Solving Image Processing Critical Problems Using Machine Learning

Ajay Sharma, Ankit Gupta, and Varun Jaiswal

Abstract The history of vision can go back around 543 BC. From fossil studies, researchers found out that no. of animal species floats out and soon develops the first animal with eyes. As evolution in bigger species around fifty percent of neurons in our body are involved in the visual processing. Vision allows human to perceive, comprehend, and understand surrounding them while computer vision aims to duplicate the impact of human vision electronically in the form of images. The history of making artificial vision starts around 1545, Gemma Frisius, and Encylopedie in eighteenth century. The foundation period of the concept lies from 1943 to 1956, during this period research is centered towards understanding the concept of machine learning intelligence, artificial neural system, and automata theory computation. The first work in the field of Artificial intelligence in computer vision and image processing was done in MIT called “Summer Vision Project”. The rise of artificial intelligence starts and it is used widely in every field of technology. As the technology advances machine learning is used in almost every field. This book chapter provides you some details how the machine learning is used in image processing and to solve the current problem and challenge in computer vision like: Denoising, Image compression, Image color enhancement, segmentation, etc. Machine learning algorithm like Neural Network, Support Vector Machine, Genetic Algorithm, Convolutional Neural Network, etc. were discussed in the book chapter.

A. Sharma (✉)
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology JUIT, Solan, Himachal Pradesh 173215, India
e-mail: ajaysharmadharmani@gmail.com

A. Gupta
Interactive Technologies Institute (ITI/LARSyS), Madeira-ITI, University of Madeira, Caminho da Penteada, 9020-105 Funchal, Portugal
e-mail: gupta.ankit894@gmail.com

V. Jaiswal
National Centre for Disease Control NCDC, New Delhi, India
e-mail: computationalvarun@gmail.com

© Springer Nature Singapore Pte Ltd. 2021
P. Kumar et al. (eds.), Machine Learning for Intelligent Multimedia Analytics, Studies in Big Data 82, https://doi.org/10.1007/978-981-15-9492-2_11
Keywords Image processing · Machine learning · Neural network · Genetic algorithm · Support vector machine · CNN

1 Introduction of Machine Learning

Machine Learning is one of the powerful and influential technology in today’s world. In 1959 Arthur Samuel an American computer scientist coined this term with the statement that “It gives computer the ability to learn without being explicitly programmed”. Artificial Intelligence and Machine Learning researches the assessment, examination and advancement of calculations that can gain up from and make prediction on given data sets. Machine learning is an interdisciplinary field that utilizes algorithmic mathematical, procedures, and a use of artificial intelligence which enable computer systems and other electronic devices to learn and get it. Machine learning investigates the examination and development of calculations that can gain from and make prediction on given data sets. Among the diverse kinds of Machine Learning methods, an urgent refinement is drawn among is supervised and unsupervised learning [1]. In Supervised machine learning the computer program is “prepared” on a pre-characterized set of “preparing models”, which at that point encourage its capacity to achieve an exact end when given new information while in unsupervised machine learning the program is given a pack of information and must discover examples and connections in that. Machine learning is divided into different classes as follows.

Classification of machine learning as follows:
1. Supervised Learning
2. Unsupervised Learning
3. Semi-supervised Learning
4. Reinforcement Learning

1. Supervised Learning: It is also called as the inductive or predictive learning model. In the supervised learning, the prediction of the future outcome is based on the historical labeled data. The machine learning models are trained on the basis of labeled data for the outcome which consists of the input parameter and we know the expected outcome form the given data set. The model can then be used to predict the samples with unknown labels by calculating the features for that sample, similar used for training the model [2]. One of the issues with the usage of these models to the data is their selection. A solution to this problem is Meta-Learning which selects the attributes from the feature set and look for correlations between those attributes and performance of algorithms [3]. For example, Classification of persons whether the person is male of female based on the given data set, machine learning algorithm will learn based on the features and some patterns in the data and classify based on the training data. Supervised learning algorithm classified into two types of category, classification, and regression algorithm.
(a) **Classification Algorithms**: This type of algorithm classifies the data into some classes of some labels within the dataset. For the classification, we have algorithm like K-nearest Neighbor (KNN) [4, 5].

(b) **Regression Algorithm**: Algorithm like Linear regression and Logistic regression determines the mathematical relationship between the different data points within the dataset [6]. These algorithms are used in the predicting the output dependency and two or more variables. Regression algorithm can be classified into the linear regression and logistic regression. Logistic regression includes Support Vector Machine (SVM) [7], Random Forests, and decision tree algorithm [8].

2. **Unsupervised Algorithm**: Unsupervised algorithm is known as the descriptive models where the output is unknown and we don’t have any idea that what the algorithm will spit out from the input data. The input data for the training is unlabeled for unsupervised machine learning. The algorithms belonging to this category are used to cluster the data based on the similarity between the data points. However, these algorithms differ in their functionality for data clustering. For example, K-means clustering, the data is clustered into the different group based on some properties between the different input variables [9, 10].

(a) **K-means Clustering**: In the k-means clustering, the collected data is grouped together in a similar cluster based on their similar properties in the data points. It is a combination of two words k and means, K refers to no. of cluster of centroids’ in the group and means refers centroid centre point with respect to that centroid or cluster formed with in the given data set [11]. The commonly used data clustering method is k-means clustering where k is the number of clusters to which data is divided [12–14]. Deciding the number of clusters is tricky and data-dependent, and is often decided by elbow method, which works on heuristics.

3. **Reinforcement Learning**: It is a new type of learning technique where the algorithm learns and behaves based on the previous action within a problem. for every decision algorithm will work on the score whether it is positive of negative based on that machine will took the decision that what to do next. This type of learning uses the concept of agent-environment. The agent takes the input and current state of the environment for changing the state by an action. The-changed state is further shared using reinforced signal. The action selected by the agent is based on the score whose value should be as large as possible. Reinforcement model works on trial and error and hence follows heuristics For the algorithm to take decision is depend upon the process called as the Markov Decision Process [15]. In recent year’s reinforcement learning algorithm works in so many areas, e.g., in 2016 AlphaGo game use the reinforcement learning for and beat the world champion [16].

4. **Semi-supervised learning**: Is a technique that uses the small amount of labeled data for the training purpose within the large amount of unlabeled data sets. By using the labeled data will allow the algorithm to reduce the pattern and helps in
the identifying the relationships with in your target and the information which your dataset contains already has [17], e.g., if you have a dataset of records where the thousands of record are known and rest of the data is unknowns then the semi-supervised learning model will automatically classify that based on the behavior on the previous learn model and the algorithm will build the quick model on the basis of labeled data and applies it on the unlabeled data set in case of credit card fraud detection of bank loan approval task in the banking system, web page classification on the web, speech recognition and some in bioinformatics for genetic sequencing [18] (Fig. 1).

2 Image Processing Challenges

Image is a multidimensional signal, which requires certain operations for the extraction of useful information. Information extraction from the various image modalities are not straightforward and have certain issues or challenges associated with it. To list few, noise associated with the images, large amount of data for transferring images from one place to other digitally. Few image processing challenges in the computer vision and machine learning are as follows: Image compression, enhancement, Recognition, and visualization [19]. With the help of using machine learning lot of work is going to development of appropriate solutions using image processing algorithms. Here are some of the image processing challenges in which the machine learning work is going on are as follows:
2.1 Image Compression

Modern information and communication technologies include transferring images from one place to another e.g., television signal digitization. Image compression enables the sending and receiving of image frames using the available bandwidth [19]. Compression deals with reducing the size of images in order to make them transferrable at a faster rate, depending on the application. Compressing an image is tricky and requires good precision so that compressed and decompressed image should represent the same information. Designing a compression algorithm is a two-step process consisting of encoding the information while compressing and decoding while decompressing.

2.2 Image Enhancement

Image enhancement is a technique of enhancing the quality of image. Every image acquisition system has a noise content associated with it. The noise makes it hard to acquire a clean image. Image with higher noise content makes it difficult to extract plateful information. Hence, image denoising is one of the predominant challenges in image processing and requires high level of expertise to deal with it. Machine learning algorithm enhances the quality of images, below there is discussion of some of the algorithm used for the image enhancements.

2.3 Image Recognition

Image recognition deals with the identification and classification of objects in the image. This is a hard problem for the computer systems to process as these systems recognize images by just their pixel intensity values. A typical image recognition system tries to identify the pattern by looking into the texture color shapes of the objects in an image and classify the objects. Building a generic image recognition system is difficult to identify large number of objects. Moreover, it is easy to figure out patterns if the objects are similar but relatively harder for alike objects with different color, texture, or shapes. Artificial intelligence algorithm finds out the patterns within the image and classify accordingly.

