

Красноярск 2024

АННОТАЦИЯ

Целью магистерской диссертации является исследование и реализация методов распознавания жестов руки на изображениях. В основе решения задачи распознавания лежат методы, основанные на нейронных сетях.

Поставлена и решена практическая задача распознавания жестов на изображении. Предложены модификации нейронных сетей для детекции и классификации жестов. Создано программное обеспечение, реализующее работу данных алгоритмов. Решена практическая задача распознавания жестов на изображении. Проведены вычислительные эксперименты по оценке работы различных методов детекции и классификации жестов по точности и качеству распознавания. Выполнен анализ полученных результатов.

Ключевые слова: распознавание жестов, нейронная сеть, классификатор, изображение, машинное обучение.

ANNOTATION

The purpose of the master's thesis is to study and implement methods of recognizing hand gestures on images. The solution to this problem is based on neural network methods.

The practical task of recognition of gestures on the image has been set and solved. Modifications of neural networks for the detection and classification of gestures are proposed. The software for the implementation of these algorithms has been created. The practical problem of recognition of gestures on the image has been solved. Computational experiments were conducted to evaluate the work of various methods of detection and classification of gestures on the basis of accuracy and quality of recognition. Analysis of the results is obtained.

Keywords: gesture recognition, neural network, classifier, image, machine learning.

ВВЕДЕНИЕ

Разработка систем компьютерного зрения является одним из наиболее востребованных и актуальных направлений в области анализа данных, искусственного интеллекта и машинного обучения. Она направлена на распознавание и классификацию объектов на изображениях различной природы. Одним из возможных направлений применения таких систем является распознавание жестов руки на изображении.

Целью магистерской диссертации является исследование и реализация методов распознавания и классификации жестов руки на изображении. Решение данной задачи состоит из этапов сегментации, выделения признаков и классификации. Этап сегментации заключается в выделении части изображения, содержащей кисть руки человека. На этапе выделения признаков из полученного ранее сегмента изображения извлекаются признаки, характеризующие жест. Наконец, на этапе классификации происходит определение жеста на основе выделенных признаков.

Для достижения цели работы были поставлены и решены следующие задачи:

- Изучить основные алгоритмы распознавания объектов на изображении и их классификации;
- Провести вычислительные эксперименты по оценке эффективности основных архитектур нейронных сетей распознавания и выбрать архитектуру нейронной сети для решения задачи классификации жестов;
- Разработать нейронные сети выбранной архитектуры для детекции руки и распознавания жестов;
- Провести модификацию методов предобработки для методов детекции и классификации;

- Решить практическую задачу распознавания жестов на изображении;
- Провести вычислительные эксперименты по сравнению алгоритмов распознавания и классификации по их быстродействию, скорости обучения и качеству распознавания.

Для решения задачи распознавания и классификации будут использоваться нейросетевые методы. В качестве набора данных для обучения и проверки качества распознавания будут использоваться цветные изображения человека, который показывает рукой жест. Задача заключается в следующем: алгоритму даётся изображение, на котором он должен выделить область, содержащую руку, и затем классифицировать жест руки по заранее заданным классам.

Теоретической значимостью исследования является исследование и сравнение методов распознавания жестов руки на изображении. Это исследование позволяет увеличить доступность данной технологии за счёт сравнения эффективности алгоритмов при использовании изображения только с одной цветной камеры, без использования дополнительных камер или специализированных датчиков, таких как датчики глубины или акселерометры.

Практическая значимость исследования заключается в реализации наиболее эффективных методов распознавания в виде программного обеспечения. Одним из востребованных применений данной технологии является естественное взаимодействие человека и компьютера. Например, распознавание жестов может использоваться для управления мультимедийной системой. Пользователь может выполнять различные команды управления с помощью определенных жестов руки: изменять громкость звука, останавливать воспроизведение видео, переключать на другое видео и т.д. Помимо взаимодействия с мультимедиа, распознавание жестов является одним из востребованных направлений

в области виртуальной реальности. Например, вместо нажатия кнопки для поднятия предмета в виртуальном мире, пользователь может просто поднять руку, и система распознает его жест и выполнит соответствующее действие. Это позволяет создать естественный и интуитивный опыт использования, так как пользователи могут взаимодействовать с виртуальным окружением так же, как они делают это в реальной жизни. Также распознавание жестов может использоваться для интерпретации жестов языка глухонемых.

