# Single-Image Super-Resolution Techniques: A Review

**Article** · April 2017



**1 author:**

Sabyasachi Moitra
Techno India University

# Single-Image Super-Resolution Techniques: A Review

**Sabyasachi Moitra**
Department of Computer Science & Engineering
RCC Institute of Information Technology, Kolkata, India

***Abstract***- *The objective of image super-resolution (SR) is to enhance the resolution of a given low-resolution (LR) image, which has always been a continuous ongoing process in image technology, through up-sampling, de-blurring, de-noising, etc. In order to restore an image into a high-resolution (HR) image correctly, it is necessary to infer high frequency components of a low-resolution image. In some applications like, video surveillance, forensic investigation, face recognition, medical diagnosis, satellite images and pattern recognition, it becomes essential to extract the useful information from the images. During this process, enlarging the image beyond a certain limit results in a blurred image with no peculiar information. Hardware limitations of sensors is one of the main cause behind this problem. Also, the main idea behind achieving the high-resolution images is not to tamper the observable quality of the image. Many super resolution techniques have been proposed to overcome the hardware limitations in order to achieve the best results. There are lots of techniques exist to increase the resolution of the input images. This paper provides comparative studies among them.*

***Keywords***- Super-resolution, single-image, interpolation-based, reconstruction-based, example-based, edge-preserving, Gaussian process regression, iterative Weiner filter.

## I. INTRODUCTION

The main objective of super-resolution is to estimate the high-resolution visual output of a corresponding low-resolution visual input, which can either be a low-resolution image (single-image) or a set of images (multi-image), for example, corresponding to frames in a video sequence. The goals range from providing better content visualization for traditional image processing application to achieving better visual recognition, including computer vision tasks. Image super-resolution is important in many applications of multimedia, such as playing a video on a higher-resolution screen. Due to some technical limitations in imaging devices and systems, like, the presence of optical distortions and lens blur, insufficient sensor sampling density and aliasing, motion blur due to low shutter speed, the presence of noise due to sensor limitations and lossy coding, super-resolution technique is actually needed. The high-resolution visual output can be obtained either by providing devices with excellent spatial resolution, at the cost of a very high market price of the imaging device or with the use of software-related tools. The former is achieved by some hardware-related tools which includes – reducing the pixel size (which unfortunately leads to an increasing appearance of shot noise as the amount of light captured by the device decreases), increasing the chip size to accommodate a larger number of pixel sensors (which unfortunately results in an increased capacitance), reducing the shutter speed (which leads to an increasing noise level), adoption of high-precision optics and sensors (which invariably results in an increase in the price of the device). The advantage of post-processing the captured visual data is that it allows us to balance computational and hardware costs. Thus, on one hand we may have a lower market price and, on the other we can work with contemporary imaging devices and systems. [4, 1, 2]

Fig. 1 shows an example of a single-image super-resolution, where enlarging a low-resolution image does not affect the image quality, i.e., no loss of information.

The single-image SR problem arises in a number of real-world applications [25], one of the applications is found in web pages with images. To shorten the response time of browsing such web pages, images are often shown in LR forms (thumbnail images). An enlarged, HR image is only shown if the user clicks on the corresponding thumbnail. This approach requires the HR image to be stored on the web server and downloaded to the user's machine on demand, thus requiring much storage space. To save the storage space and communication bandwidth, it would be desirable if the LR image is downloaded and then enlarged on the user's machine through a SR technique, or by embedding a SR technique within the web application , which will take the LR image as input, on clicking the thumbnail, and will produce the corresponding enlarged HR image as output, thus there is no need of storing the HR image on the web server.


Figure 1. An example of a single-image super-resolution

The main features of a super-resolution technique are image interpolation and image restoration. Image interpolation changes the dimension of an image and image restoration recovers a degraded image. In other words, image restoration tries to fix the image to get back to the real, true image, even if that image doesn't look good (unlike image enhancement which makes a picture look better, without regard to how it really truly should look). It might even be worse than what you started with but it is the actual true starting values before it got degraded. But it might look clearer and sharper. Thus, image super-resolution is a technique that restores the degraded image and also increases the size of the image. [1, 2, 65]

Increasing image resolution is found to be important in various image processing applications (microscopic image analysis [61, 64], documented text image [62], surveillance [70], MRI image [66], and other [67]).

As stated earlier, super-resolution methods are also used to recover a HR image from more than one LR input images (multiframe super-resolution). In multiframe super-resolution, the information contained on multiple low-resolution captures of the same scene is combined to estimate the high-resolution visual output. [4, 49]

In this paper, we make a review on some well-known approaches (or techniques) leading to single-image super-resolution. Section II reviews on the three major categories of single-image super-resolution approaches, section III makes a review on other single-image super-resolution approaches, section IV makes a comparative study among the single-image super-resolution techniques, and section V the conclusion.

## II. SINGLE-IMAGE SR APPROACHES

The super-resolution of a given single low-resolution input image can be achieved through various techniques, which are better than the other in terms of their individual performances. In this section we will briefly summarize their principles.

Based on the application point of view single-image super-resolution approaches can be classified into three major categories [3] – Interpolation-Based, Reconstruction-Based, and Example-Based. Fig. 2 depicts this classification.

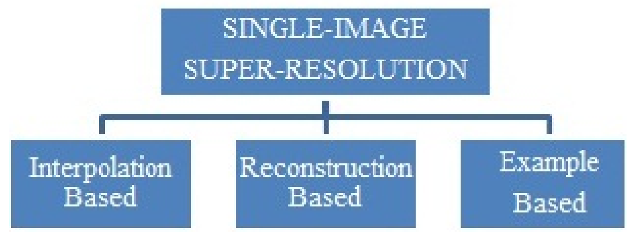


Figure 2. Approaches of single-image SR

### A. Interpolation-Based Super-Resolution

The simplest way to provide super-resolution is to apply interpolation on the sampled visual data acquired from the sensor. It is one of the early super-resolution algorithms based on resampling (a mathematical technique used to create a new version of the image with a different width and/or height in pixels). This approach, for example, which is present in digital cameras via the digital zoom, ultimately relies on the operations based on linear filtering. [4]

This technique has advantage of less computational complexity due to its simplicity and also real-time applications are possible. The low-frequency (LF) band of the high-resolution image will be more or less accurately reconstructed. [4]

However there are some limitations like, it is not possible to obtain the high-frequency information in the resized image due to low-pass behaviour of interpolation filters (bilinear, bicubic, Lanczos, illustrated in Fig. 3 and Table I), it is unable to find the missing spectral contents of the resampled image, and the results often appear over-smooth (softer) and have jagged (rough) artifacts along the edges, which can be overcome in example-based SR (discussed in the section C). [4, 50]

As pointed out in [54], the interpolation-based methods [6-9] are fast but the results are lack of fine details.

