

На правах рукописи

**Найнг Мин Тун**

**СИНТЕЗ СТРУКТУРЫ И АЛГОРИТМОВ ФУНКЦИОНИРОВАНИЯ  
КРОСС-ДОМЕННОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ЛИЦ ДЛЯ  
УСЛОВИЙ НИЗКОЙ ОСВЕЩЕННОСТИ**

Специальность: 05.13.01 – Системный анализ, управление и обработка  
информации (технические системы)

**АВТОРЕФЕРАТ**

диссертации на соискание учёной степени  
кандидата технических наук



Москва – 2021 г.

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана).

Научный руководитель:

**Гаврилов Александр Игоревич**  
кандидат технических наук, доцент

Официальные оппоненты:

**Дивеев Асхат Ибрагимович**  
доктор технических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» РАН.

**Внуков Андрей Анатольевич**

кандидат технических наук, доцент, Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Высшая школа экономики»

Ведущая организация:

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технологический университет «СТАНКИН»

Защита диссертации состоится «21» сентября 2021 г. в 16 час. 30 мин. на заседании диссертационного совета Д 212.141.02 при МГТУ им. Н.Э. Баумана по адресу: 105005, Москва, Госпитальный пер., 10, ауд. 613м.

С диссертацией можно ознакомиться в библиотеке МГТУ Н.Э. Баумана и на сайте <https://www.bmstu.ru/>.

Отзывы и замечания по автореферату в двух экземплярах, заверенные печатью, просьба высыпать по адресу: 105005, Москва, ул. 2-я Бауманская, д. 5, стр. 1 на имя ученого секретаря диссертационного совета.

Автореферат разослан «\_\_\_» \_\_\_\_ 2021 г.

Учёный секретарь  
диссертационного совета Д 212.141.02



И. В. Муратов

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Технологии распознавания лиц вызывают большой интерес у многих исследователей, поскольку не предполагают необходимости в дорогостоящем оборудовании и в кооперации со стороны человека, который будет идентифицирован. Задача распознавания лиц в условиях низкой освещенности является ключевым вопросом в развивающихся и слаборазвитых странах, в которых характерны криминалистические проблемы в условиях слабой освещенности и полной темноты.

Для решения этой проблемы необходимо использовать камеры ночного видения, формирующие изображение в различных диапазонах спектра излучения. Визуализация в ближнем инфракрасном диапазоне обеспечивает недорогое и эффективное решение для получения высококачественных изображений в условиях низкой освещенности или полной темноты. Однако многие приложения требуют, чтобы регистрация шаблонов лиц основывалась на видимых изображениях, таких как онлайн-регистрация и предварительная регистрация с использованием паспорта или удостоверения личности. При сопоставлении изображений лица в ближнем инфракрасном и видимом диапазонах существует разница в модальности изображения, дающая неудовлетворительные результаты распознавания лиц.

Задачи различия в модальности изображения решаются методами кросс-доменного (гетерогенного) распознавания лиц. Несмотря на то, что существует множество методов гетерогенного распознавания лиц, метод циклических генеративных состязательных сетей (ЦГСС) позволяют решить проблему различия модальности изображения обучением модели без входа-выходных пар данных.

Одной из важнейших проблем, связанных с базовой моделью ЦГСС, является необходимость генерации высококачественных синтезированных изображений для использования в системе распознавания лиц.

Недостатки базовой модели ЦГСС для использования в кросс-доменной системе мотивировали исследователей изучать различные структуры, преимущественно на основе глубоких нейронных сетей. Проведено большое число исследований по изучению возможности использования алгоритмов искусственного интеллекта в области гетерогенных систем распознавания лиц. Исследователи использовали различные структуры, объединяющие модули искусственного интеллекта с остальными частями системы распознавания лиц.

Решением задач повышения точности кросс-доменной системы распознавания лиц на основе базовой модели ЦГСС занимаются Wang H., Wang L., Yang X., Yu L., Zhang H., Wu F., Jeremy S. S., You W., Lu W., Zhang B. и другие.

Таким образом, актуальной задачей является повышения точности и надежности распознавания лиц в условиях низкой освещенности.

**Целью** диссертационной работы является разработка структуры и алгоритмов функционирования гетерогенной системы распознавания лиц по изображениям, а также видеопоследовательностям, полученным в условиях недостаточной освещенности. Для достижения цели сформулированы и решены следующие задачи:

1. На основе аналитического обзора существующих методов и алгоритмов, применяющихся в системах гетерогенного распознавания лиц, предложить состав системы и подход к распознаванию лиц в условиях низкой освещенности;
2. Разработать структуру системы гетерогенного распознавания лиц, включающую в себя модель шумоподавления, модель преобразования модальности изображения и модель распознавания лиц;
3. Разработать модель шумоподавления на основе нейронных сетей;
4. Разработать модифицированную математическую модель на основе циклических генеративных состязательных сетей (ЦГСС), позволяющую преобразовать модальность изображения с учетом требуемого формата;
5. Разработать гетерогенную систему распознавания лиц для условий недостаточной освещенности и проблемно-ориентированное ПО;
6. Провести сравнение эффективности разработанного подхода с использующимися на сегодняшний день альтернативными методами гетерогенного распознавания лиц.

### **Объект исследования**

Объектом исследования диссертационной работы является система компьютерного зрения, осуществляющая распознавание лиц на изображениях и в видеопоследовательности в режиме реального времени.

### **Предмет исследования**

Предметом исследования является комбинированный подход к решению задачи распознавания и алгоритмы распознавания лиц.

### **Методы исследования**

В работе использовались методы цифровой обработки изображений, математической статистики, теории глубокого обучения, вероятностных моделей.

**Научная новизна** работы заключается в следующем:

1. Предложен комбинированный подход к решению задачи распознавания лиц на основе метода глубокого обучения, отличающийся наличием модифицированного автоэнкодера, предназначенного для эффективного шумоподавления в ближнем ИК диапазоне, и позволяющий повысить качество работы системы распознавания лиц по сравнению с другими современными альтернативными методами гетерогенного распознавания;

2. Предложена оригинальная структура гетерогенной системы распознавания лиц для работы в условиях пониженной освещенности и зашумленной окружающей среды;

3. Предложена модифицированная математическая модель на основе существующей модели циклических генеративных состязательных сетей (ЦГСС) для преобразования модальности изображения, отличающаяся от известных функцией потерь, и обеспечивающая повышение точности распознавания лиц;

4. Разработан новый алгоритм, основанный на применении предложенного комбинированного подхода, дающий возможность распознавать личность на статических изображениях и на видеопоследовательности в режиме реального времени.

