



Министерство науки и высшего образования Российской Федерации
Мытищинский филиал
Федерального государственного бюджетного образовательного
учреждения
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МФ МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ _____ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ_____

КАФЕДРА _____СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ_____

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

НА ТЕМУ:

Предсказание дерматологической
болезни по изображению с помощью transfer
learning

Студент ИУ5-31М
(Группа)

(Подпись, дата) М.В.Кашурин
(И.О.Фамилия)

Руководитель курсовой работы

(Подпись, дата) Ю.Е. Гапанюк
(И.О.Фамилия)

2024 г.

**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)**

УТВЕРЖДАЮ
Заведующий кафедрой ИУ5
(Индекс)
В.И. Терехов
(И.О.Фамилия)
« 04 » сентября 2024 г.

**ЗАДАНИЕ
на выполнение научно-исследовательской работы**

по теме Предсказание дерматологической болезни по изображению с помощью transfer learning

Студент группы ИУ5-31М

Кашурин Максим Владимирович
(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)
ИССЛЕДОВАТЕЛЬСКАЯ

Источник тематики (кафедра, предприятие, НИР) КАФЕДРА

График выполнения НИР: 25% к ____ нед., 50% к ____ нед., 75% к ____ нед., 100% к ____ нед.

Техническое задание _Выполнить обработку и анализ данных принятых с борта космического аппарата_

Оформление научно-исследовательской работы:

Расчетно-пояснительная записка на 40 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Дата выдачи задания « 04 » сентября 2023 г.

Руководитель НИР

(Подпись, дата)

Ю.Е. Гапанюк

(И.О.Фамилия)

Студент

(Подпись, дата)

М.В. Кашурин

(И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

Введение

Человеческий фактор присутствует всегда. Зачастую это может исказить любую достоверную информацию, особенно когда дело касается точной области, такой как медицина. В ней иногда ошибочный диагноз может стоить жизни человеку. С другой стороны человеческий ресурс тоже не бесконечный. Если вдруг случится эпидемия то на каждого врача занимающейся этими болезнями, будет приходить больше тысячи пациентов, что достаточно трудоемко для сотрудника, также фактор ошибок повышается. Основным механизмом, лежащим внутри сервиса будет свёрточная нейронная сеть, которая призвана решать задачи компьютерного зрения[1].

Датасет

Данные для обучения собраны из различных источников и выложены на платформу Kaggle, содержащие 10 классов, но для обучения и последующей классификации использовались 5 кожных заболеваний.[2]

Melanoma 15.75k файлов

Basal Cell Carcinoma (BCC) 3323 файлов

Melanocytic Nevi (NV) 7970 файлов

Benign Keratosis-like Lesions (BKL) 2079 файлов

Warts Molluscum and other Viral Infections 2103 файлов

Было отобрано по 2 тысячи изображений для каждого класса, чтобы устранить дисбаланс классов.

Обработка изображений

В качестве обработки использовалась приведение к размеру 300x300, стандартизация изображений для лучшей сходимости нейронной сети, аугментации для обучения (случайный поворот, аффинные преобразования, изменения яркости и контраста).

Аугментация изображений — это метод увеличения объема данных для обучения моделей машинного обучения, особенно в области компьютерного зрения. Этот процесс включает в себя модификацию существующих изображений с целью создания новых, что позволяет улучшить качество и обобщающую способность модели.[3].

