A REVIEW ON IMAGE PROCESSING USING ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS

THESIS SUBMITTED

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE OF

MASTER OF ENGINEERING

IN

ELECTRONICS & TELECOMMUNICATION ENGINEERING

By

SOUMIK BANERJEE

Examination Roll No: M4ETC19017

Registration No: 140700 of 2017-18

Under the guidance of

Prof. SHELI SINHA CHOUDHARI

Department of Electronics & Telecommunication Engineering

Jadavpur University, Kolkata – 700 032

West Bengal, India

May-2019

Faculty of Engineering & Technology
Jadavpur University

This is to certify that the thesis titled **—A REVIEW ON IMAGE PROCESSING USING ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS** has been carried out by **SOUMIK BANERJEE** (Class Roll No: 001710702015, Examination Roll No.: M4ETC19017 and Registration No: 140700 of 2017-18) under my guidance and supervision and may be accepted in partial fulfillment of the requirement for the degree of Master of Electronics & Telecommunication Engineering.

Prof. Sheli Sinha Chaudhuri
Supervisor
Department of Electronics and
Telecommunication Engineering
Jadavpur University
Kolkata-700032

Prof. Sheli Sinha Chaudhuri
Head of the Department
Electronics and Telecommunication
Engineering
Jadavpur University
Kolkata-700032

Prof. Chiranjib Bhattacharjee
Dean
Faculty Council of
Engineering and Technology
Jadavpur University
Kolkata-700032

Faculty of Engineering & Technology

Jadavpur University

CERTIFICATE OF APPROVAL

The forgoing thesis titled **—A REVIEW ON IMAGE PROCESSING USING ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS** is here by approved as a creditable study of an engineering subject carried out and presented in a manner satisfactory to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or accept every statement made, opinion expressed or conclusion drawn there in, but approve the thesis only for the purpose for which it has been submitted.

Committee on Final Examination for
Evaluation of the Thesis:

Additional Examiner

Superviser

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that this thesis contains literature survey and original work by the undersigned candidate, as part of his Master of Electronics and Telecommunication studies.

All information in this document have been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name	**:**	**SOUMIK BANERJEE**
Exam Roll No	**:**	**001710702015**
Thesis Title	**:**	**A REVIEW ON IMAGE PROCESSING USING ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS**

Signature of Candidate

CHAPTER 1

INDRODUCTION

Neural networks and artificial intelligence are becoming progressively important in the present time due to rapid technological advancements. Artificial intelligence is being installed in robots and these robots are being trained all over the world using neural networks. Now these robots can almost work like human beings. After the various researches done by several universities in the United States of America, the secretariat office and the Oxford University has finally declared that fifty percent of the jobs in the United States of America will be eaten by robots by the year two thousand thirty. The jobs that involve automation are only going to sustain i.e. the jobs that can be done by robots are only going to sustain and remaining jobs are going to vanish. The people who can club the dynamics of human beings and computers are only going to be successful. The government of the United States of America is trying to sensitize science, technology, engineering and mathematics. The automation process has already started in the United States of America and shall reach India within twenty years .The jobs that are physically predictable i.e. the jobs that doesn’t involve more stress i.e. data processing, data collection jobs etc will be lost and the jobs that are physically unpredictable i.e. the jobs involving more stress and the jobs that involve people skills, managerial skills etc are going to sustain. The companies that have technical potential for automation are going to survive and remaining companies are going to shut down. Recently Amazon go company has brought a revolution. Suppose if a customer wants to buy something from the Amazon store, he/she can enter into the Amazon store using the Amazon mobile app, pick up anything he/she wants to buy, put it in the basket and leave the store. After the customer leaves the store, the money will be automatically deducted from the customer’s bank account. All the companies will be hiring robots for their work in future because it will not be required for them to give monthly salary to the robots and they will be saving lot of money. New jobs like up gradation jobs, technical maintenance jobs, and technical management jobs of robots will be generated. Artificial intelligence is a profound science and the dream of a faster society, where a work can be executed within seconds is trying to be accomplished by using artificial intelligence. The speed of execution and speed of taking decisions will be at its peak by using artificial intelligence so that life becomes very easy and a big project gets completed in very less time. A project that takes three years to complete will take only thirty days to complete by robots. Robots are also being used in medical sciences. Seven and half lakh robotic surgeries have already been done in the year 2017. The robots will also be used in military so that they can fight war. Due to this, human beings will not have to lose their lives. The main aim of this thesis is to get acquainted with artificial intelligence and neural networks and its application to image processing so that survival in future may become easy.