**Научная ценность диссертационной работы** заключается в предложенном оригинальном комбинированном подходе к распознаванию лиц в условиях низкой освещенности, основанном на использовании методов обнаружения лиц, шумоподавления, преобразования модальности изображения, детектирования и классификации признаков, повышающем качество работы системы распознавания лиц как на статических изображениях, так и на видеопоследовательности в реальном времени.

**Практическая ценность диссертационной работы** состоит в том, что разработанные алгоритмы в виде программного обеспечения могут применяться в системах правоохранительных органов для распознавания личности, осуществления судебной экспертизы. Реализованные алгоритмы представлены для гетерогенной системы распознавания лиц на статических изображениях и на видеопоследовательности в режиме реального времени.

**Достоверность полученных результатов проведённых исследований** подтверждается экспериментальными результатами, полученными применении системы распознавания лиц, разработанной на основе предложенных методов и алгоритмов.

#### **Основные научные положения, выносимые на защиту:**

1. Комбинированный подход к решению задачи распознавания личности с использованием оригинальных моделей обнаружения лиц, шумоподавления, преобразования домена изображения, детектирования и классификации признаков, позволяющий повысить качество работы гетерогенной системы распознавания лиц в условиях низкой освещенности;

2. Алгоритм шумоподавления для существенно зашумленных изображений, отличающийся настройками автоэнкодера, и позволяющий повысить качество шумоподавления по сравнению с другими современными альтернативными алгоритмами;

3. Модифицированная математическая модель ЦГСС для преобразования модальности изображения, отличающаяся от известных функцией потерь. Средства моделирования с использованием языка программирования Python, библиотеки глубокого обучения Keras и Tensorflow;

4. Оригинальные структура и алгоритмы функционирования гетерогенной системы распознавания лиц, позволяющие реализовать предложенную методику определения сходства изображений в ближнем ИК диапазоне (NIR) с визуальными эталонами (VIS), хранящимися в базе данных.

#### **Внедрение результатов**

Разработанные варианты реализации гетерогенной системы распознавания лиц внедрены на кафедре информационных технологий машиностроения из

технологического университета (Тханлин) под руководством министерства образования Союза Мьянмы.

### **Апробация работы**

Результаты диссертационной работы докладывались на следующих конференциях:

1. Международная научно-техническая конференция «Пром-Инжиниринг», (Сочи, 2019).
2. IEEE Conference of Russia Young Researchers in Electrical and Electronic Engineering (2020 EIConRus), organized by Moscow, National Research University of Electronic Technology «МИЭТ», (Зеленоград, 2020).
3. Международная научно-техническая конференция «Пром-Инжиниринг», (Сочи, 2020).
4. Всероссийская студенческая конференция «Студенческая научная весна», (Москва, 2018).
5. Всероссийская студенческая конференция «Студенческая научная весна», (Москва, 2019).

### **Публикации по работе**

По теме диссертации опубликовано 7 научных работ, в числе которых 2 работы в журналах, входящих в список, утвержденный ВАК, 4 в материалах конференций, включенных в базу данных SCOPUS и тезисы доклада на научной конференции.

### **Структура диссертации**

Диссертация содержит введение, четыре главы, заключение, список сокращений, список литературы, содержащий 131 наименования. Объем диссертации изложен на 130 страницах, включая 60 рисунок и 15 таблиц и 3 листинга.

## **СОДЕРЖАНИЕ РАБОТЫ**

**Во введении** содержится обоснование актуальности темы диссертации, формулируются цели и задачи исследования, научная новизна, практическая ценность результатов, а также кратко изложено содержание разделов диссертации.

**В первой главе** содержатся общие сведения о биометрических системах. Описываются процессы формирования нового биометрического шаблона на основе данных, основные принципы построения систем распознавания лиц и процедура решения биометрической системой задачи «список наблюдения».

Рассматриваются различные проблемы, с которыми приходится сталкиваться в задаче распознавания лиц. Рассматриваются способы решения проблемы вариаций освещения и отсутствия освещения. Для решения проблемы низкого освещения проведен анализ источников, обеспечивающих спектральные системы визуализации. В качестве подсветки выбирается ближний

инфракрасный диапазон. Анализируются методы, обеспечивающие распознавание лиц с использованием изображений в ближнем ИК-диапазоне.

Гетерогенные системы распознавания лиц обеспечивают соответствие между различными системами визуализации, либо различными спектральными полосами, либо модальностями. Рассматриваются области применения гетерогенного распознавания лиц.

Проведен анализ методов и алгоритмов, используемых в задаче преобразования ближнего инфракрасного домена в визуальный домен. Проведен анализ методов шумоподавления с использованием нейронных сетей в качестве предварительной обработки. Представлена формализация задачи системы распознавания лиц в соответствии данной работы.

**Вторая глава** посвящена методами и алгоритмами решения задач гетерогенного распознавания лиц. Проведена разработка архитектуры системы распознавания лиц, обеспечивающая возможность использования в зашумленной и темной окружающей среде (см. Рисунок. 1). Приводится описание методов, используемых в построении предложенной гетерогенной системы распознавания лиц.