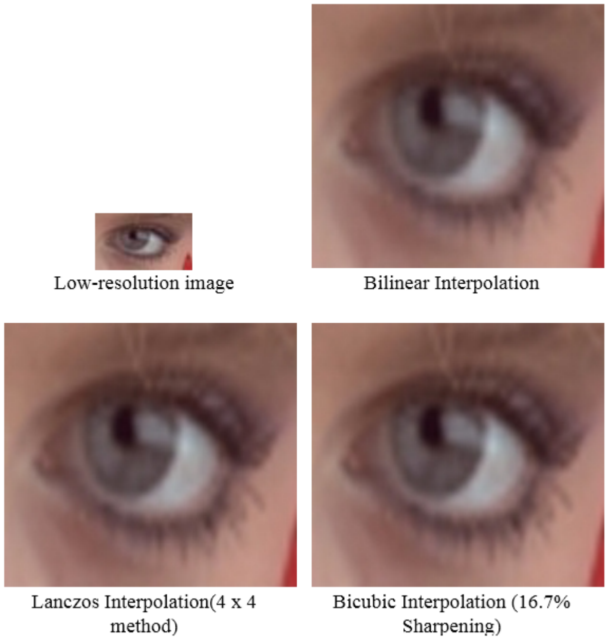


Figure 3. Behaviour of different interpolation filters

### B. Reconstruction-Based Super-Resolution

The reconstruction-based super-resolution approaches [3, 10-16, 17-20] apply various smoothness priors [21, 16, 3] and impose the constraint that when properly downsampled, the high-resolution image should reproduce the original low-resolution image. [54]

The reconstruction-based super-resolution approaches are formulated under the framework of

Table 1.Comparison of different interpolation filters (bilinear, bicubic, lanczos)

| Interpolation Filters | Description | Advantages | Limitations |
|---|---|---|---|
| Bilinear | Computes new pixels using linear interpolation [65], operates on 2-by-2 cell of pixels surrounding each new pixel location (during up-sampling) | Smoother than nearest neighbour method [5, 68] | Not sharp or smooth as bicubic interpolation (as shown in Fig. 3) |
| Bicubic | Computes new pixels using linear interpolation, operates on 4-by-4 cell of pixels surrounding each new pixel location (during up-sampling), has an extra parameter for sharpening the image during resampling | Recommended resampling method as it represents a good balance between accuracy and speed | No sharpening gives the smoothest result, excessive sharpening can produce halos and more jagged diagonals [5] (sharpening value of 16.7% is usually a good compromise, as shown in Fig. 3 and [5]) |
| Lanczos | Computes new pixels using *Lanczos-windowed sinc function* [65], operates on 4-by-4, 6-by-6 or 8-by-8 cell of pixels surrounding each new pixel location (during up-sampling) | The 6 x 6 and 8 x 8 methods slightly produce more accurate results than bicubic | The 4 x 4 method is nearly identical to bicubic method with no sharpening (as shown in Fig. 3), the 6 x 6 and 8 x 8 methods take a little longer time to compute |

enforcing a reconstruction constraint and imposing a prior knowledge on high-resolution images. The reconstruction constraint requires that the high-resolution image via smoothing and down-sampling should be as close as possible to the low-resolution image. [50]
In these methods, the imposed prior is not true for arbitrary images and many ringing artifacts may appear in the high-resolution image, which is one of the limitations of the reconstruction-based super-resolution technique. [50]

### C. Example-Based Super-Resolution

The second, and arguably one of the most successful, alternative is to estimate the high-resolution version of a low-resolution single-image by exploiting examples. The algorithms of example-based super-resolution problems are based on machine learning models exploiting available examples. [4]

For example, we are given a low-resolution single-image of a landscape with the lower part of the image showing the sea and the upper part showing the blue sky with some clouds. If we are having a database of high-resolution sea and sky images, which we can, for example, downscale to the resolution of our visual input, look for the low-resolution examples that best resembles or matches the two parts of our given image and reconstruct the corresponding high-resolution version of the given image with a montage (technique of selecting, editing and piecing together separate sections of film to form a continuous whole) of the sea and sky high-resolution parts. [4]

Fig. 4 illustrates the above example diagrammatically.

The example-based super-resolution technique can be broadly classified into two categories [4] – parametric and non-parametric methods. The non-parametric methods can be further classified into internal learning and external learning methods. Internal learning method can be classified into high-frequency transfer and neighbor embedding approaches, whereas external learning method can be classified into sparse coding, anchored regression, regression trees and deep learning approaches.

Fig. 5 depicts this hierarchical classification of example-based super-resolution approach.

However there are some limitations like, due to insufficient training examples, high-frequency artifacts often appear in the result, rely on training set. [50, 32]

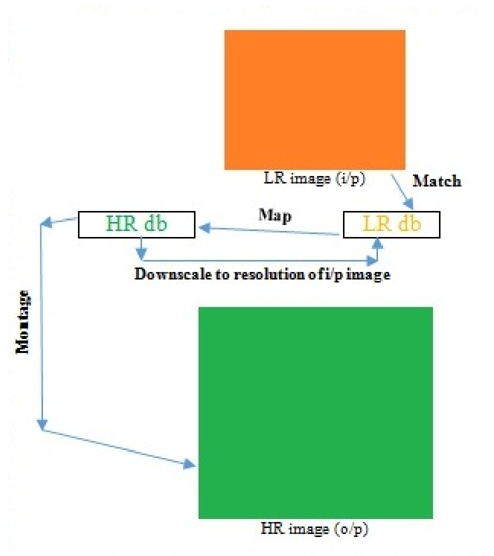


Figure 4. General idea behind example-based super-resolution

Example-based methods are also termed as learning-based approaches, as super-resolution of an image is being carried out by learning examples. The category of learning-based approaches [22-33, 49] is also known as image hallucination techniques [27].