2.4 Visualization

It is one of the subdomains of Computer Graphics and deals with generation of models and objects. The challenging part of the visualization is to model the dynamic objects.
Dynamic objects are the objects that changes as a function of time like moving hairs, curtains, trees, etc. Secondly, the generated models should be genuine which is computationally expensive and requires high computational configuration.

3 Image Processing Task, Noise Removing/Denoising of Images

Before understanding the technique of denoising of image, we have an idea about the noise and what are its cause and how it effect in the digital images. Let us understand the concept of noise. In the digital image processing noise is a topic, which makes lot of confusion in the students, research scholars or any other person who is hearing this term noise. Just in case of sound noise, refer to the audible distortion, in case of image it is visual distortion in the image. For example: if you took photograph with a closed cap on the lens, the resulting photo isn’t black totally it is approximate to black there must be some hinge, random brightness or some kind of distorted pixels in the image. In digital sensor noise cause by the sensors when attempting to capture tiny white light. In the real word scenario our camera is trying to focus on capturing the dark matter on the image as a result the electrical signal windup and as a fleck or noise [20]. Machine Learning algorithms are capable of removing the noise with in the image and helps in the denoising in the images. There are different type of noise found in the image while capturing the image and this can be removed with the help of using machine learning.

3.1 Types of Noise in Digital Image

In every image, there is always some kind of noise due to physical properties of noise, so technically we can say that every image contains some kind of noise.

- **Fixed Pattern Noise**: This sort of noise incorporates what is called as “hot pixels,” which are characterized as such when a pixel’s intensity far outperforms suppresses that of random noise arbitrary variances/fluctuations. Fixed pattern noise generally appears for the most part and shows up in long exposures and is exacerbated by higher temperatures. Fixed example clamor is special in that it will show nearly a similar circulation of hot pixels whenever taken under similar conditions (temperature, length of exposure, ISO speed). The fixed pattern noise is more offensive and is easier to remove from the captured image due to its repetitive nature. Camera internal electronic circuits has to just reveal the repetitive pattern of noise and just applying some algorithm subtract this noise reveal out the true color image [21].
- **Random Noise**: Random noise is also called shot noise/photon noise, which is categorized by intensity and color fluctuations above and below the actual
image intensity. The light composed of photons, which are discrete, random in nature. There will consistently be some random, irregular noise at any exposure introduction length and it is most impacted by ISO speed in camera. The random pattern noise commotion changes regardless of whether the exposure camera settings are similar or identical. The light emits and reflects everything you see the image. However, this thing is not fixed every time there is always graininess in the image. E.g. a very dim light bulb may emit on an average of 500 photons of light per second, but every time individual second will be a bit different—486 photons, 528 photons, 466 photons, 481 photons, 560 photons, and so on. On the off chance that, if you are taking a one-second long image of this light, you won’t get the very same outcome each time. This is what photographers call “shot noise/random noise/photon noise” in an image. Random noise is less pragmatic than the fixed pattern noise in the latest imaging devices, but a slight amount of noise could create more errors than the fixed pattern noise and difficult to remove [22].

• **Bending Noise:** This type of noise is also called as the digital noise and it is depended on the digital image device, internal electronics of cameras in which image is being captured. This noise is also visible in the high ISO speed and shadows in an image, when an image is extremely brightened or white balance on an image depending upon the camera model [23].

**Different type of noise reduction model in image processing are**

Unwanted of unnecessary information within an image like as blurred objects, corners, unseen lines, edges, artifacts, disturbed background sense contribute the noise. This noise is produced form the CCD (Charged Coupled Device), CMOS (Complementary Metal Oxide Semi-Conductor) sensors with in the digital image devices. To reduce this redundant noise, prior dectection and learning of model needs to be develop. In some cases mathematical function like Modulation transfer function (MTF), point spreading function (PSF) are used for quantitative and qualitative analysis of mode of noise model [23].

**4 Color Enhancement**

For visual perception of human being color image provide the more and accurate information as compare to the grayscale images. Color enhancement is the key issue in the high-quality pictures and some other HDTV cameras. The quality of images is easily affected by the light and weather condition which leads to suffer from the loss of information. The are some color enhancement technique such as gamma correction technique, histogram equalization, contrast stretching, Contrast-limited adaptive histogram equalization which are older one and give the mean performance in the Root Mean Square deviation (RMSE), Peak signal to noise ratio (PSNR) and mean absolute error. The newly developed algorithms are used with the machine
learning model give us the better accuracy for the color enhancement. These are the algorithm such as Retinex, Homomorphic and Wavelet Multiscale technique are popular for the image enhancement [24].

With the help of machine learning color enhancement is done in various domains like underwater animal detection and classification [25], Iris Recognition [26], Smart laser surgery [27], Brain image enhancement [28, 29], Ultrasound image enhancement [30] etc.

Colored image provides and hold a vast amount of information regarding object in image. It is difficult for the human eye to recognize and analyze each and every little potion of image such as color intensity and texture with in an image [31]. In colored images we need an algorithm which segments out all the parts with in an image and gives us the better results especially in case of the satellite image because they contain lot of noise with it and it is a trivial task of computer. Color enhancement is more difficult in case of color image rather than the gray scale image due to the following reasons.

1. In case of color we need to consider vector instead of scalar image due to presence of color.
2. Complexity of image is also a problem in case of color image.

Color or contrast enhancement is another technique in which the intensity of pixel is changes to maximize or enhance the density of pixel with in image. The term “contrast” used for the separation of the dark and bright areas within an image. Due to contrast enhancement we are able to remove the dark areas from the image and we are able understand the content within an image [32]:

\[ s = (r - c) \left( \frac{b - a}{d - c} \right) + a \]

Mathematical equation: \(a\) and \(b\) are the upper limits for the 8-bit gray scale picture, \(c\) and \(d\) are lower and upper value of histogram.

Image Enhancement Techniques: This provide the information for human to understand the image by interpretation and perception of viewer against image via providing better input for automated image processing techniques. Image enhancement also classified into the two type of category:

- **Spatial and Frequency Domain:** In the special domain we directly deal with the image pixels. People do the manipulation with the pixels with for the enhancement in the image. And in case of the frequency domain we change the frequency of pixels and change the image into the frequency. So, this states that we have need to apply the Fourier transform within the image computed first and after that we have to apply the inverse Fourier transform in order the get the resultant image [33].

- **Histogram Equalization:** is a technique which normalize the contrast with in the image by using the histogram equalization. It is based on the idea of near-uniform
probability distribution function. It redistributes the probability distribution function within an image. For example, if in an image there are peaks and valley, after applying the histogram it will normalize the peaks and valleys will be shifted to normal curve. This results in improving the image contrast. There are types of the Histogram equalization namely Global Histogram Equalization (GHE), Adaptive Histogram Equalization (AHE), Block based Histogram Equalization (BHE) [34, 35]. In GHE each pixel with in an image assigned a new value based on the cumulative distribution function. To perform Global Histogram Equalization (GHE), the first original histogram of the grayscale picture should be adjusted or equalized. The aggregate/cumulative histogram from the input picture should be evened or equalized out to 255 by making a new force an incentive by applying [34].

\[ s = T(r) = \int p_r(w)dw \]

\( Pr(r) \) is probability distribution function, \( r \) = gray level of input image

\[ p_s(s) = \begin{cases} 
1 & \text{for } 0 \leq s \leq 1 \\
0 & \text{otherwise} 
\end{cases} \]

\[ S_k = T(r_k) = \sum_{j=0}^{K} \frac{n_j}{N^2} = \sum_{j=0}^{k} P_s(r_j), \quad 0 \leq r_k \leq 1, k = 0, 1, \ldots, L - 1 \]

\( n_j \) = count of jth pixel, \( N \) is the no of pixels in horizontal and vertical direction in an image.

- **Homomorphic Filter**: Homomorphic filter work in the principle of normalization of brightness and increase the contrast across the image. This will help the user to see the dark area within the image. For the removal of multiplicative noise, Illumination and reflectance we use the homomorphic filter. Since illumination and reflectance are combined multiplicatively, the components that are made added substance by taking the logarithm of the picture intensity, so that therein frequency domain linearity is formed. Illumination or brightening varieties can be thought of as a multiplicative noise and can be decreased by separating in the log domain area.

- **Retinex**: In 1964 this theory was first proposed by Edwin Land. Multi-Scale Retinex (MSR) and Multi-Scale Retinex with modified color restoration are the other technique in the Retinex [36].

\[ L = R/E \]

\( L \) is the incident of light, \( R \) is the object reflection, \( E \) is the reflected light.
There is another technique called the single-scale Retinex method proposed by the Jobson, in which estimate the illumination by using the low pass filter for an input color image.

