|  |
| --- |
| Пермский филиал федерального государственного автономного образовательного учреждения высшего образования  «Национальный исследовательский университет  «Высшая школа экономики»  *Факультет экономики, менеджмента и бизнес-информатики* |
|  |
| Мамедов Нурлан Рагим оглы  **ОБНАРУЖЕНИЕ ОЧАГОВ ОСТРОГО ИНСУЛЬТА НА СНИМКАХ МРТ ПОСРЕДСТВОМ НЕЙРОННЫХ СЕТЕЙ И АЛГОРИТМА C ИСПОЛЬЗОВАНИЕМ ГИСТОГРАММ НАПРАВЛЕННЫХ ГРАДИЕНТОВ.**  *Курсовая работа*  по направлению подготовки *09.03.04 Программная инженерия*  образовательная программа «Программная инженерия»   |  |  | | --- | --- | |  | Руководитель  Доктор наук, доцент НИУ ВШЭ в Перми  Софья Петровна  Куликова | |

Пермь, 2019 год

# Содержание

[Введение 2](#_Toc35891971)

[Обзор существующих решений 3](#_Toc35891972)

[Этапы классификации 5](#_Toc35891973)

[Результаты решения 5](#_Toc35891974)

[Данные и методы исследования 7](#_Toc35891975)

[Нейронные сети 7](#_Toc35891976)

[Принцип работы нейронной сети 7](#_Toc35891977)

[Функции активации 8](#_Toc35891978)

[Обучение нейронной сети 9](#_Toc35891979)

[Метод обратного распространения ошибки 11](#_Toc35891980)

[Свёрточные нейронные сети 12](#_Toc35891981)

[Сбор данных для обучения 14](#_Toc35891982)

[Гистограммы направленных градиентов 16](#_Toc35891983)

[Принцип работы 16](#_Toc35891984)

[Программная реализация 16](#_Toc35891985)

[Требования к функциональности 16](#_Toc35891986)

[Выбор инструментов и технологий 17](#_Toc35891987)

[План реализации 17](#_Toc35891988)

[Схема работы распознавания очага 18](#_Toc35891989)

[Проектирование графического интерфейса программы 18](#_Toc35891990)

[Результаты 19](#_Toc35891991)

[Проблемы, возникшие во время разработки 19](#_Toc35891992)

[Метод выделения краёв очагов инсульта 20](#_Toc35891993)

[Сравнение результата с существующими решениями 21](#_Toc35891994)

# Введение

Инсульт – это острое нарушение кровообращения мозга, представляющее собой одну из основных причин смертности и инвалидности в мире. Умение точно и своевременно выявлять очаги повреждений чрезвычайно важно для назначения правильного лечения и при выборе реабилитационных подходов. В современном мире для этих целей широко используется магнитно-резонансная томография (МРТ).

В настоящее время при выявлении и анализе очагов инсульта используется практика обведения вручную на каждом срезе МРТ изображения экспертом врачом-радиологом. Такой подход требует много времени, а также специальной подготовки и характеризуются высокой вариабельностью результатов в связи с различным опытом экспертов.

Создание решения в области автоматического распознавания очагов инсульта на снимках МРТ позволило бы:

1. уменьшить время обнаружения очага и определения его границ, а значит, снизить стоимость и ускорить лечение и реабилитацию пациентов.
2. снизить стоимость и ускорить процесс профессиональной подготовки специалистов в области «неврология» и/или «радиология», за счет использования цифровых ассистентов.
3. уменьшить вариабельность результатов, связанную с различной экспертной подготовкой, и повысить их воспроизводимость.
4. автоматизировать обработку большого объема данных. Вручную это было бы слишком времязатратно.

Все вышеобозначенные преимущества обуславливают необходимость разработки программных инструментов для автоматического распознавания очагов острого инсульта на МРТ снимках. Однако существующие решения пока не достигли достаточной точности для их рутинного применения в клинической практике. Одним из направлений для улучшения методов автоматического распознавания очагов является использование сверточных нейронных сетей (СНС), которые широко используются для автоматического распознавания образов и, следовательно, задач сегментации, в том числе в области анализа медицинских изображений. В то же время, ни одна из архитектур СНС, разработанных до настоящего времени, не достигла высокой точности при сегментации поражений инсультом, по причине их неоднородности по расположению, форме, размеру, интенсивности изображения и текстуре. Целью данной работы является реализация алгоритмов распознавания очагов острого инсульта и исследование возможности улучшения точности работы алгоритмов на основе.

# Обзор существующих решений

Многообещающее решение было представлено в статье «Stroke Lesion Segmentation in FLAIR MRI Datasets Using Customized Markov Random Fields»[[1]](#footnote-1). Исследователи используют для обнаружения очагов на снимке МРТ Байесовскую классификацию на основе текстурной информации Габора и случайные поля Маркова.