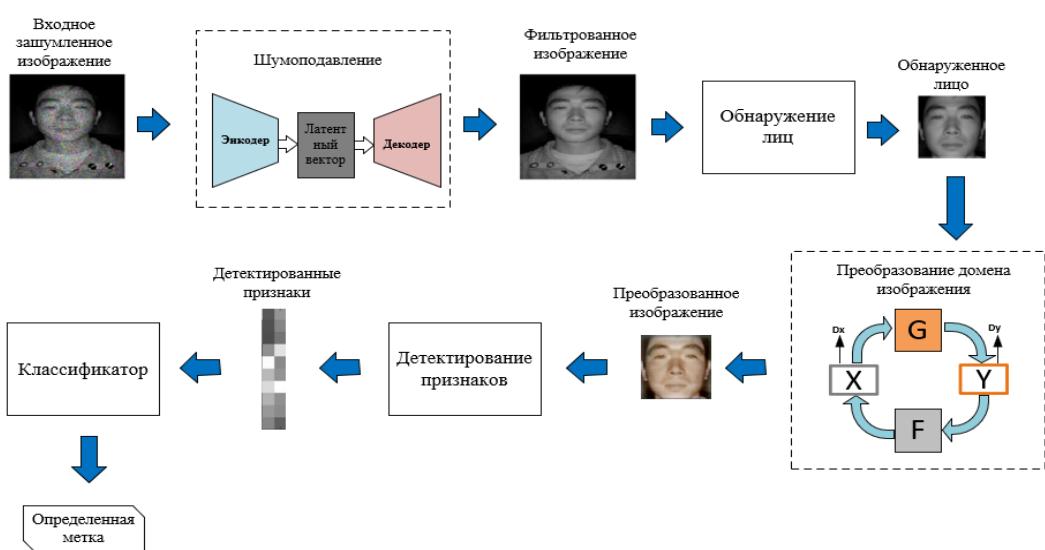


Рисунок 1.

Процедура функционирования гетерогенной системы распознавания лиц.

Разработана модель шумоподавления на основе автоэнкодера и её структура функционирования (Рисунок 2).

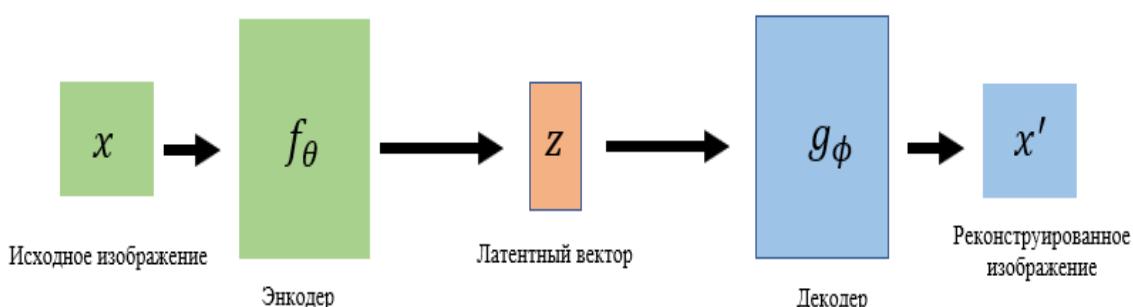


Рисунок 2.

Функционирование автоэнкодера по шумоподавлению изображения.

Математическая модель автоэнкодера имеет следующий вид:

$$\begin{aligned}\theta: x \rightarrow z \\ \phi: z \rightarrow x' \\ \theta, \phi = \arg \min_{\theta, \phi} \|x - (\theta \circ \phi)x'\|^2\end{aligned}$$

где  $\theta, \phi$  – функции энкодера и декодера.  $x, x', z$  – входное изображение, реконструированное изображение и латентный вектор, соответственно.  $\theta$  сопоставляет входное изображение  $x$  с латентным вектором  $z$ .  $\phi$  сопоставляет латентный вектор с выходное изображение  $x'$ .

Сеть энкодера может быть представлена стандартной функцией нейронной сети через функцию активации:

$$z = \sigma(Wx + b)$$

Аналогично, сеть декодера может быть представлена в той же форме, но с другим весом, смещением и используемой функцией активации.

$$x' = \sigma'(W'z + b')$$

где,  $\sigma', W', b'$  параметры сети декодера могут быть не связанны, а  $\sigma, W, b$  – параметры сети энкодера.

Разработана кросс-доменная модель на основе существующей модели ЦГСС для решения задачи преобразования модальности изображения. Результаты преобразования модальности изображения, генерируемые существующей моделью ЦГСС, являются не высокими по качеству формирования изображения, что может привести к плохой производительности системы идентификации лиц (см. Рисунок 3).

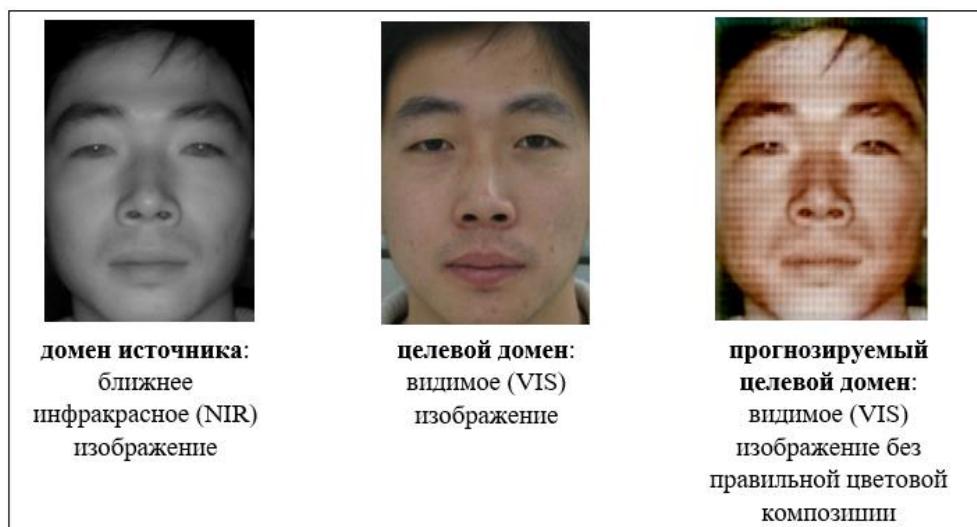


Рисунок 3.

Пример результата преобразования модальности изображения с помощью существующей модели ЦГСС.

На Рисунке 4 представлена разработанная кросс-доменная модель с использованием существующей модели ЦГСС. Модифицированная часть предлагаемой модели представлена в красном пунктирном прямоугольнике.

