



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 ИУ «Информатика и системы управления»

КАФЕДРА ИУ-7 «Программное обеспечение ЭВМ и информационные технологии»

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА  
К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ  
НА ТЕМУ:**

***«Классификация известных методов увеличения  
разрешения видеопотока»***

Студент      ИУ7-73Б      \_\_\_\_\_ Марченко В.

Руководитель НИР      \_\_\_\_\_ Тассов К. Л.

Рекомендуемая руководителем НИР оценка \_\_\_\_\_

## РЕФЕРАТ

Отчет X с., X рис., X табл., X источн., X прил.

ВИДЕО, ВИДЕОПОТОК, ВИДЕОИЗОБРАЖЕНИЕ, РАЗРЕШЕНИЕ, НЕЙРОННЫЕ СЕТИ

Объектом исследования являются методы увеличения разрешения видеопотока.

Цель работы: классификация известных методов увеличения разрешения видеопотока.

В результате исследования было проведено сравнение ... по ... критериям.

Область применения результатов — ...

Результат работы...

# СОДЕРЖАНИЕ

<b>ВВЕДЕНИЕ</b>	<b>6</b>
<b>1 Анализ предметной области</b>	<b>8</b>
1.1 Суперразрешение видеопотока . . . . .	8
1.2 Понижение разрешения . . . . .	9
1.3 Подходы к увеличению разрешения видео . . . . .	9
1.4 Частотная область . . . . .	10
1.5 Пространственная область . . . . .	11
1.5.1 Методы, основанные на интерполяции . . . . .	11
1.5.2 Методы, основанные на регуляризации . . . . .	12
1.6 Методы, основанные на использовании нейронных сетей . . . . .	13
1.6.1 Нейронная сеть, использующая динамические фильтры повышения разрешения без явной компенсации движения	13
1.7 Параметры оценки производительности . . . . .	16
<b>2 Классификация методов увеличения разрешения видеопотока</b>	<b>17</b>
2.1 Критерии оценки методов увеличения разрешения видеопотока	17
2.2 Сравнение методов увеличения разрешения видеопотока . . . . .	17
<b>ЗАКЛЮЧЕНИЕ</b>	<b>18</b>
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>19</b>
<b>ПРИЛОЖЕНИЕ А Презентация</b>	<b>20</b>

# ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

В настоящем отчете о НИР применяют следующие сокращения и обозначения:

VSR	Суперразрешение видео (Video Super-Resolution)
SISR	Суперразрешение фото (Single-Image Super-Resolution)
DFT	Дискретное преобразование Фурье (Discrete Fourier Transform)
DCT	Дискретное косинусное преобразование (Discrete Cosine Transform)
DWT	Дискретное вейвлет-преобразование (Discrete Wavelet Transform)
NEDI	New Edge-Directed Interpolation
GBA	Grouped Bees Algorithm
POCS	Проецирование в выпуклые множества (Projections onto Convex Sets)
IBP	Interval Bound Interpolation
RLS	Рекуррентный метод наименьших квадратов (Recursive Least Squares)
MAP	Оценка апостериорного максимума (Maximum a posteriori Probability)
MLE	Метод максимального правдоподобия (Maximum Likelihood Estimation)
MRF	Марковское случайное поле (Markov Random Field)
SSIM	Индекс структурного сходства (structure similarity)

## ВВЕДЕНИЕ

Суперразрешение — это способ получения изображения или видеоизображения с высоким разрешением из изображений с низким разрешением [1]. В отличие от суперразрешения одного изображения (SISR), основная цель суперразрешения видео — не только восстановить больше мелких деталей при сохранении крупных, но и сохранить согласованность движения.

Во многих областях, работающих с видео, люди имеют дело с различными типами деградации видео, включая понижение разрешения. Разрешение видео может снизиться из-за несовершенства измерительных устройств. Плохое освещение и погодные условия добавляют шум. Движение объектов и камеры также ухудшает качество видео. Методы суперразрешения помогают восстановить исходное видео. Это полезно в широком спектре приложений, таких как [2]:

- 1) видеонаблюдение (для улучшения качества видео, снятого с камеры, а также распознавания номеров автомобилей и лиц);
- 2) медицинская визуализация (чтобы лучше обнаружить некоторые органы или ткани для клинического анализа и медицинского вмешательства);
- 3) судебно-медицинская экспертиза (для помощи в расследовании в ходе уголовного процесса);
- 4) астрономия (для улучшения качества видео звезд и планет);
- 5) дистанционное зондирование (для облегчения наблюдения за объектом);
- 6) микроскопия (для усиления возможностей микроскопов).