*Байесовский классификатор* – это классификатор, основанный на формуле Байеса о вероятности [(1)](#Формула_Байеса) .

|  |  |
| --- | --- |
|  |  |

Где P (A) – это вероятность гипотезы А;

P (B) – полная вероятность события B;

P (A | B) – вероятность гипотезы A при событии B;

P (B | A) – вероятность наступления события B при гипотезе A.

*Текстурная информация Габора* – это информация, полученная путём применения фильтра Габора на изображение.

*Фильтр Габора* – линейный фильтр, представленный формулой [(2)](#Фильтр_Габора). Используется для выделения краёв в области изображения.

|  |  |
| --- | --- |
|  |  |



Рисунок 1 Пример применения фильтра Габора

*Случайные поля Маркова –* это графовая модель, которая используется для представления совместных распределений набора нескольких случайных переменных. Неориентированный граф, где каждая вершина является случайной переменной Х и каждое ребро представляет собой зависимость между случайными величинами u и v.

## Этапы классификации

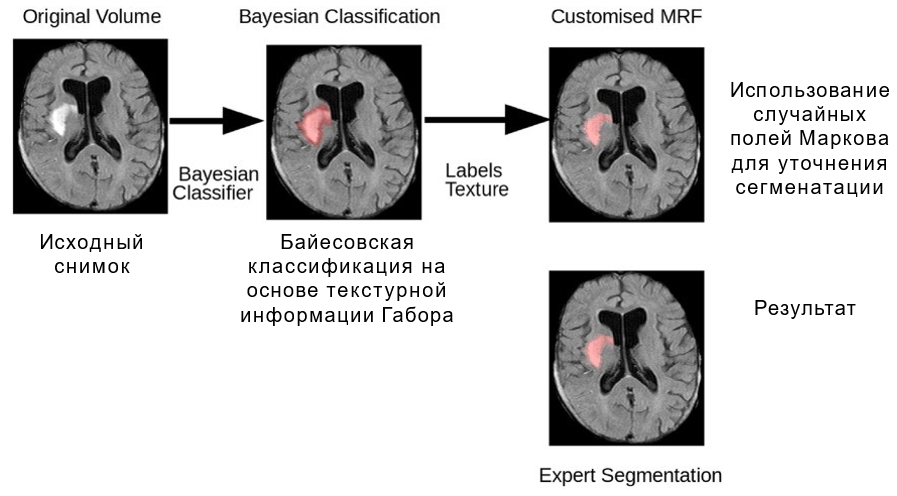
****

Рисунок Этапы классификации

## Результаты решения

Исследователи сравнили коэффициент Дайса с другими решениями на основе различных наборов изображений.

*Коэффициент Дайса* – это удвоенная область перекрытия изображений области распознавания и самого объекта, делённая на общее число пикселей на обоих изображениях.

|  |  |  |  |
| --- | --- | --- | --- |
| Набор изображений | Количество изображений | Предложенный в исследовании метод | СНС |
| SFB | 23 | 0.621 ± 0.260 | 0.572 ± 0.230 |
| I-Know | 102 | 0.583 ± 0.236 | 0.535 ± 0.301 |
| ISLES | 26 | 0.544 ± 0.280 | 0.541 ± 0.270 |
| Wtd. Avg. | 151 | 0.582 ± 0.250 | 0.541 ± 0.272 |

Таблица 1 Сравнение предложенного метода на различных наборах изображений

Также стоит упомянуть решения: «Novel Screening Tool for Stroke Using Artificial Neural Network» 24 апреля 2017 г. [[2]](#footnote-2); «Stroke Lesion Detection Using Convolutional Neural Networks», представленный в марте 2018 на конференции «2018 International Joint Conference on Neural Networks» [[3]](#footnote-3); «Classification of stroke disease using convolutional neural network» представленный в “Journal of Physics: Conference Series” [[4]](#footnote-4). Все эти исследования используют в качестве основного инструмента различные вариации свёрточных нейронных сетей.

На данный момент в сообществе крайне мало решений, позволяющих распознавать очаги инсульта на снимках МРТ типа FLAIR.

