

УДК. 004.8

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ ЧЕЛОВЕЧЕСКИХ ЛИЦ С ПОМОЩЬЮ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Костецкая Г.Ю., Федяев О.И.

Донецкий национальный технический университет г.Донецк

Кафедра прикладной математики и информатики

E-mail: squiiirrel@gmail.com

Аннотация

Костецкая Г.Ю., Федяев О.И. Распознавание изображений человеческих лиц с помощью свёрточной нейронной сети. В статье рассматривается распознавание изображений человеческих лиц с помощью сверточной нейронной сети, которое реализовано в виде программной модели.

Общая постановка проблемы. В настоящее время всё более широкое распространение получают биометрические системы идентификации личности. Распознавание человека по изображению лица выделяется среди биометрических систем тем, что, во-первых, не требуется специальное или дорогостоящее оборудование, а, во-вторых, не нужен непосредственный контакт с устройством. На данный момент проблеме распознавания человека по изображению лица посвящено множество работ, однако в целом она ещё далека от решения. Основные трудности состоят в том, чтобы распознать человека по изображению лица независимо от ракурса, положения, условий освещённости и т.д.

Постановка задачи исследования. Для решения задачи распознавания лиц были предложены различные методики, среди которых можно выделить подходы, основанные на искусственных нейронных сетях. Особого внимания заслуживает новый подход к распознаванию с помощью сверточных нейронных сетей. В данной работе необходимо было разработать программную модель сверточной нейронной сети для распознавания изображений человеческих лиц.

Решение задачи и результаты исследований. В ходе исследования была изучена работа сверточной нейронной сети и разработана структура её программной модели. Программная модель реализована в среде Java. Для настройки модели выбрано обучающее множество, обладающее достаточной репрезентативностью.

В настоящее время основной топологией нейронной сети используемой для решения задачи локализации лица на изображении является полносвязная нейронная сеть без обратных связей, так называемый многослойный персептрон. Но его применение к решению данной задачи представляет несколько трудностей.

Во-первых, как правило, изображения имеют большую размерность, соответственно вырастает размер нейронной сети (количество нейронов и т.п.). Большое количество параметров увеличивает вместимость системы и соответственно требует большей тренировочной выборки, что увеличивает время и вычислительную сложность процесса обучения.

Во-вторых, недостатком полносвязной архитектуры является то, что топология ввода полностью игнорируется. Входные переменные могут быть представлены в любом порядке, не затрагивая цель обучения.

На преодоление этих недостатков направлена архитектура сверточных нейронных сетей. Сверточные нейронные сети обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям[1].

Сверточная нейронная сеть представляет собой особый класс НС, наилучшим образом подходящий для интеллектуальной обработки визуальных данных. Идея сверточных

нейронних сетей заключается в чередовании сверточных слоев (C-layers), субдискретизирующих слоев (S-layers) и наличии полносвязных (F-layers) слоев на выходе. Они объединяют в себе три архитектурных идеи для достижения инвариантности к сдвигу и искажению исходного изображения: локальные поля восприятия, разделяемые веса и пространственная субдискретизация[2].

Локальное восприятие подразумевает, что на вход одного нейрона подается не всё изображение (или выходы предыдущего слоя), а лишь некоторая его область. Такой подход позволил сохранять топологию изображения от слоя к слою. Концепция разделяемых весов предполагает, что для большого количества связей используется очень небольшой набор весов. Суть субдискретизации и S-слоев в сверточных нейронных сетях заключается в уменьшении пространственной размерности изображения.

Типичная структура сверточной нейронной сети представлена на рис. 1.

