

УДК 004.89

## ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА ПРИ РАЗРАБОТКЕ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ РУССКОГО ЖЕСТОВОГО ЯЗЫКА

Граник А. В. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Кугаевских А. В. (ИТМО)

**Введение.** Русский жестовый язык – важный способ коммуникации глухих и слабослышащих. Разработка нейронной сети для распознавания русского жестового языка облегчит процесс общения для его носителей, которых насчитывается более 200 000. Основной задачей такой нейронной сети будет являться выделение и классификация отдельных жестов. Для повышения качества распознавания речи в дополнение к классификатору возможно использовать инструменты обработки естественного языка.

**Основная часть.** Распознавание жестовых языков – комплексная задача, эффективное решение которой без использования инструментов машинного обучения невозможно. Достижения в области глубокого обучения послужили причиной активного развития данной области и привели к появлению большого числа распознавателей различных жестовых языков [1]. Количество предложенных моделей для распознавания русского жестового языка относительно невелико. Большинство современных решений основываются на применении сверточных нейронных сетей и концептуально предполагают отнесение каждого отдельного жеста к тому или иному классу, которому может быть сопоставлено слово или иная языковая единица.

Одной из проблем разработки качественного распознавателя жестового языка является наличие в них визуально схожих жестов. Корректное отнесение их к нужному классу на основе исключительно визуальной информации затруднительно. Для решения этой проблемы предлагается использование технологий обработки естественного языка [2]. Задача предсказания следующего жеста на основе предыдущих фактически идентична задаче предсказания слова на основе контекста.

Предлагаемая модель распознавателя русского жестового должна позволить эффективно обрабатывать поступающий видеоряд и переводить жестовую речь в текст. С точки зрения архитектуры в нейронной сети могут быть выделены 2 основных модуля:

1. классификатор жестов, выделяющий в видео отдельные жесты и определяющий вероятности их отнесения к различным классам;
2. анализатор текста, окончательно относящий жест к конкретному классу на основании полученных от классификатора вероятностей и контекста (уже переведенных в текст жестов).

**Выводы.** Проведен анализ существующих решений проблемы распознавания жестовых языков, на основании которого предложена комбинированная архитектура. Проведено исследование эффективности предложенной архитектуры.

### Список использованных источников:

1. Razieh Rastgoo, Kouros Kiani, Sergio Escalera Sign Language Recognition: A Deep Survey // Expert Systems with Applications. 2021. Volume 164
2. R. Zuo, F. Wei, B. Mak Natural Language-Assisted Sign Language Recognition // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada, 2023. P. 14890-14900.