# Данные и методы исследования

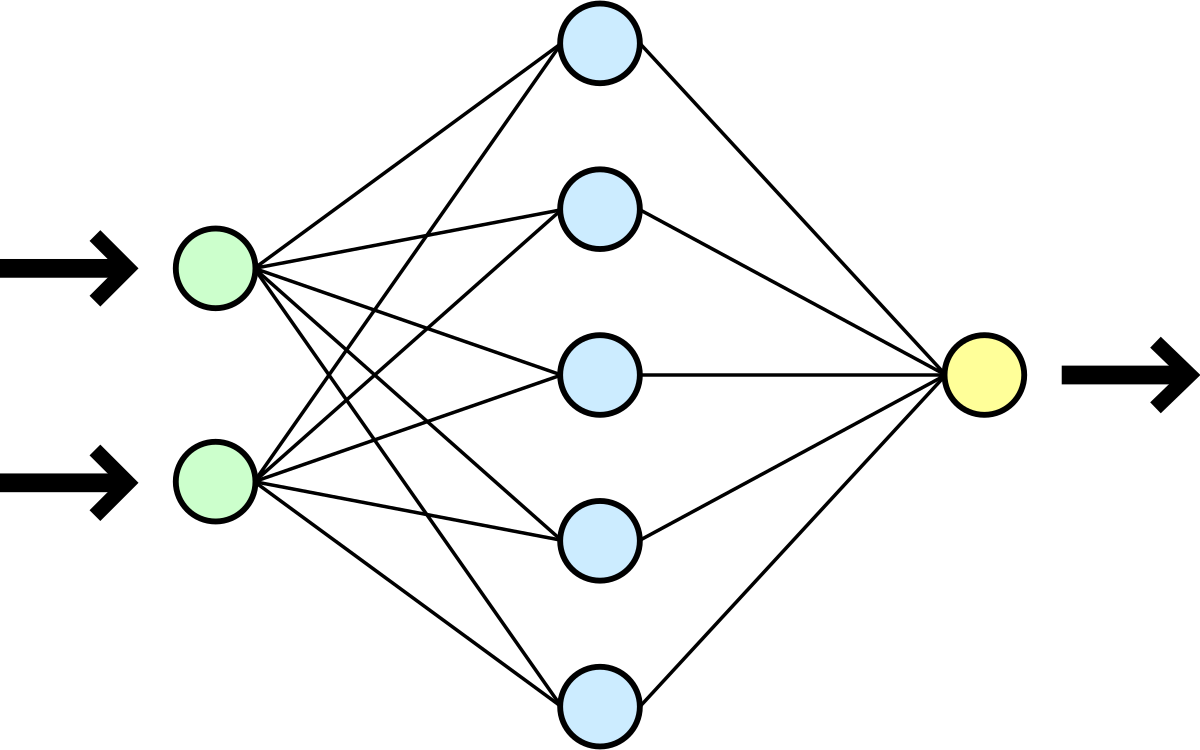
## Нейронные сети

Основным инструментом для создания системы по распознаванию очагов острого инсульта являются свёрточные нейронные сети.

Для начала стоит определить понятие *нейронная сеть*.

*Нейронная сеть* – это математическое представление сети биологических нейронов, целью которых является преобразование множества сигналов в один сигнал (решение). Каждый нейрон обладает дендритами и аксонами. Дендриты обеспечивают сбор сигналов от других нейронов, нейрон обрабатывает их, по аксонам сигнал передаётся в другие нейроны. Сигнал, передающийся по аксону, обрабатывается нейроном и уже в зависимости от его «функции активации», которая производит манипуляции с сигналом и нормализует его. Также каждый нейрон обладает «весом», значение которого подбирается во время обучения и влияет на усиление/ослабление сигнала.

Рисунок 3 Упрощенная схема нейронной сети



Выходной слой

Слой нейронов

Входной слой

## Принцип работы нейронной сети

Информация поступает на входной слой, разделяется, затем каждая часть поступает каждому нейрону и в нём вычисляется значение. К примеру, на [рисунке 4](#Рисунок_Схема_Многослойной_НС) в нейроне значение будет равно , где – это взвешенное значение, а F – это функция активации. В остальных нейронах первого слоя значения получается соответственно, далее эти значения передадутся следующему слою и так до тех пор, пока не будет вычислено значение выходного слоя, на основе которого уже будет сделано решение.

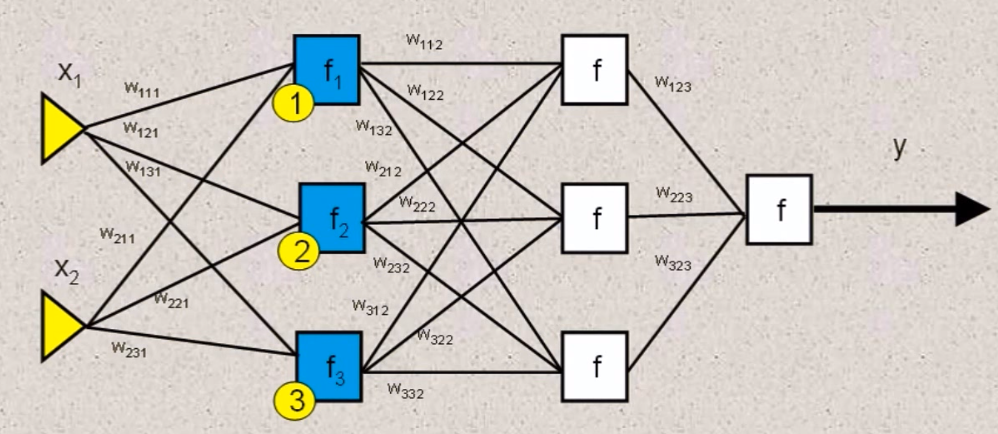


Рисунок 4 Схема многослойной нейронной сети

### Функции активации

Функция активации вычисляет выходное значение нейрона в зависимости от взвешенной суммы. Функция может быть разной и настраивается разработчиком.

