НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
июль–август 2023 Том 23 № 4 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
July–August 2023 Vol. 23 No 4 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

doi: 10.17586/2226-1494-2023-23-4-734-742

Brain MRT image super resolution using discrete cosine transform and convolutional neural network

Pooja Singh[1]✉, Dinesh Ganotra[2]

[1,2] Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India

[1] pujasingh0409@gmail.com✉, https://orcid.org/0000-0002-8603-5954
[2] dinesh_ganotra@hotmail.com, https://orcid.org/0000-0002-3720-8716

Abstract

High Resolution (HR) images have numerous applications, such as video conferencing, remote sensing, medical imaging, etc. Furthermore, a few challenges with the super resolution algorithms of magnetic resonance brain images are now obtainable, namely, low sensitivity, significant frequency noise as well as poor resolution. To fix these problems, a Convolutional Neural Network (CNN) based Discrete Cosine Transform (DCT) singular frame quality improvement method is described. There are two stages in this proposed method, involving training and testing. During the training stage, the HR, and Low Resolution (LR) pictures are employed as input, and they are preprocessed to create blocks of images. The histogram and DCT are used for extracting the features from the LR and HR blocks, and these extracted features are assigned with class id. The CNN, which extracts the features and allocates class id, receives its feature extractor as its final input. An LR input image is once more divided into [2 × 2] blocks during the testing stage, so each block histogram and DCT feature are estimated. Each feature vector is fed into the neural network as well as the results are contrasted with a set of feature vectors that have been recorded, in addition to the class id that has been allocated to a certain vector. In order to generate a Super resolution image with an LR image, a relevant HR block is then swapped out for this LR block. These results indicated that the initial dataset can achieve 22.4 and 19.5 Peak Signal to Noise Ratio (PSNR) and Root Mean Square Error (RMSE) values while measuring the effectiveness of this proposed method using RMSE and PSNR. Then, the second dataset illustrates that the PSNR and RMSE values are 20.1 and 25.5. For the third dataset, the values are 45.7 and 12.3, respectively. However, the presented method works better than the neural method of Super Resolution Channel Spatial Modulation Network and resolution enhancement technique.

Keywords

high resolution, low resolution, discrete cosine transform, resolution enhancement, RMSE, PSNR, convolutional neural network

For citation: Singh P., Ganotra D. Brain MRT image super resolution using discrete cosine transform and convolutional neural network. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 734–742. doi: 10.17586/2226-1494-2023-23-4-734-742

УДК 004.032.26

Сверхвысокое разрешение изображения магнитно-резонансной томографии головного мозга с использованием дискретного косинусного преобразования и сверточной нейронной сети

Пуджа Сингх[1]✉, Динеш Ганотра[2]

[1,2] Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия

[1] pujasingh0409@gmail.com✉, https://orcid.org/0000-0002-8603-5954
[2] dinesh_ganotra@hotmail.com, https://orcid.org/0000-0002-3720-8716

Аннотация

Изображения с высоким разрешением (High Resolution, HR) имеют широкое применение, например при проведении видеоконференций, дистанционного зондирования, медицинской визуализации и других. С использованием алгоритмов сверхвысокого разрешения появилась возможность решить несколько проблем

магнитно-резонансной томографии изображений мозга, связанных с низкой чувствительностью, значительным частотным шумом, а также низким разрешением. Чтобы устранить данные проблемы, предложен метод улучшения качества сингулярного кадра на основе сверточной нейронной сети (Convolutional Neural Network, CNN) с дискретным косинусным преобразованием (Discrete Cosine Transform, DCT). Метод состоит из двух этапов, включающих обучение и тестирование. На этапе обучения изображения HR и низкого разрешения (Low Resolution, LR) используются в качестве входных данных и проходят предварительную обработку для создания блоков изображений. Для извлечения признаков из блоков LR и HR применены гистограмма и DCT. Извлеченным признакам присваивается идентификатор класса. CNN извлекает функции DCT, назначает идентификатор класса и получает свой экстрактор функций для окончательного ввода. Входное изображение LR на этапе тестирования повторно делится на блоки [2 × 2], с помощью гистограммы оценивается каждый блок и функции DCT. Каждый вектор признаков передается в нейронную сеть, полученные результаты сравниваются с набором векторов признаков, которые были записаны, в дополнение к идентификатору класса и назначены определенному вектору. Для генерации изображения сверхвысокого разрешения с изображением LR соответствующий блок HR заменяется на блок LR. Полученные результаты показали, что эффективность предложенного метода исходного набора данных достигла значений отношения пикового сигнала к шуму (PSNR) и среднеквадратичной ошибки (RMSE) 22,4 и 19,5 соответственно. Второй набор данных показал значения PSNR и RMSE, равные 20,1 и 25,5, а третий набор — 45,7 и 12,3. Таким образом, представленный метод работает лучше, чем нейронная сеть пространственной модуляции канала сверхвысокого разрешения и метод повышения разрешения.

Ключевые слова

высокое разрешение, низкое разрешение, дискретное косинусное преобразование, повышение разрешения, среднеквадратическая ошибка, PSNR, сверточная нейронная сеть

Ссылка для цитирования: Сингх П., Ганотра Д. Сверхвысокое разрешение изображения магнитно-резонансной томографии головного мозга с использованием дискретного косинусного преобразования и сверточной нейронной сети // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 734–742 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-734-742

Introduction

High Resolution (HR) images have numerous applications, such as video conferencing, remote sensing, medical imaging, etc. Low Resolution (LR) pictures are converted into HR pictures through an algorithm using resolution improvement techniques [1, 2]. It is relatively easier to retrieve an HR image from numerous LR images rather than from only one LR photo. Computer vision is highly dependent upon the challenge of retrieving an HR image from a single LR image. This problem is known as an under-determined inverse problem having no unique solution. Most often, the resolution to this issue helps to limit the possible solutions using solid previous knowledge.

