

**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

**«Начальная настройка весов полностью связанных и сверточных
нейронных сетей 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	7
Показатели качества	7
Результаты экспериментов	8

Цель работы

Использовать методы обучения без учителя для настройки начальных значений весов сетей. Использовать полученные веса в полностью связанных и сверточных нейронных сетях, разработанных ранее, используя библиотеку MXNet.

Задачи

- Выбрать несколько архитектур сверточных и полностью связанных нейронных сетей, разработанных ранее.
- Выбрать и применить метод обучения без учителя для выполнения настройки начальных весов сетей.
- Обучить выбранные нейронные сети.
- Протестировать выбранные нейронные сети и собрать результаты.

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



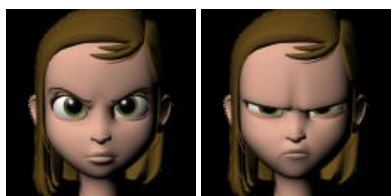
Рис.4 Персонаж Malcolm. Все эмоции



Рис.5 Персонаж Mery. Все эмоции



Рис.6 Персонаж Ray. Все эмоции



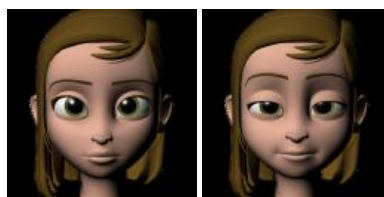
А)



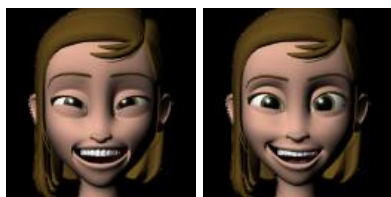
Б)



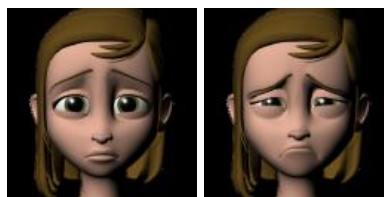
В)



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

Для каждого слоя нейронной сети строилась двухслойная нейронная сеть, состоящая из кодирующего слоя (рассматриваемый слой исходной сети) и декодирующего слоя (слой, обратный к кодирующему). Этой нейронной сети на вход подавался вход соответствующего слоя исходной сети, и сеть обучалась тождественному преобразованию этого входа (то есть ожидаемый выход равен входу). После обучения веса с кодирующего слоя переносятся на соответствующий слой исходной сети, а выход с кодирующего слоя подаётся на вход следующему слою, и процедура настройки весов повторяется для следующего слоя.

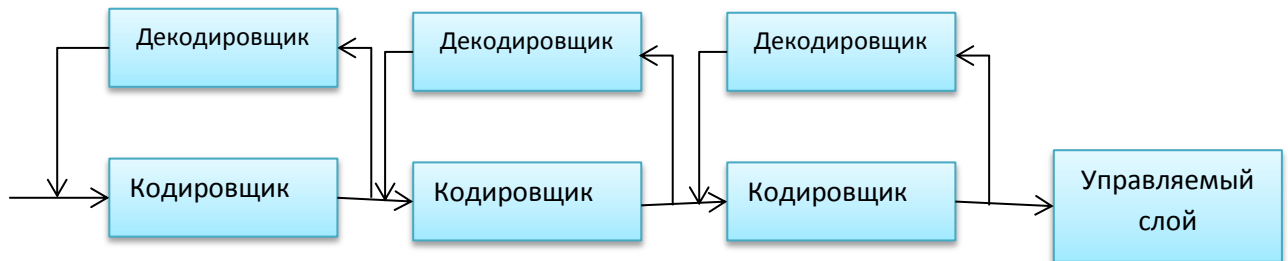


Рис.1 Стек автокодировщиков

После этой процедуры запускается обучение обычной полносвязной нейронной сети с настроенными в ходе процедуры весами.

Ниже представлена схема конфигурации полносвязной нейронной сети с 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: $f = \frac{e^{s_j}}{\sum_{j=1}^n e^{s_j}}$

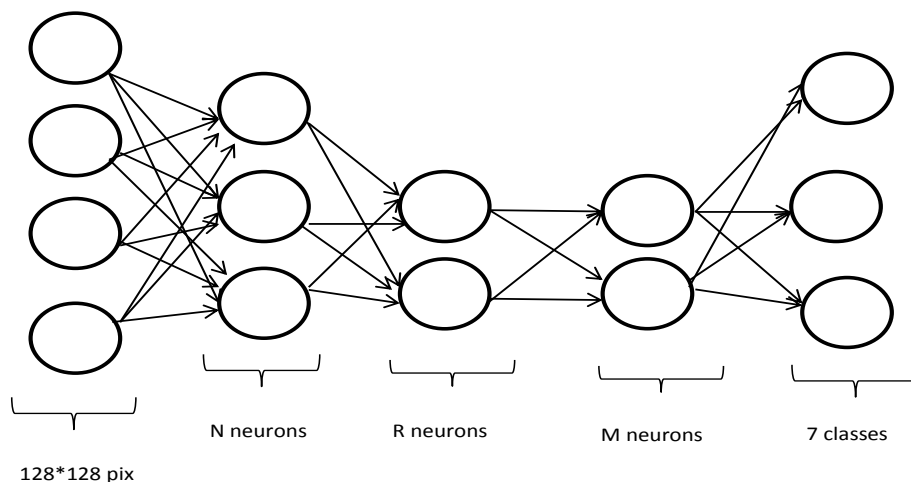


Рис.2 Полносвязная нейронная сеть с 3-мя скрытыми слоями

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

Число скрытых слоев	Кол-во нейронов на скрытом слое	Функции активации	Точность тестирования	Среднее время тестирования (сек)
1	2000	sigmoid		
1	2000	relu		
2	2000-800	sigmoid-sigmoid		
3	2500-1000-500	tanh-tanh-tanh		
2	2000-1000	sigmoid-sigmoid		