Одной из самых распространенных функций является [*сигмоида*](#Формула_Сигмоиды)[(3).](#Формула_Сигмоиды)

|  |  |
| --- | --- |
|  |  |

Функция принимает на входе вычисленное значение, а на выходе дает вещественное число в интервале от 0 до 1. Сигмоида получила широкое распространение, поскольку имеет легкую интерпретацию: 0 – отсутствие активации, 1 – полная активация. Недостатком использования функции сигмоиды является то, что в крайних значениях 0 и 1, градиент становится близок нулю. Это нежелательный исход, поскольку при обучении нейронной сети часто используется метод обратного распространения ошибки, а в нём локальный градиент умножается на общий градиент. Следовательно, он фактически обнуляет общий градиент, а значит сигнал не будет проходить к его весам и обратно к данным. Подробно про обучение описано в разделе[*Метод обратного распространения ошибки*](#_Метод_обратного_распространения)*.*

Следующая распространённая функция [гиперболический тангенс (4)](#Формула_Гиперболического_тангенса)

|  |  |
| --- | --- |
|  |  |

Принимает на вход произвольное вещественное число, но в отличие от сигмоиды, на выходе даёт число от минус 1 до 1. Следовательно, выход функции центрирован относительно нуля.

## Обучение нейронной сети

Классическое определения понятия обучаемости нейронной сети даёт Томас Митчелл в книге «Машинное обучение»: *«Компьютерная программа обучается по мере накопления опыта относительно некоторого класса задач T и целевой функции P, если качество решения этих задач (относительно P) улучшается с получением нового опыта».*

Это определение даёт нам понять, что главное в обучении не данные, а функция. Она определяет данные для обучения в сети в дальнейшем.

Существует несколько форм обучения нейронной сети:

1. Обучение с учителем
2. Обучение без учителя
3. Обучение с подкреплением

При обучении с учителем формируется набор тренировочных данных (на англ. *training set* или *dataset*). Задача состоит в формировании новых ответов на основе имеющегося опыта. Предполагается, что новые данные будут похожими на данные из тренировочного набора.

Задачи обучения с учителем обычно делятся на задачи *классификации* и задачи *получения новых данных*. В задаче классификации требуется определить входной объект в один из конечного числа классов. К примеру, определение области, в которой на изображении находится машина или дорожный знак. К задачам получения новых данных можно отнести *предсказание.* Это предполагает получение новых данных на основе входной информации. К примеру, сделать прогноз погоды на основе количества осадков в предыдущие дни. Все эти задачи требуют предварительно размеченного набора данных. На рисунке 7 изображен пример выделения очага инсульта. Слева исходный снимок, справа снимок с выделенной областью очага. Именно эта форма обучения используется в программной реализации системы распознавания очагов инсульта на снимках МРТ.

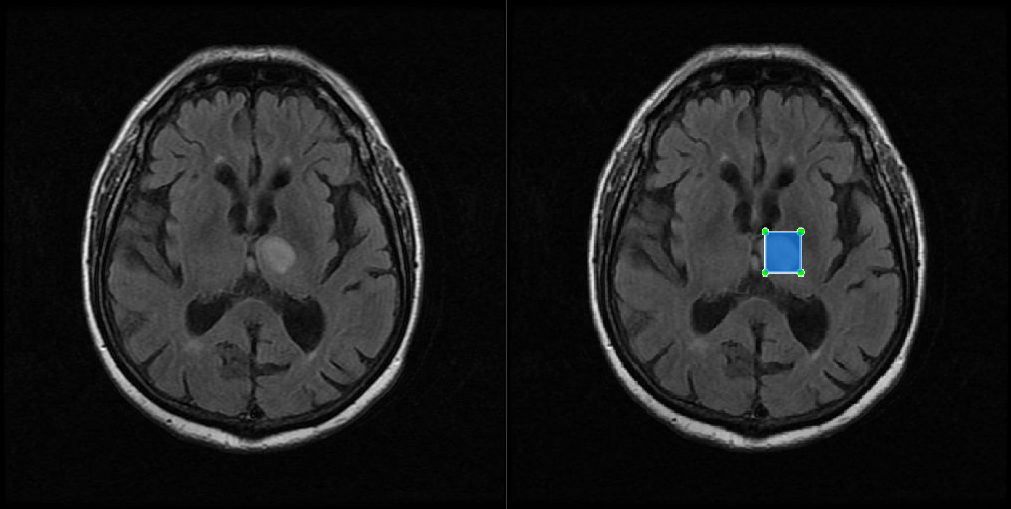


Рисунок Пример размеченных данных

Задачи обучения без учителя обычно появляются, когда размеченного набора данных нет. То есть, нейронной сети требуется найти какие-либо схожие признаки объектов и классифицировать их.

