regime of a pretrained neural network affects performance in a high-complexity recognition task under a fixed architecture and identical experimental conditions. The working hypothesis assumed that partial fine-tuning of the upper layers of a pretrained convolutional neural network would outperform a regime in which the convolutional



backbone remains frozen and only the final classification block is trained. The open PlantVillage dataset, containing 54,303 images and 38 classes, was used as a standardized benchmark; its subject domain was treated as a convenient testbed for complex classification of visually similar states. MobileNetV3Small served as the base model. Two adaptation regimes were compared: a frozen convolutional backbone and partial fine-tuning of the upper part of the feature extractor. The main gain was achieved with partial fine-tuning: validation accuracy increased from 0.9707 to 0.9816, while validation loss decreased from 0.0929 to 0.0576. Class-wise analysis on the independent test split showed that the F1-score exceeded 0.95 for 33 of 38 classes, whereas the lowest values, 0.8889 and 0.9078, were observed in groups with high visual similarity. The scientific novelty does not lie in the general idea of partial fine-tuning of a pretrained convolutional network, since this approach is widely used in computer vision, but in the controlled experimental comparison of two MobileNetV3Small adaptation regimes under identical conditions on the PlantVillage dataset. The results show that, for the selected architecture, fixed data split, and visually proximate class classification task, partial fine-tuning of the upper part of the feature extractor improves recognition quality and produces an interpretable error structure.

Keywords: transfer learning, neural network adaptation regimes, convolutional neural networks, partial fine-tuning, image classification, visually proximate classes, error analysis, inter-class confusion, benchmark dataset, PlantVillage

For citation: Muthana, A.S.A., Lyapunsova, E.V. (2026). Adaptation regimes of pretrained neural networks in visually proximate class classification: evidence from the PlantVillage benchmark. *Modelling and Data Analysis*, 16(2), 127–145. (In Russ.). <https://doi.org/10.17759/mda.2026160207>

Введение

Классификация визуально близких категорий относится к числу наиболее трудных задач компьютерного зрения, когда различия между классами малы, а вариативность внутри класса заметна даже при стандартизированной съемке. В обзорах (Ramanjot et al., 2023) и (Pascal et al., 2024) подобные постановки рассматриваются как характерная зона напряжения для современных моделей глубокого обучения, поскольку итоговое качество здесь определяется не только архитектурой сети, но и способом ее адаптации к целевому набору данных.

Открытые коллекции изображений листьев растений удобны в качестве экспериментального полигона для анализа таких режимов адаптации. Наборы этого типа объединяют достаточно большое число классов, выраженную межклассовую близость и хорошо формализованную разметку, благодаря чему позволяют исследовать поведение модели в задаче тонкого различения визуальных состояний. В обзоре (Zhao et al., 2025) подчеркивается, что именно сходство текстурных, цветовых и морфологических признаков делает подобные данные полезными для проверки устойчивости алгоритмов распознавания.



Перенос обучения в рассматриваемой постановке важен прежде всего как механизм повторного использования ранее сформированного признакового пространства. Практический вопрос заключается не в самой возможности переноса, а в выборе режима адаптации заранее обученной сети: оставить сверточную основу неизменной либо открыть часть слоев для последующей настройки. Работа (Richter, Kim, 2025) показывает, что глубина такой адаптации заметно влияет на итоговое качество даже при использовании одних и тех же открытых наборов данных, а (Shafik et al., 2024) связывает эффективность подхода с тем, насколько точно модель подстраивается под специфику целевого визуального материала.

Стандартизированные наборы изображений дают удобную среду для сопоставления моделей, однако интерпретация результатов требует аккуратности. Контролируемый фон, крупный план объекта и ограниченное число внешних помех упрощают задачу классификации по сравнению с более сложными сценариями распознавания, где присутствуют доменный сдвиг, вариативность ракурсов и неоднородное освещение. По этой причине исследовательский интерес в данной работе сосредоточен на сравнении режимов адаптации в одинаковых условиях, а предметная область набора данных используется как воспроизводимый пример задачи высокой визуальной сложности.

Цель исследования заключается в оценке влияния режима адаптации заранее обученной нейросетевой модели на качество классификации визуально близких классов при фиксированной архитектуре и единых экспериментальных условиях. Рабочая гипотеза состоит в том, что частичное дообучение верхних слоев заранее обученной сверточной нейронной сети обеспечит более высокое качество распознавания по сравнению с режимом, при котором изменяется только завершающий классификационный блок. Для проверки гипотезы сопоставление проводится на одной архитектуре, одном стандартизированном наборе изображений, фиксированном разбиении выборки и общей системе показателей качества.

Задачи исследования:

1. Охарактеризовать применимость обучения с переносом к задаче классификации визуально близких классов на основе недавних научных публикаций (Ramanjot et al., 2023; Pacal et al., 2024; Zhao et al., 2025; Shafik et al., 2024; Richter, Kim, 2025);
2. Реализовать два режима адаптации заранее обученной сверточной модели, включающие замороженную сверточную основу и частичное дообучение верхней части сети;
3. Сравнить полученные модели по точности, функции потерь и поклассовым показателям качества, а также выявить группы классов, в которых межклассовое смешение проявляется наиболее отчетливо.

