Министерство образования Республики Беларусь

Учреждение образования

"Белорусский Государственный университет информатики

и радиоэлектроники"

Лабораторная работа №4

“**Реализация приложения по распознаванию номеров домов”**

по учебной дисциплине “Машинное обучение”

|  |  |
| --- | --- |
| Выполнил: | Студент гр. 956241 Дубовик Н.О. |
|  |  |

Минск 2020

**Данные:** Набор изображений из Google Street View с изображениями номеров домов, содержащий 10 классов, соответствующих цифрам от 0 до 9.

* 73257 изображений цифр в обучающей выборке;
* 26032 изображения цифр в тестовой выборке;
* 531131 изображения, которые можно использовать как дополнение к обучающей выборке;
* В двух форматах:
* Оригинальные изображения с выделенными цифрами;
* Изображения размером 32 × 32, содержащих одну цифру;
* Данные первого формата можно скачать по ссылкам:
* <http://ufldl.stanford.edu/housenumbers/train.tar.gz> (обучающая выборка);
* <http://ufldl.stanford.edu/housenumbers/test.tar.gz> (тестовая выборка);
* <http://ufldl.stanford.edu/housenumbers/extra.tar.gz> (дополнительные данные);
* Данные второго формата можно скачать по ссылкам:
* <http://ufldl.stanford.edu/housenumbers/train_32x32.mat> (обучающая выборка);
* <http://ufldl.stanford.edu/housenumbers/test_32x32.mat> (тестовая выборка);
* <http://ufldl.stanford.edu/housenumbers/extra_32x32.mat> (дополнительные данные);
* Описание данных на английском языке доступно по ссылке: <http://ufldl.stanford.edu/housenumbers/>

Результат выполнения заданий опишите в отчете.

**Задание 1.**

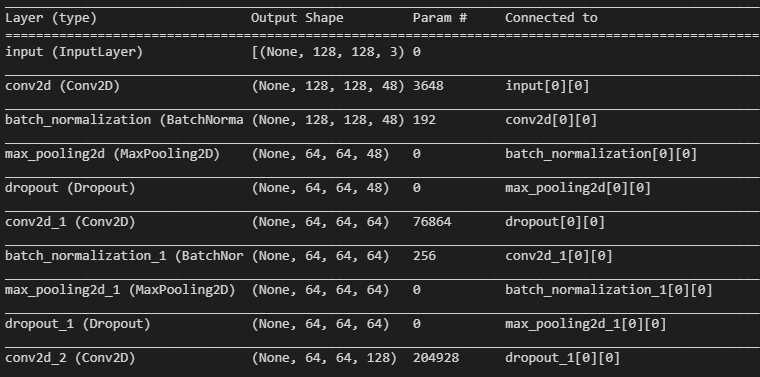
Реализуйте глубокую нейронную сеть (полносвязную или сверточную) и обучите ее на синтетических данных (например, наборы MNIST (http://yann.lecun.com/exdb/mnist/) или notMNIST).

Ознакомьтесь с имеющимися работами по данной тематике: англоязычная статья ([http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/42241.pdf](http://static.googleusercontent.com/media/research.google.com/en/pubs/archive/42241.pdf)), видео на YouTube (https://www.youtube.com/watch?v=vGPI\_JvLoN0).

Для решения поставленной задачи была создана следующая архитектура нейронной сети:

* 8 сверточных слоев;
* 2 полносвязных слоя;
* 6 выходящих слоев;
* первый слой определяет количество цифр в номере дома;
* второй, третий, четвёртый, пятый и шестой выходящий слой определяет первую цифру, вторую, третью, четвёртую и пятую цифру в номере соответственно;
* использовался оптимизатор Adam;
* функция потерь “sparse\_categorical\_crossentropy”.

На рисунке 1 представлена таблица, описывающая получившеюся архитектуру. Данная таблица была получена с помощью метода model.summary().