In most cases, an illustration method is used to compare the most recent and prior province methods. Some methods can focus on an image internal similarity [3] or outside LR and HR pairings to identify mapping operations [4]. These methods based on single-image Super Resolution (SR) are categorized into two classes: sparse coding [5, 6] based on deep learning [7, 8]. Spare coding is an example of one of the typical external example-based resolution enhancement techniques [9, 10]. The coding coefficients are identical in the patch space of each pair of patches in HR and LR images. The overlapped patch were densely cropped and pre-processed from source image (e.g., subtracting mean and normalization).

A LR dictionary then encodes these patches. An HR dictionary is used to store the sparse coefficients for reconstructing HR patches. To generate the final output, the overlapping reconstructed patches are aggregated. One of the earlier approaches that use Convolutional Neural Network (CNN) for resolution enhancement is Super Resolution Convolutional Neural Network (SRCNN). It is connected to the conventional resolution enhancement method focused on sparse coding, according to the nature of the network structure. This network comprises three convolutional layers [11]. The scale of the convolution kernel is respectively [1 × 1], [3 × 3], ... [9 × 9]. While SRCNN would be an end-to-end method that proceeds completely forwards, their speed is better compared to conventional techniques.

Deep convolution networks have been shown to produce better results than the conventional resolution enhancement approach [12]. As improved network length, deep neural networks have become more effective. In order to improve the network size, both numbers of layers as well as the units at every layer are grown. Specifically, when working with such a big amount of training data set, this is a rapid and effective method for developing predictive method [13]. If there are only few identified samples in the training set, having a greater variety of parameters makes it harder for the network training to convergence and renders the enlarged networks highly susceptible to overloading.

Direct end-to-end mapping from LR photos to HR images is the target of deep learning based methods. In this paper, the designed method consists of two phases: training and testing. At first, the HR and LR images are taken as input and are preprocessed for dividing the image pixels into blocks such as [2 × 2] and [4 × 4]. The histogram and Discrete Cosine Transform (DCT) are used for extracting the features from the LR and HR blocks. Each and every extracted feature vector from LR and HR blocks using the DCT is aligned with the class id. In the end, the DCT and histogram characteristics are merged, but this acts as an input for the CNN. Through the use of a class id which is aligned, the features from the input element are extracted to create the SR image by swapping out the LR values into HR values. The LR image should be first compressed to a [2 × 2] block in the testing stage, before the histogram and DCT are evaluated, and the LR block is therefore swapped out for the proper HR block to generate the superior image

quality. An earlier resolution improvement method for head imaging used to have a drawback that the proposed scheme has fixed.

The following list contains the main goals of the proposed method.

— LR Magnetic Resonance Tomography (MRT) brain scans can be improved to produce HR MRT scans utilizing CNN.
— Blocks made from the image pixels are employed to effectively features extracted.
— The histogram and DCT methods are used to extract the features in order to minimize the time training in CNN.

The main accomplishments of the proposed method include training the CNN with HR brain MRT images to transform LR MRT brain images into SR MRT brain images. In order to extract features using HR and LR images, the image pixels are split into blocks. Both histogram and DCT method are then used to retrieve the coefficient of feature vectors from LR and HR pictures.

Literature Review

This section addresses some of the essential studies related to the Brain MRT SR image for transforming the LR image to HR image previously proposed utilizing diverse deep learning approaches.

Dong et al. [14] used an end-to-end mapping between the LR–HR images which is straightforwardly learned in a deep learning method that has been proposed. Considering LR image input and HR image output, this mapping is represented as a deep CNN. They also demonstrated that conventional resolution enhancement methods based on sparse-coding could also be observed as a deep convolutional network. This proposed method optimizes all layers, unlike conventional approaches that treat each variable separately, and proposed deep CNN has a smaller network but shows state-of-the-art quality in restoration. However, CNN end performance is poor due to its capacity limitations.

Aharon et al. [15] proposed a method using the representation of sparse signals for single-image SR. An over-complete dictionary describes image patches using a sparse linear combination of components. For the LR image, sparse representation is calculated for each patch, and the HR image is produced with the help of this coefficient. These outcomes demonstrate that signals that were down-sampled can still accurately retrieve the excess representation. It is possible to determine exactly comparable the dense representation of the LR and HR image patch pairings can be with regard to respective vocabularies through simultaneously training 2 dictionaries for the LR and HR image patches. When compared to earlier approaches, the learned vocabulary pairing samples a greater number of picture patch pairings, which minimizes the computational expense. HR images are generated by using this algorithm that is comparable or better in quality to images formed by other similar resolution enhancement approaches. Furthermore, this approach is inherently robust to noise.

Rueda et al. [16] introduced an SR method based on sparse coding that integrates low and high-frequency information to generate an HR image from an LR brain MRT image which was adapted for conveniently integrating prior knowledge. With a 3D HR reconstruction, the proposed method considerably improves computationally both speed and precision by incorporating a whole-image multi-scale edge assessment with dimensionality reduction technique. Reconstructed and interpolated reconstructed types of 29 MRT brain images were compared to the novel images learned in a 3T scanner to validate the method, yielding a 70 % reduction in Root Mean Square Error (RMSE) and a 10.3 dB increase in Peak Signal to Noise Ratio (PSNR). Its method outperforms a present state algorithm, demonstrating it will have a big impact on voxel-based morphometry analyses. These approaches are quick and simple to understand and use, but they struggle with edges and complex textures.

