



УДК 004.93

ОБЗОР ИМЕЮЩИХСЯ РЕШЕНИЙ АЛГОРИТМОВ ПОСТРОЕНИЯ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ОБРАЗОВ

А.С. Бельшева, *belysheva.nastya@yandex.ru*, **С.П. Воробьев**, *vsp1999@yandex.ru*

Южно-Российский государственный политехнический университет (НПИ)

имени М.И. Платова, г. Новочеркасск

В данной работе представлен обзор существующих решений алгоритмов построения нейронной сети различного вида для распознавания образов изображения. Представлено подробное описание методов построения и детализированы их особенности: вид нейронной сети (такие как многослойный персептрон, нейронная сеть глубокого доверия, свёрточные нейронные сети), способы построения нейронной сети, обучающая выборка, принципы ее формирования. Целью данной работы является выявление общих решений построения нейронной сети для решения задачи распознавания образа изображения и определение ее эффективности. Одним из важных вопросов, рассматриваемых в представленной статье, является определение того, как выбор того или иного вида нейронной сети влияет на результат в задаче распознавания образов.

Ключевые слова: Нейронная сеть, алгоритм построения, распознавание образов, сегментация изображения, структуризация обучающего множества.

REVIEW OF AVAILABLE SOLUTIONS OF NEURAL NETWORK ALGORITHMS FOR PATTERN RECOGNITION

A.S. Belysheva, S.P. Vorobyov

Platov South-Russian State Polytechnic University (NPI), Novocherkassk

This paper provides an overview of existing solutions of various types of neural network construction algorithms for pattern recognition. A detailed description of the construction methods is presented and their features are detailed: the form of a neural network (such as a multilayer perceptron, deep confidence neural network, convolutional neural networks), methods for constructing a neural network, a training sample, the principles of its formation. The aim of this work is to identify common solutions for constructing a neural network to solve the problem of image image recognition and determine its effectiveness. One of the important issues considered in this article is the determination of how the choice of a particular type of neural network affects the result in the pattern recognition problem.

Keywords: Neural network, construction algorithm, pattern recognition, image segmentation, structuring of the training set.

Введение

В настоящее время современная наука добилась огромных успехов, используя информационные технологии, которые применяются в любой сфере нашей жизни: экономике, медицине, освоение космоса, образовании и многих других. Особенно удивительным достижением являются искусственные нейронные сети, созданные по подобию человеческого мозга. Они так же могут обучаться, анализировать, прогнозировать, классифицировать. Но чаще всего нейронные сети применяются для распознавания признаков и свойств каких-либо образов на изображениях в системах безопасности, контроля и управления доступом, в системах видеонаблюдения и т.д. Главными особенностями получения положительного результата при распознавании образов являются выбор вида нейронной сети, ее архитектуры и обучающей выборки.

Обзор алгоритмов распознавания образов

В работе [1] рассматривается метод организации архитектуры нейронной сети, предназначенной для распознавания образов. В качестве вида нейронной сети была выбрана нейронная сеть глубокого доверия из-за ее отличительной особенности - алгоритма обучения, за счет которого происходит увеличение количества



обрабатывающих слоев нейронной сети посредством уменьшения ошибки обучения. Для реализации нейросети автор использует базу рукописных символов *MNIST* (*Mixed National Institute of Standards and Technology*), которая состоит из 70000 изображений рукописных цифр, где 60000 изображений необходимы для обучения, 10000 – для тестирования.

Разработанный метод построения нейронной сети основывается на сравнении данного образа с последующими для выявления их схожести «(одна цифра будет сравниваться с каждой)». В своей работе автор немного модернизировал стандартный алгоритм обучения, добавив ограничитель по времени t , в течение которого нейросеть должна обучиться. Обучающая выборка – это множество чисел, состоящее из нулей и единиц. Принцип процесса обучения основывается на алгоритме обратного распространения ошибки, причем здесь используется его измененная версия:

1. Устанавливается шаг обучения и время работы алгоритма.
2. Случайно задаются весовые коэффициенты и пороговые значения нейронной сети.
3. Последовательно подаются векторы данных $\{x_1 \dots x_n\}^i$ на вход нейронной сети, при этом для каждого входного вектора выполняются следующие действия:
 - вычисляется выходная активность всех нейронных элементов сети,

$$y_j = F\left(\sum_{i=1}^n w_{ij} * y_i - T_j\right),$$

где индекс j характеризует нейроны следующего слоя относительно слоя i ;

- определяется ошибка нейронных элементов сети для выходного слоя:

$$Es_j = Y_j - t_j,$$

и скрытого:

$$Es_j = \sum_{i=1}^n Es_i * F'(S_i) * w_{ji},$$

где индекс j характеризует нейроны следующего слоя относительно слоя i , t^j

– эталонное значение.