Задачи обучения с подкреплением одни из самых близких живым организмам. К примеру, дрессировка животных. Чтобы обучить собаку сидеть по команде, требуется давать ей лакомство каждый раз, когда она сидит и в это время говорить команду. Собаке не требуется тысяча изображений сидящих собак. Она сначала случайным образом садиться и получает вознаграждение, тем самым понимает, что если выполнять какое-то действие, то можно получать больше лакомства . Если перевести это на нейронную сеть, то ей требуется сначала случайным образом взаимодействовать с данными и получать за них «вознаграждение», если они близки к нужному результату. Затем нейронная сеть меняет свои методы – целевую функцию и максимизирует награду.

### Метод обратного распространения ошибки

Метод обратного распространения ошибки – это один из методов вычисления значения весов нейронов.

Работает следующим образом: нейронная сеть получает какой-либо результат и затем сравнивает с результатом на размеченных данных. После этого вычисляется ошибка – она равна текущий результат минус ожидаемый результат. Далее высчитывается дельта весов – это ошибка, умноженная на дифференциал производной функции активации. Чтобы откорректировать веса, нужно идти в обратную сторону от результата. Допустим, что мы корректируем нейрон, находящийся перед результатом. Тогда его вес будет равен текущий вес минус значение нейрона, умноженное на дельту весов и скорость обучения. Значение скорости обучения подбирается опытным путём и выбирает его разработчик. Далее соответственно формула применяется на остальные значения весов в нейронной сети. Тем самым вычисляются оптимальные значения для них.

|  |  |
| --- | --- |
|  |  |

## Свёрточные нейронные сети

*Свёрточные нейронные сети* – это сети, основная идея которых заключается в повторном использовании одних и тех же частей нейронной сети для работы с небольшими участками входных данных. В основном используются для распознавания объектов на изображении.

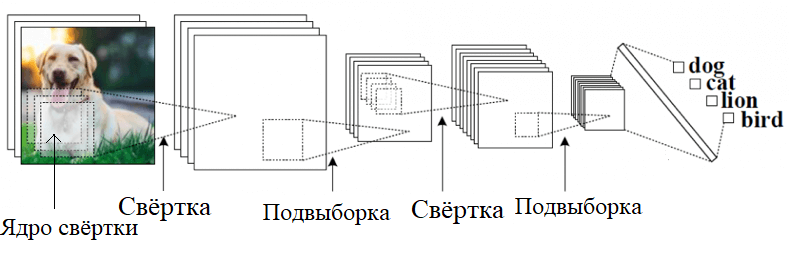


Рисунок 6 Схема свёрточной сети 1

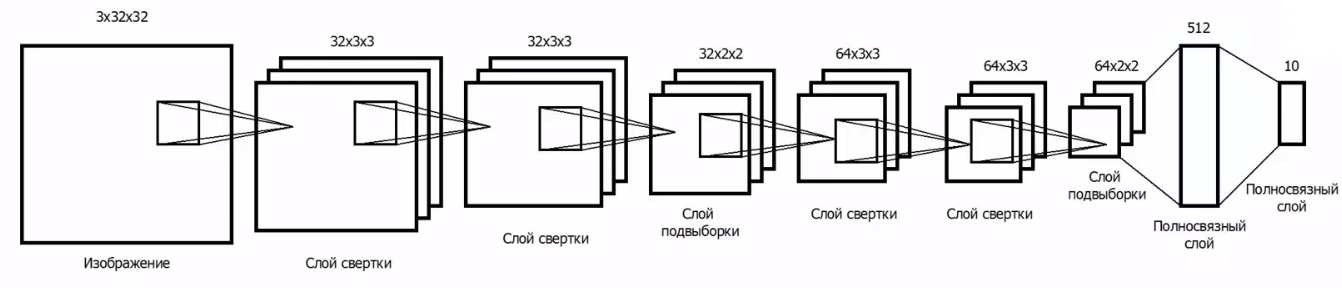


Рисунок 7 Схема свёрточной нейронной сети 2

Принцип работы заключается в проходе изображения по свёрточным слоям, слоям объединения и полносвязным слоям, а затем генерируется вывод [(рис. 7).](#Схема_СНС)

Свёрточный слой – это слой, который работает с небольшой частью изображения. Эта часть перемножается с ядром свёртки, полученные значения суммируются, тем самым получается одно значение [(рис. 8).](#Схема_работы_сверт_слоя)

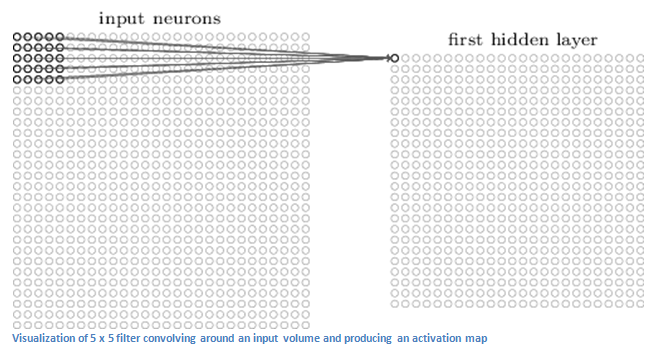


Рисунок 8 Схема работы свёрточного слоя (из книги “Neural Networks and Deep Learning” автора Майкла Нилсена)

Суть свёрточного слоя заключается в том, что если часть изображения в общих чертах похожа на искомую фигуру, которая определяется ядром свёртки, то и её сумма умноженных значений будет больше.