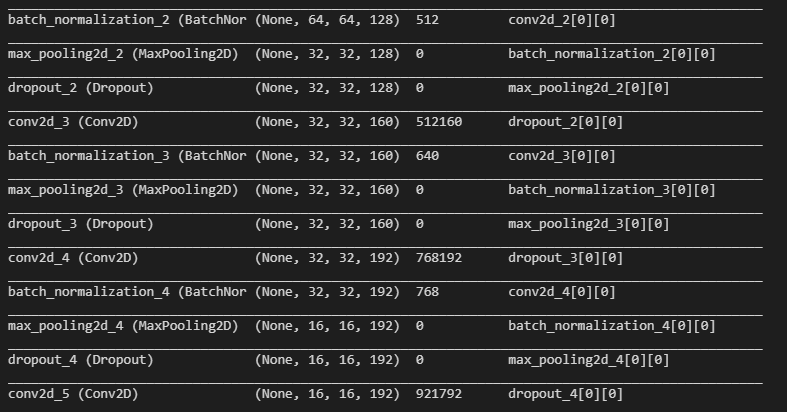
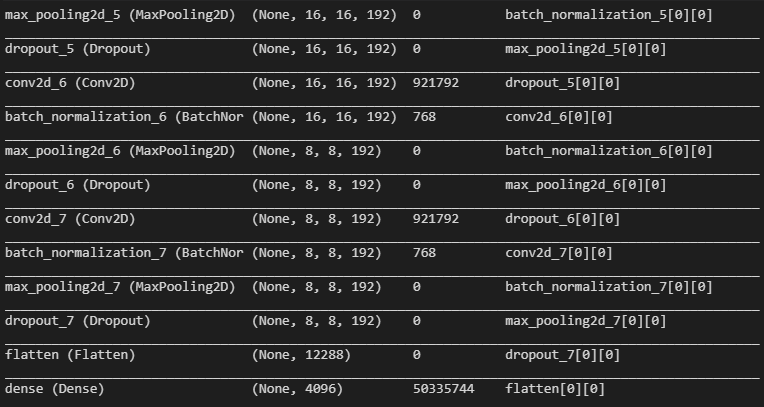


Рисунок 1 – Архитектура созданной нейронной сети



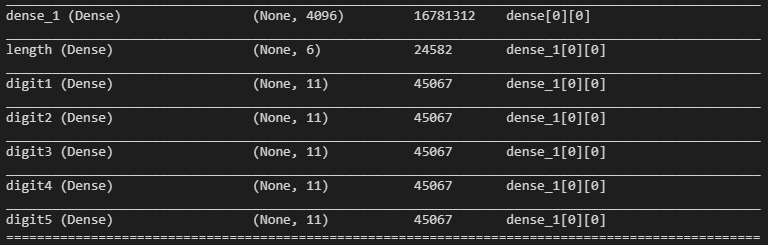


Рисунок 1, лист 2

**Задание 2.**

После уточнения модели на синтетических данных попробуйте обучить ее на реальных данных (набор Google Street View). Что изменилось в модели?

После обучения на реальных данных были получены следующие результаты:

Точность на обучающей выборке составляет 98.64% для первого слоя; 95.60% для второго слоя; 94.35% для третьего слоя; 96.99 % для четвертого слоя; 99.48% для пятого слоя; 99.97% для шестого слоя.

Точность на валидационной выборке составляет 89.47% для первого слоя; 85.35% для второго слоя; 81.51% для третьего слоя; 88.48% для четвертого слоя; 97.72% для пятого слоя; 99.98% для шестого слоя.

Точность на тестовой выборке составляет 80.76% для первого слоя; 68.69% для второго слоя; 62.91% для третьего слоя; 86.03% для четвертого слоя; 99.12% для пятого слоя; 99.98% для шестого слоя.

**Задание 3.**

Реализуйте приложение для ОС Android, которое может распознавать цифры в номерах домов, используя разработанный ранее классификатор. Какова доля правильных классификаций?

Нейронная сеть была использована для распознавания четырех изображений с камеры телефона.

В результате только два изображения из четырех были распознаны верно, рисунок 2.



Рисунок 2 – Распознавание сфотографированных номеров домов