



Прикладные проблемы оптимизации

<https://doi.org/10.25559/SITITO.021.202503.411-419>

УДК 004.81

Экспериментальный анализ моделей трансферного обучения для классификации изображений

Т. А. Самойлова

Оригинальная статья

ФГБОУ ВО «Смоленский государственный университет», г. Смоленск, Российская Федерация

Адрес: 214000, Российская Федерация, г. Смоленск, ул. Пржевальского, д. 4

tatsamoilova2020@gmail.com

Аннотация

Настоящее время характеризуется значительным повышением производительности в задачах классификации изображений, обусловленным использованием методов глубокого обучения. Но, как правило, эти методы требуют больших наборов аннотированных данных, что затрудняет их применение при небольших наборах. Многообещающей альтернативой преодолению проблемы стал метод трансферного обучения, используемый для оптимизации моделей глубокого обучения. Сущность оптимизации состоит в адаптации предварительно обученной нейронной сети к новой задаче средствами небольшого набора данных конкретной предметной области. Цель данной работы – сравнение производительности моделей трансферного обучения в задачах классификации изображений для разных способов их оптимизации. Выполнена квалитетическая оценка трех отобранных моделей на двух целевых наборах данных – простом и сложном. Первая модель – MobileNetV3 от компании Google, предназначенная для устройств с ограниченными ресурсами, адаптировалась простым методом извлечения признаков. Вторая – ResNet-50, представленная компанией Microsoft Research, адаптировалась методом тонкой настройки параметров в модель под названием BigTransfer. Третья – гибридная модель, адаптировалась методом извлечения признаков, объединяет предварительно обученную ResNet-50 с преобразователем компьютерного зрения Vision Transformer для сложных изображений целевого набора. Работа содержит описание экспериментов, где построены графики производительности указанных моделей. С целью измерения производительности использовались такие оценочные метрики как точность, потери и время обучения. Результаты экспериментов выявили, что для сложного набора самой точной была гибридная модель, имеющая при этом наибольшее время обучения. Простой набор продемонстрировал лучшие значения точности в модели Big Transfer. MobileNetV3 имела схожую с ней точность, но меньшее время обучения и большую простоту использования. Проведенное исследование показало, что выбор метода оптимизации модели трансферного обучения в задачах классификации изображений на заданных целевых наборах – это компромисс между достижением высокой точности и малым временем обучения.

Ключевые слова: трансферное обучение, оптимизация моделей, классификация изображений, MobileNetV3, ResNet-50, BigTransfer, Vision Transformer

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Для цитирования: Самойлова Т. А. Экспериментальный анализ моделей трансферного обучения для классификации изображений // Современные информационные технологии и ИТ-образование. 2025. Т. 21, № 3. С. 411-419. <https://doi.org/10.25559/SITITO.021.202503.411-419>

© Самойлова Т. А., 2025



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Applied Optimization Problems

Experimental Analysis of Transfer Learning Models for Image Classification

T. A. Samoilova

Original article

Smolensk State University, Smolensk, Russian Federation

Address: 4 Przhevalsky St., Smolensk 214000, Russian Federation

tatsamoilova2020@gmail.com

Abstract

Currently, there is a significant increase in performance in image classification tasks due to the use of deep learning methods. However, these methods typically require large sets of annotated data, which makes them difficult to use for small sets. A promising alternative to overcome this problem is the transfer learning method used to optimize deep learning models. The essence of optimization is to adapt a pre-trained neural network to a new task using a small set of data from a specific subject area. The purpose of this paper is to compare the performance of transfer learning models in image classification tasks for different optimization methods. A qualitative evaluation of three selected models was performed on two target datasets - simple and complex. The first model, MobileNetV3 from Google, designed for devices with limited resources, was adapted using a parameter fine-tuning method into a model called BigTransfer. The second model, ResNet-50, presented by Microsoft Research, was adapted using a parameter fine-tuning method into a model called BigTransfer. The third is a hybrid model adapted using feature extraction, which combines a pre-trained ResNet-50 with the Vision Transformer for complex target images. The paper contains a description of the experiments where the performance graphs of these models are constructed. The performance metrics used to measure the performance were accuracy, loss, and training time. The results of the experiments revealed that for the complex set, the hybrid model was the most accurate, while having the longest training time. For the simple set, the best accuracy values were demonstrated by the Big Transfer model. MobileNetV3 had similar accuracy, but shorter training time and greater ease of use. The study showed that the choice of an optimization method for a transfer learning model in image classification problems on given target sets is a compromise between achieving high accuracy and short training time.

Keywords: transfer learning, model optimization, image classification, MobileNetV3, ResNet-50, BigTransfer, Vision Transformer

Conflict of interests: The author declares no conflict of interests.