Wang and Jiang [17] proposed a research method of the Adaptive Regulation of Thought-Rational resolution enhancement reconstruction method. In order to train a more advanced end-to-end deep threshold recurrent neural network, these authors also added a recurrent layer to the basic CNN during the matching step. During the recruitment process, the neighbourhood embedding approach is employed to constitute for the lost image data. According to the experimental outcomes, the methods has a great potential for photo restoration and has achieved good outcomes in calculating metrics for both PSNR and Structural Similarity Index. However, the quality of reconstruction begins to decline.

Liu et al. [18] have suggested an improvement of image resolution used in medicine. A method was proposed for increasing resolution that first retrieves an HR picture from its counterpart using lower ranking as well as nonlocal self-similar consistency. By iteratively applying a subsampling accuracy restriction on low rank restoration, then they improve the recovered HR value. Results obtained on Magnetic Resonance and Computer Tomography images illustrate that the proposed method works better than standard estimation techniques and therefore it is comparable with current state-of-the-art technology in terms of both measures and picture quality. It results in a large computational cost, though.

Liu et al. [19] had developed a Positron Emission Tomography (PET) Imaging Image Enhancement Using Artificial Intelligence (AI). Excessive noises as well as a restricted spatial resolution are two important influencing factors that limit the qualitative and statistical efficiency of PET scans. AI methods for image denoising and de-blurring are growing in popularity for post-reconstruction enhancement of PET images. The proposed method offers a comprehensive analysis of prior AI-based PET image enhancement efforts, with a focusing on network architectures, file formats, loss functions and evaluation criteria.

According to the mentioned study, brain imaging super resolution faces a number of important difficulties. Due to its capacity limitations, EES end performance is poor (Dong et al. [14]). This approach is inherently robust to noise Aharon et al. [15], quick and simple to understand and use, but they struggle with edges and complex textures used by Rueda et al. [16], and the quality of reconstruction

begins to decline Wang et al. [17]; it leads to a high computational cost of Liu H. et al. [18]. CNN was proposed for brain image SR as a solution to this issue.

Proposed Method for Brain Image Super Resolution

Resolution is generated by image sensors or image collection hardware. The size or quantity of sensor elements affects an images spatial resolution. Converting LR photos to HR is termed as brain image improvement. For improving picture quality from extracted features in this method, CNN is suggested. Video conferencing, remote sensing, medical imaging, and several other uses are available for HR images. LR photos are transformed into HR images by an algorithm utilizing resolution enhancement techniques. A CNN is suggested in this method because previous Brain-computer interface technology mainly depends on artificial stimuli for resolution improvement.

The process flow of the proposed method is illustrated in Fig. 1. The HR [256 × 256] and LR [128 × 128] images are initially selected as input for the training phase, then they pre-processed to split the image pixels into blocks. After that, the histogram and DCT methods are utilized to extract the features from HR and LR blocks. The CNN receives these extracted features as input and utilizes them to extract the features and establish a class id. The histogram and DCT feature of an LR input image were computed for each block of the source images throughout testing. While feeding the neural network with this feature data, the output is assessed against a previously established recorded feature vector which is given a class id. In order to create a SR image from an LR image, each LR block is again swapped out with the equivalent HR block. Three distinct dataset of brain MRT images are utilized to test the proposed method.

Data Augmentation

When including slightly modified versions of either existing data or brand-new synthetic data that is derived from previous data, data augmentation in data analysis is utilized to expand the quantity of data. Less data is employed in the datasets for such method design for training and testing the classifier. Consequently, the data augmentation technique increases the dataset size.

Pre-processing

The pixels are divided into blocks using pre-processing in this method. HR and LR images are utilized for the pre-processing for it. From the first dataset[1]. LR images of size [128 × 128] with matching HR images of size [256 × 256] from the SR of Brain MRT Images Dataset Download was employed for training. The LR image is separated into blocks of size [2 × 2] but each HR image is split into blocks of size [4 × 4] pixels, obtaining 4096 blocks of [2 × 2] as well as [4 × 4] pixel resolution by each LR and HR image, correspondingly. For 21 LR images total [21 × 4096] blocks were used for training.

Another dataset: 1 of 49 images was used for training method land from another database [2 × 46] images were used for training, resulting in [49 × 4096] and [46 × 4096] blocks respectively.

1. Feature extraction

The process of feature extraction is fundamental to the improvement of images. These features are extracted of the pre-processed images using the DCT method. Each LR and HR block was utilized to extract the histogram with 2D DCT. Even though the input to the 2D DCT was a matrix of size [2 × 2], the output of the 2D DCT function was chosen as a matrix of size [3 × 3] to improve the image quality and reduce the complexity. For the histogram, 16 bins were chosen. This shape of each HR block now consists of a feature vector with a size of [1 × 25]. The sixteen bins were from 0–15, 16–31, …, 240–255 grayscale values.

It is possible to add cosine functions which fluctuate at specific frequencies to describe a continuous sequence of points using a DCT. Extraction of DCT features occurs in two phases. First, the entire image is subjected to the DCT to generate DCT coefficients, while in the following phase, feature vectors are built using a subset of the coefficients that have been chosen in the initial process. Approximately the same size to the input image is the DCT coefficient matrix [20].

[1] Available at: https://projecttunnel.com/Super-Resolution-of-Brain-MRI-Images-Dataset-Download (accessed: 14.07.2023).