4. Вычисляется суммарная среднеквадратическая ошибка нейронной сети:

$$E = \frac{1}{2} \sum_{k=1}^L \sum_j (y_j^k - t_j^k)^2,$$

где L – количество обучающих векторов.

Далее создается переменная $Error = E$, включается счетчик времени t_1 , задается условие, при котором значение $t > t_1$, иначе алгоритм заканчивается.

Весовые коэффициенты и пороги нейронных элементов изменяются по следующим правилам:

$$\begin{aligned} w_{ij}(t+1) &= w_{ij}(t) - alf * Es_j * F'(S_j) * y_i, \\ T_j(t+1) &= T_j(t) + alf * Es_j * F'(S_j). \end{aligned}$$

Выполняются пункты 3 и 4. Если получившаяся ошибка меньше $Error$, то $Error = E$, весовые коэффициенты W и порог T запоминаются и происходит переход к проверке значения t , в противном случае сразу к проверке значения.

После обучения нейронной сети происходит ее тестирование на обучающей выборке, в ходе которого алгоритм распознал одну неверную цифру на изображении из 10000 имеющихся.



Предложенный в данной работе метод проектирования архитектуры нейросети, основывающийся на усовершенствованном алгоритме обучения, в котором используется ограничитель по времени t , заслуживает внимания, поскольку полученное значение ошибки распознавания было близко к нулевому.

В работе [2] рассматривается проблема влияния обучающей выборки в задаче распознавания образов нейронной сети. Для распознавания автор взял базу изображений рукописных цифр (ноль и один) в виде бинарного кода. В ней всего 1783 изображения, из них 1702 изображения – это нули, 81 – это единицы. В качестве нейронной сети он выбрал многослойную нейронную сеть с одним скрытым слоем, а для обучения использовал алгоритм обратного распространения ошибки.

В качестве эксперимента автор решил обучить нейронную сеть на четырех вариантах обучающей выборки, т.е. составляющие выборки были одинаковыми, менялся их порядок. Размер обучающей выборки равен количеству пикселей в изображении. Например, изображение имеет размер 20×20 , тогда нейрон входного слоя соответствует одному пикселю и имеет вид:

$$Training_{it} = \begin{bmatrix} Image(1)_{1,1} & Image(2000)_{1,1} \\ \vdots & \vdots \\ Image(1)_{1,20} & Image(2000)_{1,20} \\ Image(1)_{20,1} & Image(2000)_{20,1} \\ \vdots & \vdots \\ Image(1)_{20,20} & Image(2000)_{20,20} \end{bmatrix},$$

где $Training_{it}$ – обучающая выборка, $i = 1 \dots 400$ – количество нейронов входного слоя, t – количество обучающих выборок.

Были рассмотрены варианты обучающей выборки с различной последовательностью чисел во множестве. В результате экспериментального исследования было доказано, что вид обучающей выборки напрямую влияет на обучение и работоспособность нейронной сети.

В статье [3] реализована автоматизация построения нейронной сети в рамках объектно-ориентированного подхода. Нейронная сеть определяется как иерархическое древо классов, где в первом классе описываются общие характеристики сети, а затем он делится на более мелкие подклассы с более точным описанием данной модели сети. В данной работе акцентируется внимание на проблеме выбора архитектуры сети и параметров, входящих в ее состав. Для ее решения проиллюстрирован объектно-ориентированный подход построения нейронной сети.

Алгоритм построения нейронной сети состоит из определенной последовательности шагов: создание объекта общего класса, определение основных характеристик нейросети, построение необходимого количества слоев, формирование объектов в каждом слое и вычисление весового коэффициента нейрона, потенциала нейрона.

Основным фактором создания нейронной сети является определение количества слоев в сети, количества нейронов входного (зависит от характера и размерности входных данных), скрытого (зависит от спецификации поставленной задачи) и выходного слоев. Также важно задание параметров нейрона в сети, который принадлежит к тому или иному слою. В статье нейроны входного слоя имеют значение веса, равное 1, поскольку они не участвуют в процессе обработки сигналов, а лишь обеспечивают их передачу в следующий слой.



Для обучения используется алгоритм обратного распространения ошибки, при котором значение квадратичной ошибки сводится к минимуму:

$$E = \frac{1}{2} \sum_{k=1}^p \sum_j (d_j^k - y_j^k)^2,$$

где p – количество примеров обучающей выборки, индекс i пробегает все выходы многослойной сети, d_k^i – желаемый выход нейрона i , y_k^i – текущий выход нейрона j последнего слоя.