Ядро свёртки (фильтр свёртки) – это матрица такого же размера, что и рассматриваемая небольшая часть изображения. Ядра свёртки определяются автоматически во время обучения.

После первого свёрточного слоя будут проверяться признаки, на основе предыдущего слоя, следовательно более обобщенные данные, пример на [рисунке 9.](#Признаки_на_разных_уровнях)



Рисунок 9 Признаки на разные уровнях свёрточной сети

Данные со свёрточных слоёв поступают на полносвязный слой.

Полносвязный слой – это слой, который выводит данные на K-мерный вектор, где K – это количество классов распознавания. Далее программа выбирает нужный. Допустим, есть сеть, которая распознаёт кошек, собак и жаб на изображении. Тогда соответственно итоговый вектор будет [0.6, 0.99, 0.1], если на фотографии собака. Это означает, что сеть считает вероятность существования собаки на изображении – 99%, кошки 60% и жабы 10%.

Преимущества использования свёрточной нейронной сети:

1. Одно ядро весов используется для всего изображения, вместо каждого пикселя.
2. Свёртка сохраняет структуру входа, так как применяется к каждому входу по отдельности.
3. Возможность распараллеливания вычислений.
4. Для обучения используется метод обратного распространения ошибки.
5. Стойкость к небольшим поворотам и изменениям изображения.
6. Множество готовых архитектур для прикладного использования.

## Сбор данных для обучения

Данные очень важны для обучения, от них зависит, как нейронная сеть будет реагировать на различный вход. В задачах классификации на изображении это способность сети распознавать объекты под разными углами, тенями, освещением, бликами и деформациями. Входными данными может быть всё, что можно представить в цифровом варианте – изображения (как матрицы), таблицы, массивы чисел и прочее. Во время набора данных важно проследить за тем, чтобы признаки, выделяемые на различных изображениях, были разными, иначе набор не будет иметь никакой статистической значимости. Очень важно, чтобы данные были вариативны и содержали в себе случаи с искажениями и неточностями. К примеру, набор данных для распознавания карт и их масти представлен на [рисунке 8](#Пример_Набора). Для создания системы распознавания очагов инсульта использовались снимки МРТ, то есть изображения.



Рисунок Пример набора данных

Изображения для обучения с учителем обязательно должны быть размечены. Это означает, что на изображении должно быть указано, какой объект на изображении и где он находится. Кроме этого, для обучения требуется набор тестовых данных, который обычно является 20% от всего набора. Они позволяют нейронной сети уже во время обучения проверять прогресс и корректировать показатели.

Для задачи распознавания очагов острого инсульта на снимках МРТ было получены снимки реальных пациентов и из них отобраны и размечены 706 снимков. Пример размеченных данных был представлен на [рисунке 7](#Пример_размеченного_снимка).

## Гистограммы направленных градиентов

Гистограмма направленных градиентов – это способ представления графического объекта на основе выделяющихся признаков. Используется в компьютерном зрении для выделения краёв и распознавания объектов на изображении.

### Принцип работы

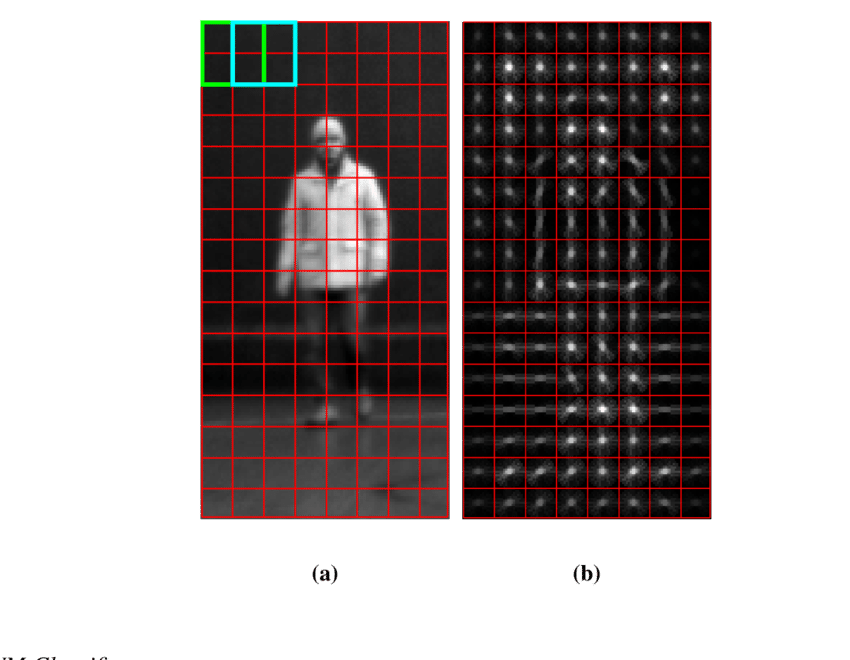
Принцип работы заключается в разделении изображения на множество небольших участков и высчитывания направления пикселей в нём.