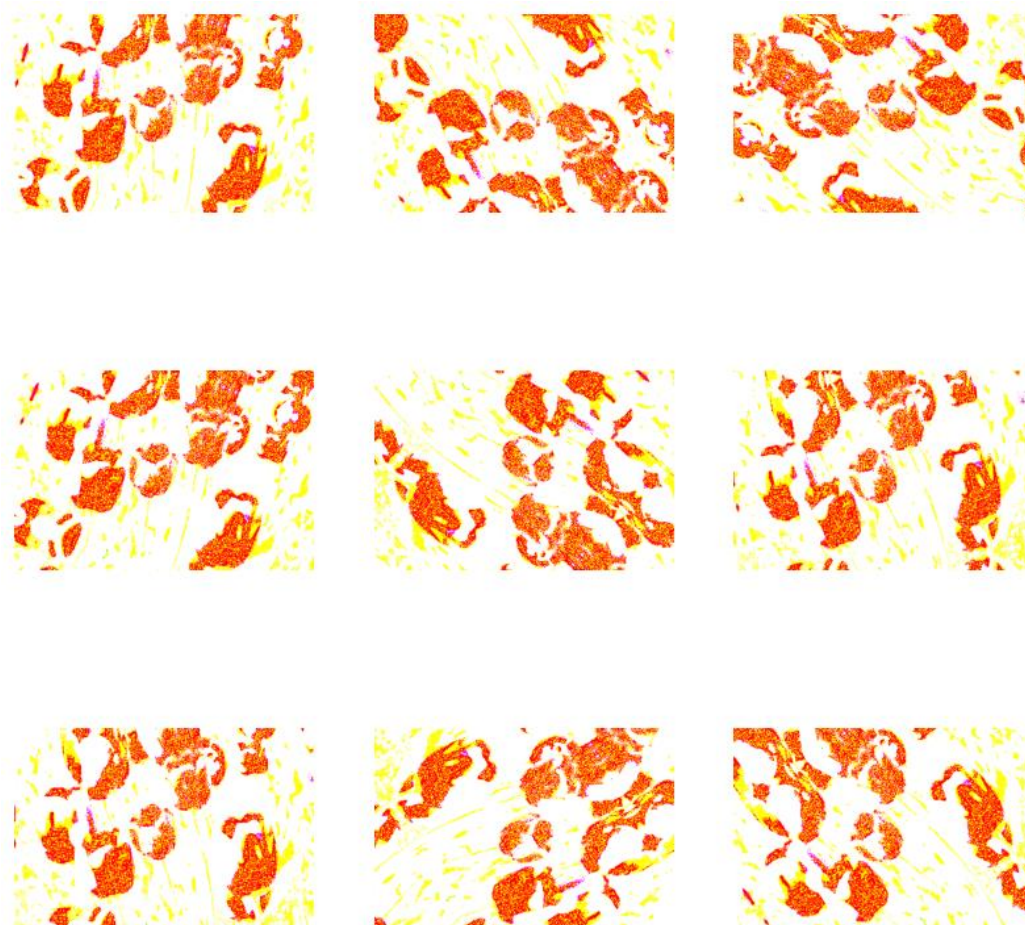


Рис. 1. Аугментация изображения при помощи поворотов

Сверточная нейронная сеть

Сверточные нейронные сети широко применяются для обработки изображений, которые являются тензорами 3-го ранга[4].

В качестве каждого слоя сверточных сетей используют ядра свертки для выявления зависимостей пикселей в исходном изображении, нахождения признаков в разных каналах. Для уменьшения размера тензора используют различные варианты pooling'а, суть которых заключается в том, чтобы фильтровать по определенному признаку активации предыдущего слоя.