Материалы и методы

Каталог TensorFlow Datasets (TensorFlow, 2024a) фиксирует для набора PlantVillage 54 303 изображения, распределенные по 38 классам. В настоящем исследовании этот набор использовался не как источник данных для прикладной агрономической



экспертизы, а как стандартизированный испытательный полигон для анализа режимов адаптации нейросетей в задаче классификации визуально близких категорий. Работа (Rahman, Islam, Islam, 2024) также показывает, что PlantVillage удобен для сопоставления заранее обученных сверточных сетей благодаря воспроизводимой разметке, однородной подаче объекта и достаточному числу классов. В исследование была включена версия набора без категории Background_without_leaves, используемая в реализации TensorFlow Datasets.

Таблица 1 / Table 1

**Характеристики экспериментального набора
данных и схема разбиения выборки**
Characteristics of the experimental dataset and sample splitting scheme

Показатель	Значение
Набор данных	PlantVillage
Общее число изображений	54 303
Число категорий	38
Исходный размер изображения	$256 \times 256 \times 3$
Рабочий размер изображения	$224 \times 224 \times 3$
Обучающая часть	38 012
Проверочная часть	8 146
Итоговая часть	8 145
Доля обучающей части, %	70,0
Доля проверочной части, %	15,0
Доля итоговой части, %	15,0
Фиксированное начальное состояние генератора	42

Примечание: Численность частей выборки получена при однократном воспроизводимом разбиении полного набора данных.

Визуальный материал набора демонстрирует удобную для сравнительного эксперимента структуру: объект занимает основную часть кадра, фон остается сравнительно однородным, а межклассовые различия задаются сочетанием текстурных, цветовых и морфологических признаков. Подобная конфигурация делает PlantVillage полезным стандартным набором для проверки того, как режим адаптации нейросети влияет на качество классификации в условиях высокой визуальной близости классов. Сопоставимые выводы о роли стандартизированных открытых наборов данных в задачах тонкого различения изображений приводятся в (Yang et al., 2024; Sambana et al., 2025). Фрагменты исходных изображений приведены на рис. 1.

Предобработка включала изменение размера изображений до 224×224 точек, приведение массива пикселей к типу float32 и сохранение естественного диапазона яркости от 0 до 255. Официальная документация TensorFlow для MobileNetV3Small (TensorFlow, 2024b) указывает, что при включенной встроенной предобработке модель принимает именно такой диапазон входных значений, поэтому отдельная



внешняя нормализация до интервала от -1 до 1 не выполнялась. Результат предобработки показан на рис. 2; после уменьшения разрешения классово значимые визуальные признаки, включая границы локальных изменений, цветовые переходы и особенности текстуры, сохранились в различимом виде.

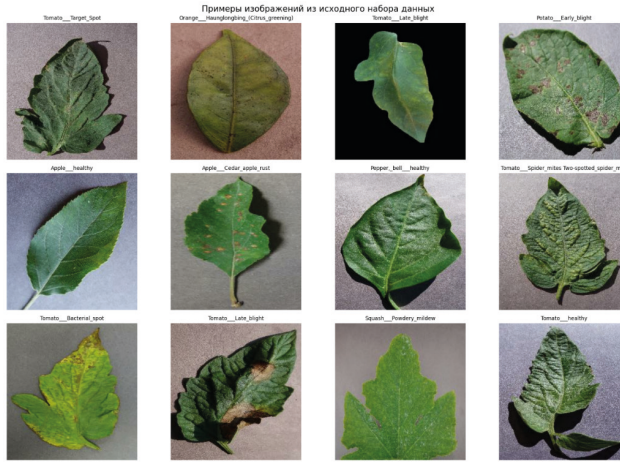


Рис. 1. Примеры исходных изображений листьев из экспериментального набора данных PlantVillage

Fig. 1. Examples of original leaf images from the experimental PlantVillage dataset

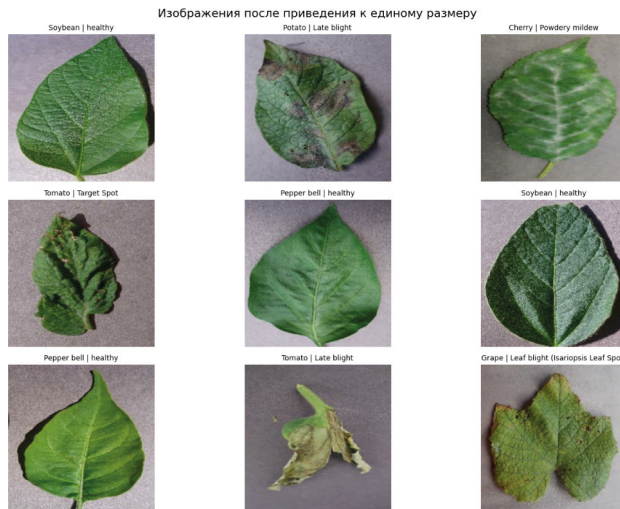


Рис. 2. Примеры изображений после предобработки и приведения к единому размеру

Fig. 2. Examples of images after preprocessing and resizing to a unified format



Разбиение на обучающую, проверочную и итоговую части проводилось однократно, в воспроизводимом порядке, с фиксированным начальным состоянием генератора случайных чисел. Контроль репрезентативности выполнялся через максимальное отклонение доли класса в каждой части выборки от доли того же класса в полном наборе данных. Величина такого отклонения вычислялась по формуле (1); для обучающей части она составила 0,131 процентного пункта, для проверочной 0,448, для итоговой 0,608, что позволяет считать распределение классов достаточно близким к исходному.

$$D_s = \max_{1 \leq c \leq C} |p_{c,s} - p_c| \quad (1)$$

где D_s — максимальное отклонение доли класса в части выборки s ; C — число классов; $p_{c,s}$ — доля класса c в части выборки s ; p_c — доля класса c в полном наборе данных.