For citation: Samoilova T.A. Experimental Analysis of Transfer Learning Models for Image Classification. *Modern Information Technologies and IT-Education*. 2025;21(3):411-419. <https://doi.org/10.25559/SITITO.021.202503.411-419>



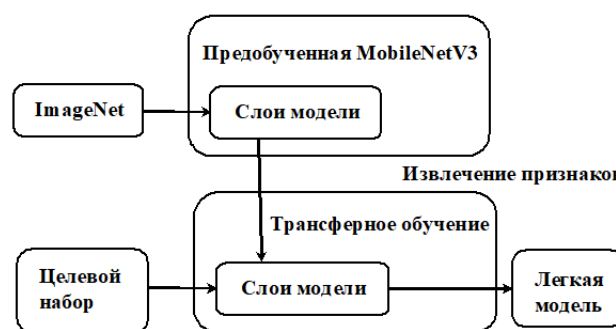
1 Введение

Большинство современных подходов к классификации изображений основано на обучении глубоких сверточных нейронных сетей (*Convolutional Neural Network, CNN*). Этот процесс требует значительного набора аннотированных данных, что затрудняет применение CNN для данных небольшого объема. Альтернативой преодолению этих проблем стали стратегии трансферного обучения [1, 2], сущность которых состоит в оптимизации готовой, предварительно обученной сети средствами целевых данных – изображений малого объема конкретной предметной области. Наиболее востребованы эти стратегии в таких областях, как медицина, биология, сельское хозяйство, распознавание дорожных знаков [3–6]. Существуют разные подходы к оптимизации моделей классификации изображений, наиболее распространенные стратегии – извлечение признаков [7] и тонкая настройка. В случае извлечения признаков предварительно обученная модель используется только с целью извлечения признаков, представляющих полную информацию об обучающих изображениях. Эти предварительно обученные слои, содержащие особенности миллионов входных изображений, замораживаются, а последние классификационные слои заменяются новыми, которые дообучаются на малом наборе данных целевой задачи. В этом подходе предобученная модель сохраняется неизменной и её веса не обновляются во время обучения классификатора. Тонкая настройка начинается, как и в случае извлечения признаков, с замораживания слоев предварительно обученной модели, но затем некоторые из них, обычно верхние, размораживаются, что позволяет им переобучаться в соответствии с небольшим входным набором изображений. Таким образом, модель адаптируется, обновляя только конкретные признаки, связанные с новой задачей. Сущность этого подхода состоит в том, чтобы использовать знания, представленные весами модели, полученные в первом процессе обучения, в качестве отправной точки дообучения в целевой задаче. Некоторые исследования [8], предлагают использовать тонкую настройку для извлечения признаков, показывая, что тонкая настройка достигает в этом случае более высокой производительности. Но есть результаты [9], демонстрирующие, что простое извлечение признаков достигает сопоставимой производительности, требуя меньше вычислительных ресурсов для обучения. Цель настоящей работы – сравнить и оценить производительность предварительно обученных моделей для разных способов их оптимизации, направленных на эффективную адаптацию моделей к новой области. Извлечение признаков выполняется с адаптацией двух моделей: легкой MobileNet V3, предобученной на данных относительно небольшого объема, и гибридной ResNet50 + Vision Transformers (ViT)¹. Тонкая

настройка – для предобученной на большом наборе модели BigTransfer [10]. Полученные результаты обсуждаются с использованием квалитетических оценок качества: точности, потерь и времени адаптации. Целевые данные для оптимизации – две выборки изображений разного характера из интернета: простая состоит из 3670 изображений растений 5 классов, сложная включает 9518 изображений больших листьев 4 классов.

2 Модели экспериментов

С целью проведения экспериментов выбраны три предварительно обученные модели глубокого обучения разной архитектуры: легкая, тяжелая и гибридная. Первая – легкая модель CNN MobileNetV3, её архитектура обеспечивает эффективное развертывание на мобильных устройствах с ограниченными вычислительными ресурсами [11], имея при этом конкурентоспособную точность по сравнению с более крупными моделями. Небольшой размер модели обеспечивает быстрое время работы, что делает её подходящей в приложениях реального времени. Исследователи Google разработали MobileNetV3 как усовершенствование оригинальной модели MobileNet, она быстрее и точнее предыдущей версии MobileNetV2 именно в задачах классификации. MobileNetV3 предварительно обучена на ImageNet-1k [12] – подмножестве ImageNet, состоящем 1,4 млн изображений, и может классифицировать изображения по 1000 категориям объектов, таких как растения, фрукты, овощи, одежда, животные и пр. В эксперименте эта сеть оптимизируется методом извлечения признаков, при этом ее слои копируются вместе с параметрами, а классификационные слои дообучаются на небольшом целевом наборе.



Р и с. 1. Создание легкой модели извлечения признаков
F i g. 1. Creation of a lightweight feature extraction model

Источник: здесь и далее в статье все рисунки составлены автором.
Source: Hereinafter in this article all figures were drawn up by the author.

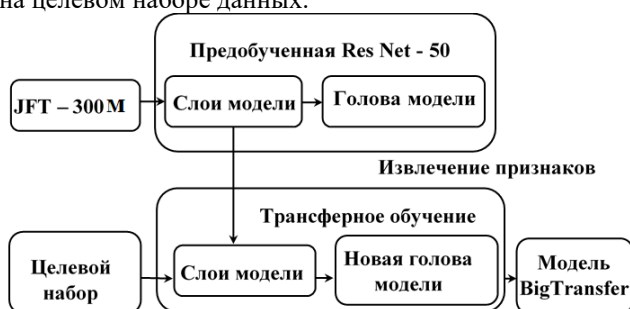
На рисунке 1 показано, как получив извлеченные признаки из предварительно обученной модели, создается новая легкая модель.

Вторая крупная модель ResNet-50, предварительно

¹ mobilenet_v3 [Электронный ресурс] // Google, 2025. URL: <https://www.kaggle.com/models/google/mobilenet-v3> (дата обращения: 21.07.2025).

