

Секция «Искусственный интеллект и цифровая трансформация в бизнесе и
государственном управлении»

**Анализ медицинских изображений для диагностики вирусной и
бактериальной пневмонии с использованием свёрточных нейронных сетей**

Научный руководитель – Бобоев Шараф Асрорович

Хамидов Сиёвуш Халифабобоевич

Студент (бакалавр)

Филиал Московского государственного университета имени М.В.Ломоносова в городе
Душанбе, Душанбе, Таджикистан
E-mail: *siyovushchik14@gmail.com*

Пневмония является одной из ведущих причин смертности, особенно в странах с ограниченными медицинскими ресурсами. Ранняя диагностика заболевания играет ключевую роль в снижении летальности и осложнений. Использование методов машинного обучения и глубоких нейронных сетей позволяет автоматизировать процесс выявления пневмонии на рентгеновских снимках, что может значительно повысить эффективность диагностики.

В данной работе рассматривается применение свёрточных нейронных сетей (CNN) для классификации рентгеновских изображений грудной клетки. Модель обучалась на специально размеченном наборе данных, содержащем снимки пациентов с диагнозом *пневмония* и *норма*. Основное внимание уделяется архитектуре модели, процессу обучения и оценке её эффективности.

Функция `preprocess_image` принимает изображение в формате OpenCV и выполняет следующие шаги:

- 1) Изменение размера изображения: Изображение изменяется до размера 150×150 пикселей.
- 2) Нормализация пиксельных значений: Пиксельные значения нормализуются в диапазон $[0, 1]$ путем деления на 255.

Математически это можно представить следующим образом:

Пусть I — исходное изображение, $I_{resized}$ — изображение после изменения размера, и $I_{normalized}$ — нормализованное изображение. Тогда:

$$I_{resized} = \text{resize}(I, (150, 150))$$

$$I_{normalized} = \frac{I_{resized}}{255}$$

Архитектура свёрточной нейронной сети (CNN) представлена следующим образом:

```
model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(150, 150, 1)),
    BatchNormalization(),
    MaxPooling2D(pool_size=(2, 2)),
    Dropout(0.25),
    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(1, activation='sigmoid')
])
```

Модель состоит из пяти сверточных слоев с функцией активации ReLU:

$$f(x) = \max(0, x)$$

После каждого сверточного слоя применяются слои пакетной нормализации и подвыборки (MaxPooling), что позволяет уменьшить размерность входных данных и повысить обобщающую способность модели. Для предотвращения переобучения использовались слои Dropout с вероятностью $p = 0.5$.

Обучение модели происходило с использованием функции потерь бинарной кросс-энтропии:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)]$$

где y_i — истинная метка класса, $\hat{y}_i$ — предсказанная вероятность принадлежности к классу пневмонии, N — количество обучающих примеров.

Для минимизации функции потерь использовался оптимизатор RMSprop, с адаптивным выбором коэффициента обучения. Скорость обучения адаптировалась с помощью механизма *ReduceLROnPlateau*.

Датасет был разделен на обучающую, валидационную и тестовую выборки. Обучение проводилось в течение 11 эпох с размером батча 32. Для увеличения вариативности обучающих примеров использовался метод аугментации данных, включающий повороты, масштабирование и отражения изображений.

Результаты обучения модели показывают точность 96.66% на обучающем наборе данных и 75.00% на валидационной выборке. Для окончательной проверки модель тестировалась на независимом наборе данных, где показала точность 90.22% при значении функции потерь 0.2470.

Для оценки производительности модели использовался отчет классификации, содержащий метрики точности, полноты и F_1 -меры. В таблице 1 представлены основные результаты:

Метрика	Пневмония	Норма
Точность	93%	86%
Полнота	92%	88%
F_1 -мера	92%	87%

Таблица 1: Результаты классификации модели на тестовой выборке.

Результаты показывают эффективность модели в задаче диагностики пневмонии на основе рентгеновских снимков. Несмотря на высокую точность, наблюдается небольшая разница между точностью на обучающем и валидационном наборах данных, что может свидетельствовать о частичном переобучении. Один из путей решения этой проблемы — расширение тренировочного набора данных и применение более агрессивных методов регуляризации.