1.1 Review of Image Processing

Every image can be represented in the form of a two dimensional function f(x, y). If the values of x and y are continuous in nature i.e. x and y exists for the values between $-\infty$ to ∞ then it is said to be an analog image. Any point (x, y) of the function f(x, y) is the intensity of the pixel in the image. In this two dimensional analog function, x and y varies between $-\infty$ to ∞. Therefore, it can be said that an analog image will have infinite number of intensity values varying between $-\infty$ and ∞ . In order to process an analog image, a digital computer is needed and it is impossible to represent infinite number of intensity values with infinite number of bits in a digital computer. Therefore, digitization of an image is needed for its processing. The digitization process is done with the help of sampling and quantization which is the process of converting an analog signal into a digital signal. In the digital domain, instead of taking every point of x and y as it was in the case of an analog domain, integral values of x and y are taken, that too in finite number and it is represented in the form of a finite dimensional matrix. For example a digital image can be represented in the form of an N dimensional matrix as shown below

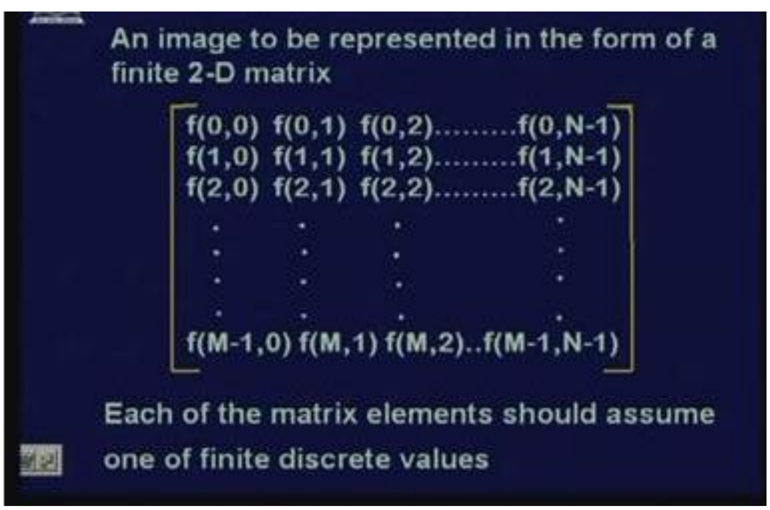

Figure 1: Representation of Digital Image

Digital image processing is motivated by three major applications. One is improvement of pictorial representation for human perception i.e. whatever image is obtained; the quality of image is improved. The second application is for automatic machine applications. The third application for digital image processing is efficient storage and transmission. For example if an image has to be stored in computer it needs certain amount of disk space. By using digital image processing, the disk space required for storing the image will be less compared to analog image as the number of bits required is finite. Image processing can also be used to transmit the image or video signal over a low bandwidth communication channel. Typical applications of image processing to improve the pictorial representation for human perception is Noise filtering, content enhancement which includes contrast enhancement and de blurring and remote sensing. In some of the cases the images that are obtained may be very noisy, so the noise filtering is done to remove noise from the image.