21.07.2025); resnet-50 [Электронный ресурс] // Google, 2025. URL: <https://www.kaggle.com/models/tensorflow/resnet-50> (дата обращения: 21.07.2025); vision_transformer [Электронный ресурс] // Google, 2025. URL: <https://www.kaggle.com/models/spsayakpaul/vision-transformer/tensorFlow2/vit-r50-l32-fe/1?tfhub-redirect=true> (дата обращения: 21.07.2025).

обучена на большом наборе JFT 300M, содержащим более 300 млн изображений, носит название Big Transfer (BiT). Она принадлежит к семейству ResNet (*Residual Networks*) и хорошо известна своей надёжностью и высокой точностью в задачах классификации изображений [13]. Архитектура ResNet50 включает 48 сверточных слоев, каждый из которых содержит признаки входных изображений. Эта архитектура позволяет модели различать тонкие текстурные и цветовые различия, что особенно необходимо для таких задач, как классификация растительности или обнаружение лесных пожаров, где окружающая среда изменчива. Ключевая особенность ResNet-50 – использование остаточных связей, известных как короткие пути. Короткие пути позволяют ей передавать важную информацию не через все слои, а только выборочно, что делает процесс обучения более быстрым и надежным. Недавнее исследование [14] обнаружили, что сильное предварительное обучение на больших наборах данных, таких как JFT-300M [15] может превзойти существующие современные модели классификации изображений. Главное в создании целевой модели Big Transfer – тонкая настройка весов предобученной модели. На рисунке 2 показан процесс тонкой настройки целевой модели трансферного обучения. Подобно извлечению признаков, загружаются предварительно обученные слои модели и слои классификации. Размораживается только последний слой – голова модели, что позволяет обновлять её веса во время дообучения. Созданная таким образом модель дообучается (квалификационные слои и голова модели) на целевом наборе данных.



Р и с. 2. Создание модели Big Transfer
F i g. 2. Creation of the Big Transfer model

Гиперпараметры Big Transfer, такие как скорость обучения, размер пакета, оптимизатор нахождения оптимального набора параметров модели, определяются экспериментальным путем. Python – пакет `keras.optimizers.schedules` предоставляет классы определения графиков скорости обучения, позволяя динамически изменять скорость во время обучения. В эксперименте используются средства указанного пакета с целью корректировки скорости обучения с течением времени, что улучшает сходимость и производительность модели. При наличии достаточных вычислительных ресурсов возможна автоматическая оптимизация гиперпараметров [16].

Третья – гибридная модель CNN + ViT объединяет предварительно обученную CNN ResNet-50 и преобразователь зрения Vision Transformers (ViT). ViT

был предложен в октябре 2020 года [17] как конкурентоспособная альтернатива CNN в задачах машинного зрения, включая классификацию изображений. При традиционном обучении ViT разлагает каждое входное изображение обучающего набора на ряд неперекрывающихся прямоугольных фрагментов (патчей). Затем каждый патч сериализуется в вектор и отображается в меньшее измерение с помощью умножения матриц. После преобразования патчей в векторное представление, они проходят через несколько слоев трансформера – кодера, в каждом из которых применяется механизм внутреннего внимания для выявления взаимосвязей между патчами.



Р и с. 3. Создание гибридной модели CNN + ViT
F i g. 3. Creation of a hybrid CNN + ViT model

Эти извлеченные трансформером – кодером связи преобразуются в прогнозы классов. В трансферном обучении, с целью создания гибридной модели эксперимента (рисунок 3), ViT предварительно настраивается методом извлечения признаков из предобученной CNN ResNet-50. Извлеченные признаки напрямую поступают на вход модуля трансформера. Целевой набор изображений, размером 224×224 , преобразуется в патчи 32×32 , которые также подаются на вход трансформера, выполняющего операцию внутреннего внимания. Механизм внимания позволяет трансформеру в ходе дообучения обращать внимание на разные части изображений целевого набора и находить наиболее важные. Выходные данные трансформера пропускаются через многослойный перцептрон для классификации.

Трансформер – кодировщик [18], использующий механизмы самовнимания, в отличие от CNN, хорошо обрабатывает сложные данные, где изображения мелкие, имеют пятна, неровные края и содержатся в небольшом пространстве с разным освещением. Подобный набор сложных данных выбран с целью дообучения гибридной модели, которая, сочетая достоинства ViT и CNN, в последнее время стала применяться в диагностике болезней (медицина, сельское хозяйство) [19, 20], где необходима классификация сложных данных.

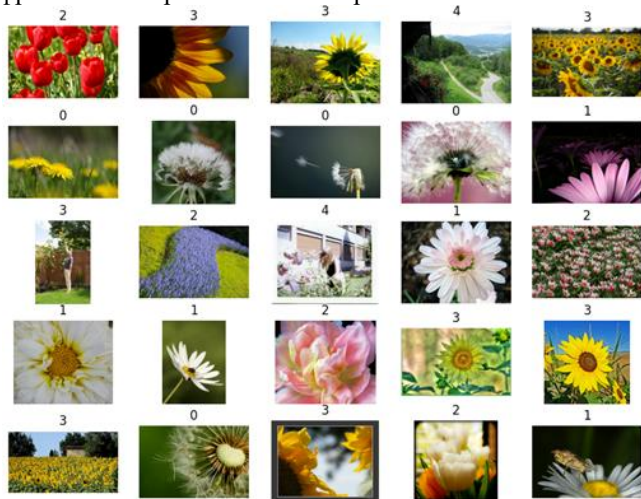
Модели для проведения эксперимента отобраны из репозитория TensorFlow Hub, предоставляющего библиотеку предварительно обученных моделей трансферного обучения².

