

АННОТАЦИЯ
диссертационной работы Муханова С.Б. «Разработка и применение
высокоточных методов распознавания образов», представленной на
соискание степени доктора философии (PhD)
по специальности 6D070400 – Вычислительная техника
и программное обеспечение.

В настоящее время всё больше исследований направлено на решение задач с применением библиотек компьютерного зрения и инструментов искусственного интеллекта. Наиболее частыми являются решения и подходы с использованием моделей машинного и глубокого обучения искусственных нейронных сетей для распознавания жестов казахского жестового алфавита на основе методов обучения с учителем и глубокого обучения для обработки последовательных данных. Объектом исследования является казахский жестовый алфавит для построения коммуникации между людьми с ограниченными возможностями. Предметом исследования – методы машинного обучения и модели искусственных нейронных сетей и глубокого обучения для классификации и распознавания жестов. Методы исследования – Data Science, Machine Learning, Deep Learning, Neural networks и Computer Vision.

Распознавание образов — это изображение, на котором расположен объект. Так как объект является абстрактным (объектом может быть любая форма, изображенная на картине). Мы решили исследовать одно из актуальных направлений – распознавание жестов. Для распознавания казахского жестового языка, первым делом нужно изучить казахский жестовый алфавит. Для того, чтобы обучить нейронную сеть распознавать казахский язык жестов, необходимо собрать данные (датасеты) в формате изображений обозначенные жестами рук. Распознавание жестов – это задача классификации, которая является одним из направлений видов распознавания образов. Фундаментальной основой распознавания – является теория распознавания образов.

Целью исследования. Разработать и применить высокоточные методы распознавания образов. Для этого необходимо использовать инструменты искусственного интеллекта, а именно алгоритмы и библиотеки (язык программирование Python) для работы с машинным обучением и нейронными сетями глубокого обучения. Собрать данные для обучения и предварительно обработать для обучения нейронных сетей. Важно произвести математические расчеты для вычисления градиента функций и алгоритма обучения обратного распространения ошибки (backpropagation) и функций активации для каждого нейрона. Применить эмпирический подход к разработке собственной архитектуры моделей нейронных сетей для обучающих параметров и гиперпараметров. Архитектура данной модели может варьироваться как глубиной, так шириной нейронной сети, таким образом мы произвольно определяем количество обучающихся параметров. В диссертационной работе основное внимание уделяется изучению методов обучения с учителем, метода глубокого обучения, задаче классификации и распознавания жестов, обученных на собственных данных (изображения, полученные и разделенные на кадры из

видеопоследовательности, снятые веб-камерой или мобильным устройством). Эти методы позволяют значительно расширить спектр задач, которые можно эффективно решать в реальном времени в области распознавания жестов. Разработка программного обеспечения позволит протестировать эффективность работы обученной модели и применять ее в лабораторных целях, вносить корректировки для улучшения данной модели. Достижение поставленной цели приведет к повышению эффективности и расширению возможностей современных систем компьютерного зрения и распознавания жестов.

Для достижения поставленной цели необходимо было решить следующие **задачи:**

1. Собрать данных (изображения букв, представленные в виде жестов рук казахского жестового алфавита) и предварительная обработка в целях подготовки обучающего набора для методов машинного и глубокого обучения.

2. Анализ архитектур глубокого обучения и применение методов для распознавания жестов.

3. Обучить нейронную сеть распознавать казахский жестовый алфавит на собственных данных (изображениях).

4. Сделать сравнительный сверточной и рекуррентной нейронной сети глубокого обучения.

5. Разработать программное обеспечение для распознавания жестов (казахского жестового алфавита) в реальном времени.

Объектом исследования является обучение искусственной нейронной сети для распознавания казахского жестового алфавита с помощью инструментов искусственного интеллекта, а именно методов машинного обучения.

Предметом исследования является программное обеспечение системы распознавания жестов, разработанное на базе (основе) обученной модели глубоких нейронных сетей.

Научная новизна исследования определяется:

1. Собраны данные в формате изображений жестов рук и для обучения моделей искусственных нейронных сетей.

2. Обучены модели глубокого обучения для распознавания казахского жестового (дактильного) алфавита.

3. Построена собственная архитектура гибридной модели на основе метода обучения с учителем сверточной нейронной сети и метода глубокого обучения рекуррентной нейронной сети.

4. Разработано программное обеспечение для распознавания казахского жестового алфавита.

Научная положения, выносимые на защиту:

1. Собраны данные в формате изображений жестов рук для обучения моделей искусственных нейронных сетей.

2. Обучены модели машинного и глубокого обучения для распознавания казахского жестового алфавита.

3. Построена собственная архитектура гибридной модели на основе метода обучения с учителем сверточной нейронной сети и метода глубокого обучения рекуррентной нейронной сети.