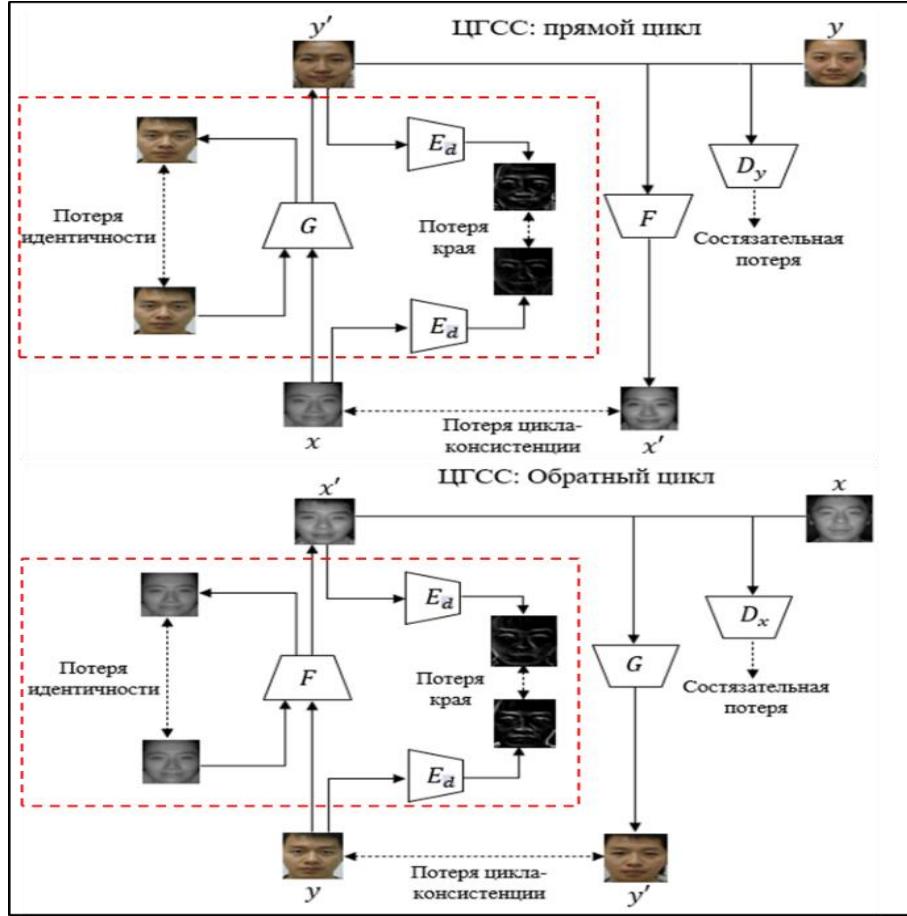


Рисунок 4.

Предложенная кросс-доменная модель для преобразования модальности изображения.

Существующая модель ЦГСС состоит из двух генераторов ( $G, F$ ) и двух дискриминаторов ( $D_x, D_y$ ). Каждый генератор имеет соответствующую модель дискриминатора. Функция генератора состоит в создании синтезированных изображений, а функция дискриминатора состоит в определении различий между реальными изображениями и синтезированными изображениями.

Генератор  $G$  производит синтезированные изображения и дискриминатор  $D_y$  пытается различить эти синтезированные изображения и реальные изображения из домена  $Y$ ; дискриминатор  $D_x$  различает реальные изображения из домена  $X$  и синтезированные изображения, произведенные генератором  $F$ .

Для генератора  $G: X \rightarrow Y$  существующей модели ЦГСС и связанного дискриминатора  $D_y$ , состязательная потеря формулируется следующим образом:

$$L_{adv}^G(G, D_y, X, Y) = \mathbb{E}_{x \sim p_{data(x)}} \left[ (D_y(G(x)) - 1)^2 \right]$$

$$L_{adv}^D(G, D_y, X, Y) = \mathbb{E}_{y \sim p_{data(y)}} \left[ (D_y(y) - 1)^2 \right] + \mathbb{E}_{x \sim p_{data(x)}} \left[ D_y(G(x))^2 \right]$$

где  $x, y$  – исходные данные из доменов  $X$  и  $Y$ ,  $x \in X, y \in Y$ .  $L_{adv}^G(G, D_y, X, Y), L_{adv}^D(G, D_y, X, Y)$  – состязательные потери генератора  $G$  и дискриминатора  $D_y$  для прямого цикла модели ЦГСС.  $\mathbb{E}_x, \mathbb{E}_y$  – матожидание для реальных данных и синтезированных данных.  $p_{data(x)}$  и  $p_{data(y)}$  обозначают

выборочные распределения в доменах  $X$  и  $Y$ , соответственно. Для генератора  $F: Y \rightarrow X$  и связанного с ним дискриминатора  $D_x$ , состязательная потеря представлена следующим образом:

$$L_{adv}^F(F, D_x, Y, X) = \mathbb{E}_{y \sim p_{data(y)}} \left[ (D_x(F(y)) - 1)^2 \right]$$

$$L_{adv}^D(F, D_x, Y, X) = \mathbb{E}_{x \sim p_{data(x)}} [(D_x(x) - 1)^2] + \mathbb{E}_{y \sim p_{data(y)}} \left[ D_x(F(y))^2 \right]$$

где  $L_{adv}^F(F, D_x, Y, X), L_{adv}^D(F, D_x, Y, X)$  – состязательные потери генератора  $F$  и дискриминатора  $D_x$  для обратного цикла модели ЦГСС.

Таким образом, общая потеря состязательности может быть записана следующим образом:

$$L_{adv}^{G,F} = L_{adv}^G(G, D_y, X, Y) + L_{adv}^F(F, D_x, Y, X)$$

$$L_{adv}^D = L_{adv}^D(G, D_y, X, Y) + L_{adv}^D(F, D_x, Y, X)$$

$$L_{adv} = L_{adv}^{G,F} + L_{adv}^D$$

где  $L_{adv}^{G,F}$  – общая состязательная потеря для генератора  $G$  и генератора  $F$ .  $L_{adv}^D$  – общая состязательная потеря для дискриминатора  $D_y$  и дискриминатора  $D_x$ .  $L_{adv}$  – общая состязательная потеря для всех состязательных потерь генераторов и дискриминаторов.

Для восстановления исходного изображения после цикла преобразования, в качестве второго набора функций потерь, потери цикла-консистенции представляется следующим образом:

$$L_{cyc} = \mathbb{E}_{x \sim p_{data(x)}} \left[ \|F(G(x)) - x\|_1 \right] + \mathbb{E}_{y \sim p_{data(y)}} \left[ \|G(F(y)) - y\|_1 \right]$$

где  $L_{cyc}$  – потеря цикла-консистенции.