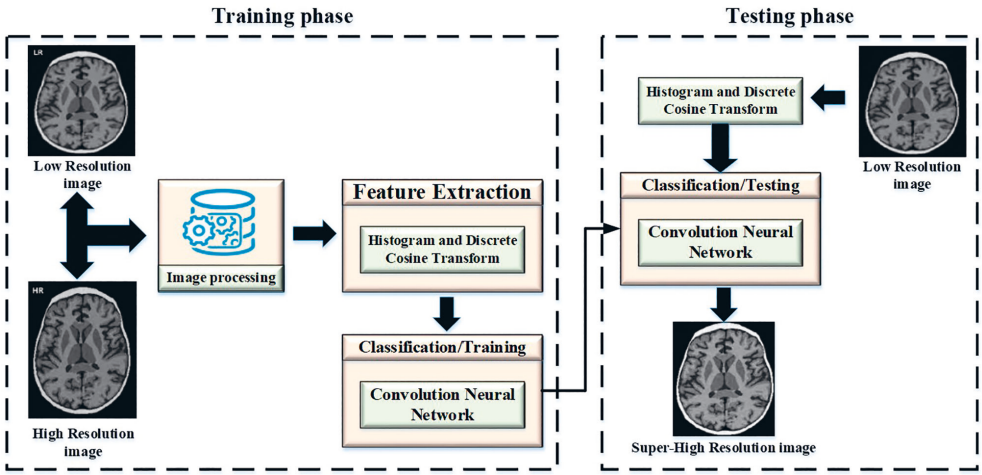

Fig. 1. Architecture of the proposed method

For image of $M \times N$ size, in which each photo relates to a 2D matrix, DCT coefficients are calculated.

$$\mathbf{F}(u, v) = \frac{1}{\sqrt{MN}}\alpha(u)\alpha(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y) \times$$
$$\times \cos\left(\frac{(2x+1)u\pi}{2M}\right)\cos\left(\frac{(2y+1)v\pi}{2N}\right),$$

where $u = 0, 1... M - 1$ and $v = 0, 1... N - 1$, $\alpha(w)$ is defined by,

$$\alpha(w) = \begin{cases} \frac{1}{\sqrt{2}}, w = 0 \\ 1, otherwise \end{cases}.$$

The image brightness function is known as $f(x, y)$, and the 2D matrix $\mathbf{F}(u, v)$ includes the DCT coefficients. Applying the DCT to the entire image gives the frequencies coefficients matrices of the same dimensions. A method is utilized for feature extraction that has a significant impact on accurate identification is the second phase. The purpose of selecting the coefficient is used to minimise the reconstruction error.

2. Convolutional Neural Network

In this method, CNN is proposed for extracting features from the previously extracted image. After using CNN for extracting the features, the LR values are replaced by the HR values. An already-trained version of the networks, which has been trained on more than a million pictures, can be loaded from the ImageNet dataset. Evaluate this proposed method to handle recognizing various kinds of sign images.

As inputs to the Convolution layers, the feature vector of dimension [1 × 25] was utilized. Using [2 × 2] kernel in the CNN layer, [4 × 4] feature vector is obtained. This feature vector sent to the max-pooling layer results in [3 × 3] size. After CNN, each feature vector has size [3 × 3]. Each corresponding input and output feature vector pair were assigned a number referred to as class id. For the same feature vector, the same class id is assigned.

The convolution layer, which is the initial layer of the CNN, executes a convolution operation on the input, as illustrated in Fig. 2. A total result is then filtered using an activation function, such as Rectified Linear Unit, and sent to the pooling layer which is the subsequent layer. The following equation will be utilized to estimate the convolution layer operation [21]:

$$O_{i(x, y)} = f'(\sum_{i=0}^{m-I}\sum_{j=0}^{m-I} w(i, j)I(x + i, y + j) + bias),$$

where O is the basis function of fully connected layer; f' is the representation of function; I is the representation of input image; x and y are the mapping row and column of image; m is the maximum size of filter in coevolution; *bias* is the bias value.

This equation includes map dimension as well as the network parameters which are minimized through the use of the pooling layer. The features, such as edges and points, are extracted using max pooling. In order to link the neurons from the current layer to the following layer, a completely connected layer is therefore used. An input image is classified into different classes using the training dataset. The neural network output y is identified as in [21]

$$y = \sigma(w^L .. \sigma(w^2\sigma(w^1x + b^1) + b^2) \ldots + b^L),$$

where σ is an activation function, w is a network parameter, b is a bias, and x is input, L is layer number.

3. Training

Depending on the class id of the feature vector for LR block, blocks were recognized during training. The trained vector, which has a size of [86,016 × 9] for the primary dataset, has size [200,704 × 9] for second dataset and [188,416 × 9] for third dataset, is created by obtaining the class ids for the LR and HR photo blocks.

A histogram feature with 16 bins and a DCT of 9 values are used in the proposed method. Thus, we have an input vector of 25 elements. This 25-element vector is placed through a CNN method, and the resulting 9-element vector includes what the initial hidden layer will use as input. This output of the initial hidden layer has 128 items and also the 2 items hidden layer includes another 9 elements. Finally, there is only one output element in the output layer which denotes the class id. At the input to the hidden layer and the hidden layer output layer, sigmoid functions are employed as activation functions. In order to identify the correct class id for the LR block, the neural network obtains its inputs feature vector. It is replaced with an HR block that matched this LR block.

4. Testing

This stage again divides the LR input image into [2 × 2] blocks for testing, so each block histogram and DCT features are obtained.

