**Министерство образования и науки Российской Федерации**

**(МИНОБРНАУКИ РОССИИ)**

**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (ТГУ)**

**Автономные магистерские программы**

**Интеллектуальный анализ больших данных**

**ОТЧЕТ**

**о выполнении технического задания**

**Выполнил**

**студент группы №148БИ**

**\_\_\_\_\_\_\_\_\_\_\_\_\_О.А. Майер**

**Проверил д-р. техн. наук**

**доцент кафедры**

**теоретических основ**

**информатики**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_А.В.Замятин**

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

**Отчет защищен**

**с оценкой\_\_\_\_\_\_\_\_\_\_\_**

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

**Томск 2018**

Оглавление

[Введение 3](#_Toc531111934)

[Предметная область 3](#_Toc531111935)

[Описание исходных данных 4](#_Toc531111936)

[Предобработка исходных изображений 5](#_Toc531111937)

[Построение архитектуры модели 6](#_Toc531111938)

[Интерпретация полученных результатов 7](#_Toc531111939)

[Выводы. 8](#_Toc531111940)

# Введение

Актуальность выполняемой задачи

Предметная область

Цель выполняемой работы, Система поддержки принятия решений при постановке анамнеза СППА

Краткое описание выполняемых работ

## Предметная область

Виды поражений кожи, для определения моделью

1) Actinic keratoses and intraepithelial carcinoma / Bowen's disease - Актинические кератозы, интраэпителиальная карцинома и болезнь Боуэна

2) basal cell carcinoma – Базальноклеточная карцинома: постоянное обновление клеточного состава за счет непрерывного размножения клеток базального (самого глубокого) слоя.

3) benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses)

4) dermatofibroma – Фиброма кожи: Фиброма кожи — это доброкачественная опухоль, которая состоит из соединительной ткани. Она располагается под верхним слоем кожи. Фиброма практически не представляет опасности в плане перерождения в злокачественную опухоль. Но она является косметическим дефектом и доставляет человеку дискомфорт.

5) melanoma - злокачественная опухоль, развивающаяся из меланоцитов — пигментных клеток, продуцирующих меланины. Наряду с плоскоклеточным и базальноклеточным раком кожи относится к злокачественным опухолям кожи. Преимущественно локализуется в коже, реже — сетчатке глаза, слизистых оболочках (полость рта, влагалище, прямая кишка). Одна из наиболее опасных злокачественных опухолей человека, часто рецидивирующая и метастазирующая лимфогенным и гематогенным путём почти во все органы.

6) melanocytic nevi - Меланоцитарные невусы (или родинки) представлены многими разновидностями и есть почти у всех людей – они выявляются у ¾ части населения земного шара. Такие образования на коже являются доброкачественными опухолями

7) vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage) - К сосудистым поражениям относятся приобретенные (например, пиогенная гранулема) и врожденные или проявляющиеся вскоре после рождения (сосудистые родимые пятна). К сосудистым родимым пятнам относятся сосудистые опухоли (например, врожденная гемангиома) и сосудистые мальформации.

Виды диагностики поражений кожи:

Пока не понятно как использовать эти данные.

1) Гистопатология - это раздел микроскопического изучения поражённой ткани, является важным инструментом патоморфологии (патологическая анатомия), так как точный диагноз рака и других заболеваний обычно требует гистопатологического исследования образцов.

2) Последующая проверка – потверждение

3) заключение экспертного консенсунса

4) Конфокальная Микроскопия - в настоящее время под этим термином чаще всего подразумевают Конфокальную Лазерную Сканирующую Микроскопию (КЛСМ, Confocal laser scanning microscopy, CLSM), которая представляет собой разновидность световой оптической микроскопии, обладающей значительным контрастом и пространственным разрешением по сравнению с классической световой микроскопией, что достигается использованием точечной диафрагмы (пинхол, pinhole), размещённой в плоскости изображения и ограничивающей поток фонового рассеянного света излучаемого не из фокальной плоскости объектива[1]. Это позволяет получить серии изображений на различных глубинах фокальной плоскости внутри образца (т. н. оптическое секционирование образца по глубине), и затем реконструировать трехмерное изображение образца из этих серий.

