



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет  
имени Н. Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н. Э. Баумана)

---

ФАКУЛЬТЕТ «Информатика и системы управления»

---

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

---

# РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

## *К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ*

### *НА ТЕМУ:*

*«Разработка метода повышения разрешения  
изображения по нескольким кадрам»*

Студент ИУ7-23М  
(Группа)

\_\_\_\_\_  
(Подпись, дата)

П. Ю. Сироткина  
(И. О. Фамилия)

Руководитель НИР

\_\_\_\_\_  
(Подпись, дата)

М. В. Филиппов  
(И. О. Фамилия)

*2024 г.*

## РЕФЕРАТ

Расчетно-пояснительная записка 20 с., 9 рис., 21 источн., 1 прил.

**Ключевые слова:** цифровое изображение, суперразрешение, свертка, сверточные нейронные сети, контрастно – ограниченная эквализация гистограмм.

Объектом исследования является повышение разрешения изображения по нескольким кадрам низкого разрешения.

Предметом исследования является разработка модификации классической архитектуры сверточной нейронной сети для решения поставленной задачи.

Целью работы являлась разработка метода повышения разрешения изображения по нескольким кадрам.

Для достижения поставленной цели были выполнены следующие задачи:

- изложены особенности предлагаемого метода;
- описаны основные этапы разрабатываемого метода в виде детализированной диаграммы IDEF0 и схем алгоритмов;
- спроектирована структура программного обеспечения для реализации разрабатываемого метода.

В результате была сформулирована архитектура модифицированной сверточной нейронной сети: на первом этапе работы нейронной сети применяются несколько параллельных сверточных слоев с возможным уплотнением карт признаков для каждого входного изображения, затем полученные карты объединяются. Полученные карты признаков передаются в остаточные слои (англ. Residual Blocks) с целью проведения более эффективного обучения сети. В выходном слое к каждой карте также применяется свертка с нелинейной функцией активации, формируя итоговое изображение.

На основе проведенной работы также определено направление дальнейших исследований метода, заключающееся в использовании различных способов объединения карт признаков входных изображений.

# СОДЕРЖАНИЕ

<b>РЕФЕРАТ</b>	<b>3</b>
<b>ВВЕДЕНИЕ</b>	<b>5</b>
<b>1 Конструкторский раздел</b>	<b>6</b>
1.1 Требования и ограничения к разрабатываемому методу . . . . .	6
1.2 Требования к разрабатываемому программному обеспечению . .	7
1.3 Основные этапы разрабатываемого метода . . . . .	8
1.3.1 IDEF0 – диаграмма уровня A0 . . . . .	8
1.3.2 Алгоритм обработки входных изображений . . . . .	9
1.3.3 Алгоритм контрастно – ограниченной эквализации гистограммы . . . . .	10
1.3.4 Свертка (конволюция) . . . . .	11
1.3.5 Алгоритм конкатенации карт признаков . . . . .	13
1.3.6 Архитектура разрабатываемой нейронной сети . . . . .	14
1.3.7 Структура разрабатываемого программного обеспечения	15
1.4 Методы оценки качества восстановления цифрового изображения	15
<b>ЗАКЛЮЧЕНИЕ</b>	<b>17</b>
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>18</b>

## ВВЕДЕНИЕ

Существует множество способов повышения разрешения по нескольким кадрам, таких как: частотные методы, методы на основе интерполяции, методы на основе теории множеств, методы на основе решения обратной задачи с применением регуляризации, примеро – ориентированные методы, методы с использованием сверточных нейронных сетей [1–9].

Выбор метода суперразрешения зависит от конкретной задачи и имеющихся вычислительных и временных ресурсов. Методы на основе глубокого обучения часто дают наилучшие результаты, но они требуют большого объема данных и мощных вычислительных ресурсов для обучения сетей. Аналитические методы более вычислительно эффективны, но могут не давать такие высококачественные результаты, а также не учитывать структурные особенности изображений.

На основе проведенной классификации и сравнительного анализа для дальнейшей разработки было выбрано направление оптимизации сверточных нейронных сетей для повышения разрешения изображения по нескольким кадрам, т.к. в рамках поставленной задачи ключевым фактором является качество получаемого результата, другие же факторы (вычислительная сложность, необходимость пост – или предобработки) являются второстепенными.

Целью работы является разработка метода повышения разрешения изображения по нескольким кадрам.

Для достижения поставленной цели необходимо выполнить следующие задачи:

- изложить особенности предлагаемого метода;
- описать основные этапы разрабатываемого метода в виде детализированной диаграммы IDEF0 и схем алгоритмов;
- спроектировать структуру программного обеспечения для реализации разрабатываемого метода.

# 1 Конструкторский раздел

В данном разделе сформулированы требования и ограничения к разрабатываемому методу. Разработан метод повышения разрешения изображения по нескольким кадрам. Описаны основные этапы разработки в виде детализированной диаграммы IDEF0 и схем алгоритмов, а также изложены особенности предлагаемого метода. Спроектировано программное обеспечение для реализации разрабатываемого метода.