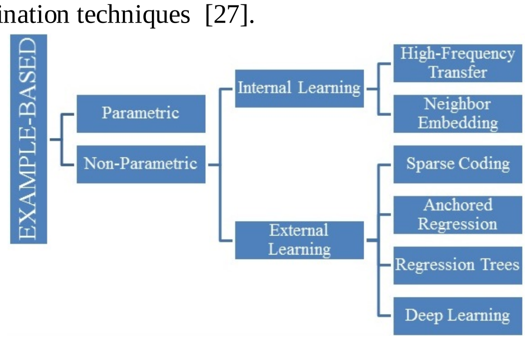


Figure 5. Classification of Example-Based Super-Resolution

**Parametric vs. Non-Parametric Methods**

Parametric super-resolution methods model the super-resolution problem with mapping functions which are controlled by a relatively compact amount of parameters. These parameters are learned from the examples that do not necessarily come from the input image. Parametric models are comparatively more powerful than non-parametric models in terms of efficiency, as, less data are needed to estimate the models. [4]

Non-parametric models differ from the parametric models in that the model structure is not specified a priori. It is directly determined from the available training data. The term non-parametric does not imply that they completely lack parameters, but rather that the number and nature of the parameters are flexible and can depend upon the training data. The non-parametric models have a very desirable property of making fewer assumptions. The particular feature of a non-parametric model is that the input low-resolution image is decomposed into overlapping patches, and the reconstructed high-resolution image is obtained by combining the contributions of the computed overlapping output patches. [4]

**Internal vs. External Learning**

The main idea behind the methods under internal learning category is that the patch examples are obtained directly from the input image (i.e., examples are internal, extracted from the input image), exploiting the cross-scale self-similarity property [] of natural images. This learning method has advantages of implicit adaptivity to the image contents, robust to noise, and can able to obtain better upscale results (including sharper edges and better preservation of texture). However there are some limitations like, this method of learning requires iterative application for providing large up-scaling factors, the cross-scale self-similarity property of natural images degrades with large scale differences, and the computational cost is high. [4]

In case of external learning methods the examples are external, that is, extracted from an external database. The most relevant research aspect to external learning methods is that of ensuring generalization with offline-trained models that provide efficient inference pipelines. [4]

**1. High-Frequency Transfer**

In the high-frequency transfer technique a coarse (rough) version of the high-resolution image is initially obtained through the interpolation method. Since the obtained coarse version lacks high-frequency components, we can exploit cross-scale self-similarity by computing the coarse version of the input low-resolution image. The difference between the original input image and its coarse version is the high-frequency band that we need to transfer to the high-resolution image. For each patch in the high-resolution coarse image, we find the closest patch in the low-resolution coarse image. Then the high-frequency band or missing detail, for this patch, can be assumed to be the patch at the same position in the low-resolution, high-frequency band. [4]

Let,

$$Xl = Hs * (Y \uparrow s)$$

where, Xl is the up-scaled version of the input low-resolution image Y containing low-frequency band of the spectrum, Hs is the linear interpolation kernel and ↑s is the up-scaling factor.

The input image Y can be analysed into two separate bands by using the same linear interpolation kernel employed for up-scaling. Thus,

$$Yl = Hs * Y \text{ and } Yh = Y - Yl$$

where, Yl is the low-frequency band and Yh is the high-frequency band of the input image Y.

Thus, we are generating pairs of low-frequency references (in Yl) and their corresponding high-frequency examples (inYh). It should be noted that Yl has the same normalized bandwidth as Xl and, that the cross-scale self-similarity property is also present between these two images.

Let, xl,i be a patch within Xl. We look for the best matching patch (say yl,i) in Yl. The location of yl,i is also the location of the corresponding high-frequency example yh,j.

The local estimate of the high-frequency band corresponding to a patch xl,i can be directly regressed as xh,i = yh,j.

The resultant high-frequency band Xh may contain some low-frequency spectral components. In order to improve the spectral compatibility between Xl and Xh, we can subtract the low-frequency spectral component from Xh before adding it to the low-frequency band to compose the reconstructed image,

$$X = Xl + Xh - Hs * Xh$$

Thus we can say that, after analysing the input image into a coarse or low-frequency band and a corresponding fine detail or high-frequency counterpart, the high-frequency is transferred for each patch in the HR image to produce SR image. [4]

The high-frequency transfer technique is an unsupervised learning [69] method. Fig. 6 shows the general idea behind high-frequency transfer method.

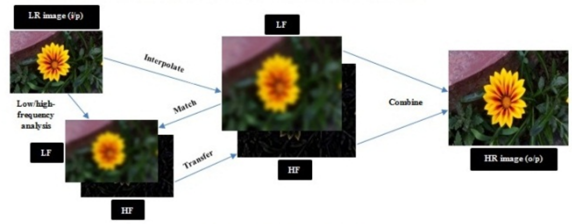


Figure 6. High-Frequency Transfer technique

## 2. Neighbour Embedding

The general idea behind the neighbour embedding framework is that each input data vector (i.e., a low-resolution patch) can be described as a linear combination of its nearest neighbors on the natural image manifold of low-resolution patches. [4]

In the neighbour embedding single-image super-resolution approach the low-resolution input image is downscaled. The large patches in the original input image and the small patches in the downscaled version of the input image are used to create databases of the related HR and LR patches respectively. Each small patch in the original image is described as a weighted combination of its nearest neighbors in the low-resolution database, and the large high-resolution patch is computed as the same weighted combination from the corresponding patches in the high-resolution database. [4]

Let, Y be a low-resolution input image for which we want to estimate a high-resolution version X with an up-scaling factor s.

The input image is first downscaled by the same factor s to generate the corresponding downscaled version of the low-resolution input image (say Yd),

$$Yd = (Y * Hd) \downarrow s$$

where, Hd is the downscaling filter.

The input image is divided into a set of overlapping patches {yp}. For each patch yp, the set of nearest neighbors (NN) in Y' is {y'p,i} (i = 1, …, k), where Y' is the database created from the small patches in the downscaled version of the input image, Yd.

