

УДК 004.424.4

Исследование влияния количества слоев нейронов и их количества на поведение ошибки скорость обучения сети

*Живых С.Ю., студент
кафедра «Система обработки информации и управление»
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана*

*Научный руководитель: Ревунков Г.И., к.т.н., доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
revunkov@bmstu.ru*

На сегодняшний день многослойный персептрон - одна из самых популярных и используемых нейросетей. Одно из главных преимуществ многослойного персептрона, это возможность решать алгоритмически неразрешимые задачи или задачи, для которых алгоритмическое решение неизвестно, но для которых возможно составить репрезентативный набор примеров с известными решениями. При обучении нейросеть, за счёт своего внутреннего строения, выявляет закономерности в связи входных и выходных образов, тем самым как бы "обобщает" полученный на обучающей выборке опыт. В этой способности к обобщению и состоит основа привлекательности многослойного персептрона. Мы сами можем и не знать какова зависимость между входными и выходными образами, достаточно иметь большой набор векторов для которых известен ожидаемый выход. Задачами, где успешно применялся многослойный персептрон являются задачи распознавания, экспертные системы, анализ изображений и прогнозирование рынка.

В общем случае, количество нейронов и слоев связано с:

- требуемым количеством входов и выходов сети;
- сложностью задачи;
- количеством данных для обучения
- памятью и быстродействием сети

Проведем исследование по изменению ошибки обучения и распознавания многослойной нейронной сетью распознавание и обучение некоторого множества изображений. Для этого был создан исследовательский макет. Данный программный

продукт позволяет распознавать цветные и черно-белые изображения, полученные из файла или нарисованные пользователем в специальном окне. На экранной форме отображается ход эксперимента, итоговые веса в градациях серого в качестве визуализации эксперимента, уменьшение ошибки с течением времени. На экране доступна начальная выборка для эксперимента. Программа работает в двух режимах- обучение и распознавание. Количество слоев нейронной сети и нейронов в них задается пользователем.

Приведем общий алгоритм обучения на Рис.1.

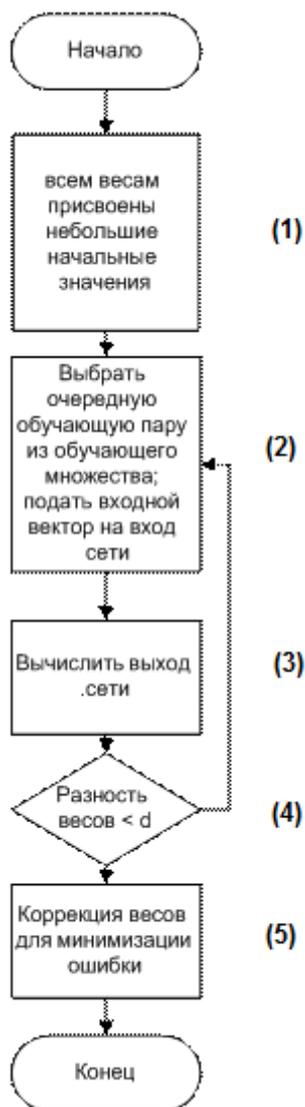


Рис.1. Общий алгоритм работы программы в режиме обучения

При распознавании выравниваем количество нейронов, а затем ищем с максимальным откликом.

На шаге (1) алгоритма обучения создаем матрицу весов связей нейронов скрытого слоя и матрицу весов выходного слоя. Придадим им в начальный момент случайные значения $\text{rand}()$ до 1.

Пусть на некоторый входной образ сеть вырабатывает выходной вектор. При этом для скрытого слоя произведем суммирование произведения весов связей нейронов скрытого слоя и входного вектора, а для выходного- суммирования произведения выходного вектора и весов матрицы выходного слоя (Шаг 2, Рис.1).

