|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ \_\_\_\_\_\_\_\_\_ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ\_\_\_\_\_\_\_(ИУ)\_\_\_

КАФЕДРА \_\_\_\_СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ\_\_\_(ИУ5)\_\_

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

***К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ***

***НА ТЕМУ:***

***\_\_\_\_\_\_Исследование возможностей нейронных сетей для распознавания продуктов питания по фотографии\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

Студент \_\_\_**ИУ5-31М**\_\_\_\_ **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_ Ю.Б.Саватюгина\_\_\_**

(Группа) (Подпись, дата) (И.О.Фамилия)

Руководитель **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_ Гапанюк Ю.Е. \_\_\_**

(Подпись, дата) (И.О.Фамилия)

Консультант **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

(Подпись, дата) (И.О.Фамилия)

*2021 г.*

**Министерство науки и высшего образования Российской Федерации**

**Федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технический университет имени Н.Э. Баумана**

**(национальный исследовательский университет)»**

**(МГТУ им. Н.Э. Баумана)**

УТВЕРЖДАЮ

Заведующий кафедрой \_\_ИУ-5\_\_

(Индекс)

\_\_\_\_\_\_\_\_\_\_\_\_\_ В.М.Черненький

(И.О.Фамилия)

« \_\_\_\_\_ » \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г.

**ЗАДАНИЕ**

**на выполнение научно-исследовательской работы**

по теме Исследование возможностей нейронных сетей для распознавания продуктов питания по их фотографии в автоматизированных информационных системах обеспечения качества питания на основании физических параметров спортсменов

Студент группы ИУ5-31М

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Саватюгина Юлия Борисовна\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)

\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_учебная\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Источник тематики (кафедра, предприятие, НИР) \_\_\_НИР\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

График выполнения НИР: 25% к 4 нед., 50% к 8 нед., 75% к 12 нед., 100% к 16 нед.

Техническое задание  **\_\_**Произвести исследование возможностей нейронных сетей для распознавания продуктов питания по их фотографии для использования их в информационных системах обеспечения качества питания на основании физических параметров спортсменов.

***Оформление научно-исследовательской работы:***

Расчетно-пояснительная записка на 21 листе формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Семь (7) рисунков, девять (9) таблиц.\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дата выдачи задания « 8 » октября 2021 г.

**Руководитель НИР**  \_\_\_\_\_\_\_\_\_ Ю.Е. Гапанюк

(Подпись, дата И.О.Фамилия)

**Студент**  \_\_\_\_\_\_\_\_\_ Ю.Б.Саватюгина

(Подпись, дата И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

**ОГЛАВЛЕНИЕ**

[ЦЕЛЬ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ 3](#_Toc73211343)

[1. Изучение факторов, влияющих на использование приложений для обеспечения качества питания на основании физических параметров спортсменов 4](#_Toc73211344)

[2. Сравнение наиболее популярных приложений 11](#_Toc73211345)

[3. Расчет затрат, длительности и стоимости разработки 15](#_Toc73211346)

[3.1. Расчет затрат (с указанием значения масштабных факторов и параметров разработки). 15](#_Toc73211347)

[3.2. Расчет длительности и стоимости разработки. 16](#_Toc73211348)

[3.3. Зависимости между параметрами модели. 17](#_Toc73211349)

[3.4. График зависимости затрат и стоимости от варьируемого параметра модели. 18](#_Toc73211350)

[3.5. Расчет уменьшения затрат за счет функционала. 18](#_Toc73211351)

[3.6. Расчет уменьшения затрат за счет параметров модели 19](#_Toc73211352)

[ВЫВОДЫ 20](#_Toc73211353)

[СПИСОК ЛИТЕРАТУРЫ 21](#_Toc73211354)

# **ЦЕЛЬ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ**

Целью этой научно-исследовательской работы является изучение возможностей нейронных сетей для распознавания продуктов питания по их фотографии.

# **Введение**

Тема правильного питания вошла в обычную жизнь почти каждого человека. Люди придерживаются определенного пищевого поведения для поддержания фигуры и здоровья, тем самым обеспечивая себе высокое качество жизни.