Once the weights wp,i are defined following the equation,

$$wp,i = yp \,.\, y'p,i \,/\, \textstyle\sum k\ j = 1\ yp \,.\, y'p,j$$

the corresponding output patch xp can be reconstructed using the following equation,

$$xp = \textstyle\sum k\ i = 1\ wp \,.\, x'p,i$$

where, {x'p,i} (i = 1, …, k) are the high-resolution counterparts to {y'p,i} in X' (X' is the database created from large patches in the original input image). [4, 25]
Fig 7 shows the general idea behind the neighbor embedding approach, in a block diagram format, inherited from the example as pointed out in [4].

The neighbor embedding technique is also an unsupervised learning method. As pointed out in [4] the neighbor embedding framework can be used for both internal and external learning, and the external learning approach can perform better than the internal learning one.

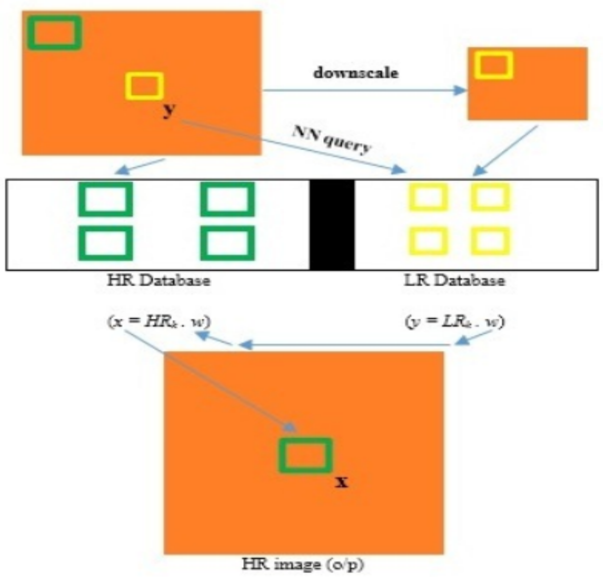


Figure 7. Neighbor Embedding approach

## 3. Sparse Coding

The sparse coding based single-image super-resolution approach is one of the external learning methods. It is an improved version of neighbor embedding technique. This technique provides a compact or sparse dictionary that can be obtained by unsupervised learning on either the low resolution examples (using the K-SVD algorithm [34]) or on both the low and high resolution pairs.

During the training stage, a large set of patches extracted from the natural image and its downscaled (low-resolution) version are used to generate two coupled sparse dictionaries (high-resolution and low-resolution dictionaries respectively). [4]

During the inference stage, decompose each LR image patch as a sparse linear combination of base patches or atoms from the over-complete LR dictionary. Over-complete dictionaries are dictionaries where the number of exemplars or manifold samples is much larger than the dimensionality of the underlying sampled manifold. Same coefficients are used with the atoms in the HR dictionary to generate the HR patch. [4]

Let, Y be a low-resolution input image for which we want to estimate a high-resolution version X with an up-scaling factor s.

The input image is first downscaled by the same factor s to generate the corresponding downscaled version of the low-resolution input image (say Yd),

$$Yd = (Y * Hd) \downarrow s$$

where, Hd is the downscaling filter.

During the training stage [4, 33, 31] a large set of patches are being extracted from both the input image and its downscaled version. These extracted patches are being used to generate two coupled sparse dictionaries, one for high-resolution Dh, and another one for low-resolution Dl.

For each patch of the input LR image, a sparse decomposition is computed as a linear combination of a small number of entries in a low-resolution dictionary Dl.
The high-resolution reconstruction is obtained by applying the same mixing coefficients onto the corresponding high-resolution dictionary Dh,

$$x = Dh \, .\alpha$$

where, x is the computed HR patch, Dh is the over-complete dictionary containing HR patches that is being extracted from the input image, as previously mentioned, and α is the sparse representation of x with a small amount of non-zero entries. In order to recover x, the sparse representation α is computed from the observable low-resolution patches y with respect to a dictionary containing the corresponding low-resolution patches Dl and possibly using a feature transformation. [4, 33, 31, 2]

Thus, the high-resolution image can be computed by combining the contributions of each patch.
Fig. 8 illustrates the sparse coding SR approach.

The sparse coding technique follows the principles of neighbor embedding approach with the additional constraint of a compact and optimized dictionary (obtained through training process , unlike neighbor embedding technique [4, 25]) that provides robust nearest-neighbor decomposition. Further specialization can be achieved by selecting the training images that are more representative of the input image and compute a compact dictionary.

However there is a limitation of high computational cost.

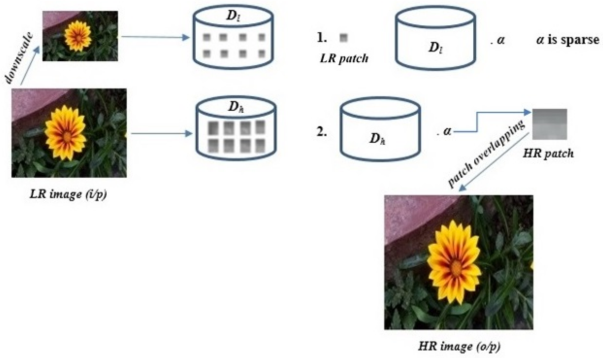


Figure 8. Sparse Coding approach

## 4. Anchored Regression

The anchored regression based single-image super resolution technique is also one of the external learning methods. It overcomes the high computational cost of the neighbour embedding and sparse coding techniques.

In the anchored regression approach, an external database composed of a low-resolution dictionary and a set of linear regression matrices to map the low-resolution examples to their high-resolution counterparts. During the inference stage, the input image is divided into patches; each patch is matched to the nearest atom (say α) in the low-resolution dictionary (Dl) and the corresponding regression matrix (say Rα) is used to generate the high-resolution patch. Each linear regression matrix is anchored to each patch or atom in the sparse optimized dictionary, and is built by using the high and low resolution parts of the most similar atoms within the dictionary. [4, 37]

Fig. 9 depicts the basic idea behind the anchored regression approach.

One of the main advantages of anchored regression over sparse coding approach is that the running time is greatly reduced by removing the sparsity constraint from the inference stage in sparse coding technique [37, 4].
Further improvement of the approach can be achieved by using all the training examples for computing the regression matrices, and also by using efficient anchor search schemes, as pointed out in [4].