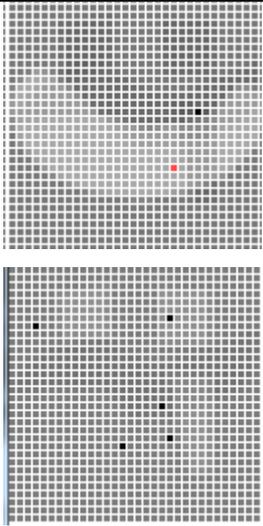
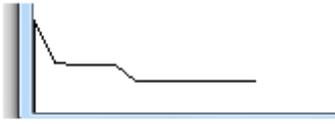
В качестве функционала подлежащего минимизации (выход сети) (Шаг 3, Рис.1) методом градиентного спуска возьмём среднеквадратичное отклонение реального выхода сети от предполагаемого. Процесс минимизации начнем с выходного слоя методом градиентного спуска, регулируя темп обучения. В качестве функции активации возьмем сигмоидальную функцию. Далее переходим на скрытый слой (Шаг 4,Рис.1) и производим коррекцию ошибки.

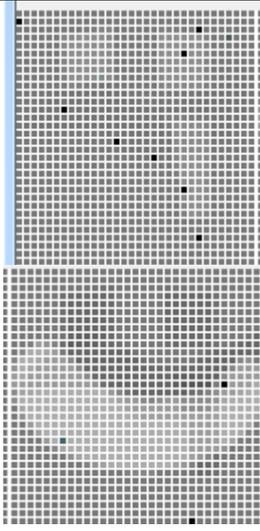
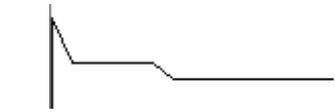
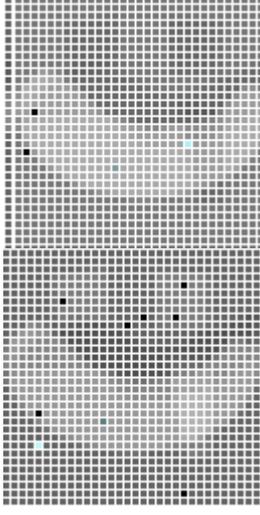
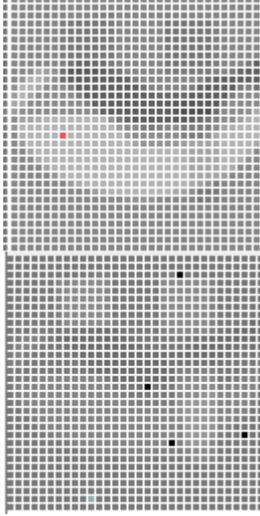
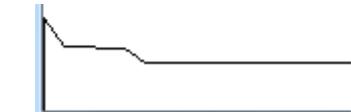
Исследуем зависимость числа слоев и нейронов на реальных изображениях, с визуализацией весов и графика изменения ошибки при корректировке значений.

Пусть имеется 2 слоя. Протокол экспериментов представлен Таблице 1.

Таблица 1

Эксперимент с двухслойной нейронной сетью по распознаванию изображений

Количество нейронов В 1 слое	Количество нейронов Во 2 слое	Матрица весов (В градациях серого)	График ошибки	Процент
2	1			25%

3	1			23%
4	1			20%
5	1			22%

В эксперименте было проведено распознавание образов, одинаковых для различного количества нейронов в скрытом слое. Видим, что наилучшие результаты обучения были достигнуты при двухслойной нейронной сети при 4 нейронах в скрытом

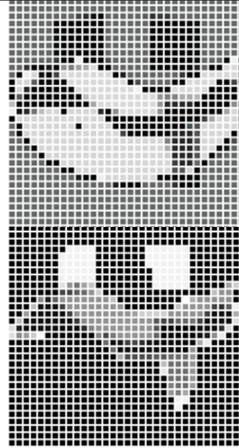
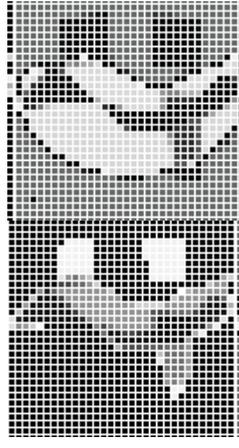
Молодежный научно-технический вестник ФС77-51038