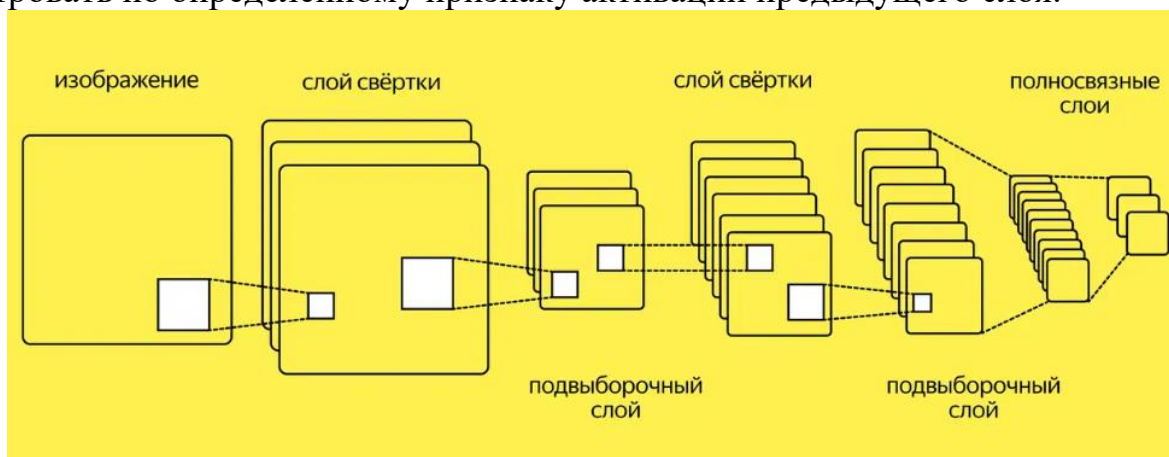


Рис. 2. Сверточная нейронная сеть (CNN)

Извлечение признаков из изображения

Основные этапы извлечения признаков:

Сверточные слои: В CNN используются сверточные слои, которые применяют фильтры (или ядра) к входным изображениям. Эти фильтры сканируют изображение, выявляя локальные паттерны, такие как края, текстуры и формы. Каждый фильтр отвечает за извлечение определенного типа признака. **Активация:** После применения свертки к изображению, результат проходит через функцию активации (например, ReLU), которая добавляет нелинейность в модель. Это позволяет сети учиться более сложным признакам.[5]

Пулинг: Для уменьшения размерности и повышения устойчивости к смещениям используется операция пулинга (например, максимальный или средний пулинг). Это помогает сохранить наиболее значимые признаки, уменьшая количество данных, которые нужно обрабатывать.

Глубокие слои: С увеличением глубины сети, извлекаемые признаки становятся все более абстрактными. Первые слои могут выявлять простые формы, в то время как более глубокие слои могут распознавать сложные объекты и концепции.

Полносвязные слои: В конце сети, после извлечения признаков, обычно располагаются полносвязные слои, которые используют извлеченные признаки для классификации или других задач.

Таким образом, сверточные нейронные сети автоматически извлекают и обрабатывают признаки, что делает их мощным инструментом для анализа изображений и других визуальных данных.

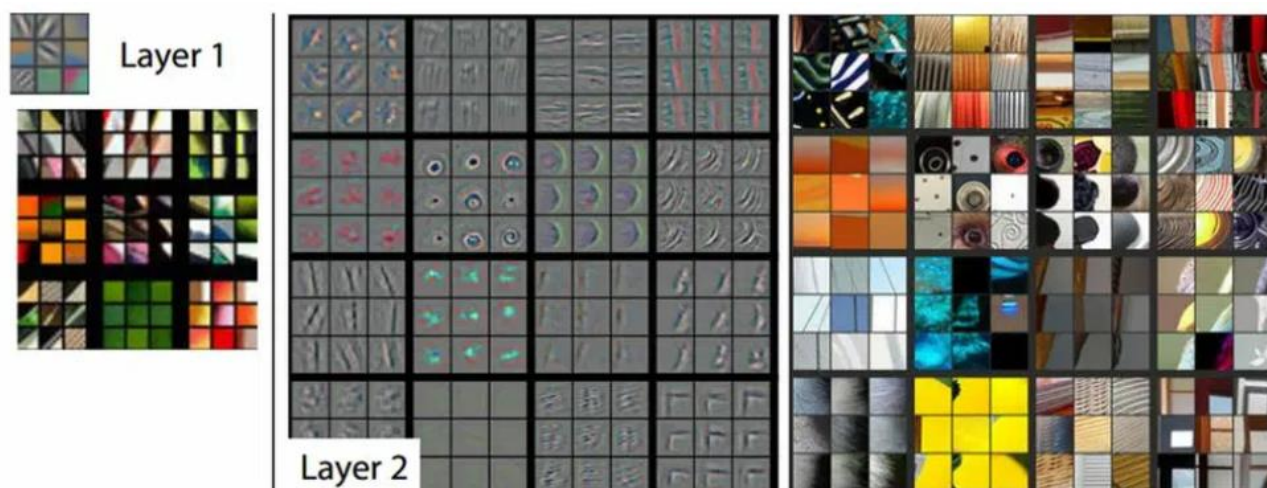


Рис. 3. Извлеченные признаки на слоях CNN

Transfer learning

Transfer learning (перенос обучения) - это метод машинного обучения, при котором знания, полученные в результате решения одной задачи или на одном наборе данных, используются для улучшения производительности модели на другой связанной задаче или наборе данных. Этот метод особенно полезен в

