

УДК 004.8

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА РАСПОЗНАВАНИЯ ЛИЦ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ И ТЕХНОЛОГИИ УЛУЧШЕНИЯ КАЧЕСТВА ВХОДНЫХ ИЗОБРАЖЕНИЙ

Красновский Д.С., магистрант гр. ПИМ-231, II курс,
Негров Ю.В., магистрант гр. ПИМ-231, II курс
Научный руководитель: Пимонов А.Г., профессор, д.т.н.
Кузбасский государственный технический университет
имени Т.Ф. Горбачёва, г. Кемерово

Сегодня нейронные сети, машинное зрение и искусственный интеллект стали неотъемлемой частью современного мира, оказывая значительное влияние на различные сферы жизни. Эти технологии находят применение в медицине, финансах, транспорте, развлечениях и многих других областях [1]. Примеры использования и назначения:

1. Компьютерное зрение: распознавание образов, классификация изображений, детектирование объектов, распознавание лиц и эмоций, сегментация изображений, распознавание жестов, автоматический поиск.
2. Обработка естественного языка: классификация текста, машинный перевод, генерация текста, голосовые помощники.
3. Финансовые приложения: анализ данных, торговые стратегии, безопасность, автоматизация процессов.
4. Промышленность: оптимизация производства, автоматизация, управление качеством.
5. Игры: генерация уровней, оптимизация игровой механики, улучшение алгоритмов балансировки игры, создание сюжетов.
6. Медицина: диагностика заболеваний, мониторинг состояния пациентов, персонализированная медицина.
7. Транспорт: автономные транспортные средства, прогнозирование спроса, управление транспортными потоками.
8. Развлечения: рекомендации, создание контента, интерактивные игры.

Для разных задач используются различные типы нейросетей, каждая из которых имеет свои особенности и преимущества [2]:

1. Свёрточные нейронные сети (CNN) – обработка изображений, распознавание объектов, классификация изображений.
2. Рекуррентные нейронные сети (RNN) – обработка последовательных данных (текст, временные ряды).
3. Генеративно-состязательные сети (GAN) – генерация реалистичных изображений, видео и аудио.

4. Долгосрочные кратковременные памяти (LSTM) – обработка длинных последовательных данных (тексты и аудио).

5. Трансформеры (Transformers) – обработка естественного языка, машинный перевод, генерация текста.

В рамках разработки интеллектуальной системы был реализован комплексный подход, объединяющий несколько передовых нейронных сетей, каждая из которых выполняет ключевую функцию:

1. ArcFace (рис. 1) – глубокая свёрточная нейронная сеть, которая использует специальную «угловую» функцию потерь для распознавания лиц. Главное назначение в том, что он заставляет нейросеть разделять разные лица в особом «угловом пространстве», добавляя специальный зазор между классами [3].

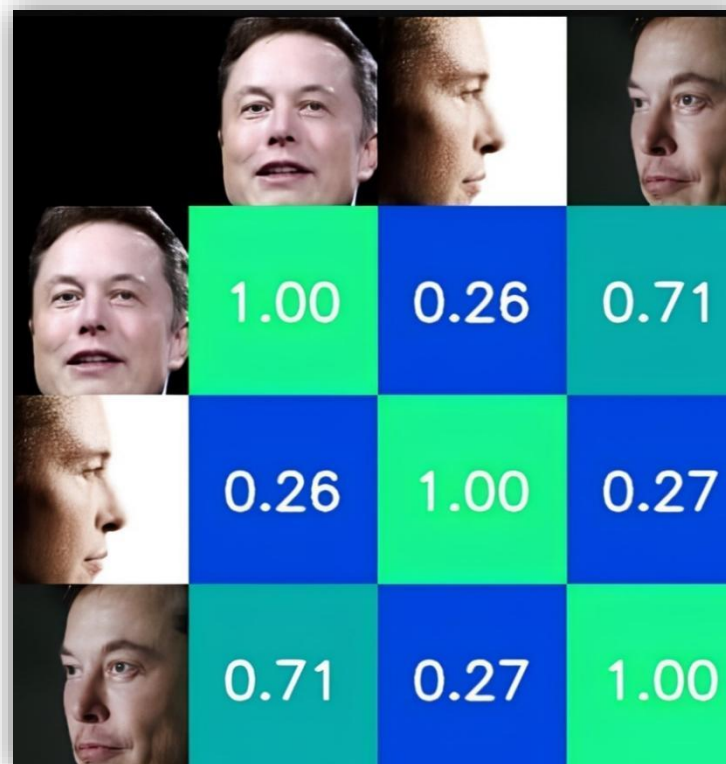


Рисунок 1 – ArcFace

2. GFPGAN (рис. 2) – генеративно-состязательная нейронная сеть, предназначенная для восстановления лиц на фотографиях с повреждениями или низким качеством. Сеть использует предварительную модель лица для анализа его черт и восстановления лиц на фотографиях с повреждениями или низким качеством. Сеть использует предварительную модель лица для анализа его черт и восстановления недостающих частей изображения [4].