² TensorFlow Hub is a repository of trained machine learning models [Электронный ресурс] // Google, 2025. URL: <https://www.tensorflow.org/hub> (дата обращения: 21.07.2025).



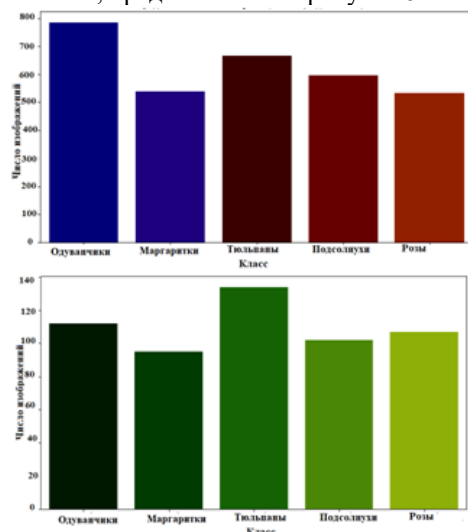
3 Данные экспериментов

Эксперименты с легкой моделью, основанной на CNN MobileNetV3, и большой моделью BigTransfer выполнены на целевом наборе «Цветы»³. Он взят из лицензированной Python – библиотеки tensorflow-datasets под именем 'tf_flowers' и состоит из 3670 изображений цветов с 5 возможными метками классов: 0-одуванчики, 1-маргаритки, 2-тюльпаны, 3-подсолнухи и 4-розы. На рисунке 4 представлен фрагмент изображений из набора:



Р и с. 4. Фрагмент изображений из целевого набора «Цветы»
F i g. 4. Fragment of images from the target set "Flowers"

Данные разделяются на обучающие и тестовые в пропорции 85% и 15% соответственно. Модель дообучается на обучающих данных, а затем выполняется оценка, насколько хорошо модель работает на данных, которые она никогда не видела – тестовом наборе. Столбчатые диаграммы, соответствующие числу обучающих и тестовых данных каждого класса, представлены на рисунке 5.



Р и с. 5. Диаграммы числа обучающих и тестовых данных набора «Цветы»

F i g. 5. Diagrams of the number of training and test data in the "Flowers" set

³ tf_flowers [Электронный ресурс] // Google, 2025. URL: https://www.tensorflow.org/datasets/catalog/tf_flowers (дата обращения: 21.07.2025).

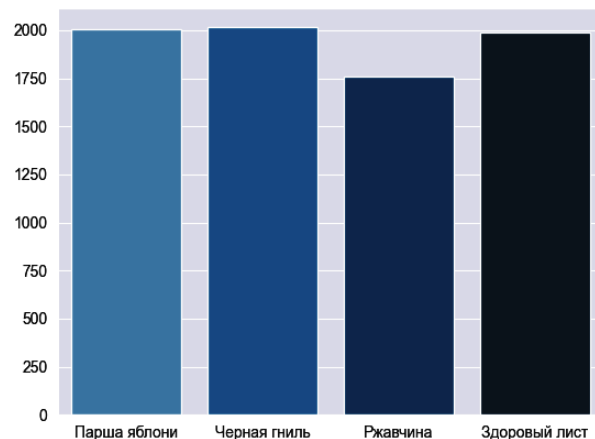
Эксперименты с гибридной моделью выполнены на дообучающем целевом наборе «Яблоня»⁴. Он отобран из библиотеки tensorflow-datasets под именем 'plant_village', включает 7771 обучающих и 1747 тестовых изображений больных и здоровых листьев яблони, сфотографированных вручную с различным освещением, углами, поверхностями и шумом.

Набор, фрагмент которого представлен на рисунке 6, включает четыре класса растений, соответствующих названию болезни или ее отсутствию: 'парша яблони', 'черная гниль', 'ржавчина', 'здоровый лист'.



Р и с. 6. Фрагмент целевого набора «Яблоня»
F i g. 6. Fragment of the "Apple Tree" target set

Количества обучающих данных каждого класса, представлены диаграммой на рисунке 7.



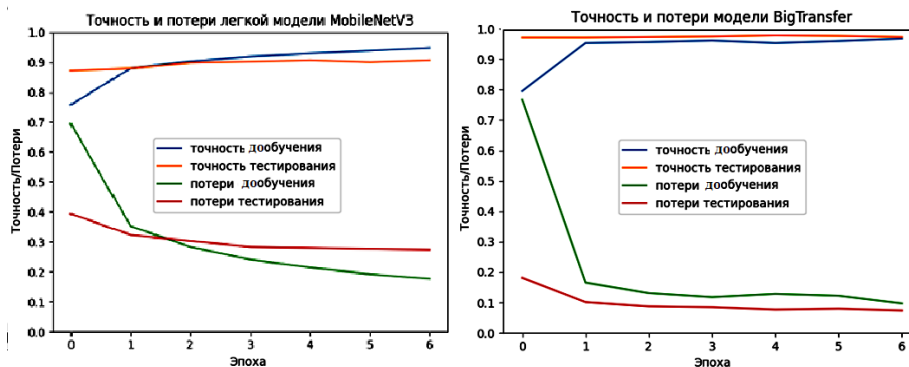
Р и с. 7. Диаграмма количества обучающих данных набора «Яблоня»
F i g. 7. Diagram of the amount of training data in the "Apple Tree" set

4 Результаты экспериментов

Модели оптимизировались на процессоре Intel(R) Core(TM) i7 CPU с 16 ГБ оперативной памяти. Программные средства чтения целевых наборов данных, загрузки предобученных моделей из библиотеки TensorFlowHub и их адаптация реализованы Python – библиотеками, включая фреймворки TensorFlow и Keras⁵ для построения нейронных сетей.