Рисунок 11 Пример гистограммы направленных градиентов

После вычисления гистограммы становится несложным выделение краёв объекта. В программной реализации системы распознавания очагов инсульта использовались гистограммы направленных градиентов для уточнения краёв и выделения цветом самих очагов.

# Программная реализация

## Требования к функциональности

1. Загрузка снимка МРТ в формате jpg.
2. Обработка фотографии – выделение очагов на снимке цветом.
3. Возможность скачивать обработанное изображение.

## Выбор инструментов и технологий

В качестве языка программирования для программы используется Python 3.6.

Выбор языка обусловлен следующими факторами:

1. Python имеет широкое сообщество людей в теме нейронных сетей и машинного обучения.
2. Python обладает важными для сложных вычислений библиотеками NumPy, SciPy и TensorFlow.
3. Библиотека компьютерного зрения и работы с изображениями OpenCV предоставляет удобный интерфейс взаимодействия с ней.
4. Python реализует протокол буфера памяти, который позволяет различным типам данных получать доступ к другим типам на уровне байтов. Примером использования является тип данных изображения OpenCV, который представляет доступ к изображению через массив NumPy. Это позволяет ускорить вычисления и оптимизировать работу.
5. Python обладает механизмом совместимости с языками C/C++, что позволяет ускорить вычисления и при этом сохранить лаконичность написания алгоритмов.

## План реализации

1. Определение набора инструментов для программной реализации.
2. Выбор модели для обучения свёрточной нейронной сети.
3. Настройка программного окружения для разработки.
4. Сбор снимков МРТ для тренировки свёрточной нейронной сети.
5. Разметка данных – выделение очага на снимке МРТ и сохранение данных о выделении на всех снимках.
6. Обучение нейронной сети на размеченных данных.
7. Сохранение обученной модели.
8. Создание тестовой программы для проверки нейронной сети.
9. Проектировка интерфейса.
10. Привязка функций к элементам интерфейса.
11. Добавление функцию сохранения обработанного изображения.

## Схема работы распознавания очага

Загрузка изображения

Поиск областей очагов на снимке

Уточнение краёв в области очага

Выделение очагов цветом

## Проектирование графического интерфейса программы

Интерфейс должен содержать следующие элементы:

1. Кнопку для загрузки изображения. При нажатии на неё должно открываться диалоговое окно с выбором файла.
2. Поле для вывода загруженного изображения.
3. Кнопку для начала процесса обработки. При нажатии на неё должен запускаться процесс распознавания очагов на загруженном снимке.
4. Поле для вывода обработанного изображения.
5. Поле для вывода количества найденных очагов на изображении.
6. Кнопку для скачивания изображения. При нажатии на неё должно открываться диалоговое окно с выбором файла.

