|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **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)** |

ФАКУЛЬТЕТ \_\_\_\_\_\_\_ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

КАФЕДРА \_\_\_\_\_ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ (ИУ5)\_\_\_\_\_

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

***К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ***

***НА ТЕМУ:***

***Классификация CIFAR10 с помощью нейронных сетей \_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

Студент \_\_\_ ИУ5-65Б\_\_ **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А. C. Ищенко\_\_**

(Группа) (Подпись, дата) (И.О.Фамилия)

Руководитель **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_Ю.Е. Гапанюк\_\_\_**

(Подпись, дата) (И.О.Фамилия)

*Москва, 2023 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)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(И.О.Фамилия)

« \_\_\_\_\_ » \_\_\_\_\_\_\_\_\_\_\_\_ 20 \_\_\_\_ г.

**ЗАДАНИЕ**

**на выполнение научно-исследовательской работы**

по теме \_\_\_\_Классификация CIFAR10 с помощью нейронных сетей\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Студент группы \_\_\_\_ИУ5-65Б\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Ищенко Анастасия ?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Фамилия, имя, отчество)

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

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Исследовательская\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Источник тематики (кафедра, предприятие, НИР) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_НИР\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Техническое задание Исследовать архитектуры нейронных сетей для классификации изображений из датасета CIFAR10\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Оформление научно-исследовательской работы:***

Расчетно-пояснительная записка на \_20\_ листах формата А4.

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

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дата выдачи задания «13» февраля 2023 г.

**Руководитель НИР**  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_ **Ю.Е. Гапанюк**\_\_\_

(Подпись, дата) (И.О.Фамилия)

**Студент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_ А.C. Ищенко** \_\_

(Подпись, дата) (И.О.Фамилия)

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 4](#_Toc131275276)

[1 Постановка задачи 5](#_Toc131275277)

[2 Описание исходных данных и методов 6](#_Toc131275278)

[3 Выполнение работы 7](#_Toc131275279)

[ЗАКЛЮЧЕНИЕ 19](#_Toc131275280)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 20](#_Toc131275281)

# **ВВЕДЕНИЕ**

Распознавание образов – одна из основных задач глубокого обучения, имеющая большое значение при разработке автоматизированных систем управления.

Для задач классификации изображений разработано множество архитектур нейронных сетей, каждая из которых имеет ряд преимуществ и недостатков.

Целью данной научно-исследовательской работы является исследование архитектур нейронных сетей для классификации изображений из датасета CIFAR10 [1], обучение соответствующих моделей и оценка их качества.

# **1 Постановка задачи**

Датасет CIFAR10 является стандартом при тестировании новых нейросетевых алгоритмов, предназначенных для работы с изображениями, поскольку, с одной стороны, имеет высокое качество (идеальный баланс классов, отсутствие пропусков и дублирования данных и т.д.), а с другой – небольшой размер, что позволяет быстро обучать модели. Задачей работы является поиск наилучшей архитектуры нейронной сети для классификации CIFAR10, её обучение с использованием алгоритма Adam и оценка качества полученной модели.

# **2 Описание исходных данных и методов**

Датасет содержит 60000 цветных изображений размером 32 на 32 пиксела. Значения цветовых компонент представлены типом данных uint8 (byte), каждое изображение представляет собой трёхмерный массив в формате CHW, последовательность цветовых каналов - RGB. Датасет изначально разделён на обучающую выборку, содержащую 50000 изображений в 5 файлах (10000 изображений в каждом), и тестовую, содержащую 10000 изображений в одном файле. В каждом файле датасета хранится словарь Python, упакованный средством сериализации Pickle. Изображения хранятся в массиве NumPy с ключом «data», метки классов хранятся в списке целых чисел Python с ключом «labels».

Изображения датасета разделены на 10 классов: «самолёт», «автомобиль», «птица», «кошка», «олень», «собака», «лягушка», «лошадь», «корабль», «грузовик».

