

## КЛАССИФИКАЦИЯ МЕТОДОВ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ ИЗОБРАЖЕНИЯ ПО НЕСКОЛЬКИМ КАДРАМ

П.Ю. Сироткина

sirotkinapyu@student.bmstu.ru

*МГТУ им. Н.Э. Баумана, Москва, Российская Федерация*

Статья посвящена обзору методов повышения разрешения изображения (так называемой задаче суперразрешения) по нескольким кадрам одного и того же объекта. Кратко рассмотрены предметная область и основные подходы к решению задачи: методы на основе интерполяции, методы реконструкции с регуляризацией, методы на основе теории множеств, примеро-ориентированные методы (выделение патчей), методы оптимизации сверточных нейронных сетей. Проведен критический анализ методов повышения разрешения изображения по нескольким кадрам с определением областей наиболее эффективного использования на основе следующих критериев: вычислительная сложность, качество обработки, необходимость пост- или предобработки. Сделан вывод об области применимости рассмотренных методов. Для достижения наилучшего качества обработки целесообразно применять методы, основанные на сверточных нейронных сетях.

**Ключевые слова:** цифровое изображение, обработка сигналов, суперразрешение, сверточные нейронные сети, интерполяция, реконструкция, регуляризация, теория множеств, патчи

**Введение.** Для многих задач, связанных с областью цифровой обработки сигналов, высокое разрешение изображения является ключевым аспектом, позволяющим проводить более качественный анализ и последующую обработку полученной информации. Изображение высокого разрешения содержит больше сведений о деталях, учет которых может быть критически важен во многих областях, таких как медицина, астрономия и многие другие.

Одним из способов решения рассматриваемой проблемы является использование более качественной техники, однако при таком подходе быстро достигается лимит стоимости, массы и размеров оборудования, что делает этот подход нецелесообразным с точки зрения использования ресурсов. Также понижение качества часто может быть связано с физическими ограничениями (например, при регистрации спутников).

В связи с этим существует широкое разнообразие альтернативных методов повышения разрешения изображения, систематизация которых является актуальной задачей для систем фотосъемки.

**Модель наблюдения изображения низкого разрешения.** В контексте данной работы под термином «кадр» понимается предоставляемое цифровое изоб-

ражение низкого разрешения (НР), используемое для синтеза изображения высокого разрешения (ВР). Каждый кадр является изображением одного и того же объекта, незначительно смещенного относительно других кадров из набора. Модель наблюдения изображения низкого разрешения представлена на рис. 1.



Рис. 1. Модель наблюдения изображения низкого разрешения

Задача повышения разрешения изображения по одному или нескольким кадрам низкого разрешения получила название суперразрешения (*англ.* super resolution). В общем случае эта задача некорректна, поскольку в реальных задачах сведения об искажениях изображения заведомо неизвестны.

Представим желаемое изображение ВР размером  $N = L_1 N_1 \times L_2 N_2$  как вектор  $x = [x_1, x_2, \dots, x_N]^T$ . Параметры  $L_1$  и  $L_2$  представляют собой коэффициенты понижения дискретизации в наблюдении для горизонтального и вертикального направлений соответственно. Таким образом, каждое изображение НР имеет размерность  $M = N_1 \times N_2$ .

Аналогично представим каждое изображение НР в виде вектора  $y_k = [y_{k_1}, y_{k_2}, \dots, y_{k_M}]^T$ , где  $k = \overline{1, p}$ ,  $p$  — размер входного набора.

Для получения изображения высокого разрешения  $x$  методом суперразрешения рассматривается следующая система уравнений [1]:

$$A_k x = y_k,$$

где оператор  $A_k x$  в общем случае представляется в виде выражения:

$$A_k x = D B_k M_l x + n_k.$$

Здесь  $D$  — матрица движения фотосистемы (сдвиг, поворот и т. д.);  $B_k$  — матрица размытия (может быть вызвано оптической системой, относительным движением между фотосистемой и сценой, а также функцией рассеяния в датчике фотосистемы);  $M_l$  — матрица понижения размерности (генерирует изображение НР путем наложения, т. е. алиасинга);  $n_k$  — вектор шума.