Суперразрешение видео также помогает решить задачу обнаружения объектов, распознавания лиц и символов (в качестве этапа предварительной обработки).

Существует множество подходов к решению этой задачи, но она по-прежнему остается популярной и сложной.

Цель научно-исследовательской работы: провести обзор известных методов увеличения разрешения видеопотока и классифицировать их по сформулированным критериям.

Задачи научно-исследовательской работы:

- 1) исследовать предметную область увеличения разрешения видеопотока;
- 2) проанализировать известные методы увеличения разрешения видеопотока;
- 3) сформулировать критерии для сравнения этих методов;
- 4) сравнить методы увеличения разрешения видеопотока по сформулированным критериям.

# 1 Анализ предметной области

## 1.1 Суперразрешение видеопотока

Суперразрешение — это набор действий с целью получения изображения (или последовательности изображений) высокого разрешения из группы изображения низкого разрешения. Концепция суперразрешения представлена на рисунке 1.1. Суперразрешение позволяет получить изображение или видео повышенного качества с большим количеством деталей на сцене, что важно для точного анализа [2].



Рисунок 1.1 – Концепция суперразрешения [2]

Суперразрешение может быть оптическим и геометрическим. В оптических методах используются характеристики оптики, датчиков и компонентов дисплея устройства визуализации, которые отвечают за ухудшение качества или разрешения изображения. Улучшение пространственного разрешения устройства визуализации может быть достигнуто путем модификации аппаратного обеспечения двумя способами [2]: увеличить количество пикселей (но есть ограничения, т. к. это уменьшает отношение сигнал/шум (ОСШ) и увеличивает время получения изображения) и увеличить размер чипа, необходимого для получения изображений высокого разрешения (такие чипы достаточно дорогие) [1].

Хорошей альтернативой обоим подходам является использование метода автономного улучшения разрешения, то есть геометрического суперразрешения. В этом типе суперразрешения для восстановления и реконструкции изображения используются методы цифровой обработки изображений [2].

Благодаря широкой применимости концепции суперразрешения это одна из наиболее быстро развивающихся областей исследований в области обработки изображений [3].

## 1.2 Понижение разрешения

На рисунке 1.2 показан процесс понижения разрешения изображения.

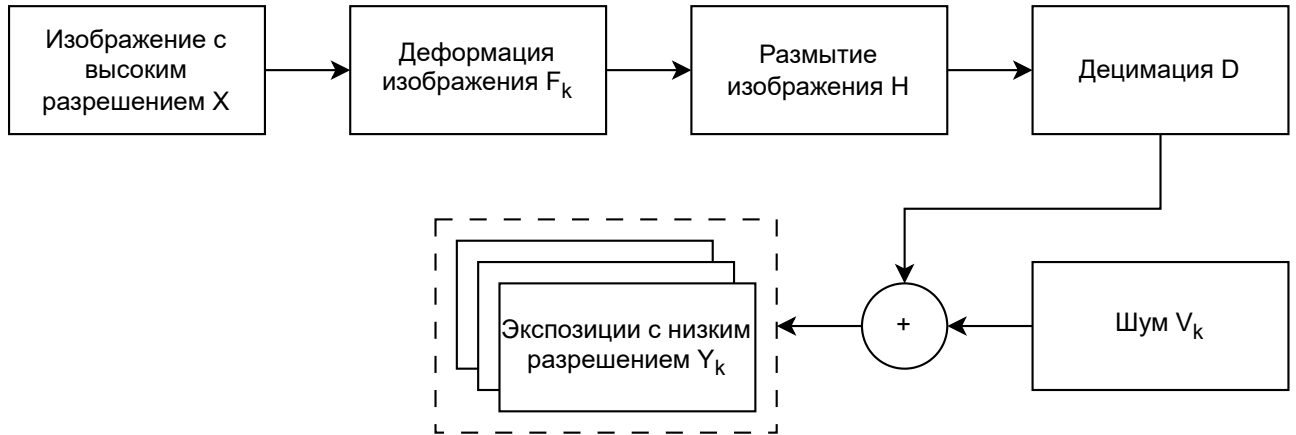


Рисунок 1.2 – Процесс понижения разрешения изображения [2]

