

УДК 004.931

## ПРОГРАММНАЯ МОДЕЛЬ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

Умяров Н.Х., Костецкая Г.Ю., Федяев О.И.

Донецкий национальный технический университет

кафедра прикладной математики и информатики

E-mail: nailumyarov@mail.ru

### Аннотация

Умяров Н.Х., Костецкая Г.Ю., Федяев О.И. Программная модель свёрточной нейронной сети для распознавания лиц. На основании ранее составленных параметрической, логической и физической моделей свёрточной нейронной сети разработана программная модель сети.

### Постановка задачи

В работах [1, 2] были составлены параметрическая, логическая и физическая модели свёрточной нейронной сети. В данной статье речь пойдет в первую очередь о программной реализации свёрточной нейронной сети и связанных с ней проблемах.

### Архитектура свёрточной нейронной сети

Архитектура свёрточной нейронной сети, которая применялась для распознавания лиц, состоит из пяти слоёв (рис.1). В сетях такого типа используются слои двух типов: свёрточные (Convolutional, C-Layer) и подвыборочные (Subsampling, S-Layer). Свёрточные и подвыборочные слои чередуются друг с другом.

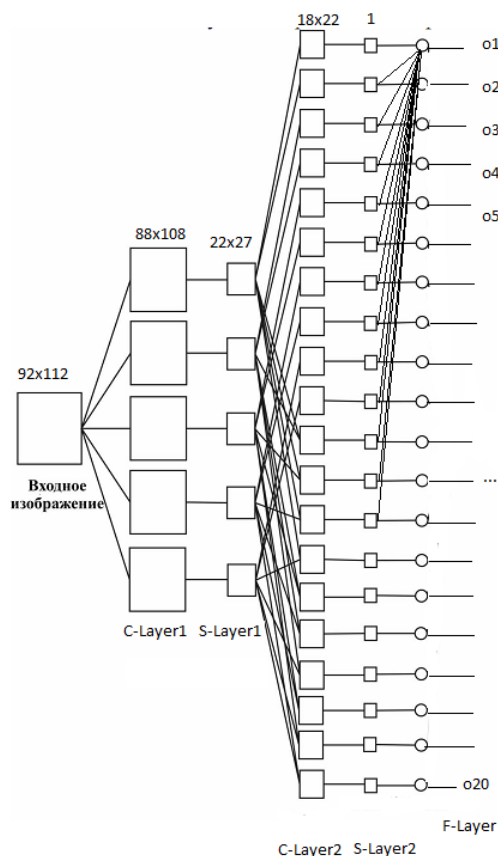


Рисунок 1 - Архитектура свёрточной нейронной сети

В каждом слое имеется набор из нескольких плоскостей, причём нейроны одной плоскости имеют одинаковые веса, ведущие ко всем локальным участкам предыдущего слоя. Изображение предыдущего слоя сканируется небольшим окном и пропускается сквозь набор весов, а результат отображается на соответствующий нейрон текущего слоя.

Входной слой размером 92 на 112 нейронов не несёт какой-либо функциональной нагрузки и служит лишь для подачи входного образа в нейронную сеть.

Следом за входным слоем находится свёрточный слой C-Layer1. Каждый нейрон в плоскости свёрточного слоя получает свои входы от некоторой подобласти предыдущего слоя (локальное рецептивное поле), пропускает их сквозь набор весов и результат отображается на соответствующий нейрон свёрточного слоя.

Процесс функционирования нейрона свёрточного слоя задается выражением:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^K \sum_{t=1}^K w_{k,s,t} x^{((i-1)+s, (j-1)+t)},$$

где  $y_k^{(i,j)}$  – нейрон  $k$ -ой плоскости свёрточного слоя;  $b_k$  – нейронное смещение  $k$ -ой плоскости;  $K$  – размер рецептивной области нейрона;  $w_{k,s,t}$  – элемент матрицы синаптических коэффициентов;  $x$  – выходы нейронов предыдущего слоя.

Слой C-Layer1 состоит из 5 свёрточных плоскостей и выполняет свёртывание входного изображения с помощью маски размером 5 на 5, таким образом, слой C-Layer1 осуществляет 5 свёрток входного изображения.