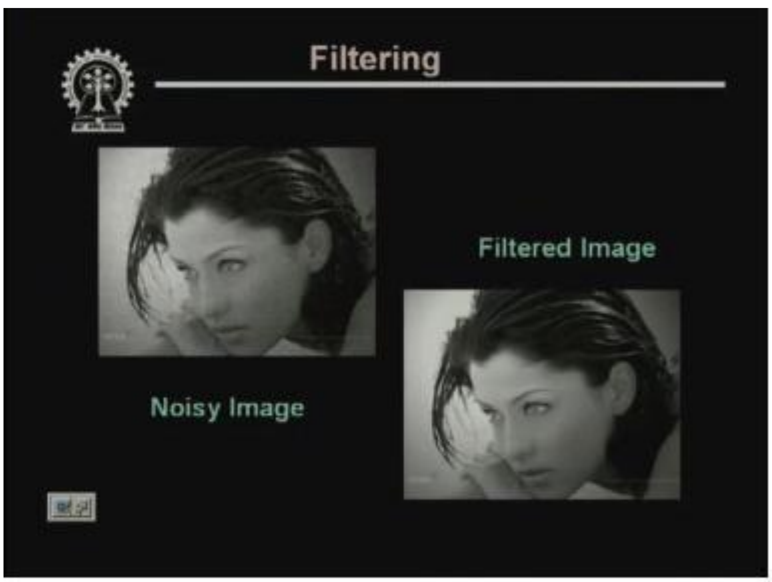

Figure 1.1: Filtering of noise

In some other kind of applications, the image has to be enhanced and for this, different applications like contrast enhancement, where the images having very poor contrast are enhanced. The image may also look blurred due to improper setting of camera or due to de focused lens. So image processing can be used to de blur the image. The other kind of blurring may take place due to moving platform i.e. the images might be taken from a moving car or moving train and due to this the image looks blurred. This type of blurring is called motion blur.

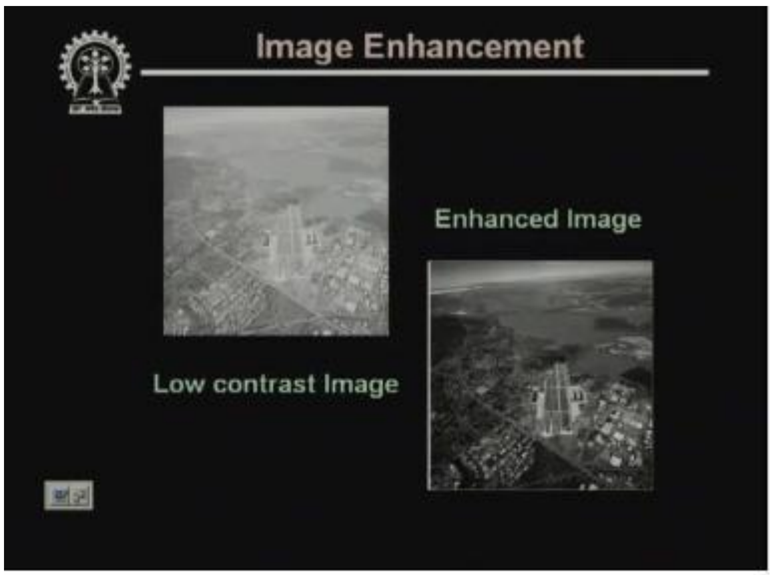

Figure 1.2: Image Enhancement

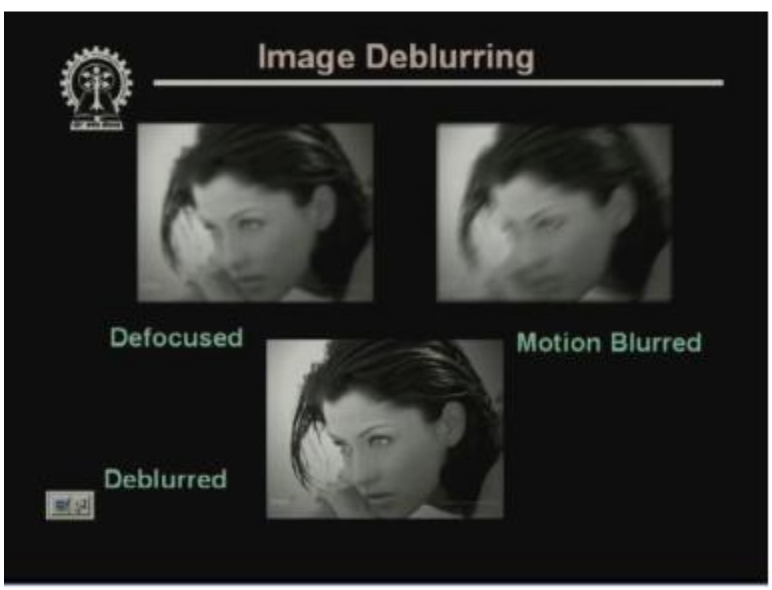

Figure 1.3: Image De blurring

The other kind of image processing application is the remote sensing where the image is an aerial image, taken from the satellite.