Приведенный процесс можно записать с помощью формулы:

$$Y_k = D * H * F_k * X + V_k, \quad (1.1)$$

где  $Y_k$  —  $k$ -я экспозиция сцены с низким разрешением,  $H$  — коэффициент размытия, которое появляется из-за особенностей камеры,  $D$  — коэффициент децимации,  $F_k$  — деформация, а  $V_k$  — коэффициент шума [2].

В приведенной выше формуле факторами деградации являются  $F_k$ ,  $H$ ,  $D$  и  $V_k$ . Если эти коэффициенты известны разработчику, то система называется системой с предварительно известными данными, а изображение с высоким разрешением получается путем решения математического уравнения 1.1 [2].

## 1.3 Подходы к увеличению разрешения видео

Самый простой способ реализовать суперразрешение видео — покадровый запуск суперразрешения фото. Однако, поскольку методы суперразрешения фото не учитывают временные отношения между кадрами, существует высокая вероятность того, что последовательные кадры не будут соединены естественным образом, что приведет к мерцающему артефакту [4].

Суперразрешение осуществляется или покадрово, или используя сразу несколько кадров. Субпиксельный сдвиг между последовательными кадрами используется для восстановления кадров высокого разрешения в многокадровых методах суперразрешения. Однокадровые методы стремятся улучшить



качество изображения без добавления размытия. Алгоритмы суперразрешения работают в двух областях — частотной и пространственной. На рисунке 1.3 представлены некоторые методы суперразрешения видео [2].