глубоком обучении, так как он позволяет обучать глубокие нейронные сети с использованием относительно небольшого количества данных.[6]

В переносе обучения модель, уже обученная на одной задаче, используется в качестве отправной точки для обучения модели на второй задаче. Это может быть полезно, когда вторая задача похожа на первую или когда для второй задачи доступно ограниченное количество данных.

Перенос обучения имеет множество применений, от решения задач регрессии в области науки о данных до обучения моделей глубокого обучения. Он особенно привлекателен для глубокого обучения, так как требуется огромное количество данных и вычислительных ресурсов для обучения глубоких нейронных сетей.

Также использовалась техника fine tuning.

Fine-tuning (тонкая настройка) в машинном обучении - это подход к обучению модели, при котором параметры предварительно обученной модели настраиваются на новых данных. Вместо обучения модели с нуля, мы используем предварительно обученные веса модели и затем настраиваем их на новую задачу или новый набор данных.

В процессе fine-tuning можно настроить все слои нейронной сети или только некоторые из них. Если мы заморозим (не обновляем) некоторые слои, то модель сохранит знания, полученные во время предварительного обучения, и будет использовать их для новой задачи. Это особенно полезно, когда у нас есть ограниченное количество данных для новой задачи.

Была выбрана архитектура EfficientNet b3.

Также использовались в качестве эксперимента: 3хслойная Conv NN, ResNet, MobileNet. На рис.4 видно, что из большинства моделей лучше всех при наименьших параметрах работает EfficientNet .

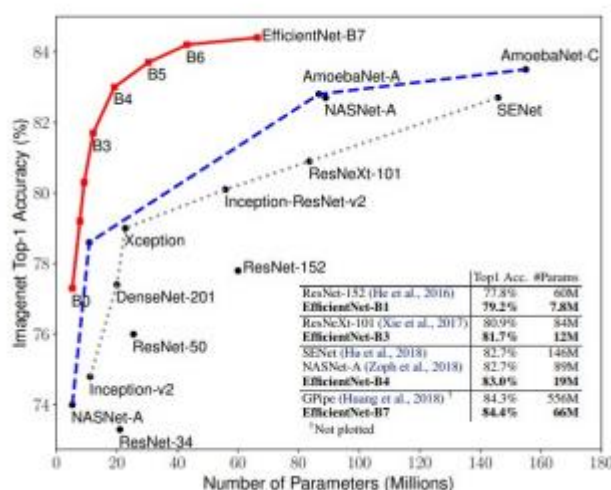


Рис. 4. График зависимости точности разных моделей нейронных сетей от количества параметров

Почему EfficientNet v3 лучше других архитектур в данной задаче

EfficientNet v3 — это последняя версия архитектуры EfficientNet, которая была разработана для достижения высокой точности классификации изображений при минимальных вычислительных затратах. Она предлагает несколько ключевых преимуществ по сравнению с предыдущими версиями EfficientNet, а также с другими популярными архитектурами, такими как ResNet и MobileNet.[7]

Преимущества EfficientNet v3:

Оптимизация производительности: EfficientNet v3 использует методы автоматического масштабирования, что позволяет достигать высокой точности при меньших размерах модели и меньших вычислительных затратах. Это делает её более эффективной по сравнению с предыдущими версиями и другими архитектурами.

Улучшенная точность: Эта архитектура демонстрирует улучшенные результаты по сравнению с ResNet и MobileNet, обеспечивая более высокую точность классификации при аналогичных или меньших вычислительных затратах. Например, EfficientNet-B0 уже показывает точность, сопоставимую с ResNet-34, но с меньшим количеством параметров и операций.

