

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КубГУ»)

Факультет компьютерных технологий и прикладной математики
Кафедра анализа данных и искусственного интеллекта

КУРСОВАЯ РАБОТА

**ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ
РАСПОЗНАВАНИЯ 2D ОБЪЕКТОВ**

Работу выполнил _____ К.И. Варламов
(подпись)

Направление подготовки 09.03.03 Прикладная информатика

Направленность (профиль) Прикладная информатика в экономике

Научный руководитель
канд. пед. наук, доц. _____ В.А. Акиньшина
(подпись)

Нормоконтролер
канд. физ.-мат. наук. _____ Г.В. Калайдина
(подпись)

Краснодар
2022

СОДЕРЖАНИЕ

Реферат	3
Введение	4
1 Знакомство с нейронными сетями.....	6
1.1 Архитектура нейронных сетей	6
1.2 Искусственные нейронные сети	7
1.3 Где используются сверточные нейронные сети	11
2 Распознавание 3D объектов.....	13
2.1 Задача детекции объектов.....	13
2.2 Проблемы распознавания.....	14
3 Методы распознавания объектов	15
3.1 Основные методы (по признакам и исследовании конструкций).....	15
3.2 Структурные нейронные сети.....	16
4 Анализ предметной области и метода решения задачи	18
4.1 Постановка задачи.....	18
4.2 Построение математической модели.....	18
4.3 Реализация программы.....	20
Заключение	24
Список использованных источников	25

РЕФЕРАТ

Курсовая работа 19 с., 4 рис., 6 источников.

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ 2D ОБЪЕКТОВ

Объектом исследования являются сверточные нейронные сети, предметом их применение в обработке изображений.

Целью работы является изучение различных способов обработки изображений средствами сверточных нейронных сетей.

Цель работы определяет следующие задачи:

- Изучить строение нейронных сетей
- Установить и решить задачу распознавания 3D объектов
- Узнать о методах распознавания объектов
- Используя язык Python построить сверточную нейронную сеть

ВВЕДЕНИЕ

Развитие и распространение компьютерного зрения приводит к изменениям в различных профессиональных областях. Сверточные нейронные сети используются для распознавания объектов и лиц, анализа медицинских снимков, автопилотов автомобилей с автономными системами, систем безопасности и других областях. С возрастанием вычислительной мощности и появлением баз изображений стало возможным обучать глубокие нейронные сети (deep neural networks). При распознавании изображений используются свёрточные нейронные сети (Convolutional Neural Networks).

Одной из задач, которую решает машинное обучение, является классификация изображений. При этом, классификация заключается в определении номера категории, к которой принадлежит объект на изображении. Для оценки качества работы алгоритмов машинного обучения часто используются размеченные базы данных изображений, такие как CIFAR-10, ImageNet, PASCAL VOC. Однако, на изображениях базы данных, например, ImageNet, может быть изображено несколько объектов, при этом размечено (аннотировано) только один. В связи с этим, основным критерием оценки ошибки является top-5 ошибка, которая означает, что алгоритм считается не ошибившимся, если правильная категория объекта находится среди пяти категорий, наиболее вероятных по мнению алгоритма. Следовательно, многие нейронные сети, используемые для классификации изображений, оцениваются с помощью ошибки top-5.

В настоящее время существует множество технологий, которые позволяют автоматически распознавать изображения и объекты, восстанавливать цветность, восстанавливать старые фотографии и устранять дефекты. Одним из методов классификации изображений является использование искусственного интеллекта, что позволяет достичь более точной классификации изображений с меньшими затратами времени и усилий. Для решения таких задач, как оптическое распознавание образов,

классификация изображений, детектирование предметов, семантическая сегментация и другие, используются сверточные нейронные сети.

1 Знакомство с нейронными сетями

1.1 Архитектура нейронных сетей

В нейронных сетях с полным соединением каждый нейрон передает свой выходной сигнал всем остальным нейронам, включая себя. Входные сигналы подаются на все нейроны сети, а выходные сигналы могут быть получены после нескольких циклов работы сети.

В многослойных нейронных сетях нейроны объединяются в слои, где каждый слой содержит нейроны с общими входными сигналами. Количество нейронов в слое может быть любым и не зависит от других слоев. Обычно нейронные сети состоят из нескольких слоев, расположенных слева направо. На вход нейронной сети подаются внешние входные сигналы, которые передаются на входные нейроны нулевого слоя, а выходные сигналы получают на выходных нейронах последнего слоя. Кроме входного и выходного слоев, многослойная нейронная сеть может содержать один или несколько скрытых слоев. Связи от выходов нейронов некоторого слоя q к входам нейронов следующего слоя $(q+1)$ называются последовательными.

