



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

# **Многокадровый метод повышения разрешающей способности изображений посредством сверточных нейронных сетей**

**Студент:** Светличная Алина Алексеевна ИУ7-83Б

**Научный руководитель:** Филиппов Михаил Владимирович

Москва, 2024

# Цель и задачи

**Цель:** разработать многокадровый метод повышения разрешающей способности изображений с использованием сверточных нейронных сетей.

## **Задачи:**

- описать предметную область
- проанализировать существующие методы повышения разрешения
- спроектировать алгоритмы необходимые для подготовки данных, обучения модели и ее использования
- по спроектированным алгоритмам реализовать программное обеспечение
- провести исследование применимости разработанного метода

# Анализ предметной области

*Растровое изображение* – двумерный массив, элементы которого (пиксели) содержат информацию о цвете.

*Полноцветное изображение* – изображение, непосредственно хранящее информацию о цвете на основе компонентов цветовой модели.

*Разрешающая способность* – количество пикселей (точек) на единицу площади изображения.

# Классические методы

*Интерполяция* — процесс определения значений между известными точками данных.

*Ступенчатость* — видимые лестничные ступеньки или резкие перепады яркости вдоль контуров.

*Размытие* — потеря деталей и четкости изображения при сглаживании краев.

*Граничное гало* — светлые или темные области, окружающие объекты или контуры на изображении.

1. **Метод ближайшего соседа** учитывает только один пиксель ближайший к точке интерполяции.
2. **Билинейная интерполяция** рассматривает квадрат 2x2 известных пикселей, окружающих неизвестный.
3. **Бикубическая интерполяция** рассматривает массив из 4x4 окружающих пикселей.
4. **Метод Ланцоша** основан на применении нормированной функции *sinc()*.



# Нейронные методы

**Сверточные нейронные сети** – автоматическое извлечение иерархии признаков из входных данных.

$$2. \quad F_1(y) = \max(0, W_1 * y + B_1)$$

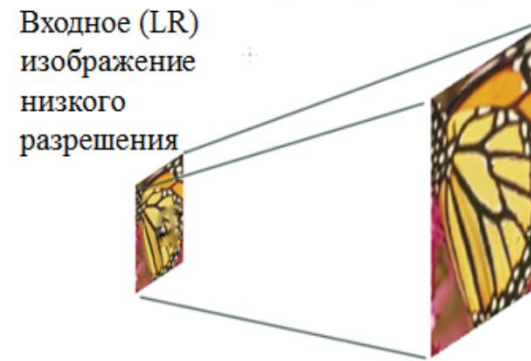
$$3. \quad F_2(y) = \max(0, W_2 * F_1(y) + B_2)$$

$$4. \quad F(y) = W_3 * F_2(y) + B_3$$

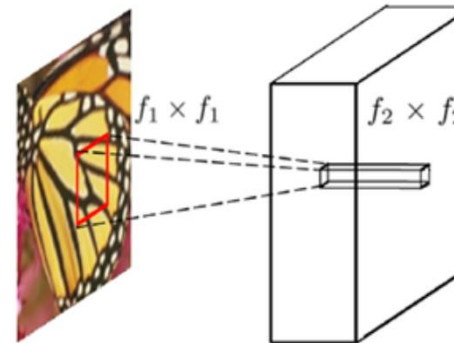
$\max$  – функция активации ReLU,  $*$  – операция свертки  
 $W$  – фильтр,  $B$  – смещение

**Генеративно-сопоставительные нейронные сети** – система из двух нейронных сетей, выполняющих роли генератора и дискриминатора.

1 Увеличение размера изображения



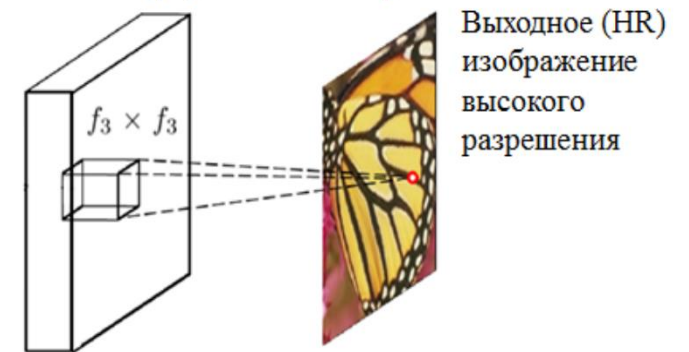
2 Выделение патчей и их представление



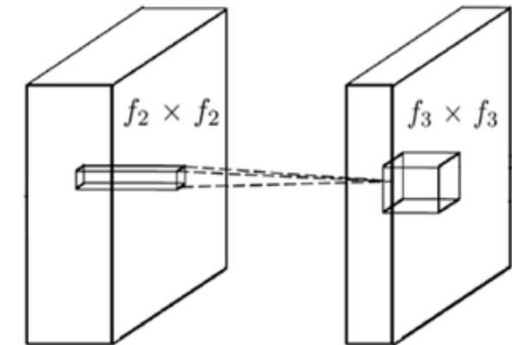
Вектор сигналов изображения LR



4 Генерирование изображения



3 Нелинейное преобразование



Вектор сигналов изображения HR

# Классификация методов

## Классические методы:

- $K_1$  – ступенчатость
- $K_2$  – размытие
- $K_3$  – граничное гало
- $K_4$  – вычислительная сложность

## Нейронные методы:

- $K_1$  – время обучения
- $K_2$  – вычислительная сложность
- $K_3$  – память
- $K_4$  – сложность реализации
- $K_5$  – пиковое отношение сигнала к шуму
- $K_6$  – индекс структурного сходства

## Классические и нейронные типы методов:

- $K_1$  – вычислительная сложность
- $K_2$  – сложность реализации
- $K_3$  – качество

Метод	$K_1$	$K_2$	$K_3$	$K_4$
Ближайшего соседа	4	1	1.5	1
Билинейный	3	3	1.5	2
Бикубический	1	4	3.5	3
Ланцоша	2	2	3.5	4

Метод	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$K_6$
Сверточный	1	1	1	1	1	2
Генеративно-состязательный	2	2	2	2	2	1

