

Магистрант ИКИТ создал алгоритм для распознавания речи по видео

Магистрант Института космических и информационных технологий Сибирского федерального университета Антон Дзюба научил нейронные сети распознавать речь по артикуляции губ на видео. Экспериментальные исследования проведены с использованием самостоятельно полученных видеороликов с русскоговорящими спикерами.



Распознавание речи выполняется в два этапа. Сначала осуществляется поиск лица и выделение области губ в отдельном кадре видеопоследовательности с использованием признаков Хаара. Далее последовательность кадров поступает на вход свёрточной и рекуррентной нейронных сетей глубокого обучения для распознавания виземы речи.

При проведении экспериментальных исследований использован набор данных, содержащий 768 различных высказываний, произнесённых различными спикерами. Видеоролики получены самостоятельно, так как требовался набор данных русскоязычных спикеров. Высказывания помечены теми же метками, что и набор данных для обучения. Тестовыми словами были: «бегу», «пила», «милый», «усы», «вулкан», «банан», «тонуть». Лучшая точность распознавания речи по артикуляции составила 93,7 % для слова «банан», а средняя точность — 68 %.

«Визуальное распознавание речи является важнейшей задачей при общении людей с нарушениями слуха. Распознавание речи по артикуляции применяется также в областях, не связанных с медициной, например, в правоохранительной деятельности. Виземы и фонемы не имеют однозначного соответствия. В русском языке насчитывается 42 фонемы. Из них 6 гласных и 36 согласных. Часто несколько фонем соответствуют одной виземе и выглядят одинаково на лице говорящего человека. В дальнейшем мы планируем усовершенствовать алгоритм, повысить точность его работы и расширить количество распознаваемых слов», — рассказал **Антон Дзюба**.



Добавим, научным руководителем магистранта является доцент кафедры систем искусственного интеллекта ИКИТ СФУ Анна Пятаева. Результаты работы были представлены на VIII Международной научной конференции «Региональные проблемы дистанционного зондирования Земли».

[Пресс-служба СФУ](#), 4 октября 2021 г.

© Сибирский федеральный университет. Редакция сайта: +7 (391) 246-98-60, info@sfu-kras.ru.

Адрес страницы: <https://news.sfu-kras.ru/node/25317>