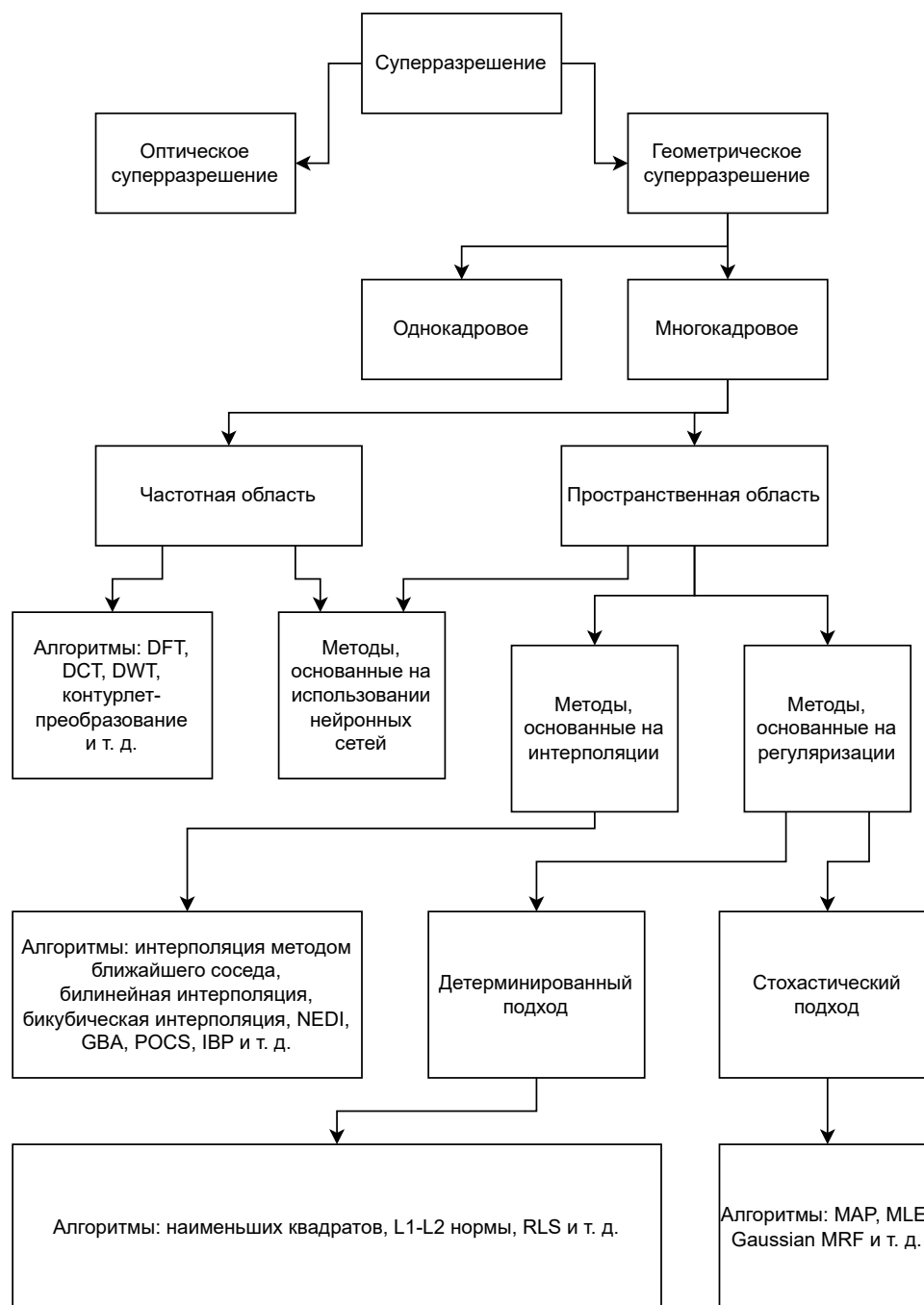


Рисунок 1.3 – Некоторые методы суперразрешения видеопотока [2]

## 1.4 Частотная область

Подходы с частотной областью рассматривают частотную составляющую как признак изображения. Преобразование области сигнала изображения/видео в частотную область осуществляется с помощью дискретного

преобразования Фурье, дискретного косинусного преобразования и дискретного вейвлет-преобразования. Метод частотной области точно использует алиасинг, существующий в каждом изображении низкого разрешения для восстановления изображения высокого разрешения [2].

Подходы с частотной областью базируются на трех принципах [5]:

- 1) свойство временного сдвига преобразования Фурье;
- 2) отношение алиасинга между непрерывным преобразованием Фурье оригинального изображения с высоким разрешением и дискретным преобразованием Фурье изображений низкого разрешения;
- 3) оригинальное изображение высокого разрешения ограничено диапазоном частот.

Вейвлет-преобразование дает частотные компоненты с их временной информацией, которая отвечает за более многообещающие результаты, чем другие преобразования [2].

## **1.5 Пространственная область**

В пространственной области процесс восстановления происходит путем обработки на уровне пикселей вместо работы с каким-либо признаком изображения. Алгоритмы, относящиеся к пространственной области, в основном делятся на алгоритмы, использующие интерполяцию или регуляризационные [2].

Итеративные методы обратного проецирования предполагают некоторую функцию между кадрами с низким и высоким разрешением и пытаются улучшить свою предполагаемую функцию на каждом этапе итеративного процесса [6]. Метод проецирования в выпуклые множества, который определяет конкретную функцию стоимости, также может использоваться для итеративных методов [7].

### **1.5.1 Методы, основанные на интерполяции**

Самый простой способ повысить разрешение изображения — интерполяция. Процесс интерполяции — это оценка нового пикселя с помощью заданного набора пикселей. Регистрация, интерполяция и восстановление —

три основных этапа интерполяционных методов суперразрешения [5]. Геометрическое выравнивание происходит при регистрации изображений, при которой изображения низкого разрешения выравниваются по одному конкретному изображению низкого разрешения, используемому в качестве эталона. Смещения и повороты субпикселей необходимы для точной оценки параметров движения перед их объединением для создания изображения высокого разрешения [2].

Простые и базовые методы интерполяции представляют собой не что иное, как интерполяция методом ближайшего соседа, билинейная интерполяция и бикубическая интерполяция. В этих методах для интерполяции неизвестного пикселя используется либо ближайший пиксель, либо средневзвешенное значение соседних пикселей [2].