Гибкость и масштабируемость: EfficientNet v3 предлагает гибкие настройки для различных задач, что позволяет легко адаптировать модель под конкретные требования, сохраняя при этом высокую производительность.

Снижение вычислительных затрат: Архитектура EfficientNet v3 оптимизирована для работы на устройствах с ограниченными ресурсами, что делает её идеальной для мобильных и встроенных приложений, в отличие от более тяжелых моделей, таких как ResNet.

Сравнение с другими архитектурами

Сравнение с ResNet: ResNet известен своей способностью увеличивать глубину сети без потери производительности благодаря остаточным связям. Однако, несмотря на это, EfficientNet v3 достигает сопоставимой или даже лучшей точности при меньшем количестве параметров и вычислительных операций, что делает её более эффективной для практического использования.

Сравнение с MobileNet: MobileNet также ориентирован на мобильные устройства и предлагает хорошую производительность при низких вычислительных затратах. Однако EfficientNet v3 превосходит MobileNet по точности, что делает её более предпочтительной для задач, требующих высокой точности классификации.

EfficientNet представляет собой значительный шаг вперед в области архитектур нейронных сетей для классификации изображений. Она сочетает в себе высокую точность, эффективность и гибкость, что делает её отличным выбором для современных приложений в области компьютерного зрения.

Метрики

Для измерения эффективности работы модели используют различные метрики. Они бывают разные под определенный тип задач. В исследовании были использованы следующие метрики классификации: precision, recall, accuracy, fl-score.

$$precision = \frac{TP}{TP + FP}$$

Где TP (true positive) – количество верно классифицированных положительных меток.

FP (false positive) – количество ложно классифицированных положительных меток.

$$recall = \frac{TP}{TP + FN}$$

Где FN (false negative) – количество ложно классифицированных негативных меток.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Где TN (true negative) – количество верно классифицированных положительных меток

Для получения общей картины использована метрика fl-score, рассчитываемая по формуле

$$f1 = 2 * \frac{precision * recall}{precision + recall}$$

Результаты

В данной работе было проведено исследование применения нейронных сетей в целях диагностирования кожных заболеваний по изображению тела пациента.

Результаты работы для мультиклассовой классификации представлены на рисунке 5.

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 1.Melanoma | 0.99 | 0.98 | 0.98 | 100 |
| 2.BCC | 0.93 | 0.94 | 0.94 | 100 |
| 3.MN | 0.98 | 0.86 | 0.91 | 100 |
| 4.BKL | 0.86 | 0.96 | 0.91 | 100 |
| 5.WMVI | 1.00 | 1.00 | 1.00 | 100 |
| accuracy | | | 0.95 | 500 |
| macro avg | 0.95 | 0.95 | 0.95 | 500 |
| weighted avg | 0.95 | 0.95 | 0.95 | 500 |

Рис. 5 Результаты мультиклассовой классификации.

Из рисунка 5 видно, что на каких-то классах сеть показывает себя не лучшим образом. Происходит это предположительно из-за того, что в обучающей выборке мало представителей этого класса.

Так как на бинарной классификации проблема дисбаланса классов не так выражена, то результаты лучше, чем на мультиклассовой классификации.

Заключение

По результатам работы был исследован метод выявления кожных заболеваний по изображению с использованием алгоритмов глубокого машинного обучения. Выявлена наилучшая архитектура модели для обработки данных. Считаем, что для улучшения работы следует использовать модели трансформеры .

Список литературы

1. Свёрточная нейронная сеть
https://ru.wikipedia.org/wiki/Свёрточная_нейронная_сеть
2. <https://www.kaggle.com/datasets/ismailpromus/skin-diseases-image-dataset/data>
3. https://www.tensorflow.org/tutorials/images/data_augmentation?hl=ru
4. https://en.wikipedia.org/wiki/Convolutional_neural_network
5. Keiron O'Shea, Ryan Nash: An Introduction to Convolutional Neural Networks
<https://arxiv.org/abs/1904.02838>
6. Transfer Learning for Performance Modeling of Deep Neural Network Systems
<https://arxiv.org/abs/1904.02838>
7. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
<https://arxiv.org/pdf/1905.11946v3>