## 1.1 Требования и ограничения к разрабатываемому методу

К методу повышения разрешения изображения по нескольким кадрам предъявляются следующие требования:

1. Метод должен базироваться на использовании сверточных нейронных сетей с необходимыми модификациями.
2. Метод должен подавлять шум, присутствующий в отдельных кадрах, не теряя при этом полезные детали.
3. Метод должен сохранять и улучшать мелкие детали (не только увеличивать размер изображения).
4. Метод должен быть применим для работы с различными типами объектов и сцен, а также различными освещением.

Также представлен ряд ограничений для разрабатываемого метода:

1. Качество восстановления может быть неудовлетворительным, если входные изображения подвержены критичному искажению или некачественно зарегистрированы фотосистемой.
2. Качество восстановления может быть неудовлетворительным, если входные изображения с высоким уровнем шума и/или имеют высокочастотные детали (выбросы интенсивности).

## 1.2 Требования к разрабатываемому программному обеспечению

К разрабатываемому программному обеспечению предъявляются следующие требования:

1. Возможность загрузки изображений в формате PNG, JPG или BMP.
2. Возможность обработки как RGB – изображений, так и изображений в тонах серого.
3. Возможность просмотра результата повышения разрешения изображения.
4. Возможность просмотра информации по процессу обучения модели.
5. Возможность просмотра значения выбранных метрик качества для конкретного случая работы алгоритма.
6. Возможность сохранения результата в отдельный файл в формате PNG, JPG или BMP.
7. Если время выполнения программы может превышать комфортное время реакции системы для человека, необходимо обеспечить вывод предупреждающего сообщения.
8. Разрабатываемое ПО должно корректно реагировать на любые действия пользователя.

## 1.3 Основные этапы разрабатываемого метода

### 1.3.1 IDEF0 – диаграмма уровня A0

На рисунке 1.1 представлена IDEF0 – диаграмма уровня A0 для метода повышения разрешения изображения по нескольким кадрам.

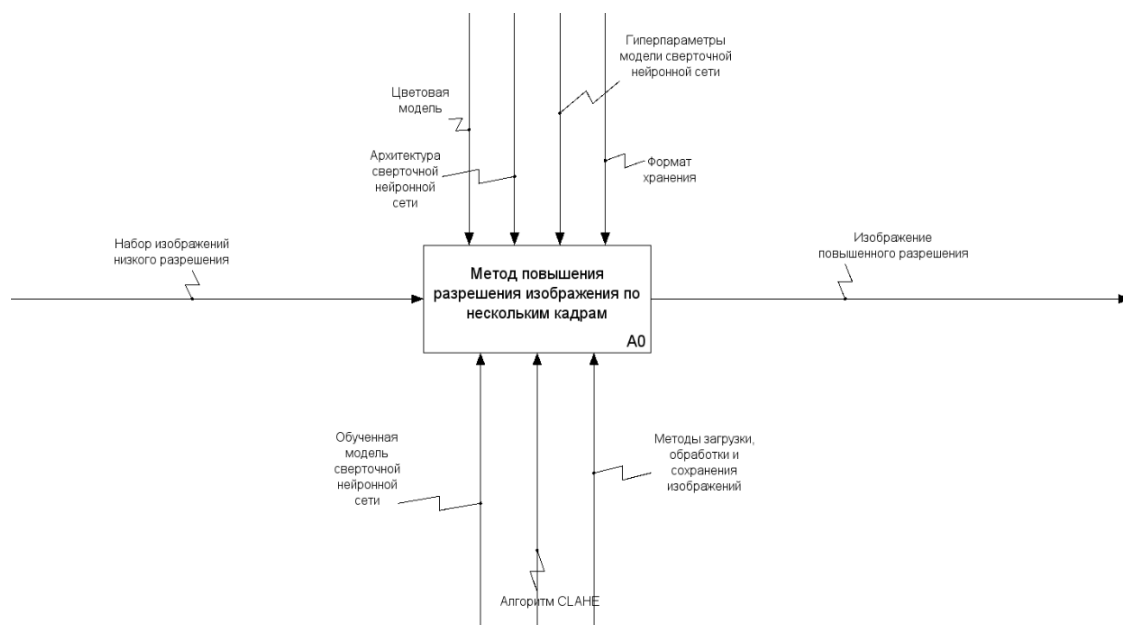


Рисунок 1.1 – IDEF0 – диаграмма уровня A0

На рисунке 1.2 представлена детализированная IDEF0 – диаграмма для метода повышения разрешения изображения по нескольким кадрам.