Базовой архитектурой выбрана MobileNetV3Small (TensorFlow, 2024b), поскольку официальная реализация TensorFlow поддерживает предварительно обученные веса ImageNet и входной размер 224×224 , а вычислительная сложность модели остается умеренной для бесплатной среды Google Colab. Работы (Rahman, Islam, Islam, 2024; Yang et al., 2024; Sambana et al., 2025) показывают, что заранее обученные сверточные сети способны сохранять высокий уровень качества и при сравнительно компактной архитектуре, если режим адаптации к целевому набору данных выбран корректно.

Таблица 2 / Table 2

Параметры двух режимов адаптации заранее обученной нейросетевой модели Parameters of two adaptation regimes for a pretrained neural network model

Параметр	Режим 1: замороженная сверточная основа	Режим 2: частичное дообучение
Базовая архитектура	MobileNetV3Small	MobileNetV3Small
Предварительно обученные веса	ImageNet	ImageNet
Размер входного изображения	$224 \times 224 \times 3$	$224 \times 224 \times 3$
Изменяемая часть сети	Только классификационный блок	Классификационный блок и верхняя часть сверточной основы
Число открываемых верхних слоев сверточной основы	0	40*
Слой глобального усреднения	Да	Да
Слой пакетной нормализации	Да	Да
Слой случайного отключения части нейронов	0,30	0,30
Выходной слой	38 классов, softmax	38 классов, softmax
Алгоритм оптимизации	Adam	Adam



Параметр	Режим 1: замороженная сверточная основа	Режим 2: частичное дообучение
Скорость обучения	0,001	0,00001
Максимальное число эпох	5	5
Критерий сохранения лучшей модели	Точность на проверочной части	Точность на проверочной части
Досрочная остановка	patience = 2	patience = 2

Примечание. При частичном дообучении для адаптации рассматривались последние 40 слоев сверточной основы MobileNetV3Small. Слои пакетной нормализации внутри этой группы сохранялись в замороженном состоянии, поэтому фактически обновлялись параметры 32 слоев сверточной основы.

Использование одной и той же базовой сети в обоих экспериментальных режимах позволило исключить влияние архитектурного фактора и свести сравнение к глубине последующей настройки модели.

Поверх сверточной основы размещались слой глобального усреднения по пространственным координатам, слой пакетной нормализации, слой случайного отключения части нейронов с вероятностью 0,30 и выходной полносвязный слой на 38 классов с функцией softmax, задающей нормированное распределение вероятностей по категориям. Первый режим обучения изменял только параметры завершающего классификационного блока. Во втором режиме для адаптации открывались последние 40 слоев сверточной основы MobileNetV3Small, однако слои пакетной нормализации внутри этой группы оставались замороженными. Поэтому число фактически дообучаемых слоев сверточной основы составило 32, что позволяет сохранить стабильность статистик Batch Normalization и одновременно адаптировать верхние признаки модели к целевому набору изображений; аналогичная рекомендация приводится в руководстве TensorFlow по дообучению (Chollet, 2023).

Проверочная часть использовалась для выбора лучшей эпохи и контроля остановки обучения, а итоговая часть полностью исключалась из настройки модели и применялась только на завершающем этапе сравнения. Сохранение лучшей версии выполнялось по точности на проверочной части, досрочная остановка включалась при отсутствии улучшения в течение двух эпох. Практика поэтапного перехода от замороженной сверточной основы к аккуратному частичному дообучению регулярно применяется и в недавних работах по диагностике болезней растений (Chollet, 2023; Sambana et al., 2025).

Оценка качества строилась на общей точности классификации, точности, полноте и мере F1 с макроусреднением, а также на значении функции потерь перекрестной энтропии. Макроусреднение выбрано по причине заметной неравномерности классов: часть категорий содержала менее 50 изображений, тогда как классы Orange | Haunglongbing и Tomato | Tomato Yellow Leaf Curl Virus включали более 800 наблюдений. Дополнительно рассчитывались взвешенные варианты показателей и покласовые метрики для лучшей модели.



$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^N I\left(\hat{y}_i = y_i\right) \quad (2)$$

где Accuracy — доля правильно распознанных изображений; N — число изображений в оцениваемой выборке; y_i — истинная метка i -го изображения; \hat{y}_i — метка, предсказанная моделью; $I\left(\hat{y}_i = y_i\right)$ — индикатор правильного распознавания.

$$\text{Precision}_{\text{macro}} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (3)$$

где $\text{Precision}_{\text{macro}}$ — точность с макроусреднением; C — число классов; TP_c — число истинно положительных решений для класса c ; FP_c — число ложноположительных решений для класса c .

$$\text{Recall}_{\text{macro}} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (4)$$

где $\text{Recall}_{\text{macro}}$ — полнота с макроусреднением; FN_c — число ложноотрицательных решений для класса c ; остальные обозначения совпадают с формулой (3).

$$F1_{\text{macro}} = \frac{1}{C} \sum_{c=1}^C \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (5)$$

где $F1_{\text{macro}}$ — мера F1 с макроусреднением; Precision_c — точность для класса c ; Recall_c — полнота для класса c ; C — число классов.

