

Классификация изображений с помощью сверточных нейронных сетей

Курсовая работа

Нефедов Д.А.

МГТУ им. Н.Э. Баумана

Научный руководитель: Фетисов Д.А.

Москва, 2025 г.

Постановка задачи

- Дано: набор данных с изображениями 10 видов животных: бабочка, паук, собака, слон, овца, кот, корова, белка, лошадь, курица.
- Объем данных: 26 тысяч изображений.
- Задача: построение классификатора на основе модели сверточной нейронной сети.
- Средства: язык **Python** и библиотека глубокого обучения **pytorch**.

Распределение данных

Вывод: изображения приведены к разрешению 256×256

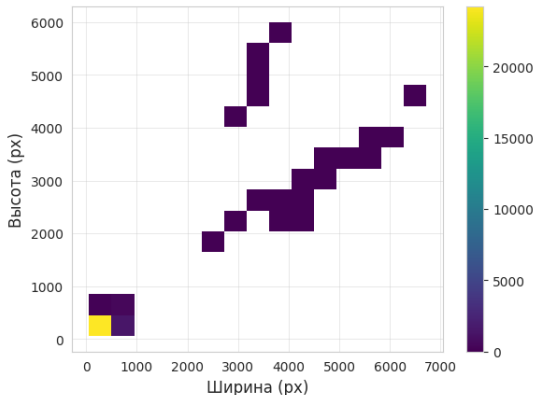


Рисунок 1 – Распределение данных по разрешениям

Распределение данных (продолжение)

Вывод: учитывался баланс классов

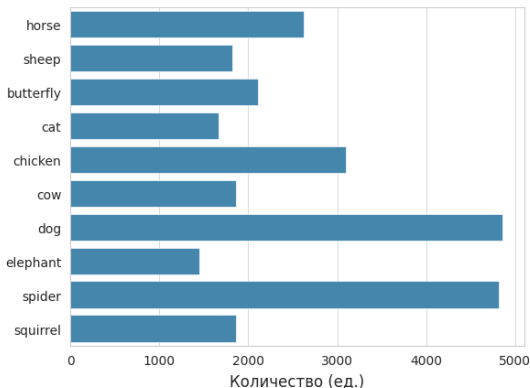


Рисунок 2 – Распределение данных по классам

Сверточная нейронная сеть

Две интерпретации свертки:

- Классическая: $(X * Y)(t, s) = \sum_{i,j} X_{i,j} \cdot Y_{t-i,s-j}$
- Взаимная корреляция: $(X \otimes Y)(t, s) = \sum_{i,j} X_{i,j} \cdot Y_{i+t,j+s}$

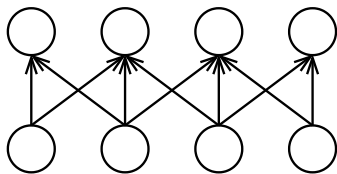


Рисунок 3 – Сверточный слой

Сверточная нейронная сеть (продолжение)

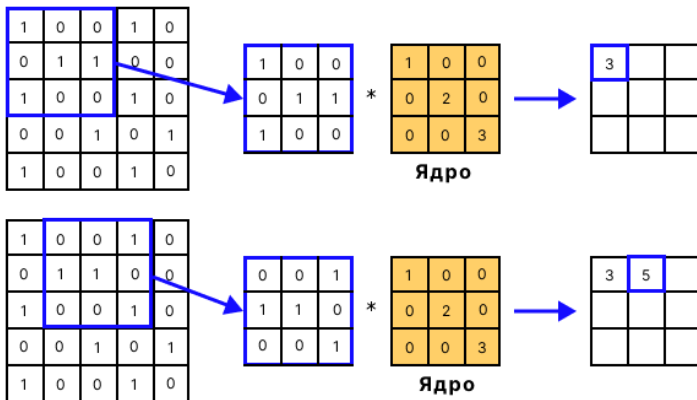


Рисунок 4 – Визуализация свертки

Модель AlexNet

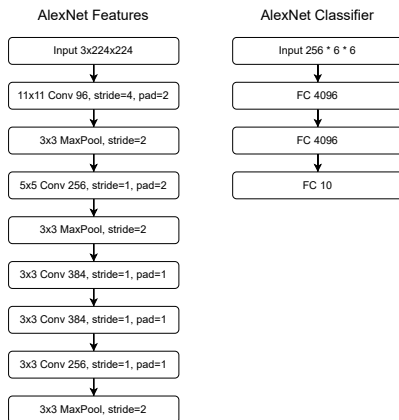


Рисунок 5 – Модель AlexNet

Результаты обучения AlexNet

Вывод: получена желаемая точность ($> 0.85\%$)