Возможна более короткая форма записи:

$$A_k x = W_k x + n_k = y_k,$$

где матрица  $W_k$  размером  $(N_1, N_2)^2 \times L_1 N_1 L_2 N_2$  представляет собой вклад пикселей изображения  $x$  в изображение  $y_k$  посредством размытия, движения и субдискретизации.

**Частотные методы.** Данная группа методов суперрезолюции ориентирована на увеличение разрешения изображения посредством анализа частотных характеристик кадров. Классическим математическим аппаратом для подобных задач является фурье-анализ. Этот подход основывается на следующих принципах:

1) свойство сдвига преобразования Фурье :  $f(\widehat{x - x_0}) = e^{-i\omega x_0} \hat{f}(\omega)$  ; ;

2) взаимосвязь между непрерывным преобразованием Фурье исходного изображения ВР и дискретным преобразованием Фурье наблюдаемых изображений НР;

3) частотный диапазон изображения ограничен.

Пусть  $x(t_1, t_2)$  — непрерывное изображение ВР,  $X(\omega_1, \omega_2)$  — соответствующее ему непрерывное преобразование Фурье,  $y_k[n_1, n_2]$  —  $k$ -е изображение НР,  $Y(\Omega_1, \Omega_2)$  — соответствующее ему дискретное преобразование Фурье. Глобальные сдвиги, которые являются единственным движением, рассматриваемым в подходе частотной группы методов, порождают следующее уравнение для изображений НР:

$$x_k(t_1, t_2) = x(t_1 + \delta_{k_1}, t_2 + \delta_{k_2}),$$

где  $\delta_{k_1}$  и  $\delta_{k_2}$  — субпиксельные сдвиги между кадрами.

Согласно свойству сдвига преобразования Фурье,

$$X_k(\omega_1, \omega_2) = \exp\left[i2\pi(\delta_{k_1} \omega_1 + \delta_{k_2} \omega_2)\right] X(\omega_1, \omega_2).$$

Смещенное изображение  $x_k(t_1, t_2)$  дискретизируется с периодами  $T_1$  и  $T_2$  для создания наблюдаемого изображения НР  $y_k[n_1, n_2]$  [2].

Таким образом, взаимосвязь между  $X_k(\omega_1, \omega_2)$  и  $Y(\Omega_1, \Omega_2)$  можно описать как

$$Y(\Omega_1, \Omega_2) = \frac{1}{T_1 T_2} \sum_{n_1=0}^{L_1-1} \sum_{n_2=0}^{L_2-1} X_k \left[ \frac{2\pi}{T_1} \left( \frac{\Omega_1}{N_1} + n_1 \right), \frac{2\pi}{T_2} \left( \frac{\Omega_2}{N_2} + n_2 \right) \right].$$

Учитывая, что функция  $X$  финитна, матрично-векторная форма уравнения может иметь вид  $Y = \Phi X$ .

**Пространственные методы. Регистрация — интерполяция — восстановление.** Линейный подход к решению задачи суперразрешения на основе интерполяции является наиболее тривиальным, однако в связи с этим имеет ряд существенных недостатков: возникновение алиасинга (эффект «лесенки»), размытия и эффекта Гиббса. Алгоритм суперразрешения на основе рассматриваемого подхода представлен на рис. 2 [3].