**Wavelet Multi-scale Transform:** In recent year, wavelet transform becomes a famous tool for the image processing. Wavelet transform (WT) is a powerful mathematical tool for the processing of images of multiple resolutions. It provides a septal and frequency characteristic with in an image. WT transform for an image the high-frequency part stores the original image details and the low frequency or imaginary part is stored with in the low-frequency part in an image [37]. The imaginary and low frequency part determine the dynamic image with in the image, but in case of the low frequency of the annotation there is the loss of information, so that why some of the information is stored in the high-frequency domain that is why the image reconstructed by the Inverse Wavelet Transform has the more details and widely used for the image processing [37, 38].

\[
\psi(t) = 2 \sin c(2t) - \sin c(t) = \frac{\sin(2\pi t) - \sin(\pi t)}{\pi t}
\]

Equation for continuous wavelet transform

\[
\psi_{m,n}(t) = \frac{1}{\sqrt{a^m}} \psi\left(\frac{t - nb}{a^m}\right)
\]

Discrete wavelet Transform equation

5 Sharpening of Images in Machine Learning

Image sharpening belongs to the category of spatial filtering in the image processing deals with the enhancement of detailed information within the images [39]. Sharpening of the images increases the brightness within the image. This could be done via using the high pass filter within the image. On the spatial regions there is small image features and image features correspond to the detailed information about high-frequency components of an image. Image sharpening is dependent on the two factors which are resolution and acutance. The resolution of the image deals with the no of pixel within the image more the pixel in the image more will celerity or sharpness within the image so this doesn’t concern with the subjective nature of image. Acutance is little bit more complicated to understand and is linked with the subjective measure of image dealing with the edge with in the image. Edge on the image has more contrast appear more accurately in the image and human visual system. So, to increase the acutance within the image there is only one way to increase the acutance within the image at the edge contrast [40, 41].
Technique developed by Google like RAISER: Rapid and Accurate Image Super-Resolution “use machine learning to produce high quality of images from low resolution of images from 10 to 100 times faster than the currently available methods”. SISR (Single Image Super Resolution) is the technique which converts the low-resolution images into the high resolution [42]. This method is based in the upscaling the image in the linear interpolator including the nearest neighbor in the image [43, 44] (Figs. 2, 3 and 4).

**Binary reconstruction of gray scale and nonlinear processing images:**
Grayscale images contain only black and with color. It permits the removal of small regions that are disjoint forms the larger objects without distorting the small features within the large objects. Neural networks have been proposed and a variety of works in binary reconstruction of grayscale images [45, 46]. This technique is mainly used in X-ray images and MRI images and helps in the detection of cancer. The interpretation of these gray scale images is difficult task and depends on the expertise and experienced radiologist. Thus, extracted features were classified with the help of Artificial Neural Network. In medical imaging diagnostic the detection of presence

![Diagram](a)

![Diagram](b)

**Fig. 2** Basic learning with global filters, low resolution to high resolution. a learning stage, b upscaling stage

\[
\begin{pmatrix}
-1/9 & -1/9 & -1/9 \\
-1/9 & 1 & -1/9 \\
-1/9 & -1/9 & -1/9
\end{pmatrix}
\]

**Fig. 3** Array of kernel in high pass filter for sharpening of image
of metal within the body which makes it important to diagnose of detect the metal object with in the body with the help of x-ray image or MRI CT images [47]. The CT images are classified into two categories one is high density metal and low-density tissues with in the human body. Within the sight of a metal object, linear attenuation coefficient at the limit of the metal changes out of suddenly, and the identification of the metal boundary is identical to transfer a signal with sharp edge. Not quite the same as regular CT recreation, where the center focuses is to infer the density appropriation inside the field of view and is frequently tested by the missing of solid projection signals behind the metal objects in body, here we estimate a paired binary picture. Mathematically, recreation of such a picture is considerably more manageable on the grounds that there is no missing information issue here. That is, the projection information close by is adequate for us to characterize the metal boundary within the patient [48, 49] (Figs. 5 and 6).

**Fig. 4** Sharpening of MRI (Magnetic resonance image) images. a MRI image of knee, b knee MRI with sharpened filter sharpened knee MRI [44]

**Fig. 5** Binary image reconstruction
Advantages

- Easy to acquire: straightforward computerized digital cameras can be utilized along with basic frame stores, or minimal low-cost scanners or thresholding might be applied to gray level pictures or images.
- Low storage: close to 1 bit/pixel, frequently this can be diminished as such pictures are entirely amiable to compression in the image (for example run-length coding).
- Simple handling and processing: the calculations by the algorithm are by and large a lot more straightforward than those applied to gray level pictures/images obtained through cameras.

Disadvantages

- Limited application: as the representation is just an outline silhouette, application is restricted to undertakings where inside detail isn’t required as a distinctive characteristic.
- Does not stretch out to 3D: The 3D idea of items can once in a while be represented to by outlines silhouettes (What could be compared to binary processing preparing utilizes voxels, spatial inhabitance occupancy of little small cubes in 3D space).
- Specialized lighting is required for outlines: it is hard to get solid, reliable binary image without limiting, restricting the environment. The simplest example among is model is an overhead projector or light box.
6 Segmentation and Edge Detection Using Neural Network

The development of technology in our society we can expect that the image sensing device detects the things around you. Image segmentation is an important step for any image processing and computer vision algorithm. Analyzing and extracting and object form and image is an important for building intelligent machines for our surroundings. Segmentation will help in the retrieval of images form a large-scale database for content-based image retrieval system software. In a region, a group of connected pixels are there with similar properties. Edge detection is segmentation by finding similar pixels on their boundary. This can be termed as the binary image classification at pixel level. To solve this problem use of fixed and adaptive and feature selection in concurrence with the use of Support vector machine [49]. In this case the first work was done by David Marr in 1970 and then first successfull work done in the David Lowe in 1987 [49]. There was improvement in the Algorithm by Shi and Malik, 1997.

The image segmentation is the process of partitioning the image into different segments. This is done through the different patterns based on the pixel’s similarity within the image. Segmentation is the way toward recognizing, distinguishing objects in the dataset from their environmental surroundings factors in order to encourage the creation of geometric models. For instance, in clinical biomedical imaging it is regularly important to measure the shape, surface area, or volume of tissues in the body. Once the dataset is segmented, those quantities are easily measured. The image segmentation is needed to identify the content within the image. Image segmentation is used to classify the data within the image. Image segmentation is used in almost every field ranging from the automated driving, medical cancer diagnosis, Handwritten digit recognition, underwater object classification/detection, food industries, etc. It has a vital role and more important among other problems in image processing. Due to introduction of machine learning algorithm computer researchers developed new algorithm to automated segmentation and edge detection within the image. Algorithm like neural network and other clustering algorithm will classify the segmented part within the image. Hence there lot of work is needed to be done for image segmentation and edge detection.

The objective of the edge detection is to find the pixels in the picture that relate to the edges of the object found in the picture captured through the cameras. This is normally finished with a first and second derivative estimation following by a test which denotes the pixel as either having a belong to the place with which an edge is there or not. The outcome is a binary image which contains just the recognized or detected object edge pixels. Edge detection is the fundamental tool for the image segmentation. There are several tools and software that are used for the image segmentation and for the edge detection also. For the image segmentation, every method is not appropriate so there are several techniques which are used for the image segmentation are as follows [50].

- Region-based Segmentation
- Edge-based Segmentation
• Threshold
• Feature-based Clustering

• **Region-Based Segmentation:** It is based on the process that the group of the pixels are based on the similar properties. Partition of the image takes place based on similar region. Region in an image may be consist of the object or some entity within the image. This type of classification is suitable for the noisy image where the border with the object are difficult to detect or border are blur. Fair level of accuracy is achieved in the care of region-based segmentation [51].

• **Edge-Based Segmentation:** Segmentation within the image based on the discontinuous similar intensity values in an image. Edge detection can be found where there are abrupt changes within the brightness and the intensity in the pixel and a problem in the primary value in the image analysis. Hence by detection the edge with in the image we do the image segmentation. Edge-based Segmentation is further classified into sub-category: Filtering and Enhancement, Detection of edge points, and edge localization [51].

• **Filtering and Enhancement:** So as to encourage the identification of edges, it is fundamental to repress as much noise as could reasonably possible be expected and decide changes in power in the area of a point, without destroying the true edges.

• **Detection of Edge Points:** Figure out which edge pixels ought to be disposed as of noise and which ought to be retained (as a rule, thresholding gives the criterion used for the identification or detection).