4. Разработано программное обеспечение для распознавания казахского жестового алфавита.

Теоретическая и практическая значимость. Теоретическая значимость исследования заключается в разработке и обосновании нового подхода к решению задачи классификации для распознавания казахского жестового алфавита, обученного на собственных собранных данных для обучения модели искусственной нейронной сети на алгоритмах/методах машинного обучения.

Практическая значимость заключается в разработке программного обеспечения для распознавания алфавита казахского жестового языка в реальном времени.

Достоверность результатов. Достоверность результатов диссертационной работы подтверждается правильностью поставленных исследовательских задач и применением математических формул и расчетов, а также вычисления и получения результатов посредством обучения собственной нейронной сети на принципе работы архитектуры многослойного перцептрона, разработав гибридную архитектурную модель сверточной и рекуррентной нейронной сети и получив обученные и тестовые результаты на основе таких метрик как: accuracy, recall, precision и f1-score, confusion matrix (матрицы ошибок), а также программной реализацией и вычислительными экспериментами, проведенными в среде разработки PyCharm и Jupyter Notebook на языке программирования Python и библиотек для работы с компьютерным зрением и машинным обучением.

Апробация диссертационной работы. Получены акты внедрения по результатам исследования диссертационной работы в таких организациях как ТОО «Verigram» и ТОО «Smart-edu.kz». ТОО «Verigram» создает комплексные решения в области распознавания и верификации документов, лиц, объектов, а также внедряет технологии OCR и биометрии для улучшения качества обслуживания клиентов, защиты от мошенничества. ТОО «Smart-edu.kz» в свою очередь представляет образовательные курсы, предназначенные для профессионального развития и обучения взрослых лиц в различных областях и профессиях, например таких как компьютерная графика, Web-программирование и т. д. В данных компаниях в исследовательских целях пройдены тестирование программного продукта. Результаты о проделанной работе докладывались на следующих конференциях:

1. Uskenbayeva R.K., Mukhanov S.B. Contour analysis of external images, ACM International Conference Proceeding Series, 3410811, 2020.

2. Mukhanov S.B., Uskenbayeva R.K. Pattern Recognition with Using Effective Algorithms and Methods of Computer Vision Library, Advances in Intelligent Systems and Computing, 2020.

3. Mukhanov S.B., Tursunov S.A., Izteleuov N.E., Tazabekov A. (2019) data science and machine learning «Changing Kazakhstan Society Using Smart Technologies».

4. Aitulen A.D., Mukhanov S.B., Khassenova G.I. (2019) data science and machine learning «Face Recognition Through Various Facial Expressions».

5. Slyamkhan S.M., Yembergenov A.A., Bordousov N.S., Mukhanov S.B. (2019) data science and machine learning «Game Application with Machine Learning Elements».

Связь с государственными программами. Обзорная статья диссертационной работы была опубликована в Материалах Всемирного конгресса по глобальной оптимизации: 6th World Congress on Global Optimization, WCGO 2019 (Metz, France) в рамках реализации проекта программно-целевого финансирования (ИРН №BR05236517). Статья с полученными результатами была опубликована в журнале Eastern-European Journal of Enterprise Technologies, 5 (2-113), pp. 44-54 (процентиль 34) в рамках реализации проекта грантового финансирования (ИРН №AP08053034). Задачи, поставленные в данной диссертации, имеют высокую практическую значимость и непосредственно связаны с процессами внедрения цифровых технологий в производство. Вопросы цифровизации активно обсуждаются и получают особое внимание в выступлениях президента Республики Казахстан Токаева К.К. и в стратегических документах правительства, таких как "Стратегия "Казахстан-2050"" и "Государственная программа "Цифровой Казахстан".

Научные публикации:

1. Mukhanov S.B., Tursunov S.A., Izteleuov N.E., Tazabekov A. (2019) DATA SCIENCE AND MACHINE LEARNING «Changing Kazakhstan Society Using Smart Technologies».

2. Aitulen A.D., Mukhanov S.B., Khassenova G.I. (2019) DATA SCIENCE AND MACHINE LEARNING «Face Recognition Through Various Facial Expressions».

3. Slyamkhan S.M., Yembergenov A.A., Bordousov N.S., Mukhanov S.B. (2019) DATA SCIENCE AND MACHINE LEARNING «Game Application with Machine Learning Elements».

4. Aitulen A.D., Mukhanov S.B., (2019) «Обработка, Идентификация и Распознавание Личности Методом Виолы-Джонса» - «Вестник Казниту», №6.

5. Uskenbayeva R.K., Mukhanov S.B. Contour analysis of external images, ACM International Conference Proceeding Series, 3410811, 2020.

6. Mukhanov S.B., Uskenbayeva R.K. Pattern Recognition with Using Effective Algorithms and Methods of Computer Vision Library, Advances in Intelligent Systems and Computing, 2020, (Индексируется в Базе Scopus процентиль – “28”).