⁴ plant_village [Электронный ресурс] // Google, 2025. URL: https://www.tensorflow.org/datasets/catalog/plant_village (дата обращения: 21.07.2025).

⁵ Keras: The high-level API for TensorFlow [Электронный ресурс] // Google, 2025. URL: <https://www.tensorflow.org/guide/keras> (дата обращения: 21.07.2025).



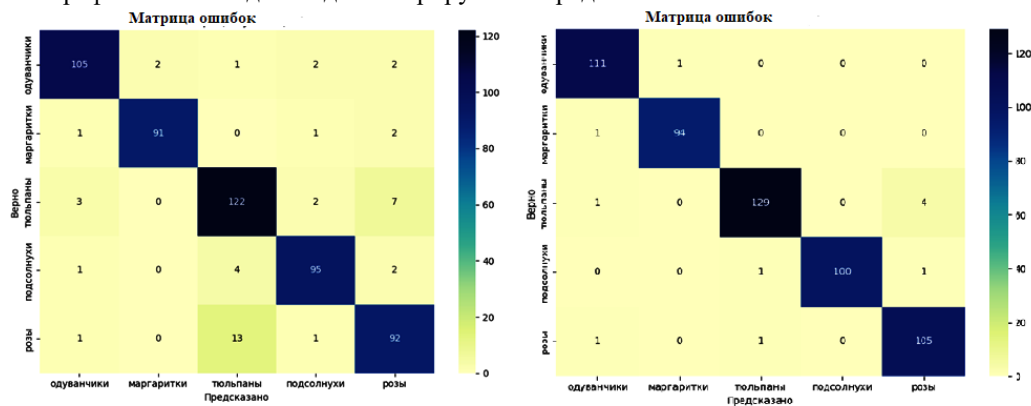
Р и с. 8. Точность и потери дообучения моделей MobileNetV3(a) и Big Transfer(b)
Fig. 8. Accuracy and loss of additional training of the MobileNetV3 (a) and Big Transfer (b) models

Показатели производительности, на основании которых сравниваются модели, – это точность, потери и время дообучения. Точность относится к её способности правильно классифицировать входные данные, мера того, как часто модель предсказывает правильную метку заданного входного набора. Она рассчитывается как отношение количества правильно предсказанных образцов к общему количеству образцов в наборе. Более высокая оценка точности указывает на то, что модель лучше предсказывает правильный результат и более надежна. Потери относятся к разнице между предсказанным результатом и фактическим результатом, мера того, насколько далеки предсказания от истинных значений. Цель дообучения модели – получить максимальную точность и минимальные потери. Процесс дообучения включает в себя несколько итераций (эпох). После каждой эпохи модель проверяется на тестовых данных с целью контроля точности и потерь. Именно результаты на тестовой выборке показывают, насколько хорошо модель будет работать в реальных условиях. Время дообучения может зависеть от разных факторов, таких как сложность модели, размер целевого набора, доступные вычислительные ресурсы и используемые гиперпараметры.

На рисунке 8 показаны графики точности и потерь в ходе дообучения и тестирования легкой модели CNN MobileNetV3 и модели Big Transfer на целевом наборе «Цветы».

Приведенные графики наглядно демонстрируют

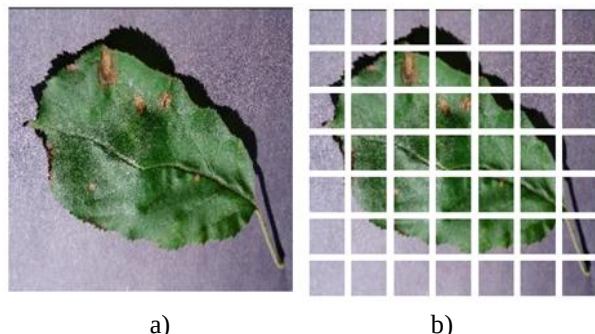
быстрое снижение потерь и увеличение точности до предельных значений, что можно объяснить тем, что обе модели инициализируются предварительно обученными весами из базовых моделей CNN. Именно поэтому уже на второй эпохе для легкой и большой моделей потери резко снижаются, а точность увеличивается. На обучающем наборе установившиеся значения точности на шестой эпохе составили 0.9342 и 0.9748, потеря – 0.1927 и 0.1583 соответственно. На тестовом наборе значения точности – 0.9035 и 0.9831, потеря – 0.2978 и 0.1328. Среднее время дообучения одной эпохи легкой и большой моделей составляет 1.2 и 4.6 минуты. Стоит отметить, что Big Transfer, дообученная методом тонкой настройки, показывает лучшие значения метрик, чем легкая модель, дообученная методом извлечения признаков. Основные гиперпараметры Big Transfer подобраны опытным путем: размер партии изображений (*batch size*) = 64, начальная скорость обучения (*learning rate*) = 0,001, оптимизатор SGD (*Stochastic Gradient Descent*). Достоинства легкой модели – простота разработки и малое время дообучения [21, 22]. Чтобы увидеть основную разницу производительности моделей, для них построены матрицы ошибок (рисунок 9). Каждая строка этой матрицы представляет экземпляры тестовой выборки в реальном классе, тогда как каждый столбец представляет экземпляры в предсказанном классе. Таким образом, диагональ матрицы представляет экземпляры, которые были правильно предсказаны.