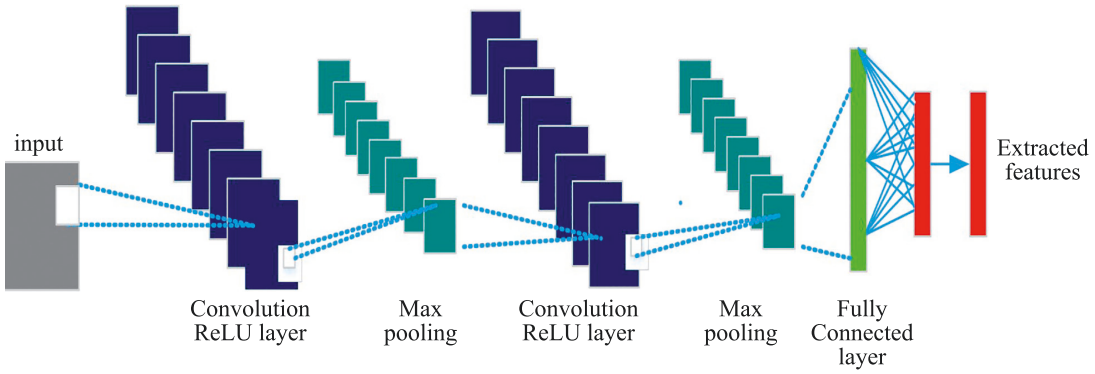

Fig. 2. Generic architecture of Convolutional Neural Network

The class id is given to the feature vector when the outputs of the neural network are evaluated with a group of feature vectors that have been recorded. In order to create an HR image with an LR image, an appropriate HR block is then exchanged for this LR block. Ten pictures from the initial dataset were utilized to test the proposed method. 40 images were used to test the method for the second and third datasets.

Result and Discussion

The purpose of this method is to transform a low quality picture into a high-quality image. This section assesses the CNN. MATLAB 2021a is utilized to execute the testing, with an Intel Core i5, NVidia GeForce GTX 1650 and 16GB of RAM. In order to evaluate if it transforms the LR image into an HR image, CNN examines the data entered. Whenever the data was first collected, it was utilized in two phases: training and testing. The CNN classifier in the present method receives its input from the DCT features that have been retrieved.

Both LR and HR pictures are utilized as input initially, then pre-processing follows. Pre-processing is used to decompose picture pixels into blocks, while extracted features serves to feature extraction from the pre-processed pictures discrete cosine transform. Block splitting is employed by the proposed method for both the HR and LR images which have pixel sizes of [256 × 256] and [128 × 128], respectfully. Different class ids are given after extracting the histogram and DCT features from such LR and HR blocks. These feature vectors are used to train the method to identify the relevant class id. CNN is utilized to train the method. Three distinct datasets of brain MRT images are employed to test the proposed framework. The proposed method results are compared to Super Resolution Channel Spatial (SRCS) and neural network-based resolution enhancement. The performance of this proposed method with neural network-based learning and the SRCS method was calculated using three datasets of brain MRT images.

Dataset description

Dataset 1: There are 62 photographs in this set, including 31 images in LR and 31 in HR. Each low-resolution image is [128 × 128] pixels in size[1]. Similarly, each HR image is 256 × 256 pixels in size.

Dataset 2: The data are collected from the Kaggle dataset. There are 155 non-tumorous data images in this dataset, and 98 tumorous images[2].

Dataset 3: the data are collected from the figshare dataset. In this dataset, 86 images are used[3].

The above chosen datasets consist of 31, 89 and 86 images of LR for training and testing using the classifier. These fewer amounts of data are not feasible for training the CNN. Instead of searching for a new dataset, the collected dataset is augmented 10 times, resulting in more data for training the CNN.

In Table 1, in the first dataset totally of 31 images are used for training and testing. For training, 21 images are used, and 10 images are taken from the total images for testing purposes.

In Fig. 3, from the first dataset of 31 brain MRT images, 21 were used as training, and 10 were used for the testing. In comparison to the SRCS and neural network based method, the PSNR values and on average, raised by 6 dB and 9.6 dB, respectively, based on the outcomes of these 10 testing images. RMSE is decreased by 12.8 dB and 22.2 dB than the neural network based and SRCS methods.

The average RMSE and PSNR of 10 test images are shown in Table 2. For the parameter PSNR, the SRCS value is 12.8, neural network based value is 18.8, and our method reaches 22.4. For RMSE, the value of SRCS is 41.7, the neural network based value reaches 28.9, and the proposed method reaches 19.5, respectively.

In Table 3, in the first dataset totally 89 images are used for training and testing: 49 images are used for training,

Table 1. First dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	31	21	10
After	310	248	62

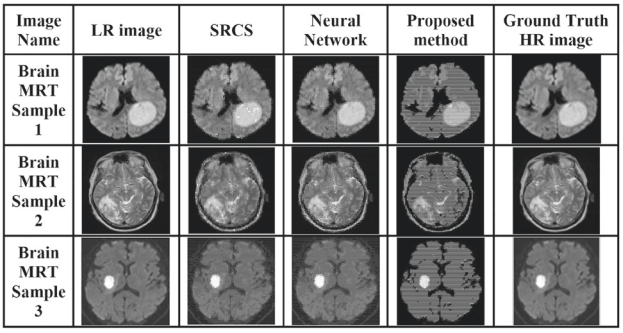

Fig. 3. Results of various techniques for Dataset 1

Table 2. Average RMSE and PSNR of 10 test images from dataset 1

Parameter	SRCS	Neural Network	Proposed Method
PSNR	12.8	18.8	22.4
RMSE	41.7	28.9	19.5

Table 3. Second dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	89	49	40
After	890	712	178

[1] Dataset 2: https://projecttunnel.com/Super-Resolution-of-Brain-MRI-Images-Dataset-Download (accessed: 14.07.2023).

[2] Chakrabarty N. (2019, April 14). Brain MRI images for Brain tumor detection. Kaggle. Available at: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection (accessed: 14.07.2023).

[3] Dataset 3: https://figshare.com/articles/dataset/brain_tumor_dataset/1512427 (accessed: 14.07.2023).

and 40 images are taken from the total images for testing purposes.

In Fig. 4, for the second dataset of 89 images, 49 were used for training the network, and the proposed method was tested on 40 images. After averaging the RMSE and PSNR values across the 40 images, it was found that the RMSE is lower by 2.2 dB and 9.3 dB than the neural network based method as well as SRCS method, respectively, and also the PSNR values are greater by 20.6 dB and 23.6 dB than the neural network based method and SRCS, respectively.