• **Edge Localization:** Not all of the points in an image are edges for a particular application. Edge localization determines the exact location of an edge. Edge thinning and linking are usually required in this step.

• **Threshold:** Thresholding is simple and powerful tool for image segmentation. This is beneficial where the light object in the shady background. It will convert the multi-level image into the grayscale image. The separation of the object forms the background is done by selecting the values threshold value of $T$. Local thresholding and Global thresholding. When the $T$ is not constant it is called as the local thresholding and when $T$ is constant global thresholding.

• **Feature-based Clustering:** Clustering the process of grouping the individual together who have similar properties and dissimilar properties into another cluster [51]. Good quality clustering produces high intracluster and low inter clustering properties. Issues arise in the feature-based clustering while dealing with the image are as follows:
  
  − How to represent the image.
  − How to organize the data from the image.
  − How to classify the image into a certain cluster (Table 1)

Clustering is further classified into the Fuzzy C-means clustering and K-means clustering. K-means is the fast, robust, and simple unsupervised learning method for the clustering. Fuzzy clustering allows the object which form different cluster with membership. It is an effective method for the pattern recognition.
### Table 1  Comparison between the different algorithms used for the segmenting and edge detection

| Parameter                          | Threshold-based method | Region-based method | Edge-based segmentation | Cluster-based method | Fuzzy C-means method |
|------------------------------------|------------------------|---------------------|-------------------------|----------------------|----------------------|
| Spatial information                | Ignored                | Considered          | Ignored                 | Considered           | Considered           |
| Region continuity                  | Fast                   | Good                | Reasonable              | Reasonable           | Good                 |
| Speed                              | Reasonable             | Slow                | Moderate                | Fast                 | Moderate             |
| Computation complexity             | Less                   | Rapid               | Moderate                | Rapid                | Moderate             |
| Automaticity                       | Semiauto               | Semiauto            | Inactive                | Automatic            | Automatic            |
| Noise resistance                   | Less                   | Less                | Less                    | Moderate             | Moderate             |
| Multiple object detection          | Poor                   | Fair                | Poor                    | Fast                 | Fair                 |
| Accuracy                           | Moderate               | Fine                | Moderate                | Moderate             | Moderate             |

**Automatic Edge Detection Using Deep Learning:**

Edge detection, identification is an image processing method used to find and distinguish the sharp discontinuities within a picture captured from the camera. Edge detection is as often as possible the initial phase in recouping data of information from pictures. One of the oldest classical examining studies in computer vision is image processing in which further most of research work is done in edge detection or identification. There are many edge detection operations available, which recognize vertical, level, horizontal, corner, and step edges in an image. The quality, nature of edges identified by these operators is profoundly subject to noise, clamor, lighting conditions, objects of similar intensities and the density and the thickness of edges in the scene [52]. There are two alternatives which are used for the detection of edges in the image as follows.

1. Classification of all the pixels that satisfy the criteria of homogeneous.
2. Detection of all the pixels that are on the border in different homogeneous areas (Fig. 7).

Instead of using the classical edge detection method neural network or deep convolutional neural network are used more widely for the classification of this purpose (Fig. 8).

\[
L = \frac{1}{N_p} \sum_{i=1}^{N_p} - \log \left( \frac{e^{f_{i,j}}}{\sum_j e^{f_{i,j}}} \right) + \lambda \sum_{j,k} W_{jk}^2
\]

Mathematical equation for the Edge Detection
Cardiovascular edge detection using the neural networks to perform the flow of blood flow simulation. In this method is used to identify the blood vessels in the entire region of image for image segmentation. This method used a fully connected convolutional neural network for classification of image segmentation of cardiovascular disease within the lung nodules [53, 54] (Fig. 9).

$$L = \frac{1}{N_p} \sum_{i=1}^{N_p} - \log \left( \frac{e^{f_{i,j}}}{\sum_j e^{f_{i,j}}} \right) + \lambda \sum_{j,k} W_{jk}^2$$

Mathematical equation for the loss functions in the deep neural network.
Brain tumor classification with the help of deep learning is possible. Convolutional neural network is used for the training the model on the brain tumor data set, testing is performed on the unlabeled data set and the accuracy is measured. Automated segmentation and detection is also possible with the help of these Deep learning based algorithms [56].

7 Image Compression

Image compression is the method of minimizing the size of the image by minimizing the sizes in bytes of graphics file with reducing the quality of the image. This will create more space and allow storing more images in an available memory size. Today on the internet ninety percent of the data is in the form of image and videos. Thus, most of the companies will use the data compression technique in order to run out of space. Machine Learning algorithms like support vector machine combines with the DCT discrete cosine transform achieve the level as compared to the RBF and multilayer preceptor [54, 57]. DCT is based on the JPEG compression algorithm one of the most widely used image compression algorithms. In the resulting value of the matrix it divides the whole image into pixel of squares and apply a discrete cosine function [57].

\[
F(u, v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\Lambda(i) \Lambda(j) \cos \left[ \frac{\pi u}{2N} (2i + 1) \right] \cos \left[ \frac{\pi v}{2M} (2j + 1) \right])
\]

Mathematical representation of Fourier transforms for discrete solution.
Image compression with the help of machine learning:

1. **Neural Network**
2. **Genetic Algorithm**

- **Neural Network**: NN generates the rules to adjust and compute the complexity of image compression. They are generally composed of an input layer, hidden neurons, and an output layer. There may be one or more layer of hidden neurons. The hidden layer networks are logic functions that, based on input from one or more neurons in preceding layers will output information to one or more neurons in subsequent layers [58, 59]. Each of the connections between the neurons is designed to weight the input into the next neuron, hence if, for example, neuron A on input layer, outputs the value 2.46 to neuron B on hidden layer, there may be a weight, or bias, between both that doubles up the value, and consequently neuron B receives 4.92 as an input. In this case, the weight for the connection (synapses) was 2, doubling up the value, but this weight can be modified by giving a network training data. If a Neural Network is provided with a sufficiently large training data set, eventually the value of the weights in the synapses will be such that the results obtained from a neural network computation will be close to the expected results. While applying neural network method in image we can reconstruct and compress the image from the following provided image data set on which machine learning model is trained [60]. Fast data transmission in industrial scale on IoT devices is possible due to machine learning algorithm for the image compression [61]. In medical science for the detection of diabetic retinopathy is possible because of the image compression and machine learning image processing plays a vital role in the development of better treatment for diabetic patients [62].

- **Genetic Algorithm**: Genetic algorithms (GA’s) are based on the processes of evolution and natural selection. They are a subset of evolutionary algorithms. Genetic algorithms take an iterative approach to problem-solving, generally following these steps [63, 64].

  - **Initialization**: A population of algorithms is generated (randomly or semirandomly) to solve a particular problem.
    
    E.g. \( x = a + b, x = 2b - 3c, \) and \( x = a + b - c \) for deriving the quadratic equation.

  - **Selection**: The algorithms to be combined are selected, based on the proceeding of fitness if it is not the first iteration.

  - **Crossover**: Parent algorithms are combined to produce new children algorithms, the operation used to combine them is a factor to be determined by the supervisor, and may have a significant effect on the computational efficiency of the process.
    
    E.g. \( x = 2b - 3c, \) and \( x = a + b - c \) are added to produce \( 2x = a + 3b - 4c \)

  - **Mutation**: A random number or factor is added to the algorithm, mimicking the process of mutation in nature. For our example, this could be \( 2x = a + 3b - 4c + 8 \). This will generate new, potentially unique, algorithms for solving the particular problem.
– **Proceeding of Fitness:** The generated algorithms are tested against a training set, determining the fitness of the solutions. The solutions that are inaccurate above a certain threshold are dropped, and the rest are iterated through the process again, going to step 2: selection. This process is repeated until a maximum number of iterations are taken or the solution is close enough to the training set.

**Combining the machine learning algorithm and compression algorithm:**
- GA and DCT
- Neural Network and DCT

**Genetic Algorithm and Discrete Cosine Transform**

Genetic Algorithm is one of the most common used optimization algorithms; so in image compression, this could be used for the optimizing the fractional compression and DCT (Discrete Cosine Transform). In DCT for optimal compression genetic algorithm could be a better solution for the optimization to improve the quantization with in the image. GA also applied in the fractional image compression. In fractal compression are searching the domain blocks and transforming the domain block to match the range blocks, it is these processes that are optimized by the genetic algorithm from Mtira and Morthy (1998) [64].

**Neural Network and DCT**

Neural networks have been used to improve image compression extensively. Both as standalone methods as well as combined with other techniques, traditional or more novel approaches.

