

## **АННОТАЦИЯ**

**на диссертационную работу Муталовой Жазиры Саткановны на тему  
«Разработка математической модели распознавания изображений в  
видеопотоках на основе нейронных сетей» на соискание степени  
доктора философии (PhD) по образовательной программе 8D06101 –  
«Аналитика больших данных»**

**Актуальность темы исследования** определяется необходимостью совершенствования математических методов и алгоритмов распознавания изображений в видеопотоках с использованием нейронных сетей для повышения точности классификации, устойчивости к внешним воздействиям и эффективности функционирования интеллектуальных систем компьютерного зрения.

В современных условиях интенсивного развития технологий искусственного интеллекта и компьютерного зрения особую актуальность приобретают задачи разработки математических моделей распознавания изображений в видеопотоках. Анализ и обработка видеоданных в режиме реального времени находят широкое применение в интеллектуальных системах видеонаблюдения, транспортной безопасности, медицине, робототехнике, промышленной автоматизации и системах мониторинга поведения человека.

Существующие методы распознавания изображений в видеопотоках не всегда обеспечивают необходимую точность и устойчивость работы в условиях изменения освещения, фона, масштаба, ракурса и динамики объектов. Кроме того, высокая размерность видеоданных и необходимость учета пространственно-временных характеристик усложняют процесс обработки информации и требуют совершенствования математического аппарата распознавания.

В связи с этим особую научную значимость приобретает разработка математической модели распознавания изображений в видеопотоках на основе нейронных сетей, позволяющей повысить точность классификации, эффективность выделения признаков и устойчивость системы к внешним помехам. Использование современных нейросетевых архитектур обеспечивает автоматизированное формирование информативных признаков и оптимизацию процессов обработки видеоданных.

**Целью диссертационного исследования** является разработка математической модели распознавания изображений в видеопотоках на основе нейронных сетей и создание системы распознавания лиц на основе предложенных моделей и алгоритмов.

**Задачи диссертационного исследования:**

– провести обзор и сравнительный анализ отечественных и зарубежных научных исследований в области распознавания изображений в видеопотоках на основе нейронных сетей;

- разработать математическую модель предварительной обработки изображений с шумовыми воздействиями и распознавания лиц в видеопотоках на основе сверточных нейронных сетей и алгоритмов Retinex;
- разработать многоэтапную модель на основе сверточной нейронной сети с целью повышения точности распознавания изображений в видеопотоках и оптимизации времени обработки кадров;
- построить архитектуру сверточной нейронной сети, направленную на повышение точности классификации данных и эффективности обработки;
- разработать систему распознавания лиц на изображениях в видеопотоках на основе предложенных математических моделей и алгоритмов.

**Объект исследования:** процессы распознавания изображений в видеопотоках.

**Предмет исследования:** математическая модель распознавания изображений в видеопотоках на основе нейронных сетей.

**Методы исследования.** Для решения поставленных в диссертационной работе задач использован комплекс теоретических и прикладных методов исследования. Для построения математической модели процесса распознавания изображений в видеопотоках эффективно применялись методы математического моделирования.

Для автоматического обнаружения и классификации изображений предложено использование метода сверточных нейронных сетей (CNN). Методологическую основу исследования составляют теория систем, системный анализ, методы статистической обработки данных, а также методы проведения вычислительных экспериментов и анализа полученных результатов.

#### **Положения, выносимые на защиту:**

1. Разработана и обоснована математическая модель предварительной обработки изображений с шумовыми воздействиями и распознавания лиц в видеопотоках на основе сверточных нейронных сетей и алгоритмов Retinex.
2. Построена многоэтапная модель на основе сверточной нейронной сети, направленная на повышение точности распознавания изображений в видеопотоках и оптимизацию времени обработки кадров.
3. Разработана архитектура сверточной нейронной сети, ориентированная на повышение точности классификации данных и эффективности обработки.

**Основные результаты исследования.** В результате проведенного исследования сформулированы следующие основные выводы:

Разработана математическая модель предварительной обработки изображений и распознавания лиц на основе сверточных нейронных сетей и алгоритмов Retinex. В рамках предложенной модели были формализованы процессы коррекции неравномерного освещения, нормализации контрастных характеристик и подавления шумовых искажений, что обеспечило повышение качества входных данных. Показано, что применение алгоритмов Retinex на этапе предварительной обработки повышает информативность пространства признаков и существенно увеличивает точность распознавания лиц, особенно в сложных условиях визуальной среды с изменяющимся освещением.

