

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

Направления подготовки: «Прикладная математика и информатика», «Фундаментальная информатика и информационные технологии»

Магистерские программы: «Системное программирование», «Компьютерная графика и моделирование живых и технических систем»

Отчет по лабораторной работе
«Разработка сверточной нейронной сети с помощью библиотеки
MXNet»

Выполнили:
студенты группы 381606м2
Сахаров Алексей
Шерстнева Анна
студенты группы 381603м4
Бастракова Ксения
Семичев Юрий

Нижний Новгород
2017

Оглавление

Цель работы.....	3
Задачи.....	3
Схемы конфигураций сетей.....	6
Тренировочные и тестовые данные	8
Показатели качества	8
Результаты экспериментов	9

Цель работы

Разработать сверточную нейронную сеть используя библиотеку MXNet.

Задачи

- Разработать нескольких архитектур сверточных нейронных сетей в формате, который принимается библиотекой.
- Обучить разработанную нейронную сеть
- Протестировать разработанную нейронную сеть

Решаемая задача

В качестве тестовой задачи для обучения полностью связанной нейронной сети использовался набор данных FERF DB. Он представляет из себя набор из 55767 цветных изображений шести анимационных персонажей. Эти изображения являются поккадровой нарезкой созданных в программе MAYA анимаций, изображающих одну из семи эмоций: гнев, отвращение, страх, радость, нейтральное выражение, печаль и удивление. Все изображения размечены именами персонажей и эмоциями, которые они выражают на том или ином изображении.

На основе данного набора размеченных изображений решалась задача распознавания эмоций персонажей, которые они на этих изображения выражают.

Ниже представлены примеры тренировочных и тестовых изображений. Порядок эмоций: злость – отвращение – страх – радость – нейтральное – грусть - удивление:

Рис.1 персонаж Aia. Все эмоции



Рис.2 персонаж Bonnie. Все эмоции



Рис.3 персонаж Jules. Все эмоции



Рис.4 персонаж Malcolm. Все эмоции



Рис.5 персонаж Mery. Все эмоции



Рис.6 персонаж Ray. Все эмоции



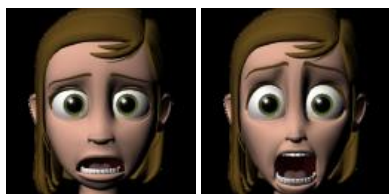
Рис.7 персонаж Aia. Максимальное различие изображений одной эмоции
а) – злость, б) - отвращение, в) – страх, г) – нейтральное, д) – радость, е) – грусть, ж) - удивление



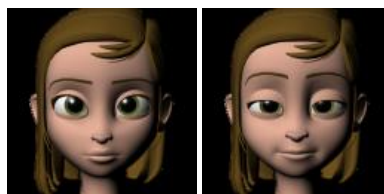
А)



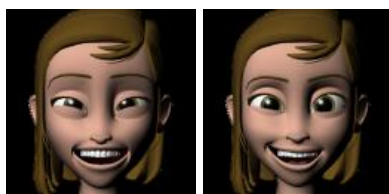
Б)



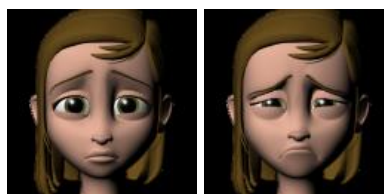
В)



Г)



Д)



Е)



Ж)

Схемы конфигураций сетей

Ниже приведены используемые конфигурации нейронных сетей с двумя и тремя сверточными слоями. Ядра сверточных слоев выбраны следующих размеров:

- 7x7
- 5x5
- 3x3

Также в свёрточных слоях фильтр не обязательно «прикладывался» к каждому фрагменту картинки, а шёл по картинке с некоторым шагом по пикселям, который называется stride'ом. Использовались следующие значения stride:

- (4, 4)
- (2, 2)
- (1, 1)

В каждом свёрточном слое к картинке может применяться не один фильтр, а несколько (то есть одна картинка на выходе со слоя может превратиться в несколько). Во всех архитектурах использовалось по 500 фильтров на каждом слое.

