

МИНОБРНАУКИ РОССИИ  
Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«Ярославский государственный университет им. П.Г. Демидова»

Кафедра дискретного анализа

Сдано на кафедру

«\_\_»\_\_\_\_\_ 2020 г.

Заведующий кафедрой,

д.ф.-м.н., профессор

\_\_\_\_\_ В.А. Бондаренко

***Выпускная квалификационная работа***

**Анализ алгоритмов глубокого машинного обучения в задачах  
распознавания изображений**

по направлению

01.03.02 Прикладная математика и информатика

Научный руководитель

к.т.н., старший преподаватель

\_\_\_\_\_ Д.В. Матвеев

«\_\_»\_\_\_\_\_ 2020 г.

Студент группы ИВТ-41БО

\_\_\_\_\_ А.С. Коротков

«\_\_»\_\_\_\_\_ 2020 г.

Ярославль, 2020

## Реферат

Выпускная квалификационная работа: 18 стр., 3 гл., 4 рис., 1 таблиц, 10 источников.

Ключевые слова: **машинное обучение, глубокие нейронные сети, распознавание изображений, TensorFlow.**

Объектом исследования являются модели глубоких нейронных сетей для решения задач распознавания изображений.

Цель работы – изучить и проанализировать применение глубокого машинного обучения в задачах распознавания изображений.

В результате работы были разработаны и реализованы модели нейронных сетей для решения ????. Проведен анализ полученных результатов и сделан вывод о качестве работы алгоритмов машинного обучения в задачах распознавания изображений.

## Содержание

|  |           |
|--|-----------|
| <b>Введение</b>                                    | <b>4</b>  |
| <b>1 Машинное обучение</b>                         | <b>6</b>  |
| 1.1 Понятие искусственной нейронной сети . . . . . | 6         |
| 1.2 Активационная функция . . . . .                | 7         |
| 1.3 Обучение нейронных сетей . . . . .             | 9         |
| 1.4 Сверточные нейронные сети . . . . .            | 10        |
| 1.5 Рекуррентные нейронные сети . . . . .          | 11        |
| <b>2 Программная реализация</b>                    | <b>12</b> |
| 2.1 Конфигурация системы . . . . .                 | 12        |
| 2.2 Средства реализации . . . . .                  | 12        |
| 2.3 Обучающая выборка . . . . .                    | 12        |
| 2.4 Сверточная нейронная сеть . . . . .            | 12        |
| 2.5 X2 . . . . .                                   | 13        |
| 2.6 X3 . . . . .                                   | 13        |
| <b>3 Анализ работы программ</b>                    | <b>14</b> |
| <b>Заключение</b>                                  | <b>15</b> |
| <b>Список литературы</b>                           | <b>16</b> |
| <b>Приложение А</b>                                | <b>17</b> |

## Введение

В настоящее время, в связи со стремительным развитием цифровых технологий, использование автоматизированных и роботизированных систем распространилось на множество областей как в промышленности, науке, так и в повседневной жизни. В следствие этого, возрастает необходимость в эффективной обработке информации, представленной, в частности, в формате видео и изображений.

На текущий момент изображения тесно влились в жизнь человека. Поэтому многие автоматизированные системы используют их в качестве основного источника информации. Нахождение, локализация, классификация и анализ объектов на изображении компьютером – сложная задача компьютерного зрения.

Компьютерное (машинное) зрение (Computer Vision) – это совокупность программно-технических решений в сфере искусственного интеллекта (ИИ), нацеленных на считывание и получение информации из изображений, в реальном времени и без участия человека.

В процессе обработки информации, получаемой из глаз, человеческий мозг پردازывает колоссальный объем работы. Человек без труда сможет описать что находится и что происходит на случайно взятой фотографии. Изображения могут нести в себе колоссальное количество деталей и отличаться множеством параметров, таких как: разрешение, цветность, качество, яркость, наличие шума и т.д. Объекты на изображениях также могут обладать множеством особенностей: масштаб, положение, цвет, поворот, наклон и т.д. Однако, в цифровом формате, каждое изображение представляет собой лишь массив числовых данных. Научить компьютер находить и классифицировать образы на изображении с учетом всех факторов – очень сложная алгоритмическая задача. Для её решения активно применяют технологии машинного обучения.