В статье автор вводит дополнительный параметр *Neuron Err*, основное предназначение которого – это хранение значения ошибки каждого нейрона, зависевшее от его весового коэффициента. Вес нейрона вычисляется по следующей формуле:

$$W_j(t + 1) = wy(t) + \Delta Wy,$$

где W_{ij} – вес от нейрона i или от элемента входного сигнала i к нейрону j в момент времени t .

Основная идея представленного подхода состоит в том, чтобы вычислять чувствительность ошибки сети к изменениям весов. Для этого нужно вычислить частные производные от ошибки по весам.

На основе всего вышесказанного можно сделать вывод о том, что с помощью данного подхода к построению нейросети можно более детально продумать все составные части нейронной сети, что отразится на эффективности ее работы.

В статье [4] представлен способ структуризации обучающего множества искусственной нейронной сети для сегментации изображения. Сегментация изображения рассматривается с точки зрения кластеризации набора пикселей по признаку цветовой близости. Для этого используются нейронные сети, однако могут возникнуть проблемы с формированием обучающего множества, а именно в случае распознавание одного изображения, т.к. для сегментации рассматриваются несколько аналогичных изображений, этой проблемы нет.

Для решения этой проблемы автор используется ряд алгоритмов, различных как по составу нейронной сети, так и по принципу действия и ориентацией на определенный вид изображения, а именно: многослойный персептрон и его модификации, генетические алгоритмы, выращивание областей, минимальные отношения различия, вейвлет-разложения исходного изображения, самоорганизующиеся карты Кохонена.

В результате исследования предложено использовать алгоритм сегментации изображения на основе метода выращивания областей, который состоит в следующем: в случайном порядке в необработанной части изображения выбирается точка, а затем к ней присоединяются еще несколько точек по цветовым характеристикам. Таким образом формируется сегмент изображения, который считается обработанным. Затем происходит присоединение граничных точек соседнего сегмента. Каждый сегмент маркируется натуральным числом, где метка у необработанного сегмента равна нулю.

Для сравнения двух точек по их цветовым характеристикам в модели RGB используется трехслойный персептрон. Входной слой имеет шесть нейронов, поскольку сравнение происходит для двух точек с соответствующими координатами. Выходной слой состоит из двух нейронов, в которых хранится информация о положительном и отрицательном решении схожести точек.



Рассматривается изображение размером $N \times M$ пикселей, которое в модели RGB может быть представлено тремя матрицами $\|Ry\|$, $\|Gy\|$, $\|By\|$ ($i = 1, \dots, N; j = 1, \dots, M$) соответствующими разным цветовым частям. Вводятся следующие обозначения:

– для пикселя с координатами (i, j) – v_{ij} . Цветовые составляющие пикселя v_{ij} представлены как его поля данных:

$$v_{ij}.r = R_{ij}, \quad v_{ij}.g = G_{ij}, \quad v_{ij}.b = B_{ij};$$

– для матрицы меток – ij ($i = 1, \dots, N; j = 1, \dots, M$);

– для метки отдельного пикселя – как поля данных: $v_{ij}.c = C_{ij}$.

Алгоритм наращивания сегмента состоит в следующем: считается, что к данному моменту выделено k сегментов. Из множества пикселей с нулевой меткой происходит выбор одного v_{ij} и изменение его метки $v_{ij}.c = k + 1$, запускается рекурсивный алгоритм.

1. Пусть выращиваемый в составе сегмента имеется r пикселей $\{v_1, \dots, v_r\}$.

2. Происходит перебор в данном сегменте всех ближайших соседей с нулевой меткой пикселей.

3. Формируется входной вектор нейросети из цветовых координат точки сегмента v и её ближайшего соседа v' , не принадлежащего сегменту

$$(v.r, v.g, v.b, v'.r, v'.g, v'.b)$$

4. Вычисляется выходное значение нейросети. При положительном результате изменяется метка пикселя ($v': v'.c = k + 1$). При отрицательном – переход к следующему соседнему пикселю из другого сегмента.

5. В случае получения отрицательного результата у всех ближайших соседей алгоритм по формированию сегмента прекращается.

Завершение процесса сегментации изображения происходит в тот момент, когда не остаётся ни одного пикселя с нулевой меткой. Для обучения нейросети используется искусственно сгенерированный импульсный шум.

Необходимо отметить то, что тестирование обученной нейронной сети проводилось в автоматическом и интерактивном режиме, т.е. точку для распознавания в сегменте задавал сам пользователь, а не алгоритм. Исходя из предложенных примеров работы нейросети, можно сделать вывод о том, результат сегментации изображения в интерактивном режиме лучше, чем в автоматическом, однако это дает хорошую почву для дальнейшего развития данного способа использования нейронной сети в области распознавания образов.