Р и с. 9. Матрицы ошибок моделей MobileNetV3(a) и Big Transfer(b)
Fig. 9. Confusion matrices of the MobileNetV3 (a) and Big Transfer (b) models

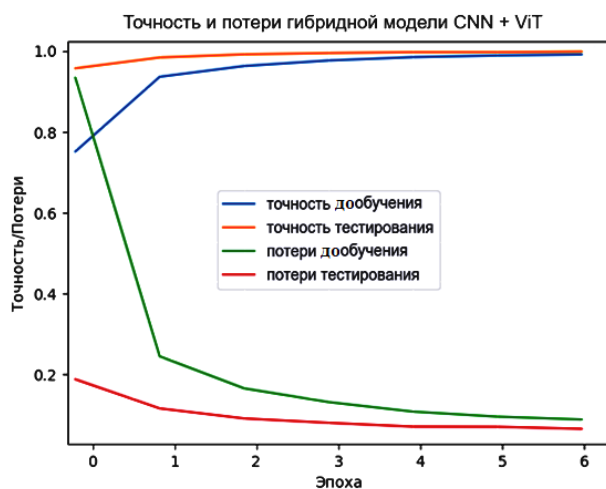


Дообучение и тестирование гибридной модели CNN + ViT выполнено на целевом наборе «Яблоня», преобразуемом в патчи с целью подачи на вход трансформера. Результат преобразования больного листа яблони в патчи представлен на рисунке 10:



Р и с. 10. Преобразование изображения больного листа (a) в патчи (b)
F i g. 10. Transformation of an image of a diseased leaf (a) into patches (b)

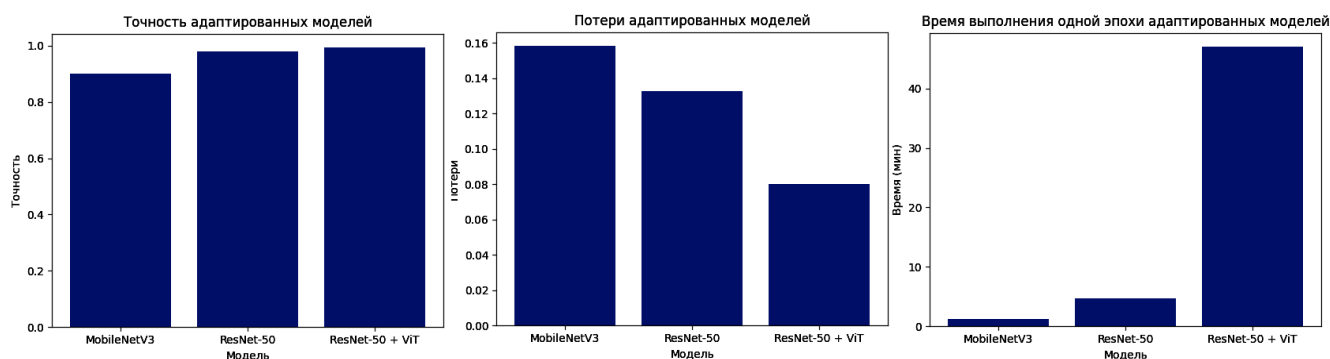
На рисунке 11 показан график точности и потерь дообучения и тестирования гибридной модели.



Р и с. 11. Точность и потери дообучения гибридной модели CNN + ViT

F i g. 11. Accuracy and loss of additional training of the hybrid CNN + ViT model

Модель CNN + ViT обеспечила точность 0.9905 на обучающей выборке и 0.9931 на тестовой, превосходя MobileNetV3 и Big Transfer. Она имеет и меньшие потери (0.1037 и 0.0815). Однако среднее время дообучения одной эпохи гибридной модели составляет



Р и с. 13. Сравнение точности, потерь и времени выполнения одной эпохи различных моделей на данных тестирования

F i g. 13. Comparison of model accuracy, loss, and training time



Р и с. 12. Матрица ошибок гибридной модели CNN + ViT
F i g. 12. Confusion matrix of the hybrid CNN + ViT model

47 минут, значительно превышая времена других моделей. С целью учета возможного дисбаланса классов этой модели, построена матрица ошибок, с результатами оценки изображений каждого класса тестовой выборки. В сферах диагностики заболеваний, например, медицине, где результаты прогноза могут иметь серьёзные последствия, анализ показателей по отдельным классам позволяет лучше оценивать такие прогнозы. Матрица ошибок тестовой выборки всех четырех классов представлена на рисунке 12.