Основные компоненты GFPGAN включают:

а. Генератор: создаёт новое изображение лица, опираясь на предварительную модель. Его задача – сделать восстановление изображения максимально реалистичным и естественным.

б. Дискриминатор: оценивает качество восстановленного изображения, сравнивая его с реальными фотографиями. Дискриминатор помогает генератору улучшать свои результаты, предоставляя обратную связь о том, насколько близко его изображения соответствует действительности.

В результате работы этих двух компонентов достигается баланс, при котором генератор способен создавать высококачественные восстановленные изображения лиц.

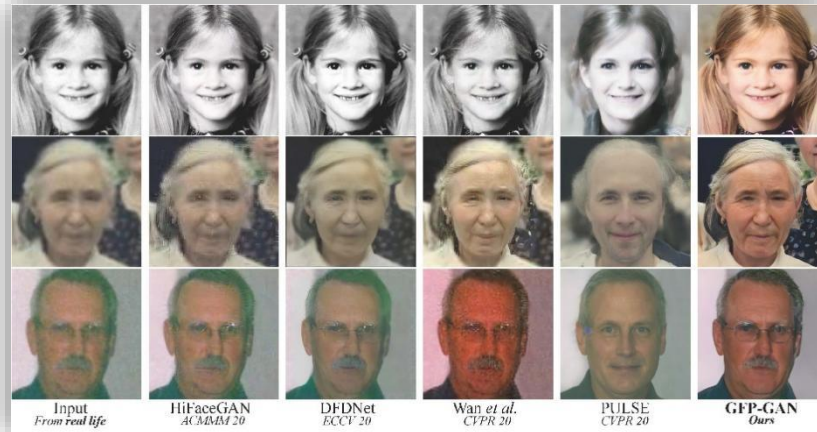


Рисунок 2 – GFP-GAN

3. ESRGAN (рис. 3) генеративно-согласительная нейронная сеть, которая улучшает качество изображений за счёт увеличения их разрешения. Также состоит из генератора и дискриминатора, но больше нацелена на улучшение общего качества картинки на изображении [5].



Рисунок 3 – ESRGAN

4. Yolo11 (рис. 4) – последняя версия серии YOLO, поставляемая с очень лёгкими моделями, которые намного быстрее и эффективнее, чем предыдущие версии Yolo. Она способна выполнять более широкий спектр задач компьютерного зрения, обнаружения объектов, сегментацию изображений, оценку позы и так далее [6].



Рисунок 4 – Yolo11

В результате совмещения этих моделей система была протестирована на наборе данных, где были изображения разных качеств:

1. фотографии с плохим качеством исходного изображения (рис. 5);
2. фотографии с недостаточным уровнем света (рис. 6);
3. фотографии с плохим качеством и обычным уровне освещения (рис. 7).



Рисунок 5 – Результат работы при плохом качестве изображения



Рисунок 6 – Результат работы при недостаточном освещении изображения



Рисунок 7 – Результат работы при плохом качестве и обычном уровне освещения изображения

В процессе обработки видео в реальном времени получилось распознавать объекты и идентифицировать их (рис. 8).



Рисунок 8 – Результат работы при поиске лиц в реальном времени

Список литературы:

1. Как применяются нейросети в современном мире?. – URL: <https://vc.ru/u/1560497-neironochka/634194-kak-primenyayutsya-neiroseti-v-sovremennom-mire> (дата обращения: 28.03.2025).
2. Нейронные сети: принцип работы, перспективы и 159 современных нейронок. – URL: <https://neiros.ru/blog/automation/neyronnye-seti-printsip-raboty-perspektivy-i-159-sovremennykh-neyronok/> (дата обращения: 28.03.2025).
3. ArcFace Loss. – URL: <https://shubham-shinde.github.io/blogs/arcface/> (дата обращения: 28.03.2025).
4. GFT-GAN Photo Restoration Tool. – URL: <https://retouchinglabs.com/gfp-gan-photo-restoration-tool/> (дата обращения: 28.03.2025).
5. Image Super-Resolution with ESRGAN using PyTorch. – URL: <https://www.geeksforgeeks.org/image-super-resolution-with-esrgan-using-pytorch/> (дата обращения: 28.03.2025).
6. YOLO11: Redefining Real-Time Object Detection. – URL: <https://learnopencv.com/yolo11/> (дата обращения: 28.03.2025).