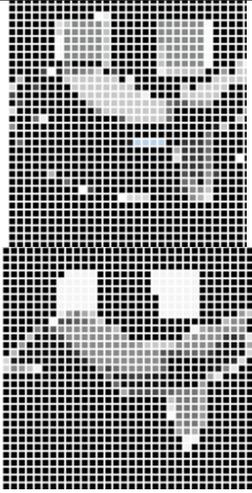
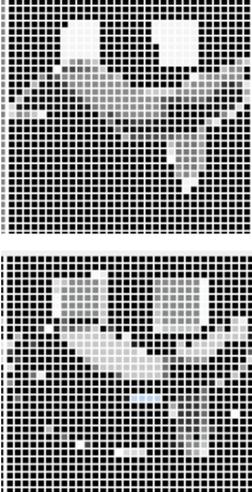
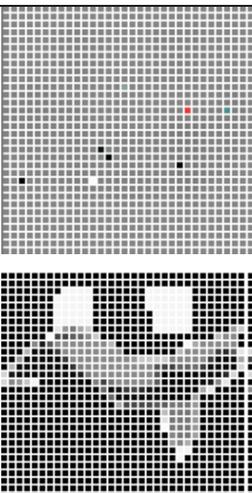
слое, т.к. ошибка в данном случае минимальна и стабильный уровень отклика достигается за минимальное время. Далее начинается процесс потери информации и рабочая точка нейросетевой системы в процессе обучения скользит по поверхности ошибок по направлению к глобальному минимуму целевой функции. В силу неровности рельефа поверхности ошибок сеть может принять локальный минимум за глобальный и «застрять» в нем, если склон этого локального минимума слишком крут, а шаг обучения мал (так называемый паралич сети).

Рассмотрим трехслойную нейронную сеть.

Таблица 2

Эксперимент с трехслойной нейронной сетью по распознаванию изображений

Количество нейронов V_1 в скрытом слое	Количество нейронов V_0 в скрытом слое	Матрица весов (В градациях серого)	График ошибки	Процент
3	5			11%
3	4			10%
4	2			10%

				
3	2			9%
7	2			8%

Видим, что чрезмерное увеличение нейронов в скрытых слоях персептрона может привести к явлению гиперразмерности (или переобучения), т.е. к тому, что хорошо обучившийся персептрон не будет проявлять обобщающих свойств, т.е. ошибка обобщения окажется чрезмерно большой.

Таким образом, в ходе экспериментов на исследовательском макете с конкретной задачей распознавания изображений были сделаны выводы о количестве нейронов и слоев
Молодежный научно-технический вестник ФС77-51038

в сети. При недостаточном количестве нейронов и слоев сеть не обучится и ошибка при работе сети останется большой, а так же на выходе сети не будут передаваться резкие колебания аппроксимируемой выходной функции.

С другой стороны, чрезмерное количество нейронов и слоев также мешает работе сети, т.к. быстродействие снижается, возможно переобучение (при распознавании сеть реагировала на несущественные детали), возможно попадание в локальный минимум.

Длительное время обучения может быть результатом неоптимального выбора длины шага. Неудачи в обучении обычно возникают по двум причинам: паралича сети и попадания в локальный минимум.

Список литературы

1. Ф. Уоссермен Нейрокомпьютерная техника: Теория и практика //evrika.tsi.lv: Научный портал «Эврика!» Россия.2013.
<http://evrika.tsi.lv/index.php?name=texts&file=show&f=410> дата обращения (27.05.2013).
2. Хайкин С. Нейронные сети: полный курс, 2-е издание. : Пер. с англ. – М.: Издательский дом «Вильямс», 2006 г. – 1104 с.
3. Алмаев Владимир, Петрунин Сергей - Особенности обучения персептронных сетей// it-claim.ru: Научно образовательный кластер CLAIM Россия.2013 http://it-claim.ru/Education/Course/AiSystem/Students/ANN_Almaev_Petrunin.pdf дата обращения (27.05.2013).
4. Александр Вежнев. Популярныe нейросетевые архитектуры. Компьютерная графика и мультимедиа. Выпуск №2(1)/2004. <http://cgm.computergraphics.ru/content/view/57> дата обращения (27.05.2013).