В методе интерполяции кубическим B-сплайном большое количество точек соединяются кривой, известной как сплайн. Кубические сплайны рассчитывают весовые коэффициенты сплайнов, которые используются для интерполяции. Метод интерполяции NEDI (New Edge-Directed Interpolation) рассматривает интерполяцию, основанную на геометрической двойственности между ковариацией низкого и высокого разрешения [2]. Метод EGI (Edge-Guided Interpolation) использует классификацию соседних пикселей на два подмножества для оценки недостающего пикселя по отдельности, а для интерполяции берется наиболее подходящая аппроксимация пикселя [8].

### 1.5.2 Методы, основанные на регуляризации

**Детерминированный подход.** Существует двусторонний априорный подход, который основан на сильной регуляризации и минимизации наименьшего абсолютного отклонения для необычных данных и шума. Этот алгоритм полезен для оценки ошибок движения, резких краев изображений и размытия.

**Стохастический подход.** В стохастическом подходе адаптируемый и удобный способ включения априорной информации обеспечивается методом оценки с помощью апостериорного максимума. В одном из методов используется формула Байеса для надежной оценки движения путем включения регуляризаций в одну модель оценки с помощью апостериорного максимума. Существует авторегрессионная модель, которая является одним из стохастических подходов. Метод основан на адаптивной модели авторегрессии,

управляемой суперпикселями. Метод разреженной регрессии используется для разрешения ключевых кадров, которые выбираются автоматически. Этот метод превосходит другие методы как с точки зрения субъективного качества изображения, так и с точки зрения объективного показателя пикового отношения сигнала к шуму. Данный метод требует меньше вычислительного времени и подходит для практических приложений [2].

## **1.6 Методы, основанные на использовании нейронных сетей**

Традиционные методы суперразрешения видео используют несколько кадров низкого разрешения в качестве входных данных и на выходе выдают кадры высокого разрешения, принимая во внимание субпиксельные движения между соседними кадрами низкого разрешения. Все методы суперразрешения видео, основанные на глубоком обучении, работают именно по этому принципу и состоят из двух этапов: оценки движения и процедуры компенсации, за которой следует процесс увеличения разрешения. Одна из проблем этого двухэтапного подхода заключается в том, что результаты во многом зависят от точной оценки движения. Другая потенциальная проблема заключается в том, что выходной кадр высокого разрешения создается путем смешивания значений из нескольких входных кадров низкого разрешения с компенсацией движения через сверточные нейронные сети, что может привести к размытому выходному кадру высокого разрешения [4].

### **1.6.1 Нейронная сеть, использующая динамические фильтры повышения разрешения без явной компенсации движения**

В этом методе вместо явного вычисления и компенсации движения между входными кадрами, информация о движении неявно используется для генерации динамических фильтров увеличения разрешения. С помощью сгенерированных фильтров кадр высокого разрешения напрямую строится путем локальной фильтрации входного центрального кадра. Поскольку этот метод не полагается на явное вычисление движений и не объединяет напрямую значения из нескольких кадров, можно создавать гораздо более четкие и

согласованные по времени видео высокого разрешения [4].

На рисунке 1.4 показан пример масштабирования пикселя  $(3, 3)$  центрального входного кадра  $X_t$  с помощью коэффициента масштабирования  $r = 4$ . Шестнадцать сгенерированных фильтров от  $F_t^{3,3,0,0}$  до  $F_t^{3,3,3,3}$  используются для создания шестнадцати пикселей в области от  $(12, 12)$  до  $(15, 15)$  кадра  $\hat{Y}_t$  высокого разрешения [4].