С целью повышения точности распознавания изображений в видеопотоках и оптимизации времени обработки кадров была разработана многоэтапная модель на основе CNN. В предложенной архитектуре использовались модели Faster R-CNN, EfficientNetB0, Self-Attention и FaceNet. Модель Faster R-CNN обеспечила высокоточное обнаружение объектов в сложных сценах и при использовании бэкбона ResNet50 FPN достигла 100% точности при минимальном значении функции потерь 0,0102. Однако полное время обучения составило 183 минуты, что свидетельствует о высоких вычислительных затратах данной архитектуры. Архитектура MobileNetV3 Large показала наибольшую скорость обучения, сократив общее время обучения до 40 минут. Вместе с тем рост значения loss до 0,1807 на последних эпохах продемонстрировал снижение устойчивости модели. Модель EfficientNetB0 с механизмом Self-Attention достигла 100% точности уже на третьей эпохе, а итоговое значение loss составило 0,0289. Использование механизма Self-Attention позволило модели фокусироваться на наиболее информативных областях изображения, что повысило качество распознавания в условиях сложного и изменяющегося освещения. Архитектура FaceNet обеспечила надежное сравнение лиц на основе евклидова расстояния, преобразуя изображения лиц в 128-мерные эмбединги. В результате предложенная многоэтапная модель обеспечила высокоточное и устойчивое распознавание лиц в видеопотоках в режиме реального времени.

С целью повышения точности классификации данных и эффективности обработки была разработана многоуровневая архитектура сверточной нейронной сети. В предложенной архитектуре Nexiface в единый конвейер обработки были интегрированы модуль предварительной обработки Retinex, детектор Faster R-CNN, модель EfficientNetB0, модуль извлечения признаков Self-Attention и модель идентификации FaceNet. Алгоритм Retinex позволил уменьшить влияние неравномерного освещения и повысить качество изображений, что обеспечило увеличение устойчивости нейронной сети. Двухэтапная архитектура детекции Faster R-CNN обеспечила высокоточную локализацию областей лица. Благодаря облегчённой архитектуре EfficientNetB0 удалось снизить вычислительные затраты и сохранить высокую скорость обработки при эффективном извлечении значимых признаков лица. Механизм Self-Attention позволил модели фокусироваться на наиболее информативных областях изображения, таких как глаза, нос и контуры лица, что повысило информативность пространства признаков. Модель FaceNet формировала 128-мерные эмбединг-векторы и за счёт использования triplet loss обеспечивала сближение векторов, принадлежащих одному человеку, и разделение векторов различных лиц, тем самым повышая точность классификации. Разработанная архитектура была адаптирована для работы в режиме реального времени и обеспечила обработку до 30 кадров в секунду с каждой камеры при сохранении общей задержки менее 100 мс.

Практическая значимость полученных результатов подтверждается их внедрением в деятельность организации QUARES. Это позволило автоматизировать процессы интеллектуальной видеоаналитики и повысить надёжность систем мониторинга в сложных эксплуатационных условиях.

Таким образом, совокупность разработанных математических моделей и алгоритмов представляет собой полноценное научно-технологическое решение актуальной задачи в области компьютерного зрения.

#### **Научная новизна исследования.**

1. На основе интеграции свёрточных нейронных сетей (CNN), основанных на глубоком обучении, и алгоритма Retinex разработана математическая модель для обработки высокообъёмных и высокоскоростных видеоданных и распознавания лиц в условиях Big Data. Было установлено, что данная математическая модель может использоваться для обработки видеoinформации и распознавания лиц, подвергшихся воздействию шума.

2. Разработана многоэтапная модель на основе свёрточных нейронных сетей, направленная на повышение точности распознавания изображений в видеопотоках в условиях сложного освещения, шумов и изменений масштаба, а также на оптимизацию времени обработки кадров.

**Теоретическая значимость** исследования заключается в развитии математических моделей распознавания изображений в видеопотоках и методов обработки изображений на основе нейронных сетей.

**Практическая значимость** обусловлена возможностью повышения эффективности распознавания изображений в режиме реального времени путем применения предложенных алгоритмов в системах видеонаблюдения и интеллектуальной видеоаналитики.

**Внедрение результатов исследовательской работы.** Результаты диссертационного исследования внедрены на практике в ТОО «QUARES» с целью повышения точности и скорости идентификации лиц в системах интеллектуальной видеоаналитики в условиях сложной освещенности.

**Апробация результатов диссертационного исследования.** Основные результаты диссертационного исследования докладывались и обсуждались на научных семинарах кафедры «Информационные системы» НАО Казахского агротехнического исследовательского университета имени Сакена Сейфуллина, г. Астана, 2026 г.