**Setiono and Lu** used a standalone neural network in order to compress images, using the image as input data for the network, and the reconstructed (compressed) image as the output. The images used where black and white, and the compression ratio achieved was about 10, producing good quality, albeit noticeably lossy, images. Figure shows Lena, uncompressed on the left, and compressed with the neural network on the right. The loss of information is particularly noticeable around the hair and the band on the hat, due to this being high-frequency areas [64].

**Parodi and Passaggio** approached the compression problem by combining DCT and Neural Networks. The approach they take to first divide the image into low and high activity area blocks, which will be compressed at different ratios and by different neural networks. The high and low activity areas are classified as such by using DCT’s; high activity areas are those with higher coefficients within the DCT transform, indicating that they contain a wider range of frequencies. Each of this image is posteriorly provided to the corresponding neural network, either high or low capacity, with a number of input neurons proportional to the size of the block to be processed. The images are then compressed by the neural network, which uses the original block as a target for the learning process. The results obtained indicate a significant improvement in image quality over standard neural network compression.
of images, without compromising compression ratio, they achieve higher PSNR and lower MSE.

8 Facial Expression Analysis

In recent years there is growing interest in emotional intelligence in the human-computer interaction. For the facial expressional analysis there is need for the computer to interact naturally with the user via speech, body gestures to display emotions. One of the important way humans display emotions is through facial expressions via vocal, visual and other physiological means like Neutral, anger, disgust, fear, joy, sadness, surprise. Support vector machine and neural network can classify the facial expression with a speed and accuracy. Data set Cohn and Kanade’s DFAT-504 have images of 100 of university students age ranging from 18 to 30 years [65] (Fig. 10).

- **Face Finding**—FaceReader finds a precise position of the face via utilizing the well-known Viola-Jones calculation.
- **Modeling**—A precise 3D modeling demonstrating of the face is made, which depicts more than 500 key points with in the face
- **Classification**—A neural network system utilizes more than 10,000 pictures to arrange the basic emotional expressions, passionate articulations, and various properties of facial expression
- **Deep face classification**—This procedure permits FaceReader software to straightforwardly order the face from picture pixels utilizing an artificial neural network system to perceive patterns with in the image. This permits the product to break down the face regardless of whether a part of it is covered up with (Fig. 11).

Fig. 10 The pi chart indicates the emotions of human facial expression analysis software works by following these consecutive steps
Deep learning algorithm for the face expression analysis:

- Multiclass Support Vector Machines (SVM)
- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Convolutional Long Short-Term Memory (ConvLSTM)

- **Multiclass Support Vector Machines (SVM):** Are supervised machine learning algorithms that investigate, analyze, group information, and classify the data, and they perform well while predicting, classifying human facial expressions. In any case, they possibly do so when the pictures are made in a controlled lab setting with steady head postures, brightening and a constant illumination [66].

- **Support Vector Machine:** SVMs perform less well when characterizing pictures are captured “in the wild,” or in unconstrained, unspontaneous, uncontrolled settings. Along these lines, the most recent training machine leering or deep learning architecture being explored are altogether profound as deep neural network systems which perform better under those other available architectures. Convolutional Neural Networks (CNN) are right now considered the go-to neural systems for picture characterization, since they get on designs small patterns of a picture, for example: curve of an eyebrow or lips style, nose style etc.

\[
f(x_i, W, b) = Wx_i + b
\]

Mathematical equation for the SVM classification

where \(x_i\) = image pixel data, \(W = C \times K\) weight matrix, \(b = \) biased term.

- **Convolutional Neural Network:** CNNs apply kernels, which are matrices smaller than the image, to chunks of the input image [67]. By applying kernels to inputs, new initiation lattices, at times referred to as feature maps, are created and gone as a contribution, pass the input to the next layer of the network. Along these
lines, CNNs process increasingly granular elements inside a picture and make a better via classification between the two similar at recognizing two comparable emotion classification. VGG16 is the state-of-the-art architecture for the CNN network with 16 fully connected layers, with 3*3 convolutional and 2*2 pooling layers [68, 69] (Fig. 12).

- **ResNet50**: It is one the most popular CNN architecture for the face recognition. It has the additional mapping capacity as compare to the VGG16 networks [71] (Fig. 13).

- **Recurrent Neural Networks (RNN)**: Utilize dynamic transient behavior while the classification of an image. This implies when an RNN forms an input information model, it doesn’t simply take a data at the information from that example model. It additionally takes a data from past sources or layer with in the training

Fig. 12  CNN architecture for facial expression analysis [70]

Fig. 13  ResNet block diagram with additional mapping identity
model from the images, which are utilized to give further output. In FER, the setting could be past picture frame of a video clip [72].

• The thought of this methodology approach is to catch and capture the transition between facial patterns example after some time, permitting these progressions to turn into extra additional information data points which focuses and support the characterization. For instance, it is conceiving possible to catch the changes in the edges of the lips as an articulation expression goes from natural to happy, cheerful by smiling, rather than the simply the edges of a smile from an individual image frame (Fig. 14; Table 2).

9 Different Machine Learning Algorithm for Image Processing

• Edge detection algorithm
• Pattern detection algorithm
• Expression analysis algorithm
• **Edge Detection Algorithm**

The point at which the image brightness changes sharply and organized into the lines or curves are termed as the edges. Edge detection is the fundamental tool in the image processing especially in the computer vision and particularly playing the role in the feature extraction. Feature detector such as Scale Invariant Feature Transform SIFT (DoG), Harris, and SUSAN are one of the good methods which yield high quality of features but due to too much computationally intensive nature for the use in real-time application of any complexity [74]. Canny edge detector is an edge detection technique which uses a multi-stage algorithm which is developed by the John F. Canny in 1986 to detect a wide range of edges in an image. Machine learning can be used to derive a feature detector which can process the PAL video less than seven percent of processing time. ML can compare with Harris detector (120%) and SIFT (300%) faster at full frame rate [75, 76] (Fig. 15).

**Working of SIFT algorithm**: SIFT algorithm is widely used and can be divided into the following parts as follows [77]:

1. Constructing a scale space This is the underlying initial arrangement. You can create internal representation of the original image to make sure that there is scale invariance. This is finished by creating a “scale-space”.
2. LoG Approximation: The Laplacian of Gaussian is incredible for finding intriguing interesting points (or key points) in a picture image. Be that as it may, it’s computationally expensive. In this way, we cheat and surmised approximate it using the representation created earlier.
3. Finding key points with the super-fast estimate approximation, we currently attempt to discover out the key points. These are maxima and minima in the Difference of Gaussian picture we have to figure out.
4. Get rid of bad key points at Edges, light illumination and contrast and low complexity areas are bad key point. Eliminating these points make an algorithm efficient, productive, and strong and robust toward the feature learning. A method like the Harris Corner Detector is used here for the edge and curve detection in objects.

5. Assigning a direction to the key focuses a direction is determined for each key point. Any further computations are done comparative with this direction. This adequately cancel out the impact of direction, making it turn invariant.

6. Generate SIFT includes Finally, with scale and rotation invariance set up, one more representation is produced. These sides particularly distinguish features of the objects in an image. The algorithm will be implemented in the Open Cv.

- **Pattern Detection Algorithm**

  Repetitive sequence within the image called pattern. Pattern in the images can be either seen physically or can be found with the help of mathematics by applying some technique. In case or colored image, there should be more than one pattern in the image. Pattern detection is the process of classification of image data based on the supervised learning or statically information extracts out form the unla- beled image data. Pattern reorganization has done some great work in the speech recognition, Speaker identification, multimedia document recognition (MDR), and medical image analysis and could be used in the biometric identification and Fraud detection also. It could be done via supervised learning (Classification, Regression), Clustering, ANN, and reinforcement learning. In the literature, so many algorithms are available for the classification like support vector machine for multi-class pattern recognition, ANN, CNN, etc (Fig. 16).

![Fig. 16 Pattern detection in medical image analysis](image-url)
• **Expression Analysis Algorithm**

Face expression technology could be used in the variety of applications such as access control and surveillance and identity authentication. Human being uses the facial expression to interact with the socially. Since the rise of artificial intelligence proceeded with research interests for empowering computer system frameworks to perceive the expressions and to utilize the emotive data inserted in them in human–machine interfaces. Research activity in face recognition have been increased in the past couple of years. SVM is recently proposed and widely used algorithm for the facial expression analysis. The Support Vector Machine (SVM) classification trains, identifies the different displacement patterns features form unknown sample images and based on the learning experiences it classifies the different facial expressions like anger, fear, joy, sorrow provided during the training period (Fig. 17; Table 3).

