

**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

**«Применение переноса обучения для решения задачи
распознавания эмоций персонажей 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	8
Результаты экспериментов	9

Цель работы

Исследование возможности переноса обучения для решения задачи распознавания эмоций персонажей, используя библиотеку MXNet.

Задачи

- Сформулировать исходную задачу близко к целевой — «распознавание эмоций персонажей».
- Найти натренированную модель для решения исходной задачи.
- Выполнить 4 типа экспериментов по переносу знаний.
- Собрать результаты экспериментов.

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 покaдровой 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



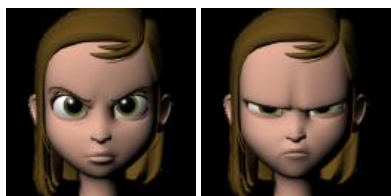
Рис.4 Персонаж Malcolm. Все эмоции



Рис.5 Персонаж Mery. Все эмоции



Рис.6 Персонаж Ray. Все эмоции



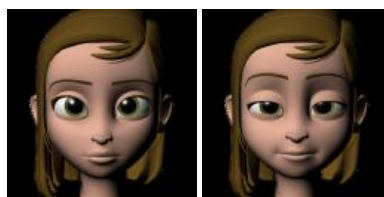
А)



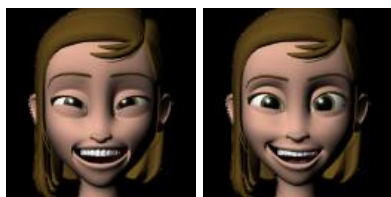
Б)



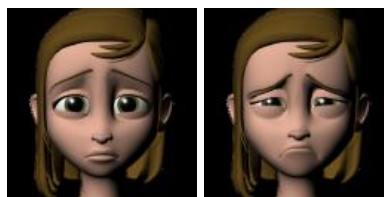
В)



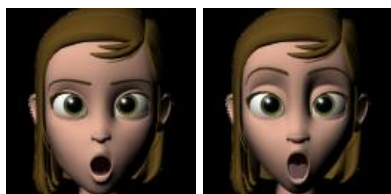
Г)



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

В качестве исходной задачи была выбрана задача распознавания рукописных цифр по изображениям из набора данных MNIST, а в качестве объекта переноса знаний была выбрана сеть LeNet-5, решающая эту задачу.

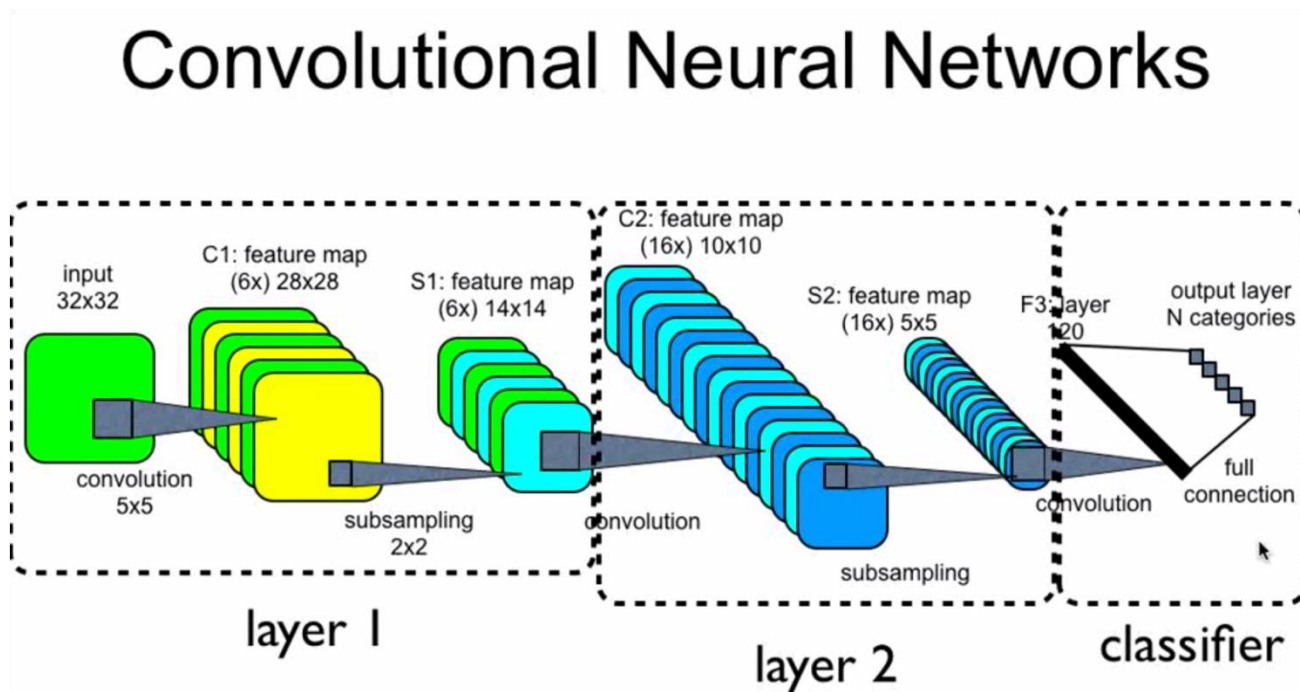


Рис.8. Сверточная сеть LeNet-5

Были использованы следующие методы переноса знаний:

- 1) Прямое использование модели LeNet-5, обученной для задачи распознавания рукописных цифр, для решения задачи распознавания эмоций

Так как размер входных и выходных данных у исходной и решаемой задач не совпадают, веса переносились только на связи между внутренними слоями сети. Для целевой задачи веса после входного слоя и перед выходным слоем дообучались.

- 2) Использование структуры нейронной сети LeNet-5 для решения задачи распознавания эмоций

Нейронная сеть LeNet-5 полностью обучалась для решения задачи распознавания эмоций.

- 3) Использование модели LeNet-5, построенной для решения исходной задачи, в качестве фиксированного метода извлечения признаков при построении модели, решающей задачу распознавания эмоций.

Для решения задачи распознавания эмоций использовалась видоизменённая структура сети LeNet-5. Структура свёрточных слоёв осталась прежней, но единственный полносвязный слой из 120 нейронов сети LeNet был заменён на 3 полносвязных слоя конфигурации 1000-500-200.

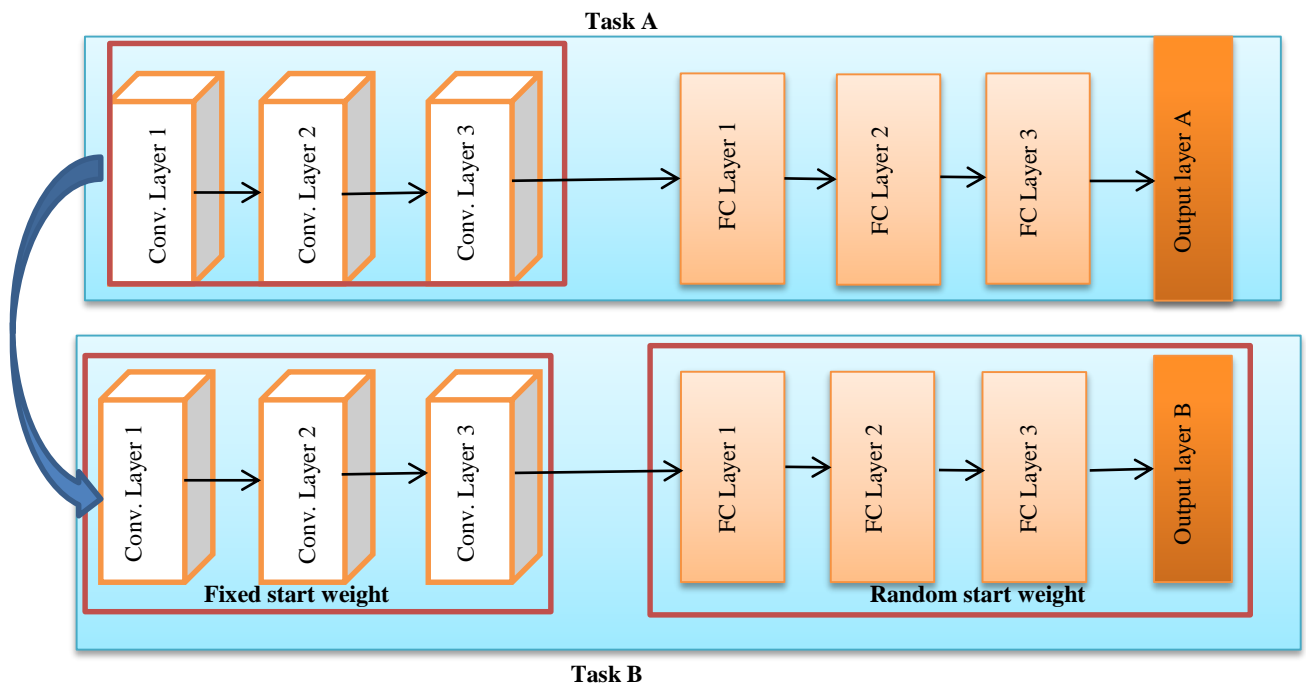


Рис.9. Схема использования модели, построенной для решения исходной задачи, в качестве фиксированного метода извлечения признаков при построении модели, решающей целевую задачу

4) Тонкая настройки параметров модели LeNet-5, построенной для решения задачи распознавания рукописных цифр

Использовалась структура сети описанная в предыдущем эксперименте, но веса для свёрточных слоёв переносились с модели LeNet-5, обученной на решение задачи распознавания цифр. Далее происходило дообучение всех весов сети.