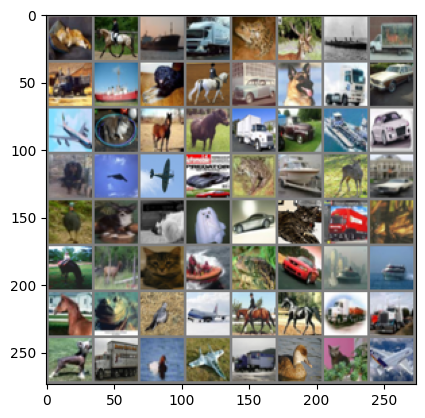


Рисунок 2.1 – некоторые изображения обучающей выборки

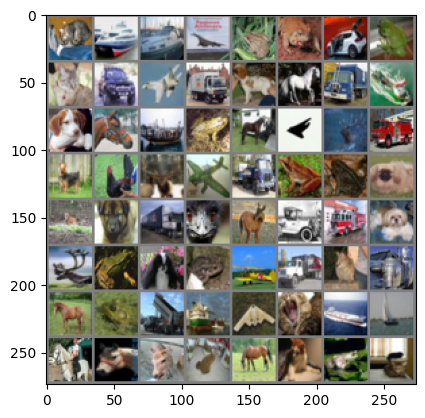


Рисунок 2.2 – некоторые изображения тестовой выборки

В работе рассмотрены следующие модели глубокого обучения (использована библиотека pytorch):

1. Блок из 10 логистических регрессий
2. Полносвязная нейронная сеть
3. Свёрточная нейронная сеть LeNet-5
4. VGG-подобная свёрточная нейронная сеть
5. Resnet-подобная остаточная нейронная сеть
6. WideResNet-подобная остаточная нейронная сеть

В качестве метрик качества моделей были использованы:

1. CrossEntropyLoss – перекрёстная энтропия (в качестве функции потери и для оценки сходимости)
2. Precision – точность

# **3 Выполнение работы**

Во всех экспериментах примем размер пакета, равный 64, и количество эпох, равное 16.

**3.1 Логистическая регрессия**

В качестве первой исследованной модели была использована логистическая регрессия. Поскольку логистическая регрессия сама по себе не может быть применена для множественной классификации, необходимо создать несколько независимых регрессий по количеству классов. В коде это реализовано одиночным полносвязным слоем с 3072 входными и 10 выходными признаками. Слой был обучен стохастическим градиентным спуском с импульсом (momentum SGD). Скорость обучения (learning\_rate) взята равной , импульс равен 0.9.

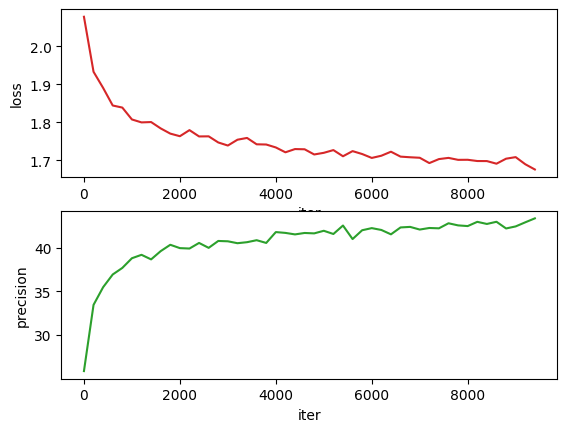
**

Рисунок 3.1.1 – графики метрик качества модели в процессе обучения

Обученная модель показала точность на тестовой выборке, равную 41%.

**3.2 Полносвязная нейронная сеть**

В работе использована классическая полносвязная нейронная сеть, состоящая из полносвязных слоёв и сигмоиды в качестве функции активации.