The average RMSE and PSNR of 40 test images are shown in Table 4. For the parameter PSNR, the SRCS value is 10.8, neural network based value is 16, and our method reaches 20.1. For RMSE, the value of SRCS is 49.1, the neural network based value reaches 34.5, and the proposed method reaches 25.5, respectively.

In Table 5, in the first dataset totally 86 images are used for training and testing: 46 images are used for training, and 40 images are taken from the total images for testing purposes [22].

In Fig. 5, for the third dataset having 86 images, 46 were used for training purposes and testing was done on 40 images. The outcomes of these 40 test images were averaged, while it was found that the RMSE was reduced by 4.4 dB and 9.7 dB compared to the neural network-based approach and also the SRCS method, respectively, and thus the PSNR values significantly improved by 10.5 dB and 26.4 dB. DCT and histogram features were used to enhance the proposed method accuracy through training.

The average RMSE and PSNR of 40 test images are shown in Table 6. For the parameter PSNR, the SRCS value is 19.3, neural network based value is 35.2, and our method reaches 45.7. For RMSE, the value of SRCS is 22.0, the neural network based reaches 16.7, and the proposed method reaches 12.3, respectively.

Table 7 shows the comparison of PSNR and RMSE for proposed and existing approach. Proposed approach attain 45.7 % PSNR value which is high when compared to other current approaches, such as Dual residual Path Block (DBPN), Channel Spatial Attention Module (CSAM) and Deep Back Projection Network (DRPB) whose values are 31.86 %, 31.42 % and 30.17 %, respectively. RMSE values for proposed approach is 12.3 % which is low when compared to the current approaches. It clearly shows that the proposed approach yield better performance when compared to other current approaches.

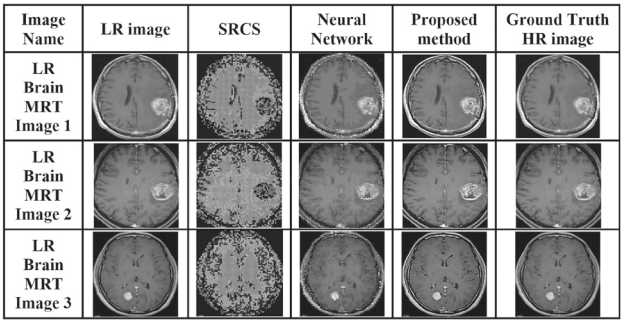

Fig. 4. Results of various techniques for Dataset 2

Table 4. Average RMSE and PSNR of 40 test images from dataset 2

Parameter	SRCS	Neural Network	Proposed Method
PSNR	10.8	16	20.1
RMSE	49.1	34.5	25.5

Table 5. Third dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	86	46	40
After	860	688	172

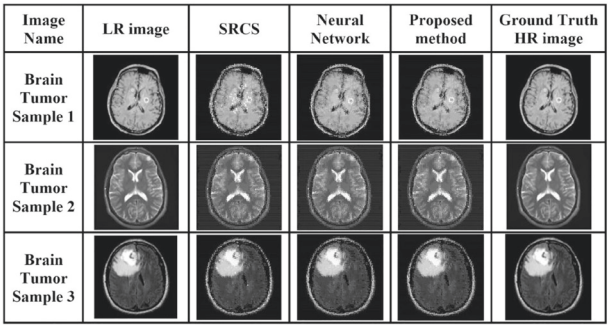

Fig. 5. Results of various techniques for Dataset 3

Table 6. Average RMSE and PSNR of 40 test images from dataset 3

Parameter	SRCS	Neural Network	Proposed Method
PSNR	19.3	35.2	45.7
RMSE	22.0	16.7	12.3

Table 7. Average RMSE and PSNR for proposed and existing approaches

Parameter	Proposed method	DBPN [23]	CSAM [24]	DRPB [25]
PSNR	45.7	31.86	31.42	30.17
RMSE	12.3	17.8	22.4	28.11

Conclusion

In this research, DCT and CNN based resolution improvement method called “Brain Image” was developed. CNN is utilized to extract the features and change the LR value with the HR value. The HR image and LR images are divided into smaller blocks. The histogram and DCT features are extracted for each HR and LR block. The method is trained using these extracted features, and the unique class id is allotted to each feature vector. The proposed method attained 22.4 and 19.5 PSNR and RMSE values. And for the second dataset, the PSNR and RMSE

values are 20.1 and 25.5. For the third dataset, the values are 45.7 and 12.3. The analysis of three various datasets of brain MRT images reveals comparisons between the proposed approach and SRCS and resolution improvement using a neural network. The RMSE and PSNR values are improved as compared to these two methods. As a result, the proposed approach may be a good improvement over the methods already being used. The processing of large-scale medical images and pixel fusion picture registration should be enhanced in future work.