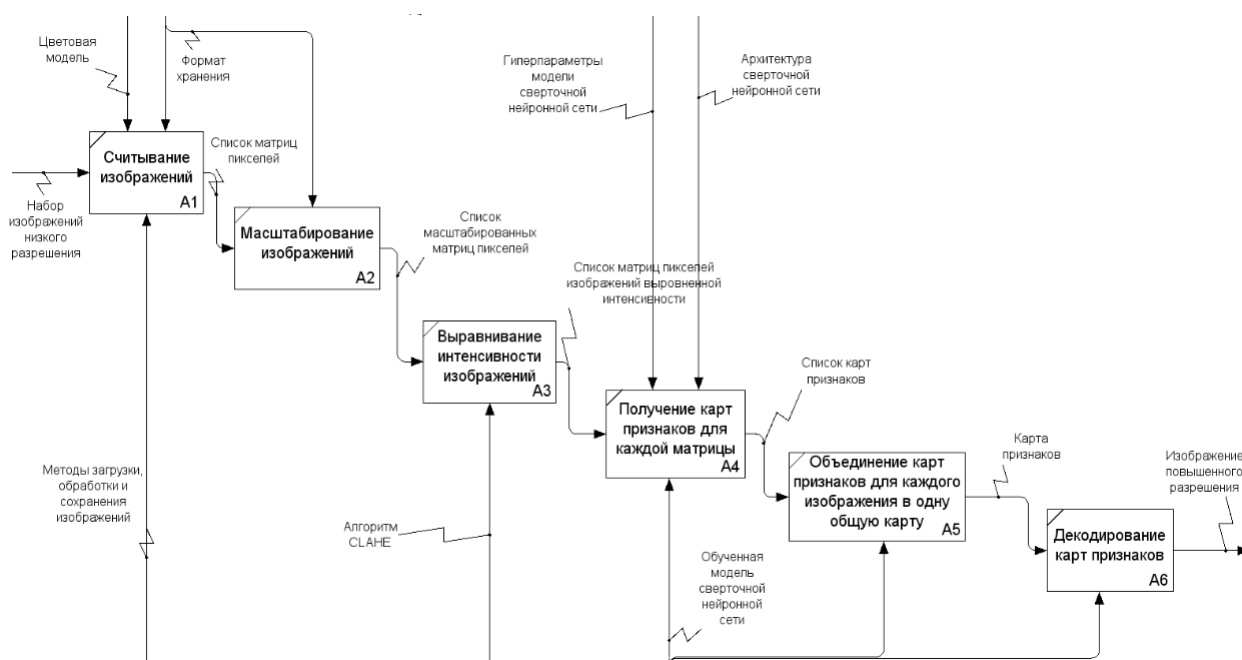


Рисунок 1.2 – IDEF0 – диаграмма уровня A1

Особенностью рассматриваемой архитектуры нейронной сети является обработка нескольких входных изображений (в отличие от одного для классической архитектуры), в связи с чем обработка входного набора изображений также требует дополнительных действий (приведение к одному размеру, выравнивание интенсивности кадров низкого разрешения, конкатенация карт признаков нескольких кадров).

### 1.3.2 Алгоритм обработки входных изображений

На рисунке 1.7 представлен алгоритм обработки входных изображений (кадров низкого разрешения).