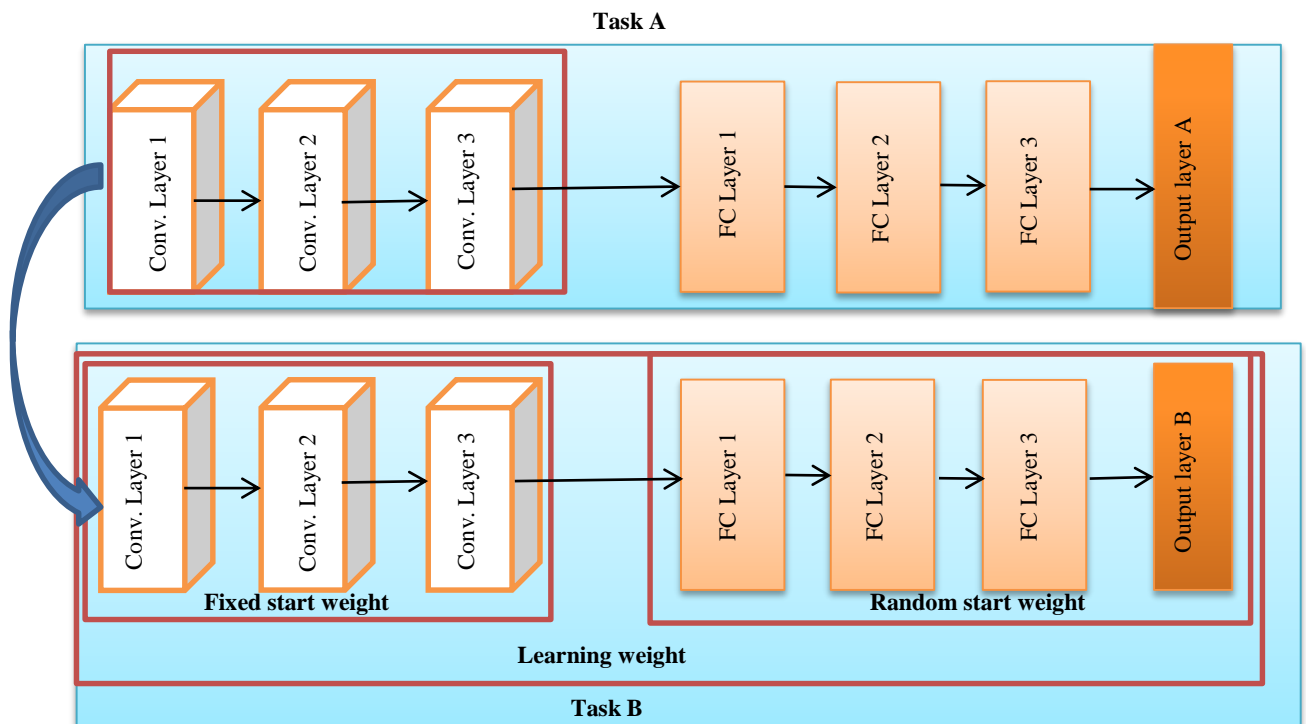


Рис.2. Схема тонкой настройки параметров модели, построенной для решения исходной «Задачи А», с целью решения целевой «Задачи В»

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

Эксперимент	Точность на тестовой выборке	Среднее время тестирования (сек)
1	0.157	189,45
2	0.132	190.16
3	0.715	477.38
4	0.157	461.37

Таблица 1. Результаты экспериментов

Выводы

По Таблице 1 видно, что успешным был только третий эксперимент – когда использовалась структура свёрточных слоёв сети LeNet-5, а структура полносвязных слоёв была изменена под решение целевой задачи. Плохой результат в первых двух экспериментах можно объяснить тем, что 120 нейронов в последнем полносвязном слое сети не хватает для того, чтобы извлечь ключевые признаки из результата свёрток, так как решаемая задача сложнее, чем задача распознавания рукописных цифр. Плохой результат в 4 эксперименте можно объяснить тем, что при решении задачи распознавания рукописных цифр свёрточные слои выделили признаки отличные от тех, которые они выделили бы для решения задачи распознавания эмоций. Таким образом, во время решения задачи распознавания эмоций с перенесёнными начальными весами на свёрточных слоях, функция ошибки могла попасть в окрестность локального минимума, далёкого от глобального и дообучение не выводит сеть из этой окрестности.