

Рисунок 2 — Результат преобразования русской речи в текст

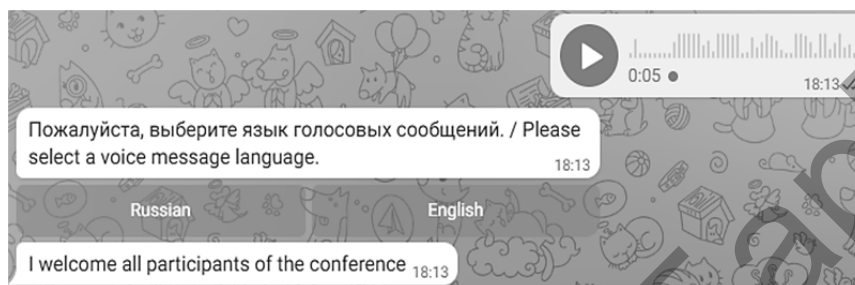


Рисунок 3 — Результат преобразования английской речи в текст

Заключение. В ходе работы был разработан чат-бот для преобразования речевого сигнала в текст. После завершения этапа тестирования, результаты, которые были получены, подтверждают работоспособность созданной системы и ее полное соответствия всем предъявляемым требованиям. В дальнейшем планируется добавить функцию эмоционального окраса сообщения.

Список цитируемых источников

1. Telegram FAQ [Электронный ресурс]. — Режим доступа : <https://telegram.org/faq#q-what-is-telegram-what-do-i-do-here> . — Дата доступа : 26.04.2022.
2. Шах, А. В. Чат-боты как современный инструмент маркетинга / А. В. Шах, Е. Г. Шапович // Стратегия и тактика развития производственно-хозяйственных систем : сб. науч. тр. / М-во образования Респ. Беларусь, Гомел. гос. техн. ун-т им. П. О. Сухого, Гомел. обл. орг. о-ва «Знание»; под ред. В. В. Кириенко. — Гомель, 2019. — С. 200—203.
3. History and License — Python 3.10.4 documentation [Электронный ресурс]. — Режим доступа : <https://docs.python.org/3/license.html>. — Дата доступа : 26.04.2022.

УДК 004.8

Е. Г. Шапович¹, М. Д. Евтух²

¹Учреждение образования «Барановичский государственный университет», Барановичи, Республика Беларусь

²Государственное учреждение образования «Средняя школа № 9 имени Героя Советского Союза Г. Н. Холостякова», Барановичи, Республика Беларусь

АННОТИРОВАНИЕ ИЗОБРАЖЕНИЙ

Введение. Цель данной работы состоит в разработке Android-приложения для описания изображений на естественном языке с использованием нейросетевых технологий.

Актуальность данного исследования состоит в том, что в настоящее время всё большую популярность завоёвывают программные продукты, которые используют нейронные сети. Также разрабатываемое приложение может стать своеобразным помощником для людей с ограниченными возможностями (слепых и слабовидящих), которые с помощью данного приложения могут узнать, что изображено на фото.

Объектом данного исследования являются подходы, методы представления изображений на естественном языке. Предметом исследования являются информационные модели искусственных нейронных сетей, а также описание алгоритма представления изображений на естественном языке.

Задачи исследования:

- 1) изучить литературу по нейронным сетям;
- 2) ознакомиться со сверточными нейронными сетями;
- 3) разработать Android-приложение в среде Android Studio на языке Java.

Научная новизна исследования заключается в использовании технологий нейросетей (информационных моделей) для реализации алгоритма представления изображений на естественном языке. Практическая значимость исследования: результаты работы могут быть использованы для помощи людям с ограниченными возможностями, а также быть своеобразным примером использования нейросетевых технологий.

Основная часть. Под искусственной нейронной сетью понимается математическая модель, а также ее программная и аппаратная реализация, построенная по принципу биологических нейронных сетей — нервных клеток живого организма. Это понятие возникло при попытке смоделировать процессы, протекающие в мозге человека [1, с. 3].

Искусственная нейронная сеть представляет собой систему простых процессоров (искусственных нейронов), соединенных и взаимодействующих между собой. Каждый из процессоров сети имеет дело с сигналами, которые периодически поступают или передаются другим процессорам. Большая сеть способна решать сложнейшие задачи в самые короткие сроки.

Если классифицирование изображений и обнаружение объектов уже сами по себе являются интеллектуальными задачами, то описание изображения на естественном языке, безусловно, является намного более сложной задачей, требующей гораздо большего интеллекта, — просто подумайте на мгновение о том, как по мере взросления с самых первых дней своей жизни мы сначала учимся распознавать объекты и определять их местоположение и только позже, в трехлетнем возрасте, учимся рассказывать о том, что видим на картинке.

Официальный термин для описания изображения на естественном языке — аннотирование (или титрование) изображения. В отличие от распознавания речи, которое имеет долгую историю исследований и разработок, аннотирование изображений (на полном естественном языке, а не только с участием ключевых слов) имело лишь короткую, но захватывающую историю исследований из-за своей сложности и благодаря инновационному прорыву в технологии глубокого обучения в 2012 году.