Большое количество информации человек получает при помощи зрения. Изображения способны хранить огромное количество информации. Как следствие их использование в компьютерных системах способствует увеличению их производительности. Однако такие технологии требуют сложных вычислений. Задачей компьютерного зрения является разработка эффективных алгоритмов, выполняющих извлечение и анализ данных из изображений или видео.

В настоящий момент, подобные технологии применяются для решения таких сложных задач как:

- OCR – Optical character recognition (Оптическое распознавание символов): преобразование текста на изображении в редактируемый.
- Фотограмметрия – технология создания трехмерной модели объекта на основе фотографий, сделанных с различных ракурсов.
- Motion capture – технология, широко применяемая в киноиндустрии, позволяющая преобразовывать движения реальных людей в компьютерную анимацию.
- Дополненная реальность (AR) – технология, позволяющая в реальном времени проецировать виртуальные объекты на изображение реального окружения.
- Медицинская диагностика – обнаружение раковых клеток на ранней стадии, увеличение качества МРТ изображений, их анализ и т.д.

В данной работе был проведен анализ алгоритмов глубокого машинного обучения для решения задач распознавания изображений. ???

Первая глава содержит основные сведения об машинном обучении.

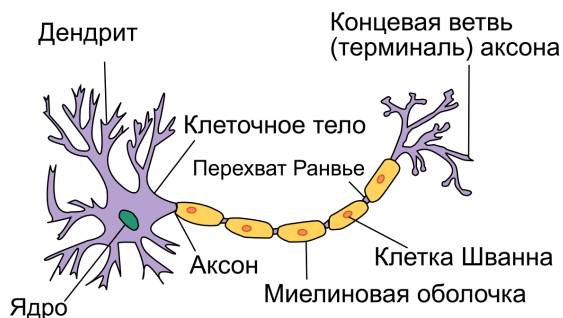
Во второй главе выполнено проектирование и реализация алгоритмов для распознавания изображений.

В третьей главе проведен сравнительный анализ работы разработанных алгоритмов.

## 1.1 Понятие искусственной нейронной сети

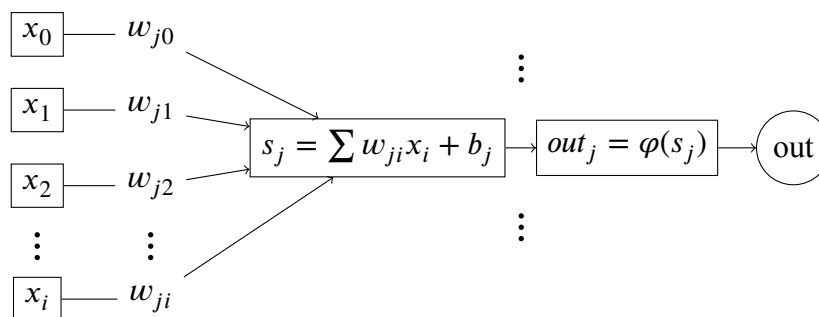
Машинное обучение (Machine Learning) – раздел исследований в сфере ИИ, в основе которых лежат методы разработки систем способных к обучению. Алгоритмы машинного обучения эффективно себя показывают в задачах, в которых требуется у заранее подготовленных (обучающих) данных определить общие признаки и по ним идентифицировать новые данные. В проектировании таких, обучающихся, систем часто применяют искусственные нейронные сети.

Искусственная нейронная сеть (ИНС) – компьютерная модель, в основе которой лежат принципы работы биологической нейронной сети - совокупности связанных между собой нервных клеток - нейронов. Каждый нейрон имеет набор входных связей - синапсов, по которым он получает информацию, представленную в виде импульсов, от других нейронов. По полученным данным нейрон формирует своё состояние и, с помощью аксона, сообщает его другим нейронам, обеспечивая функционирование системы. В процессе формирования системы одни нейронные связи укрепляются, а другие ослабляются, обеспечивая обучаемость сети.