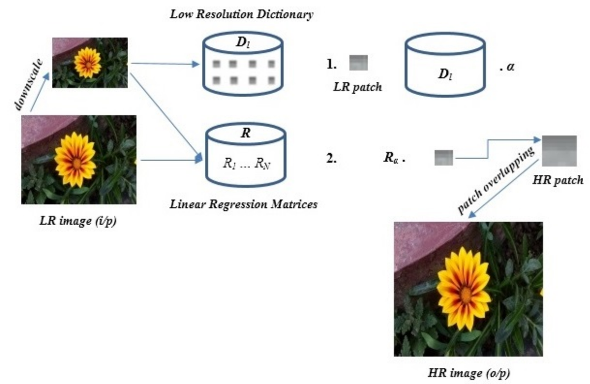


Figure 9. Anchored Regression approach

## 5. Regression Trees

The most affecting practical bottleneck of anchored regression approach is the exhaustive search stage. One of the possible solutions to this problem is the regression trees and forests based single-image super-resolution technique.

In this approach the input image is divided into patches. Each patch traverses the tree from root (top) node to the most suitable leaf node, and the corresponding regression model is used to generate the high-resolution patch. [4]

Fig. 10 illustrates the general idea behind the regression trees and forests based single-image super-resolution approach.
As pointed out in [4], this technique performs an unsupervised hierarchical clustering to learn the patch subspace. This strategy allows determining the clusters of similar patches without the need for an additional dictionary learning method, and also provides an intrinsic logarithmic-time search during the inference stage.

However there is a limitation of high memory requirements for storing the regression parameters.

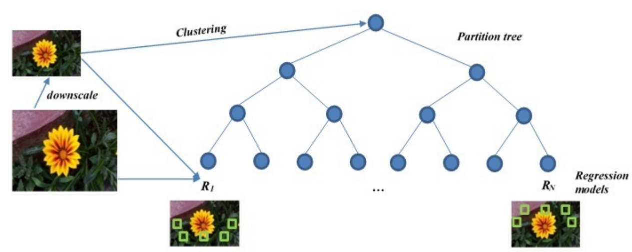


Figure 10. Regression Trees and Forests based SR approach

## 6. Deep Learning

Deep learning is one of the current alternatives with supervised learning approach, for single-image super-resolution, based on deep convolution neural networks (CNN). The main idea behind this technique is to use the power of back-propagation algorithm [69] in order to learn the hierarchical representations that allow for minimizing the error at the end of the network.

The adoption of convolution neural networks is pervasive, i.e., spreading widely in image processing, such as image classification [40, 41] and computer vision applications, such as object detection [43, 46, 47], face recognition [45], and pedestrian detection [44]. The main idea behind the convolution neural networks is to reduce the number of parameters by exploiting the properties of locality and stationarity [4] present in natural images. Another advantage of CNN is that they accept inputs of different sizes (unlike fully connected networks, where the number of inputs and outputs are fixed by the network architecture). Several factors are of central importance in this progress, viz. efficient training implementation on modern and powerful GPUs [41], proposal of the Rectified Linear Unit (ReLU) [42, 4] which makes convergence much faster while still presents good quality [41], thus improves learning, and easy access to large quantity of data [39] for training larger models.

CNN directly learns an end-to-end mapping between the low and high-resolution images, with little pre or post-processing beyond the optimization.

Let, Y be a single low-resolution input image which is first upscaled to the desired size (say, s) Yu using bicubic interpolation Hs, which is the only pre-processing being performed,

$$Yu = Hs * (Y \uparrow s)$$

The patches from the interpolated image Yu are extracted (each extracted patch is represented as a high-dimensional vectors. These vectors comprise a set of feature maps) (patch feature extraction). Each high-dimensional vector is non-linearly mapped to another high-dimensional vector (each mapped vector represents a high-resolution patch. These vectors comprise another set of feature maps) (non-linear mapping). The final high-resolution image X is obtained by overlapping the high-resolution patches (reconstruction). [4, 38]

Fig. 11 depicts the general idea behind deep learning based super-resolution approach using convolution neural network (2-layer network with a 4-neuron hidden layer with ReLU non-linear activations, a 1-neuron output layer without non-linear activation, and 5x5 spatial interactions (an additional bias parameter for each neuron is implicitly assumed)).

As pointed out in [4, 38], the traditional sparse coding based super-resolution approach can also be viewed as a deep convolutional neural network.

An interesting fact about deep learning is that the design is being inspired by biological systems . [4, 69]

As pointed out in [4], the main advantage of deep learning in comparison with the other discussed frameworks is the power to determine the hierarchical descriptions of the visual data that are directly learned from the data.

One of the major shortcomings in deep learning approach is that the fine tuning of all the parameters in the network takes a considerably large amount of time than the classical machine learning. [4]

The deep learning framework is able to solve the general learning problems, which strongly contrasts with the former machine learning models that are task-specific and present a loose tie with biological systems. [4]

By exploiting the concept of deep learning neural network (ANN) based super-resolution [48] can be achieved.

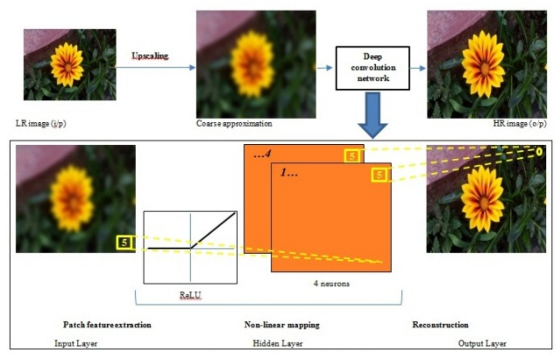


Figure 11. Deep Learning based SR approach

## III. OTHER SINGLE-IMAGE SR APPROACHES

Beside the above mentioned approaches researchers have proposed many other single-image super-resolution techniques which can able to overcome the limitations faced in the previous approaches. Here we are reviewing some of them in brief.

### A. Edge-Preserving Super-Resolution

Due to some limitations found in the previous super-resolution approaches [4], researchers have proposed this edge-preserving or edge-based super-resolution technique to overcome those limitations.