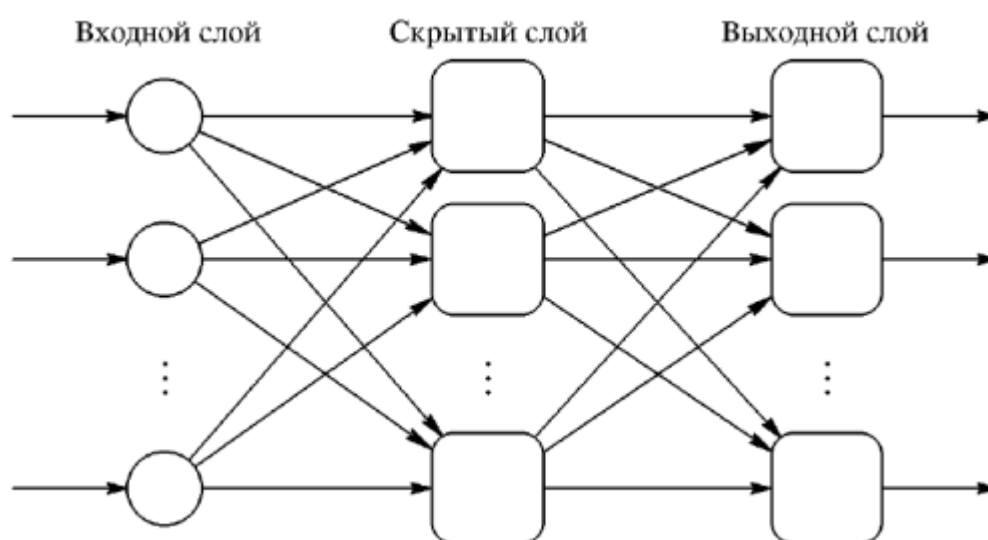


Рисунок 1 – Многослойные нейронные сети

1.2 Искусственные нейронные сети

Искусственная нейронная сеть (ИНС, НС) - это структура обработки информации, которая состоит из нейронов, связанных между собой и способна к параллельной обработке данных. Модель нейронной сети в программировании представляет собой компьютерную интерпретацию головного мозга. В отличие от центральной нервной системы человека, состоящей из спинного и головного мозга, функциональность искусственной нейронной сети осуществляется при помощи множества связанных между собой нейронов, которые могут передавать информацию посредством импульсов через синаптические связи.

Перцептрон - это математическая модель, созданная Ф. Розенблаттом в середине прошлого века, которая имитирует восприятие информации человеческим мозгом. Перцептрон стал одной из первых моделей нейронных сетей, которую впоследствии стали активно развивать и улучшать.

Нейронные сети обладают способностью самообучения, то есть они могут обобщать имеющиеся прецеденты на новые случаи, используя полученный опыт для выдачи результатов.

Искусственные нейронные сети состоят из трех типов слоев: входного, скрытого и выходного (см. рисунок 2). Каждый слой содержит нейроны в определенном количестве.

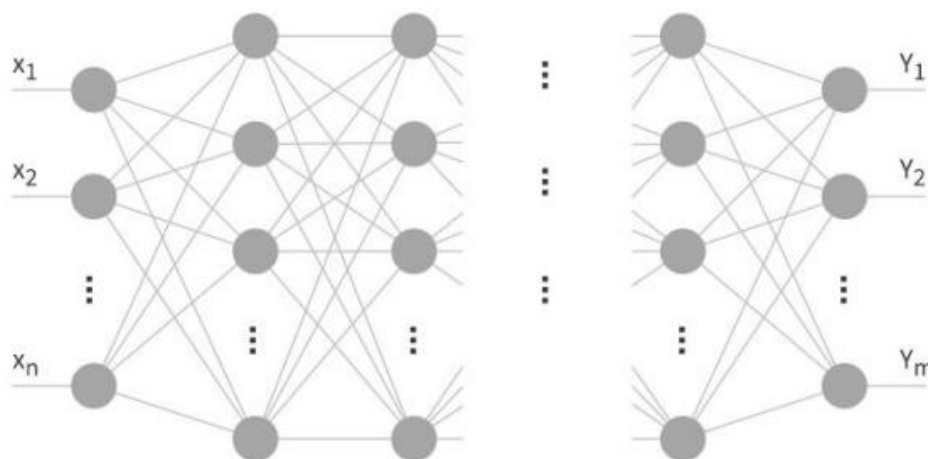


Рисунок 2 – Слой нейронной сети

Процесс обучения нейронной сети является итеративным, его шаги называют эпохами.

