

Селютин Александр Дмитриевич
студент Института прикладных информационных технологий и коммуникаций
Саратовский Государственный Технический Университет им. Гагарина Ю.А., г. Саратов
E-mail: cool.selutin99@yandex.ru

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ КЛАССИФИКАЦИИ МОЗГОВЫХ ОПУХОЛЕЙ НА СНИМКАХ МРТ

Selyutin Alexander Dmitrievich
student of Institute information technology and communications
Saratov State Technical University, Saratov

DEVELOPMENT AN INFORMATION SYSTEM FOR BRAIN TUMOR CLASSIFICATION ON MRI IMAGES

Аннотация: В данной статье рассмотрено решение проблемы ручной классификации снимков МРТ головного мозга на наличие опухолей специалистами технологами. Для решения данной проблемы разработана сверточная нейросетевая модель, а также, написан веб-интерфейс к разработанной модели, что позволяет быстро и удобно анализировать снимки в автоматическом режиме. Представлены ссылки на разработанную информационную систему, а также фрагменты кода веб-интерфейса.

Abstract: This article discusses the solution to the problem of manual classification of brain MRI images for the presence the tumors by specialist technologists. To solve this problem, a convolutional neural network model was developed, and a web interface for the developed model was written, which allows to quickly and conveniently analyze images in automatic mode. Links to the developed information system, as well a code fragments of the web interface are provided.

Ключевые слова: Сверточные нейронные сети, CNN, Flask, МРТ головного мозга, опухоль головного мозга

Keywords: Convolutional neural networks, CNN, Flask, brain MRI, brain tumor

Введение

Повсеместный прирост онкологических заболеваний вселяет серьезные опасения. За эти 12 лет этот прирост составил свыше 15% [4]. Помимо этого также возрастают показатели смертности. Различные опухоли занимают основные позиции среди заболеваний всех органов и систем. Согласно статистике, каждый день по всему миру 30 000 людей узнают о наличии у себя онкологического заболевания [6, с. 127]. В основном ситуация усугубляется поздней диагностикой опухолей, когда помощь больному уже практически бессмысленна.

Описание предметной области

Опухоль головного мозга – внутримозговое образование, которое проявляется, когда в головном мозге возникают нездоровые клетки. Классифицируют 2 основных вида опухолей: раковые, они же злокачественные, а также доброкачественные опухоли. Все виды раковых опухолей можно дифференцировать на первичные, которые возникают в мозге, а еще вторичные опухоли, распространяющиеся из других мест [3]. Они известны под названием метастазирующих опухолей. Опухоли головного мозга могут различаться симптоматикой, в зависимости от места их локализации. Симптомы могут сопровождаться головными болями, судорогами, психическими изменениями. Головная боль, как правило, усиливается по утрам и исчезает после рвоты. Другие симптомы могут проявляться трудностями при ходьбе, косноязычием. По мере прогрессирования заболевания может возникнуть потеря сознания.

Проблематика

Для проведения диагностики и своевременной помощи пациентам по выявлению и обнаружению опухолей существует магнитно-резонансная томография головного мозга. Часто снимки оказываются не лучшего качества и

определить на них опухоль технологу рентгенологу бывает затруднительно. К тому же, ручная работа утомительна и дальнейшая сегментация снимка может оказаться ненужной в случае отсутствия опухоли.

Постановка задачи

Для автоматизации работы технолога рентгенолога необходимо разработать систему, которая будет способна классифицировать поступающие на вход снимки МРТ. Система должна давать ответ на вопрос о наличии опухоли у пациента. Также информационная система должна обладать удобным и понятным интерфейсом.

Цель работы

Целью проводимой работы является разработка сверточной нейросетевой модели для классификации снимков МРТ головного мозга на наличие опухолей. Данный вид нейросетевой архитектуры лучше всего подходит для анализа, классификации, сегментации изображений и других задач, связанных с обработкой изображений. Будет необходимо обработать имеющиеся данные (набор снимков МРТ головного мозга). Также нужно провести препроцессинг снимков, чтобы разрабатываемая модель обладала достаточной точностью и верно классифицировала поступающие изображения. Помимо этого необходимо, обученной нейросетевой модели необходимо придать понятный пользовательский интерфейс. В данной статье разработка модели будет производиться с помощью языка высокого уровня Python версии 3.6.3. Будут использованы такие библиотеки для машинного обучения как TensorFlow, Keras и другие. Веб-интерфейс будет разработан на фреймворке для создания минималистичных каркасов веб-приложений Flask.

Разработка нейросетевой модели

В наше время существует огромное количество нейросетевых архитектур, выполняющих максимально разнообразные задачи. Существуют различные

методы обучения нейросетей, например, с учителем и без. Для данной задачи, связанной с распознаванием изображений оптимально будет использовать архитектуру сверточной нейронной сети. При разработке будет использоваться фреймворк Keras, содержащий в своей основе другой фреймворк для машинного обучения от компании Google под названием TensorFlow [7, с. 81]. Фреймворк Keras позволит абстрагироваться от низкоуровневых аспектов сверточных нейронных сетей таких как программирование работы ядра свертки и пр. Благодаря высокому уровню абстракции разработка модели будет протекать быстро и удобно [2].