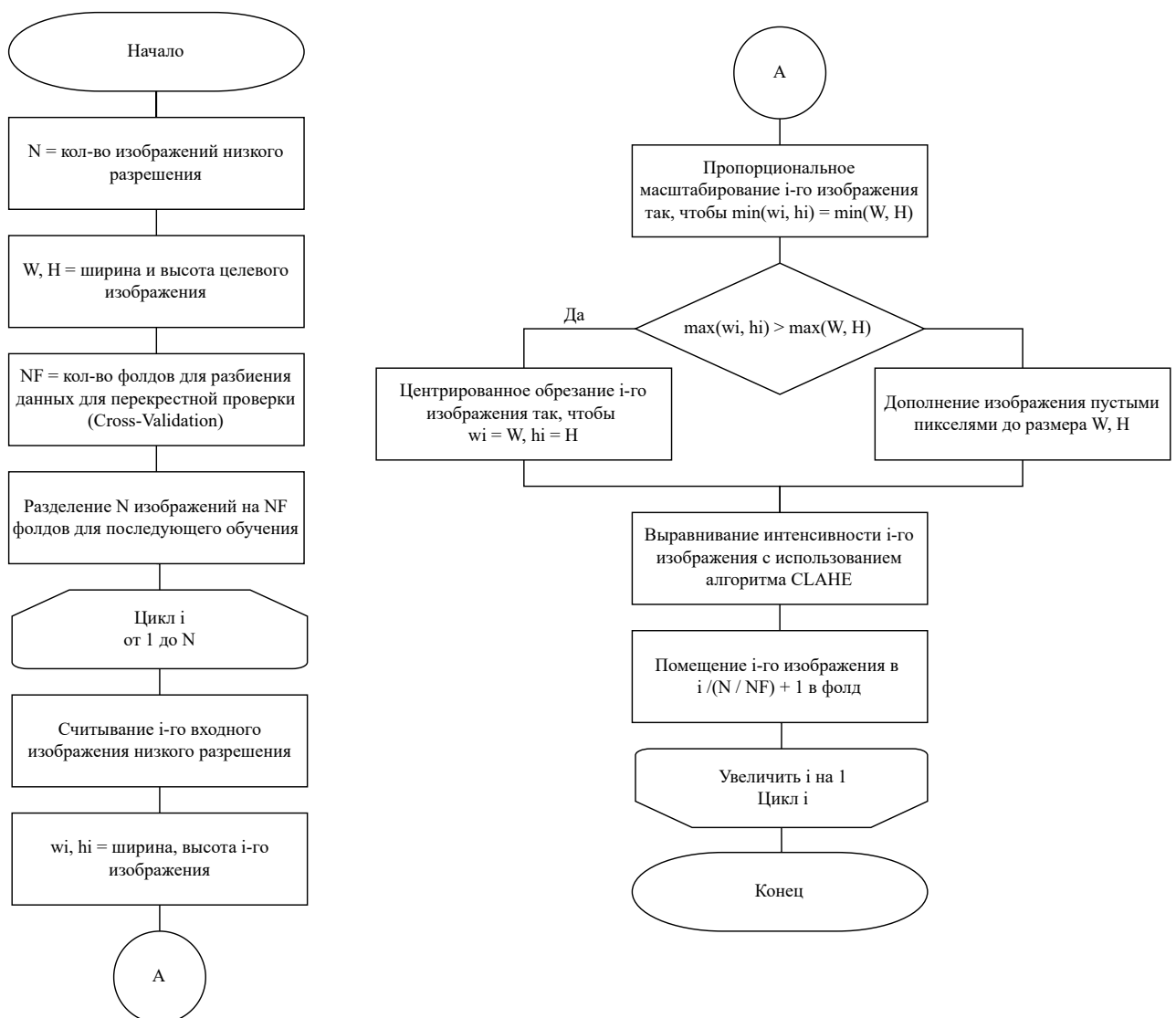


Рисунок 1.3 – Алгоритм обработки входных изображений



### 1.3.3 Алгоритм контрастно – ограниченной эквализации гистограммы

На рисунке 1.4 представлен алгоритм контрастно – ограниченной эквализации гистограммы (CLAHE — англ. Contrast Limited Adaptive histogram equalization), применяемый к каждому входному изображению с целью улучшения видимости деталей в затемненных или засвеченных областях изображения, предотвращая при этом чрезмерное усиление шума.



Рисунок 1.4 – Алгоритм контрастно – ограниченной эквализации гистограммы

В обычной адаптивной гистограммной эквализации (АНЕ) контраст может быть увеличен до такой степени, что небольшие вариации интенсивности, включая шум, становятся сильно заметными, ухудшая качество изображения. CLAHE решает эту проблему, ограничивая усиление контраста в каждой локальной области, что позволяет улучшить видимость деталей без значительного увеличения шума и артефактов.

На рисунке 1.5 представлен пример применения алгоритма CLAHE.

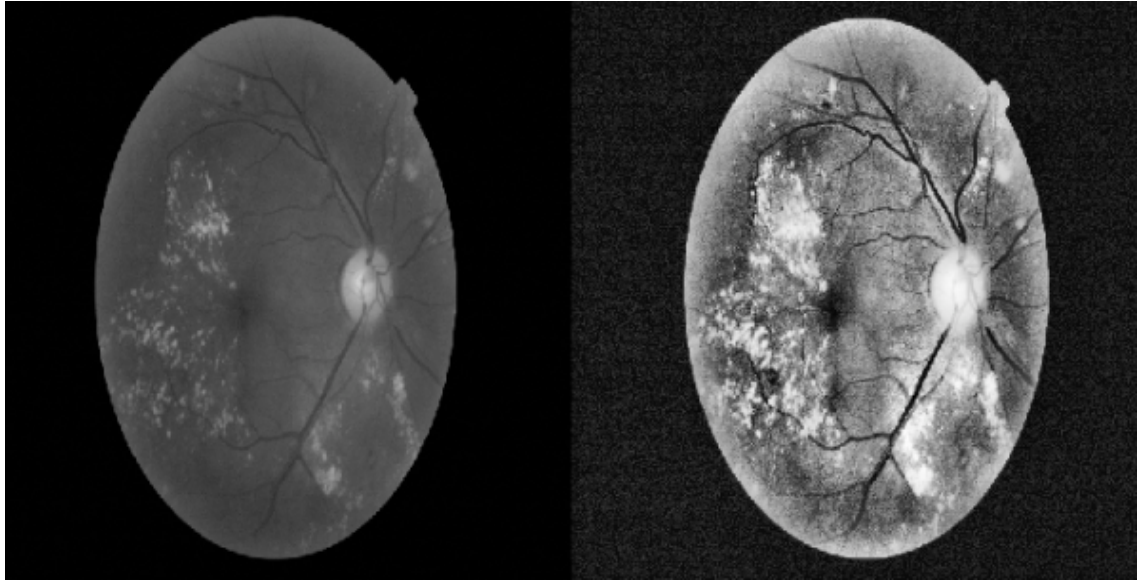


Рисунок 1.5 – Пример применения алгоритма CLAHE

### 1.3.4 Свертка (конволюция)

Применительно к обработке цифровых изображений операция свертки может быть интерпретирована следующим образом: на основе некоторого множества пикселей исходного изображения вычисляется новый пиксель результирующего (искаженного ядром свертки) изображения.

В зависимости от выбранного ядра свертки, применяемого к изображению, можно получить тот или иной эффект: размытость, повышение резкости, обнаружение контуров, граничное обнаружение и т. д.

Математически операцию двумерной свертки цифрового изображения размером  $M \times N$  в пространственной области можно описать в виде выражения (1.1) [10]:

$$f(x, y) \oplus h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cdot h(x - m, y - n). \quad (1.1)$$

На рисунке 1.6 [11] представлен пример выполнения операции свертки. Для вычисления новых значений используется т.н. ядро свертки. На представленном примере ядром является матрица серого цвета размером  $3 \times 3$  ячейки.