**RAISAR Sharp Image with Machine Learning:**

Technique developed by the Google team that incorporates the machine learning in order to produce the high-resolution images from the low scale image. Its 10–100 times better than the currently available algorithm and can be run on the mobile devices [78] (Fig. 18).

![Facial expression analysis pipeline using support vector machine SVM](image)

**Table 3** Showing the classification accuracy of different emotions and total accuracy of 87.9%

| S. No. | Emotions | Anger (%) | Disgust (%) | Fear (%) | Joy (%) | Sorrow (%) | Surprise (%) | Overall (%) |
|--------|----------|-----------|-------------|----------|---------|------------|--------------|-------------|
| 1      | Anger    | –         | 72.5        | 61.8     | 97.1    | 93.8       | 97.1         | 84.1        |
| 2      | Disgust  | 72.5      | –           | 63.2     | 94.9    | 88.9       | 100.0        | 83.9        |
| 3      | Fear     | 61.8      | 63.2        | –        | 90.9    | 66.7       | 100.0        | 76.9        |
| 4      | Joy      | 97.1      | 94.9        | 90.9     | –       | 96.8       | 97.1         | 95.3        |
| 5      | Sorrow   | 93.8      | 88.9        | 66.7     | 96.8    | –          | 100.0        | 89.4        |
| 6      | Surprise | 97.1      | 100.0       | 100.0    | 100.0   | –          | –            | 98.8        |
Fig. 18  a Image showing pre-learned filters from LR to HR. b Figure showing LR to HR with weighted average

Upsampling:

Is the process of producing an image of larger size with significantly more pixel and higher image quality from low quality of images. The term up sampling is associated with the process of resampling in Digital image signal processing (Fig. 19).

Fig. 19  Upsampling and interpolation of image signal in image
1. Algorithms: Hashing Based Learning and Upscaling

(i) Algorithm: RAISER Hashing Based Learning

**INPUTS:**
1: \( s \) – Upscaling factor
2: \( \{X_i\}_{i=1}^L \) - The ground truth HR images (Optionally sharpened/enhanced).
3: \( \{Z_i\}_{i=1}^L \) - The LR (optionally compressed) versions of \( \{X_i\}_{i=1}^L \).

**OUTPUT**
1: \( \mathcal{H}_t : (j) \mapsto h_{j,t} \) – Hash table that maps each key.

\( j = (\theta, \lambda_1, \mu) \) to the corresponding \( d^2 \) -dimensional filter \( h_{j,t} \), where \( 1 \leq t \leq s^2 \) is the pixel/patch type.

Process

\( Q_{j,t} \leftarrow 0 \), \( V_{j,t} \leftarrow 0 \) for all the hash-table keys denoted by \( j \) and pixel-type \( t \)

for \( i = 1 \) to \( L \)

compute \( y_i \), an initial interpolated version of \( z_i \)

Initialize \( A_{j,t} \) and \( b_{j,t} \) to be empty matrices for all the possible hash-tables

Denoted by \( j \) and pixel type \( t \)

foreachpixel \( k \) in \( y_i \)

Denote by \( j = (\theta_k, \lambda_1^k, \mu_k) \) the hash-table key of the pixel \( k \)

Denoted by \( p_k \in \mathbb{R}^{d^2} \) the patch extracted from \( y_i \) centred at \( k \)

Denoted by \( x_i(k) \) ground truth of the pixel of \( x_i \) located at \( k \)

Denoted by \( t \) the type of the pixel \( k \)

\( A_{j,t} \leftarrow [A_{j,t} : p_k^T] \), i.e., concatenate

\( p_k^T \) to the end of \( A_{j,t} \)

\( b_{j,t} \leftarrow [b_{j,t} : x_i(k)] \), i.e., concatenate

\( x_i(k) \) to the end of \( b_{j,t} \)

end

for eachkey \( j \), and pixel \( t \) do

\( Q_{j,t} \leftarrow Q_{j,t} + A_{j,t}^T t A_{j,t} \)

\( V_{j,t} \leftarrow V_{j,t} + A_{j,t}^T t b_{j,t} \)

end

end
(ii) Algorithm: RAISER Hashing Based Upscaling

**INPUTS**
1. $s$ – Upscaling factor.
2. $z$ – LRimage
3. $H_{t}: (j) \mapsto h_{j,t}$ – Hash-table that maps each key $j = (\theta, \lambda, \mu)$ to the corresponding $d^2$-dimensional filter $h_{j,t}$ where $1 \leq t \leq s^2$ is the pixel/patch type.

**OUTPUT**
1. $\hat{X}$ – HRestimate of $z$

**PROCESS**
Compute $y$, an initial interpolated version of $z$

```plaintext
foreach pixel $k$ in $y$
    Denote by $j = (\theta_k, \lambda_k^t, \mu_k)$ the hash-table key of the pixel $k$
    Denote by $p_k \in \mathbb{R}^{d^2}$ the patch extracted from $y$ centred at $k$
    Denote by $t$ the type of the pixel $k$
    $h_{j,t} \leftarrow H_{t}(j)$
    $\hat{X}(k) \leftarrow p_k^t h_t$
End
$\hat{X} \leftarrow $ CT-Blending $\hat{X}, y$
```

2. Algorithm: SUR-Net: Predicting the Satisfied User Ratio Curve for Image Compression with Deep Learning [79]

Let $I_1[0], I_2[0], ..., I_K[0]$ be a large training set of $K$ pristine reference images. For each pristine image $I_k[0], k \in \{1, ..., K\}$, we associate the $N$ distorted images $I_k[n], n = 1, ..., N$ corresponding to the distortion levels $n = 1, ..., N$.

Let $SUR_k$ denote the SUR function of image $I_k[0]$.

**Regression model**: $f \theta$ parameterized by $\theta$ such that: $f \theta(I_k[0], I_k[n]) \approx SUR_k(n), k = 1, ..., K, n = 1, ..., N$.

Then we calculate $\Delta f_{\text{min}}, \Delta f_{\text{max}},$ and $\Delta f_{\text{avg}}$, between the distorted images $I_k[n]$ and the pristine image $I_k[0]$, e.g., $\Delta f_{\text{min}} = f_{\text{min}}(I_k[n]) - f_{\text{min}}(I_k[0])$.

**Given the training data**: $\{(I_k[0], I_k[n], SUR_k(n))| k = 1, ..., K, n = 1, ..., N\}$,

**Mean Absolute Error (MAE) loss function**:

$$L = \frac{1}{KN} \sum_{k=1}^{K} \sum_{n=1}^{N} |f \theta(I_k[0], I_k[n]) - SUR_k(n)|$$
Prediction of SUR curve deep learning:  
$I[0]$, together with its distorted versions $I[1], \ldots, I[n]$, a sequence of predicted satisfied user ratios $\text{SUR}(1), \ldots, \text{SUR}(N)$ is obtained from the network. Assuming that the JND is normally distributed, we estimate the mean $\mu$ and variance $\sigma^2$ by least squares fitting:

$$(\hat{\mu}, \hat{\sigma}^2) = \arg\min_{\mu, \sigma^2} \sum_{n=1}^{N} \left| \Phi(n|\mu, \sigma^2) - \text{SUR}(n) \right|^2$$

3. **Algorithm: AROC-DPNet Lightweight Convolutional Neural Network and Deep Clustering for Hyperspectral Image Semi-supervised Classification with Limited Training Samples** [80]

**Input:** Small number of available labeled training samples $\{X^L, y^{(m)}\}$, $y^{(m)} \in \{1, 2\ldots M\}$, abundant unlabeled samples $X^U$ and test samples $X^T$. The number of max-iteration is $N$.

**Output:** The labels of $X^T$.

1. Pre-train 3D-DPNet with dual loss by Eqs. (1) and (2) on $\{X^L, y^{(m)}\}$.
2. Obtain deep features $F^L$ and $F^U$ for training samples ($X^L$ and $X^U$) and the primary labels $p^{(m)}$ of $X^U$, $y^{(m)} \in \{1, 2\ldots M\}$ through 3D-DPNet classification.
3. Project all training samples ($X^L$ and $X^U$) to the new deep feature space and implement AROC algorithm by Eq. (5) on $\{F^L, p^{(m)}\}$ and $\{F^U, p^{(m)}\}$, to obtain cluster labels $\{X^U, c^{(m)}\}$, $c^{(m)} \in \{1, 2\ldots M\}$.
4. Comparison of $\{X^U, c^{(m)}\}$ and $\{X^U, p^{(m)}\}$, and those with matched labels are considered to have high confidence and their corresponding pseudo labels $\{X^{Ci}, s^{(m)}\}$, $X^{Ci} \subset X^U$ are determined.
5. Fine-tune the pre-trained 3D-DPNet using $\{X^{Ci}, s^{(m)}\}$ and $\{X^L, y^{(m)}\}$.
6. Repeat 2 to 5, until $i = N$.
7. The labels of $X^T$ were obtained using the parameters of the last iterations of 3D-DPNet. Output the labels of $X^T$.