Для обучения НС чаще всего используют методы обучения с учителем, например, метод обратного распространения ошибки и его модификации. Нередко используют и обучение без учителя. Тип обучения главным образом зависит от решаемой задачи.

Обучение с учителем происходит при наличии полного набора размеченных данных. В обучающем наборе каждому примеру соответствует решение, которое сеть должна получить. Разность между правильным и полученным представляет собой ошибку, которую необходимо устранять с помощью новой настройки параметров.

Обучение без учителя проводится без контроля разработчика над процессом. На вход подается набор данных, и нейронная сеть пытается самостоятельно найти взаимосвязи. При обучении без учителя у разработчика отсутствуют правильные решения, которые сеть должна получить.

Одним из достоинств НС является решение задач в условиях неопределенности. Способность к самообучению позволяет сетям искать решение задач с неизвестными закономерностями и зависимостями между входными и выходными данными.

Сверточная искусственная нейронная сеть (СИНС, СНС) – это архитектура искусственных нейронных сетей, которая входит в состав технологий глубокого обучения. Такие сети нацелены на распознавание образов. Их модель в программировании основана на особенностях работы клеток зрительной коры головного мозга.

Для решения задач компьютерного зрения широко используется технология сверточных нейронных сетей. Преимуществом такого подхода является то, что такие сети являются одним из лучших алгоритмов по распознаванию и классификации изображений, устойчивы к повороту и сдвигам изображений, имеют гораздо меньше настраиваемых весов по

сравнению с обычной нейронной сетью. Еще одним достоинством является возможность реализации параллельно работающих алгоритмов [3], что позволяет повысить производительность системы и увеличить скорость вычислений.

Сверточные сети работают по тому же принципу, что и обычные нейронные сети, только помимо умножения матриц в них присутствует операция свертки. Свертка – это линейная операция, применяющаяся к двум функциям вещественного аргумента и возвращающая третья, которая характеризует сходство одной функции с отраженной и сдвинутой копией другой функции.

Еще одно отличие сверточных сетей – это разреженная связность.

В обычном многослойном персептроне каждый нейрон следующего слоя связан со всеми нейронами предыдущего слоя, и все связи имеют свои веса в матрице весов. А в СНС используется небольшая матрица весов, которая на каждом шаге движется по обрабатываемому слою.

Матрицу весов называют ядром свертки, в сети ядер несколько. Ядра кодируют наличие каких-либо признаков на изображении – горизонтальных и вертикальных линий, линий под углом, дуг, сложных фигур (эллипсов, треугольников, квадратов и других). Следующий слой, формирующийся в результате свертки одной из матриц весов, будет отражать существование определенного признака в обработанном слое. Таким образом формируется карта признаков, представляющая собой массив матриц.

Каждая карта содержит в себе синаптическое ядро или фильтр. Фильтр показывает наличие определенного признака. Проход каждым набором весов составляет свой экземпляр карты признаков, делая нейронную сеть многоканальной (то есть много независимых карт признаков на одном слое).

При выполнении операции свертки окно с размерностью ядра проходит с заданным шагом всю область изображения, на каждом шаге прохода умножает содержимое окна на ядро, результат суммирует и записывает в

матрицу результата – очередную карту признаков. Размеры карт признаков одного сверточного слоя одинаковы.

Операция подвыборки или пулинг (англ. Pooling) необходима для уменьшения масштаба карт признаков. Чаще всего используется операция MaxPooling – отбор наибольших значений: из нескольких соседних нейронов карты выбирают максимальный и принимают его за один нейрон новой карты признаков меньшей размерности. Благодаря этому можно значительно 20 снизить объем используемой памяти и ускорить процесс дальнейших вычислений.