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \ln p_{ic} \quad (6)$$

где L — значение функции потерь перекрестной энтропии; N — число изображений в пакете или выборке; C — число классов; y_{ic} — индикатор принадлежности i -го изображения классу c ; p_{ic} — предсказанная вероятность принадлежности i -го изображения классу c .

Матрица ошибок использовалась для интерпретации структуры неверных решений после выбора лучшей модели. Подобный разбор особенно важен в задаче распознавания листовых заболеваний, где визуально близкие симптомы, например разные виды пятнистости или ожога листа, могут давать высокую общую точность и одновременно создавать локальные зоны межклассового смешения (Yang et al., 2024; Sambana et al., 2025). Переход от общей точности к поклассовому анализу позволял рассматривать качество модели не как единое число, а как распределение успешных и проблемных случаев внутри всей таксономии заболеваний.



Результаты

Сопоставление двух режимов адаптации показало, что частичное дообучение верхней части заранее обученной сети дало более высокий результат, чем обучение одного классификационного блока. По данным табл. 3 точность на проверочной части выросла с 0,9707 до 0,9816, а значение функции потерь снизилось с 0,0929 до 0,0576. Прирост составил 1,09 процентного пункта при той же базовой архитектуре, том же наборе данных и неизменной схеме разбиения выборки. Следует учитывать, что сравнение выполнено при одном фиксированном разбиении выборки и одном воспроизводимом запуске эксперимента, поэтому полученный прирост рассматривается как практически значимый результат в рамках заданного экспериментального контура, а не как окончательная статистическая оценка устойчивости эффекта. Для более строгой проверки в дальнейшем целесообразно выполнить несколько повторных запусков с разными начальными состояниями генератора случайных чисел и рассчитать доверительные интервалы для итоговых метрик. Значения подобного порядка для контролируемых наборов изображений листьев приводятся и в недавних работах (Natarajan, Chakrabarti, Margala, 2024; Aboelenin et al., 2025), однако в них использовались другие конфигурации классов и иные экспериментальные постановки, поэтому прямое численное сопоставление допустимо только с оговорками.

Таблица 3 / Table 3

Сопоставление двух режимов переноса обучения по результатам на проверочной части Comparison of two modes of learning transfer based on the results of the testing part

Показатель	Замороженная сверточная основа	Частичное дообучение	Изменение
Общее число параметров	963 350	963 350	0
Число настраиваемых параметров	23 078	731 054	+707 976
Число ненастраиваемых параметров	940 272	232 296	-707 976
Число фактически дообучаемых слоев в сверточной основе	0	32	+32
Число верхних слоев сверточной основы, выбранных для адаптации	0	40	+40
Лучшая эпоха	4	5	+1
Точность на проверочной части	0,9707	0,9816	+1,09 п.п.
Функция потерь на проверочной части	0,0929	0,0576	-0,0353

Примечание: Для обоих режимов использовалась архитектура MobileNetV3Small; различался только способ адаптации заранее обученной сверточной основы. Во втором режиме для адаптации была выбрана верхняя группа из 40 слоев, однако слои пакетной нормализации не обновлялись, поэтому фактически дообучались 32 слоя сверточной основы.



Рост качества по эпохам имел различный характер. Режим с замороженной сверточной основой быстро вышел на плато: основная часть прироста пришлась на первые две эпохи, после чего улучшение стало минимальным. Частичное дообучение сохраняло поступательное движение до пятой эпохи, а кривая функции потерь убывала более ровно. Графики обучения, приведенные на рис. 3 и рис. 4, показывают отсутствие заметного расхождения между обучающей и проверочной частями, что указывает на устойчивую настройку модели в пределах выбранного экспериментального контура.

Поклассовый расчет для лучшей модели позволил оценить качество уже на независимой итоговой части выборки через показатели, заданные формулами (2) — (6). Расчет по данным итоговой классификации дал общую точность около 0,9843, макроусредненную меру F1 около 0,9792 и взвешенную меру F1 около 0,9842. Разница между макроусредненными и взвешенными значениями оказалась небольшой, хотя более низкая макроусредненная полнота отразила наличие нескольких классов, в которых распознавание шло заметно труднее. Обзор (Salka et al., 2025) связывает подобные локальные просадки с высокой визуальной близостью классов, неравномерностью выборки и особенностями стандартизированных наборов изображений.

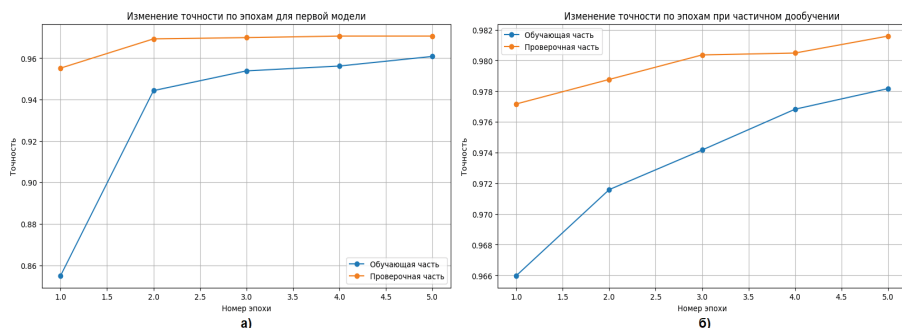


Рис. 3. Динамика точности на обучающей и проверочной частях данных при двух режимах переноса обучения: а — замороженная сверточная основа; б — частичное дообучение