Полная потеря существующей модели ЦГСС может быть представлена следующим образом:

$$L_{sum} = \lambda_1 L_{adv} + \lambda_2 L_{cyc}$$

где  $L_{sum}$  – полная функция потерь.  $\lambda_1, \lambda_2$  – весовые значения. Результаты преобразования модальности изображения не удовлетворяются при использовании существующей модели ЦГСС, как показано на Рисунке 3. Разработана модификация существующей модели ЦГСС добавлением новых функций потери, называемых потерей идентичности и потерей края. Функция потери идентичности и потеря удержания края представлены следующим образом:

$$L_{identity} = \mathbb{E}_{x \sim p_{data(x)}} [\|F(x) - x\|_1] + \mathbb{E}_{y \sim p_{data(y)}} [\|G(y) - y\|_1]$$

$$L_{edge} = \mathbb{E}_{x \sim p_{data(x)}} \left[ \|E_d(G(x)) - E_d(x)\|_1 \right] + \mathbb{E}_{y \sim p_{data(y)}} \left[ \|E_d(F(y)) - E_d(y)\|_1 \right]$$

где  $L_{identity}$ ,  $L_{edge}$  – потеря идентичности и потеря удержания края.  $E_d$  – Сеть для извлечения границ изображения с целью вычисления предполагаемой потери удержания границ.

Функция потери идентичности и потери удержания края оптимизируют существующую модель ЦГСС во время обучения. Новая модифицированная полная функция потерь представлена следующим образом:

$$L_{sum} = \lambda_1 L_{adv} + \lambda_2 L_{cyc} + \lambda_3 L_{identity} + \lambda_4 L_{edge}$$

где  $\lambda_3, \lambda_4$  – весовые значения.

Рассмотрены известные методы обнаружения лиц, метод детектирования признаков и метод классификации признаков как предварительно обученных моделей.

**В третьей главе** приведена реализация разработанного комбинированного метода для решения задачи распознавания личности в зашумленной и темной окружающей среде. Проведен выбор средств для разработки аппаратного и программного обеспечения.

Реализована модель шумоподавления на основе метода автоэнкодера (Рисунок 5).

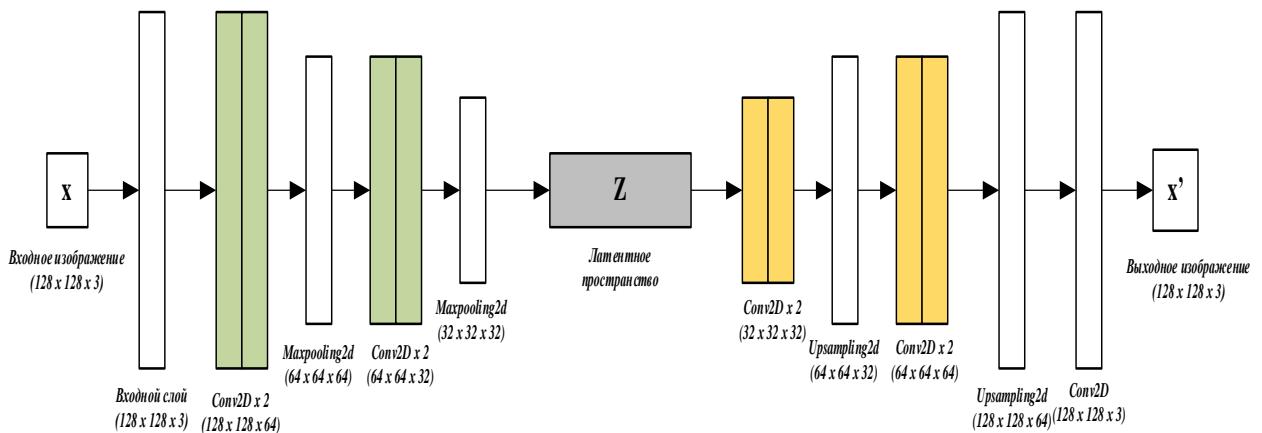


Рисунок 5.

Архитектура модели шумоподавления на основе метода модифицированного автоэнкодера.

В разработанной модели настройки были изменены путем дублирования сверточных слоев в энкодере и декодере.

Приведен сравнительный анализ разработанной модели шумоподавления с другими методами шумоподавления за счет пикового отношения сигнал / шум (PSNR). Чтобы вычислить PSNR, используется следующее уравнение:

$$PSNR = 10 \log_{10} \left( \frac{R^2}{MSE} \right)$$

где  $R$  - максимально возможное значение яркости пикселя изображения. MSE представляет собой среднеквадратичную ошибку между сжатым и исходным изображением:

$$MSE = \frac{\sum_{M,N} \left[ I_1(m,n) - I_2(m,n) \right]^2}{M * N}$$

где  $M$  и  $N$  — это количество строк и столбцов во входных изображениях.

Результаты сравнения представлены в Таблице 1.

Таблица 1.

Сравнительные результаты различных методов шумоподавления за счет пикового отношения сигнал / шум.

значение параметра помехи	С шумом	Медианный метод	Байесшринк	Висуширинк	Модифицированный автоэнкодер
$\sigma = 0.10$	19.99	29.80	29.73	26.69	<b>31.91</b>
$\sigma = 0.15$	16.45	26.62	27.89	23.73	<b>29.88</b>
$\sigma = 0.20$	14.03	24.50	26.84	21.67	<b>27.35</b>

Результаты шумоподавления на случайно выбранном тестовом изображении из ORL базы данных представлены на Рисунке 6.



Рисунок 6.

Сравнение результатов на зашумленном изображении с использованием различных методов шумоподавления.

Проведена реализация модели преобразования модальности изображения в виде программного обеспечения. Предложенная модифицированная кросс-доменная модель построена с пропускными соединениями для генераторов, U-Net для двух детекторов границ и обычные свёрточные сети для дискриминаторов. Построены модели с использованием Keras API, а оптимизатор Adam был использован для изменения атрибутов нейронной сети, таких как вес и скорость обучения, чтобы уменьшить потери.

