

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 полностью связанной 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
Решаемая задача.....	4
Схемы конфигураций сетей.....	6
Тренировочные и тестовые данные.....	8
Показатели качества.....	10
Результаты экспериментов.....	11
Вывод.....	12

Цель работы

Получить базовые навыки работы с библиотекой MXNet и разработать полносвязную нейронную сеть.

Задачи

- Установить библиотеку MXNet и проверить корректность установки.
- Выбрать практическую задачу распознавания образов для выполнения практических работ
- Разработать программы/скрипты для подготовки тренировочных и тестовых данных в формате, который обрабатывается библиотекой.
- Разработать несколько архитектур полностью связанных нейронных сетей в формате, который принимается библиотекой.
- Обучить и протестировать сеть

Решаемая задача

В качестве тестовой задачи для обучения полностью связанной нейронной сети использовался набор данных FERG DB. Он представляет из себя набор из 55767 цветных изображений шести анимационных персонажей. Эти изображения являются покадровой 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



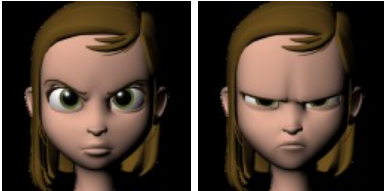
Рис.4 Персонаж Malcolm. Все эмоции



Рис.5 Персонаж Mery. Все эмоции



Рис.6 Персонаж Ray. Все эмоции



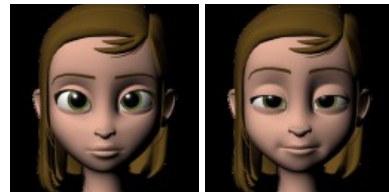
А)



Б)



В)



Г)



Д)



Е)



Ж)

Рис.7 Персонаж Aia. Максимальное различие изображений одной эмоции

а) – злость, б) - отвращение, в) – страх, г) – нейтральное, д) – радость, е) – грусть, ж) - удивление

Схемы конфигураций сетей

Ниже представлены типичные схемы конфигураций полносвязных нейронных сетей с 2-мя и 3-мя скрытыми слоями. Активационная функция на скрытых слоях выбирается одной из следующих:

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

Активационной функцией на выходном слое является Softmax:

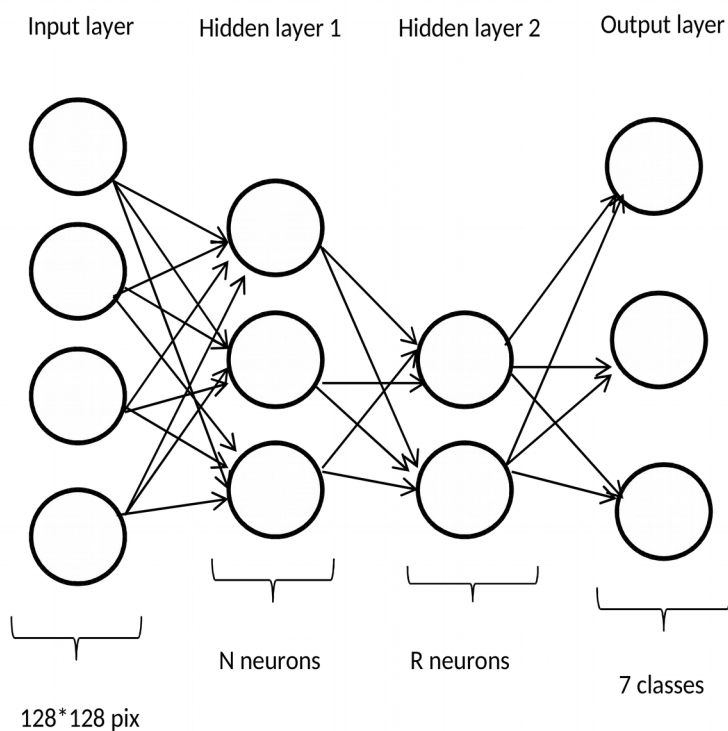
$$f = \frac{e^{s_j}}{\sum_{j=1}^n e^{s_j}}$$