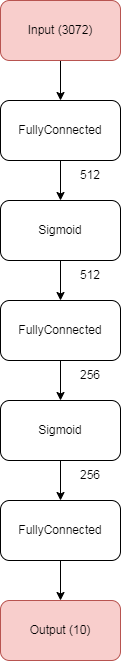


Рисунок 3.2.1 – архитектура нейронной сети

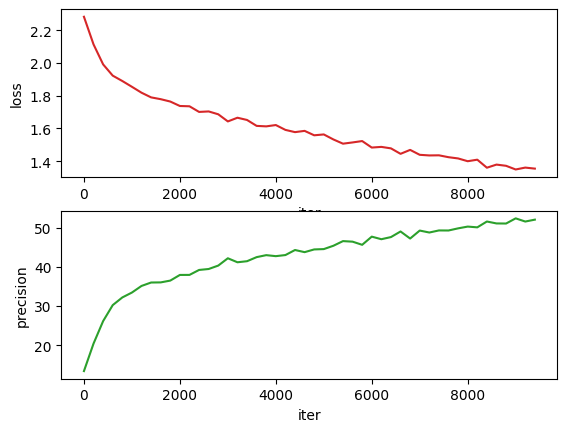


Рисунок 3.2.2 – графики метрик качества модели в процессе обучения

Обученная модель показала точность на тестовой выборке, равную 48,27%.

**3.3 Свёрточная нейронная сеть LeNet-5**

Архитектура взята из работы Яна Лекуна [2].

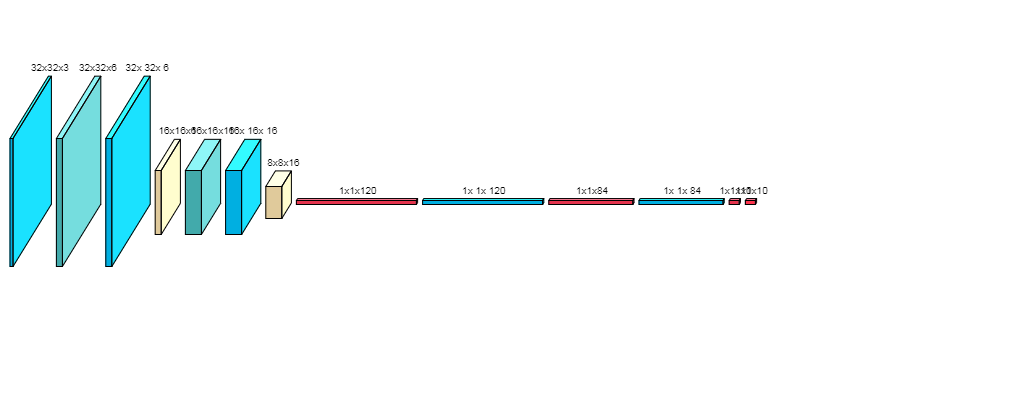


Рисунок 3.3.1 – архитектура нейронной сети

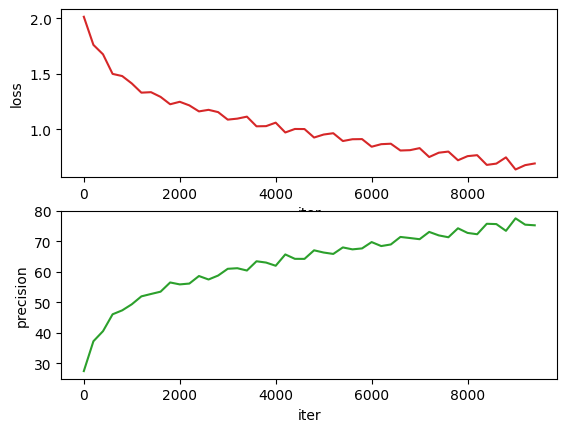


Рисунок 3.3.1 - графики метрик качества модели в процессе обучения

Обученная модель показала точность на тестовой выборке, равную 59,3%.

**3.4 VGG-подобная свёрточная нейронная сеть**

При построении нейронной сети использованы подходы, описанные в статье [3].



Рисунок 3.4.1 – архитектура нейронной сети

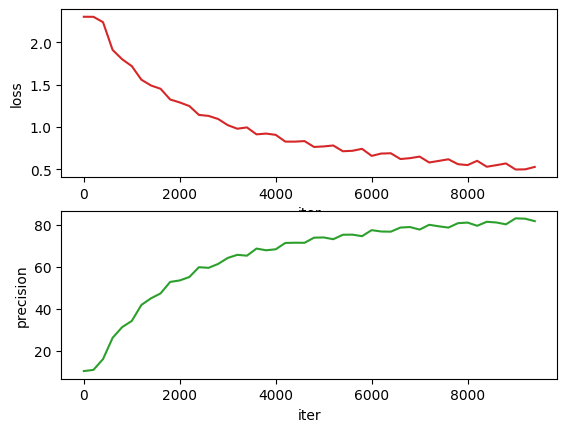


Рисунок 3.4.2 – графики метрик качества модели в процессе обучения

Обученная модель показала точность на тестовой выборке, равную 75,4%.