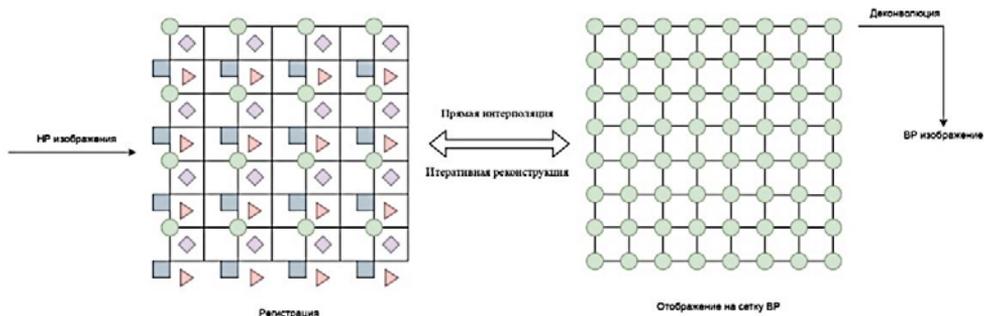


Рис. 2. Алгоритм суперразрешения на основе подхода регистрация — интерполяция — восстановление

**Примеро-ориентированные методы.** Данная группа методов реализовывает подход, основанный на глубоком обучении [4]. Создается набор изображений (изображения ВР подвергают деградации (см. рис. 1) и формируется пара (ВР; НР)), которая затем используется для анализа деталей, соответствующих отдельным областям цифрового изображения. Выявленные взаимосвязи используются для прогнозирования мелких деталей на других изображениях.

Алгоритм суперразрешения реализуется путем разбиения кадра на патчи и составления изображения ВР путем перебора соответствующих патчей в базе данных. Пример сравнения патчей представлен на рис. 3.

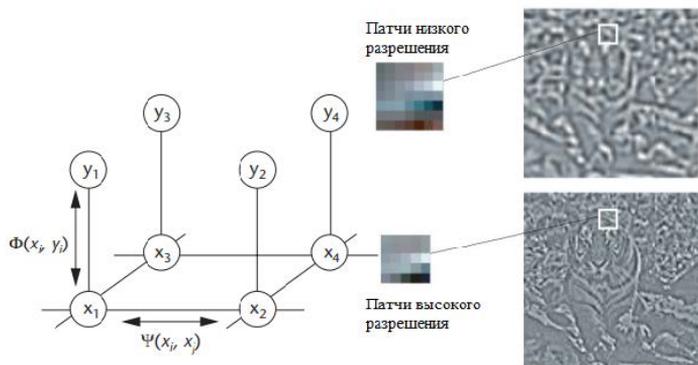


Рис. 3. Пример сравнения патчей

Пример базы данных патчей, предложенный в работе Example-Based Super-Resolution, представлен на рис. 4 [5].

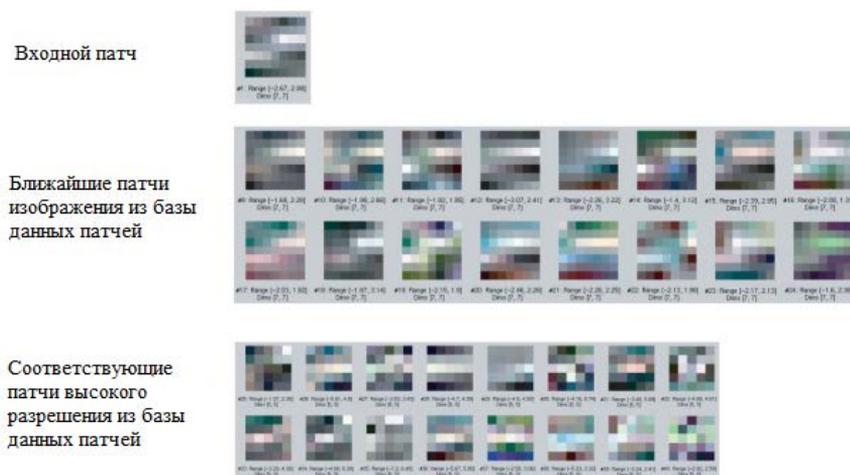


Рис. 4. Пример базы данных патчей

**Реконструкция с регуляризацией.** Задача реконструкции (решение обратной задачи) изображения ВР некорректна ввиду недостаточного количества информации о деталях и отсутствия априорной информации о функции размытия в общем случае. Методы регуляризации сводят некорректно поставленную задачу к корректной посредством введения дополнительных ограничений.