Размер свёрточной плоскости определяется в соответствии со следующими выражениями:

$$\begin{aligned} w_c &= w_u - K + 1, \\ h_c &= h_u - K + 1, \end{aligned}$$

где  $w_c$ ,  $h_c$  – ширина и высота свёрточной плоскости соответственно;  $w$ ,  $h$  – ширина и высота плоскости предыдущего слоя;  $K$  – ширина (высота) окна сканирования.

Из выражений для расчёта размера свёрточной плоскости следует, что размер плоскости свёрточного слоя C-Layer1 – 88 на 108 нейрона. Нейроны в слое сгруппированы в плоскости, в пределах которых все нейроны имеют один и тот же набор синаптических коэффициентов. Набор выходных сигналов в такой плоскости называют картой характеристик.

Следующий за слоем C-Layer1 подвыборочный слой S-Layer1 состоит из 5 карт характеристик и обеспечивает локальное усреднение и подвыборку. Этот слой также состоит из плоскостей, количество которых такое же, как и в предыдущем слое. Рецепторная область каждого нейрона – область 4x4 в соответствующей карте особенностей предыдущего слоя.

Каждый нейрон вычисляет среднее его входов, умножает на синаптический коэффициент, добавляет нейронное смещение и передаёт результат через активационную функцию. Процесс функционирования нейрона подвыборочного слоя определяется следующим соотношением:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^2 \sum_{t=1}^2 x^{((i-1)+s, (j-1)+t)}.$$

В заключении полученный результат подвыборки проходит через активационную функцию.

После операции подвыборки, точное местоположения и специфические признаки каждой особенности изображения становятся менее важными, что придаёт нейронной сети довольно большую степень инвариантности. В качестве активационной функции используется гиперболический тангенс. Каждая плоскость слоя S-Layer1 связана лишь с одной плоскостью слоя C-Layer1. Размер каждой плоскости слоя S-Layer1 – 22x27 нейронов,

что 4 раза меньше чем размер плоскости предыдущего слоя. Каждая плоскость слоя S-Layer 1 имеет единственный синаптический коэффициент и нейронное смещение.

Свёрточный слой C-Layer2 состоит из 20 плоскостей, слои S-Layer1 и C-Layer2 перекрестно связаны. Плоскости слоя C-Layer2 формируются следующим образом: каждая из 5 плоскостей слоя S-Layer1 свёрнута двумя различными синаптическими масками 5x6, обеспечивая 10 плоскостей в C-Layer2, другие 10 плоскостей C-Layer2 получены, суммируя результаты двух свёртываний на каждой возможной паре плоскостей слоя S-Layer1. Таким образом, свёрточной сети добавляется способность объединять различные виды характеристик, чтобы менее зависеть от искажений входного изображения [3].

Размер плоскости слоя C-Layer2 – 18 x 22 нейронов.

Слой S-Layer2 состоит из 20 плоскостей. Каждая плоскость слоя S-Layer2 имеет единственный синаптический коэффициент и нейронное смещение, что даёт в итоге 40 настраиваемых параметров.

На последнем этапе нейронный слой из 20 нейронов (по числу классов) получает на вход вектор из значений, полученных на предыдущем этапе, и непосредственно классифицирует изображение.

Как и в случае многослойного персептрона, начальные веса задаются случайным образом. Обучение производится по стратегии «обучение с учителем», которая модифицирована с учётом неполноты связи в свёрточных слоях. Процесс состоит из последовательности обучающих циклов и завершается когда их число достигает заданного значения [4]. Также условием остановки обучения может являться достижение необходимой точности результата действия алгоритма для обучающей выборки.

### Обучение нейронной сети

Обучается СНС стандартным методом обратного распространения ошибки [5]. Основная идея этого метода состоит в распространении сигналов ошибки от выходов сети к её входам, в направлении, обратном прямому распространению сигналов в обычном режиме работы.