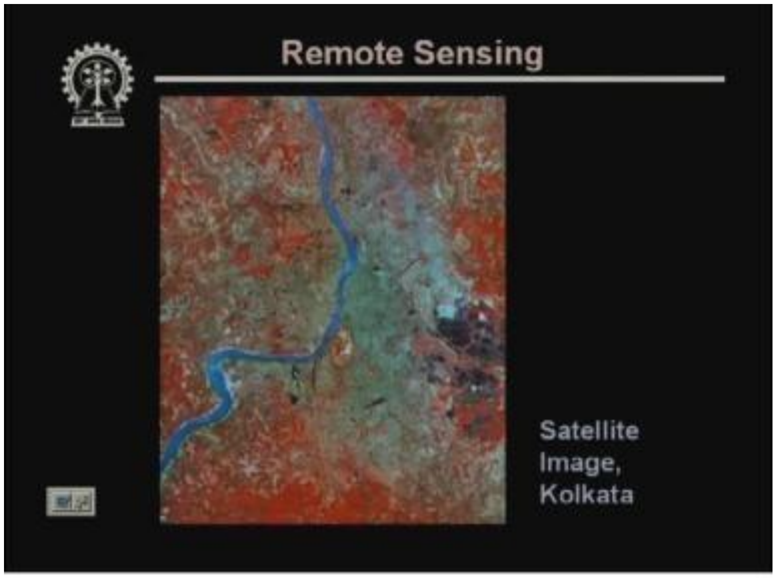

Figure 1.4: Remote Sensing of a region

The above image is a satellite image taken over the region of Kolkata and much information is present in the image. The blue thick line shows the river Ganges. The different color coding indicates different regions of Kolkata. If a remote sensing image or an aerial image of this form is present, various things can be studied like whether the river has changed its path or about the growth of vegetables in certain region or about the pollution in certain region etc. The remote sensing images can be used for planning a city. Suppose it is needed to build a city over a certain region, then these remote sensing images will be much helpful. One can determine, where a residential area has to be grown and where the industrial area has to be grown and through which regions the paths have to be formed, roads and car parking regions have to be constructed can all be planned using remote sensing images.

The Remotely sensed images can also be used for terrain mapping. The figure in the next page shows the terrain map of a hilly region which is not accessible easily. Then the images of this region can be taken with the help of satellite and then processed to find out the three dimensional terrain maps.

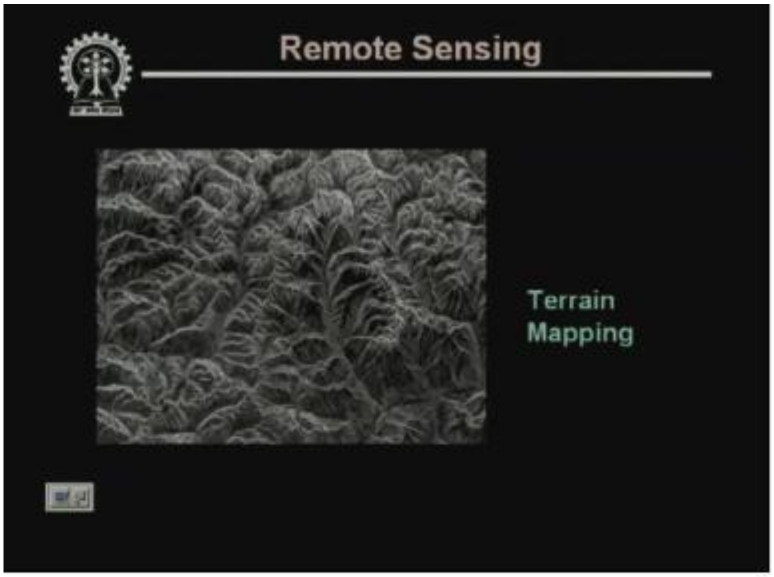

Figure 1.5: Terrain Mapping

The remote sensing images can also be used to check the region that are caught with fire and in which direction the fire is moving and once this is identified, the loss that has incurred by the fire can be found out. People can be warned beforehand about the direction of the fire, so that they can take precautionary measures against it and save their lives. The figure in the next page shows the fire that was caught in Borneo.