**3.5 ResNet-подобная остаточная нейронная сеть**

Использованы подходы из статьи [4].

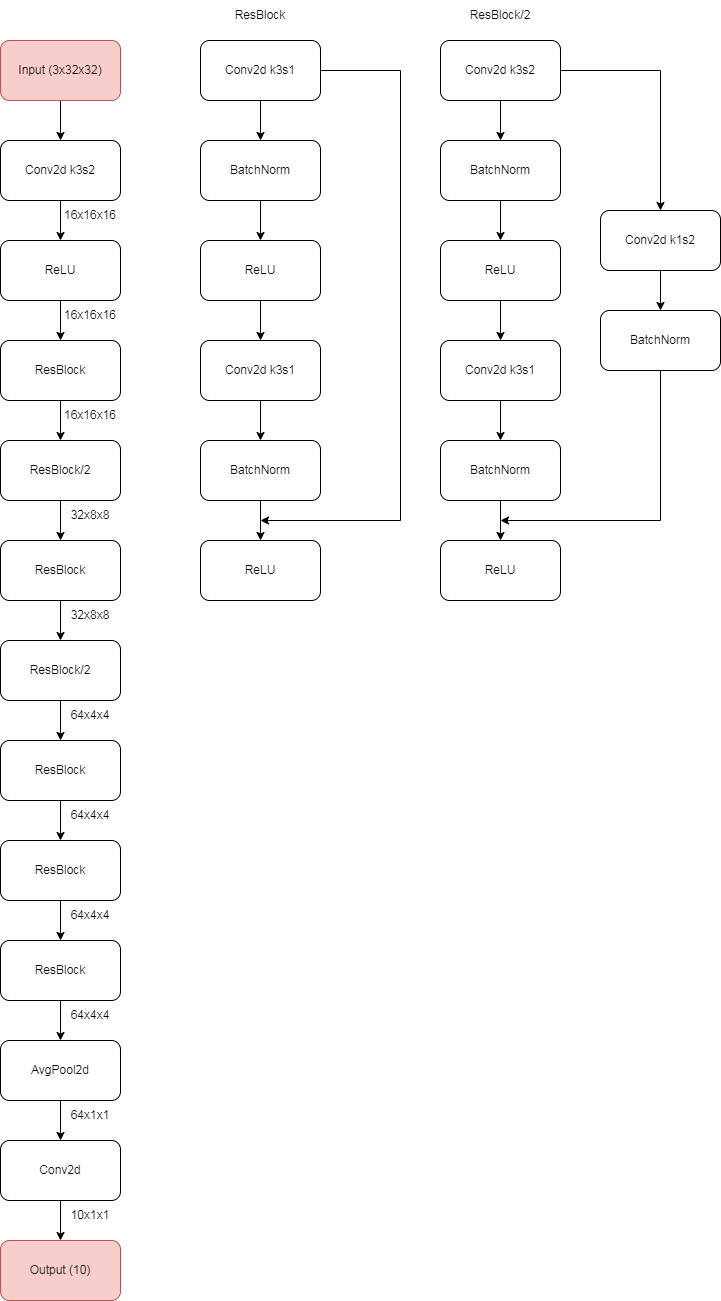


Рисунок 3.5.1 – архитектура нейронной сети

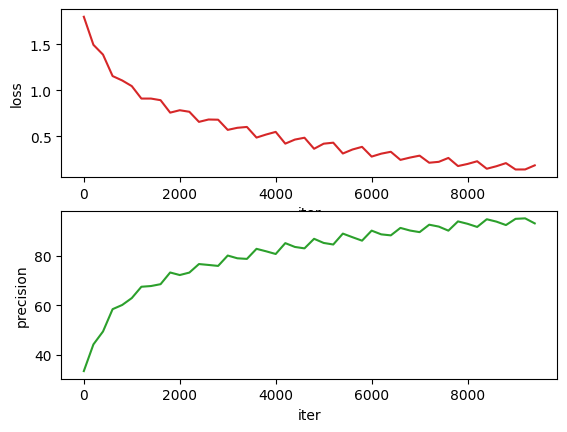


Рисунок 3.5.2 – графики метрик качества модели в процессе обучения

Обученная модель показала точность на тестовой выборке, равную 72,91%.