В данном разделе представлены детерминированные и стохастические (статистические) методы реконструкции.

**Детерминированные регуляризационные методы.** Оценив параметры регистрации, можно полностью уточнить модель наблюдения. Рассматриваемый подход позволяет решить обратную задачу на основе априорной информации о решении, которая может быть использована для корректной постановки задачи.

Рассмотрим формулировку метода наименьших квадратов [6], выбрав  $x$  для минимизации Лагранжиана:

$$\sum_{k=1}^p \|y_k - W_k\|^2 + \alpha \|Cx\|^2,$$

где оператор  $C$  — фильтр верхних частот;  $\|\bullet\|$  —  $l_2$ -норма;  $\alpha$  — параметр регуляризации (множитель Лагранжа), контролирующий компромисс меж-

ду точностью данных. Чем больше значение  $\alpha$ , тем более гладким является решение.

Основные методы детерминированной итеративной регуляризации решают следующую задачу:

$$\left[ \sum_{k=1}^p W_k^T W_k + \alpha C^T C \right] \hat{x} = \sum_{k=1}^p W_k^T y_k,$$

что приводит к формулировке следующей рекуррентной формулы для  $\hat{x}$ :

$$\hat{x}^{n+1} = \hat{x}^n + \beta \left[ \sum_{k=1}^p W_k^T (y_k - W_k \hat{x}^n) - \alpha C^T C \hat{x}^n \right],$$

где  $\beta$  — параметр сходимости.

**Стохастические регуляризационные методы.** Данная группа методов основана на байесовской статистике и используется, когда можно установить апостериорную функцию плотности исходного изображения. Оценка  $x$  с помощью метода максимального правдоподобия максимизирует апостериорную функцию плотности  $P(x|y_k)$  относительно  $x$ :

$$x = \arg \max P(x|y_1, y_2, \dots, y_p).$$

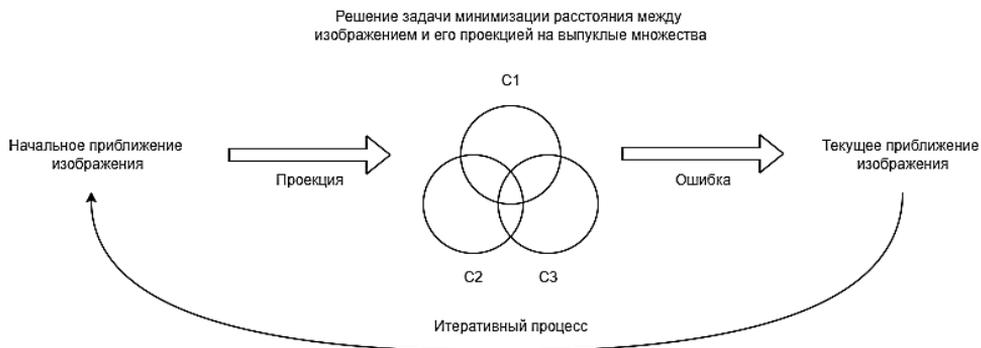
Запишем задачу максимизации, применив логарифмирование, а также теорему Байеса к условной вероятности:

$$x = \arg \max \left\{ \ln P(y_1, y_2, \dots, y_p|x) + \ln P(x) \right\}.$$

Здесь и априорная модель изображения  $P(x)$ , и условная плотность  $P(y_1, y_2, \dots, y_p|x)$  будут определяться априорной информацией об изображении ВР и статической информацией о шуме. Поскольку этот метод учитывает априорные ограничения, предоставляется стабильная оценка изображения ВР.

Главным преимуществом байесовской модели является использование предварительной модели изображения с сохранением краев изображения.

**Методы резольюции на основе теории множеств.** Согласно методу POCS (англ. Projection Onto Convex Sets — проекция на выпуклые множества) [7, 8], учет априорных знаний в решение можно интерпретировать как наложение ограничения на то, чтобы решение было членом замкнутого выпуклого множества  $C_i$ , которое определяется как набор векторов, удовлетворяющих определенному свойству. Ограничение представляет собой предположительные сведения о гладкости, структуре, текстуре и других характеристиках изображения.