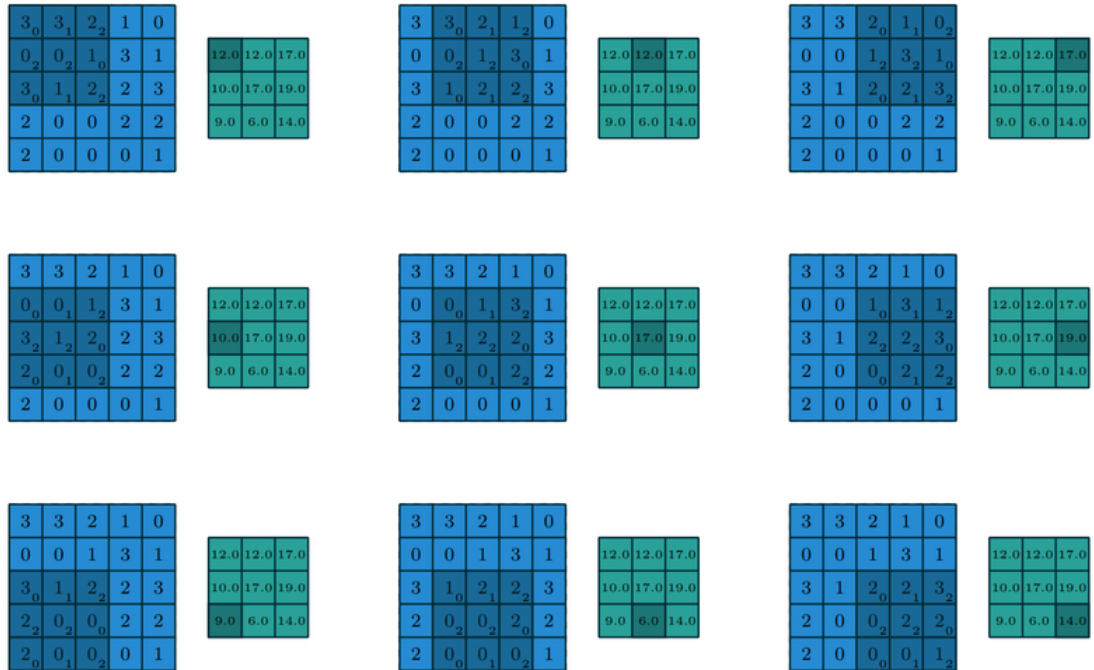


Рисунок 1.6 – Пример конволюции цифрового изображения

Использование свертки в рамках рассматриваемой задачи предоставляет инструмент для извлечения высокоуровневых характеристик изображения, которые затем используются для повышения разрешения целевого изображения посредством составления различных карт характеристик, т.е. применения сверток с различными ядрами, описывающими то или иное свойство. Начальные значения весов нейронов в сети задаются случайным образом. Начальные значения весов нейронов в сети задаются случайным образом. Обычно сверточный слой по умолчанию объединяется с нелинейной функцией активации [12].

После применения сверточного слоя также часто применяются подвыборочные слои с целью уплотнения карты признаков: считается, что обнаружение признака важнее нахождения его координат. Таким образом, карты признаков после прохождения нескольких слоев сети могут преобразоваться в вектор или скаляр.

### 1.3.5 Алгоритм конкатенации карт признаков

На рисунке 1.7 представлен алгоритм конкатенации карт признаков.