Рис.1 Полносвязная нейронная сеть с 2-мя скрытыми слоями

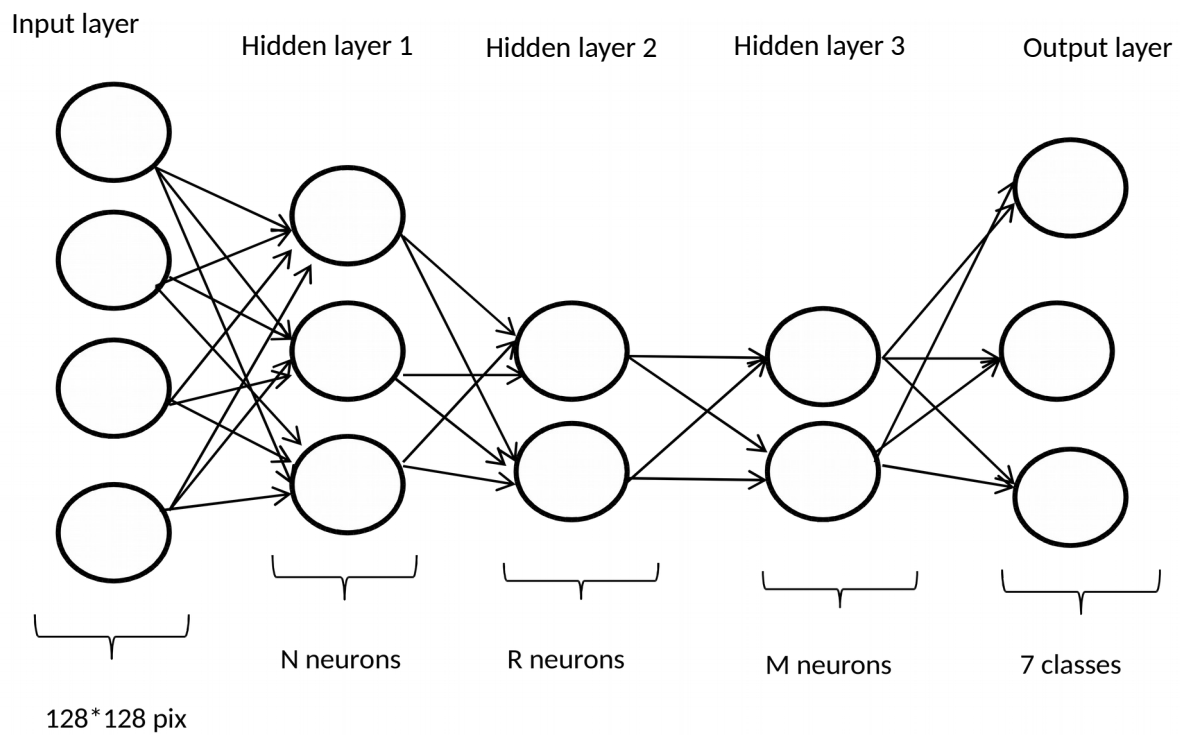


Рис.2 Полносвязная нейронная сеть с 3-мя скрытыми слоями

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

Таблица 1. Результаты экспериментов

Число скрытых слоев	Кол-во нейронов на скрытом слое	Функции активации	Точность тестирования	Среднее время тестирования (сек)
1	150	sigmoid	0,1696	
1	150	relu	0,1696	
1	250	sigmoid	0,1697	
1	250	relu	0,1696	
1	2000	sigmoid	0,2509	460
1	2000	relu	0,2493	434
2	100-100	sigmoid-sigmoid	0,1696	
2	100-100	relu-relu	0,1701	
2	2000-800	sigmoid-sigmoid	0,5389	785
3	100-100-100	sigmoid-relu-sigmoid	0,1306	
3	2500-1000-500	tanh-tanh-tanh	0,2119	1017

Выводы

Поподробнее остановимся на результатах экспериментов.

В целом можно сказать, что полносвязные нейронные сети плохо справляются с задачей распознавания эмоций (максимальная точность предсказания всего лишь 0.54). Сети с маленьким числом нейронов на каждом слое (до 500) совсем не справляются с задачей –

точность предсказания лишь немногим больше $\frac{1}{7} \approx 0.14$, то есть примерно такой же результат даёт случайное предсказание. Тем не менее, некоторые архитектуры сетей показали себя лучше, чем остальные. Среди рассматриваемых моделей с достаточно большим числом нейронов на слоях наилучший результат показала двухслойная сеть – её точность предсказания оказалась более чем в 2 раза. Это выше, чем у однослойной и трехслойной сети.

Наиболее вероятное объяснение низкой точности предсказания полносвязных сетей – их неспособность выявлять устойчивые признаки, как это делает, например, свёрточная нейронная сеть, способная акцентировать эти «фичи» с помощью свёрточных фильтров, а затем сохранить и передать эти значения дальше, выделив их из незначущей информации с помощью, например, max-pooling-операции.