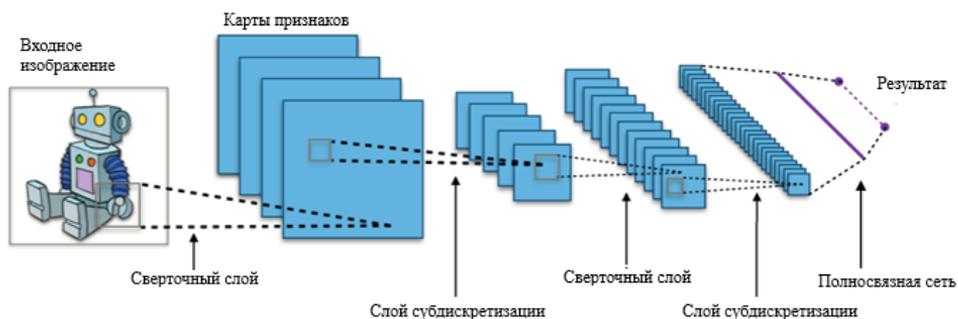


**Рис. 5.** Метод суперразрешения на основе теории множеств

Иллюстрация метода суперразрешения на основе теории множеств представлена на рис. 5.

**Методы на основе сверточных нейронных сетей.** Сверточные нейронные сети (СНС) в области распознавания образов на изображениях и их классификации нашли широкое применение, что объясняется достаточно высоким качеством решения подобных задач. Архитектура сети получила название в виду использования операции свертки.

Типовая архитектура сверточной нейронной сети представлена на рис. 6.



**Рис. 6.** Типовая архитектура СНС

Как можно увидеть, СНС состоит из трех типов слоев [9–11]:

1) слой свертки — основной блок СНС, производящий свертку входной матрицы с ядром свертки. Количество ядер равно количеству карт признаков, выделяемых на изображении;

2) слой пуллинга, или субдескрипция — данный слой принимает результат свертки предыдущего слоя в вид матрицы и сжимает данную матрицу с целью выделения низкоуровневых признаков;

3) полносвязный слой — на данный слой подается одномерный вектор от стоящего перед ним слоя, причем данный вектор получен из матрицы путем записи ее элементов построчно в одну строку.

**Оптимизация глубоких нейронных сетей.** Одна из ключевых задач в тренировке глубоких нейронных сетей — оптимизация их параметров. Вследствие большого совокупного числа параметров на всех слоях нейронной сети применение методов второго порядка (таких, как BFGS, SR1 и других квазиньютоновских методов) крайне затруднительно, а методы нулевого порядка в большинстве случаев не позволяют найти качественное решение за приемлемое время.

По этой причине для оптимизации глубоких и в том числе сверточных нейронных сетей наиболее распространены методы первого порядка, требующие лишь знания градиента сети как функции от ее параметров.

**Проекция — аппроксимация — восстановление.** Исходя из приведенных выше соображений были разработаны методы, основанные на подходе проекции-аппроксимации-восстановления [12, 13].

Под подходом «проекция — аппроксимация — восстановление» понимается следующая процедура (далее обозначаемая акронимом ПАВ):

1) *проекция*: проекция нескольких последних точек из оригинального пространства в пространство меньшей размерности посредством умножения на специальным образом построенную прямоугольную матрицу с ортонормальными строками;

2) *аппроксимация*: построение квадратичного полинома, аппроксимирующего полученные проекции точек и соответствующие им значения целевой функции;

3) *восстановление*: аппроксимация параметров целевой функции (например, Гессiana, точки минимума) в оригинальном пространстве на основе полученного полинома в пространстве меньшей размерности и прямоугольной матрицы из первого пункта.

**Классификация и сравнительный анализ рассмотренных методов.** Критерии, согласно которым можно классифицировать рассматриваемые методы повышения разрешения изображения, представлены на рис. 7.