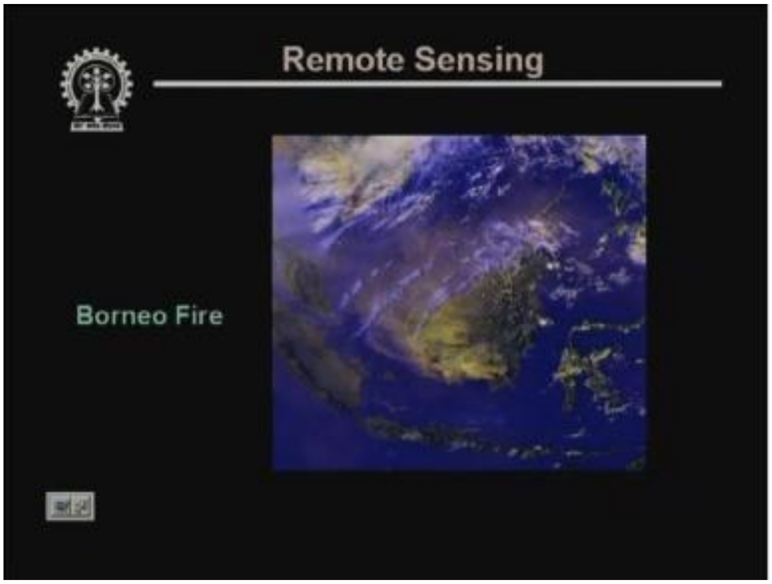

Figure 1.6: Remote sensing to detect fire

The image processing techniques are also very important for weather forecasting applications. On a television channel when the news about the weather forecasting is shown then in that case on a map the images are overlapped which tells about the cloud formation over a certain region. This gives an idea about the occurrence of rain or storm or clear weather. The image below shows the formation of hurricane over the Dennis in the year nineteen ninety. The strength of the hurricane and the precautionary measures that has to be taken to save the lives can be found out by looking at the image.

Figure 1.7: Weather Forecasting

Image processing can also be used for atmospheric study. The image below shows the formation of the ozone hole. Ozone layer is the protective layer in the stratosphere region that protects the atmosphere from ultra violet radiations. Whenever ozone hole is formed, the ultra violet rays enter the atmosphere and can cause damage to everyone. So, the image processing can be used to find the region where the ozone hole is formed so that people in that region can take some precautionary measures to protect themselves from ultra violet radiations.

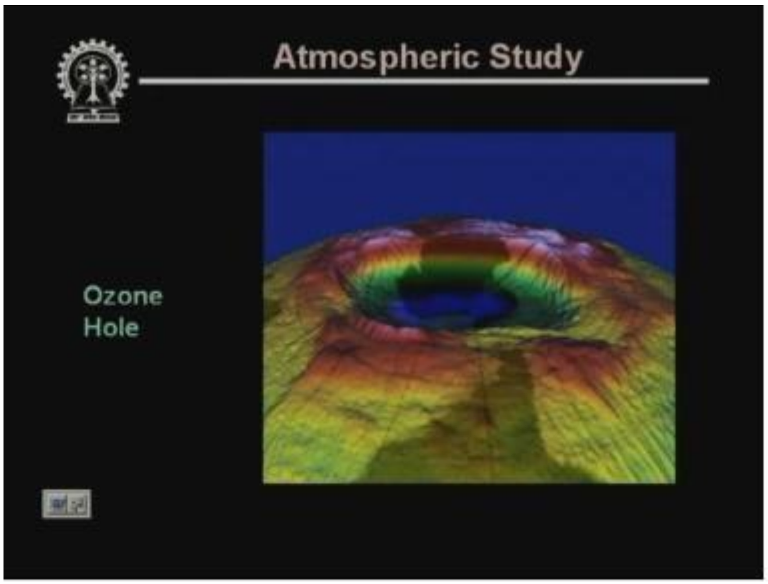

Figure 1.8: Ozone hole formation

Image processing is also useful for astronomical studies. The Image in the next page shows the star and galaxy formation.