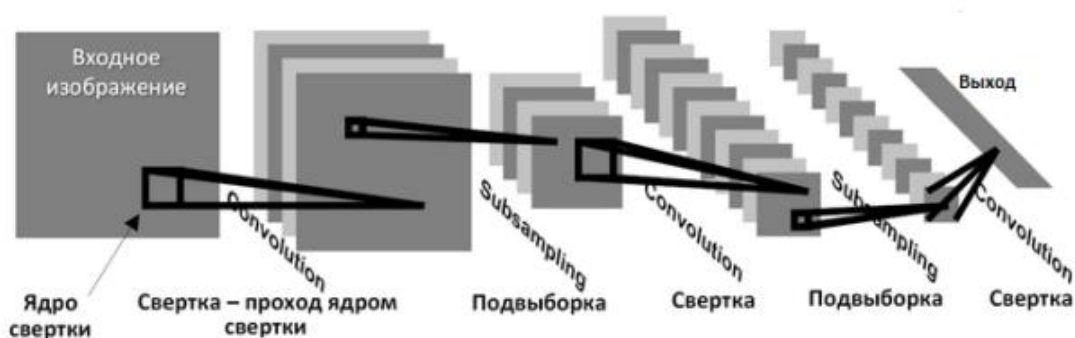


Рисунок 3 – Схема работы сверточной сети

Сигнал проходит через некоторое количество слоев (на рисунке: один входной слой, представляющий собой непосредственно само входное изображение, четыре скрытых слоя, один выходной слой), где чередуются свертка и подвыборка (пулинг). При этом составляются карты признаков. Их количество увеличивается с каждым слоем, но одновременно уменьшается разрешение карт.

На выходе получается большой набор каналов, содержащих небольшое количество данных, которые отражают признаки, выявленные на входном изображении. Затем, как правило, эти данные передаются на обычную полносвязную нейронную сеть, которая формирует конечный выходной сигнал. В итоге исходное изображение проходит через множественную фильтрацию.

Ядра свертки формируются с помощью обучения нейронной сети методом обратного распространения ошибки. Процессы свертки, выполняемые по каждой карте признаков, происходят параллельно, что является существенным преимуществом с точки зрения времени работы сети.

Количество карт определяется требованиями к ответу решаемой задачи. Если будет большое число карт, то улучшится качество распознавания, но в то же время повысится сложность вычислений. Чаще всего берется соотношение 1:2 – на одну карту предыдущего слоя приходится две карты нового слоя.

Недостатком СНС является наличие большого количества варьируемых параметров – количество слоев, количество и размерность ядер свертки, шаг сдвига ядра при обработке слоя, параметры подвыборки, а также параметры выходной полносвязной сети.

1.3 Где используются сверточные нейронные сети

Самый простой пример – «умные» плейлисты музыки (например, Яндекс.Музыка подбирает уникальный плейлист исходя из того, чтобы слушаете чаще всего) или видео на YouTube. Так работают нейронные сети, которые получают поступившую от вас, а также миллионов похожих на вас людей, и прогнозируют то, что вам может понравиться.

Список их применения очень широк. Вот только самые полезные и известные истории:

- Диагностировать заболевания сельскохозяйственных культур и других растений по фото с точностью до 99,35% (пример с комбайном). Если так дело и дальше пойдет, нейросети научатся заменить студентов на уборке картошки и капусты!

- Прогнозировать погоду. Этим занимается, например, «Яндекс.Погода»: там используется специальный алгоритм на основе нейронных сетей, который прогнозирует метеорологические изменения с точностью до минут (или почти так).

– Делать из чёрно-белых снимков цветные. Попробуйте этот сервис – местами он неидеален, но некоторые фотографии и правда делает красивее и ярче. Явно же у вас немало черно-белых снимков в архивах?

– Помогать вам объезжать пробки уже сейчас, а в будущем работать помощниками автопилотов автомобилей, постоянно анализируя получаемую информацию.

– Играя, обучать нейросети. Google запустили проект Quick, Draw!, который вроде бы и развлекательный, но при этом каждая игра обучает нейросеть и совершенствует её. Играете так – вам нужно нарисовать то, что просит сеть, а она непрерывно дает свои варианты. И угадывает! Даже Эйфелеву башню в моем варианте (это три палочки и несколько – посередине).

– Распознавать речь по губам лучше человека. Это удалось сделать разработчикам из Оксфордского университета. В качестве исходного материала авторы работы взяли базу данных Grid, в которой было собрано более 32 тысяч видеозаписей и, следуя алгоритму, обучали нейросеть. Результаты тестирования показали, что LipNet может правильно распознавать речь по губам в 93,4 процентах случаев. Это должно серьезно помочь для людей с нарушениями слуха.

– Творить искусство. Можно послушать альбом «Нейронной обороны» или хоралы, написанные нейросетью за Баха. А на Vimeo есть короткометражка, снятая по сценарию нейронной сети.

– Превращать фотографии в картины и автоматически редактировать их. Вы, конечно, слышали об этих сервисах: Prisma, MSQRD, Mlvch, deepart.io и так далее.