При сравнительном анализе методов повышения разрешения изображения были выделены следующие критерии: вычислительная сложность метода, качество обработки и необходимость пост- или предобработки.

В таблице представлен сравнительный анализ рассмотренных методов.

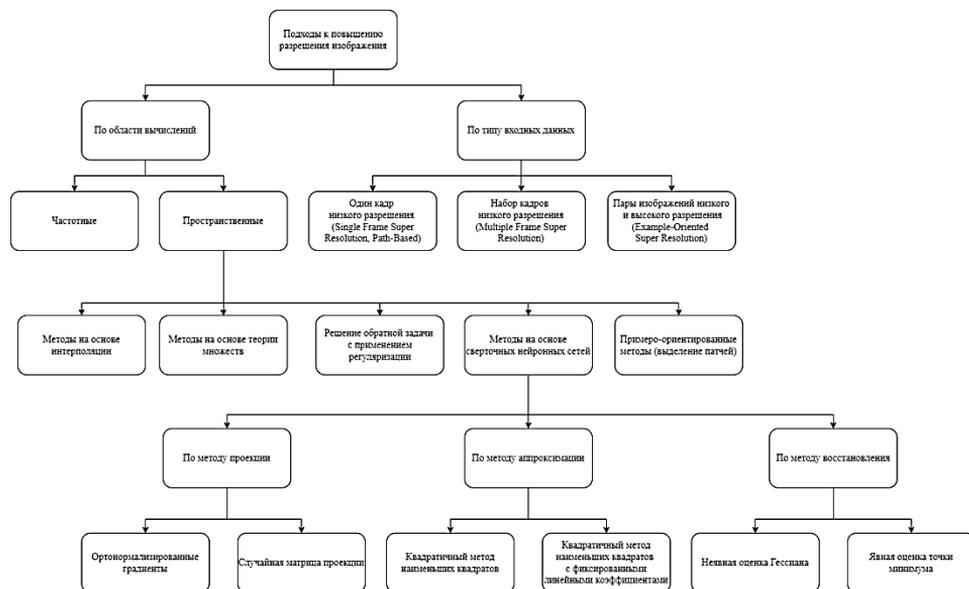


Рис. 7. Классификация подходов повышения разрешения изображения

**Сравнительный анализ рассмотренных методов**

Методы	Критерий		
	Вычислительная сложность	Качество обработки	Необходимость дополнительной обработки
Частотные	Низкая	Низкое	Нет
Интерполяционные	Низкая	Очень низкое	Нет
Методы теории множеств	Средняя	Низкое	Да
Регуляризация	Средняя	Среднее	Нет
Выделение патчей	Очень высокая	Высокое	Да
Сверточные нейронные сети	Очень высокая	Очень высокое	Да

**Заключение.** 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, поскольку в рамках поставленной задачи ключевым фактором является качество получаемого результата, а другие факторы являются второстепенными.

## Литература

- [1] Насонов А.В., Крылов А.С. Быстрое суперразрешение изображений с использованием взвешенной медианной фильтрации. *Цифровая обработка сигналов и ее применение. 12-я Междунар. конф. и выставка: сб. тр.* Москва, МГУ имени М.В. Ломоносова, 2010, т. 2, с. 101–104.
- [2] Park S.C., Park M.K., Kang M.G. Super-resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, 2003, vol. 20, no. 3, pp. 21–36. <https://doi.org/10.1109/MSP.2003.1203207>
- [3] Кокошкин А.В. и др. Оценка ошибок синтеза изображений с суперразрешением на основе использования нескольких кадров. *Компьютерная оптика*, 2017, т. 41, № 5, с. 701–711. <https://dx.doi.org/10.18287/2412-6179-2017-41-5-701-711>
- [4] Wang Q., Tang X., Shum H. Patch based blind image super resolution. *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, IEEE, 2005, vol. 1, pp. 709–716. <https://doi.org/10.1109/ICCV.2005.186>
- [5] Freeman W.T., Jones T.R., Pasztor E.C. Example-based super-resolution. *IEEE Computer graphics and Applications*, 2002, vol. 22, no. 2, pp. 56–65.
- [6] Katsaggelos A.K. *Digital image restoration*. Springer Publishing Company, Incorporated, 2012.
- [7] 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, vol. 2, pp. 333–337. <https://doi.org/10.1109/ISDA.2006.253857>
- [8] Stark H., Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections. *JOSA A*, 1989, vol. 6, no. 11, pp. 1715–1726. <https://doi.org/10.1364/josaa.6.001715>
- [9] Бредихин А.И. Алгоритмы обучения сверточных нейронных сетей. *Вестник Югорского государственного университета*, 2019, № 1 (52), с. 41–54.
- [10] Umehara K., Ota J., Ishida T. Application of super-resolution convolutional neural network for enhancing image resolution in chest CT. *Journal of digital imaging*, 2018, vol. 31, pp. 441–450. <https://doi.org/10.1007/s10278-017-0033-z>
- [11] Dong C. et al. Learning a deep convolutional network for image super-resolution. *Computer Vision-ECCV 2014: 13th European Conference. Proceed-*