Местоположение на теле

• Scalp – кожа черепа

• Ear – ухо

• Face - лицо

• Back - спина

• Trunk – грудь

• Chest – плечи

• Upper extremity - Верхняя конечность

• Abdomen – брюшная полость

• Unknown – неизвестно

• Lower extremity – нижняя конечность

• Genital – гениталии

• Neck – шея

• Hand – руки

• Foot – нога

• Acral – стопf

# Описание исходных данных

Дополнительные данные и csv

Условное обозначение целевого признака

Примеры изображений

Размерность

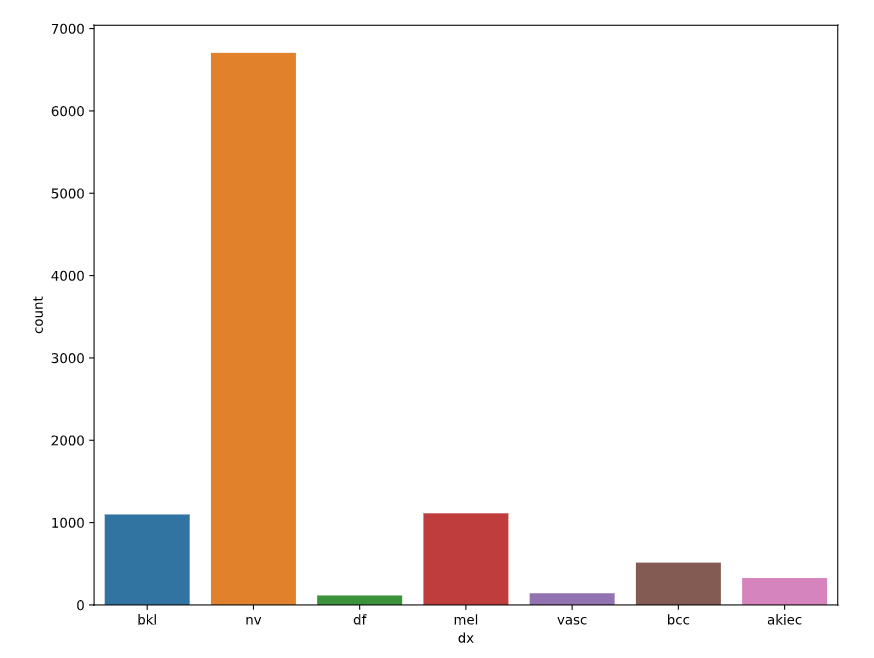


Рисунок 4. Гистограмма распределения целевого признака.

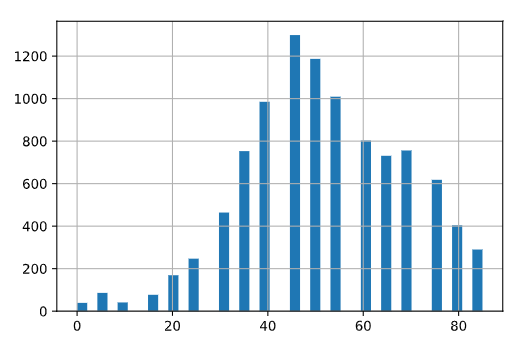


Рисунок 5. Тепловая карта корреляций.

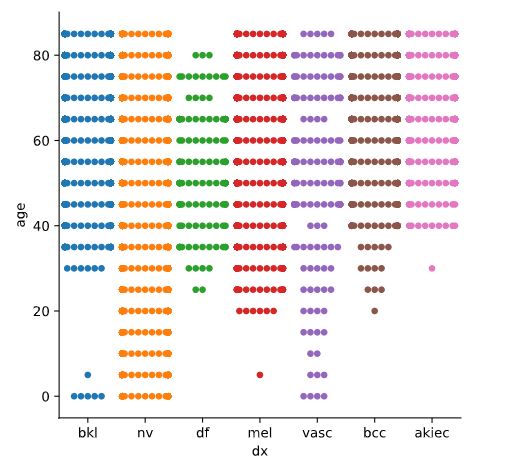


Рисунок 4. Корректно назвать .

# Предобработка исходных изображений

Считывание данных

Уменьшение количества классов для поиска ввиду – низкого количества экземпляров на 3х классах

Image Augmentaion – увеличение выборки

Трансформация выборки

Нормализация векторов изображений

Работа с билиотекой OPENCV2 для изображений

Рзделение данных на тестовую и валидационную выборку

One hot encoding для задачи классиффикации

Аугментация далее используется так как данное действие хорошо себя показывает на всех видах модели

Выделение дополнительных каналов изображений

ADAPTIVE\_THRESH\_GAUSSIAN

Channel

ADAPTIVE\_THRESH\_MEAN

Channel

Gray channel

Общая размерность выборки

Снижение размерности изображений

# Построение архитектуры модели

Определение используемых метрик модели

Определение используемых оптимизаторов модели

Adam и RMSprop

## Построение топологии нейронной сети

Процесс перехода от одной версии к другой

Добавление

## Топология нейросетевых моделей

В процессе работы было построено и протестировано 4 основных топологии нейросети

УКАЗАТЬ ИХ

Виды топологии сети

* Топлогия A – готовая топология сети, используемая для распознавания рукописного текста
* Топология B – Улучшенная версия топологии A с 2мя слоями свертки и 2 полносвязными слоями на выходе
* Топология С – Собственная топология ОПИСАНИЕ +картинка
* Топология D – Собственная топология ОПИСАНИЕ + картинка

## Проведение экспериментов

Табличка

Про каждый эксперимент, обозначенный как v1….