2 Распознавание 2D объектов

2.1 Задача детекции объектов

Нейронные сети можно рассматривать как современные вычислительные системы, которые преобразуют информацию по образцу процессов, происходящих в мозге человека. Обработка информации имеет численный характер, что позволяет использовать нейронную сеть, например, в качестве модели объекта с совершенно неизвестными характеристиками. Другие типовые приложения нейронных сетей охватывают задачи распознавания, классификации, анализа и сжатия образов.

Свыше 80% всех приложений нейронных сетей относится к так называемым многослойным сетям без обратных связей. В них сигнал пересылается в направлении от входного слоя через скрытые слои (если они имеются) к выходному слою.

Сверточные нейронные сети, СНС (англ. convolutional neural network, CNN) являются биологически вдохновленными вариантами многослойных перцептронов (multilayer perceptron) [12]. Из ранних работ Хьюбел и Визела по зрительной коре животного, было выяснено, что зрительная кора содержит сложное расположение клеток. Данные клетки чувствительны к небольшим подобластям поля зрения, называемым рецептивным полем. Субрегионы представлены сеткой, которая покрывает всё поле зрения. Данные ячейки действуют как локальные фильтры в пространстве ввода зрительной информации и хорошо подходят для использования сильной пространственной локальной корреляции, присутствующей в естественных изображениях. Если говорить простым языком, то сверточные нейронные сети представляют из себя нейронную сеть, которая умеет определять контуры и цветовое представление входного изображения. Отличительной чертой подобных сетей является сам процесс свёртки изображения. В ходе, которого

мы “сворачиваем” матрицу признаков входного изображения и получаем еще одну матрицу признаков, но уже в уменьшенном варианте.

2.2 Проблемы распознавания

К основным проблемам систем распознавания объектов относят следующие:

- низкое качество изображения (причинами могут быть сломанная камера, неверные установки съемочной аппаратуры и др.),
- плохое освещение, наличие теней,
- размытость изображения (эта проблема чаще всего присутствует в системах, которые производят видеофиксацию движущихся объектов),
- наличие предмета, который полностью или частично закрывает собой исследуемый объект.

Помимо перечисленных выше проблем можно выделить дополнительные, которые встречаются в системах распознавания ГНЗ:

- отличия между номерными знаками разных стран. Эту проблему можно решить расширением базы данных путем добавления в нее примеров номерных пластин других стран,
- наличие прицепа. В этом случае необходимо распознать две номерные пластины
 - основного транспортного средства и его прицепа,
 - наличие буксира (аналогично проблеме выше),
 - смена водителем автомобиля полосы движения в момент считывания камерой его ГНЗ. Эта проблема свойственна системам, работающим в режиме реального времени,
- закрытие некоторых символов номерного знака грязью, багажом, буксирными крюками и другими предметами.

3 Методы распознавания объектов

3.1 Основные методы (по признакам и исследовании конструкций)

Несмотря на многообразие методов распознавания образов, их можно разделить на две группы. Первая основана на понятии пространства признаков и их обработки в этом пространстве. Вторая – на исследовании конструкции рассматриваемых образов (синтаксическое распознавание).

Для первой группы методов в качестве основополагающей принята гипотеза о возможности представления образа в виде вектора, принадлежащего множеству V . Множество векторов, состоит из N таких подмножеств, что каждый вектор, отнесенный в результате классификации к j -му классу, принадлежит подмножеству E_j . Свойства множества V могут быть записаны в виде

$$\bigcup_{i=1}^N E_i = V, E_i \cap E_j = \emptyset (\forall i \neq j)$$

Синтаксический метод распознавания основан на восприятии основных элементов языка – примитивов. Они делятся на еще более мелкие составляющие – символы, являющиеся наименьшими элементами языка. Множество используемых символов называется алфавитом или словарем. Язык создается не только с помощью алфавита символов. Правила построения, преобразования и взаимодействия слов определяются грамматикой. Она представляет собой множество правил, по которым строятся фразы, а следовательно, и сам язык.

Формально грамматика может быть задана следующей записью:

$$G = \langle V_n, V_t, P, S \rangle$$

где V_n – нетерминальный словарь; V_t – терминальный словарь; P – множество правил подстановки; S – начальная аксиома ($S \in V_n$). Для грамматики характерны следующие соотношения: $V = V_n \cup V_t$ – словарь,

$Vn \cap Vt = 0$, $P = \{ \alpha_1 \rightarrow \beta_1, \alpha_2 \rightarrow \beta_2, \dots, \alpha_m \rightarrow \beta_m \}$, где $\alpha_i \in V^* - \{\lambda\}$, $\beta_i \in V^*$; $\{\lambda\}$ – пустая строка; V^* – множество всех возможных последовательностей, которые удастся построить с помощью итерационных процедур на основе данного словаря.