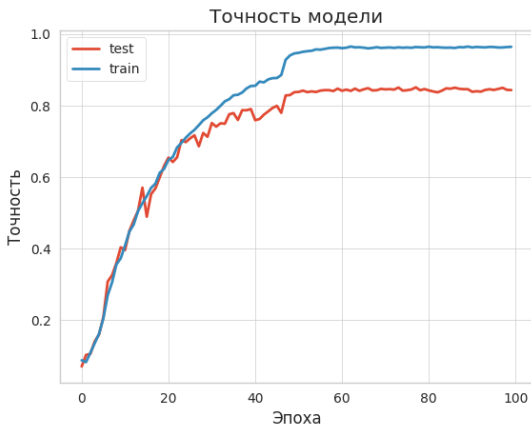


Рисунок 6 – График точности AlexNet по эпохам

Визуализация AlexNet

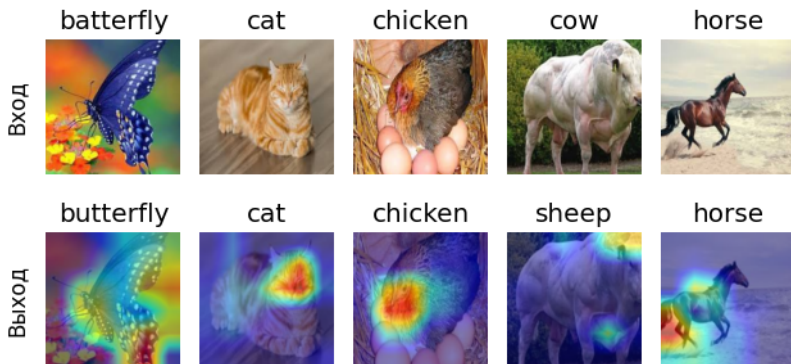


Рисунок 7 – Результаты применения Grad-CAM к модели AlexNet

Визуализация AlexNet (продолжение)

Вывод: паук предсказывается по окружению, слабая классификация собаки.

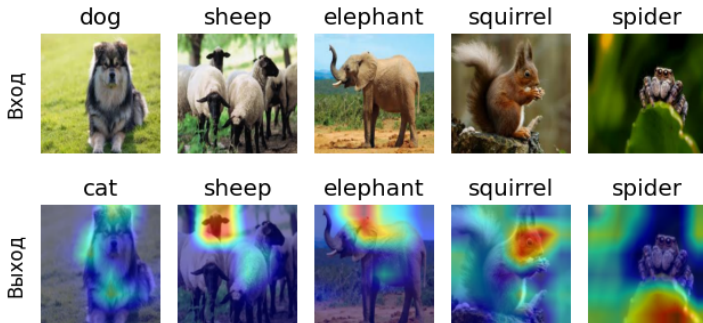


Рисунок 8 – Результаты применения Grad-CAM к модели AlexNet (продолжение)