References

1. Chen Q., Huang J., Feris R., Brown L.M., Dong J., Yan S. Deep domain adaptation for describing people based on fine-grained clothing attributes. *Proc. of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 5315–5324. https://doi.org/10.1109/cvpr.2015.7299169
2. Denton E.L., Chintala S., Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks. *Advances in Neural Information Processing Systems*, 2015, vol. 28.
3. Cui Z., Chang H., Shan S., Zhong B., Chen X. Deep network cascade for image super-resolution. *Lecture Notes in Computer Science*, 2014, vol. 8693, pp. 49–64. https://doi.org/10.1007/978-3-319-10602-1_4
4. Farhadifard F., Abar E., Nazzal M., Ozkaramanh H. Single image super resolution based on sparse representation via directionally structured dictionaries. *Proc. of the 22nd Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2014, pp. 1718–1721. https://doi.org/10.1109/siu.2014.6830580
5. Ahmed J., Memon R.A., Waqas M., Mangrio M.I., Ali S. Selective sparse coding based coupled dictionary learning algorithm for single image super-resolution. *Proc. of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 2018, pp. 1–5. https://doi.org/10.1109/icomet.2018.8346357
6. Choi J.H., Kim J.H., Cheon M., Lee J.S. Deep learning-based image super-resolution considering quantitative and perceptual quality. *Neurocomputing*, 2020, vol. 398, pp. 347–59. https://doi.org/10.1016/j.neucom.2019.06.103
7. Dong C., Loy C.C., He K., Tang X. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, vol. 38, no. 2, pp. 295–307. https://doi.org/10.1109/TPAMI.2015.2439281
8. Dong C., Loy C.C., Tang X. Accelerating the super-resolution convolutional neural network. *Lecture Notes in Computer Science*, 2016, vol. 9906, pp. 391–407. https://doi.org/10.1007/978-3-319-46475-6_25
9. Ayas S., Ekinci M. Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation. *Information Sciences*, 2020, vol. 512, pp. 1264–1278. https://doi.org/10.1016/j.ins.2019.10.040
10. Gu S., Zuo W., Xie Q., Meng D., Feng X., Zhang L. Convolutional sparse coding for image super-resolution. *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1823–1831. https://doi.org/10.1109/iccv.2015.212
11. Dosovitskiy A., Springenberg J.T., Brox T. Learning to generate chairs with convolutional neural networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1538–1546. https://doi.org/10.1109/cvpr.2015.7298761
12. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error. *Proc. of the 4th International Conference on Learning Representations (ICLR)*, 2016.
13. Alec R., Luke M., Soumith C. Unsupervised representation learning with deep convolutional generative adversarial networks. *Proc. of the International Conference on Learning Representations (ICLR)*, 2015, pp. 1–16.
14. Dong C., Loy C.C., He K., Tang X. Learning a deep convolutional network for image super-resolution. *Lecture Notes in Computer Science*, 2014, vol. 8692, pp. 184–199. https://doi.org/10.1007/978-3-319-10593-2_13
15. Aharon M., Elad M., Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, vol. 54, no. 11, pp. 4311–4322. https://doi.org/10.1109/tsp.2006.881199
16. Rueda A., Malpica N., Romero E. Single-image super-resolution of brain MR images using overcomplete dictionaries. *Medical Image Analysis*, 2013, vol. 17, no. 1, pp. 113–132. https://doi.org/10.1016/j.media.2012.09.003

Литература

1. Chen Q., Huang J., Feris R., Brown L.M., Dong J., Yan S. Deep domain adaptation for describing people based on fine-grained clothing attributes // Proc. of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 5315–5324. https://doi.org/10.1109/cvpr.2015.7299169
2. Denton E.L., Chintala S., Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks // Advances in Neural Information Processing Systems. 2015. V. 28.
3. Cui Z., Chang H., Shan S., Zhong B., Chen X. Deep network cascade for image super-resolution // Lecture Notes in Computer Science. 2014. V. 8693. P. 49–64. https://doi.org/10.1007/978-3-319-10602-1_4
4. Farhadifard F., Abar E., Nazzal M., Ozkaramanh H. Single image super resolution based on sparse representation via directionally structured dictionaries // Proc. of the 22nd Signal Processing and Communications Applications Conference (SIU). IEEE, 2014. P. 1718–1721. https://doi.org/10.1109/siu.2014.6830580
5. Ahmed J., Memon R.A., Waqas M., Mangrio M.I., Ali S. Selective sparse coding based coupled dictionary learning algorithm for single image super-resolution // Proc. of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). 2018. P. 1–5. https://doi.org/10.1109/icomet.2018.8346357
6. Choi J.H., Kim J.H., Cheon M., Lee J.S. Deep learning-based image super-resolution considering quantitative and perceptual quality // Neurocomputing. 2020. V. 398. P. 347–59. https://doi.org/10.1016/j.neucom.2019.06.103
7. Dong C., Loy C.C., He K., Tang X. Image super-resolution using deep convolutional networks // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. V. 38. N 2. P. 295–307. https://doi.org/10.1109/TPAMI.2015.2439281
8. Dong C., Loy C.C., Tang X. Accelerating the super-resolution convolutional neural network // Lecture Notes in Computer Science. 2016. V. 9906. P. 391–407. https://doi.org/10.1007/978-3-319-46475-6_25
9. Ayas S., Ekinci M. Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation // Information Sciences. 2020. V. 512. P. 1264–1278. https://doi.org/10.1016/j.ins.2019.10.040
10. Gu S., Zuo W., Xie Q., Meng D., Feng X., Zhang L. Convolutional sparse coding for image super-resolution // Proc. of the IEEE International Conference on Computer Vision (ICCV). 2015. P. 1823–1831. https://doi.org/10.1109/iccv.2015.212
11. Dosovitskiy A., Springenberg J.T., Brox T. Learning to generate chairs with convolutional neural networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 1538–1546. https://doi.org/10.1109/cvpr.2015.7298761
12. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error // Proc. of the 4th International Conference on Learning Representations (ICLR). 2016.
13. Alec R., Luke M., Soumith C. Unsupervised representation learning with deep convolutional generative adversarial networks // Proc. of the International Conference on Learning Representations (ICLR). 2015. P. 1–16.
14. Dong C., Loy C.C., He K., Tang X. Learning a deep convolutional network for image super-resolution // Lecture Notes in Computer Science. 2014. V. 8692. P. 184–199. https://doi.org/10.1007/978-3-319-10593-2_13
15. Aharon M., Elad M., Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation // IEEE Transactions on Signal Processing. 2006. V. 54. N 11. P. 4311–4322. https://doi.org/10.1109/tsp.2006.881199
16. Rueda A., Malpica N., Romero E. Single-image super-resolution of brain MR images using overcomplete dictionaries // Medical Image Analysis. 2013. V. 17. N 1. P. 113–132. https://doi.org/10.1016/j.media.2012.09.003