Указать параметры при эксперименте и провести его заново 10 раз посчитав средним все параметры

В приложении указать табличку которая пойдёт в презентацию

Основываясь на реферате начать прогонять каждую модель по параметрам

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Характеристики  Модели | V1.0 | V1.1 | V1.2 | V1.3 | V1.4 |  |  |
| Количество классов на выходе | 7 | 7 | 3 | 4 | 4 |  |  |
| Вид топологии | A | B | C | D | D |  |  |
| Gray channel | - | - | - | - | - |  |  |
| ADAPTIVE\_THRESH\_  GAUSSIAN  Channel | - | - | - | - | - |  |  |
| ADAPTIVE\_THRESH\_  MEAN  Channel | - | - | - | - | - |  |  |
| RGB Channel | + | + | + | + | + |  |  |
| Augmentation | - | - | - | + | + |  |  |
| Нормализация данных | - | - | - | - | + |  |  |
| Вид оптимизирующей функции | Adam | Adam | Adam | Adam | Adam |  |  |
| Общая размерность выборки |  |  |  |  |  |  |  |
| Размерность изображений |  |  |  |  |  |  |  |
| Общее количество связей в сети |  |  |  |  |  |  |  |

# Интерпретация полученных результатов

# Создание СППА.

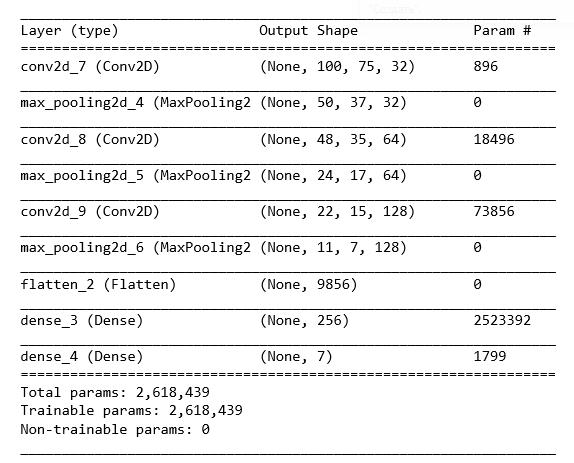
При выполнении работы были рассмотрены и применены на практике ряд наиболее используемых алгоритмов машинного обучения. Изучены их положительные и отрицательные стороны.

# Выводы.

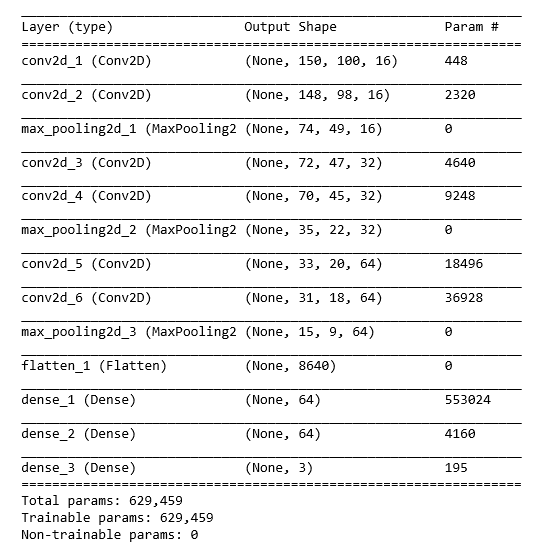
При выполнении работы были рассмотрены и применены на практике ряд наиболее используемых алгоритмов машинного обучения. Изучены их положительные и отрицательные стороны.

# Приложение А – Топологии нейронных сетей

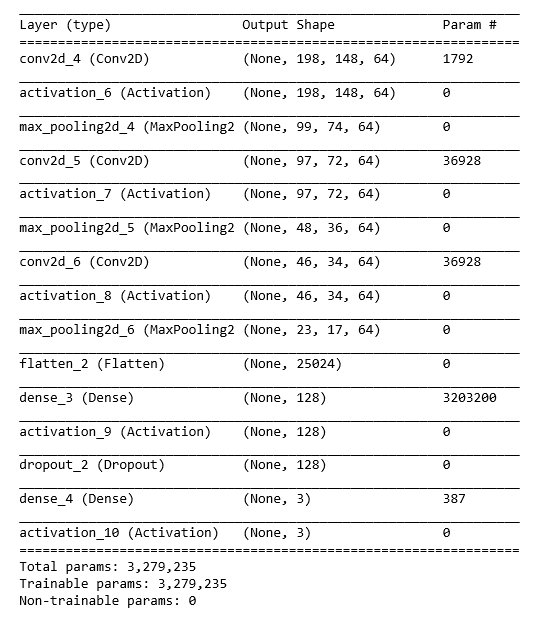
Топология A



Топология B



Топология C



Топология D