Методы	$K_1$	$K_2$	$K_3$
Классические	1	1	2
Нейронные	2	2	1

# Наборы данных

Набор данных	Объем набора в шт.	Разнообразие категорий	Вариативность одного изображения
Set5	5	Различные категории	Один экземпляр
Set14	14	Различные категории	Один экземпляр
BSD100	100	Различные категории	Один экземпляр
Urban100	100	Архитектура	Один экземпляр
DIV2K	800	Различные категории	Несколько экземпляров
Manga109	109 томов	Манга	Один экземпляр

**Тренировочный набор:** 100 изображений различных категорий обработанных собственным алгоритмом понижения разрешения

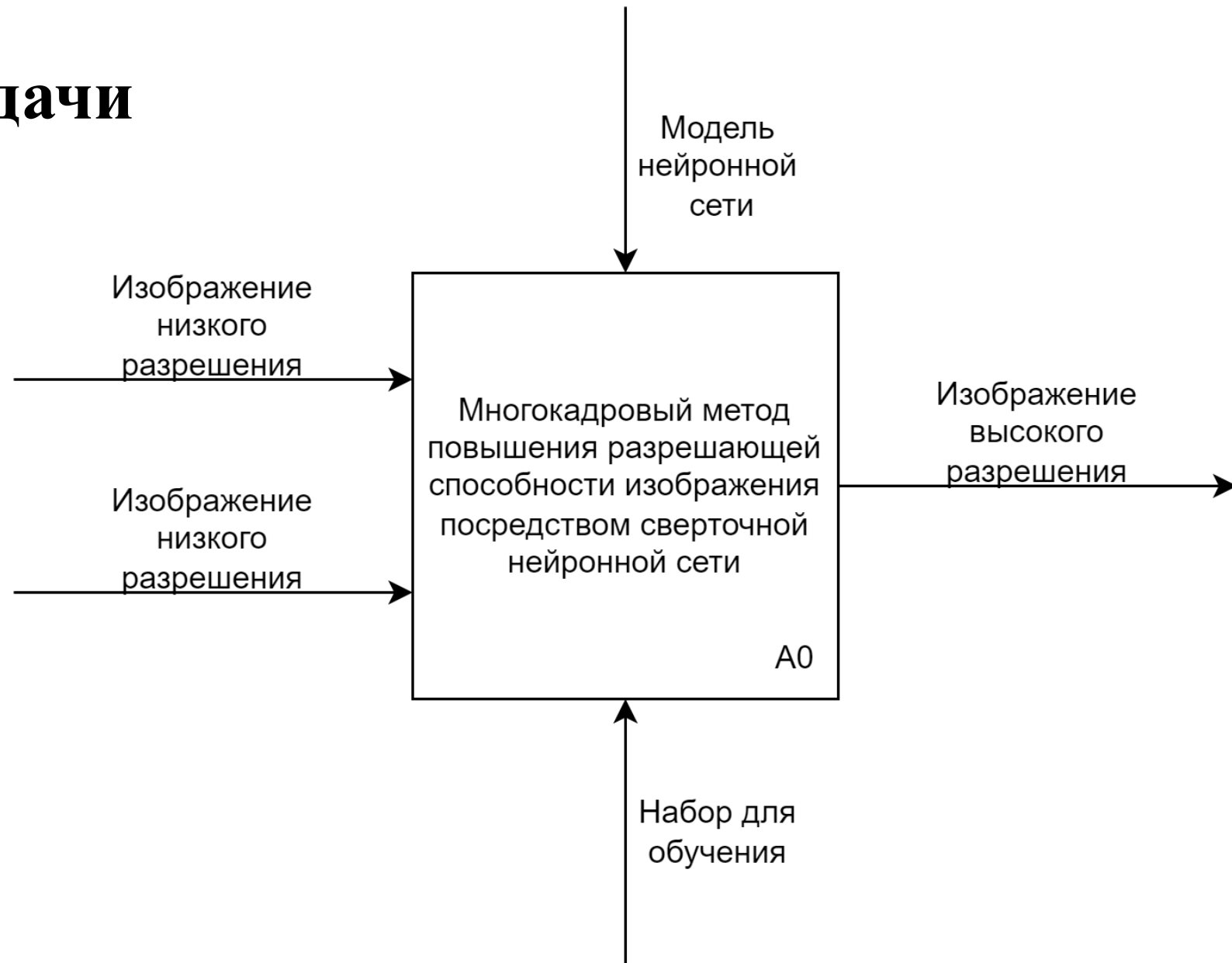
**Валидационный набор:** Set14

**Тестовый набор:** Set5

# Формализация задачи

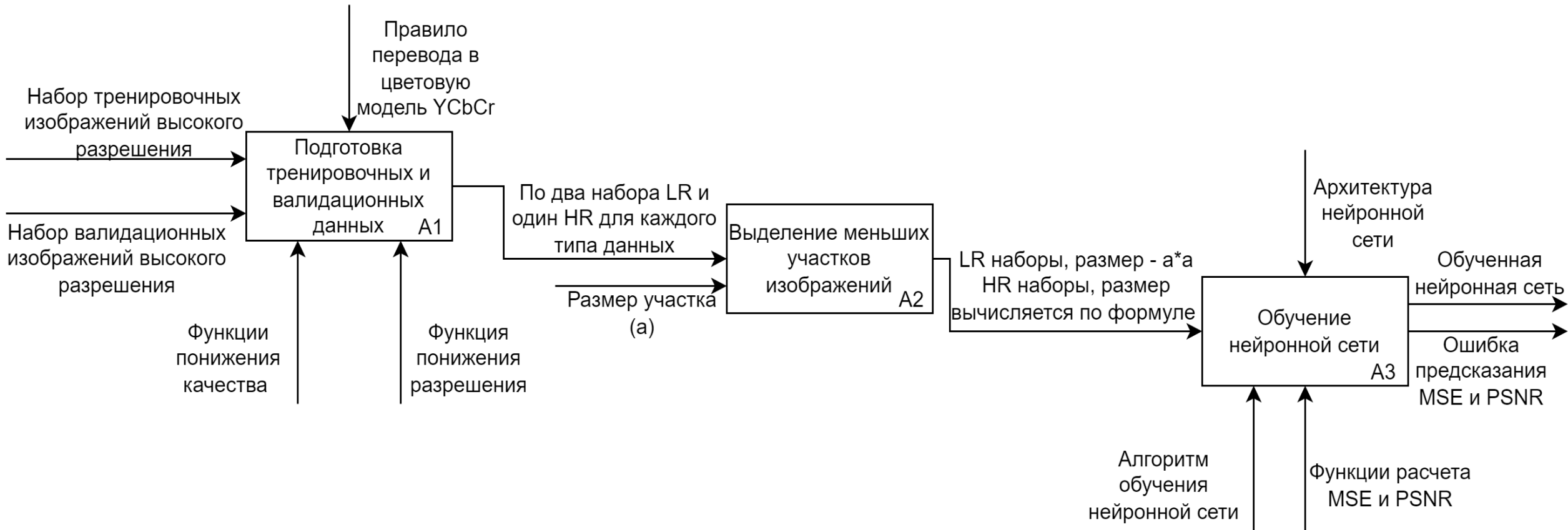
Разрабатываемый метод должен:

- получать изображения низкого разрешения
- подготавливать входные изображения
- иметь возможность обучать модель
- корректно использовать обученную модель для получения изображений высокого разрешения

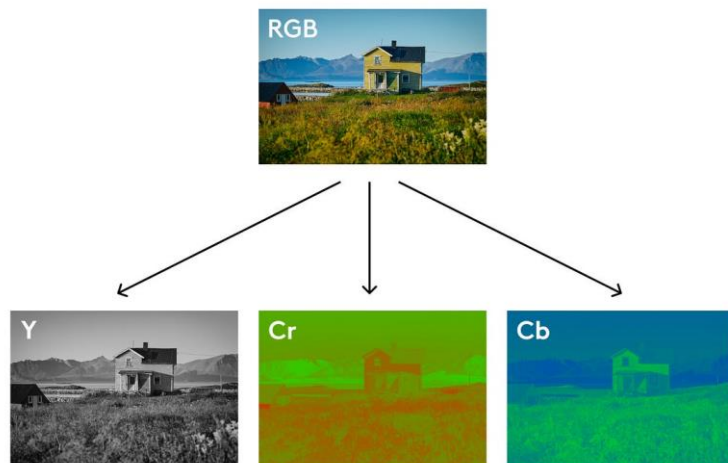




# Метод повышения разрешения



# Алгоритм предобработки изображений



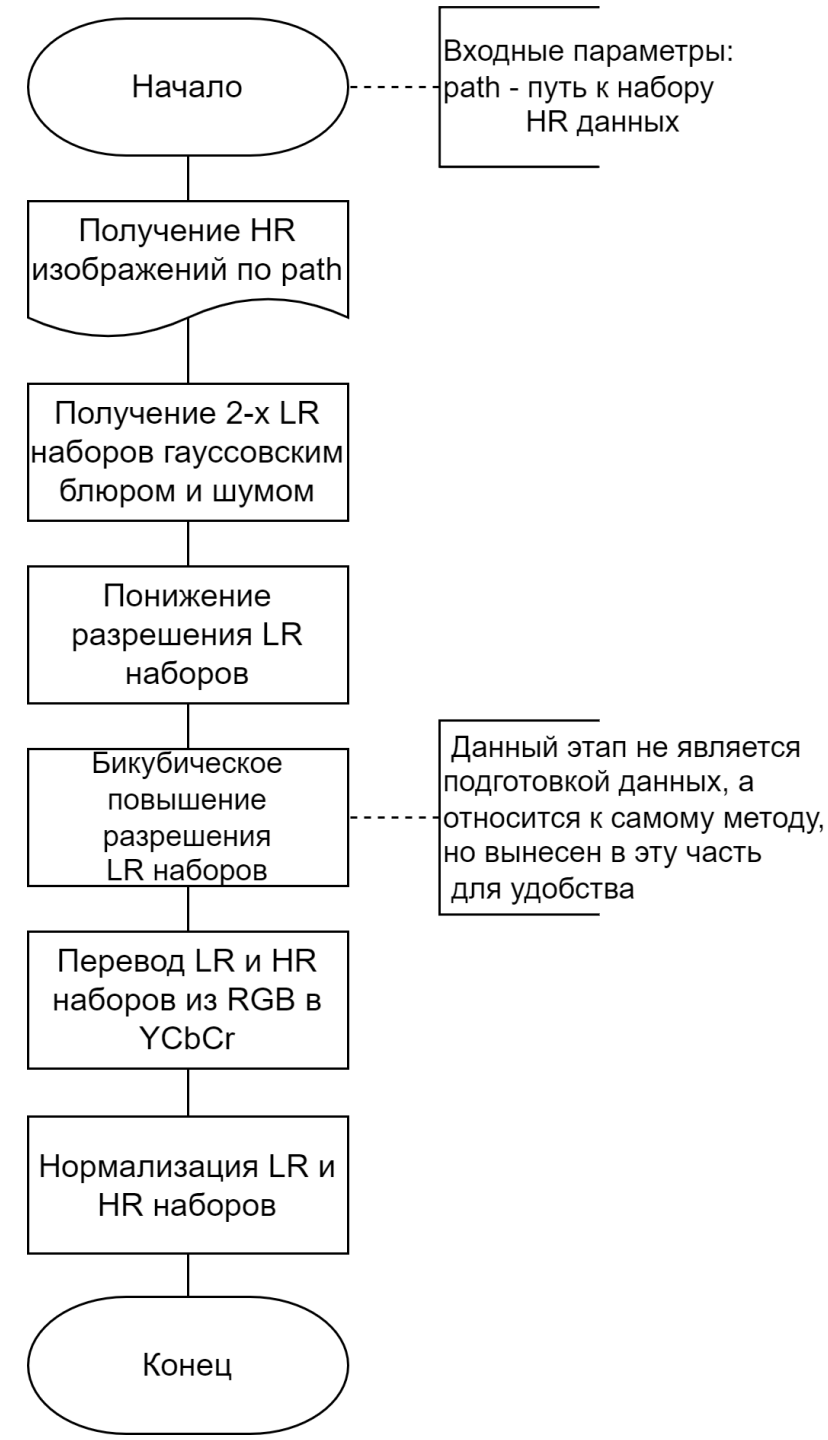
Y – яркость пикселя [16, 235]