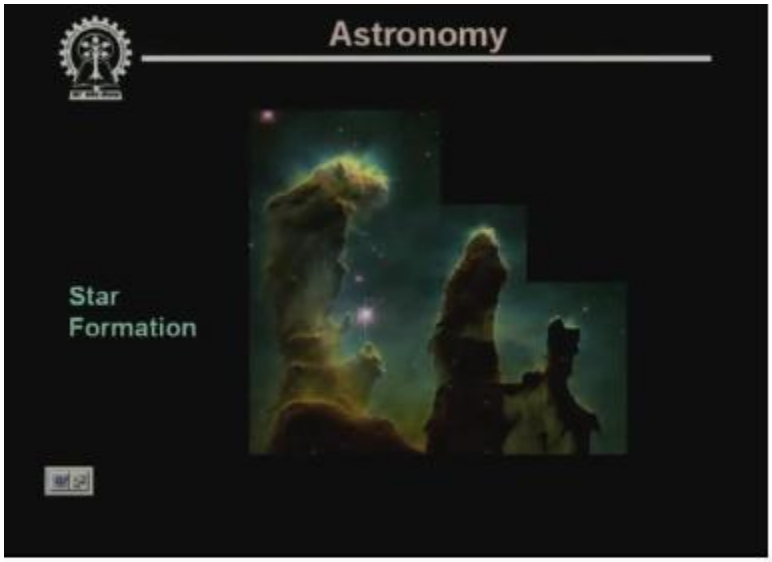

Figure 1.9: Formation of Star

Figure 1.10: Formation of Galaxy

The digital image processing techniques are also useful in the area of medicine. The image in the next page shows the CT scan images of the brain and is used to find out the location of the tumor. The left image is the original image and the middle and right images are the processed image. The yellow and red region in the processed image gives the location of tumor in the brain. These kinds of images in image processing techniques are very important because the doctors can find out the exact location of the tumor, the size of the tumor and many other things which can help the doctor to plan the operation process.

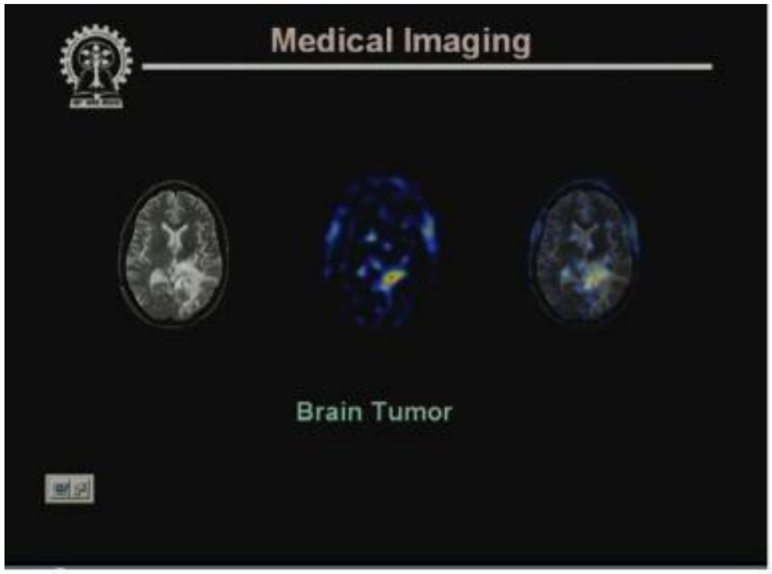

Figure 1.11: Location of brain tumor

Another application of image processing in medical field is the cancer detection. The image below shows a mammogram image which shows the presence of cancerous tissues.

