

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
**«Национальный исследовательский**  
**Нижегородский государственный университет им. Н.И. Лобачевского»**  
**(ННГУ)**

**Институт информационных технологий, математики и механики**  
**Кафедра математического обеспечения и суперкомпьютерных технологий**  
Направление подготовки «Прикладная математика и информатика»  
Магистерская программа «Системное программирование»

**Отчет по лабораторной работе**  
**«Разработка сверточной нейронной сети»**

Выполнили:  
студенты группы 381603м4  
Гладилов, Волокитин, Левин,  
Новак

---

Нижний Новгород  
2017

## **СОДЕРЖАНИЕ**

<b>1</b>	<b>ФОРМАТ ВХОДА СЕТИ</b>	<b>3</b>
<b>2</b>	<b>ТЕСТОВЫЕ КОНФИГУРАЦИИ СЕТЕЙ</b>	<b>4</b>
<b>3</b>	<b>РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ</b>	<b>8</b>

.

# 1 Формат входа сети

Для описания входа сети в библиотеке Caffe используется слой ImageData.

```
layer {
  name: "gender"
  type: "ImageData"
  top: "data"
  top: "label"
  include {
    phase: TRAIN
  }
  transform_param {
    scale: 0.00390625
  }
  image_data_param {
    source: "/home/glebg/dev/deep-
learning/train.lst"
    new_width: 150
    new_height: 150
    batch_size: 100
  }
}
```

Описание значений параметров слоя:

- top – Указывает на то, какие данные выходят из слоя, в данном случае это исходная картинка и метка класса
- phase – TRAIN ( TEST) –режим в котором используется слой
- transform\_param – описание преобразований над входными данными. В данном случае выполняется нормировка на 255
- Source – файл \*.lst где хранятся изображения и метки класса
- new\_width/new\_height – размеры входного тензора
- batch\_size – размер пачки картинок.

В данной модели мы используем RGB картинки уменьшенные до размера 150x150 и нормализованные на 255

## 2 Тестовые конфигурации сетей

- Average:

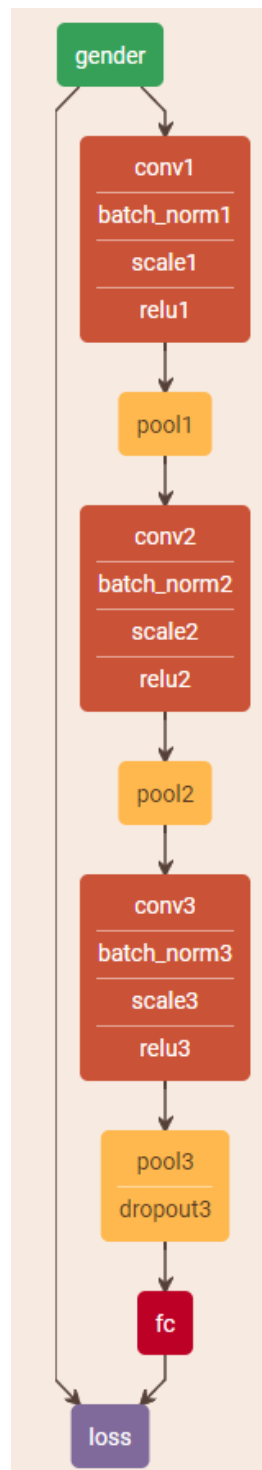


Рис. 1. Сверточная сеть с avg pool

[Визуализация в netScope](#)