7. Kenshimov, S., Mukhanov, S., Merembayev, T., Yedilkhan, D. A Comparison Of Convolutional Neural Networks For Kazakh Sign Language Recognition (2021) Eastern-European Journal of Enterprise Technologies, 5 (2-113), pp. 44-54. (Индексируется в Базе Scopus процентиль – “34”).

8. Mukhanov S.B.*, Aldanazar A.A., Uatbayeva A.M., Alimbekov A.Ye., Marat G.S COMPETITIVE LEARNING IN NEURAL NETWORKS, International Journal of Information and Communication Technologies, Vol.1, Issue 3, September, 2020 p.70.

9. Mukhanov S.B.*, Alimbekov A.Ye., Marat G.S., Uatbayeva A.M., Aldanazar A.A. AUTOMATION OF STAFF RECRUITMENT AND ASSESSMENT,

International Journal of Information and Communication Technologies, Vol.1, Issue 3, September, 2020, p.117.

10. С.Б. Муханов, А. С. ЛИ, Д. Б. ЖЕКСЕНОВ, Д. Д. ЕВДОКИМОВ, Е. Н. АМИРГАЛИЕВ, Н. К. КАЛЖИГИТОВ, Ш. КЕНШИМОВ. Сравнительный анализ нейросетевых моделей для методов распознавания жестов рук Вестник НИА РК № 2 (88) – 2023, Информационно-коммуникационные технологии.

11. Samat Mukhanov, Raissa Uskenbayeva, Young Im Cho, Kabyl Dauren, Les Nurzhan, Maqsat Amangeldi GESTURE RECOGNITION OF MACHINE LEARNING AND CONVOLUTIONAL NEURAL NETWORK METHODS FOR KAZAKH SIGN LANGUAGE. Вестник Scientific Journal of Astana IT University, VOLUME 15, SEPTEMBER 2023.

Структура и объем диссертации. Диссертация состоит из введения, 4 основных глав, заключения, списка литературы и приложений.

Во введении дано обоснование актуальности исследования темы диссертационной работы. Поставлены и сформулированы цель, объект, предмет, методы и задачи научно-исследовательской работы. Описаны результаты экспериментов, получены результаты исследований, показаны их научная новизна, практическая значимость и апробация результатов исследования диссертационной работы.

В первой главе диссертации описывается теория распознавания образов для задач классификации. Рассмотрены системы распознавания образов, среда и область применения систем распознавания и требования к развитию и улучшению данных систем. Представлены задачи распознавания автомобильных номеров, распознавания лиц и распознавания жестов. Освещена область компьютерного зрения в распознавании изображений, а именно жестов рук.

Во второй главе описаны общие понятия языка жестов и разновидность международных языков жестов, а также дактильный алфавит жестового языка. Освящен и представлен казахский язык жестов, состоящий из 42 (сорока двух) букв в формате изображений. Теоретически описаны построение жеста руки по принципу векторной модели в трехмерном пространстве Эйлера. Математически представлена конструкция модели жестов в виде уравнений в Евклидовом пространстве. Описаны обозначения ключевых точек и обнаружение кончиков пальцев по контуру ладони, а также метод построения каркасного скелета кисти рук.

Третья глава диссертации содержит применение алгоритмов/методов распознавания жестов. Применена и описана задача классификации на основе машинного обучения с использованием машины опорных векторов, рекуррентных нейронных сетей на основе последовательности, сверточных нейронных сетей, а также глубокого обучения. Применена архитектура нейронных сетей для распознавания изображений, а именно жестов рук. Освящена важная часть структуры нейронной сети, в которой применяются картины в роли объектов для обучения, матричные фильтры, параметры свертки и композиция фильтров. Математически описана (вычислено формулами) и доказана (расчетами матричных произведений векторов для каждого слоя

параметров – весов и смещений и применений сложных дифференцированных уравнений) важность применения алгоритма обучения (backpropagation) метода обратного распространения ошибок для вычисления градиента (gradient descent) и необходимость применения функции активаций для минимизации функции потерь (loss function).

Четвертая глава описывает эксперименты исследовательской работы, а именно как распознается казахский жестовый алфавит. Получены результаты с применением метрики для тестирования моделей глубокого обучения. Доказаны точность распознавания данных моделей на основе метрик точности (precision), полноты (recall), F1-score (мера) и поддержки (support) для каждого класса, а также общая точность (accuracy) и средние оценки (average). Протестирована построенная гибридная архитектурная модель нейронной сети с использованием слоев рекуррентной и сверточной нейронной сети. Разработано программное обеспечение, которое распознает жесты в режиме реального времени.

В заключениях каждой главы изложены основные результаты работы, выводы по диссертации и будущие шаги исследования.