Для построения предложенной кросс-доменной модели необходимо реализовать два генератора, два дискриминатора и два детектора границ. Сетевые архитектуры генератора, дискриминатора и детектора границ представлены на Рисунке 7, Рисунке 8 и Рисунке 9., соответственно. График потери во время обучения генераторов и дискриминаторов на основе базовой модели ЦГСС и модифицированной модели ЦГСС показаны на Рисунке 10.

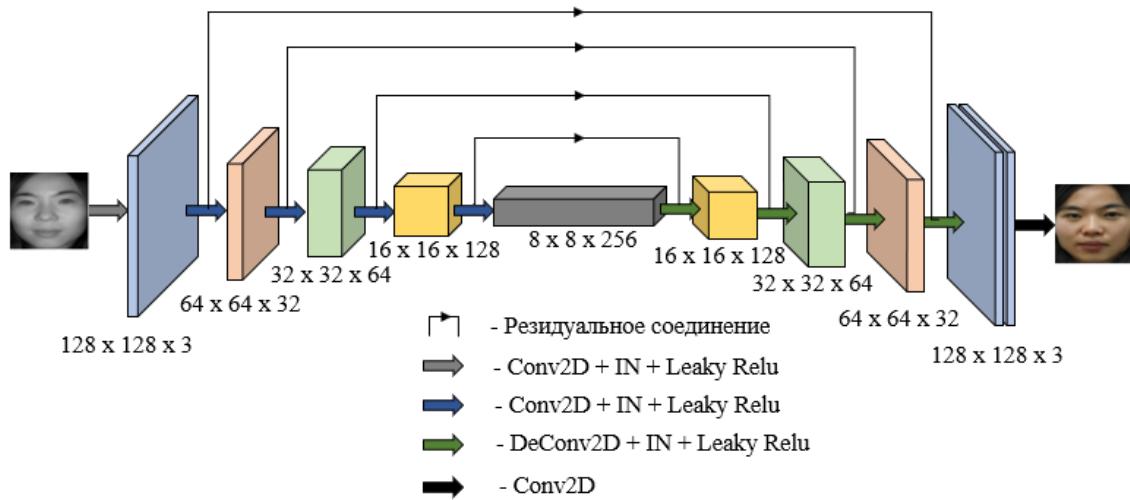


Рисунок 7.  
Сетевая архитектура генератора.

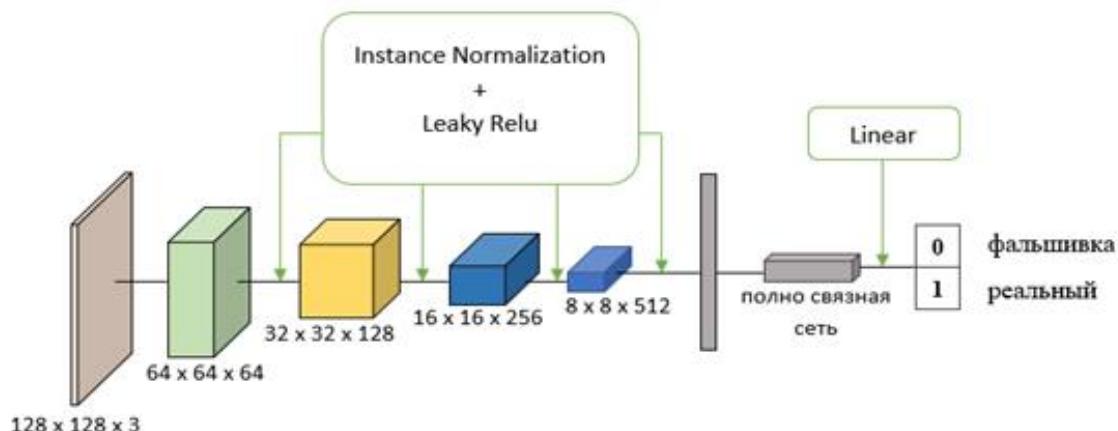


Рисунок 8.  
Сетевая архитектура дискриминатора.

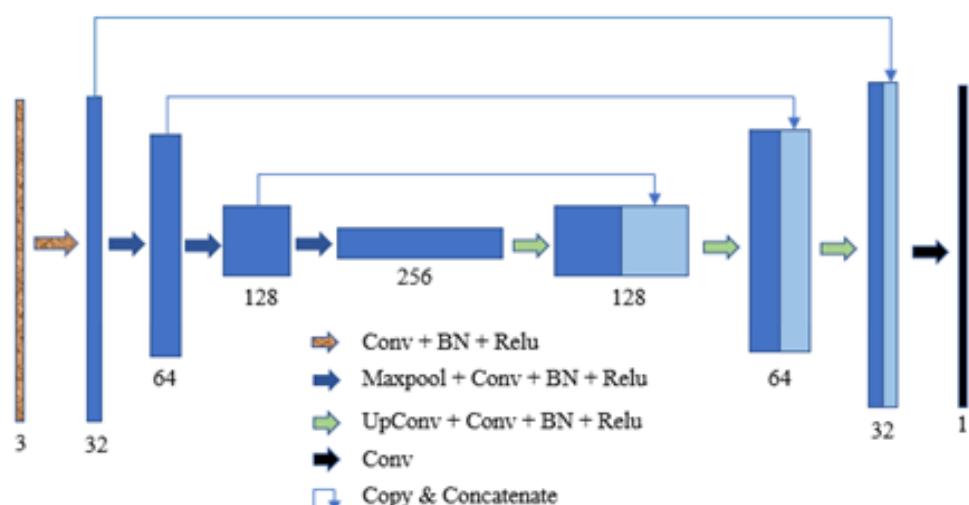


Рисунок 9.  
Сетевая архитектура краевого детектора.