Fig. 3. Accuracy dynamics on the training and validation datasets for two transfer learning modes: a — frozen convolutional base; b — partial retraining

Распределение поклассовых показателей оказалось содержательно неоднородным. Наиболее высокие значения меры F1 зафиксированы у классов Squash | Powdery mildew, Blueberry | healthy, Apple | Black rot, Grape | healthy и Orange | Haunglongbing, где распознавание приближалось к безошибочному уровню. Проблемная зона сместилась к группам с высокой визуальной близостью: минимальные значения меры F1 наблюдались у классов Corn | Cercospora leaf spot Gray leaf spot, Tomato | Early blight, Potato | healthy, Corn | Northern Leaf Blight и Tomato | Target Spot. С методической точки зрения здесь важен сам факт локального снижения качества в трудных зонах пространства признаков.

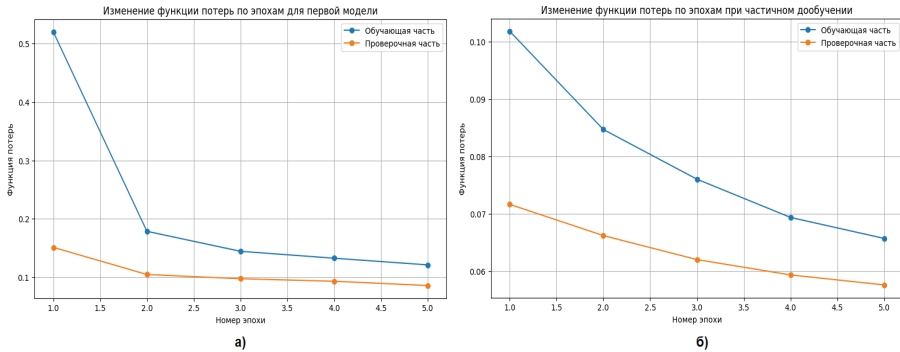


Рис. 4. Динамика функции потерь на обучающей и проверочной частях данных при двух режимах переноса обучения: а — замороженная сверточная основа; б — частичное дообучение

Fig. 4. Loss dynamics on the training and validation datasets for two transfer learning modes: a — frozen convolutional backbone; b — partial fine-tuning

Таблица 4 / Table 4

Показатели лучшей модели на итоговой части выборки
Performance of the best model on the final part of the sample

Показатель	Значение
Общая точность классификации	0,9843
Точность с макроусреднением	0,9812
Полнота с макроусреднением	0,9776
Мера F1 с макроусреднением	0,9792
Точность с взвешенным усреднением	0,9843
Полнота с взвешенным усреднением	0,9843
Мера F1 с взвешенным усреднением	0,9842

Таблица 5 / Table 5

Наиболее и наименее устойчиво распознаваемые классы лучшей модели
Most and least reliably recognized classes of the best model

Наиболее устойчивые категории	Число изображений	Мера F1	Наименее устойчивые категории	Число изображений	Мера F1
Squash Powdery mildew	281	1,0000	Corn Cercospora leaf spot Gray leaf spot	86	0,8889
Blueberry healthy	213	1,0000	Tomato Early blight	145	0,9078
Apple Black rot	107	1,0000	Potato healthy	22	0,9302
Grape healthy	58	1,0000	Corn Northern Leaf Blight	153	0,9363
Orange Haunglongbing (Citrus greening)	832	0,9994	Tomato Target Spot	226	0,9417



Нормированная матрица ошибок лучшей модели представлена на рис. 5. Главная диагональ сохраняет высокую насыщенность почти по всей совокупности классов, однако локальные смещения отчетливо видны у групп с морфологически и текстурно близкими признаками. Наиболее заметный пример связан с классами Corn | Cercospora leaf spot Gray leaf spot и Corn | Northern Leaf Blight. В группе томатных классов основное число ошибок сосредоточилось между Early blight, Late blight, Septoria leaf spot, Target Spot и Spider mites Two-spotted spider mite. Работа (Salka et al., 2025) подчеркивает, что именно сходство локальных визуальных паттернов часто становится основной причиной межклассового смешения в задачах тонкой классификации изображений.

Картина ошибок в табл. 6 показывает, что модель путала прежде всего близкие классы внутри одной морфологически согласованной группы. На первом месте оказалось смешение Corn | Cercospora leaf spot Gray leaf spot с Corn | Northern Leaf Blight, давшее 14 ошибок, или 16,279 % внутри истинного класса. Следом шли пары Tomato | Target Spot и Tomato | Spider mites Two-spotted spider mite, Tomato | Late blight и Tomato | Early blight, Tomato | Early blight и Tomato | Septoria leaf spot. Подобная структура важна для общей интерпретации результатов, поскольку указывает на содержательные зоны трудности в пространстве визуальных признаков, а не на случайный характер ошибок.