**10 Conclusion and Discussion**

- As the human species evolved human learned the things via sensing organs and learns, understand the daily or new activities day by day. As the evolution in the technology human tries to mimic the human brain behavior on computers with the help of some programmed algorithm. Machine learning is such a field which when combines with the computer vision solve most of the today’s problem. Here, in this book chapter, we have discussed and tries for find out some of the problem of computer vision with the use of ML algorithms. We have discussed how image compression, image enhancement and noise removal technique, automated segmentation like problem will be solved with the help of machine learning.
• Broadly Machine learning is composed of Supervised, Unsupervised, and semi-supervised ML. Algorithms like KNN, k-means falls into this category, there will be further improvement day by day in these algorithm as we have a lot of data for the training our model on different datasets.

• Image compression is one of the important areas of research now these days as we have better quality of cameras which are capable of capturing the high-resolution images. Thus, transferring and storing these images on the storage system, transfer the data from one place to another require a development of new algorithms thus Machine learning helps a lot and we can develop a new image compression algorithm.

• Image processing use machine learning a lot. Automated image segmentation to edge detection, color enhancement, and visualization is possible due to machine learning. In Medical Science image processing use a lot in X-RAY, MRI, and CT scans also. Image Recognition, visualization is also one more task in image processing, thus visualization of underwater object, human body organs detection, cancer/lesion detection, and recognition is popular field of computer vision and use machine learning algorithms a lot. Denoising of images is done through the machine learning. Now we have such an algorithm which can do the automated de-noising of noise from the noisy image.

• Due to machine learning algorithm like Neural network, genetic algorithm is trained on the medical image data thus it is become possible to diagnose the patient automatically. As a result, more accurate and precise diagnosis and treatment are possible in the hospitals. The increase in the data and computing power makes the automated diagnosis more easily and prediction algorithms help a doctor in diagnosis.

• Images captures form the satellite, mobile phone, or digital cameras are large in size, due to ML algorithm it is able to compress the size of image automatically without any loss of information is possible. Applying ML algorithm in image processing enhances the colors of the image.

• Facial expression analysis is done with the help of Deep learning classification. On the web there are facial expression datasets are available and via using these data set we can train our models or we can develop a new algorithm also. Facial expression algorithms read our facial feature based on these they predict the emotion or mood of human body.

• Thus, in this book chapter author has tried to provide the new advancement in the field of computer vision and artificial intelligence/machine learning. As the technology is advancing day by day there is explosion in the data also and we need some algorithms to compress, classify, label this huge amount of data having knowledge and a skill of machine learning and image processing one can gain develop new algorithm also for the better understanding of content between the images or videos.

Acknowledgements During this period of COVID-19 everyone life’s get disturbed, everyone is at their home or may be in safe place God bless them all. In this pandemic era I would like to
thank Mr. Rohit Shukla for helping hand in writing the English literature (Ph.D. scholar Dept. of Bioinformatics JUIT), Mr. Sameer for helping in drawing the Mathematical Equations (Ph.D. scholar Dept. of Mathematics JUIT), Ms. Shashi Kala (Asst Prof Dept of Bioinformatics KMV) for providing some resources for reading, and my Grandfather for providing me Internet Resources during this COVID-19 pandemic.

**Author Contribution**  In the era of pandemic of COVID-19, Mr. Ajay Sharma¹ (Ph.D. Research Scholar Dept of Bioinformatics JUIT) has prepared all the Manuscript of the book Chapter Most of the data, paper collection, and written part is done by the Ajay Sharma. Mr. Ankit Gupta² (Ph.D. Research Scholar Interactive Technologies Institute (ITI/LARSyS Madaria) has written the introduction part of the book chapter and helps in mining latest research articles for this book chapter. Dr. Varun Jaiswal (Ex. Assist Director, National Centre for Disease Control NCDC New Delhi, Current Assist prof Gachon University Dept of Biotechnology) has revised the prepared manuscript and find out the correction where it will be needed.

The book chapter contains some of the novel work. While planning and designing the book chapter entitled “Solving Image Processing Critical problems using Machine Learning” author should contain in mind that this book chapter helps those readers who don’t have any or little knowledge about image processing to those who are sound in imaging. Thus, keep in mind author has divided the book into several sections like easy, intermediate, and modest. In easy part, reader can understand the theory behind the image processing algorithm and machine learning and intermediate part contains the working and the modest part contains the mathematics algorithm that could be understood by those who have sound knowledge in image processing.

**Conflict of Interest**  The author has declared that there is no conflict of interest. The book chapter entitled “Solving Image Processing Critical problems using Machine Learning” has authentic work. The literature were mined from all the internet resources and while writing the book chapter author took the keen observation that proper credit is given to all the work via referencing and citation of research paper or book chapter.