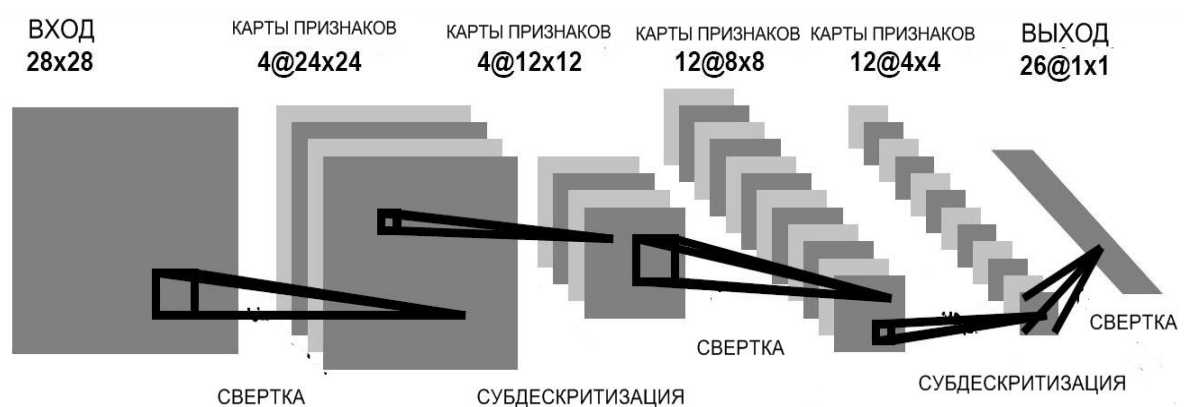


Рисунок 1 – Сверточная нейронная сеть

Сверточная нейронная сеть состоит из чередующихся слоев свертки и субдискретизации. Поступающее на вход изображение подвергается свертке с некоторым ядром свертки в соответствии с выражением[2]:

$$C\{n\}(i, j) = F \left[\sum_{k=1}^{k_s} \sum_{l=1}^{l_s} K\{n\}(k, l) \cdot S(i-k, j-l) + B(n) \right],$$

где:

C – результат свертки (карта признаков);

n – номер слоя;

i, j – индексы, определяющие положение элемента в карте признаков;

F – функция насыщения, обычно задаваемая как сигмоидальная или тангенциальная функция;

k, l – индексы определяющие положение элемента в матрице ядра свертки;

k_s, l_s – размеры ядра свертки;

K – ядро свертки;

S – входное изображение;

B – матрица смещений.

Ядро свертки является набором разделяемых весовых коэффициентов. Результатом данной операции является также некоторое изображение, которое называется картой признаков. В зависимости от выбранного ядра свертки, карта признаков будет выделять те или иные характеристики входного изображения. Для наиболее полного выделения характеристик входного изображения используется несколько различных ядер свертки так, что на выходе сверточного слоя получается несколько карт признаков.

За сверточним слоєм следует слой усреднения и субдискретизации, который понижает размерность карты признаков, тем самым понижая чувствительность выходов к сдвигам и поворотам.

Такое чередование сверточных и субдискретизирующих слоёв приводит к постепенному увеличению количества карт признаков при уменьшении их размерности от слоя к слою.

Обучение сверточной нейронной сети происходит методом обратного распространения ошибки. 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.

Для измерения качества распознавания использовалась функция среднеквадратической ошибки:

$$E^p = \frac{1}{2} * (D^p - O(I^p, W))^2$$

В данной формуле: E^p — это ошибка распознавания для p -ой обучающей пары; D^p — желаемый выход сети; $O(I^p, W)$ — выход сети, зависящий от p -го входа и весовых коэффициентов W . Задача обучения состояла в такой настройке весов W , чтобы они для любой обучающей пары (I^p, D^p) давали минимальную ошибку E^p . Чтобы посчитать ошибку для всей обучающей выборки просто бралось среднее арифметическое по ошибкам для всех обучающих пар. Для минимизации ошибки использовался стохастический [градиентный спуск](#)[3].

Сверточная нейронная сеть была реализована в виде программной модели. Эта модель в качестве обучающего множества использовала базу изображений человеческих лиц ORL Database of Faces, которая часто используется для оценивания систем распознавания. Она отвечает всем необходимым требованиям и доступна всем разработчикам. Эта база изображений содержит 400 образов по 10 в каждом классе (т.е. всего 40 различных людей). Каждый образ имеет разрешение 112x92 пикселя и 256 уровней яркости. Все лица представлены на тёмном фоне. Репрезентативность данных обеспечивается некоторыми изменениями масштаба лица, угла наблюдения и условий освещения.

Выводы: Разработанная программная модель сверточной нейронной сети была опробована на базе изображений человеческих лиц ORL Database of Faces, содержащей черно-белые изображения. Уровень ошибок распознавания на первой версии модели при 5 обучающих изображениях каждого лица составил 10-15%. Основываясь на полученных результатах можно сделать вывод, что использование сверточной нейронной сети является перспективным методом распознавания изображений человеческих лиц. Для разработанной модели необходимо продолжить исследования по нахождению оптимальных параметров модели нейросети для достижения более высоких показателей распознавания.

Список литературы

1. Y. Le Cun and Yoshua Bengio. Convolutional networks for images, speech, and time series. In Michael A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*, pages 255–258. MIT Press, Cambridge, Massachusetts, 1995.
2. Yann LeCun, J. S. Denker, S. Solla, R. E. Howard and L. D. Jackel: Optimal Brain Damage, in Touretzky, David (Eds), *Advances in Neural Information Processing Systems 2 (NIPS*89)*, Morgan Kaufman, Denver, CO, 1990
3. Y. LeCun, L. Bottou, G. Orr and K. Muller: Efficient BackProp, in Orr, G. and Muller K. (Eds), *Neural Networks: Tricks of the trade*, Springer, 1998