In this approach an up-sampling method is used to obtain the initial high-resolution image. This up-sampling method is performed by combining the soft-edge and hard-edge constraints. The soft-edge constraint is enforced via bilateral filtering [51]. The intuition of bilateral filtering is to smooth the image using pixels which are close both in the spatial domain [63, 68] and the intensity domain [68], i.e., it smoothes the image while preserving edges. The hard-edge constraint is enforced via the mean shift image segmentation algorithm [52]. For enhancing the strong edges in the obtained initial up-sampling result the complex shock filter [53] is being used, and an intermediate high-resolution image is obtained. Since the obtained intermediate high-resolution image doesn't satisfy the reconstruction constraint, the reconstruction error is minimized by back-projection [11], i.e., the final result is obtained by back-projection . [50]

It has been seen that, in this technique, after obtaining the initial up-sampling result, using the up-sampling method, the complex shock filter is being used to enhance the strong edges in the high-resolution image instead of imposing prior knowledge on the high-resolution image, which is being used in most of the previous methods [15, 13, 16, 3], due to some disadvantages – the prior is not suitable for all images, as it may introduce many artifacts to the high-resolution image or the result appears over-smooth, and the process is computationally expensive. [50]

As pointed out in [50], the technique has the following advantages – it works without any prior, thus can handle arbitrary images, and it can preserve edges excellently.

**B. Gaussian Process Regression based SR**

The Gaussian Process Regression (GPR) based single-image super-resolution approach produces a high-resolution image from a single low-resolution image without any external training set. It is a two-stage approach, where the first stage recovers an HR image based on the structural information extracted from the LR image, and the second stage refines the obtained HR image by learning from the training set constructed by the LR image itself. [54]

In the first stage, the researchers proposed a framework to learn the reverse mapping from a low-resolution image to the corresponding high-resolution image pixelwise, depending on the local structures defined by each pixel's neighborhood. Pixels in the high-resolution image are first estimated by their nearest neighbors in an initial upsampled image via Gaussian Process Regression . The result obtained from the first stage is deblurred by learning from the LR/HR patches obtained from its downsampled version and the LR input in the second stage, and the corresponding HR image is constructed. [54]

Fig. 12 depicts the block diagram of the GPR framework.

As pointed out in [54], GPR based SR can extract adequate information contained in a single low-resolution image to produce a high-resolution image with sharp edges, compared to other edge-directed and example-based [4, 22-33, 49] single-image super-resolution approaches.

The GPR technique can also be used in other SR algorithms [56].

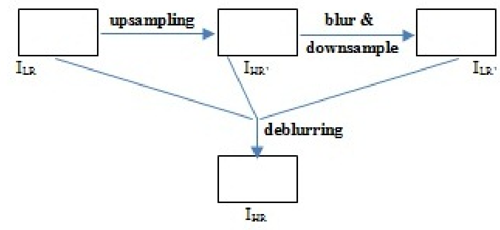


Figure 12. Block diagram of GPR framework

**C. SR using Iterative Wiener Filter**

Besides the above two approaches, researchers have also proposed an SR reconstruction framework which is based on classic Wiener filter approach [58, 65] in an iterative manner for reconstruction and feedback to correct the reconstruction errors. [56, 57]

The iterative Wiener filter can simultaneously perform interpolation and restoration by using the non-local means to directly model the correlation between the desired high-resolution image and the observed low-resolution image with successful results. The decay speed of the correlation function is controlled within the iteration loops of the Wiener filter. [57]

During the iterations, the image is decomposed into patches with similar intensities, at initial iterations, by putting the correlation function to decay quickly. As the Wiener filter iterates, the patches will be connected naturally by lowering the decay speed of the correlation function. As a result, the natural image structures will be produced. [57]

As pointed out in [57], the proposed algorithm uses a mechanism to iteratively reconstruct the natural image structures by updating the varying correlations as compared to the other available correlation-based algorithms [60, 54].

## IV. DISCUSSION

Table II shows the comparative study among the single-image super-resolution techniques (based on their PSNR (Peak Signal-to-Noise Ratio), values for the images from the set5 dataset at magnification factor 3). [4]

PSNR, is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity (degree of exactness) of its representation, i.e., how the original image is affected by added noise. It measures the quality of an image after the reconstruction. High PSNR value indicates a good quality image whereas a low PSNR value indicates a bad quality image. [65]

TABLE 2: Comparative study among the single-image super-resolution techniques (bicubic (BC), neighbor embedding (NE), sparse coding (SC), anchored regression (AR), regression trees (RT), deep learning (DL))

| BC | NE | SC | AR | RT | DL (CNN) |
|---|---|---|---|---|---|
| 30.39 | 27.72 | 31.90 | 31.92 | 32.57 | 32.39 |

## V. CONCLUSION

In this paper, we have reviewed some of the well-known single-image super-resolution techniques along with their working principles, advantages and limitations.

The goal of single-image super-resolution is to obtain a high-resolution image from a single low-resolution image. The single-image super-resolution approaches can be broadly classified into three classes - Interpolation-Based, Reconstruction-Based, and Example-Based or Learning-Based methods. The interpolation-based methods use different interpolation techniques (bilinear, bicubic, Lanczos), and are fast but the results are lack of fine details. The reconstruction-based methods apply various smoothness priors and impose the constraint that when properly downsampled, the high-resolution image should reproduce the original low-resolution image. In the example-based methods estimation of the high-resolution version of a low-resolution single-image is done by exploiting examples. High-frequency transfer, neighbor embedding, sparse coding, anchored regression, regression trees and deep learning are some of the example-based super-resolution approaches.

Edge-preserving, Gaussian process regression, iterative Weiner filter are some of the single-image super-resolution techniques proposed by the researchers other than the above mentioned approaches.

Single-image super-resolution is one of the challenging problems in the field of image processing, and the researchers are still now working on it.

## VI. ACKNOWLEDGEMENT

The author would like to acknowledge his project guide Mr. Pramit Ghosh, Asst. Prof., Computer Science & Engineering, RCC Institute of Information Technology, for his immense support, and also to his college for providing the infrastructure to carry on his work. The author is also very much grateful to all the researchers for sharing their ideas through different books, journals, papers, video lectures which had very much helped him in writing this review paper.