Первичная обработка данных

Снимки МРТ головного мозга анонимных пациентов, которые дали согласие на обработку своих персональных данных, были предоставлены местной поликлиникой. Томограммы представляют собой изображения в формате .jpg и имеют различное разрешение. В имеющемся наборе содержатся 155 изображений, классифицированных как снимки на которых имеется опухоль головного мозга, а также 74 томограммы головного мозга без опухолей.

Для эффективного обучения модели необходимо разделить имеющийся набор данных на 3 выборки по которым будет проводиться дальнейшее обучение и проверка. Был написан скрипт, который размещает в случайном порядке снимки по 3 директориям: директория для обучения модели, директория для валидации модели, а также директория проверки модели на явление переобучения. Каждая директория содержит 2 поддиректории, в которых находятся томограммы на которых нет опухолей головного мозга, и снимки на которых имеется онкологическое заболевание. После распределения снимков по соответствующим директориям в директории для обучения оказалось 74 снимка без опухолей и 119 с опухолью головного мозга, в

валидационной директории находится 31 снимок с онкологией и 19 без, в директории для проверки 5 классифицированных снимков поровну.

Преобработка снимков МРТ

После распределения томограмм по соответствующим директориям необходимо выполнить обработку самих снимков, чтобы корректно обучить модель. Первоначально необходимо произвести поиск областей интересов и обрезать снимок так, чтобы на нём содержалась лишь необходимая информация. Для этого используем фильтр Гаусса, дабы лишние и мелкие детали на изображении стали незаметны. Далее необходимо выполнить преобразование цветного изображения в двоичное [5, с. 233]. После того, как изображение будет преобразовано в бинарное нужно провести 2 итерации эрозии и такое же количество циклов для операции дилатации. Изображение будет готово для поиска на нём максимального контура, который будет представлять собой овальную границу головного мозга. В полученной области требуется выделить точки экстремума для правильной обрезки изображения. Выделенный снимок готовый к конечной обрезке представлен на Рисунке 1.

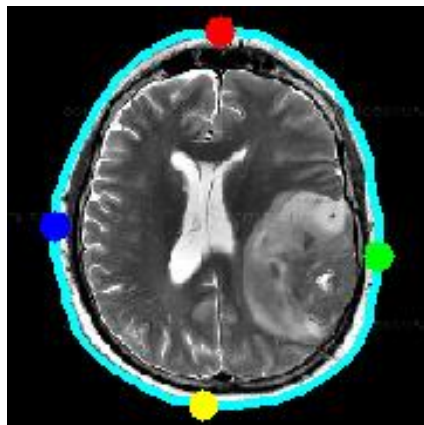


Рисунок 1. Снимок МРТ головного мозга, подготовленный для обрезки

Теперь необходимо изменить разрешение всех томограмм. Все снимки будут квадратными со стороной 224 пикселя. После требуется сохранить обработанные изображения по соответствующим директориям и начать разработку нейросетевой модели на подготовленных данных.

Обучение модели

После того как все снимки подверглись обработке и были распределены по директориям, можно начинать создавать расширенные классы для томограмм. Расширенный класс снимка представляет собой проекции исходного изображения под разными углами поворота. Для снимков, находящихся в директории с томограммами, предназначенными для дальнейшего обучения модели, необходимо создать расширенные классы.

После получения расширенных классов снимков, требуется загрузить из открытого репозитория предобученную модель сверточной сети для выделения признаков VGG16. Веса в этой модели изначально подобраны оптимальным образом, поэтому остаётся лишь корректировать их под конкретную задачу, чтобы получить максимальную точность. После загрузки модели необходимо добавить слои и активационную функцию. Благодаря фреймворку Keras сделать это очень просто. Для обучения рекомендуется назначить 30 итераций, добавить возможность раннего прерывания, если дальнейшее обучение не будет иметь смысла, а также в каждой эпохе будет 50 прогонов. Обучение модели – процесс длительный. На рабочей станции с процессором Intel Core i3 3.7GHz, 16 GB RAM, видеоадаптером Intel HD Graphics 530 весь процесс обучения занял 2 часа. После этого модель была сохранена в формате h5 для дальнейшего использования [1]. На обучающем множестве была достигнута точность 92%. На тестовом множестве точность составила 80%. Весьма неплохой результат! Теперь необходимо создать удобный интерфейс пользователя, чтобы модель было удобно эксплуатировать.

Разработка веб-интерфейса для нейросетевой модели

Для сохраненной модели возможно разработать различный графический интерфейс пользователя. Был сделан выбор в пользу веб-интерфейса в связи с гетерогенностью платформ и преимуществом в виде единообразия веб-платформы.