Алгоритм обратного распространения ошибки применяется для многослойного персептрона. У сети есть множество входов  $x_1, \dots, x_n$ , множество выходов Outputs и множество внутренних узлов. Перенумеруем все узлы (включая входы и выходы) числами от 1 до  $N$  (сквозная нумерация, вне зависимости от топологии слоёв). Обозначим через  $w_{i,j}$  вес, стоящий на ребре, соединяющем  $i$ -й и  $j$ -й узлы, а через  $o_i$  — выход  $i$ -го узла. Если известен обучающий пример (правильные ответы сети  $t_k$ ,  $k \in Outputs$ ), то функция ошибки, полученная по методу наименьших квадратов, выглядит так:

$$E(\{w_{i,j}\}) = \frac{1}{2} \sum_{k \in Outputs} (t_k - o_k)^2$$

Алгоритм реализовывает стохастический градиентный спуск, то есть веса изменяются после каждого обучающего примера. Чтобы "добраться" до минимума ошибки, необходимо "двигаться" в сторону, противоположную градиенту, т. е. на основании каждой группы правильных ответов добавлять к каждому весу  $w_{i,j}$  смещение

$$\Delta w_{i,j} = -\eta \frac{\partial E}{\partial w_{i,j}}$$

где  $0 < \eta < 1$  - множитель, задающий скорость "движения".

Получающийся алгоритм представлен на рис. 2.

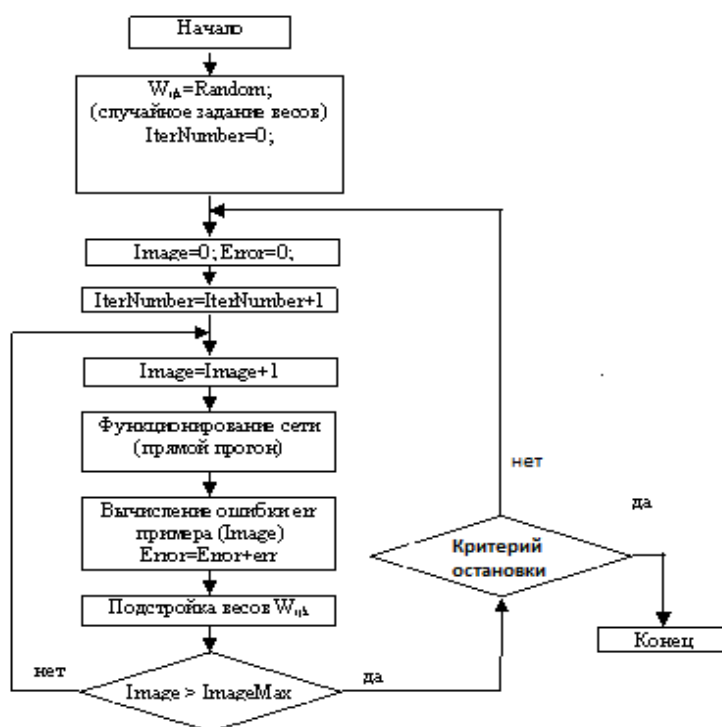


Рисунок 2 – Алгоритм обучения нейронной сети

На каждой итерации алгоритма весовые коэффициенты нейронной сети модифицируются так, чтобы улучшить решение одного обучающего примера. Таким образом, в процессе обучения циклически решается задача оптимизации.

### Описание программной модели

Общая структура модели системы представлена на рис. 3.

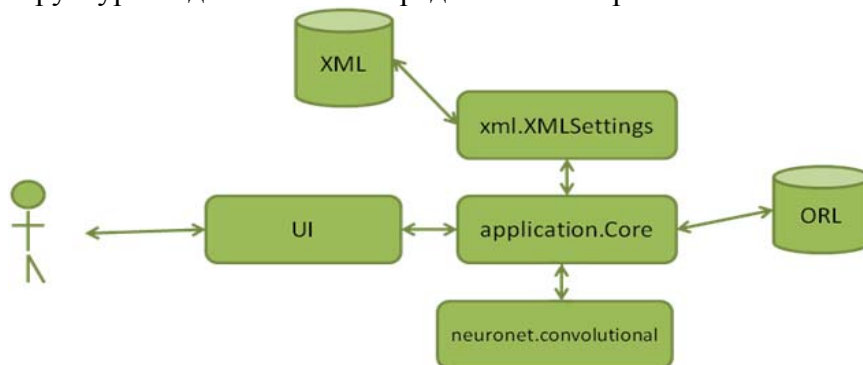


Рисунок 3 – Структура программной модели

Пакет UI отвечает за связь с пользователем и реализует графический интерфейс для настройки, обучения и тестирования системы. В нём собраны классы для работы с пользовательским интерфейсом. Application.Core - это контроллер, обеспечивающий взаимодействие пользовательского интерфейса и моделей. В пакете neuronet.convolutional реализована программная модель свёрточной нейронной сети. В пакете xml находятся xml-модели, используемые для чтения и записи данных в xml-файлы. Структура и настройки нейросети, а также обучающее множество хранятся в файлах в формате XML.