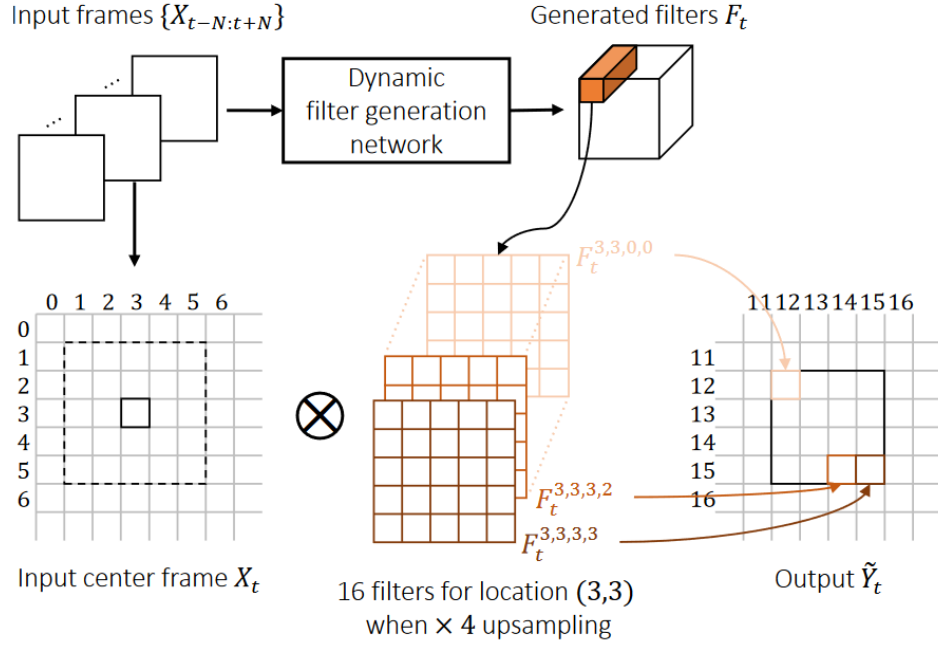


Рисунок 1.4 – Пример масштабирования пикселя [4]

Цель суперразрешения видео — оценить кадры  $\{\hat{Y}_t\}$  высокого разрешения по последовательности кадров  $\{X_t\}$  низкого разрешения. Кадры  $\{X_t\}$  низкого разрешения — это субдискретизированные исходные кадры  $\{Y_t\}$ , где  $t$  — шаг по времени. С предложенной нейронной сетью  $G$  и параметрами сети  $\theta$  задача суперразрешения видео определяется как:

$$\hat{Y}_t = G_{\theta}(X_{t-N:t+N}), \quad (1.2)$$

где  $N$  — временной радиус. Форма входного тензора для  $G$  —  $T \times H \times W \times C$ , где  $T = 2N+1$ ,  $H$  и  $W$  — высота и ширина входного кадра низкого разрешения, а  $C$  — количество цветовых каналов. Соответствующая форма выходного тензора —  $1 \times rH \times rW \times C$ , где  $r$  — коэффициент масштабирования [4].

Нейронная сеть  $G$  на выходе дает два значения для генерации конечного кадра высокого разрешения  $\hat{Y}_t$  из множества кадров низкого разрешения

$\{X_{t-N:t+N}\}$ : динамические фильтры  $F_t$  увеличения разрешения и остаток  $R_t$ . Входной центральный кадр  $X_t$  сначала локально фильтруется с помощью динамических фильтров  $F_t$  увеличения разрешения, а затем остаток  $R_t$  добавляется к результату для окончательного вывода  $\hat{Y}_t$ .

На рисунке 1.5 показана архитектура нейронной сети.

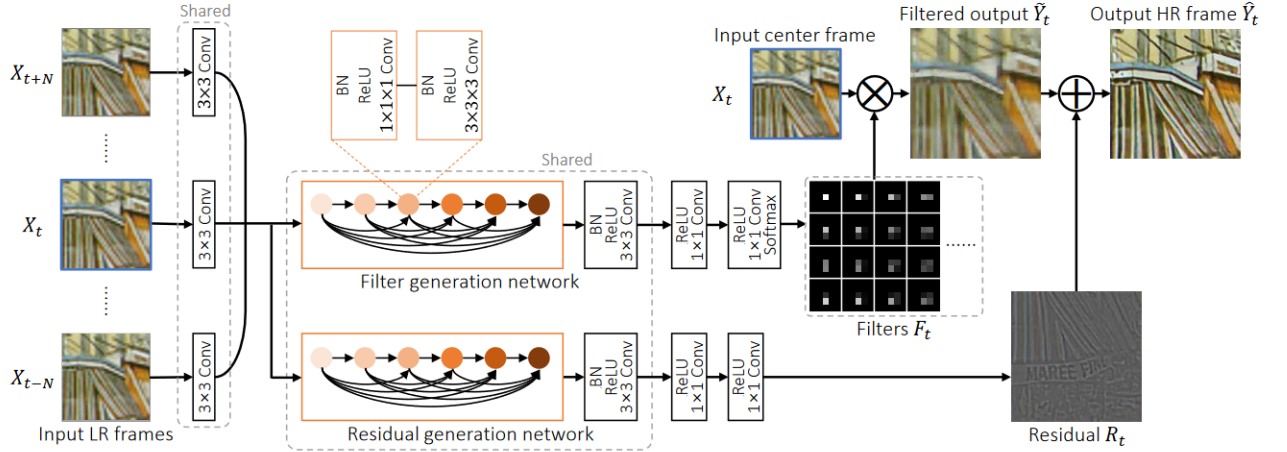


Рисунок 1.5 – Архитектура нейронной сети [4]