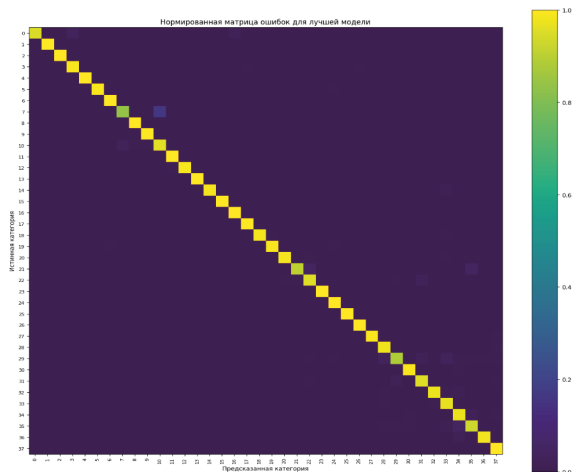


Рис. 5. Нормированная матрица ошибок лучшей модели на итоговой части выборки

Fig. 5. Normalized error matrix of the best model on the final part of the sample



Таблица 6 / Table 6

**Наиболее частые межклассовые смешения
у лучшей модели в зонах визуальной близости**

Most frequent inter-class confusions of the best model in zones of visual proximity

Истинная категория	Предсказанная категория	Количество ошибок	Доля ошибок внутри истинной категории, %
Corn Cercospora leaf spot Gray leaf spot	Corn Northern Leaf Blight	14	16,279
Tomato Target Spot	Tomato Spider mites Two-spotted spider mite	10	4,425
Tomato Spider mites Two-spotted spider mite	Tomato Target Spot	5	2,000
Tomato Late blight	Tomato Early blight	5	1,779
Tomato Early blight	Tomato Bacterial spot	4	2,759
Tomato Early blight	Tomato Septoria leaf spot	4	2,759
Corn Northern Leaf Blight	Corn Cercospora leaf spot Gray leaf spot	4	2,614
Potato Late blight	Tomato Late blight	4	2,614
Apple Apple scab	Apple healthy	3	3,125
Tomato Early blight	Tomato Late blight	3	2,069

Обсуждение результатов

Прирост точности на 1,09 процентного пункта при переходе от замороженной сверточной основы к частичному дообучению имеет методическое значение, поскольку сравнение проводилось на одной архитектуре, одном наборе изображений и при одинаковой схеме разбиения данных. Полученный выигрыш указывает на преимущество более глубокой адаптации модели к целевой задаче в пределах проведенного эксперимента. При этом результат не следует трактовать как универсальное доказательство превосходства fine-tuning для всех наборов изображений и архитектур: устойчивость выявленного эффекта требует проверки на нескольких случайных инициализациях, альтернативных разбиениях и внешних данных. В статье (Dong et al., 2024) влияние режима дообучения также рассматривается как отдельный фактор надежности распознавания, а в работе (Ifikhar et al., 2024) донастройка признакового ядра связывается с улучшением качества на сложных изображениях.

Характер обучения в двух режимах оказался не менее важен, чем итоговые числа. Замороженная сверточная основа быстро вышла на плато, что обычно наблюдается в ситуациях, когда заранее обученные признаки дают сильный старт, однако



перестают достаточно тонко различать близкие категории целевого набора. Частичное дообучение сохранило рост качества до конца обучения, и подобная динамика хорошо согласуется с выводами (Dong et al., 2024), где выбор парадигмы донастройки заметно менял качество классификации в тонких межклассовых различиях. Обзор (Shoaiб et al., 2025) также связывает успешность переноса обучения с глубиной адаптации модели к специфике целевого визуального материала.

Распределение ошибок показывает, что слабые места модели сосредоточены в группах с близкой морфологией, текстурой и цветовым рисунком. Наиболее проблемными оказались пары классов, где совпадают локальная структура изменения поверхности, конфигурация очагов и характер цветовых переходов. Обзорные работы (Shoaiб et al., 2025; Shafay et al., 2025) описывают ту же закономерность: наилучшие показатели обычно достигаются на контрастных и хорошо отделимых категориях, тогда как сходные по визуальному паттерну классы продолжают создавать зоны межклассового смешения даже у сильных моделей. С методической точки зрения полученная матрица ошибок подтверждает, что эксперимент действительно проверял режимы адаптации в трудной задаче тонкой классификации.

Интерпретация полученных значений требует аккуратности из-за свойств самого набора PlantVillage. Контролируемый фон, крупный план объекта и относительно чистая визуальная сцена упрощают задачу по сравнению с более сложными сценариями распознавания, где присутствуют тени, перекрытия, неоднородное освещение и естественные помехи. В работе (Salman, Muhammad, Nan, 2025) классификация в реальной съемке разбирается именно как случай доменного сдвига, то есть изменения распределения изображений между учебной и практической средой, а обзор (Shafay et al., 2025) рассматривает подобный разрыв как одно из центральных ограничений современной литературы по автоматическому анализу изображений. По этой причине полученный результат корректнее трактовать как доказательство эффективности выбранного режима адаптации в стандартизированной постановке.

Закрытый характер классификации задает еще одну границу применимости проведенного эксперимента. Обученная модель выбирала один из заранее известных классов, тогда как в более общей практике анализа изображений встречаются редкие, смешанные и ранее не представленные в обучении состояния. Исследование (Dong et al., 2024) показывает, что распознавание неизвестных классов требует специальной постановки задачи и отдельной проверки устойчивости, а работа (Salman, Muhammad, Nan, 2025) демонстрирует, насколько сильно меняется поведение модели при переходе к данным из естественной среды. Следующий шаг здесь связан с проверкой лучшего режима дообучения на более сложных внешних наборах данных и с оценкой способности системы корректно выделять случаи, для которых уверенное отнесение к известным классам недопустимо.