**3.6 WideResNet-подобная остаточная нейронная сеть**

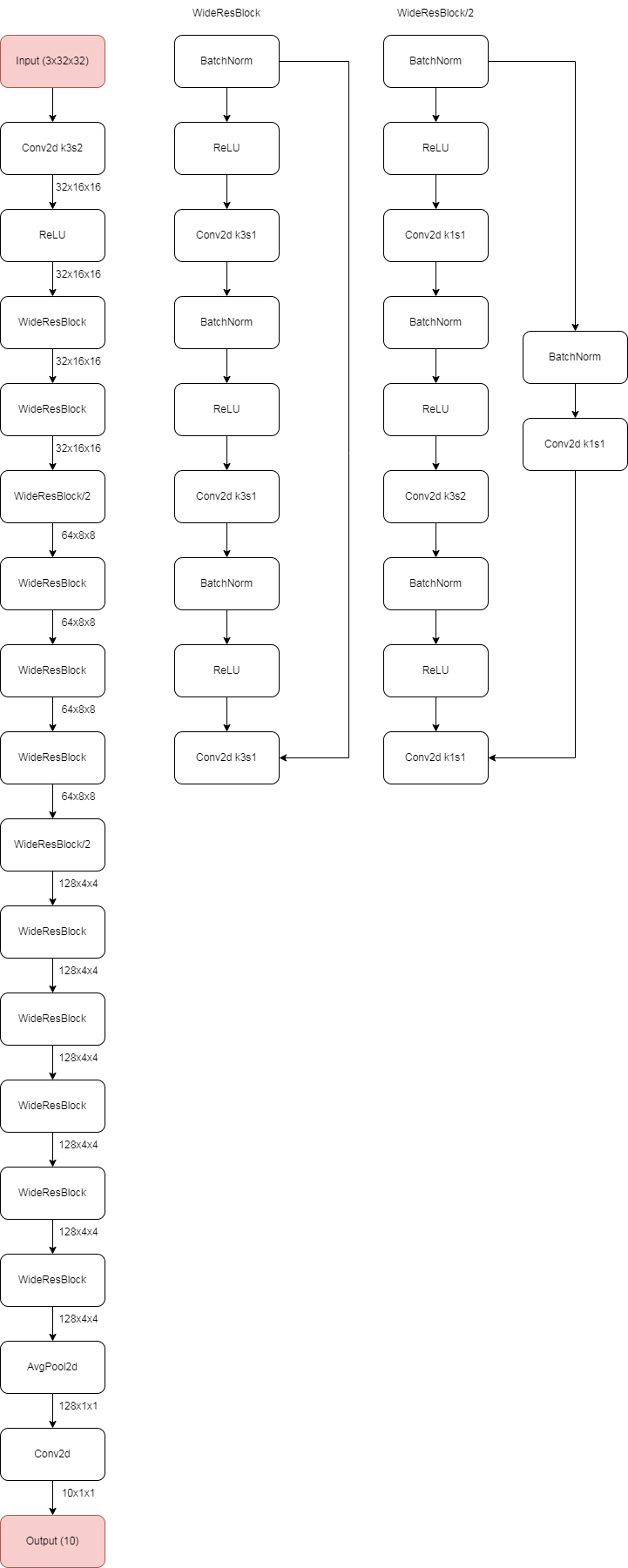
****

Рисунок 3.6.1 – архитектура нейронной сети

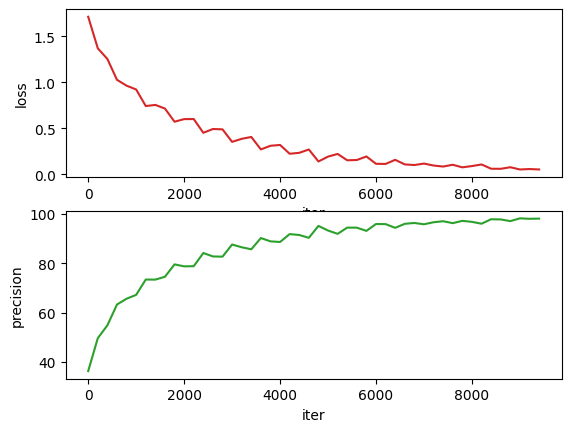
****

Рисунок 3.6.2 – графики метрик качества модели в процессе обучения

**3.7 Выбор наилучшей модели**

Сравним значения метрик качества для обученных моделей.