**В международных изданиях с ненулевым импакт-фактором, входящим в базу данных Scopus – 2 статьи.** Вклад автора состоит в формулировке научной гипотезы, сборе и предварительной обработке данных, разработке и технической реализации экспериментальной части исследования, интерпретации полученных результатов и подготовке публикации:

1. Mutalova, Z.; Shaushenova, A.; Nurpeisova, A.; et.al. Development of a mathematical model for detecting moving objects in video streams in real-time. IEEE Access. 2024, 10, 136. <https://doi.org/10.1109/ACCESS.2024.3487783>.

2. Nurpeisova, A., Shaushenova A., Kuznetsov O., Ispussinov A., Mutalova Z., Kassymova A. Deep Residual Learning for Face Anti-Spoofing: A Mathematical Framework for Optimized Skip Connections. Technologies, 2025, Volume 13, Issue 9, 413; <https://doi.org/10.3390/technologies13090413>.

**Статьи в журналах, рекомендованных Комитетом по обеспечению качества в сфере науки и высшего образования Министерства науки и высшего образования РК – 2 статьи.** В данных публикациях автор разрабатывал концепцию исследования и методологию анализа данных, участвовал в сборе данных, интерпретации результатов и подготовке научных статей к публикации:

1. Муталова Ж.С., Шаушенова А.Г., Исакова Г.О., Нурпейсова А.А., Онгарбаева М.Б., Абдыкаликова Г.А. Метод распознавания человека по изображению лица на основе перемещения точки по направляющим. Известия Национальной Академии Наук Республики Казахстан. Физико-математическая серия, ISSN 1991-346X Том 3. Номер 351 (2024). 118-137 <https://doi.org/10.32014/2024.2518-1726.296>.

2. Муталова Ж. С., Шаушенова А.Г., Нурпейсова А.А., Айнагулова А. С., Онгарбаева М. Б. Нейрондық желілер негізінде окклюзиялы бет-әлпетті тану әдістерін салыстырмалы талдау. Университет еңбектері. – 2026. №1 (102) 374-380 [https://doi.org/10.52209/1609-1825\\_2026\\_1\\_374](https://doi.org/10.52209/1609-1825_2026_1_374).

**В журналах, входящих в национальную библиографическую базу данных РИНЦ— 2 статьи:**

1. Муталова Ж. С., Шаушенова А. Г., Нурпейсова А. А., Онгарбаева М. Б. Исследование алгоритма распознавания лиц на основе YOLO // Актуальные вопросы современной экономики. — 2025. — № 6. [Электронный ресурс]. URL: <https://avse-journal.ru/Editions>.

2. Муталова Ж. С., Шаушенова А. Г., Нурпейсова А. А., Онгарбаева М. Б. Сверточная нейронная сеть, основанная на функции распознавания лиц // Актуальные вопросы современной экономики. — 2025. — № 6. [Электронный ресурс]. URL: <https://avse-journal.ru/Editions>.

**В международных конференциях – 1 статья:**

1. Муталова Ж. С., Шаушенова А. Г., Нурпейсова А. А. Алгоритм сверточной нейронной сети, основанный на улучшении функций глубокого обучения // Материалы Международной научно-практической конференции «Интеграция искусственного интеллекта в образовательный процесс в эпоху глобальной цифровизации: стратегии, инновации и вопросы кибербезопасности», посвященной 70-летию академика Национальной инженерной академии Республики Казахстан, доктора технических наук, профессора Ахметова Бакытжана Сражатдиновича. — Алматы, 5 апреля 2024 г. — С. 1148–1153.

**Патент на полезную модель – 1:**

1. Муталова Ж. С., Нурпейсова А. А., Шаушенова А. Г., Испусинов А. М., Жумалиева Л. Д. Система распознавания лиц на основе нейронной сети. Патент на полезную модель Республики Казахстан № 9870, 2024. [Электронный ресурс]. URL: <https://gosreestr.kazpatent.kz/Utilitymodel/Details?docNumber=409046>.

**Личный вклад автора.** Исследование включает постановку задачи распознавания изображений в видеопотоках, разработку концепции применения нейронных сетей для построения математических моделей, а также анализ научной литературы в области компьютерного зрения, глубокого обучения и обработки изображений. Автором разработана архитектура интеллектуальной системы обработки видеоданных, предложены математические модели обработки изображений, направленные на снижение влияния шума и искажений, а также обоснован многоэтапный подход, объединяющий алгоритмы Retinex и сверточные нейронные сети. Кроме того, разработаны модели глубоких нейронных сетей, направленные на повышение точности распознавания изображений, математические методы улучшения качества распознавания лиц, а также алгоритмы, повышающие эффективность

обработки видеок кадров. В рамках исследования выполнена программная реализация прототипа системы, проведена подготовка и обработка данных, осуществлены вычислительные эксперименты и выполнен анализ полученных результатов. Все основные результаты диссертационной работы получены автором лично.