Заключение

Частичное дообучение верхней части заранее обученной сверточной сети обеспечило более высокое качество классификации, чем режим, в котором изменялся



только завершающий классификационный блок. При одинаковой архитектуре, едином наборе данных и фиксированном разбиении выборки точность на проверочной части возросла с 0,9707 до 0,9816, а значение функции потерь снизилось с 0,0929 до 0,0576. Поклассовый анализ лучшей модели показал, что наибольшая устойчивость достигается у хорошо различимых классов, тогда как основные ошибки концентрируются в группах с высокой визуальной близостью.

Научная новизна исследования связана с контролируемым экспериментальным сопоставлением двух режимов адаптации MobileNetV3Small в задаче классификации визуально близких классов набора PlantVillage. В работе не утверждается принципиальная новизна частичного дообучения как метода, поскольку fine-tuning является распространенной практикой применения заранее обученных сверточных сетей. Новым для данной постановки является показанное на единой экспериментальной схеме преимущество частичного дообучения верхней части признакового ядра над режимом извлечения признаков без настройки сверточной основы. Полученный результат важен не только по величине общей точности, но и по структуре ошибок: улучшение проявилось на независимой итоговой части выборки и сопровождалось содержательно объяснимым распределением смещений в трудных зонах пространства признаков. Обзоры последних лет (Upadhyay et al., 2025) и работы по расширению сложных наборов изображений (Zhang et al., 2025) подтверждают, что вопрос обобщающей способности модели и переносимости результата на более сложные данные остается центральным, поэтому вывод о предпочтительности частичного дообучения имеет методическую ценность для построения аналогичных систем анализа изображений.

Практическая значимость работы определяется возможностью использовать предложенную схему адаптации при проектировании систем автоматической классификации изображений, где требуется различать близкие визуальные категории в условиях ограниченного объема целевых данных. Эксперимент на наборе PlantVillage показывает, что даже компактная заранее обученная архитектура способна обеспечить высокий уровень качества, если режим ее последующей настройки выбран корректно. Наиболее осторожного применения требуют группы классов со сходной морфологией и текстурой, поскольку именно там сохраняется основная доля межклассовых смещений.

Ограничения исследования связаны с тремя обстоятельствами. Во-первых, итоговая проверка выполнялась на стандартизированном наборе изображений PlantVillage, где фон, масштаб объекта и условия съемки заметно упрощают задачу по сравнению с более сложной визуальной средой. Во-вторых, эксперимент проводился при одном фиксированном разбиении выборки и одном воспроизводимом запуске модели, поэтому величина прироста качества требует дополнительной проверки на повторных запусках с разными начальными состояниями генератора случайных чисел. В-третьих, число изображений в отдельных тестовых классах оставалось небольшим, из-за чего поклассовые оценки для редких категорий получали более дискретный характер и сильнее зависели от единичных ошибок. Расширение пула изображений для редких классов, расчет доверительных интервалов для метрик



и проверка модели на дополнительных независимых источниках данных рассматриваются как первое направление продолжения работы. Второе направление связано с переносимостью результата на более сложные условия съемки: при переходе к данным с тенями, перекрытиями, неоднородным освещением и естественными помехами потребуется повторная проверка устойчивости модели, а при изменении состава классов и структуры входного потока — регулярная переоценка качества на новых данных (Upadhyay et al., 2025; Zhang et al., 2025).

Список источников / References

1. Aboelenin, S., Elbasheer, F.A., Eltoukhy, M.M., El-Hady, W.M., Hosny, K.M. (2025). A hybrid Framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. *Complex & Intelligent Systems*, 11, Article 142. <https://doi.org/10.1007/s40747-024-01764-x>
2. Chollet, F. (2023). Transfer learning & fine-tuning. TensorFlow Core. URL: https://www.tensorflow.org/guide/keras/transfer_learning(дата обращения: 28.03.2026).
3. Dong, J., Fuentes, A., Zhou, H., Jeong, Y., Yoon, S. et al. (2024). The impact of fine-tuning paradigms on unknown plant disease recognition. *Scientific Reports*, 14, Article 17900. <https://doi.org/10.1038/s41598-024-66958-2>
4. Iftikhar, M., Kandhro, I.A., Kausar, N. et al. (2024). Plant disease management: a fine-tuned enhanced CNN approach with mobile app integration for early detection and classification. *Artificial Intelligence Review*, 57, Article 167. <https://doi.org/10.1007/s10462-024-10809-z>
5. Natarajan, S., Chakrabarti, P., Margala, M. (2024). Robust diagnosis and meta visualizations of plant diseases through deep neural architecture with explainable AI. *Scientific Reports*, 14, Article 13695. URL: <https://doi.org/10.1038/s41598-024-64601-8>
6. Pacal, I., Kunduracioglu, I., Alma, M.H. [et al.]. A systematic review of deep learning techniques for plant diseases // *Artificial Intelligence Review*. 2024. Vol. 57. Art. 304. DOI: 10.1007/s10462-024-10944-7
7. Rahman, M.M., Islam, M.S., Islam, M.N. (2024). Classification of Various Plant Leaf Disease Using Pretrained Convolutional Neural Network On Imagenet. *The Open Agriculture Journal*, 18. <https://doi.org/10.2174/0118743315305194240408034912>
8. Ramanjot, Mittal U., Wadhawan, A. [et al.]. Plant Disease Detection and Classification: A Systematic Literature Review // *Sensors*. 2023. Vol. 23, No. 10. Art. 4769. DOI: 10.3390/s23104769
9. Richter, D.J., Kim, K. Assessing the performance of domain-specific models for plant leaf disease classification: a comprehensive benchmark of transfer-learning on open datasets // *Scientific Reports*. 2025. Vol. 15. Art. 18973. DOI: 10.1038/s41598-025-03235-w
10. Salka, T.D., Hanafi, M.B., Rahman, S.M.S.A., Zulperi, D.B.M., Omar, Z. (2025). Plant leaf disease detection and classification using convolution neural networks model: a review. *Artificial Intelligence Review*, 58, Article 322. <https://doi.org/10.1007/s10462-025-11234-6>
11. Salman, Z., Muhammad, A., Han, D. (2025). Plant disease classification in the wild using vision transformers and mixture of experts. *Frontiers in Plant Science*, 16, Article 1522985. <https://doi.org/10.3389/fpls.2025.1522985>