На диаграммах рисунка 13 представлены результаты производительности дообучения всех трех рассмотренных трансферных моделей на тестовой выборке, по которым можно сравнить их точность, потери и время выполнения эпохи. Видно, что все три модели обеспечили высокие значения точности и минимальные потери. Однако, лучшая по этим параметрам гибридная модель CNN + ViT, дообученная на сложном целевом наборе, уступает по времени моделям CNN.

Основываясь на сравнении трансферных моделей, дообученных разными методами, можно утверждать, что выбор архитектуры и метода оптимизации необходимо тщательно продумывать с учётом требуемой производительности и особенностей целевого набора. В частности, различные архитектуры



продemonстрировали предпочтения в отношении определённых стратегий оптимизации, представляющие ценную практическую информацию по выбору стратегии в задачах классификации изображений [23-25].

6 Заключение

В данном исследовании дан анализ стратегий оптимизации предварительно обученных моделей классификации изображений средствами двух целевых наборов данных небольшого объема – простом и сложном. На простом наборе оценка и сравнение стратегий выполнена для моделей CNN MobileNetV3 и ResNet-50 (Big Transfer). На сложном наборе оценка выполнялась для гибридной архитектуры CNN ResNet-50 + ViT. В ходе экспериментов оцениваются производительности моделей с учетом таких факторов, как точность, потери и время обучения. Основные ключевые выводы проведенного анализа следующие.

Выбор стратегии оптимизации между извлечением признаков и тонкой настройки непосредственно влияет на баланс между точностью и временем дообучения модели.

Легкая модель CNN MobileNetV3, дообученная методом извлечения признаков, показала наилучшее быстродействие = 1.2 минуты выполнения эпохи и точность = 0.93 на простом обучающем наборе.

Большая модель Big Transfer, дообученная методом тонкой настройки, обеспечила более высокую точность = 0.97 и время выполнения эпохи = 4.6 минуты на простом обучающем наборе.

Гибридная модель CNN + ViT, дообученная методом извлечения признаков, обеспечила наилучшую точность = 0.99 и время выполнения эпохи = 47 минут на сложном обучающем наборе.

Таким образом, полученные результаты показывают сильные и слабые стороны каждой модели, подчеркивая важность выбора подходящей архитектуры модели трансферного обучения в задачах классификации изображений. Необходимо тщательно продумывать стратегию оптимизации с учетом необходимых характеристик и особенностей целевого набора. Следует заметить, что различные модели экспериментов продемонстрировали свои предпочтения в отношении определённых стратегий оптимизации, что дает ценную практическую информацию к выбору наилучшей стратегии для определенной области. Тем не менее, проведенное исследование имеет ограничения, связанные с качеством и размером целевого набора, нерассмотренными комбинациями архитектур и стратегий оптимизации. Несмотря на то, что точность реализованных моделей хорошая, их можно доработать с целью повышения точности, разморозив некоторые слои обучения. Время дообучения легко сократить, используя параллельную обработку на графических ускорителях (GPU) или тензорных процессорах (TPU). Дальнейшее направление работ может включать в себя квалитетические исследования, оценивающие отношение качества систем трансферного обучения к необходимым для этого затратам, что позволит находить оптимальный вариант решения широкого класса задач.

References

1. Gholizade M., Soltanizadeh H., Rahmamanesh M., et al. A review of recent advances and strategies in transfer learning. *International Journal of System Assurance Engineering and Management*. 2025;16:1123-1162. <https://doi.org/10.1007/s13198-024-02684-2>
2. Gupta J., Pathak S., Kumar G. Deep learning (CNN) and transfer learning: a review. *Journal of Physics: Conference Series*. 2022;2273(1):012029. <https://doi.org/10.1088/1742-6596/2273/1/012029>
3. Petrov A.I., Belov Yu.S. Transfer learning in neural network processing of medical image data. *E-Scio*. 2023;(6):88-98. (In Russ., abstract in Eng.) EDN: LOUHOC
4. Theodoris C.V., Xiao L., Chopra A., Chaffin M.D., Al Sayed Z.R., Hill M.C., Mantineo H., Brydon E.M., Zeng Z., Liu X.S., Ellinor P.T. Transfer learning enables predictions in network biology. *Nature*. 2023;618(7965):616-624. <https://doi.org/10.1038/s41586-023-06139-9>
5. Hossen M.I., Awrangieba M., Pan S., et al. Transfer learning in agriculture: a review. *Artificial Intelligence Review*. 2025;58:97. <https://doi.org/10.1007/s10462-024-11081-x>
6. Parfenov D.I., Bolodurina I.P., Grishina L.S., Zhigalov A.Yu., Tolmachev S.V. Investigation of Transfer Learning Efficiency in Adversarial Attacks. *Modern Information Technologies and IT-Education*. 2022;18(4):861-871. (In Russ., abstract in Eng.) <https://doi.org/10.25559/SITITO.18.202204.861-871>
7. Jha R., Bhattacharjee V., Mustafi A. Transfer learning with feature extraction modules for improved classifier performance on medical image data. *Scientific Programming*. 2022;2022(1):4983174. <https://doi.org/10.1155/2022/4983174>
8. Li X., et al. Improving feature extraction from histopathological images through a fine-tuning ImageNet model. *Journal of Pathology Informatics*. 2022;13:100115. <https://doi.org/10.1016/j.jpi.2022.100115>
9. Mputu H.S., Abdel-Mawgood A., Shimada A., Sayed M.S. Tomato Quality Classification Based on Transfer Learning Feature Extraction and Machine Learning Algorithm Classifiers. *IEEE Access*. 2024;12:8283-8295. <https://doi.org/10.1109/ACCESS.2024.3352745>
10. Kolesnikov A., et al. Big Transfer (BiT): General Visual Representation Learning. In: Vedaldi A., Bischof H., Brox T., Frahm J.M. (eds.) *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*. Vol. 12350.