Сравнение с другими подходами показало, что предложенная архитектура превосходит традиционные методы обработки изображений, но уступает сложным ансамблевым моделям. Тем не менее, высокая скорость работы и компактность модели делают её подходящей для клинических условий, особенно в регионах с ограниченными вычислительными ресурсами.

Также анализ интерпретируемости модели с использованием активационных карт Grad-CAM показал, что модель может визуализировать области, на основе которых принимает решения, что повышает доверие медицинских специалистов.

В данной работе предложена модель сверточной нейронной сети для автоматического выявления пневмонии по рентгеновским снимкам. Проведенное исследование показало, что предложенная архитектура достигает точности около 90% на независимых данных.

Несмотря на хорошие результаты, остаются открытые вопросы, такие как улучшение генерализации модели и уменьшение вероятности ложных срабатываний. Дальнейшие исследования могут быть направлены на увеличение объема данных и применение методов самообучения. Также перспективным направлением является интеграция алгоритмов объяснимого ИИ для повышения доверия к модели в медицинском сообществе.

Источники и литература

- 1) J. Grus, Data Science from Scratch: First Principles with Python, O'Reilly Media, 2019
- 2) A. Artasanchez, P. Joshi, Artificial Intelligence with Python - Second Edition, Packt Publishing, 2020.
- 3) M. Sewak, M. R. Karim, P. Pujari, Practical Convolutional Neural Networks, Packt Publishing, 2018.

Иллюстрации

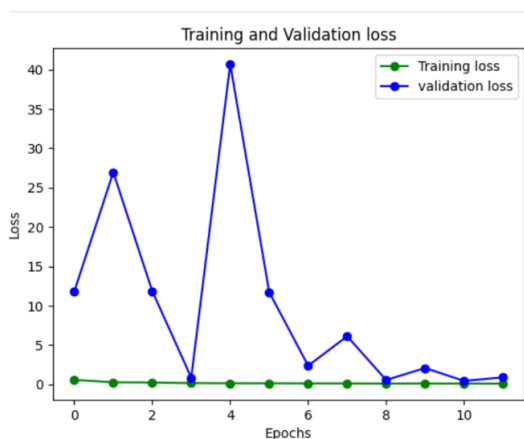


Рис. : Потери при обучении и валидации

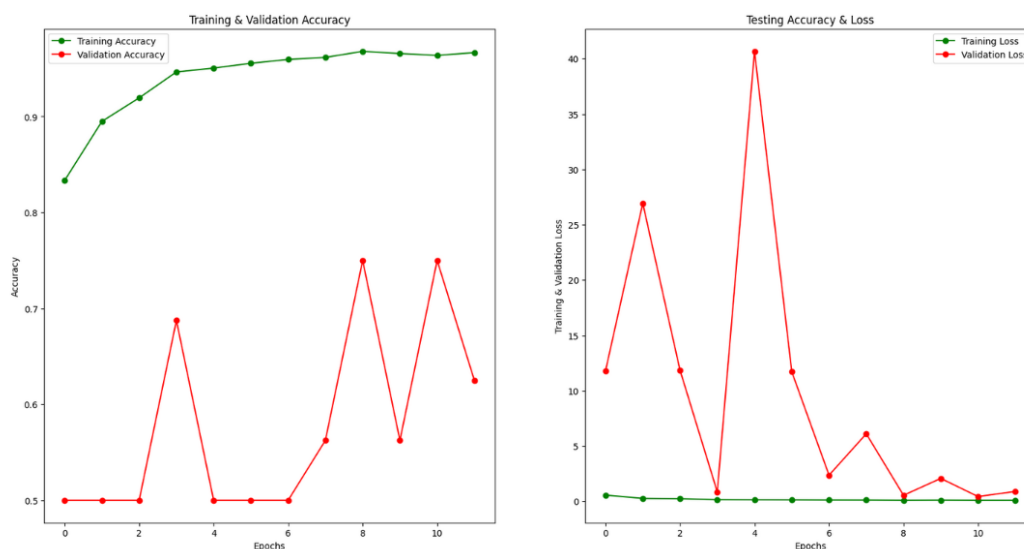


Рис. : Тренировочная и Валидационная точности (слева), Тестовые точность и потери (справа)