## REFERENCES

[1] Nirali Harsoda, Namarata Joshi. “A Review on Image Super Resolution Techniques”. International Journal of Engineering Research and Technology (IJERT), Vol. 3, Issue 2, February 2014.

[2] A. J. Shah, S. B. Gupta. “Image super resolution - A Survey”. 1st international conference on emerging technology and trends in electronics, communication and networking, IEEE, 2012.

[3] J. Sun, Z. Xu, and H. Y. Shum. “Image super-resolution using gradient profile prior”. In CVPR, 2008.

[4] Jordi Salvador. “Example-Based Super Resolution”. Academic Press, 2017.

[5] Jonathan Sachs. “Image Resampling”. Digital Light and Color, 2011.

[6] H. S. Hou, H. C. Andrews. “Cubic splines for image interpolation and digital filtering”. IEEE Trans. Acoustics, Speech & Signal Proc., ASSP-26:508–517.

[7] J. Allebac, P. W. Wong. “Edge-directed interpolation”. In Proc. ICIP, 1996.

[8] X. Li, M. T. Orchard. “New edge-directed interpolation”. IEEE Transactions on Image Processing, 10:1521–1527, 2001.

[9] W. S. Tam, C. W. Kok, W. C. Siu. “A modified edge directed interpolation for images”. Journal of Electronic Imaging, 19(1), 013011:1–20, 2010.

[10] S. Baker, T. Kanade. “Limits on super-resolution and how

to break them". IEEE Trans. on PAMI, 24:1167–1183, 2002.

[11] M. Irani, S. Peleg. "Motion analysis for image enhancement: Resolution, occlusion and transparency". Journal of Visual Communication and Image Representation, 4(4):324–335, 1993.

[12] B. S. Morse, D. Schwartzwald. "Image magnification using level-set reconstruction". In CVPR, volume 1, pages 333–340, 2001.

[13] Q. Shan, Z. Li, J. Jia, C. K. Tang. "Fast image/video upsampling". ACM Transactions on Graphics (SIGGRAPH ASIA), 2008.

[14] K.W. Hung, W. C. Siu. "New motion compensation model via frequency classification for fast video super-resolution". In ICIP, 2009.

[15] S. Dai, M. Han, W. Xu, Y.Wu, Y. Gong. "Soft edge smoothness prior for alpha channel super resolution". In Proc. CVPR, 2007.

[16] R. Fattal. "Upsampling via imposed edges statistics". In ACM SIGGRAPH, 2007.

[17] Jian Sun, ZongbenXu, Heung-Yeung Shum. "Gradient Profile Prior and Its Applications in Image Super-Resolution and Enhancement". IEEE Trans. on Image Processing, vol.20, no.6, pp.1529-1542, June 2011.

[18] M. Protter, M. Elad. "Super Resolution With Probabilistic Motion Estimation". IEEE Trans. on Image Processing, vol.18, no.8, pp.1899-1904, Aug. 2009.

[19] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, A. K. Katsaggelos. "SoftCuts: a soft edge smoothness prior for color image superresolution". IEEE Trans. on Image Process., vol. 18, no. 5, pp. 969-981, May 2009.

[20] A. Marquina, S. J. Osher. "Image super-resolution by TV-regularization and Bregman iteration". J. Sci. Comput., vol. 37, pp. 367-382, 2008.

[21] H. A. Aly, E. Dubois. "Image up-sampling using total variation regularization with a new observation model". IEEE Trans. on IP, 14(10):1647–1659, 2005.

[22] C. Liu, H. Shum, W. Freeman. "Face hallucination: Theory and practice". 75(1):115–134, 2007.

[23] X. Gao, K. Zhang, D. Tao, X. Li. "Joint Learning for Single Image Super-resolution via Coupled Constraint". IEEE Trans. on Image Processing, Vol. 21, Issue 2, pp. 469 – 480, 2012.

[24] W. T. Freeman, T. R. Jones, E. C. Pasztor. "Example-based super-resolution". IEEE Computer Graphics and Applications, pp. 56 – 65, Vol. 22, Issue 2, 2002.

[25] Hong Chang, Dit-Yan Yeung, YiminXiong. "Super-Resolution Through Neighbor Embedding". Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, 2004.

[26] Q.Wang, X. Tang, H. Y. Shum. "Patch based blind image super-resolution". In Proc. ICCV, 2005.

[27] W. Freeman, E. Pasztor. "Learning low-level vision". IJCV, 40(1):25–47, 2000.

[28] J. Sun, N.-N. Zheng, H. Tao, H.-Y. Shum. "Image hallucination with primal sketch priors". In Proc. CVPR, 2003.

[29] X. Wang, X. Tang. "Hallucinating face by eigen transformation". IEEE Trans. on SMC, Part C, 35(3):425–434, 2005.

[30] W. Liu, D. Lin, X. Tang. "Hallucinating faces: Tensor patch super-resolution and coupled residue compensation". In Proc. CVPR, 2005.

[31] J. Yang, J. Wright, T. Huang, Y. Ma. "Image super-resolution as sparse representation of raw image patches". In Proc. CVPR, 2008.

[32] Z. Lin, J. He, X. Tang, C.-K. Tang. "Limits of learning-based superresolution algorithms". In Proc. ICCV, 2007.

[33] Jianchao Yang, John Wright, Thomas Huang, Yi Ma. "Image Super-Resolution via Sparse Representation". IEEE Transactions on Image Processing, pp. 2861 – 2873, Vol. 19, Issue 11, 2010.

[34] M. Aharon, M. Elad, A. Bruckstein. "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation". IEEE Trans. on Sig. Proc., 54(11), 2006.

[35] Lei Zeng, Xiaofeng Li, Jin Xu. "An improved joint dictionary training method for single image super resolution". pp. 89 – 92, 2012 International Conference

on Computational Problem-Solving (ICCP), 2012.

[36] Jianchao Yang, Zhaowen Wang, Zhe Lin, Scott Cohen, Thomas Huang. “Couple Dictionary Training for Image Super-resolution”. IEEE Transactions on Image Processing, pp. 3467 – 3478, Vol. 21, Issue 8, 2012.

[37] RaduTimofte, Vincent De, Luc Van Gool. “Anchored Neighborhood Regression for Fast Example-Based Super-Resolution”. 2013 IEEE International Conference on Computer Vision, pp. 1920 – 1927, 2013.