После каждой свёртки применялась функция активации (во всех используемых нами архитектурах это был арктангенс), а после активации – операция пространственного объединения (pooling). Ещё одной парой параметров слоя являлись величина шаблона и stride пространственного объединения. Они выбирались из того же списка, что и размер фильтра и stride свёртки.

После свёрточных слоёв следует как минимум один полносвязный слой.

Активационная функция на полносвязных слоях выбирается одной из следующих:

- Sigmoid: $f = \frac{1}{1+e^{-s_j}}$
- Relu: $f = \max(x, 0)$
- Tanh: $f = \arctan(x)$

Активационной функцией на выходном полносвязном слое является Softmax: $f = \frac{e^{s_j}}{\sum_{j=1}^n e^{s_j}}$

Рис. 1 Нейронная сеть с 2-мя сверточными и 2-мя полносвязными слоями

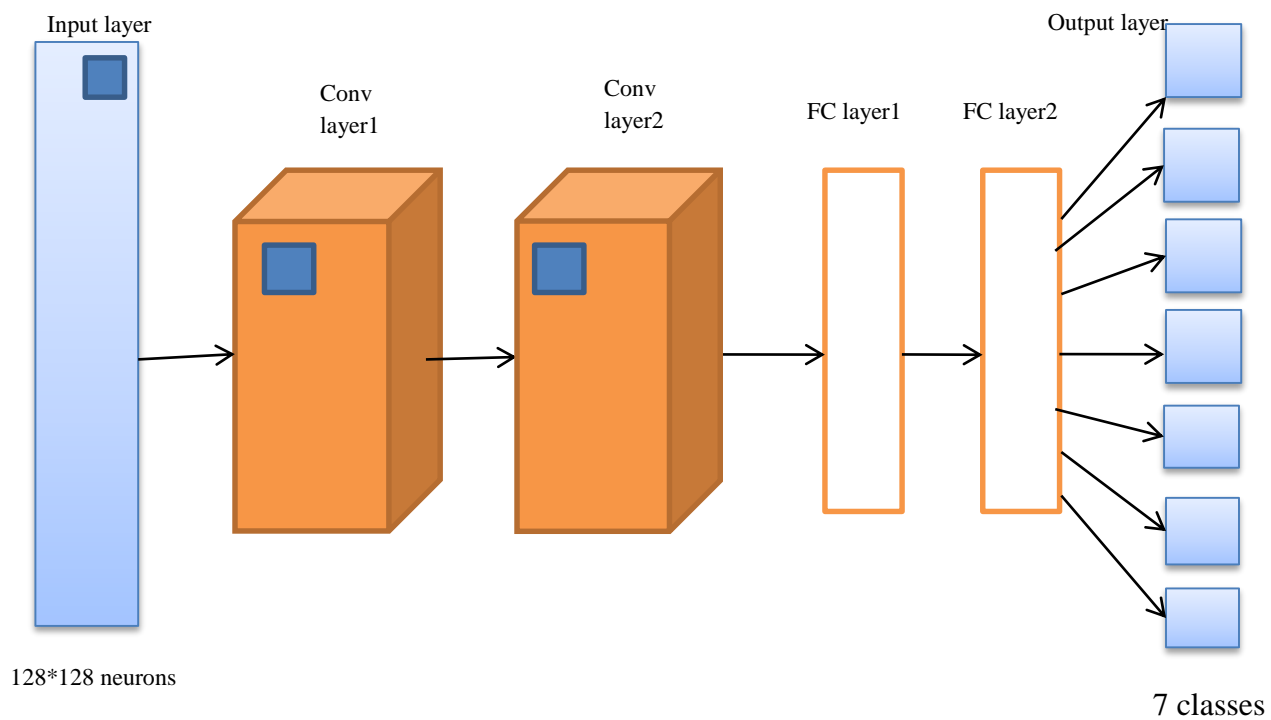
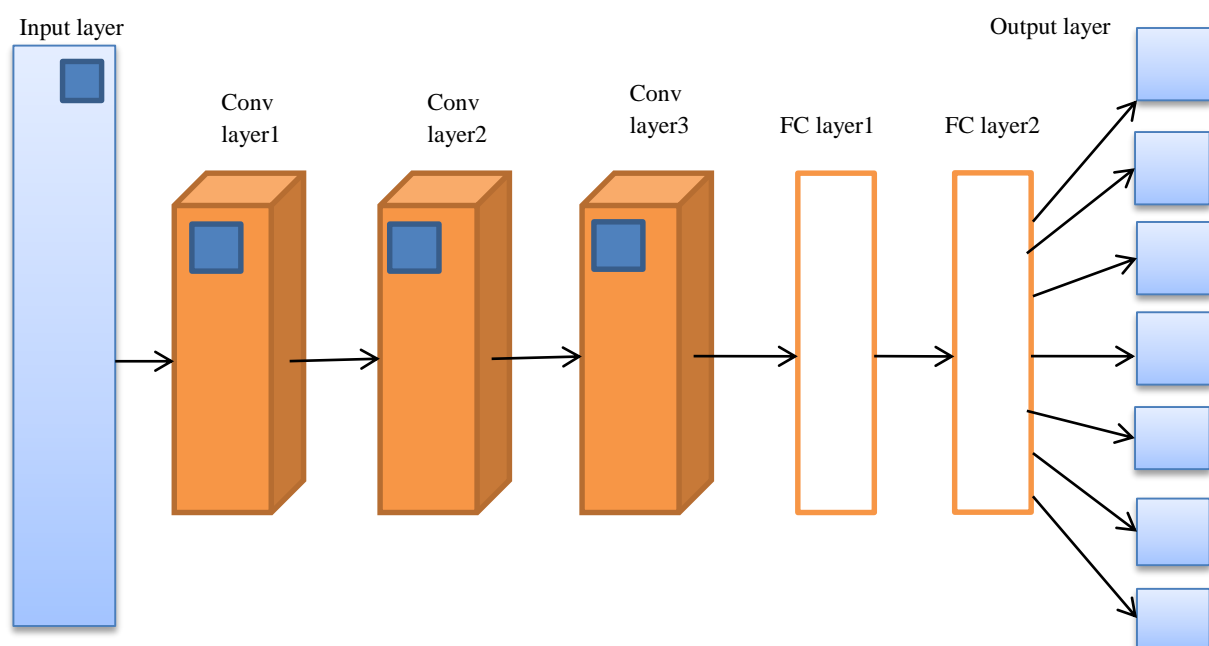


Рис. 2 Нейронная сеть с 3-мя сверточными и 2-мя полносвязными слоями



Тренировочные и тестовые данные

В качестве тренировочного набора данных используем 80% от всего набора изображений с эмоциями 6 персонажей. В качестве тестового набора – 20% всех изображений (оставшиеся). Такое разделение выбрано потому, что разнообразие входных данных невелико – лица персонажей смоделированы с помощью 3-D инструмента, многие изображения набора данных отличаются друг от друга незначительно.

Показатели качества

В качестве метрики точности решения используем отношение угаданных эмоций ко всем в тестовой выборке, т.е :

$$Accuracy = \frac{Correctly\ answers\ count}{Images\ count}$$

Результаты экспериментов

Результаты экспериментов представлены ниже в Таблице 1

Число сверточных слоев	Число полносвязных слоев	Конфигурация	Функции активации	Точность тестирования	Среднее время тестирования (сек)
3	2	C7x7-P5x5-C5x5-P3x3-C3x3-P3x3-512-7	tanh	0.8136	798
3	3	C7x7-P5x5-C5x5-P3x3-C3x3-P3x3-1024-512-7	tanh	0.6269	855