Для того, чтобы узнать количество необходимых нутриентов необходимо тратить на расчёты большое количество свободного времени. Именно поэтому большинство людей для этой цели используют автоматизированные информационные системы обеспечения качества питания на основании физических параметров человека.

Подобные приложения также используются для подсчета полученных из пищи нутриентов, то есть человеку необходимо в списке продуктов искать необходимый, но эта задача сильно упростится при замене этого алгоритма на нейронную сеть, способную определить какой продукт перед пользователем.

# **Основные проблемы решения задачи распознавания продукта по фотографии**

Для начала поиска возможных решений данной идеи было решено выявить явные проблемы для машинного зрения в данной задаче.

Для точного расчета количества нутриентов в потребленной пище необходимо точно знать вес или объем и калорийность каждого продукта. Например, обычный помидор очень похож на томат черри, следовательно, программа может воспринять их как один продукт, однако объем их будет разным. Так появляется первая проблема данной задачи: распознавание объема или веса продукта по фотографии.

Так же очень важно знать калорийность продукта, однако, калорийности схожих по внешнему виду продуктов, например, мясо индейки и мясо курицы или филе семги и филе форели, сильно отличаются. Получается вторая проблема: продукты, чей внешний вид схож, будут неправильно распознаваться программой.

Основной же проблемой данной задачи является частое изменение внешнего вида тех или иных продуктов, их ребрендинг, что не позволяет быстро обучать нейросеть распознавать новые обертки продуктов.

А также, фотографии товаров могут содержать товары-аналоги, которые по характеристикам и стоимости соответствуют самому товару, но не являются им, и в этом случае содержат такую же метку, как и основной товар. Это означает, что:

1. Мы заставляем модель выучивать внешний вид сразу нескольких разных товаров.
2. Разные товары могут иметь общие аналоги и, соответственно, мы показываем нейронной сети один и тот же товар в разных классах и путаем её.
3. Менеджеры торговой сети сами решают какие у товара аналоги и часто их меняют, или могут начать этот товар продавать у себя, поэтому он перестанет быть аналогом.

Таким образом становится ясно, что для решения данной задачи необходимо использовать более сложный алгоритм, чем классификатор.

# **Исключение добавления новых товаров от релиза модели**

Для классических методов классификации требуется большой датасет, а обучать модель при таком количестве изображений можно достаточно долго. Поэтому необходимо придумать альтернативы.

Основная идея заключается в том, чтобы не решать каждый раз задачу классификации, а один раз обучить модель, которая будет возвращать векторное представление объекта, которое можно использовать для поиска в базе данных с уже рассчитанными векторами для каждого класса. Делать это можно с помощью различных вариаций [ArcFace](https://arxiv.org/pdf/1801.07698.pdf) и triplet loss. Такой подход называется Metric Learning.