Cb – интенсивность синего по отношению к зеленому [16, 240]

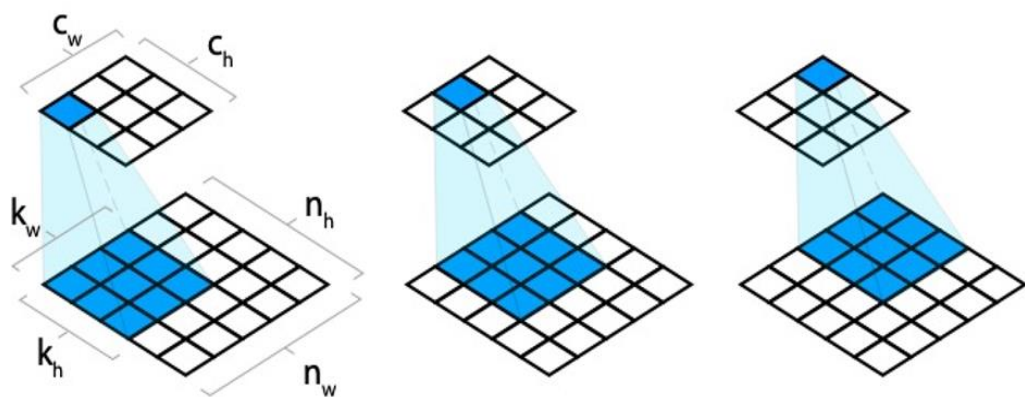
Cr – интенсивность красного по отношению к зеленому [16, 240]

*Нормализация* – процесс приведения значений пикселей к определенному диапазону или распределению, в данном случае  $[0, 1]$  за счет деления на 255

*Денормализация* – обратный процесс, в данном случае  $[0, 255]$  за счет умножения на 255



# Алгоритм разделения изображений на участки

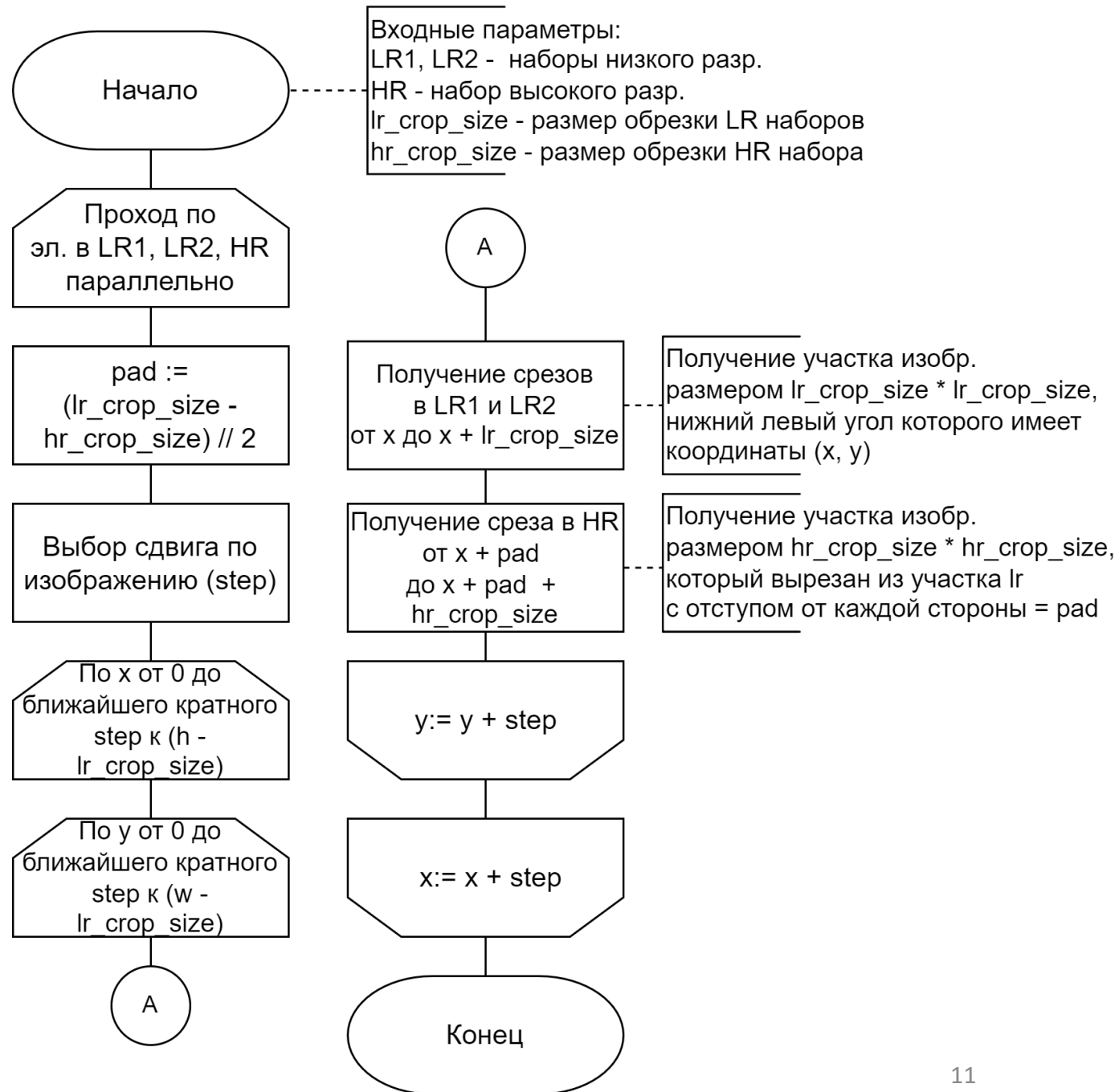


Размерность матрицы выходного слоя:

$$c_h = n_h - k_h + 1$$

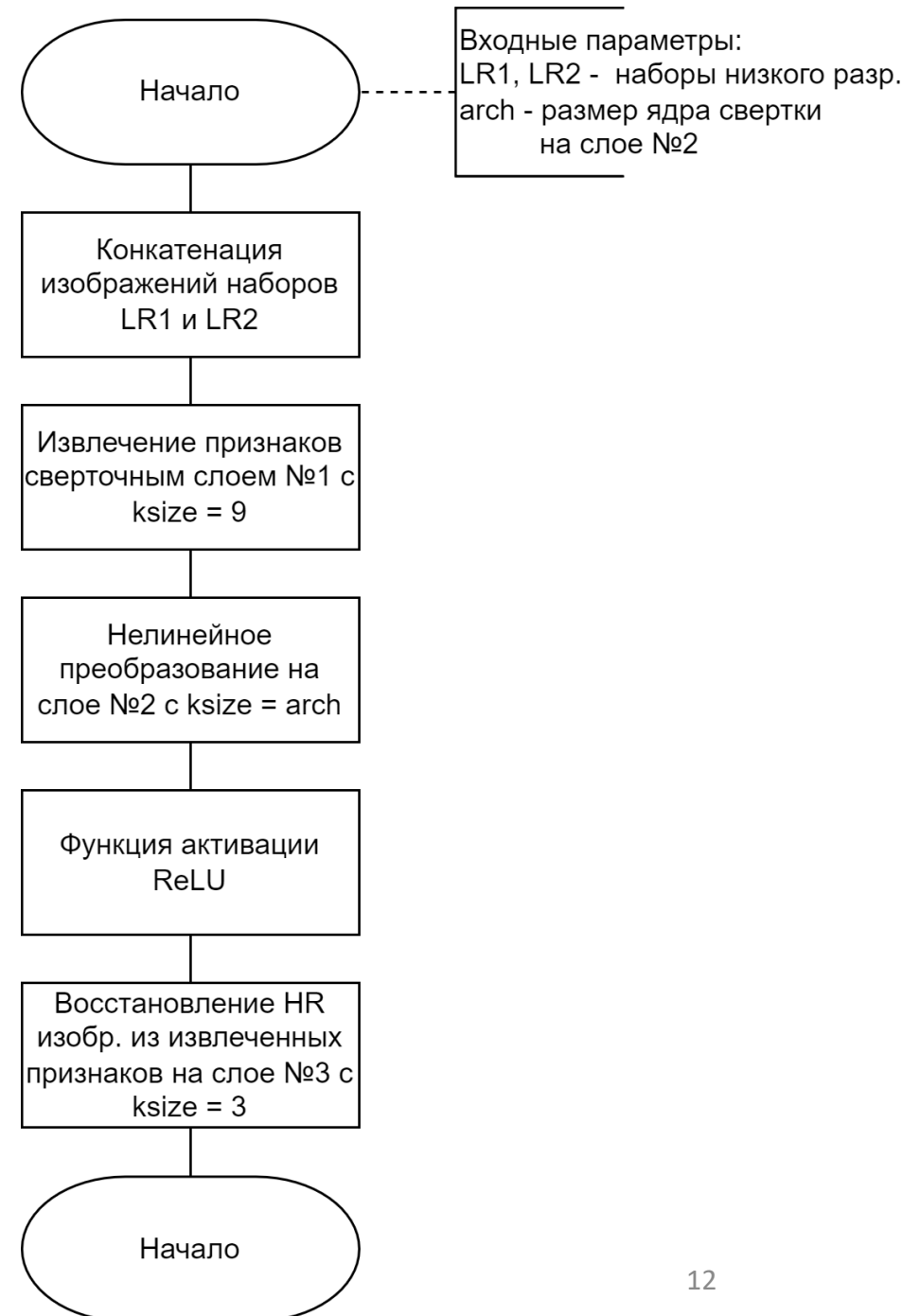
Размерность выходного изображения:

$$h' = h - \sum_{i=1}^n k_{h_i} + n$$

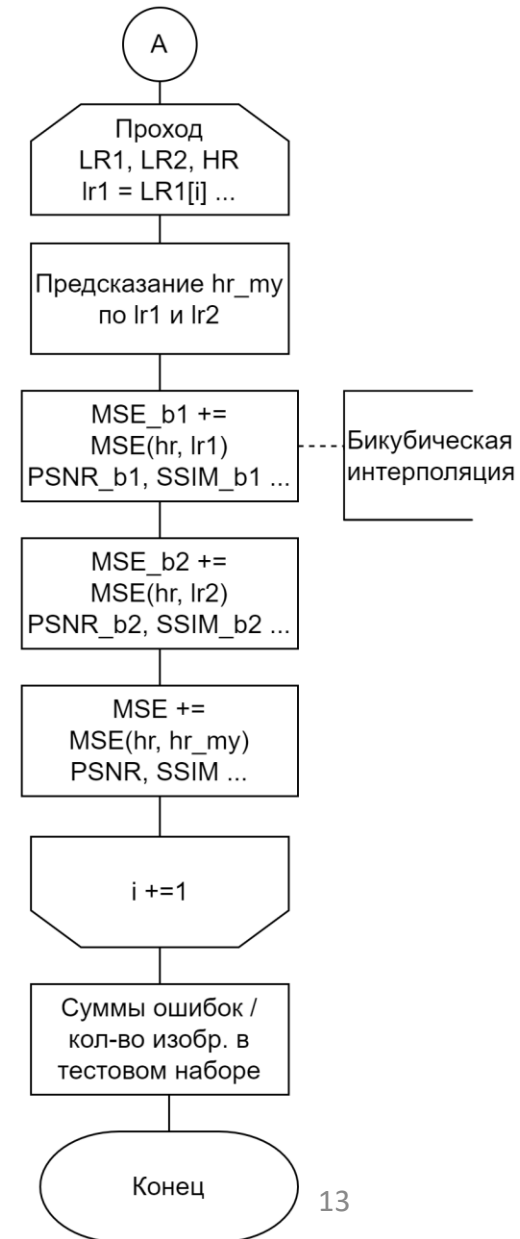
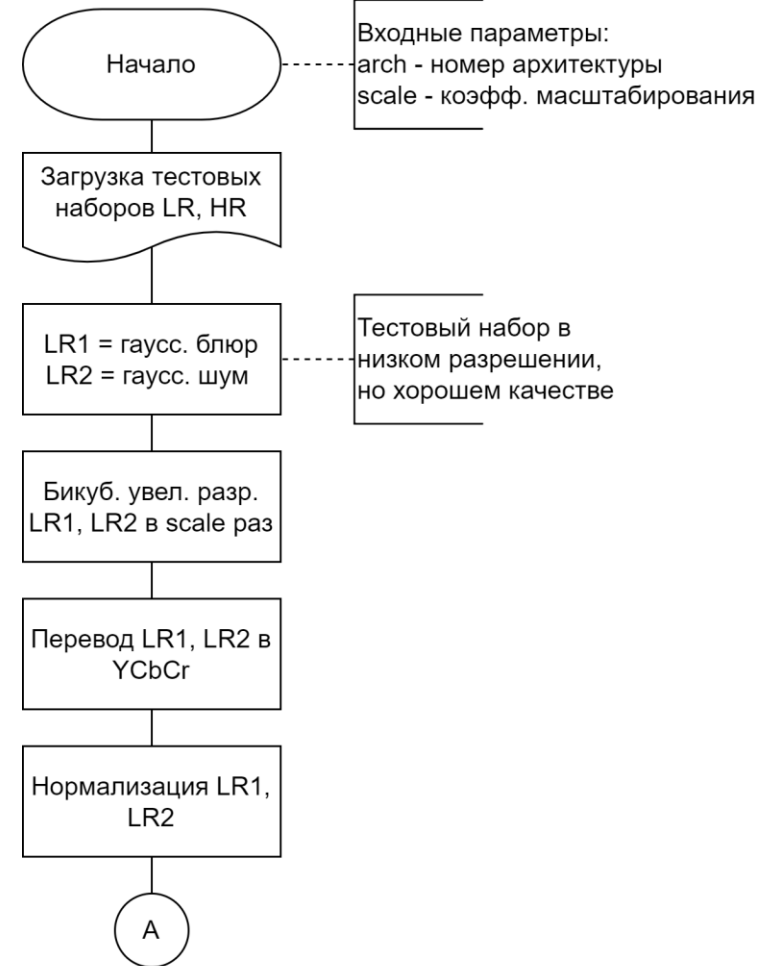
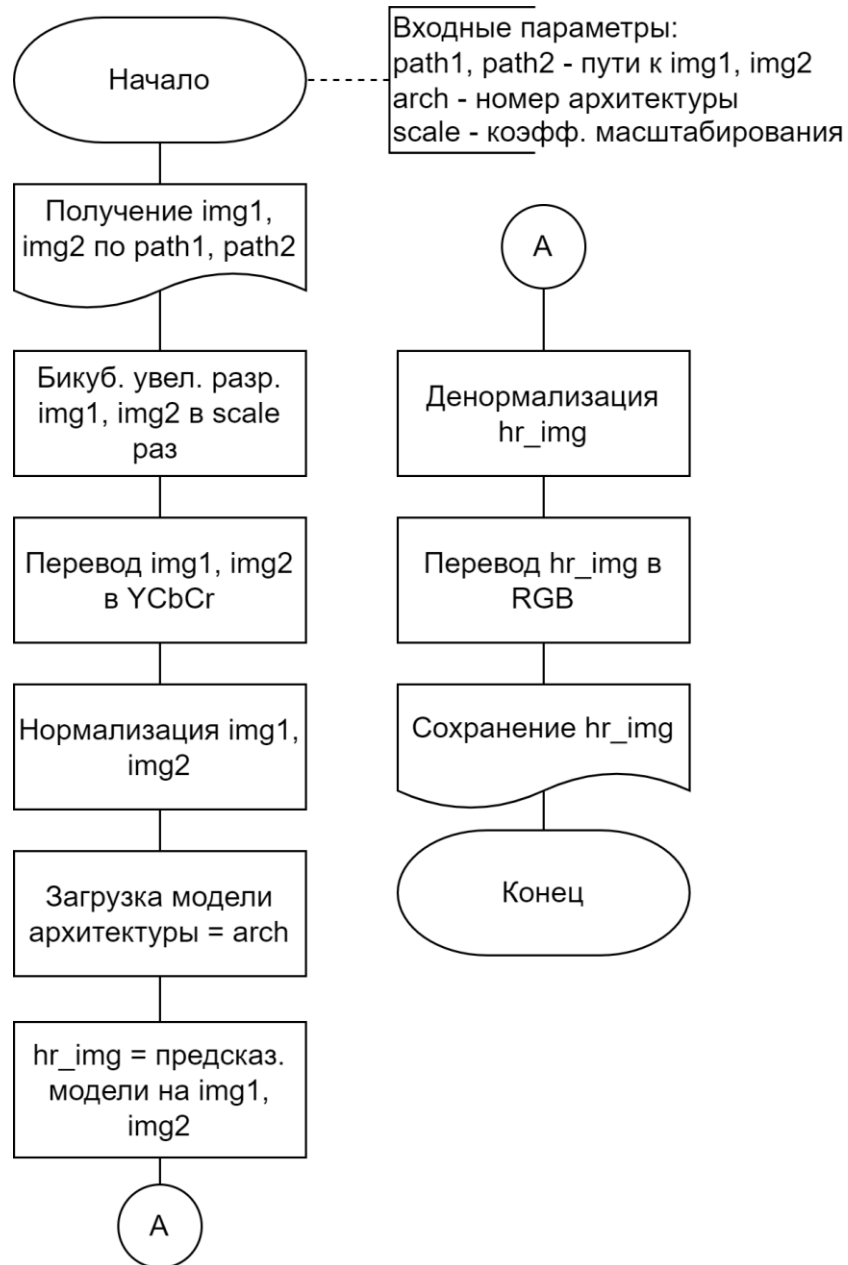