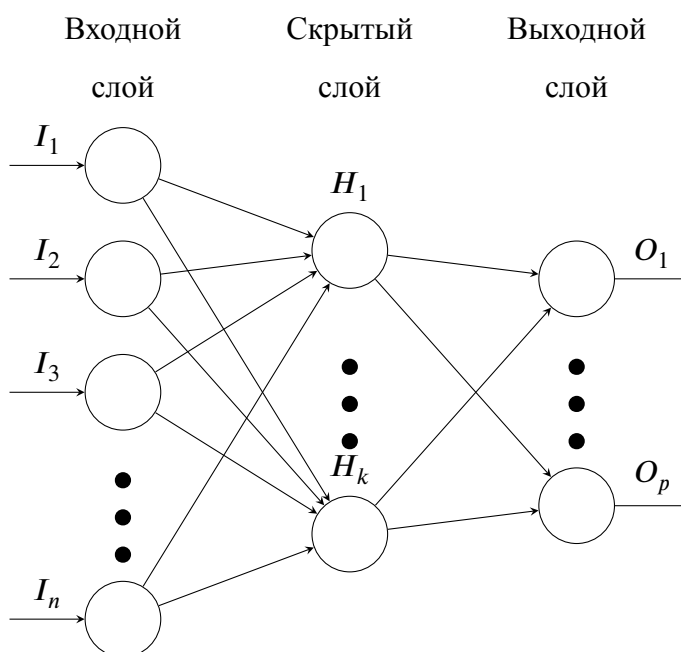
**Рис. 1** – Типичная структура биологического нейрона

Искусственный нейрон представляет собой упрощенную модель биологического нейрона. Принцип его работы представлен на рисунке 2. Сначала нейрон получает  $n$ -мерный вектор входных значений  $X = (x_1, \dots, x_n)$  и вектор весов  $W = (w_1, \dots, w_n)$ , обозначающий «укрепленность» межнейронных связей. Вычисляется сумма произведения входных значений и весов  $s_j$ . Затем к полученному результату применяется функция активации  $\varphi$ . В некоторых случаях, к сумме прибавляется величина смещения  $b_j$ .



**Рис. 2** – Схема искусственного нейрона

Множества нейронов формируют слои, слои в свою очередь формируют нейронную сеть. Входной слой получает данные, обрабатывает и передает нейронам скрытого слоя. Аналогично сработает каждый последующий слой вплоть до выходного.



**Рис. 3** – Схема простой нейронной сети

Нейросети с большим количеством скрытых слоев называется глубокими. Область машинного обучения, в которой используются глубокие нейронные сети называется - глубоким обучением (Deep Learning).

## 1.2 Активационная функция

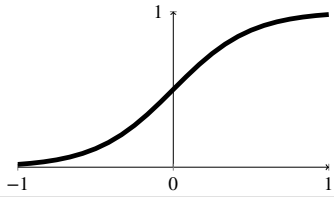
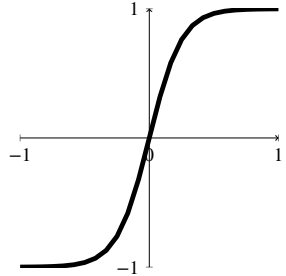
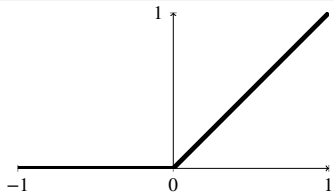
Взвешенная сумма входов представляет собой линейную комбинацию, из чего следует, что независимо от количества слоев, значения выходного слоя зависят только от входов первого слоя. Активационная функция нейрона обеспечивает нормализацию

посчитанной суммы и нелинейность нейронной сети. Для многих моделей нейронных сетей также требуется, чтобы активационная функция была монотонной и непрерывно-дифференцируемой на всей области определения.