- ings*, Zurich, Switzerland, Springer International Publishing, 2014, part IV, pp. 184–199. [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13)
- [12] Senov A. Projective approximation based quasi-Newton methods. *Proc. of International Workshop on Machine Learning, Optimization, and Big Data*, Springer, Cham, 2017, pp. 29–40. [https://doi.org/10.1007/978-3-319-72926-8\\_3](https://doi.org/10.1007/978-3-319-72926-8_3)
- [13] Сенов А.А. Глубокое обучение в задаче реконструкции суперразрешения изображений. *Стохастическая оптимизация в информатике*, 2017, т. 13, № 2, с. 38–57.

**Поступила в редакцию 03.06.2024**

**Сироткина Полина Юрьевна** — студентка кафедры «Программное обеспечение ЭВМ и информационные технологии», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

**Ссылку на эту статью просим оформлять следующим образом:**

Сироткина П.Ю. Классификация методов повышения разрешения изображения по нескольким кадрам. *Политехнический молодежный журнал*, 2024, № 05 (94). URL: [https://ptsj.bmstu.ru/catalog/icec/inf\\_tech/1002.html](https://ptsj.bmstu.ru/catalog/icec/inf_tech/1002.html)

## CLASSIFICATION OF METHODS IN INCREASING IMAGE RESOLUTION OVER SEVERAL FRAMES

**P.Yu. Sirotkina**

sirotkinapyu@student.bmstu.ru

*Bauman Moscow State Technical University, Moscow, Russian Federation*

The paper is devoted to reviewing the methods in increasing the image resolution (the so-called, super-resolution problem) based on several frames of the same object. It considers in brief the subject area and main approaches to solving the problem. They include interpolation-based methods, reconstruction methods with regularization, methods based on the set theory, example-oriented methods (patch extraction), and methods for optimizing the convolutional neural networks. The paper provides a critical analysis of methods increasing image resolution based on several frames with identifying the areas of most efficient use based on the following criteria: computational complexity, processing quality, need for post- or pre-processing. The paper concludes on the applicability area for the considered methods. To achieve the best processing quality, it is advisable to introduce the methods based on the convolutional neural networks.

**Keywords:** digital image, signal processing, super-resolution, convolutional neural networks, interpolation, reconstruction, regularization, set theory, patches

---

***Received 03.06.2024***

**Sirotkina P.Yu.** — Student, Department of Computer Software and Information Technology, Bauman Moscow State Technical University, Moscow, Russian Federation.

### **Please cite this article in English as:**

Sirotkina P.Yu. Classification of methods in increasing image resolution over several frames. *Politekhnicheskiy molodezhnyy zhurnal*, 2024, no. 05 (94). (In Russ.). URL: [https://ptsj.bmstu.ru/catalog/icec/inf\\_tech/1002.html](https://ptsj.bmstu.ru/catalog/icec/inf_tech/1002.html)