# Архитектура нейронной сети

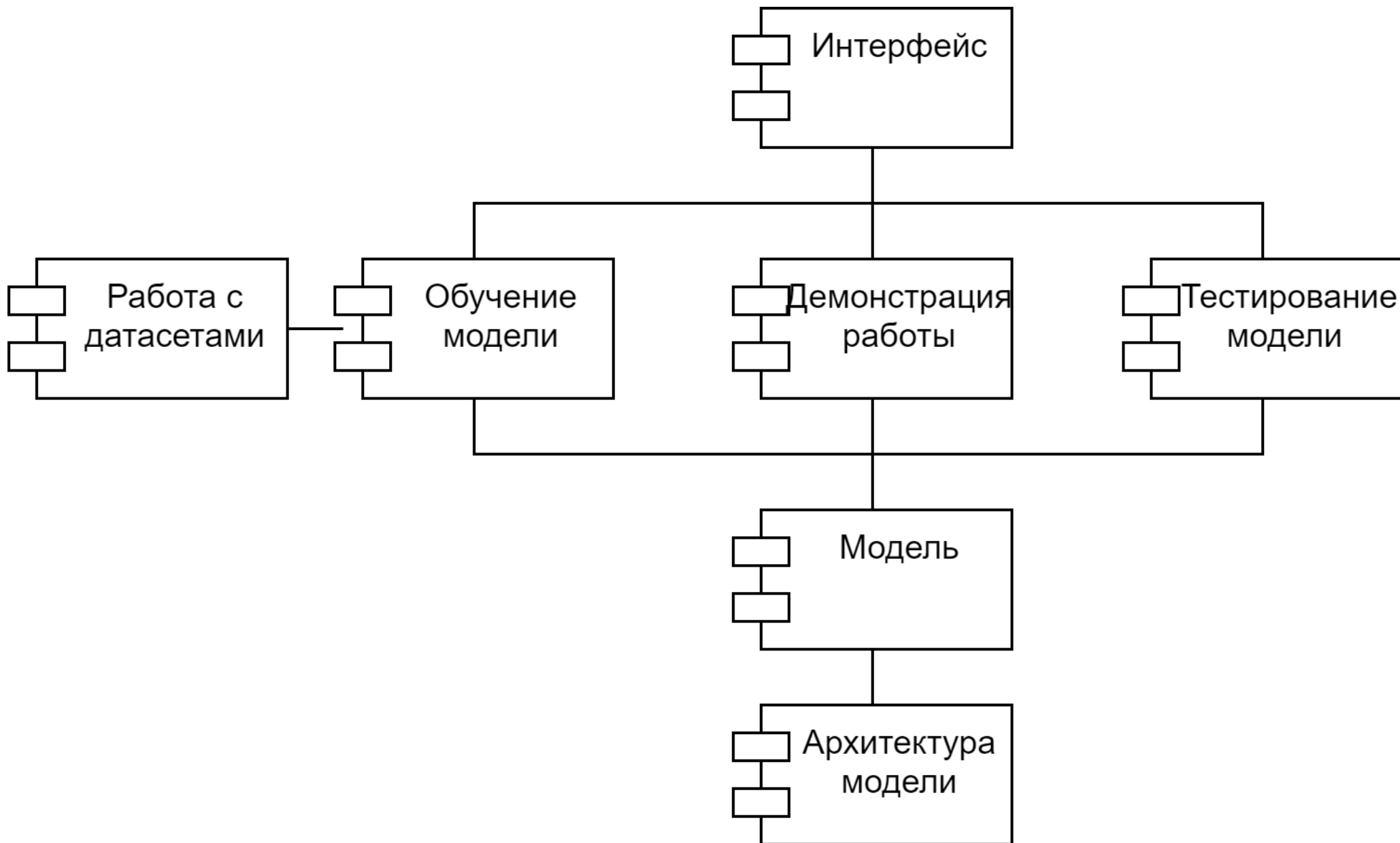
Конкатенация входных изображений будет такой, словно изображение имеет 6 каналов цвета. В модели RGB порядок каналов будет RGBRGB, а в модели YCbCr – YCbCrYCbCr.



# Алгоритмы применения и тестирования модели



# Структура программного обеспечения



# Средства реализации

Язык программирования – **Python**:

- большое количества библиотек для работы непосредственно с нейронными сетями
- широкий выбор вспомогательных библиотек для работы с массивами, изображениями и т. д.
- наличие большого числа как официальной документации, так и дополнительной информации от сообщества разработчиков

Библиотека глубокого обучения – **PyTorch**:

- вычислительный граф определяется во время выполнения программы, поэтому возможно использовать любой инструмент отладки
- легко и естественно интегрируется с другими популярными библиотеками, например, NumPy

# Ограничения данных

1. Входные изображения должны быть в форматах PNG или JPEG.

*PNG* (Portable Network Graphics) – формат, подходящий для текста и графики, так как использует сжатие без потерь.

*JPEG* (Joint Photographic Experts Group) – формат, подходящий для реалистичных изображений с плавными переходами яркости и цвета, так как использует сжатие с потерями.