Существует большое количество функций активации. Наиболее распространенные из них представлены в таблице 1.1.

Таблица 1.1

### Популярные активационные функции

| Название                | Функция   | Вид  |
|-------------------------|---|--|
| Сигмоидная              | $\sigma(x) = \frac{1}{1+e^{-x}}$                                |    |
| Гиперболический тангенс | $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$                      |  |
| ReLU                    | $f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0. \end{cases}$ |  |

Отдельно стоит упомянуть функцию Softmax. Эта функция часто применяется на последнем слое глубоких нейронных сетей в задачах классификации. Пусть последний слой сети содержит  $N$  нейронов, каждый из которых соответствует некоторому классу. Тогда значение выхода  $i$ -го нейрона вычисляется по формуле:

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^N e^{z_j}}$$

Таким образом, результат каждого нейрона будет принимать значения из диапазона  $[0, 1]$ , а их сумма равна 1. Тем самым сеть выдаст вероятности отношения входных данных к заданным классам.



### 1.3 Обучение нейронных сетей

Под обучением нейронных сетей подразумевают подбор значений весов связей для эффективного решения поставленной задачи. Изначально, веса устанавливаются случайно. Затем, в процессе прогона через сеть тестовых данных, веса корректируются так, чтобы в конечном итоге сеть выдавала правильные ответы.

Для того, чтобы контролировать процесс обучения необходимо как-то оценивать работу сети. Для этого используется функция потерь (функция стоимости), которая формирует некоторое численное значение, характеризующее точность работы сети. Таким образом обучение сети подразумевает минимизацию функции потерь.

Существует несколько подходов к обучению нейронных сетей:

#### Обучение с учителем

Обучение с учителем - это такой тип обучения, в котором сеть получает набор входных данных с заранее известными правильными ответами. Веса корректируются в зависимости от того, правильный ли ответ дала сеть.

Одним из популярных методов обучения с учителем является алгоритм обратного распространения ошибки [4], который часто применяется в обучении глубоких нейронных сетей.

Пусть сеть имеет  $L$  слоев,  $a^l, w^l, b^l$  - векторы значений, весов и смещений нейронов на  $l$ -м слое. Также имеется  $N$  обучающих пар  $(x, y)$ . В процессе обучения циклично происходят следующие итерации:

а) На вход сети подается вектор  $x$  из обучающего множества, для каждого слоя вычислить значения:

$$z^l = w^l a^{l-1} + b^l a^l = \sigma(z^l)$$

б) Вычислить значение функции стоимости:

$$C = \frac{1}{2} \sum_j (y_j - a_j^L)^2$$

в) Вычислить значения ошибок выходного слоя:

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} \sigma'(z_j^L)$$

г) Вычислить ошибки для каждого предыдущего слоя:

$$\delta_j^l = \sum_k w_{kj}^{l+1} \delta_k^{l+1} \sigma'(z_j^l)$$

д) Вычислить градиент функции стоимости:

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l$$

е) Обновить веса связей:

$$w_{ij}^l = w_{ij}^l - \mu \frac{\partial C}{\partial w_{jk}^l}, \quad 0 < \mu \leq 1$$

#### Обучение с подкреплением

Данный метод подразумевает наличие некоторой окружающей среды в которой действует сеть. Такая среда реагирует на действия модели и подает ей определенные сигналы.

### **Обучение без учителя**

Обучение при котором сеть заранее не располагает правильными ответами и самостоятельно ищет общие и отличительные признаки у входных данных.