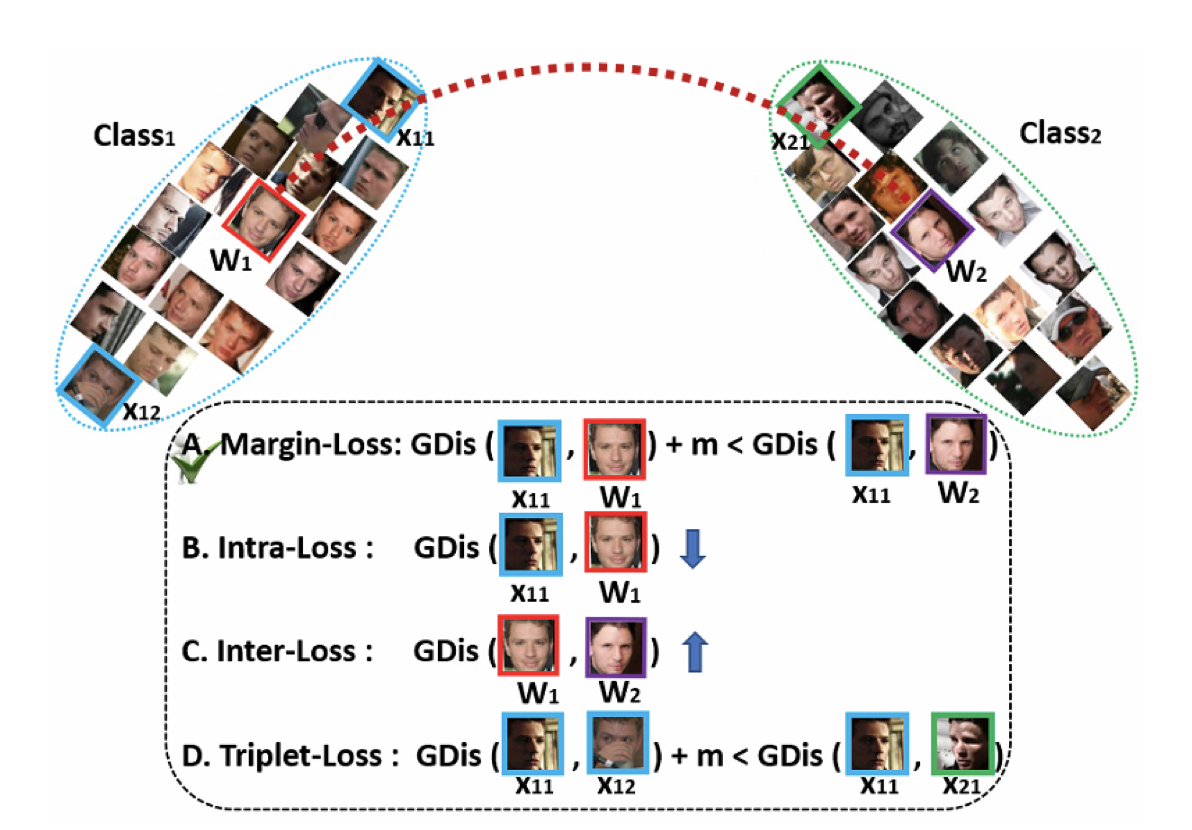


Рисунок 1. Metric Learning

Во избежание технических неудобств с выбором негативных и позитивных примеров для triplet loss, необходимо взять arcface, поскольку он хорошо показывает себя в аналогичной задаче в области распознавания лиц.

Использование Metric Learning позволяет уменьшить количество необходимых данных для обучения до 500 изображений на класс. Релиз новых товаров в системе происходит с помощью добавления их векторного представления в базу. Но их качество распознавания ниже, чем у тех, на которых модель обучалась. Такой подход позволяет делать прогноз даже тогда, когда модель не видела товар, а обучать ее можно гораздо быстрее.

Однако, есть проблема, которую даже Metric Learning методами тяжело решить — определение объемов. В повседневной жизни мы можем это сделать путем сравнения объектов между собой. Научить этому модель довольно тяжело. Порой одни и те же товары разного объема выглядят почти одинаково, а обучать нейронную сеть разделению таких изображений на два класса неправильно. Поэтому фотографии с разным объемом объединяются в один кластер, а в дальнейшем добавляется постобработка по уточнению объема товара. Например, уточнять объем можно по цене, либо по отношению размеров с другими объектами на фотографии.

# **Решение проблемы с товарами-аналогами**

Для быстрой размет необходимо объединять товары в группы по внешнему виду. В open source инструментах такого функционала нет, поэтому необходимо разработать свою платформу разметки. Ключевой должна быть функция “Найти похожие”.

Мы определяем некачественную фото по следующим признакам:

* Фотография размыта
* Фотография сделана с плохого ракурса
* Товар находится за дверью в холодильнике
* Товар находится не по центру кадра, дизайн упаковки сильно обрезан, невозможно идентифицировать товар

Для группировки товаров необходимо разработать функционал “Найти похожие”.  В одной группе могут быть разные вкусы, фото одно товара-аналога и тд. Эта часть позволяет сортировать фотографии в папке по “схожести” с выбранным изображением.

И с помощью такого функционала получается отделить товары-аналоги от основного товара.

После внедрения Metric Learning и описанной платформы разметки приходим к следующему порядку работы.