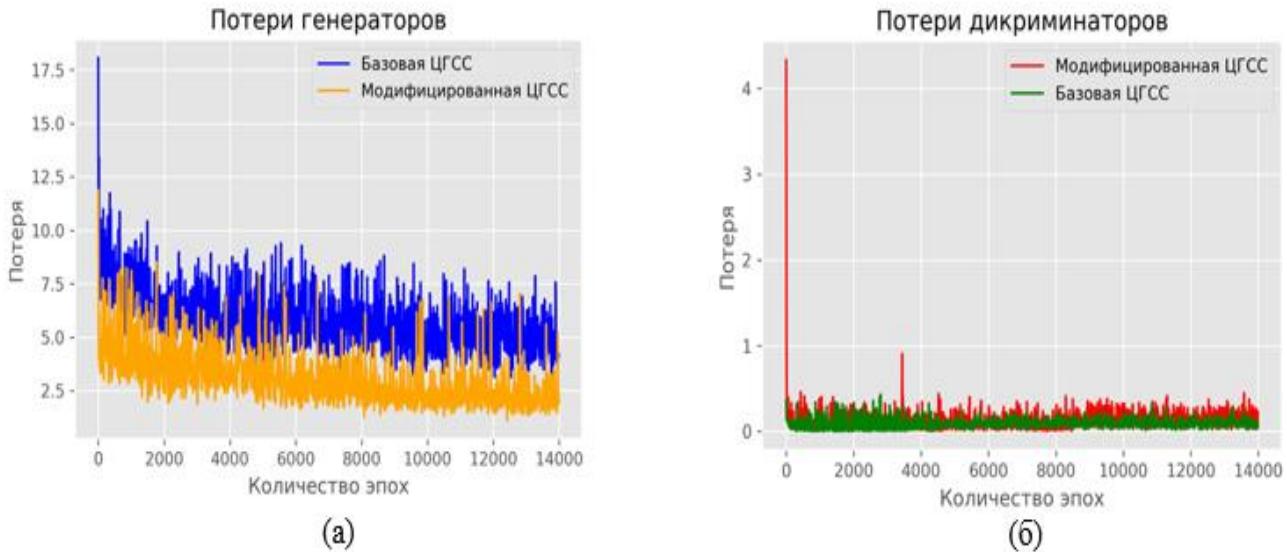


Рисунок 10.

График потери при обучении базовой модели ЦГСС и предложенной кросс-доменной модели (а) потери генераторов (б) потери дискриминаторов.

Разработанная кросс-доменная модель была обучена на базе данных CASIA NIR-VIS 2.0 после построения сетей генераторов и дискриминаторов, результаты представлены на Рисунке 11.



Рисунок 11.

Результаты преобразования домена изображения с использованием предложенной кросс-доменной модели.

Осуществлена реализация методов обнаружения лиц, детектирования признаков и классификации признаков в виде программного обеспечения. Произведен сравнительный анализ трех известных методов обнаружения лиц и представлен обзор преимуществ и недостатков этих методов.

Использована модель "FaceNet" в качестве предварительно обученной модели детектирования признаков. Для классификации личностей использован метод опорных векторов.

**В четвертой главе** приводятся результаты экспериментальных исследований описанных методов. В начале главы проведено планирование эксперимента. Задача идентификации решается для трех различных условий: статические изображения без шумов, статические изображения с шумами и идентификация лиц на видеопоследовательности в режиме реального времени.

Тестирование алгоритма предложенного комбинированного метода проводилось на следующих наборах данных:

- База данных CASIA NIR-VIS 2.0 включает в себя в общей сложности 725 различных людей. На одного человека приходится 1-22 изображения лица в видимом спектре и 5-50 изображений лица в ближнем инфракрасном спектре. База данных использована для тестирования статистических изображений.

- Собственная база данных содержит набор изображений лица 10 разных людей. В базе содержится набор видимого спектра и набор инфракрасного спектра (по 10 изображений на одного человека). База данных использована с целью тестирования видеопоследовательности в режиме реального времени.

Для оценки эффективности системы идентификации лиц с использованием предложенного комбинированного метода применяются следующие критерии:

- критерий полноты системы (Recall), показывающий отношение количества верных идентификаций лиц к общему числу лиц на изображениях.

$$Recall = \frac{TP}{TP + FN}$$

где  $TP$ ,  $FN$  – истинно-положительное решение и ложноотрицательное решение, соответственно.

- критерий точности системы (Precision), показывающий долю верных идентификаций лиц среди всех результатов идентификации.

$$Precision = \frac{TP}{TP + FP}$$

где  $FP$  – ложноположительное решение.

- кривая точности и полноты (precision and recall curve - PRC), используемая для оценки качества классификатора, что является полезным показателем успешности прогнозирования, когда классы изображений несбалансированные.

- общая точность, определяющаяся как доля количества правильной классификации во всем числе образцов.

Результаты эксперимента идентификации лиц на статических изображениях при отсутствии и наличии шумов представлены на Рисунке 12, сравнение с другими альтернативными методами гетерогенного распознавания лиц представлено в Таблице 2.

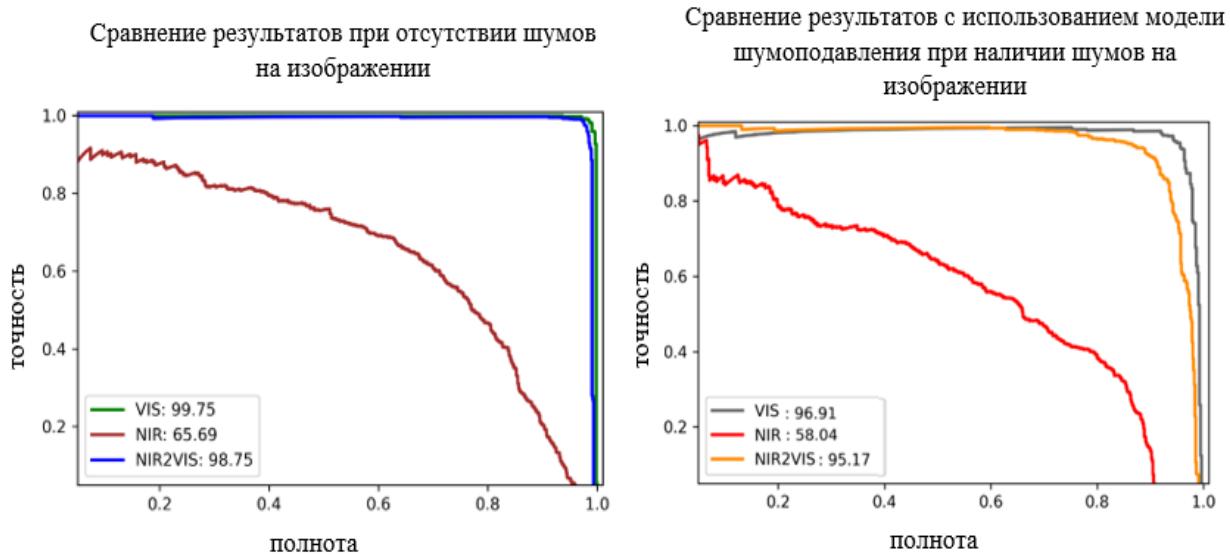


Рисунок 12.