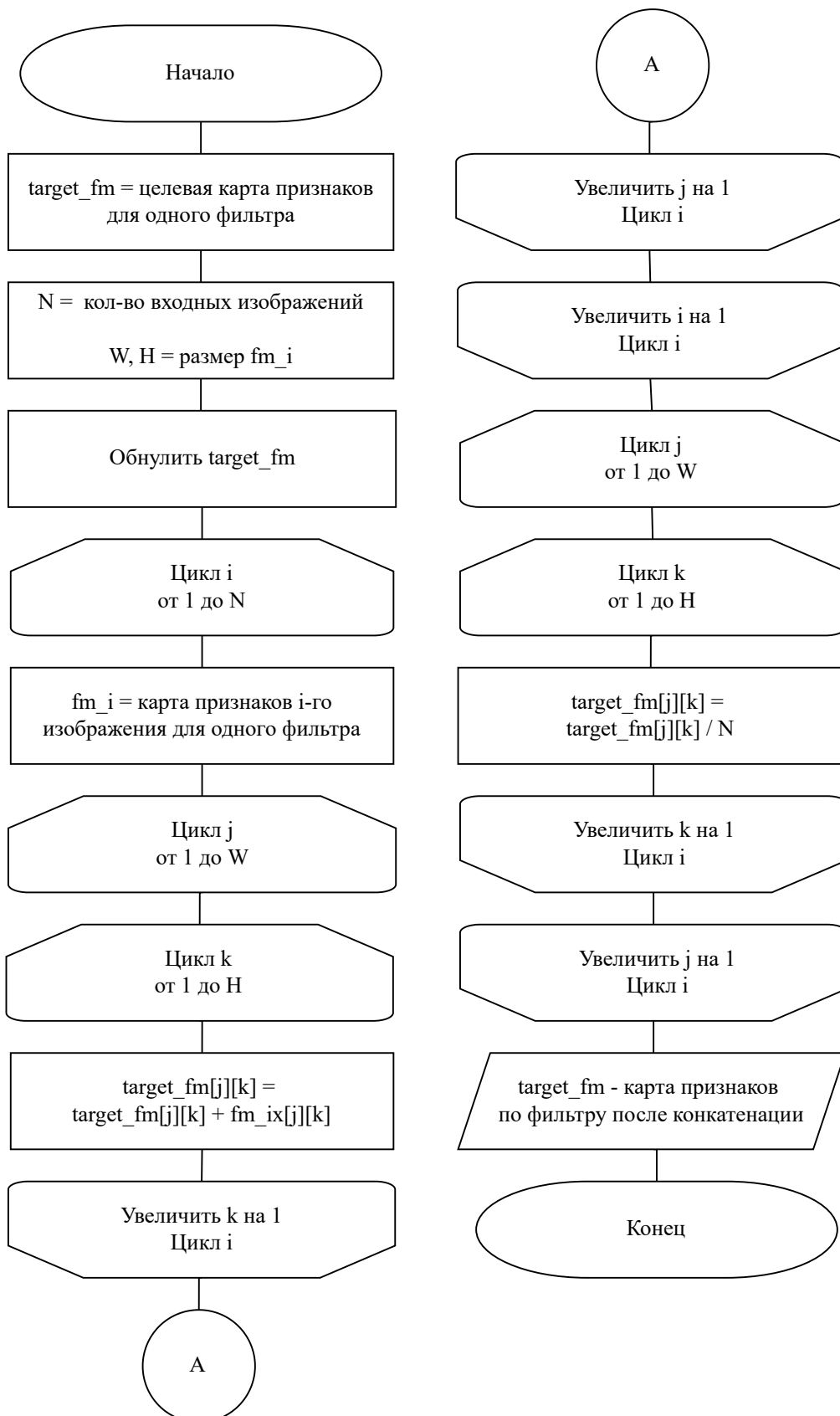


Рисунок 1.7 – Алгоритм конкатенации карт признаков

### 1.3.6 Архитектура разрабатываемой нейронной сети

На рисунке 1.8 представлена базовая архитектура [13; 14] разрабатываемой нейронной сети на данный момент.

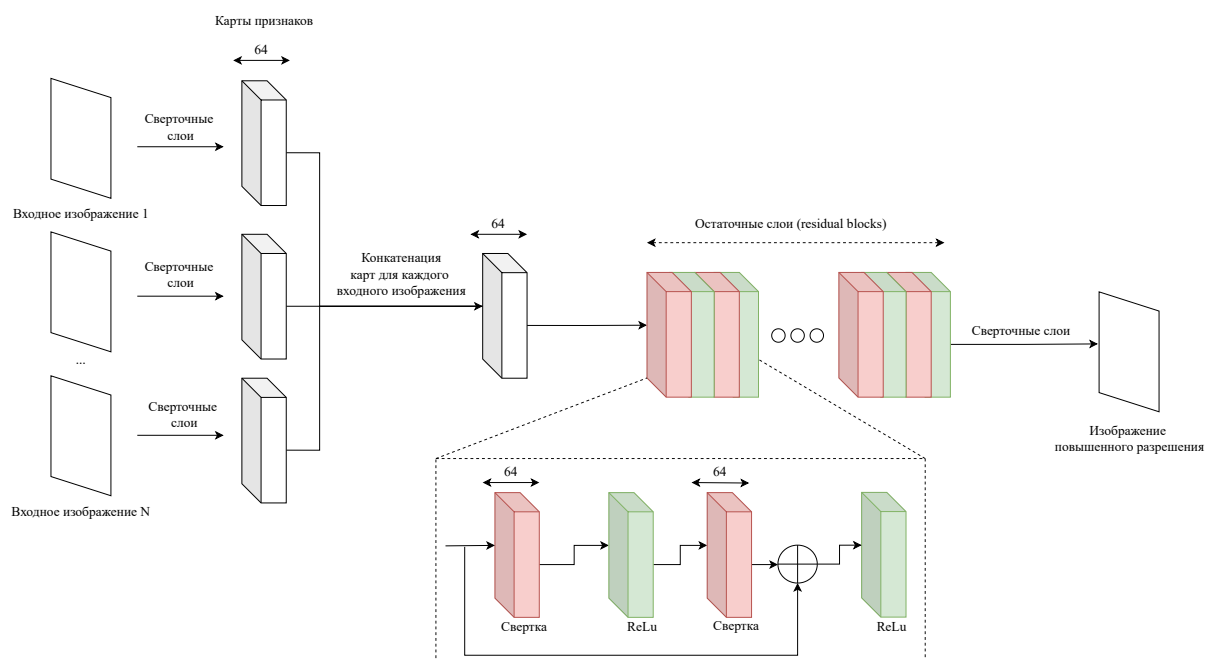


Рисунок 1.8 – Базовая архитектура разрабатываемой нейронной сети для задачи суперразрешения

В настоящее время для обучения нейронных сетей, в том числе глубоких, используется алгоритм обратного распространения ошибки (англ. error backpropagation algorithm), основанный на методе градиентного спуска.