- Cham: Springer; 2020. p. 491-507. https://doi.org/10.1007/978-3-030-58558-7_29
11. Shahriar T. Comparative Analysis of Lightweight Deep Learning Models for Memory-Constrained Devices. *arXiv:2505.03303*. 2025. <https://doi.org/10.48550/arXiv.2505.03303>
 12. Russakovsky J., Deng H., Su J., Krause S., Satheesh S., Ma Z., Huang A., Karpathy A., Khosla M., Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*. 2015;115:211-252. <https://doi.org/10.1007/s11263-015-0816-y>
 13. Liao J., et al. A machine learning-based feature extraction method for image classification using ResNet architecture. *Digital Signal Processing*. 2025;160:105036. <https://doi.org/10.1016/j.dsp.2025.105036>
 14. Joshi A.L., Veetil I.K., Premjith B., Sowmya V., Gopalakrishnan E.A., Vinayakumar R. Performance Analysis of Big Transfer Models on Biomedical Image Classification. In: Pham H. (eds.) *Analytics Modeling in Reliability and Machine Learning and Its Applications. Springer Series in Reliability Engineering*. Cham: Springer; 2025. p. 141-160. https://doi.org/10.1007/978-3-031-72636-1_7
 15. Sun C., Shrivastava A., Singh S., Gupta A. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. In: 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE Press; 2017. p. 843-852. <https://doi.org/10.1109/ICCV.2017.97>
 16. Samoilova T.A. Comparative Analysis of Hyperparameter Optimization Using Optuna and Hyperopt for Convolutional Neural Networks. *Modern Information Technologies and IT-Education*. 2024;20(3):656-665. (In Russ., abstract in Eng.) <https://doi.org/10.25559/SITITO.020.202403.656-665>
 17. Dosovitskiy A., et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv:2010.11929*. 2020. <https://doi.org/10.48550/arXiv.2010.11929>
 18. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Red Hook, NY, USA: Curran Associates Inc.; 2017. p. 6000-6010.
 19. Tharani Pavithra P., Baranidharan B. A hybrid ViT-CNN model premeditated for rice leaf disease identification. *International Journal of Computational Methods and Experimental Measurements*. 2024;12(1):35-43. <https://doi.org/10.18280/ijcmem.120104>
 20. Pantelaio D., Theofilou P.-A., Tzouveli P., Kollias S. Hybrid CNN-ViT Models for Medical Image Classification. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). Athens, Greece: IEEE Press; 2024. p. 1-4. <https://doi.org/10.1109/ISBI56570.2024.10635205>
 21. Netrapalli P. Stochastic Gradient Descent and Its Variants in Machine Learning. *Journal of the Indian Institute of Science*. 2019;99:201-213. <https://doi.org/10.1007/s41745-019-0098-4>
 22. Sun H., Cai Y., Tao R., Shao Y., Xing L., Zhang C., Zhao Q. An Improved Reacceleration Optimization Algorithm Based on the Momentum Method for Image Recognition. *Mathematics*. 2024;12(11):1759. <https://doi.org/10.3390/math12111759>
 23. Huang X. High Resolution Remote Sensing Image Classification Based on Deep Transfer Learning and Multi Feature Network. *IEEE Access*. 2023;11:110075-110085. <https://doi.org/10.1109/ACCESS.2023.3320792>
 24. Kim J.I., Baek J.W., Kim C.B. Hierarchical image classification using transfer learning to improve deep learning model performance for amazon parrots. *Scientific Reports*. 2025;15:3790. <https://doi.org/10.1038/s41598-025-88103-3>
 25. Shu M. Utilizing Transfer Learning for Deep Learning Based Image Classification. *Journal of Intelligence Technology and Innovation (JITI)*. 2025;3(1):58-73. <https://doi.org/10.30212/JITI.202503.004>

Поступила 21.07.2025; одобрена после
рецензирования 10.09.2025; принята к публикации
04.10.2025.

Submitted 21.07.2025; approved after reviewing
10.09.2025; accepted for publication 04.10.2025.

Об авторе:

Самойлова Татьяна Аркадьевна, доцент кафедры прикладной математики и информатики физико-математического факультета, ФГБОУ ВО «Смоленский государственный университет» (214000, Российская Федерация, г. Смоленск, ул. Пржевальского, д. 4), кандидат технических наук, **ORCID:** <https://orcid.org/0000-0002-3712-327X>, tatsamoilova24@gmail.com

Автор прочитал и одобрил окончательный вариант рукописи.

About the author:

Tatyana A. Samoilova, Associate Professor of the Chair of Applied Mathematics and Informatics, Faculty of Physics and Mathematics, Smolensk State University (4 Przhevalsky St., Smolensk 214000, Russian Federation), Cand. Sci. (Eng.), **ORCID:** <https://orcid.org/0000-0002-3712-327X>, tatsamoilova24@gmail.com

The author has read and approved the final manuscript.