3.2 Структурные нейронные сети

В методе распознавания образов по набору их признаков, описанном выше, используются только числовые значения. Для объектов со сложной структурой этого может быть недостаточно. Тогда необходимо к имеющимся характеристикам объекта добавить дополнительные. Такие методы и называются структурными, так как структура объекта не может быть описана только с помощью числовых параметров. Объект, как правило, разбивают на более простые элементы и определяют отношения между ними.

Сложные образы состоят из более простых подобразов. Описание отношений между подобразами можно представить в виде иерархии, например, в графовой структуре. Вершинами графа являются части и элементы образа, а ребрами – отношения между ними. Отношения могут описываться логическими или математическими функциями.

Можно сказать, что элементы, отношения и ориентированный граф в совокупности образуют один из типов семантических сетей – функциональную сеть. Формат функциональных сетей предполагает использование целостных образов, которые связаны между собой. Связи представлены в виде графа. В графах допускается множественность связей (ребер) между парами вершин.

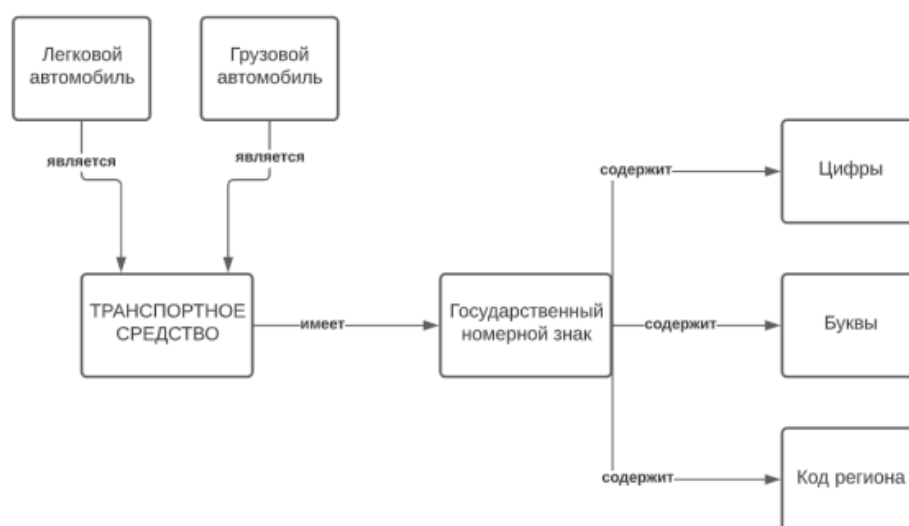


Рисунок 4 – Фрагмент семантической сети, соответствующей предметной области «Транспортные средства»

Структурный метод распознавания дает возможность описывать большое количество сложных объектов, используя относительно небольшую базу с несложными образами.

Распознавание проводится в три этапа.

На первом этапе осуществляется подготовка объекта к распознаванию – проверка его наличия на изображении, определение границ и краев.

Далее, на втором этапе, объект представляют в виде структуры: проводится сегментация, объект делится на части – элементы, между ними определяются отношения. Элементы и отношения берутся из базы правил, отражающей семантику предметной области.

На третьем этапе выполняется анализ полученной структуры. На этом шаге принимается решение, является ли представление объекта верным, и в зависимости от результата решения объект либо исключается из рассмотрения, либо определяется в класс образов.

4 Анализ предметной области и метода решения задачи

4.1 Постановка задачи

В настоящее время распознавание 2D объектов является одной из важнейших задач в области компьютерного зрения. Она находит применение в различных областях, таких как медицина, автоматическое управление, робототехника, банковское дело и многих других. Одним из наиболее распространенных примеров распознавания 2D объектов является распознавание банковских карт.

Хорошо реализованная программа может избавить человека от необходимости вводить данные вручную. На сегодняшний день большинство компаний отдают свое предпочтение бесплатному SDK, которое не обновлялось уже более 4 лет. Поэтому создание сверточной нейронной сети, которая бы корректно распознавала данные с пластиковых карт, является актуальной задачей.

Целью данной курсовой работы является исследование возможности использования сверточных нейронных сетей для распознавания 2D объектов на примере данных банковских карт «Сбербанк».