В этом исследовании мы будем использовать модель аннотирования изображений, основанную на глубоком обучении, которая одержала победу в конкурсе Microsoft COCO 2015 Image Captioning по аннотированию изображений (где используется крупномасштабный набор данных для обнаружения). Поскольку данная модель включает в себя компьютерное зрение и обработку естественного языка, мы продемонстрируем совместную работу сетей CNN и RNN, двух основных глубоких нейросетевых архитектур.

Для обучения нейронной сети использовалась библиотека TensorFlow от Google, распространяемая по лицензии Apache 2.0. Поддерживает вычисления на CPU, GPU и специально разработанными компанией Google TPU (Tensor processing units). Фронтенд на Python. Одна из самых популярных, хорошо документированных и развитых библиотек [2].

Основной причиной выбора данной библиотеки состоит в том, что она обладает простотой запуска оболочки на мобильных устройствах, оптимизирована под GPU и достаточно популярна. Задачей исследования является приложение для аннотирования изображений на базе платформы Android.

Для представления каждого слова в выводе на целевом естественном языке используется метод встраивания слов. Встраивание слов (word embedding) — это попросту векторные представления слов. На веб-сайте TensorFlow имеется хорошее руководство по данной теме (<https://www.tensorflow.org/tutorials/word2vec>) с описанием построения модели, получающей векторные представления слов. Для решения поставленной задачи было решено использовать интегрированную среду разработки Android Studio и объектно-ориентированный язык высокого уровня Java. Разработанное приложение имеет вид, представленный на рисунке 1.

Для того, чтобы получить описание исходного изображения необходимо нажать на кнопку «Описать изображение». После этого результат на английском языке будет представлен во всплывающем сообщении, а сверху будет представлен результат на русском языке (рисунок 2). В связи с тем, что используется бесплатный машинный перевод (старая версия API Google Translate) перевод не самый лучший. Однако это не мешает понять смысл того, что изображено на фото.

Если нажать на кнопку с изображением микрофона, то полученное описание будет воспроизведено с помощью библиотеки TextToSpeech. Также на рисунке 2 представлен результат описания на естественном языке другого изображения.



Рисунок 1 — Главное окно приложения



Рисунок 2 — Результат аннотирования

Заключение. Реализация интеллектуальной системы, способной описать изображение на естественном языке — это решение технически и математически сложной задачи, требующей углубленного изучения предметной области, собственного анализа, умения и навыков при обработке и выборе экспериментальных данных, глубокие знания в области дискретной математики, программирования, а также биологии и многих других областях науки.

Была выбрана информационная модель на основе сверточной нейронной сети. Данная модель оптимальна по внутренней структуре и способу управления информационными потоками между нейронами. Кроме того, такая модель способна минимизировать число входных элементов. Модель на основе сверточной нейронной сети является универсальной моделью и подходит для решения задач разного уровня сложности, в том числе и для решения задачи аннотирования изображений. Для разработки приложения использовались материалы книги «Умные мобильные проекты с TensorFlow» [3], а также библиотеки TextToSpeech и Google Translate.

На основании полученных результатов можно сделать вывод о том, что выбранная информационная модель искусственной нейронной сети соответствует модели реальной системы и учитывает все ее характеристики. Все поставленные задачи были выполнены в полном объеме, следовательно, главная цель была достигнута. Разработанное приложение описывает изображения на естественном языке. Оно позволит слабовидящим людям узнать, что представлено на изображении. В дальнейшем планируется внедрить в приложение работу с камерой мобильного телефона.

Список цитируемых источников

1. Головкин, В. А. Нейронные сети : обучения, организация и применение / В. А. Головкин. — М. : ИПРЖР, 2008. — 256 с.
2. TensorFlow [Электронный ресурс]. — Режим доступа : <https://www.tensorflow.org/>. — Дата доступа : 26.04.2022.
3. Танг, Дж. Умные мобильные проекты с TensorFlow / Дж. Танг. — М. : ДМК Пресс, 2019. — 384 с.

УДК 004.9

С. А. Ярохович, О. Д. Кравчук

Учреждение образования «Барановичский государственный университет», Барановичи, Республика Беларусь

РАЗРАБОТКА ИГРЫ «СТРЕЛОК»

Введение. Компьютерные игры прочно вошли в нашу жизнь, заняв почетное место лидера среди множества способов организации молодежного отдыха. Виртуальная реальность манит своими безграничными возможностями, от которых просто невозможно отказаться. Дети с самых малых лет не представляют свою жизнь без компьютера, планшета и телефона, которые, как правило, используют для видеоигр.

Некоторые, кто уже играл в игры, возможно, хотели бы придумать какие-нибудь свои, не похожие ни на какие другие игры. Много захватывает в таком творчестве. И не сам процесс игры, а разработка игровой вселенной, ее проектирование и реализация. Когда можно слить воедино сценарий, графику, музыку, искусно задуманный и умело запрограммированный алгоритм.