Динамические фильтры увеличения разрешения. Сначала множество входных кадров  $\{X_{t-N:t+N}\}$  низкого разрешения попадают на вход сети генерации динамических фильтров. Обученная сеть выдает множество  $r^2HW$  фильтров  $F_t$  увеличения разрешения определенного размера, которые затем используются для генерации новых пикселей отфильтрованного кадра  $\hat{Y}_t$ . Далее создаются выходные пиксели высокого разрешения с помощью локальной фильтрации входного кадра  $X_t$  с помощью соответствующего фильтра:

$$\hat{Y}_t(yr + v, xr + u) = \sum_{j=-2}^2 \sum_{i=-2}^2 F_t^{y,x,v,u}(j+2, i+2) X_t(y+j, x+i), \quad (1.3)$$

где  $y$  и  $x$  — координаты сетки низкого разрешения,  $v$  и  $u$  — координаты каждого выходного блока  $r \times r$  ( $0 \leq v, u \leq r-1$ ). Эта операция аналогична деконволюции, поэтому данную сеть можно обучать сквозным образом, поскольку она допускает обратное распространение ошибки [4].

Добавление остатка. Результату после применения динамических фильтров увеличения разрешения не хватает резкости, поскольку он представляет собой взвешенную сумму входных пикселей. Могут быть детали, которые невозможно восстановить с помощью линейной фильтрации. Чтобы решить

эту проблему, дополнительно оценивается остаточное изображение, чтобы увеличить детализацию [4].

## **1.7 Параметры оценки производительности**

Пиковое отношение сигнала к шуму. Индекс структурного сходства [2].

## **2 Классификация методов увеличения разрешения видеопотока**

### **2.1 Критерии оценки методов увеличения разрешения видеопотока**

### **2.2 Сравнение методов увеличения разрешения видеопотока**



## ЗАКЛЮЧЕНИЕ

В ходе выполнения научно-исследовательской работы была достигнута поставленная цель, а также решены все задачи:

- 1) исследована предметная область увеличения разрешения видеопотока;
- 2) проанализированы известные методы увеличения разрешения видеопотока;
- 3) сформулированы критерии для сравнения этих методов;
- 4) проведено сравнение методов увеличения разрешения видеопотока по сформулированным критериям.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Park S. C., Park M. K., Kang M. G.* Super-resolution image reconstruction: a technical overview // IEEE Signal Processing. — 2003. — С. 21—36.
2. *Mrunmayee D. V., Sachin R. D.* Video Super Resolution: A Review // Department of Electronics Engineering, Walchand College of Engineering, Sangli, Maharashtra, India. — 2021. — С. 6.
3. Image super-resolution: The techniques, applications, and future / L. Yue [и др.] // IEEE Signal Processing. — 2016. — Т. 128. — С. 389—408.
4. Deep Video Super-Resolution Network Using Dynamic Upsampling Filters Without Explicit Motion Compensation / J. Younghyun [и др.] // IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — С. 3224—3232.
5. A performance comparison among different super-resolution techniques / D. Thapa [и др.] // Computers and Electrical Engineering. — 2016. — Т. 54. — С. 313—329.
6. *Cohen B., Avrin V., Dinstein I.* Polyphase back-projection filtering for resolution enhancement of image sequences // 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. — 2000. — С. 2171—2174.
7. *Katsaggelos A. K.* An iterative weighted regularized algorithm for improving the resolution of video sequences // Proceedings of International Conference on Image Processing. — 1997. — С. 474—477.
8. *Zhang L., Wu X.* An edge-guided image interpolation algorithm via directional filtering and data fusion // IEEE Trans Image Process. — 2006. — Т. 15. — С. 2226—2238.

## ПРИЛОЖЕНИЕ А