В первой главе диссертации дается постановка задачи и приводятся основные понятия и постановка задачи распознавания образов и исследуются этапы её решения. Во второй главе рассматриваются используемые методы предобработки.

В третьей главе описываются методы, основанные признаках Хаара, используемые для распознавания и классификации жестов руки.

В четвертой и пятой главе описываются нейросетевые методы, используемые для распознавания и классификации жеста. Проводится вычислительный эксперимент по выбору архитектуры нейросети для классификации жестов. В работе проведены вычислительные эксперименты по оценке работы различных методов детекции руки и классификации жестов по точности и качеству распознавания.

В шестой главе описаны проведенные эксперименты, а также – набор данных, на котором выполнялось сравнение алгоритмов. Создано программное обеспечение, реализующее работу данных алгоритмов, и решена практическая задача, для которой были, и создан программный модуль, реализующий работу всех методов. Также выполняется сравнительный анализ получившихся результатов. Состав разработанного программного средства и примеры его работы приводятся в седьмой главе магистерской диссертации.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Местецкий, Л. М. Математические методы распознавания образов / Л. М. Местецкий. – Москва: МГУ, 2004. – 85 с.
2. Lookup tables // Image Analysis Training Resources – [Электронный ресурс] URL: neubias.github.io/training-resources/lut/index.html (дата обращения: 23.05.2024).
3. Штарк, Г. Г. Применение вейвлетов для ЦОС / Г. Г. Штарк. –Москва: Техносфера, 2007. – 192 с.
4. Гафаров, Ф.М. Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. университета, 2018. – 121 с.
5. Subject independent facial expression recognition with robust face detection using a convolutional neural network // ScienceDirect URL: <https://www.sciencedirect.com/science/article/pii/S0893608003001151> (дата обращения: 25.05.2024).
6. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Application // arXiv [Электронный ресурс] – URL: <https://arxiv.org/abs/1704.04861> (дата обращения: 25.05.2024).
7. Giuseppe, P. Aggregated Residual Transformations for Deep Neural Networks // IEEE Transactions on image processing. 2007. – PP. 2449–2461.
8. Kaiming, H. Deep Residual Learning for Image Recognition // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – PP. 770-778.
9. Saining, X. Aggregated Residual Transformations for Deep Neural Networks // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – PP. 5987-5995.
10. Andrew, H. Aggregated Residual Transformations for Deep Neural Networks // IEEE/CVF International Conference on Computer Vision (ICCV). –

2019. – PP. 1314-1324.

11. Sign Language MNIST: сайт. – [Электронный ресурс]

URL: kaggle.com/datasets/datamunge/signlanguage-mnist

12. Kapitanov, A. HaGRID - HAnd Gesture Recognition Image Dataset –

[Электронный ресурс] URL: [/github.com/hukenovs/hagrid](https://github.com/hukenovs/hagrid)

13. Официальная документация библиотеки PyTorch [Электронный ресурс]

– URL: <https://pytorch.org/docs/stable/index.html>

Министерство науки и высшего образования РФ
Федеральное государственное автономное
образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт математики и фундаментальной информатики
Базовая кафедра вычислительных и информационных технологий

УТВЕРЖДАЮ

Заведующий кафедрой

 / В.В. Шайдуров

«20» 06 2024 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

**ИССЛЕДОВАНИЕ МЕТОДОВ РАСПОЗНАВАНИЯ ЖЕСТОВ
РУКИ НА ИЗОБРАЖЕНИЯХ**

Направление 02.04.01 Математика и компьютерные науки

Магистерская программа 02.04.01.02 Вычислительная математика

Руководитель


20.06.24

доцент, кандидат физико-
математических наук

И.В. Баранова

Выпускник


20.06.24

А.А. Судакова

Нормоконтролер


21.06.24

Т.Н. Шипина

Красноярск 2024