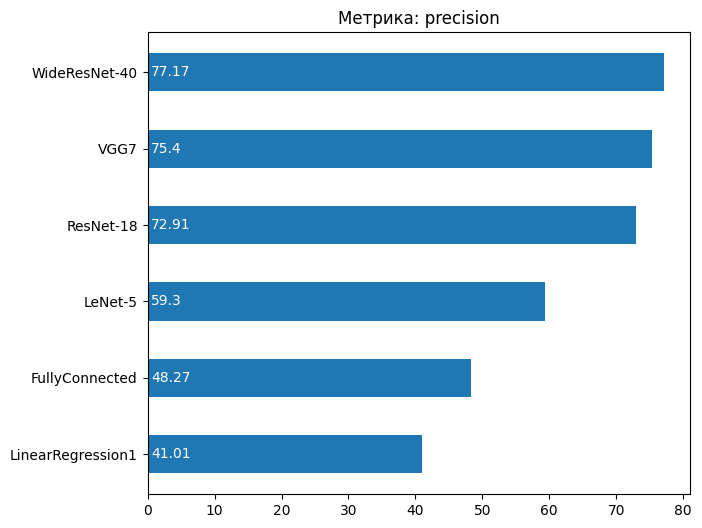


Рисунок 3.7.1 – значения метрики precision

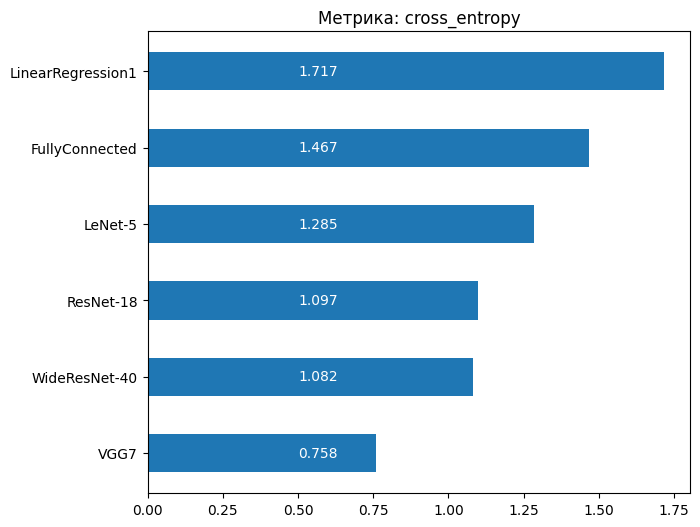


Рисунок 3.7.2 – значения перекрёстной энтропии

Можно видеть, что VGG-подобная и WideResNet-подобная модели работают лучше всего, однако, поскольку значение метрики precision для WideResNet больше, чем для VGG, далее будем работать с этой моделью.

**3.9 Оптимизация алгоритмом Adam**

Вместо подбора гиперпараметров градиентного спуска можно применить оптимизатор Adam.

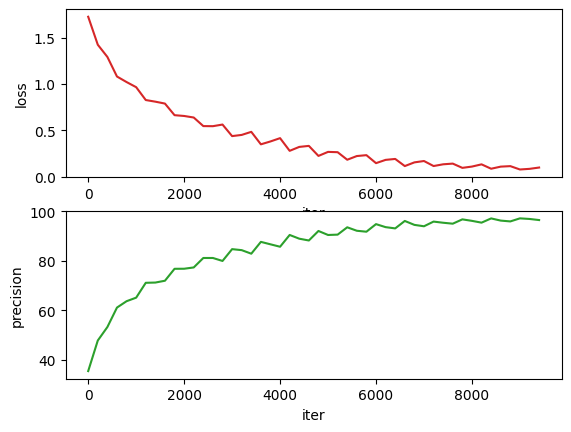


Рисунок 3.7.3 – графики метрик качества модели в процессе обучения

Сравнение полученной модели с baseline дало следующие результаты:

|  |  |
| --- | --- |
| **baseline** | **Best model** |
| precision: 77.17 | precision: 78.52 |
| CrossEntropy: 1.082 | CrossEntropy: 0.951 |