ORL Database - это база фотографий лиц, использовавшаяся для создания обучающей и тестовой выборки. База содержит 400 фронтальных изображений лиц 40 человек, размера 92 на 112 пикселей. Файлы сохраняются в формате gif.

На рис. 4 показан процесс обучения свёрточной нейронной сети методом обратного распространения ошибки. Из рисунка видно, что погрешность с течением времени уменьшается. Графическое отображение информации хорошо информирует пользователя о степени обученности нейронной сети.

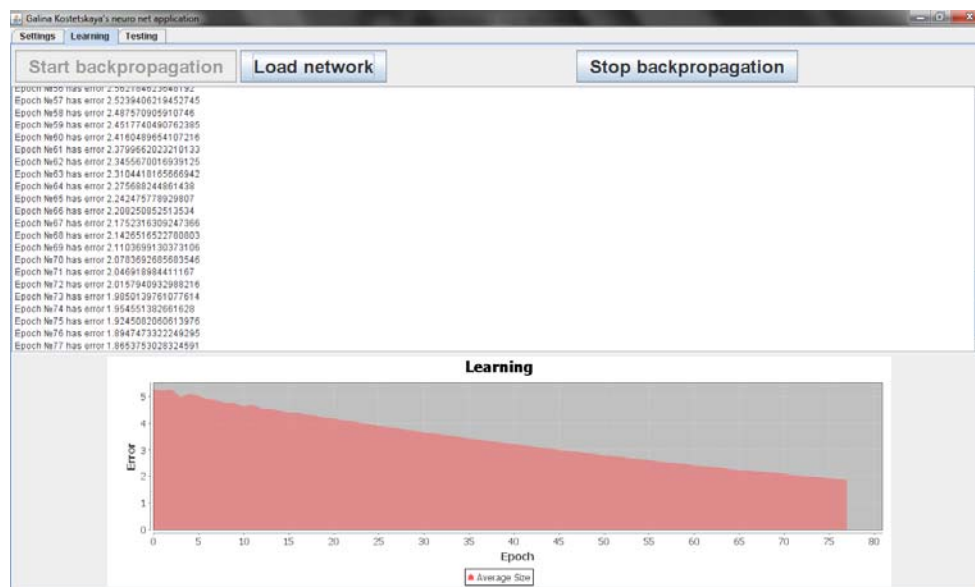


Рисунок 4 - Процесс обучения нейронной сети

### Выводы

На текущий момент создана программная модель свёрточной нейронной сети, которая реализует все функции системы распознавания и предоставляет расширенные возможности настройки для анализа и проведения исследований процессов нейросетевого распознавания лиц.

### Список литературы

1. Умяров Н.Х., Федяев О.И. Параметрическая модель свёрточной нейронной сети // VI международная научно-техническая конференция студентов, аспирантов и молодых научных работников «Информатика и компьютерные технологии»: Т. 2 – Донецк, ДонНТУ, 2010. – 292с.
2. Умяров Н.Х., Федяев О.И. Логическое и физическое представление архитектуры свёрточной нейронной сети // II Всеукраїнська науково-технічна конференція студентів, аспірантів та молодих вчених, 11-13 квітня 2011 р., м. Донецьк : зб. доп. у 3 т./ Донец. націонал. техн. ун-т; редкол.: Є.О. Башков (голова) та ін. – Донецьк: ДонНТУ, 2011. - Т.3, с. 81-85
3. Krueger N. An Algorithm for the Learning of Weights in Discrimination Functions Using a Priori Constraints. IEEE Transactions on Pattern Analysis and Machine Intelligence/ N.Krueger.// IEEE – 1997. - Vol. 19. - p. 764-768.
4. Esme B. Facial feature extraction using genetic algorithms/ B. Esme, B. Sankur, B. Anarim // 8-th European Signal Processing Conference, Trieste, - 1996. - P. 1511-1514.
5. Саймон Хайкин Нейронные сети: полный курс, 2-е издание. / Саймон Хайкин. – М. : Издательский дом. «Вильямс», 2006.