4.2 Построение математической модели

Конкретные задачи, которые необходимо решить в рамках работы, включают в себя:

- Сбор и подготовка набора данных банковских карт "Сбербанк" для обучения сверточной нейронной сети.
- Разработка и обучение сверточной нейронной сети для распознавания банковских карт «Сбербанк».
- Оценка качества работы сверточной нейронной сети на тестовом наборе данных.

– Анализ результатов и выводы о применимости сверточных нейронных сетей для распознавания 2D объектов на примере данных банковских карт «Сбербанк».

В качестве исходных данных будем использовать банковские карты: личные или найденные в сети «Интернет».

Чтобы построить сверточную нейронную сеть, будем использовать возможности языка Python. Опишем модель нейронной сети: она содержит 5 слоев свертки, где будет обрабатываться и обрезаться изображение, и 1 полносвязный слой.

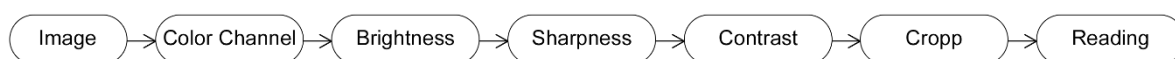


Рисунок 5 – Модель сверточной нейронной сети

В модели присутствуют три библиотеки: Pillow, OpenCV и Tesseract. Pillow и OpenCV будем использовать для обработки изображения (настройка контрастности, яркости, резкости, обрезка изображения), Tesseract – для считывания данных.

Библиотека Pillow – это форк библиотеки Python Imaging Library (PIL), которая предоставляет множество функций для работы с изображениями в Python. Она позволяет открывать, изменять и сохранять изображения в различных форматах, включая JPEG, PNG, BMP, GIF и многие другие.

С помощью Pillow можно изменять размер изображений, обрезать их, поворачивать, изменять цветовые пространства, применять фильтры и многое другое. Библиотека также поддерживает работу с анимированными GIF-изображениями.

В OpenCV реализованы многие алгоритмы компьютерного зрения, такие как обнаружение лиц, распознавание объектов, оптический поток, сегментация изображений, стереозрение и многие другие. Кроме того,

библиотека предоставляет множество функций для обработки изображений, таких как изменение размера, поворот, обрезка, наложение фильтров и т.д.

Tesseract – это библиотека оптического распознавания символов (OCR) с открытым исходным кодом, разработанная компанией Google. Она позволяет распознавать текст на изображениях и сканированных документах и конвертировать его в текстовый формат.

В Python данная библиотека может быть использована с помощью библиотеки pytesseract, которая предоставляет интерфейс для вызова функций Tesseract из Python. С помощью pytesseract можно легко распознавать текст на изображениях, а также управлять различными параметрами распознавания, такими как язык, настройки шрифта и многое другое.

4.3 Реализация программы

В качестве примера возьмем изображение банковской карты компании «Сбербанк»:

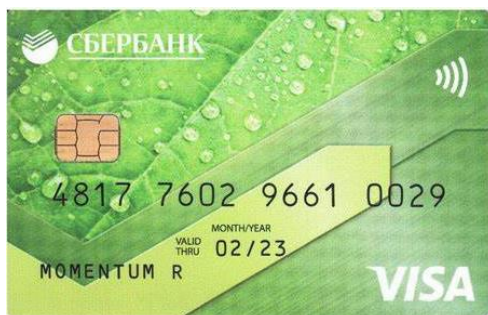


Рисунок 6 – Изображение банковской карты

Предположим, что фото было сделано на телефон. Соответственно, для начала работы стоит обрезать его по границе объекта. Используя контуры, находим объект на изображении, обрезаем и подгоняем под нужные размеры (в нашем случае, 327пикс X 205пикс).

Используя библиотеку Pillow разобьем изображение на цветовые каналы.

Данная библиотека предоставляет возможность работать с цветными изображениями, которые состоят из трех цветовых каналов: красного (R), зеленого (G) и синего (B). Каждый канал представляет собой матрицу пикселей, состоящую из значений от 0 до 255, где 0 – это черный цвет, а 255 – это белый цвет.

Например, если мы рассмотрим пиксель в координатах (x, y), то его цвет можно представить в виде тройки значений (R, G, B), где R – значение красного канала, G – значение зеленого канала, а B – значение синего канала. Таким образом, каждый пиксель на изображении представлен тройкой значений цветовых каналов.

Для доступа к каждому из цветовых каналов в Pillow можно использовать методы `split()` и `merge()` объекта `Image`. Метод `split()` разбивает изображение на отдельные каналы, а метод `merge()` объединяет каналы обратно в изображение.