Различные точности идентификации в виде кривой PRC при отсутствии и наличии шумов.

Таблица 2.

Сравнение точности (%) на базе данных CASIA NIR-VIS 2.0 и сравнение времени обработки на видеопоследовательности.

Метод	Точность	Время обработки (кадров в секунду)
DSIFT	$73.3 \pm 1.10$	-
CDFL	$71.5 \pm 1.40$	-
Gabor+RBM	$86.2 \pm 0.98$	-
LCFS	$35.4 \pm 2.80$	-
H2(LBP3)	43.8	-
CEFD	85.6	-
HFR-CNN	$85.9 \pm 0.90$	$\approx 1.18$
TRIVET	$95.7 \pm 0.52$	$\approx 1.43$
IDNet	$87.1 \pm 0.88$	$\approx 1.67$
IDR-128	$97.3 \pm 0.43$	$\approx 1.53$
ADFL	$98.2 \pm 0.34$	$\approx 2.23$
VGG	$62.1 \pm 1.88$	$\approx 1.25$
SeetaFace	$68.0 \pm 1.66$	$\approx 1.82$
CenterLoss	$87.7 \pm 1.45$	$\approx 2.01$
Light CNN	$96.7 \pm 0.23$	$\approx 3.34$
<b>Предложенный метод</b>	<b><math>98.7 \pm 0.39</math></b>	<b><math>\approx 4.02</math></b>

В конце главы для идентификации людей используется видеопоследовательность в режиме реального времени. В этом эксперименте рассматривались вопросы использования разработанной кросс-доменной модели и отказа от использования этой модели в темной среде. Результаты приведены на Рисунке 13.

При отказе от использования разработанной кросс-доменной модели в предложенном комбинированном методе



При использовании разработанной кросс-доменной модели в предложенном комбинированном методе



Рисунок 13.

Результаты тестирования на видеопоследовательности при отказе от использования разработанной кросс-доменной модели и при использовании её в предложенном комбинированном методе.

По результатам идентификации лиц на видеопоследовательности в режиме реального времени, можно заметить, что предложенный комбинированный метод без разработанной кросс-доменной модели дал низкую производительность идентификации лиц, а при использовании разработанной кросс-доменной модели в предлагаемом комбинированном методе можно получить надежную производительность для идентификации лиц.

## ЗАКЛЮЧЕНИЕ

В диссертационной работе решена актуальная научно-техническая задача исследования и разработки алгоритмов идентификации лиц на статических изображениях и на видеопоследовательностях в условиях слабого освещения или абсолютно тёмной окружающей среды.

Представленная диссертационная работа в соответствии своими с целями и задачами содержит следующие результаты:

1. Разработаны алгоритмы распознавания личности в условиях слабого освещения, или полного отсутствия освещения с применением методов шумоподавления, обнаружения лиц, преобразования домена изображения, детектирования и классификации признаков.
2. Разработана модель шумоподавления на основе сети автоэнкодера, обеспечивающая уменьшение шума в изображениях, позволяющая распознавать личность в зашумленных сценариях.
3. Разработана модифицированная кросс-доменная модель на основе существующей модели ЦГСС, позволяющие генерировать видимые изображения лучшего качества из изображений в ближнем ИК диапазоне.
4. Разработаны структура и алгоритмы функционирования системы на основе совокупности разработанных моделей и предварительно

обученных моделей, позволяющие идентифицировать личность в условиях слабого освещения.

5. Разработано и внедрено в автоматизированные системы программное обеспечение для распознавания лиц в условиях слабого освещения или отсутствия освещения для статических изображений и видеопоследовательностей.

## **ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ**

1. Найнг Мин Тун, Гаврилов А. И., Ньян Линн Тун, Пья Пью Паинг, Хтет Аунг. Автоматизированная система шумоподавления с использованием метода глубокого обучения для распознавания лиц // Современная наука: Актуальные проблемы теории и практики. Сер. Естественные и Технические Науки, 2020, №06, С. 136-142. (0,72 п.л./0,5 п.л.).
2. Найнг Мин Тун, Гаврилов А. И., Пья Пью Паинг., Ньян Линн Тун., Хтет Аунг Тху. Обнаружения объектов на основе глубоких нейронных сетей в задаче навигации автономного мобильного робота // Современная наука: Актуальные проблемы теории и практики. Сер. Естественные и Технические Науки, 2021, - №04. – С. 128-134. (0,74 п.л./0,5 п.л.).
3. Naing Min Tun, Gavrilov A. I., Pyae Phyo Paing. Human face recognition using combination of zca feature extraction method and deep neural network // 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), IEEE, 2019, P. 1-5. (0,29 п.л./0,19 п.л.).
4. Naing Min Tun, Gavrilov A. I., Nyan Linn Tun, Pyae Phyo Paing. Face Identification System using Deep Learning Method at Low Lighting Resources // 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EICoRus), IEEE, 2020, P. 2001–2004. (0,31 п.л./0,23 п.л.).
5. Naing Min Tun, Gavrilov A. I., Nyan Linn Tun. Facial image denoising using convolutional autoencoder network // 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), IEEE, 2020, P. 1-5. (0,27 п.л./0,18 п.л.).
6. Nyan Linn Tun, Gavrilov A. I., Naing Min Tun. Multi-classification of satellite imagery using fully convolutional neural network // 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), IEEE, 2020, P. 1-5. (0,25 п.л./0,1 п.л.).
7. Найнг Мин Тун. Классификация изображений на основе сверточной нейронной сети // Всероссийская студенческая конференция «студенческая научная весна». Сборник тезисов докладов, 2018, С. 233-234. (0,13 п.л.).