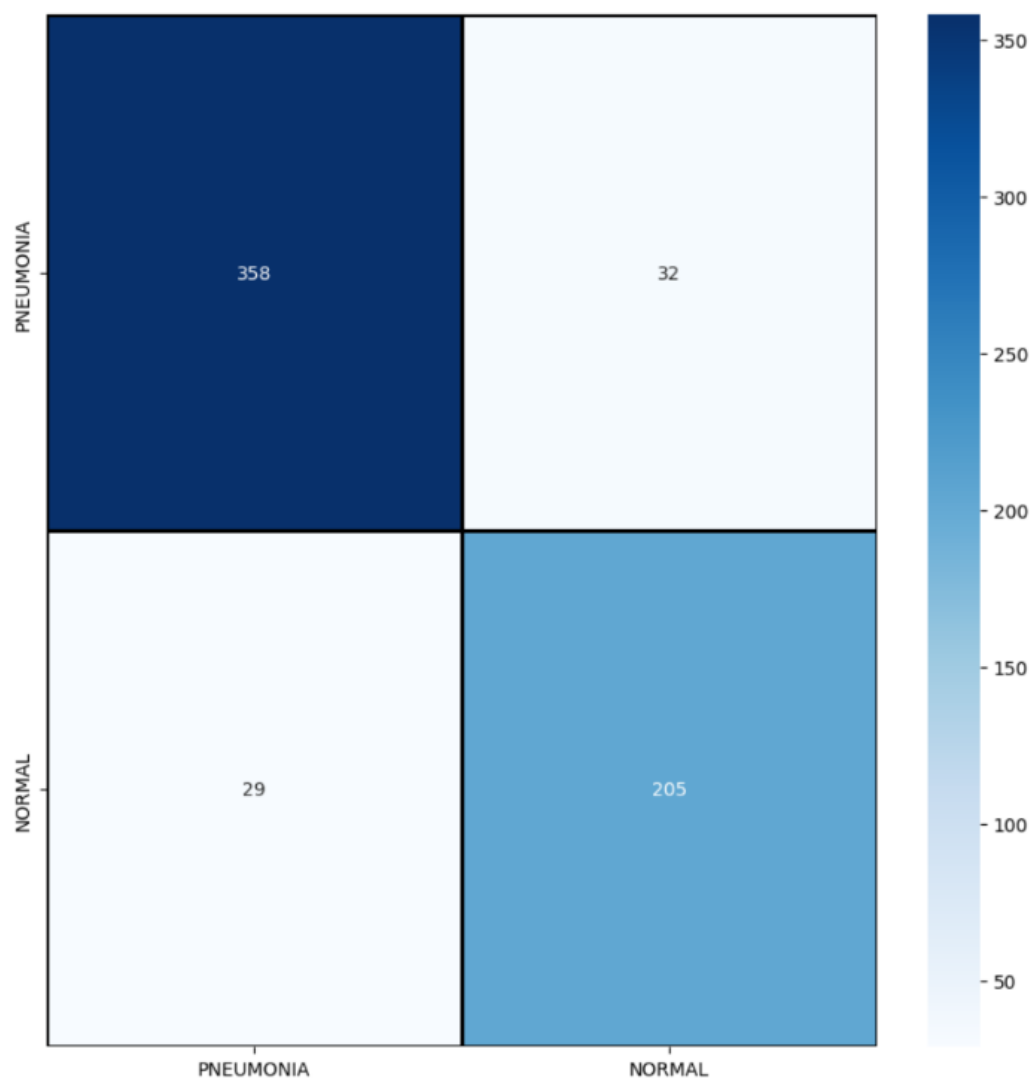


Рис. : Матрица несоответствий, построенная для детального анализа производительности по классам