Получим результат:



Рисунок 7 – Цветовые каналы (Красный, Зеленый, Синий)

Далее нужно сделать выбор: так как шрифт на банковской карте черного цвета, то следует использовать вариант такой, который «высветляет» дизайн карты. Красный и синий каналы затемняют фон, что приводит к трудностям с распознаванием текста. В это время зеленый канал делает наоборот. Используем именно его.

Обработаем изображение: в результате должны получить «черный текст на белом фоне». Используем усилители той же библиотеки (контрастность, яркость, резкость).

Усилители (enhancers) в библиотеке Pillow в Python – это инструменты, которые позволяют улучшить качество изображения путем изменения его яркости, контрастности и резкости.

Для работы с усилителями в Pillow необходимо использовать модуль ImageEnhance. В этом модуле определены классы Contrast, Brightness, Color и Sharpness, каждый из которых позволяет улучшить изображение в соответствующем аспекте.

Присвоив значения усилителей: яркость = 2, контраст = 2, получим результат:



Рисунок 8 – Результат после обработки усилителями

Данные на банковских картах всегда находятся на одном и том же месте, поэтому для обрезки изображения по контуру цифр не требуется поиск их границ. Воспользуемся библиотекой OpenCV и обрежем фото по размеру цифр (300пикс X 30пикс).

Остается последний этап: через библиотеку Tesseract считываем данные с картинки. Получаем отличный результат:



Рисунок 9 – Результат

Конечно, на начальных этапах обучения сети возникнут проблемы распознавания. Вот один пример из них:

Была выбрана карта иного дизайна: синий фон, былые символы иного шрифта.



Рисунок 10 – Банковская карта иного дизайна

В отличие от предыдущего примера, надо сделать фон изображения максимально черным, чтобы светлые цифры были читаемы. Но даже при таком условии будут проблемы: пересветы и неподходящий шрифт. Из-за них получится некорректный результат.



Рисунок 11 – Результат работы программы

ЗАКЛЮЧЕНИЕ

По итогу работы мы изучили различные способы обработки изображений средствами сверточных нейронных сетей.

Также определили следующие задачи:

- Изучили строение нейронных сетей
- Установили и решили задачу распознавания 3D объектов
- Узнали о методах распознавания объектов
- Используя язык Python построили сверточную нейронную сеть

Таким образом использование сверточных нейронных сетей для распознавания 2D объектов является эффективным и перспективным подходом, который может быть применен в различных областях, таких как компьютерное зрение, автоматическое распознавание лиц, медицинская диагностика и другие. Дальнейшие исследования в этой области могут привести к улучшению точности и скорости работы моделей, а также расширению возможностей их применения.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Разлацкий, С.А. Использование сверточных нейронных сетей для распознавания трехмерных объектов в реальном времени / С.А. Разлацкий, П.Ю. Якимов [Электронный ресурс] – Режим доступа: repo.ssau.ru/bitstream/Informacionnye-tehnologii-i-nanotehnologii/Ispolzovanie-svertochnyh-neironnyh-setei-dlya-raspoznavaniya-trehmernyh-obektov-v-realnom-vremeni-64091/1/paper%20302_1672-1677.pdf (дата обращения: 17.11.2022)
2. Рутдовская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы [Электронный ресурс] / Д. Рутдовская, М. Пилиньский, Л. Рутковский – Режим доступа: <https://reader.lanbook.com/book/11843#19> (дата обращения: 17.11.2022)
3. Что такое нейронные сети и как они, развлекая, меняют нашу жизнь? [Электронный ресурс] – Режим доступа: dit.urfu.ru/ru/blog/28689/ (дата обращения: 05.12.2022)
4. MachineLearning [Электронный ресурс]. – Режим доступа: www.machinelearning.ru (дата обращения: 05.12.2022).
5. Прокопеня, А.С. Сверточные нейронные сети для распознавания изображений / А.С. Прокопеня, И.З. Азаров [Электронный ресурс] – Режим доступа: https://libeldoc.bsuir.by/bitstream/123456789/39033/1/Prokopenyu_Svertochnyye.pdf (дата обращения: 26.12.2022)
6. Гафаров, Ф.М. Искусственные нейронные сети и их приложения / Ф.М. Гафаров, А.Ф. Галимянов [Электронный ресурс] – Режим доступа: https://kpfu.ru/staff_files/F1493580427/NejronGafGal.pdf (дата обращения: 26.12.2022)

Название ▾	Дата загрузки ▾	Оригинальность
DOCX Варламов_090303_курсовая	✓ 29 Мая 2023 09:35	82%