В работе [6] рассматривается метод сегментации мультиспектральных снимков. В качестве объекта сегментации выступает множество изображений, полученных со спутников дистанционного зондирования Земли и используется свёрточная нейронная сеть, в которой имеются свёрточный слой, слой выборки и полносвязного слоя. В первом, свёрточном, слое находится матрица, выполняющая функцию фильтра изображений. На каждом шаге обхода одного изображения в 1 пиксель происходит поэлементное умножение значения фильтра на соответствующие значения яркости пикселя, затем путем их суммирования получается итоговое число, передающееся в новое изображение размером меньше исходного. Так происходит алгоритм свёртки в первом слое.

После свертки происходит выборка в слое выборки. Прямоугольные блоки небольшого размера обрабатываются для создания одного выхода из этого блока, т.е. полученное уменьшенное изображение из свёрточного слоя делается еще



меньше для выделения каких-либо признаков. Оба слоя строят карту признаков отдельно для каждого изображения, поскольку всякое изображение имеет свой специфичный признак.

Последним слоем является полносвязный, который принимает карту признаков и строит N -мерный вектор, состоящий из N классов. Выделение класса происходит путем определения характерных ему свойств, взятых с выхода предыдущего слоя. После полносвязного слоя свёрточных слоев не может быть.

Для тестирования только что созданной нейронной сети используется изображение размером 32×32 пикселя. Выбор изображения такого размера автор объясняет тем, что во фрагменте изображения имеется участок текстуры, которая не занимает большую часть всего фрагмента. Размер фильтра свёртки в первом слое составляет 3×3 , в последующих – 2×2 , в слое выборки – 2×2 .

Обучение нейронной сети производилось на множестве исходных снимком Земли со спутника. Для подготовки обучающей выборки на каждом изображении были выделены области определенного типа, принадлежащие к одному из шести цветовых классов различной поверхности. Нарезка снимков производилась с помощью специально созданной программы. После разбиения изображений на фрагменты выборка состояла из 2000 мультиспектральных изображений. Определение ошибки обучения являлось отношением неправильно классифицированных изображений из случайного множества к общему числу изображений в этом множестве и составляла 2.42%.

Процесс сегментации происходит следующим образом: Снимок обходится маской, делится на равные фрагменты и посылается в нейронную сеть на распознавание, которая в свою очередь вычисляет для каждого фрагмента класс и вероятность принадлежности к нему. Поскольку фрагменты исходного изображения не пересекаются, то может возникнуть проблема с неправильным присвоением класса: в процессе распознавания участка на фрагменте при выявлении большей площади какого-либо объекта нейросеть может стереть меньший участок, который является важным, и соотнести к неправильному классу. Иногда нейросеть может распознать и классифицировать меньший участок фрагмента. В этом случае необходимо учитывать при проверке вероятности, чтобы ее значение было не ниже установленного порога, и только тогда принимать решение о классификации или же делать распознавание заново.

В своей статье автор делает вывод о том, что с помощью этого метода сегментации можно избежать потери объектов на изображении, появления ложных, неточности определения их границ; нейронная сеть работает эффективней, и результаты получаются более точными.

Заключение

В данной работе был представлен обзор статей, в которых описано, как разные виды нейронных сетей применяются в задаче распознавания образов на изображении. Рассмотрены модели построения архитектуры нейросети, проанализированы возникающие проблемы при формировании обучающего множества и методы их решения. Полученные знания будут учтены и применены в ходе дальнейшего исследования для построения своей нейронной сети в рамках информационной системы салона флористики.



Список цитируемой литературы

1. Татьянкин В.М. Способ идентификации образов с использованием нейронных сетей глубокого доверия // Всероссийская научная конференция по проблемам управления в технических системах, 2015. с.92-94.
2. Татьянкин В.М., Дюбо И.С. Обучающая выборка в задаче распознавания образов при использовании нейронных сетей // Вестник Югорского государственного университета, Выпуск 2 (37). С. 94–98. 2015.
3. Евдокимов Е.А., Солодовников В.И. Автоматизация построения нейронной сети в рамках объектно-ориентированного подхода // Новые информационные технологии в автоматизированных системах, 2015. С.89-97.
4. Ларионов С.Б., Белим С.В. Алгоритм формирования обучающего множества искусственной нейронной сети для сегментации изображения // Математические структуры и моделирование №2(46). 2018. С. 79–85.
5. Иванов Е.С., Тищенко И.П., Виноградов А.Н. Сегментация мультиспектральных снимков с применением свёрточных нейронных сетей // Современные проблемы дистанционного зондирования Земли из космоса 2019. Т. 16. № 1. С. 25–34.

© А.С.Белышева, С.П.Воробьев 2019