**References**

1. D. Michie, D.J. Spiegelhalter, C. Taylor, Machine learning. Neural Stat. Classif. 1994(13), 1–298 (1994)
2. S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Supervised machine learning: a review of classification techniques. Emerg. Artif. Intell. Appl. Comput. Eng. 160, 3–24 (2007)
3. A. Kalousis, J. Gama, M. Hilario, On data and algorithms: understanding inductive performance. Mach. Learn. 54(3), 275–312 (2004)
4. M.-L. Zhang, Z.-H. Zhou, A k-nearest neighbor based algorithm for multi-label classification, in 2005 IEEE International Conference on Granular Computing (IEEE, 2005)
5. Y. Liao, V.R. Vemuri, Use of k-nearest neighbor classifier for intrusion detection. Comput. Secur. 21(5), 439–448 (2002)
6. D. Böhning, Multinomial logistic regression algorithm. Ann. Inst. Stat. Math. 44(1), 197–200 (1992)
7. B. Schölkopf et al., Shrinking the tube: a new support vector regression algorithm, in Advances in Neural Information Processing Systems (1999)
8. J. Ali et al., Random forests and decision trees. Int. J. Comput. Sci. Issues (IJCSI) 9(5), 272 (2012)
9. D. Yarowsky, Unsupervised word sense disambiguation rivaling supervised methods, in 33rd Annual Meeting of the Association for Computational Linguistics (1995)
10. G.E. Hinton et al., The “wake-sleep” algorithm for unsupervised neural networks. Science 268(5214), 1158–1161 (1995)
11. J.A. Hartigan, M.A. Wong, Algorithm AS 136: a k-means clustering algorithm. J. Roy. Stat. Soc. Ser. C (Appl. Stat.) 28(1), 100–108 (1979)
12. H. Li et al., Illumination compensation and image denoising for low-light images based on deep learning, in Proceedings of the 3rd International Conference on Vision, Image and Signal Processing (2019)
13. M. Poostchi et al., Image analysis and machine learning for detecting malaria. Transl. Res. 194, 36–55 (2018)
14. A. Kan, Machine learning applications in cell image analysis. ImmunoI. Cell Biol. 95(6), 525–530 (2017)
15. R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction (MIT Press, Cambridge, 2018)
16. J.X. Chen, The evolution of computing: AlphaGo. Comput. Sci. Eng. 18(4), 4–7 (2016)
17. O. Chapelle, B. Scholkopf, A. Zien, Semi-supervised learning (chapelle, o. et al., eds.; 2006) [book reviews]. IEEE Trans. Neural Netw. 20(3), 542–542 (2009)
18. X.J. Zhu, Semi-Supervised Learning Literature Survey (University of Wisconsin-Madison Department of Computer Sciences, 2005)
19. T.S. Huang, K. Aizawa, Image processing: some challenging problems. Proc. Natl. Acad. Sci. 90(21), 9766–9769 (1993)
20. P. Patidar et al., Image de-noising by various filters for different noise. Int. J. Comput. Appl. 9(4), 45–50 (2010)
21. B.T. Hannigan, B.A. Bradley, G.B. Rhoads, Removal of fixed pattern noise and other fixed patterns from media signals. Google Patents (2003)
22. S.O. Rice, Mathematical analysis of random noise. Bell Syst. Tech. J. 23(3), 282–332 (1944)
23. A.K. Vishwakarma, A. Mishra, Color image enhancement techniques: a critical review. Indian J. Comput. Sci. Eng. 3(1), 39–45 (2012)
24. V. Lopez-Vazquez et al., Video image enhancement and machine learning pipeline for underwater animal detection and classification at cable observatories. Sensors 20(3), 726 (2020)
25. M. Liu et al., Fuzzified image enhancement for deep learning in iris recognition. IEEE Trans. Fuzzy Syst. 28(1), 92–99 (2019)
26. Y.A. Bayhaqi et al., Fast optical coherence tomography image enhancement using deep learning for smart laser surgery: preliminary study in bone tissue, in Fourth International Conference on Applications of Optics and Photonics (International Society for Optics and Photonics, 2019)
27. J. Fenshia Singh, V. Magudeeswaran, A machine learning approach for brain image enhancement and segmentation. Int. J. Imaging Syst. Technol. 27(4), 311–316
28. K.J. Halupka et al., “One-size fits all” OCT image enhancement via deep learning. Invest. Ophthalmol. Vis. Sci. 60(9), 1513–1513 (2019)
29. M. Abdel-Nasser, O.A. Omer, Ultrasound image enhancement using a deep learning architecture, in International Conference on Advanced Intelligent Systems and Informatics (Springer, Berlin, 2016)
30. L. Ruihua, Z. Quan, X. Huachao, An image compression processing method based on deep learning, in 2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP) (IEEE, 2019)
31. T. Celik, Spatial entropy-based global and local image contrast enhancement. IEEE Trans. Image Process. 23(12), 5298–5308 (2014)
32. A.M. Ali, G.C. Karmakar, L.S. Dooley, Review on fuzzy clustering algorithms. J. Adv. Computations 2(3), 169–181 (2008)
33. S.M. Pizer et al., Adaptive histogram equalization and its variations. Comput. Vis. Graph. Image Process. 39(3), 355–368 (1987)
34. Y.-T. Kim, Contrast enhancement using brightness preserving bi-histogram equalization. IEEE Trans. Consum. Electron. 43(1), 1–8 (1997)
35. E.H. Land, J.J. McCann, Lightness and retinex theory. Josa 61(1), 1–11 (1971)
37. E.P. Simoncelli et al., Shiftable multiscale transforms. IEEE Trans. Inf. Theory 38(2), 587–607 (1992)
38. G. Beylkin, R. Coifman, V. Rokhlin, Fast wavelet transforms and numerical algorithms I. Commun. Pure Appl. Math. 44(2), 141–183 (1991)
39. D. Williams, P.D. Burns, Measuring and managing digital image sharpening, in Archiving Conference (Society for Imaging Science and Technology, 2008)
40. F. Russo, An image enhancement technique combining sharpening and noise reduction. IEEE Trans. Instrum. Meas. 51(4), 824–828 (2002)
41. J.D. Buhr et al., Digital photo-finishing system including scene balance, contrast normalization, and image sharpening digital image processing. Google Patents (2000)
42. W. Xu et al., Single image super-resolution based on global dense feature fusion convolutional network. Sensors 19(2), 316 (2019)
43. Q. Zhang et al., Super-resolution of single multi-color image with guided filter. J. Vis. Commun. Image Represent. 58, 277–284 (2019)
44. C. Xing et al., Gradation measurement of asphalt mixture by X-ray CT images and digital image processing methods. Measurement 132, 377–386 (2019)
45. R. Bostanabad, Reconstruction of 3D microstructures from 2D images via transfer learning. Comput. Aided Des. 102906 (2020)
46. M. Iliadis, L. Spinoulas, A.K. Katsaggelos, DeepBinaryMask: learning a binary mask for video compressive sensing. Digital Signal Process. 96, 102591 (2020)
47. C.D. Bahadrir, A.V. Dalca, M.R. Sabuncu, Learning-based optimization of the under-sampling pattern in MRI, in International Conference on Information Processing in Medical Imaging (Springer, Berlin, 2019)
48. M. Van Dael et al., Combination of shape and X-ray inspection for apple internal quality control: in silico analysis of the methodology based on X-ray computed tomography. Postharvest Biol. Technol. 148, 218–227 (2019)
49. G. Fang et al., Identification of corrosion products and 3D distribution in reinforced concrete using X-ray micro computed tomography. Constr. Build. Mater. 207, 304–315 (2019)
50. N.R. Pal, S.K. Pal, A review on image segmentation techniques. Pattern Recogn. 26(9), 1277–1294 (1993)
51. K. Sun, J. Wang, Robust local region-based active contour for inhomogeneous image segmentation, in Recent Developments in Intelligent Computing, Communication and Devices (Springer, Berlin, 2019), pp. 333–340
52. B.A. El-Rahiem et al., An efficient deep convolutional neural network for visual image classification, in International Conference on Advanced Machine Learning Technologies and Applications (Springer, Berlin, 2019)
53. C.L. Wong, Automated septal defect detection in cardiac computed tomography images. Google Patents (2019)
54. C. Jácome et al., Convolutional neural network for breathing phase detection in lung sounds. Sensors 19(8), 1798 (2019)
55. Z. Li, Rapid methods for 2D and 3D T1 estimation using highly undersampled radial MRI data (2019)
56. K.U. Devi, R. Gomathi, Brain tumour classification using saliency driven nonlinear diffusion and deep learning with convolutional neural networks (CNN). J. Ambient Intell. Humanized Comput. 1–11 (2020)
57. M. Zhang, Q. Zhou, Y. Hu, A lossless data hiding scheme in JPEG images with segment coding. arXiv preprint arXiv:1901.11203 (2019)
58. T.L. Saaty, The Brain: Unraveling the Mystery of How It Works: The Neural Network Process (Rws Publications, Pittsburgh, 2019)
59. D.S. Levine, Theory of the brain and mind: visions and history, in Artificial Intelligence in the Age of Neural Networks and Brain Computing (Elsevier, Amsterdam, 2019), pp. 191–203
60. P. Kubal, S. Doyle, Deep learning method for pathology image compression and reconstruction, in Laboratory Investigation (Nature Publishing Group, New York, 2020)
61. B. Sujitha et al., Optimal deep learning based image compression technique for data transmission on industrial Internet of things applications. Trans. Emerg. Telecommun. Technol. e3976 (2020)
62. M.Y. Yip et al., Effect of image compression and number of fields on a deep learning system for detection of diabetic retinopathy. Invest. Ophthalmol. Vis. Sci. 60(9), 1438–1438 (2019)
63. A. Mosavi, A. Bahmani, Energy consumption prediction using machine learning: A review. Preprints (2019)
64. B.S. Neysiani et al., Improve performance of association rule-based collaborative filtering recommendation systems using genetic algorithm. Int. J. Inf. Technol. Comput. Sci. 2, 48–55 (2019)
65. M. Pantic, L.J.M. Rothkrantz, Automatic analysis of facial expressions: the state of the art. IEEE Trans. Pattern Anal. Mach. Intell. 22(12), 1424–1445 (2000)
66. J. Weston, C. Watkins, Multi-class support vector machines. Citeseer (1998)
67. A. Raghuvanshi, V. Choksi, Facial expression recognition with convolutional neural networks. CS231n Course Projects, 362 (2016)
68. S. Haykin, Neural Networks: A Comprehensive Foundation (Prentice Hall, USA, 1994)
69. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in Proceedings of the International Workshop on Cognition for Technical Systems (2008)
70. X. Zhao, X. Shi, S. Zhang, Facial expression recognition via deep learning. IETE Techn. Rev. 32(5), 347–355 (2015)
71. S. Hershey et al., CNN architectures for large-scale audio classification, in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (IEEE, 2017)
72. A. Graves et al., Facial expression recognition with recurrent neural networks, in Proceedings of the International Workshop on Cognition for Technical Systems (2008)
73. A. Savoiu, J. Wong, Recognizing facial expressions using deep learning (2017)
74. Y. Bengio, Learning deep architectures for AI. Found. Trends Mach. Learn. 2(1), 1–127 (2009)
75. W. Shen et al., Deepcontour: A deep convolutional feature learned by positive-sharing loss for contour detection, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
76. T. Liu, G. Wang, Q. Yang, Real-time part-based visual tracking via adaptive correlation filters, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
77. P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4(7), 1073 (2009)
78. Y. Romano, J. Isidoro, P. Milanfar, RAISR: rapid and accurate image super resolution. IEEE Trans. Comput. Imaging 3(1), 110–125 (2016)
79. C. Fan et al., SUR-Net: Predicting the satisfied user ratio curve for image compression with deep learning, in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX) (IEEE, 2019)
80. B. Fang et al., Collaborative learning of lightweight convolutional neural network and deep clustering for hyperspectral image semi-supervised classification with limited training samples. ISPRS J. Photogrammetry Remote Sens. 161, 164–178 (2020)