### **Генетические алгоритмы обучения**

Алгоритмы, имитирующие эволюционные механизмы развития биологической популяции, выступают как альтернатива алгоритму обратного распространения ошибки. Значение произвольного весового коэффициента в нейронной сети называется геном. Гены формируют хромосомы, а хромосомы - популяцию. Далее, в пределах одного цикла (эпохи) с определенными вероятностями происходит:

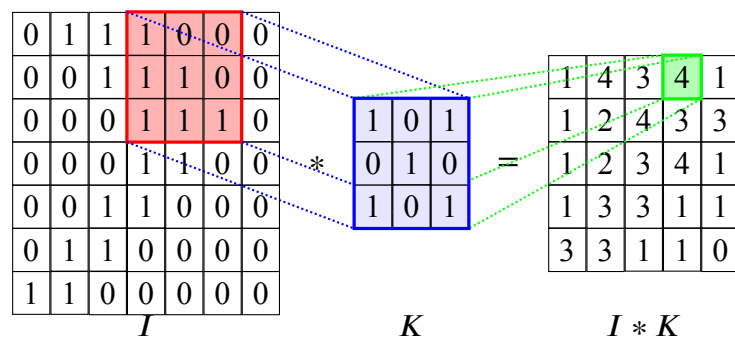
- Скрещивание хромосом - формирование новой хромосомы из генов двух других
- мутация - случайное изменение произвольного гена
- приспособление - (хромосомы показавшие худшие результаты уничтожаются из популяции.

+/- ???

## **1.4 Сверточные нейронные сети**

Стандартные нейронные сети состоят из полносвязных слоев - слоев, в которых каждый нейрон связан с каждым нейроном следующего слоя, что значительно увеличивает вычислительную сложность системы при увеличении количества нейронов. В типовых сверточных нейронных сетях к полносвязным добавляются сверточные и подвыборочные слои.

Сверточные слои характеризуются использованием матриц весов, называемых фильтрами или ядрами, которые обладают размерностью меньше исходных данных. Такое ядро с определенным шагом проходит по набору входных данных ( $I$ ) и вычисляет суммы произведений соответствующих значений ячеек и весов, формируя карту признаков ( $I * K$ ). Один сверточный слой может содержать несколько ядер и соответственно несколько карт признаков.



**Рис. 4** – Операция свертки

Так как признаки уже обнаружены, для упрощения дальнейших вычислений можно снизить детализацию входных данных. Это обеспечивает подвыборочный (пулинговый) слой, сжимая карты признаков, полученные от сверточного слоя, что снижает количество параметров, используемых в дальнейших вычислениях сети.

СНС может иметь несколько пар чередующихся сверточных и подвыборочных слоев. Завершается СНС стандартными полносвязными слоями.

## 1.5 Рекуррентные нейронные сети

Особенностью рекуррентных нейронных сетей, в отличие от обычных, является наличие обратных связей, которые позволяют передавать на нейроны предыдущее состояние системы. Подобные модели применяются в задачах где существенно важен порядок передаваемых данных, например при анализе текста или обработке видеоряда.

???

## **2 Программная реализация**

### **2.1 Конфигурация системы**

Все модели были протестированы на системе со следующими характеристиками:

- Операционная система - Ubuntu 19.10
- Процессор - Intel Core i5 9400F CPU 2.90 Ghz
- Объем оперативной памяти - 8 Гб
- Видеокарта - Nvidia GeForce GTX 1660

### **2.2 Средства реализации**

Python 3 - гибкий и мощный язык программирования, эффективно выполняющий задачи анализа и обработки данных.

TensorFlow - многофункциональный фреймворк с открытым исходным кодом, разработанный компанией Google для решения задач машинного обучения. TensorFlow позволяет проектировать и обучать различные архитектуры нейронных сетей.

cuDNN - библиотека глубоких нейронных сетей от Nvidia позволяющая использовать для вычислений мощности графического процессора.

### **2.3 Обучающая выборка**

Создание обучающей выборки в задачах такого типа является сложным и долгим процессом. Необходимо большое количество времени и помощь большого количества людей для того, чтобы создать и разметить даже небольшую выборку. Поэтому, в качестве источника обучающих и тестовых данных была использована готовая база данных: ???

### **2.4 Сверточная нейронная сеть**

Первой моделью для реализации была выбрана сверточная нейронная сеть, содержащая:

- 3 сверточных слоя с активационной функцией ReLU
- 3 пулинговых слоя
- 1 выравнивающий слой, преобразующий двумерную матрицу значений в вектор

- 1 полносвязный слой с активационной функцией ReLU
- 1 полносвязный слой с активационной функцией Softmax

По итогам обучения точность сети составила 88.01%

Время обучения: ???