**Публикации по теме диссертационного исследования.** По теме диссертационного исследования было опубликовано 8 (восемь) научных трудов, из них 2 (две) статьи в журналах, рекомендованных Комитетом по обеспечению качества в сфере науки и высшего образования Министерства науки и высшего образования РК, 2 (две) статьи в международном научном издании, входящих в базу Scopus, в журналах, входящих в национальную библиографическую базу данных РИНЦ — 2 статьи, в международных конференциях – 1 статья, также имеется 1 (один) патент на полезную модель.

**Структура и объём диссертационной работы.** Диссертация состоит из введения, пяти разделов, заключения, списка использованной литературы и приложений. Объём диссертационной работы составляет 148 страницы (56 рисунков и 7 таблиц).

Во введении обоснована актуальность темы исследования, связанная с необходимостью разработки и совершенствования методов распознавания изображений в видеопотоках на основе нейронных сетей, направленных на повышение точности и эффективности обработки видеоданных в условиях изменяющихся внешних факторов и зашумлённой среды.

В первой главе были рассмотрены основные методы распознавания изображений в видеопотоках, включая традиционные подходы и методы глубокого обучения. Установлено, что традиционные методы обладают высокой вычислительной эффективностью, однако их возможности адаптации к сложным условиям данных ограничены. В задачах обработки в режиме реального времени нейронные сети продемонстрировали более высокую точность и устойчивость. Также было отмечено, что трансформерные модели обладают высоким потенциалом, однако требуют значительных вычислительных ресурсов. Проведен анализ основных типов архитектур нейронных сетей, применяемых для обработки изображений, а также принципов их работы. Установлено, что сверточные нейронные сети (CNN) являются одними из наиболее эффективных моделей для решения задач распознавания, классификации изображений и обнаружения объектов.

Во второй главе была разработана математическая модель предварительной обработки изображений и распознавания лиц на основе сверточных нейронных сетей и алгоритмов Retinex. В рамках предложенной модели были формализованы процессы повышения качества входных данных за счёт коррекции неравномерного освещения, нормализации контрастных характеристик и подавления шумовых искажений. Доказано, что применение алгоритма Retinex на этапе предварительной обработки повышает информативность пространства признаков и существенно увеличивает точность распознавания лиц, особенно в сложных условиях визуальной среды с изменяющимся освещением. Также была разработана многоэтапная модель, направленная на повышение точности распознавания изображений в видеопотоках и сокращение времени обработки. Предложенный подход включает комплексный конвейер, состоящий из этапов предварительной

обработки, обнаружения областей лица, извлечения признаков и идентификации. Интеграция алгоритма Retinex с механизмами Faster R-CNN, EfficientNetB0 и Self-Attention позволила повысить качество изображений и точность распознавания. Формирование эмбеддингов осуществлялось с использованием модели FaceNet.

В третьей главе была разработана многоуровневая архитектура сверточной нейронной сети, направленная на повышение точности классификации данных и эффективности обработки. В предложенной архитектуре Nexiface в единый конвейер обработки были интегрированы модуль предварительной обработки Retinex, детектор Faster R-CNN, модель EfficientNetB0, модуль извлечения признаков Self-Attention и модель идентификации FaceNet. Алгоритм Retinex обеспечил уменьшение влияния неравномерного освещения и повышение качества изображений, что способствовало увеличению устойчивости нейронной сети. Двухэтапная архитектура детекции Faster R-CNN позволила с высокой точностью определять области лица. Благодаря облегчённой архитектуре EfficientNetB0 удалось снизить вычислительные затраты при сохранении высокой скорости обработки и эффективного извлечения значимых признаков лица. Механизм Self-Attention обеспечил фокусировку на наиболее информативных областях изображения, таких как глаза, нос и контуры лица, что повысило информативность пространства признаков. В результате была разработана высокопроизводительная и устойчивая архитектура Nexiface, эффективность которой была подтверждена в ходе практической реализации.

В заключении приведены основные результаты исследования и сформулированы выводы, подтверждающие достижение поставленной цели и решение задач диссертационной работы. В приложениях представлены патент на полезную модель и акт внедрения.

Диссертационная работа выполнена в рамках научного проекта AP23486538 «Исследование и разработка системы распознавания образов в видеопотоках на основе искусственного интеллекта» грантового финансирования МНВО РК на 2024–2026 гг.», реализуемого в соответствии с приоритетными направлениями научно-технологического развития Республики Казахстан.