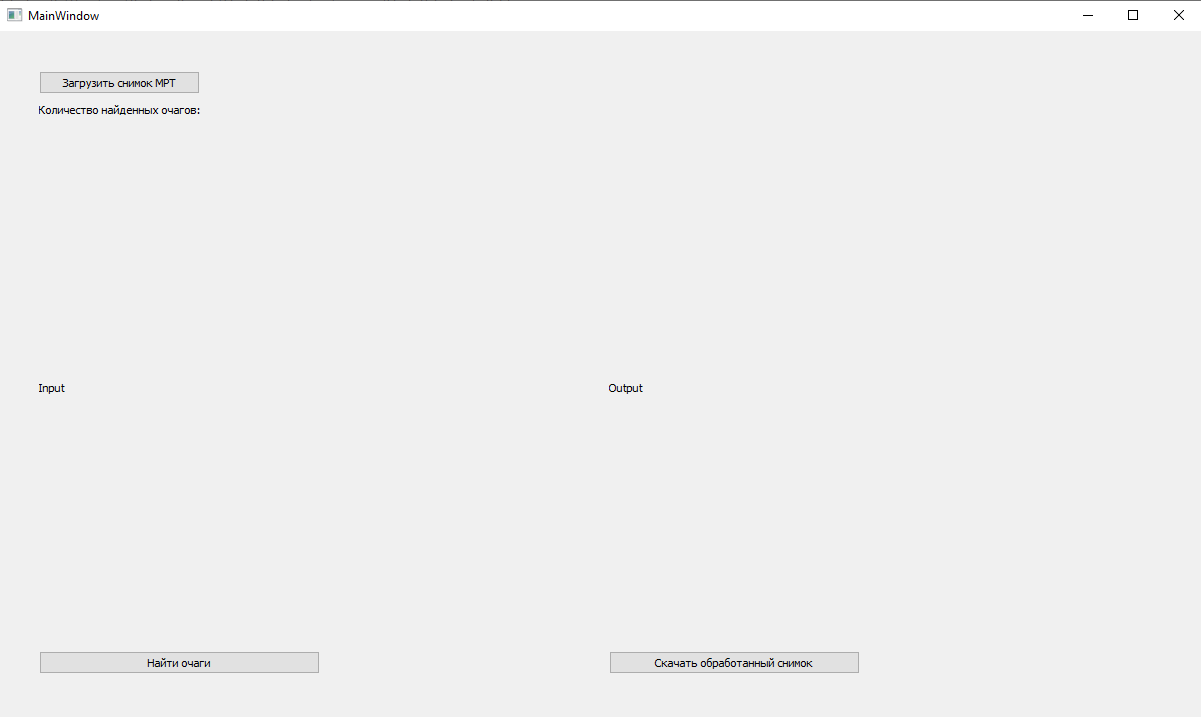


Рисунок Интерфейс приложения

# Результаты

## Проблемы, возникшие во время разработки

Во время тестирования выяснилась неэффективность использования метода гистограмм направленных градиентов для уточнения краёв очага. Это вызвано тем, что разрешение снимка невысоко, а программа к тому же работает только с её частью, то есть с еще меньшим разрешением. Так как гистограммы направленных градиентов основаны на вычислении значения в пикселях, результаты выходят нечёткие, пример на рисунке 13. Слева область снимка с очагом, справа выделение краёв гистограммами направленных градиентов.

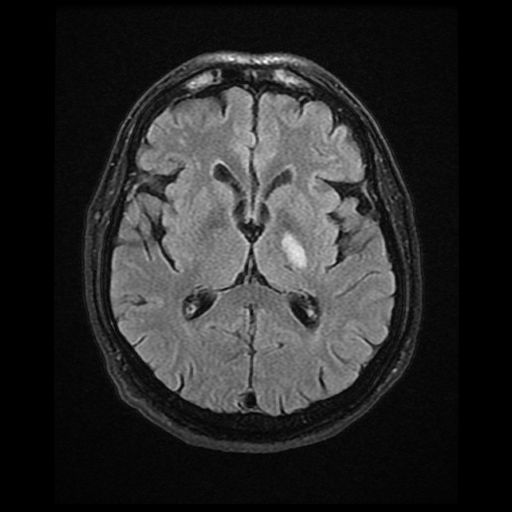
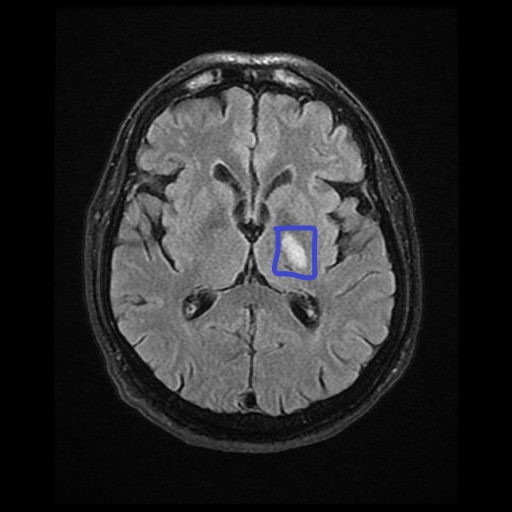
 

Рисунок 13 Слева исходный снимок, справа снимок с выделенной вручную областью очага.



Рисунок 14 Выделение краёв очага на области снимка.

### Метод выделения краёв очагов инсульта

Был разработан метод на основе интенсивности пикселей в области очага. Он заключается в двух шагах:

1. Высчитывание среднего значения интенсивности среди всех пикселей в выделенной области.
2. Закрашивание всех пикселей, интенсивность которых выше среднего значения.



Рисунок 15 Уточнение краёв очага инсульта на выделенной области

На том же снимке тот же очаг с более точными краями. Снимки с тестирования представлены в Приложении А.

## Сравнение результата с существующими решениями

1. Stroke Lesion Segmentation in FLAIR MRI Datasets Using Customized Markov Random Fields -<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6542951/> [↑](#footnote-ref-1)
2. Novel Screening Tool for Stroke Using Artificial Neural Network - http://www.ahajournals.org/doi/full/10.1161/STROKEAHA.117.017033?url\_ver=Z39.88-2003&rfr\_id=ori:rid:crossref.org&rfr\_dat=cr\_pub%3dpubmed [↑](#footnote-ref-2)
3. Stroke Lesion Detection Using Convolutional Neural Networks - https://www.researchgate.net/publication/323915650\_Stroke\_Lesion\_Detection\_Using\_Convolutional\_Neural\_Networks [↑](#footnote-ref-3)
4. Classification of stroke disease using convolutional neural network - https://iopscience.iop.org/article/10.1088/1742-6596/978/1/012092/pdf [↑](#footnote-ref-4)