Сверточная нейронная сеть не имеет обратных связей и является многослойной [15].

В глубокой нейронной сети с несколькими скрытыми слоями производится расчет ошибки, которая передается от одного слоя к другому. На первом этапе рассчитывается значение ошибки на выходе нейронной сети, для которого известны правильные ответы. Затем рассчитывается ошибка на входе в выходной слой сети, которая будет использоваться как ошибка на выходе скрытого слоя [16; 17].

Для решения проблемы затухающего градиента во время обучения используется подход использования остаточных блоков (англ. residual blocks [18]): входной сигнал пропускается не только через обычные слои, но и напрямую добавляется к выходу этих слоёв, обходя один или несколько слоёв, т.о. создавая короткий путь для градиентов во время обратного распространения.

### 1.3.7 Структура разрабатываемого программного обеспечения

На рисунке 1.9 представлена диаграмма компонентов [19] разрабатываемого программного обеспечения.

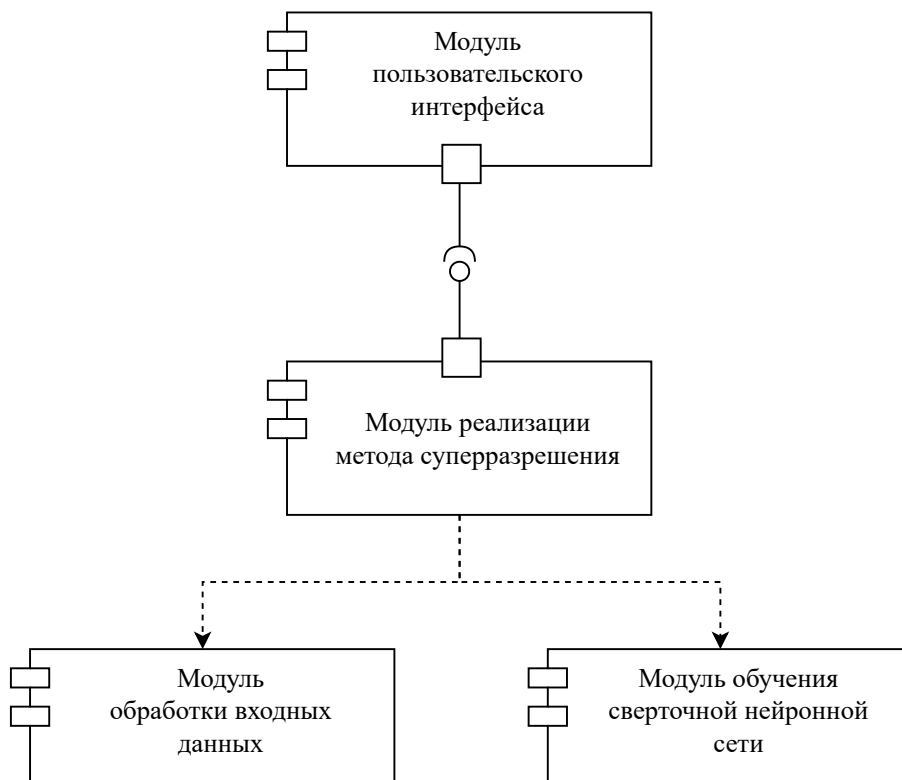


Рисунок 1.9 – Диаграмма компонентов разрабатываемого программного обеспечения

## 1.4 Методы оценки качества восстановления цифрового изображения

Существует несколько методов оценки качества восстановленного изображения [20; 21]:

- *Среднеквадратичная ошибка* (англ. MSE — Mean Square Error) — наиболее широко используемый, а также самый простой полный эталонный показатель, который рассчитывается как квадрат разности интенсивностей пикселей искаженного и эталонного изображения:

$$\text{MSE}(f, g) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f_{ij} - g_{ij})^2, \text{ где } f \text{ и } g \text{ — целевое и восстановленное изображения, } M, N \text{ — размер изображения.}$$

Встречается также вариация метрики — корень из MSE (англ. RMSE Root Mean Square Error).

- *Пиковое отношение сигнал – шум* (англ. PSNR — Peak Signal-to-Noise Ratio) — скалярная метрика, обозначающая соотношение между максимальной мощностью сигнала и шума. Для вычисления PSNR чаще всего используется среднеквадратичная ошибка (англ. Mean Square Error). Метрика выражается в логарифмической шкале и измеряется в децибеллах:

$$\text{PSNR}(f, g) = 10 \cdot \log_{10} \frac{255^2}{MSE(f, g)}.$$

Является основным количественным критерием для оценки эффективности восстановления сигналов в виду, однако имеет ряд ограничений, основным из которых является тот факт, что PSNR (и MSE в частности) слабо коррелируют с физиологией человеческого восприятия.

- *Индекс SSIM* (англ. Structure Similarity Index) характеризует структурное сходство между оригинальным и восстановленным сигналами, принимает значение в диапазоне от -1 до 1, где 1 соответствует совпадению оригинального сигнала с восстановленным, а -1 — полному различию сигналов. Метрика учитывает яркость, контрастность и структуру изображения.