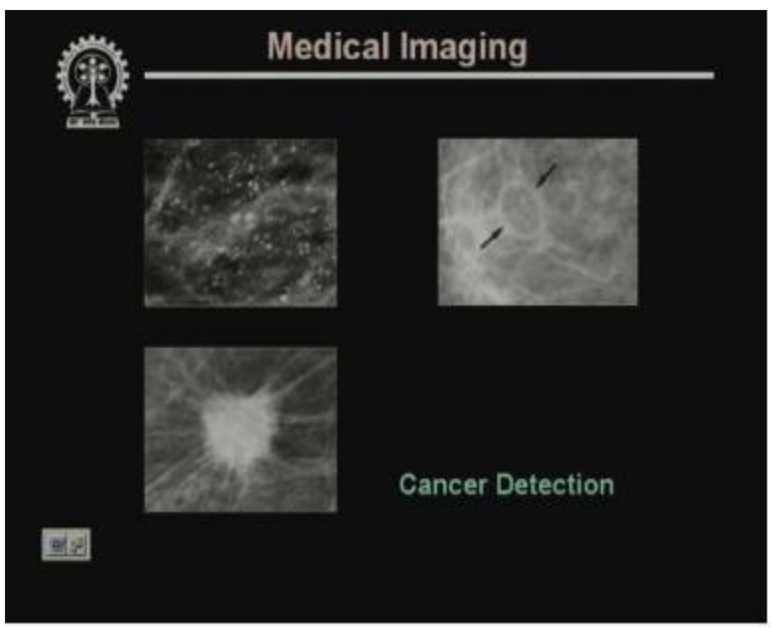

Figure 1.12: Detection of Cancer

The next application of image processing techniques is the machine vision application. It is found that all the earlier applications seen in the previous paragraphs has shown the purpose of visualization, the improvement of the visual quality of the image so that it becomes better for human perception but when it comes to machine vision application, the purpose of image processing techniques is different. Here, improving the visual quality of the image is not of much interest but the processing of the image is done to extract some description or some features which can be used for further processing by a digital computer and such kind of processing can be applied in industrial machine vision for product, assembly and inspection. It can be used for automated target detection and tracking. This can also be used for finger print recognition and for processing of aerial and satellite images for weather prediction, crop assessment and many other applications. The image below shows the application of image processing in bottling plant automation.

Figure 1.13: Bottling Plant Automation

The plant fills up the bottle with some liquid and after the bottle is filled up; it is carried by the conveyer belts and finally sent to the customers. So here, checking the quality of the product is very important and in this particular application, the quality of the product indicates whether the bottles are filled properly or some bottles are coming out empty or partially filled. So naturally, the application is, if it can be found out that some bottles are partially filled or some bottles are empty. Then naturally those bottles need not to be delivered to the customers because if the customer gets such bottles, then the reputation of that company will be lost. So, detection of the empty bottles or partially filled bottles is very important and here image processing techniques can be used to automate this particular process. From the image in figure 1.13 the bottle in the middle is partially filled and remaining bottles are completely filled. So naturally, this particular bottle has to be detected and removed from the production line so that finally when the bottles goes to the customer, no empty bottle or no partially filled bottle are given to the customers.

Another application of machine vision is the boundary information. The image below gives the information about the boundary of an animal. There is no other information available in the image except the boundary contours and if it is asked to identify this animal then everyone can say that this is a giraffe. So, even when there is no other information available except the boundary of the animal then also it can be identified. Therefore, the boundaries contain most information about the objects present in a scene