1. Проводим аналитику и собираем новые товары или те, которые плохо распознаются.
2. Выделяем из новых фото группы товаров.
3. Если нужно, дообучаем модель.
4. Векторизуем все группы товаров и добавляем в базу. Аналоги векторизуются отдельно, но им проставляется метка основного товара.
5. Когда приходит фото на распознавание, мы извлекаем из него вектор признаков и сравниваем с теми, которые есть в базе, и выбираем товар с наиболее похожим вектором.
6. И проверяем насколько большое расстояние у этого товара и отправленного изображения. Если оно больше определенного порога, мы понимаем, что такого товара у нас нет в базе.

И по итогу мы теперь можем на лету обновлять базу распознаваемых товаров и аналогов, а также увеличили качество распознавания товаров до 87,5%. А для устранения проблемы с объемами мы стали анализировать ценники, но об этом расскажем в другой раз.

# **Методы распознавания QR-кода продукта**

Еще одним способом избежать путаницы нейронной сети является простое распознавание QR-кода (баркода) необходимого продукта.

Существует 2 вида кодов:

* 1. Линейный (одномерный). Наиболее распространенный представитель — (в народе называемый «штрихкодом»).
  2. Двумерный. В свою очередь он делится на:

2.1. многоуровневый (stacked);

2.2. матричный (matrix).

Штрих-коды можно считывать с помощью оптического сканера штрих-кода, но здесь, для начала, напишем скрипт Python для считывания и декодирования штрих-кодов с рисунков, где они изображены.

Не столь часто используемый, но, как мне кажется, более ценный тип — тип матричного штрих-кода, который представляет собой машиночитаемую оптическую этикетку, которая содержит информацию об элементе, к которому она прикреплена. [На практике QR-коды](http://waksoft.susu.ru/53644d/CRFMEUtYSklbUVZGGAALCx8XEBlXDVREWF0fG0NcXw0NA1kCTA0XHxxSWFtRTFFUCFZXBg==/) часто содержат данные для локатора, идентификатора или трекера, который указывает на веб-сайт или приложение и т.д.

Штрихкод состоит из чередования объектов: полоски, точки и белое пространство. Все эти фигуры обладают разными геометрическими характеристиками. Получается матрица, которая может быть превращена в цифровой код при обработке.

**Алгоритм расшифровки QR-кода**

Полученный методами компьютерного зрения массив нулей и единиц будет расшифрован с помощью алгоритма, описанного в спецификации ISO/IEC 18004 Информационная технология - Автоматическая идентификация и сбор данных - Символика штрихового кода - QR Код. Опишем данный алгоритм:

1) получение массива "0" и "1" бит;

2) считывание информации о формате QR кода. Извлечение информации о маске (рисунок 2);

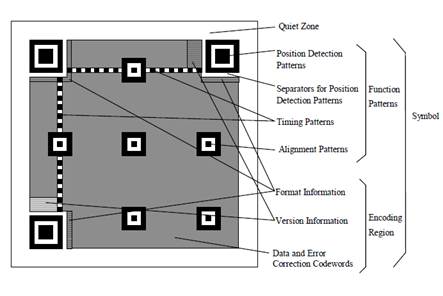


Рисунок 2. Состав QR кода

3) считывание информации о версии (где применимо), а затем определение версии кода;

4) применение к каждой ячейки массива маски посредством исключающего или. Всего существует 8 видов масок, значения бит которых определяются по формулам (1) - (8);

(i + j) mod 2 = 0, (1)

(i)mod 2 = 0, (2)

(j)mod 3 = 0, (3)

(i + j) mod 3 = 0, (4)

((i div 2) + (j div 3)) mod 2 = 0, (5)

(i j) mod 2 + (i j) mod 3 = 0, (6)

((i j) mod 2 + (i j) mod 3) mod 2 = 0, (7)

((i j) mod 3 + (i+j) mod 2) mod 2 = 0, (8)

где i и j - номера ячеек.