17. Wang H., Jiang K. Research on image super-resolution reconstruction based on transformer. *Proc. of the 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 2021, pp. 226–230. https://doi.org/10.1109/aiid51893.2021.9456580
18. Liu H., Guo Q., Wang G., Gupta B.B., Zhang C. Medical image resolution enhancement for healthcare using nonlocal self-similarity and low-rank prior. *Multimedia Tools and Applications*, 2019, vol. 78, no. 7, pp. 9033–9050. https://doi.org/10.1007/s11042-017-5277-6
19. Liu J., Malekzadeh M., Mirian N., Song T.A., Liu C., Dutta J. Artificial intelligence-based image enhancement in PET imaging: Noise reduction and resolution enhancement. *PET Clinics*, 2021, vol. 16, no. 4, pp. 553–576. https://doi.org/10.1016/j.cpet.2021.06.005
20. Dabbaghchian S., Ghaemmaghami M.P., Aghagolzadeh A. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *Pattern Recognition*, 2010, vol. 43, no. 4, pp. 1431–1440. https://doi.org/10.1016/j.patcog.2009.11.001
21. Liew W.S., Tang T.B., Lin C.H., Lu C.K. Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches. *Computer Methods and Programs in Biomedicine*, 2021, vol. 206, pp. 106114. https://doi.org/10.1016/j.cmpb.2021.106114
22. Timofte R., De V., Van Gool L. Anchored neighborhood regression for fast example-based super-resolution. *Proc. of the IEEE International Conference on Computer Vision*, 2013, pp. 1920–1927. https://doi.org/10.1109/iccv.2013.241
23. Haris M., Shakhnarovich G., Ukita N. Deep back-projectinetworks for single image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, vol. 43, no. 12, pp. 4323–4337. https://doi.org/10.1109/tpami.2020.3002836
24. Niu B., Wen W., Ren W., Zhang X., Yang L., Wang S., Zhang K., Cao X., Shen H. Single image super-resolution via a holistic attention network. *Lecture Notes in Computer Science*, 2020, vol. 12357, pp. 191–207. https://doi.org/10.1007/978-3-030-58610-2_12
25. Lan R., Sun L., Liu Z., Lu H., Pang C., Luo X. MADNet: a fast and lightweight network for single-image super resolution. *IEEE Transactions on Cybernetics*, 2021, vol. 51, no. 3, pp. 1443–1453. https://doi.org/10.1109/tcyb.2020.2970104

17. Wang H., Jiang K. Research on image super-resolution reconstruction based on transformer // Proc. of the 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID). 2021. P. 226–230. https://doi.org/10.1109/aiid51893.2021.9456580
18. Liu H., Guo Q., Wang G., Gupta B.B., Zhang C. Medical image resolution enhancement for healthcare using nonlocal self-similarity and low-rank prior // Multimedia Tools and Applications. 2019. V. 78. N 7. P. 9033–9050. https://doi.org/10.1007/s11042-017-5277-6
19. Liu J., Malekzadeh M., Mirian N., Song T.A., Liu C., Dutta J. Artificial intelligence-based image enhancement in PET imaging: Noise reduction and resolution enhancement // PET Clinics. 2021. V. 16. N 4. P. 553–576. https://doi.org/10.1016/j.cpet.2021.06.005
20. Dabbaghchian S., Ghaemmaghami M.P., Aghagolzadeh A. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology // Pattern Recognition. 2010. V. 43. N 4. P. 1431–1440. https://doi.org/10.1016/j.patcog.2009.11.001
21. Liew W.S., Tang T.B., Lin C.H., Lu C.K. Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches // Computer Methods and Programs in Biomedicine. 2021. V. 206. P. 106114. https://doi.org/10.1016/j.cmpb.2021.106114
22. Timofte R., De V., Van Gool L. Anchored neighborhood regression for fast example-based super-resolution // Proc. of the IEEE International Conference on Computer Vision. 2013. P. 1920–1927. https://doi.org/10.1109/iccv.2013.241
23. Haris M., Shakhnarovich G., Ukita N. Deep back-projectinetworks for single image super-resolution // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. V. 43. N 12. P. 4323–4337. https://doi.org/10.1109/tpami.2020.3002836
24. Niu B., Wen W., Ren W., Zhang X., Yang L., Wang S., Zhang K., Cao X., Shen H. Single image super-resolution via a holistic attention network // Lecture Notes in Computer Science. 2020. V. 12357. P. 191–207. https://doi.org/10.1007/978-3-030-58610-2_12
25. Lan R., Sun L., Liu Z., Lu H., Pang C., Luo X. MADNet: a fast and lightweight network for single-image super resolution // IEEE Transactions on Cybernetics. 2021. V. 51. N 3. P. 1443–1453. https://doi.org/10.1109/tcyb.2020.2970104

Authors

Pooja Singh — Magister, Researcher, Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India, sc 57225030639, https://orcid.org/0000-0002-8603-5954, pujasingh0409@gmail.com

Dinesh Ganotra — PhD, Associate Professor, Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India, sc 6506229541, https://orcid.org/0000-0002-3720-8716, dinesh_ganotra@hotmail.com

Авторы

Сингх Пуджа — магистр, исследователь, Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия, sc 57225030639, https://orcid.org/0000-0002-8603-5954, pujasingh0409@gmail.com

Ганотра Динеш — PhD, доцент, Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия, sc 6506229541, https://orcid.org/0000-0002-3720-8716, dinesh_ganotra@hotmail.com

Received 23.01.2023
Approved after reviewing 12.06.2023
Accepted 24.07.2023

Статья поступила в редакцию 23.01.2023
Одобрена после рецензирования 12.06.2023
Принята к печати 24.07.2023