Метрика может быть вычислена согласно следующей формуле:  $\text{SSIM}(f, g) = [l(f, g)]^\alpha \cdot [c(f, g)]^\beta \cdot [s(f, g)]^\gamma$ , где  $l$  — яркость (используется для сравнения яркости между двумя изображениями),  $c$  — контраст (используется для различия диапазонов между самой яркой и самой темной областью двух изображений),  $s$  — функция сравнения структур, которая измеряет коэффициент корреляции между двумя изображениями, а  $\alpha, \beta, \gamma$  — положительные константы.

- *Метрика оценки восприятия человека* (англ. HPM — Human Perception Metric) характеризует качество восстановленного сигнала на основе психофизиологических особенностей восприятия сигналов человеком. Является наиболее трудозатратной, т.к. требует участие людей в проведении эксперимента и анализе данных, однако является наиболее достоверной.

## ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена задача разработки метода повышения разрешения изображения по нескольким кадрам разной интенсивности.

Для достижения поставленной цели была формализована постановка задачи в виде диаграммы IDEF0 верхнего уровня, описаны основные этапы разрабатываемого метода в виде детализированной диаграммы IDEF0 и схем алгоритмов. Была спроектирована структура программного обеспечения для реализации разрабатываемого метода.

Для оценки качества полученного результата предварительно были выбраны метрики MSE, PSNR, SSIM (и модификации), NPM.

Также были сформулированы требования и ограничения к разрабатываемому методу, требования к разрабатываемому программному обеспечению. Изложены особенности предлагаемого метода.

На основе проведенной работы была сформулирована первичная архитектура нейронной сети, решающей задачу суперразрешения. В рамках дальнейшего развития предполагается реализация предложенного метода и сопутствующая корректировка архитектуры сети, а также выработка подхода к обучению для разрабатываемой архитектуры в соответствии с полученными практическими результатами.



## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Park S. C., Park M. K., Kang M. G. Super-resolution image reconstruction: a technical overview //IEEE signal processing magazine. – 2003. – Т. 20. – №. 3. – С. 21-36.
2. Wang Q., Tang X., Shum H. Patch based blind image super resolution //Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. – IEEE, 2005. – Т. 1. – С. 709-716.
3. Freeman W. T., Jones T. R., Pasztor E. C. Example-based super-resolution //IEEE Computer graphics and Applications. – 2002. – Т. 22. – №. 2. – С. 56-65.
4. Katsaggelos A. K. Digital image restoration. – Springer Publishing Company, Incorporated, 2012.
5. Elad M., Feuer A. Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images //IEEE transactions on image processing. – 1997. – Т. 6. – №. 12. – С. 1646-1658.
6. Fan C. et al. POCS Super-resolution sequence image reconstruction based on improvement approach of Keren registration method //Sixth International Conference on Intelligent Systems Design and Applications. – IEEE, 2006. – Т. 2. – С. 333-337.
7. Бредихин А. И. Алгоритмы обучения сверточных нейронных сетей //Вестник Югорского государственного университета. – 2019. – №. 1 (52). – С. 41-54.
8. Senov A. Projective approximation based quasi-Newton methods// In: Proc. of International Workshop on Machine Learning, Optimization, and Big Data. 2017. P. 29–40.
9. Сенов А. А. Глубокое обучение в задаче реконструкции суперразрешения изображений //Стохастическая оптимизация в информатике. – 2017. – Т. 13. – №. 2. – С. 38-57.

10. Воропаева Н. В. и др. Дискретное преобразование Фурье в обработке сигналов // Самара: Изд-во "Самар. ун. – 2015. – Т. 2015.
11. Convolution [Электронный ресурс]. Режим доступа: <https://paperswithcode.com/method/convolution> (дата обращения 22.05.2023).
12. Прокопеня А. С., Азаров И. С. Сверточные нейронные сети для распознавания изображений. – 2020.
13. Маршалко Д. А., Кубанских О. В. Архитектура свёрточных нейронных сетей // Ученые записки Брянского государственного университета. – 2019. – №. 4 (16). – С. 10-13.
14. Sun Y. et al. Convolutional neural network based models for improving super-resolution imaging // Ieee Access. – 2019. – Т. 7. – С. 43042-43051.
15. Youm G. Y., Bae S. H., Kim M. Image super-resolution based on convolution neural networks using multi-channel input // 2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). – IEEE, 2016. – С. 1-5.
16. Созыкин А. В. Обзор методов обучения глубоких нейронных сетей // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. – 2017. – Т. 6. – №. 3. – С. 28-59.
17. Бредихин А. И. Алгоритмы обучения сверточных нейронных сетей // Вестник Югорского государственного университета. – 2019. – №. 1 (52). – С. 41-54.
18. Lin H. J., Tokuyama Y., Lin Z. J. Residual learning based convolutional neural network for super resolution // image. – 2019. – Т. 11. – С. 14.
19. UML Component Diagrams [Электронный ресурс]. Режим доступа: <https://www.uml-diagrams.org/component-diagrams.html> (дата обращения 25.05.2024).

20. Sara U., Akter M., Uddin M. S. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study //Journal of Computer and Communications. – 2019. – T. 7. – №. 3. – C. 8-18.
21. Hore A., Ziou D. Image quality metrics: PSNR vs. SSIM //2010 20th international conference on pattern recognition. – IEEE, 2010. – C. 2366-2369.