12. Sambana, B., Nnadi, H.S., Wajid, M.A., Fidelia, N.O., Camacho-Zuñiga, C., Ajuzie, H.D., Onyema, E. M. (2025). An efficient plant disease detection using transfer learning approach. *Scientific Reports*, 15, Article 19082. <https://doi.org/10.1038/s41598-025-02271-w>
13. Shafay, M., Hassan, T., Owais, M., Hussain, I., Khawaja, S.G. et al. (2025). Recent advances in plant disease detection: challenges and opportunities. *Plant Methods*, 21, Article 140. <https://doi.org/10.1186/s13007-025-01450-0>
14. Shafik W., Tufail A., De Silva Liyanage, C. [et al.]. Using transfer learning-based plant disease classification and detection for sustainable agriculture // *BMC Plant Biology*. 2024. Vol. 24. Art. 136. DOI: 10.1186/s12870-024-04825-y
15. Shoaib, M., Sadeghi-Niaraki, A., Ali, F., Hussain, I., Khalid, S. (2025). Leveraging deep learning for plant disease and pest detection: a comprehensive review and future directions. *Frontiers in Plant Science*, 16, Article 1538163. <https://doi.org/10.3389/fpls.2025.1538163>
16. TensorFlow. (2024). `plant_village`. TensorFlow Datasets. URL: https://www.tensorflow.org/datasets/catalog/plant_village(дата обращения: 28.03.2026).
17. TensorFlow. (2024). `tf.keras.applications.MobileNetV3Small`. TensorFlow v2.16.1 API Documentation. URL: https://www.tensorflow.org/api_docs/python/tf/keras/applications/MobileNetV3Small(дата обращения: 28.03.2026).
18. Upadhyay, A., Chandel, N.S., Singh, K.P., Chakraborty, S.K. Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review*. 2025. Vol. 58. Art. 92. URL: <https://doi.org/10.1007/s10462-024-11100-x>
19. Yang, B., Li, M., Li, F., Wang, Y., Liang, Q., Zhao, R., Li, C., Wang, J. (2024). A novel plant type, leaf disease and severity identification framework using CNN and transformer with multi-label method. *Scientific Reports*, 14, Article 11664. <https://doi.org/10.1038/s41598-024-62452-x>
20. Zhang, H.-W., Wang, R.-F., Wang, Z., Su, W.-H. DLCPD-25: A Large-Scale and Diverse Dataset for Crop Disease and Pest Recognition. *Sensors*. 2025. Vol. 25, No. 22. Art. 7098. URL: <https://doi.org/10.3390/s25227098>
21. Zhao, J., Xu, L., Ma, Z. [et al.]. A review of plant leaf disease identification by deep learning algorithms // *Frontiers in Plant Science*. 2025. Vol. 16. Art. 1637241. DOI: 10.3389/fpls.2025.1637241

Информация об авторах

Мутхана Али Салем Али, аспирант, Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСИС»» (НИТУ МИСИС), ORCID: <https://orcid.org/0000-0003-4304-7469>, e-mail: adamdam265@gmail.com

Елена Вячеславовна Ляпунцова, доктор технических наук, профессор кафедры «Инновационное предпринимательство» Московский государственный технический университет имени Николая Эрнестовича Баумана (ФГАОУ ВО «МГТУ им. Н.Э. Баумана»), Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-3420-3805>, e-mail: lev86@bmstu.ru



Information about the authors

Muthana Ali Salem Ali, graduate student, National University of Science and Technology MISIS (NUST MISIS), ORCID: <https://orcid.org/0000-0003-4304-7469>, e-mail: adamdam265@gmail.com

Elena Vyacheslavovna Lyapunsova, Doctor of Technical Sciences, Professor of the Department of Innovative Entrepreneurship, Bauman Moscow State Technical University (FSAEI HE Bauman Moscow State Technical University), Moscow, Russian Federation, ORCID: <https://orcid.org/0000-0002-3420-3805>, e-mail: lev86@bmstu.ru

Вклад авторов

Все авторы приняли участие в обсуждении результатов и согласовали окончательный текст рукописи.

Contribution of the authors

All authors participated in the discussion of the results and approved the final text of the manuscript.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest

The authors declare no conflict of interest.

Поступила в редакцию 30.03.2026

Поступила после рецензирования 11.05.2026

Принята к публикации 20.05.2026

Опубликована 30.06.2026

Received 2026.03.30

Revised 2026.05.11

Accepted 2026.05.20

Published 2026.06.30