2. Разрешение входных изображений не должно быть более 1024

# Тестирование программного обеспечения

**Обучение**  
Architecture: ☒ 1 ☐ 3 ☐ 5  
Steps:   
Batch\_size:   
Save\_every:   

Запуск

**Демонстрация**  
Architecture: ☒ 1 ☐ 3 ☐ 5  
Image\_path1:   
Image\_path2:   
Scale: ☒ 2 ☐ 3 ☐ 4  

Запуск

**Тестирование**  
Architecture: ☒ 1 ☐ 3 ☐ 5  
Scale: ☒ 2 ☐ 3 ☐ 4  

Запуск

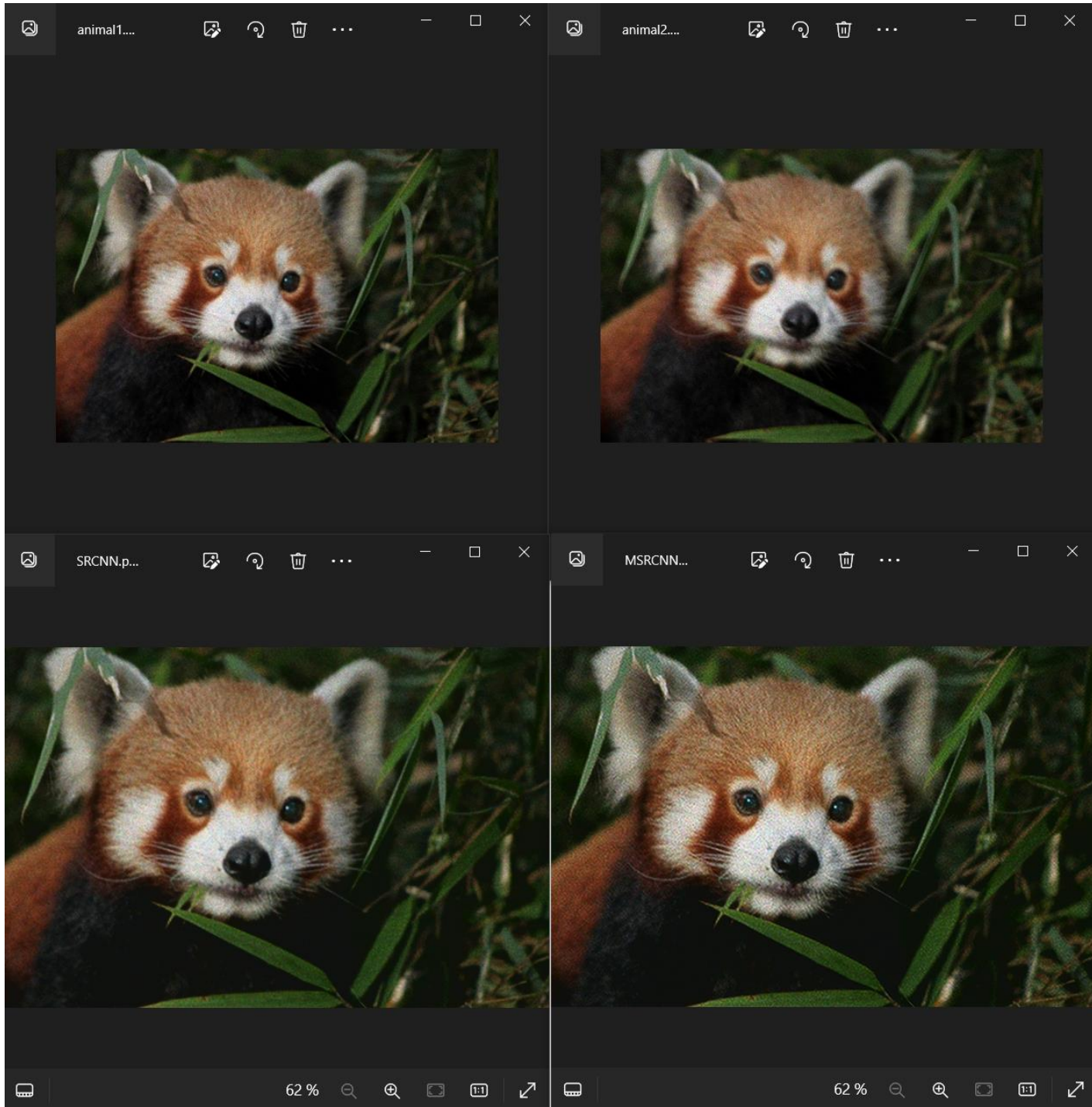
Действие	Описание ситуации	Результат
<b>Training</b>	Запуск действия	«Обучение завершено, модель сохранена по стандартному пути»
<b>Demonstration</b>	Запуск действия с корректными параметрами	Получение фотографий с названиями «bicubic1.png», «bicubic2.png», «hr.png»
Demonstration	В Image_path[№] указан набор символов, не являющийся путем в стандартном понимании	«Image_path[№] не существует»
Demonstration	В Image_path[№] указан не существующий путь	«Image_path[№] не существует»
Demonstration	Image_path[№] ведет к изображению недопустимого расширения	«Формат изображения [№] не поддерживается»
Demonstration	Image_path[№] ведет к изображению разрешением больше 1024	«Изображение [№] не может быть обработано»
Demonstration	Изображения имеют разное разрешение	«Входные изображения имеют разное разрешение»
<b>Testing</b>	Запуск действия	Получение таблицы сравнения MSE, PSNR, SSIM



# Демонстрация работы

Разрешение входных изображений:  
 $510 * 339$

Коэффициент увеличения: 2



# Постановка задачи исследования

Изображения можно классифицировать.

## По качеству:

- хорошего качества
- размытые изображения
- зашумленные изображения
- поврежденные несколькими способами

## По содержимому:

- животные
- люди
- архитектура
- природные объекты
- другое



С использованием Google Форм был составлен опрос, включающий два блока сравнение изображений по качеству и по содержимому.

«По качеству» включают по два изображения на каждый вопрос – полученное SRCNN и многокадровым методом.

«По содержимому» включают по три изображения на вопрос – два получены методом SRCNN, а одно многокадровым SRCNN.

Задающийся вопрос «Какое из представленных выше фото наиболее качественное?».

Вопросы являются выбором с одним вариантом ответа, где кроме номеров фотографий есть еще вариант «Не вижу никакой разницы».

# Результаты исследования

Тип	SRCNN	Многокадр. SRCNN	Нет разницы
Качественное	48,8%	46,3%	4,9%
Размытое	12,2%	9,8%	78%
Зашумленное	75,6%	14,6%	9,8%

Тип	SRCNN (1 фото)	SRCNN (2 фото)	Многокадр. SRCNN	Нет разницы
Животное	41,5%	—	58,5%	—
Архитектура	—	31,7%	68,3%	—
Природа	58,5%	—	41,5%	—
Лицо	70,7%	7,3%	22%	—
Текст	31,7%	—	68,3%	—

## Выводы:

- метод показал себя лучше, если получал на вход два различных изображения
- иногда метод дает излишне зашумленные результаты
- метод плохо справляется с размытием изображений

# Заключение

**Цель достигнута:** разработан многокадровый метод повышения разрешающей способности изображений с использованием сверточных нейронных сетей.

**Задачи решены:**

- описана предметная область
- проанализированы существующие методы повышения разрешения
- спроектированы алгоритмы необходимые для подготовки данных, обучения модели и ее использования
- по спроектированным алгоритмам реализовано программное обеспечение
- проведено исследование применимости разработанного метода

# Дальнейшее развитие

- реализовать возможность использования более двух входных изображений
- спроектировать и реализовать алгоритм совмещения фотографий по объектам
- добавить обработку большего числа входных форматов