**2.5 X2**

**2.6 X3**

### **3      Анализ работы программ**

По итогам работ были получены следующие результаты:

## Заключение

В соответствие с поставленной целью - анализ алгоритмов глубокого машинного обучения в задачах распознавания изображений - были реализованы алгоритмы X1, X2 и X3, а также проведен анализ качества их работ.

Модель X1 продемонстрировала ??? результат.

X2 ???

X3 ???

Анализ показал, что глубокое машинное обучение может быть эффективным в задачах распознавания изображений. X1 X2 X3 ???

## Список литературы

1. *Farinella G. M., Battiato S., Cipolla R.* Advanced Topics in Computer Vision. — Springer, 2013. — С. 475.
2. *Jamil M., Sharma S., Singh R.* Fault detection and classification in electrical power transmission system using artificial neural network // SpringerPlus. — 2015. — Июль. — Т. 4. — С. 334. — DOI: 10.1186/s40064-015-1080-x.
3. *Michelucci U.* Applied Deep Learning: A Case-Based Approach to Understanding Deep Neural Networks. — Apress, 2018. — С. 431.
4. *Nielsen M. A.* Neural Networks and Deep Learning. / Determination Press. — 2015. — URL: <http://neuralnetworksanddeeplearning.com/>.
5. *Singh P., Manure A.* Learn TensorFlow 2.0: Implement Machine Learning and Deep Learning Models with Python. — Apress, 2020. — С. 195.
6. *Жерон О.* Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. — ООО "Альфа-книга", 2018. — С. 688.
7. *Клетте Р.* Компьютерное зрение. Теория и алгоритмы. — ДМК Пресс, 2019. — С. 506.
8. *Конец Д.* Классические задачи Computer Science на языке Python. — Питер, 2020. — С. 256.
9. *Нишант Ш.* Машинное обучение и TensorFlow. — Питер, 2019. — С. 336.
10. *Шолле Ф.* Глубокое обучение на Python. — Питер, 2018. — С. 400.



## Приложение А

### Листинг 1. Сверточная нейронная сеть

```
1 from tensorflow_core.python.keras import Sequential
2 from tensorflow_core.python.keras.callbacks import ModelCheckpoint,
  TensorBoard
3 from tensorflow_core.python.keras.layers import Conv2D, MaxPooling2D,
  Flatten, Dense
4
5
6 class CNN:
7     name = "Convulation_Neural_Network"
8     model = None
9     callbacks_list = []
10
11     def __init__(self):
12         model = Sequential()
13
14         model.add(Conv2D(32, (3, 3), padding='same', input_shape
15             =(300, 300, 3), activation='relu'))
16         model.add(MaxPooling2D(pool_size=2))
17         model.add(Conv2D(64, (3, 3), padding='same', activation='relu
18             '))
19         model.add(MaxPooling2D(pool_size=2))
20         model.add(Conv2D(32, (3, 3), padding='same', activation='relu
21             '))
22         model.add(MaxPooling2D(pool_size=2))
23         model.add(Flatten())
24         model.add(Dense(64, activation='relu'))
25         model.add(Dense(6, activation='softmax'))
26
27         filepath = "trained_cnn.h5"
28         checkpoint = ModelCheckpoint(filepath, monitor='val_acc',
29             verbose=1, save_best_only=True, mode='max')
30         # tensorboard_callback = TensorBoard(log_dir='./logs')
31         self.callbacks_list = [checkpoint]
```

```
28         model.summary()
29         model.compile(loss='categorical_crossentropy', optimizer='
           adam', metrics=['acc'])
30         self.model = model
31
32     def learn(self, train_generator, test_generator):
33         self.model.fit(train_generator,
34                         epochs=100,
35                         steps_per_epoch=2276//32,
36                         validation_data=test_generator,
37                         validation_steps=251//32,
38                         callbacks=self.callbacks_list)
```