Интерфейс должен обладать следующим минимальным функционалом: возможность выбрать изображение, отменить выбор, отправить снимок на классификацию.

Для разработки был выбран фреймворк Flask. Данная технология позволит быстро и качественно развернуть приложение на рабочий сервер, написав немного кода, построить точную маршрутизацию запросов, а также с помощью шаблонизатора Jinja 2, сделать интерфейс визуально приятным для пользователя.

На Листинге 1 представлен программный код, который отвечает за загрузку изображения в формате .jpg на сервер:

```
@app.route('/upload', methods=['GET', 'POST'])
def upload_file():
    if request.method == 'POST':
        file = request.files['file']

        if file and allowed_file(file.filename):
            filename = secure_filename(file.filename)
            file.save(os.path.join(app.config['UPLOAD_FOLDER'], filename))
            return redirect(url_for('uploaded_file', filename=filename))

    return redirect(url_for('index'))
```

Листинг 1. Функция, отвечающая за загрузку снимков МРТ на сервер

Также необходимо разработать функцию, которая будет отлавливать переадресацию с данного метода на адрес, который указан в операторе возврата. Потребуется выделить отдельный поток для данной функции, так как действие по обработке и классификации томограммы должно обрабатываться параллельно с HTTP-запросом. На Листинге 2 приведён пример такой функции:

```
@app.route('/uploads/<filename>')
def uploaded_file(filename):
    global graph
    with graph.as_default():
        image_path = UPLOAD_FOLDER + filename

        img = image.load_img(image_path, target_size=IMG_SIZE)
        x = image.img_to_array(img)
        x = np.expand_dims(x, axis=0)

        images = np.vstack([x])
        classes = model.predict_classes(images, batch_size=10)
        print('Predicted class is: ' + str(classes[0][0]))

    return render_template('result.html', filename=filename, result=classes[0][0])
```

Листинг 2. Функция, отвечающая за классификацию снимков МРТ

После тестирования системы, весь исходный код был загружен на облачный хостинг Heroku. На данный момент система доступна по адресу: <https://brain-tumor-classifier.herokuapp.com/>

На Рисунке 2 представлен интерфейс готовой информационной системы, опубликованной на хостинге Heroku:

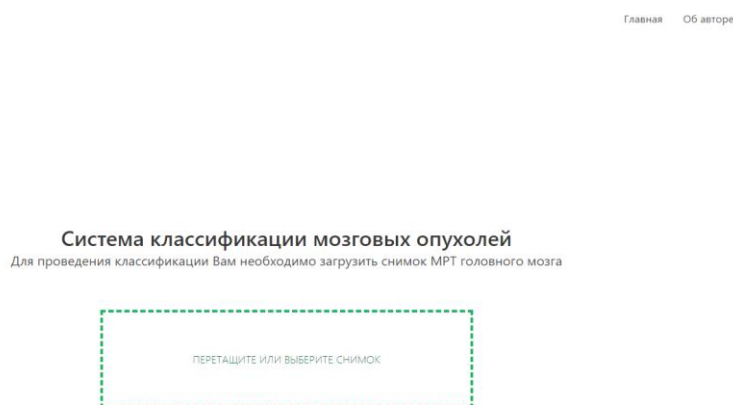


Рисунок 2. Веб-интерфейс информационной системы по классификации мозговых опухолей

Заключение

В ходе работы была смоделирована сверточная нейросеть, выполняющая функцию классификации снимков МРТ головного мозга на наличие опухолей. К разработанной модели был создан удобный веб-интерфейс, который упрощает работу с системой, делая взаимодействие с системой приятным и удобным.

Разработанная система может применяться технологами рентгенологами для повышения производительности труда. Процесс классификации снимков МРТ головного мозга автоматизирован. Исполнителю стоит заняться лишь дальнейшей сегментацией в случае наличия опухоли у пациента.

Список литературы:

1. Brain Tumor Detection [Электронный ресурс] - URL: <https://www.kaggle.com/ruslankl/brain-tumor-detection> (дата обращения: 12.08.2019)
2. Глубокое обучение для новичков: распознаем изображения с помощью сверточных сетей [Электронный ресурс] - URL: <https://habr.com/ru/company/wunderfund/blog/314872/> (дата обращения: 11.08.2019)
3. Опухоль головного мозга [Электронный ресурс] - URL: <http://navigator.mosgorzdrav.ru/diseases/onkologiya/opukhol-golovnogo-mozga/> (Дата обращения: 11.08.2019).
4. Признаки опухоли головного мозга [Электронный ресурс] - URL: <https://doctor-neurologist.ru/priznaki-opuxoli-golovnogo-mozga> (Дата обращения: 11.08.2019).
5. Редько В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики / В.Г. Редько. - Москва: РГГУ, 2017. – 233 с.
6. Справочник по онкологии. - М.: Здоров'я, 2009. – 127 с.
7. Тархов, В.Г. Нейросетевые модели и алгоритмы: Справочник / Д.А. Тархов. – М.: Радиотехника, 2014. – 81 с.