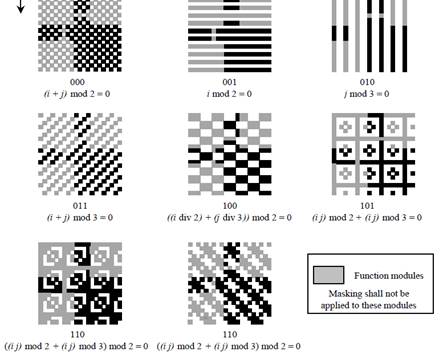


Рисунок 3. Виды масок

Применяем маску к полученному нами массиву посредством операции XOR (рисунок 4);

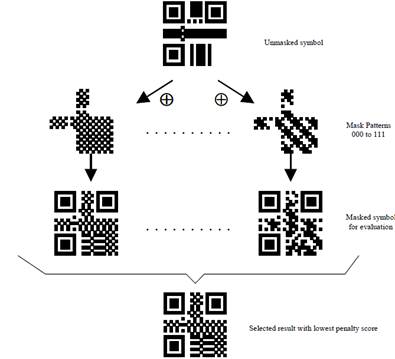


Рисунок 4. Применение маски

5) определение режима кодирования и считывание кода в соответствии с правилами размещения и восстановление данных и кодовых слов в сообщении.

Узнаем индикатор режима по четырем битам в правом нижнем углу и соответствующий ему режим по таблице 1.

В данной работе мы реализуем два режима: числовой и 8-битный. Цифровой режим кодирует данные из десятичного набора цифр (0 - 9) (ASCII значения 30HEX к 39HEX) при нормальной плотности 3 символа данных на 10 бит. В 8-битный режиме кодируется 8-битный Latin/Kana набор символов в соответствии с JIS X 0201 (символьные значения 00hex к FFhex). В этом режиме плотность данных составляет 8 бит / символ;

6) обнаружение ошибок с помощью коррекции ошибок кодовых слов. Если любая ошибка обнаружена, то исправить ее;

Таблица 1 - Режимы кодирования

|  |
| --- |
|  |
| Режим | индикатор |
| ECI | 0111 |
| Числовой | 0001 |
| Буквенно-цифровой | 0010 |
| 8-битный | 0100 |
| Kanji | 1000 |
| Structured Append | 0011 |
| FNC1 | 0101 |
| Разделитель | 0000 |

7) разделение кодовых слов данных на сегменты в зависимости от режима и показателями количество символов (рисунок 5).

Считывание происходит змейкой, начиная с правой нижней ячейки (рисунок 6). Старший значащий бит (бит 7) каждого кодового блока находится в первом доступном положении модуля;

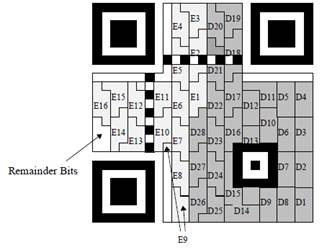


Рисунок 5. Сегменты с данными

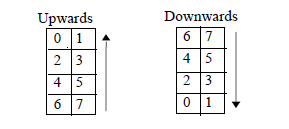


Рисунок 6. Считывание бит

8) наконец, декодирование данных в соответствии с режимом.

# **Заключение**

Задача упрощения работы пользователя с системой стояла в нахождении способов санирования или фотографирования продуктов, принимаемых в пищу. В ходе данного исследования было получено несколько вариантов решения данной проблемы, а точнее: определение продукта по фотографии (все еще не самый точный метод) и сканирование QR-кода продукта, что часто используется в аналогах автоматизированной информационной системы обеспечения качества питания на основании физических параметров спортсменов.

# **Список литературы**

1. Гапанюк Ю.Е. Машинное обучение; Лекционный материал, 2019.
2. Микеев К.К. Computer vivion в помощь декодирования QR-кода на Python. /NTA., Спб, 2021. – 5 с.
3. Тарик Рашид Создаем нейронную сеть./Диалектика.,М, 2020. – 272 с.
4. Рассел С., Норвиг П. Искусственный интеллект: современный подход./Williams., М. 2019. – 1408 с.
5. Дэвенпорт Т. Внедрение искусственного интеллекта в бизнес-практику: преимущества и сложности./Альпина Паблишер., М. 2021. – 316 с.
6. Джоши П. Искусственный интеллект с примерами на Python./Диалектика., М., 2019. – 448 с.