Модель VGG13

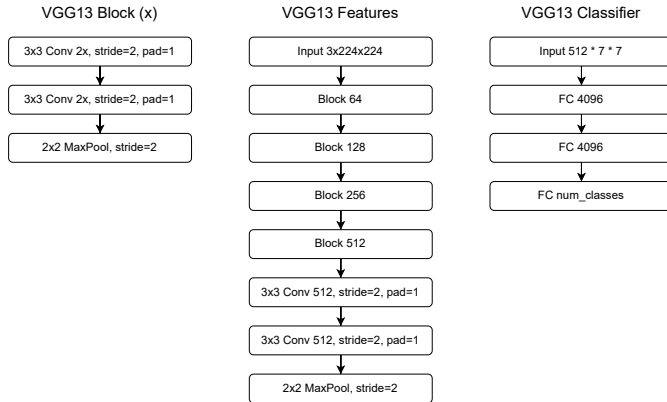


Рисунок 9 – Модель VGG13

Модификация модели VGG13

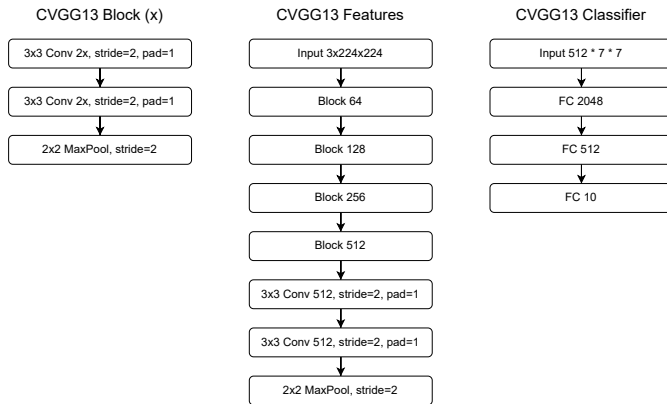


Рисунок 10 – Модификация VGG13

Результаты обучения CVGG13

Вывод: получена лучшая точность ($\approx 0.92\%$)

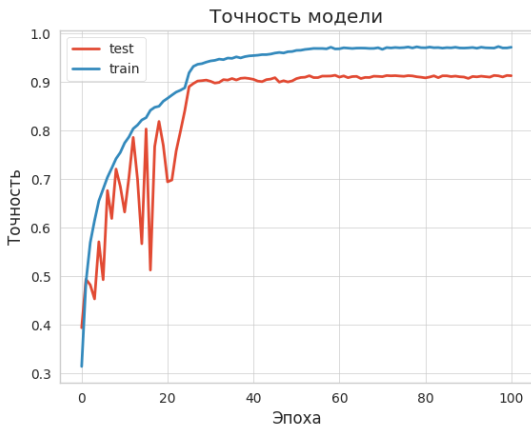


Рисунок 11 – График точности CVGG13 по эпохам

Визуализация CVGG13

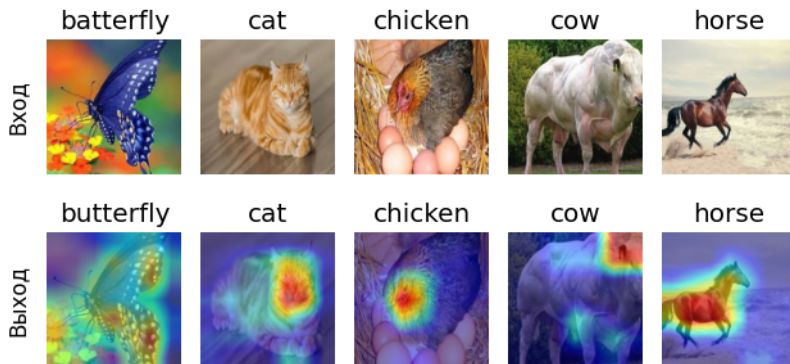


Рисунок 12 – Результаты применения Grad-CAM к модели CVGG13

Визуализация CVGG13 (продолжение)

Вывод: корректно выделенные признаки, в отличие от AlexNet

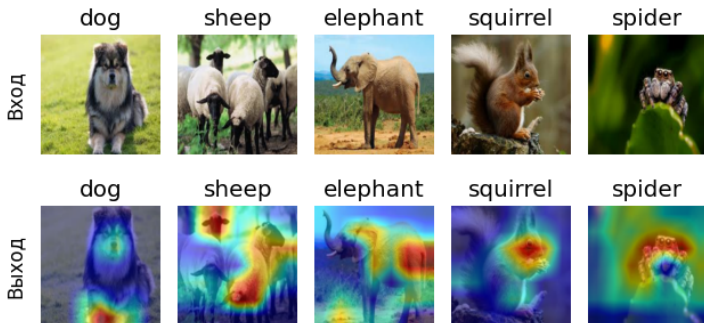


Рисунок 13 – Результаты применения Grad-CAM к модели CVGG13 (продолжение)

В рамках работы были решены следующие задачи:

- Был изучен принцип работы сверточной нейронной сети
- Были изучены и реализованы архитектуры AlexNet и VGG
- Классификаторами была достигнута желаемая точность ($> 80\%$)
- Проведен анализ поведения модели (Grad-CAM)

Таким образом, была достигнута **цель** работы.