- **MaxPool**

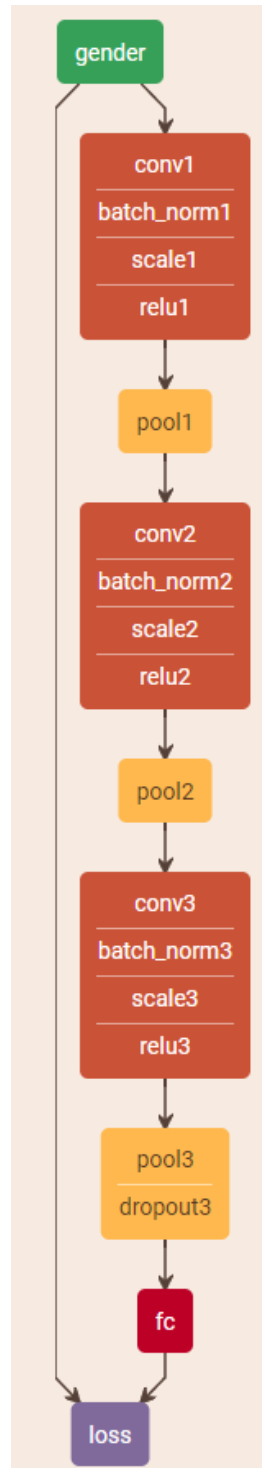
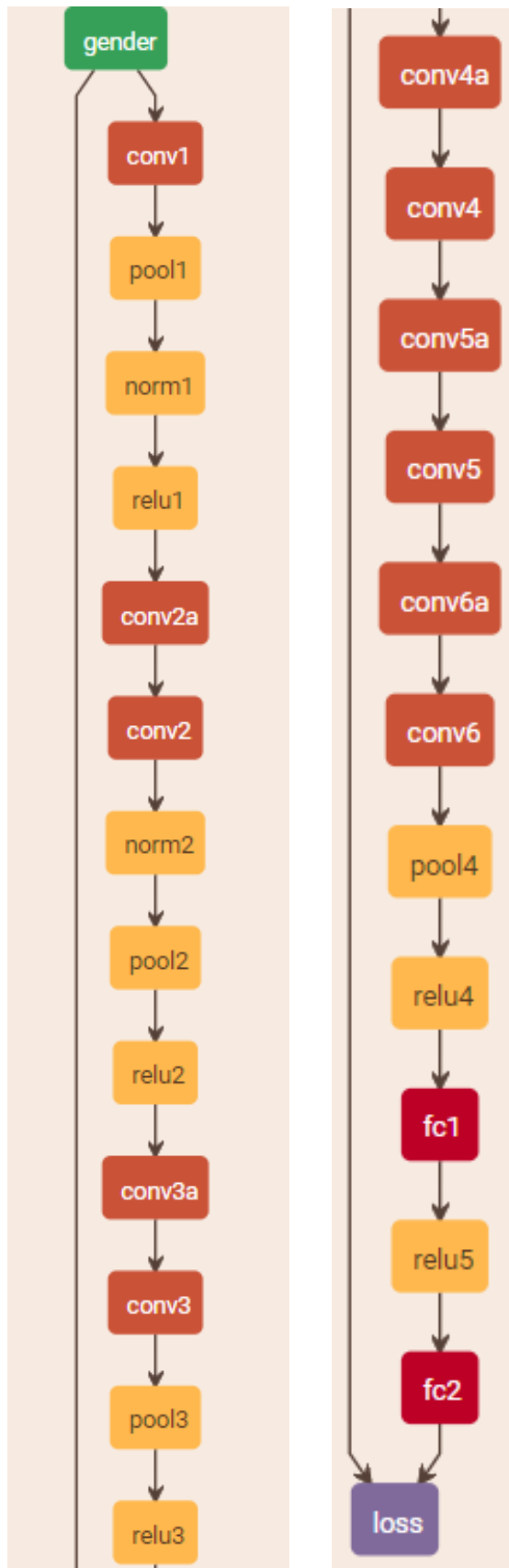


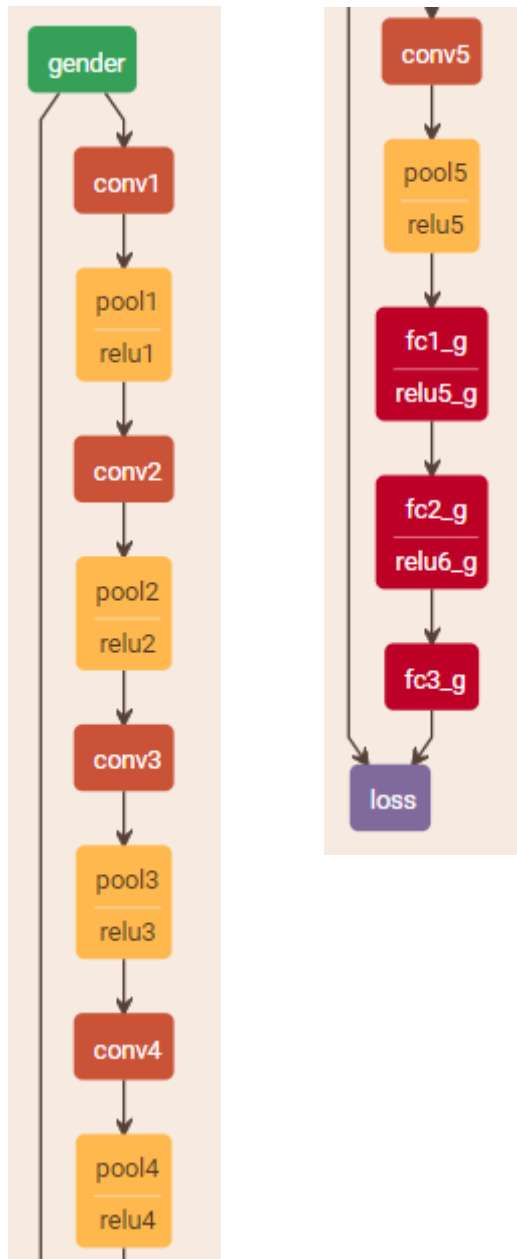
Рис. 2. MaxPool.

[Визуализация в netScope](#)

- FaceNet ([netScope](#))



- Three FC ([netScope](#))



Описание некоторых слоев:

**Scale** – В общем случае, вычисляет произведение двух входных тензоров. В частном случае, второй вход может быть опущен, тогда множитель становится параметром слоя. В нашем случае слой scale будет работать следующим образом:

$$res_{tensor} = input_{tensor} * a + b$$

Так как в параметрах слоя выставлено `bias_term: true`

### 3 Результаты экспериментов

Конфигурация сети	Точность
<b>Avg (10000)</b> (GeForce GTX 1080)	<b>0.908333</b>
<b>Max (10000)</b> (GeForce GTX 1080)	<b>0.9004</b>
<b>FaceNetWithoutNorm (12000)</b> (Tesla K20X)	<b>0.84</b>
<b>Three FC (12000)</b> (Tesla K20X)	<b>0.84</b>



## **4 Итоги**

В лабораторных работах нами было рассмотрено семейство сверточных нейронных сетей. Данный тип показал высокий результат по точности классификации пола человека. Максимально достигнутый результат составил 90% , что на 10% выше, чем точность решения аналогичной задачи с использованием полностью связанной сети.