РАЗРАБОТКА CNN-ОСНОВАННОГО АЛГОРИТМА КЛАССИФИКАЦИИ ДЛЯ ОДНОВРЕМЕННОГО РАСПОЗНАВАНИЯ ЛИЧНОСТИ И ВЫРАЖЕНИЯ ЛИЦА

Камилов Мирзаян Мирзаахмедович
Ташкентский университета информационных технологий имени Мухаммада Аль-Хоразми профессор кафедры и академик «Программное обеспечение информационных технологий»
Саъдуллаев Тохиржон Азаматович
Студент Ташкентский университета информационных технологий имени Мухаммада Аль-Хоразми

Qabul qilindi: 16.12.2025
Ko‘rib chiqildi: 17.12.2025
Tasdiqlandi: 18.12.2025

Ключевые слова: распознавание лиц, выражения лица, CNN, мультимодальные данные, мультизадачное обучение, landmarks.

Аннотация. Рассматривается задача одновременного распознавания личности человека и классификации выражения лица по изображению. Предлагается мультизадачная CNN-архитектура, объединяющая две модальности: (1) визуальные признаки, извлекаемые из нормализованного изображения лица, и (2) структурные признаки, формируемые на основе ключевых точек (landmarks). В отличие от традиционных решений, где идентификация и эмоции решаются раздельно, модель обучается совместно и использует единое мультимодальное представление признаков. Экспериментальная проверка на открытых наборах данных показывает преимущество совместного обучения и интеграции ключевых точек по сравнению с одномодальными базовыми подходами. Полученные результаты подтверждают перспективность объединения информации о лице и его выражении в рамках единого CNN-основанного алгоритма.

1. Введение

Распознавание личности по лицу и анализ выражений лица являются важными задачами компьютерного зрения и биометрии, востребованными в системах контроля доступа, видеонаблюдении, цифровой безопасности, а также в человеко-машинных интерфейсах и мониторинге эмоционального состояния. Классический подход предполагает раздельное решение двух задач, что ограничивает использование взаимосвязанной информации

изображения лица: геометрии, мимики и локальных деформаций ключевых областей (глаза, рот, брови). Современные CNN позволяют строить мультизадачные и мультимодальные модели, где несколько задач решаются в рамках единой архитектуры.

Цель работы: разработать CNN-алгоритм, который совместно определяет личность и выражение лица, объединяя признаки изображения и ключевых точек, и сравнить его с одномодальными базовыми моделями.

Постановка задачи: Необходимо разработать алгоритм, который по входному изображению лица одновременно:
определяет личность человека;
определяет выражение лица.
При этом модель должна учитывать две модальности:
визуальную информацию (изображение лица);
структурную информацию (ключевые точки лица)

Предлагаемый метод: Подготовка данных включает обнаружение и выравнивание лица, извлечение ключевых точек и их нормализацию. Далее формируется мультимодальное представление: изображение подается в CNN, а ключевые точки используются как дополнительный источник информации (например, вектор координат или карты признаков).
Архитектура модели состоит из:
визуальной ветки (CNN), извлекающей признаки из изображения;
структурной ветки, обрабатывающей ключевые точки;
блока слияния признаков, формирующего единое представление;
двух выходных классификаторов: для личности и для выражения лица.
Обучение выполняется в режиме мультизадачного обучения, что позволяет модели совместно оптимизировать обе задачи.

Экспериментальная часть: Оценка проводится на открытых наборах данных для распознавания лиц и выражений лица. Сравниваются:
базовые модели, обученные отдельно на каждую задачу;
предложенная модель, обученная совместно и использующая изображение ключевые точки.
Используются метрики Accuracy, F1-measure и матрицы ошибок для анализа качества и типичных ошибок классификации.

Заключение: Предложен CNN-основанный алгоритм, который объединяет распознавание личности и выражения лица в рамках одной мультимодальной мультизадачной модели. Использование изображения лица совместно с ключевыми точками и совместное обучение двух задач позволяет улучшить качество классификации по сравнению с раздельными одномодальными подходами. Результаты подтверждают эффективность и перспективность предложенного метода для биометрических и интеллектуальных систем.

Литература

[1] Parkhi O.M., Vedaldi A., Zisserman A. Deep Face Recognition
[2] Schroff F., Kalenichenko D., Philbin J. FaceNet: A Unified Embedding for Face Recognition and Clustering
[3] Deng J., Guo J., Xue N., Zafeiriou S. ArcFace: Additive Angular Margin Loss for Deep Face Recognition
[4] FER-2013 Facial Expression Recognition Dataset (Kaggle)
[5] Mollahosseini A., Hasani B., Mahoor M.H. AffectNet: A Database for Facial Expression, Valence, and Arousal in the Wild