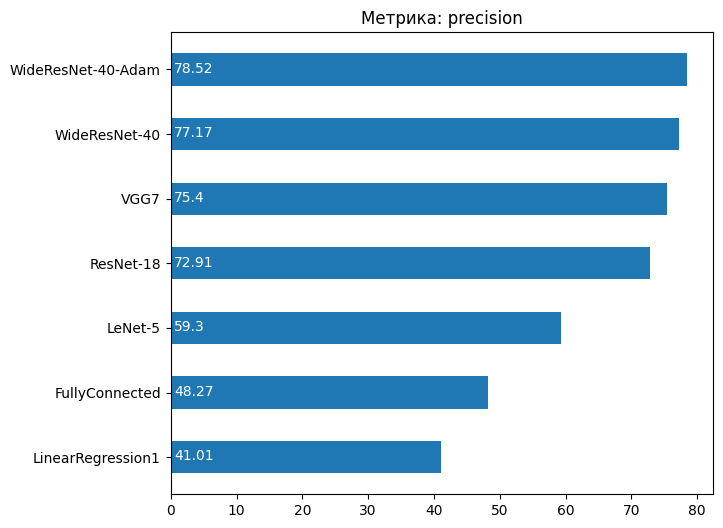


Рисунок 3.7.4 – значения метрики precision

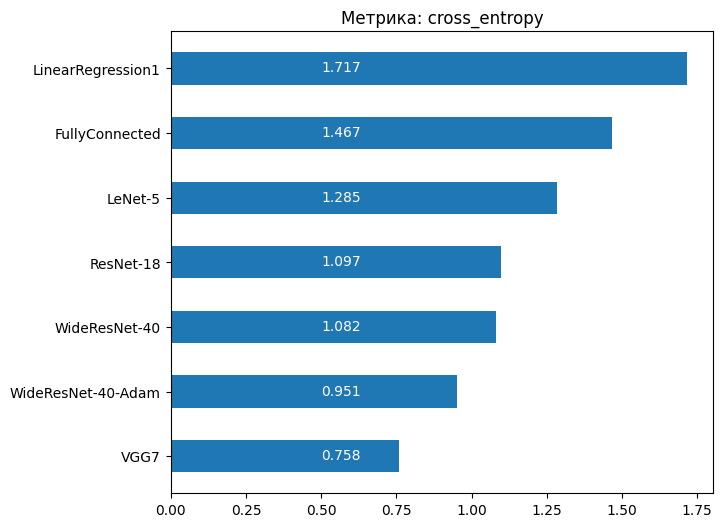


Рисунок 3.7.5 – значения перекрёстной энтропии

# **ЗАКЛЮЧЕНИЕ**

В результате выполнения данной научно-исследовательской работы были выполнены следующие задачи:

1. Были изучены архитектуры полносвязных, свёрточных и остаточных нейронных сетей. Были исследованы возможности алгоритма оптимизации Adam для повышения точности обучения нейронной сети.
2. Была обучена модель для решения задачи CIFAR10

# **СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Alex Krizhevsky, Learning Multiple Layers of Features from Tiny Images: [Электронный ресурс]. // www.cs.toronto.edu. 2009. Дата обновления: 2009. URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf (Дата обращения: 19.05.2023).
2. Yann LeCun, Leon Bottou, Yoshua Bengio and Patrick Haffner, GradientBased Learning Applied to Document Recognition: [doi:10.1109/5.726791]
3. Karen Simonyan, Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition: [https://doi.org/10.48550/arXiv.1409.1556]
4. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition: [https://doi.org/10.48550/arXiv.1512.03385]
5. Sergey Zagoruyko, Nikos Komodakis, Wide Residual Networks: [https://doi.org/10.48550/arXiv.1605.07146]