[38] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang. “Image Super-Resolution Using Deep Convolutional Networks”. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 295 – 307, Vol. 38, Issue 2, 2016.

[39] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei. “ImageNet: A large-scale hierarchical image database”. IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255 (2009).

[40] K. He, X. Zhang, S. Ren, J. Sun. “Spatial pyramid pooling in deep convolutional networks for visual recognition”. European Conference on Computer Vision, pp. 346–361 (2014).

[41] A. Krizhevsky, I. Sutskever, G. Hinton. “ImageNet classification with deep convolutional neural networks”. Advances in Neural Information Processing Systems. pp. 1097–1105 (2012).

[42] V. Nair, G.E. Hinton. “Rectified linear units improve restricted Boltzmann machines”. International Conference on Machine Learning. pp. 807–814 (2010).

[43] W. Ouyang, P. Luo, X. Zeng, S. Qiu, Y. Tian, H. Li, S. Yang, Z. Wang, Y. Xiong, C. Qian, et al.: “Deepid-net: multi-stage and deformable deep convolutional neural networks for object detection”. arXiv preprint arXiv:1409.3505 (2014).

[44] W. Ouyang, X. Wang. “Joint deep learning for pedestrian detection”. IEEE International Conference on Computer Vision. pp. 2056–2063 (2013).

[45] Y. Sun, Y. Chen, X. Wang, X. Tang. “Deep learning face representation by joint identification-verification”. Advances in Neural Information Processing Systems. pp. 1988–1996 (2014).

[46] C. Szegedy, S. Reed, D. Erhan, D. Anguelov. “Scalable, highquality object detection”. arXiv preprint arXiv:1412.1441 (2014).

[47] N. Zhang, J. Donahue, R. Girshick, T. Darrell. “Part-based RCNNs for fine-grained category detection”. European Conference on Computer Vision. pp. 834–849 (2014).

[48] N. Kumar, P.K. Deswal, J. Mehta. “Neural Network based single image super Resolution”. In Proceedings of 11th Symposium on Neural Network Applications in Electrical Engineering, pp. 213 – 218, 20-22 Sept. 2012, Belgrade.

[49] D. Glasner, S. Bagon, M. Irani. “Super-Resolution from a single image”. In Proceedings of the IEEE International Conference on Computer Vision (ICCV09). pp. 349–356, 2009.

[50] Qiang Zhou, Shifeng Chen, Jianzhuang Liu, Xiaoou Tang. “Edge-preserving single image super-resolution”. MM '11: Proceedings of the 19th ACM international conference on Multimedia, November 2011.

[51] C. Tomasi, R. Manduchi. “Bilateral filtering for gray and color images”. In Proc. ICCV, 1998.

[52] D. Comaniciu, P. Meer. “Mean shift: a robust approach toward feature space analysis”. IEEE Trans. on PAMI, 24(5):603–619, May 2002.

[53] G. Gilboa, N. A. Sochen, Y. Y. Zeevi. “Regularized shock filters and complex diffusion”. Lecture Notes in Computer Science, pages 399–413, 2002.

[54] He He, Siu Wan-Chi. “Single image super-resolution using Gaussian process regression”. In Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.449-456, 20-25 June, 2011.

[55] C. E. Rasmussen, C. K. I. Williams. “Gaussian Processes for Machine Learning”. The MIT Press, 1 edition, 2006.

[56] Aiman M. Mousa, XinboGao, Mohamed Elmahalawy. “Single Image Super-resolution Reconstruction with Neural Network and Gaussian Process Regression”. ICIMCS '14: Proceedings of International Conference on Internet Multimedia Computing and Service, June 2014.

[57] Kwok-Wai Hung and Wan-Chi Siu. “Single-Image Super-Resolution Using Iterative Wiener Filter”. In Proceedings

of IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP 2012), pp. 1269-1272, 25-30, March, 2012, Kyoto, Japan.

[58] Lim, S. Jae. “Two-Dimensional Signal and Image Processing”. Englewood Cliffs, NJ, Prentice Hall, 1990, p. 548, equations 9.44 – 9.46.

[59] A. Buades, B. Coll, J. M. Morel. “A non-local algorithm for image denoising”. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2, pp. 60 – 65, 2005.

[60] R. Hardie. "A Fast Image Super-Resolution Algorithm Using an Adaptive Wiener Filter," IEEE Trans. on Image Processing, vol.16, no.12, pp.2953-2964, Dec. 2007.

[61] Pramit Ghosh, et al. "Automatic white blood cell measuring aid for medical diagnosis." Process Automation, Control and Computing (PACC), 2011 International Conference on. IEEE, 2011.

[62] Soma Datta, Nabendu Chaki, Sankhayan Choudhury. "Information Density Based Image Binarization for Text Document Containing Graphics." IFIP International Conference on Computer Information Systems and Industrial Management. Springer International Publishing, 2016.

[63] Courtesy of TutorialsPoint - Digital Image Processing - Quick Guide. URL: http://www.tutorialspoint.com/dip/.

[64] Pramit Ghosh, Debotosh Bhattacharjee, Mita Nasipuri. "Blood smear analyzer for white blood cell counting: a hybrid microscopic image analyzing technique." Applied Soft Computing 46 (2016): 629-638.

[65] Courtesy of MathWorks - MATLAB and Simulink for Technical Computing. URL: http://www.mathworks.com/.

[66] Soma Datta, D. S. K. Bandyopadhyay. "Segmentation of MRI brain images for automatic diagnosis." International Journal of Innovative Research in Computer and Communication Engineering 1 (2013): 501-505.

[67] Pramit Ghosh, et al. "Child obesity control through computer game." Biomed. Eng. Res 2.2 (2013): 88-95.

[68] C. Gonzalez, E. Woods, L. Eddins. “Digital Image Processing Using Matlab”, 2e. Gatesmark Publishing, 2010.

[69] S. Rajasekaran, G. A. V. Pai. “Neural Networks, Fuzzy Logic, and Genetic Algorithm”. PHI.

[70] Soma Datta, Debotosh Bhattacherjee, Pramit Ghosh. "Path Detection of a Moving Object." International Journal of Recent Trends in Engineering 2.3 (2009): 37-39.