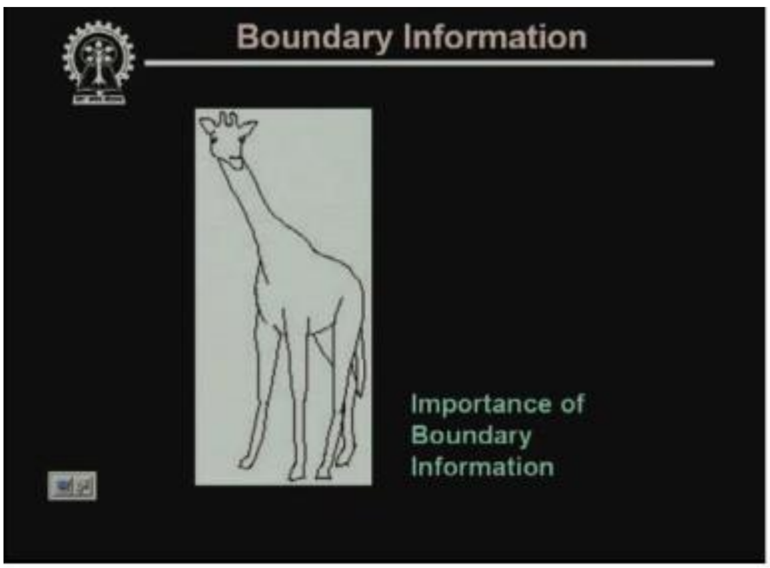

Figure 1.14: Boundary detection of a Camel

From the boundary information, various applications of image processing techniques can be developed. The figure below contains four different images. The first one is the original image; the second one is the threshold image or segmented image. The third image from the left bottom is a field image. The second image is not very smooth and there are number of black patches over the white region. The fourth image shows the boundary of the object. Therefore, the separation of background from the object is done in the image and this is very important in image processing applications.

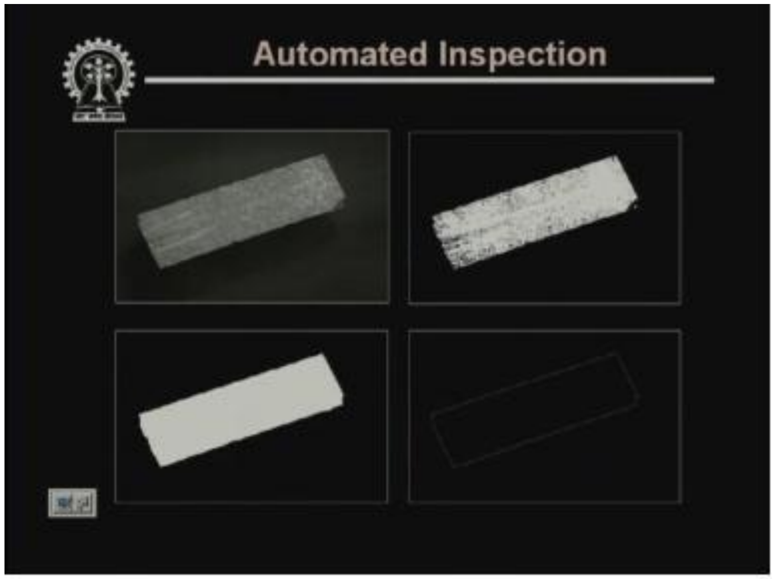

Figure 1.15: Separation of Image from boundary

One of the applications for automatic inspection can be inspection of IC manufacturing shown in figure below. The first image has a broken bond, whereas in the second image some bond is missing which should have been there. So, naturally these are the defects which ought to be identified because otherwise if this IC is made, then the IC will not function properly.

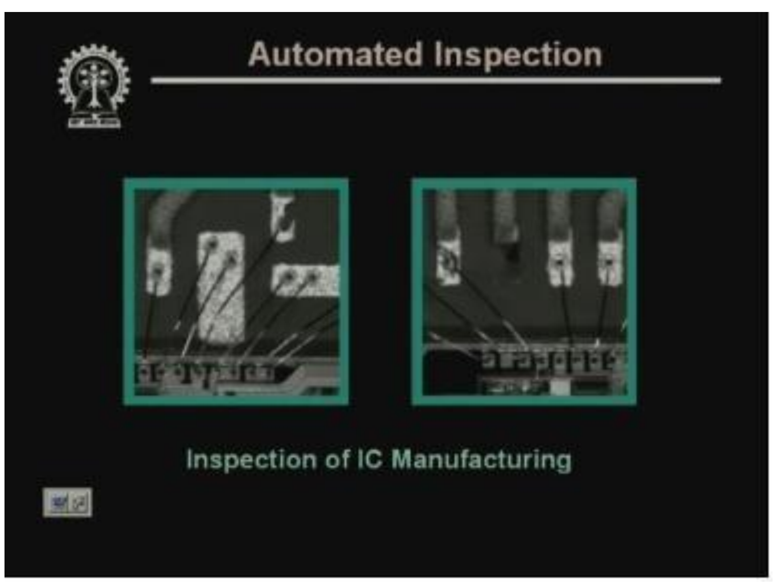

Figure 1.16: Inspection of Defective IC

Another application of image processing is the video sequence processing. When the image frames are displayed one after another and if there is any movement in the image, that movement is clearly detected. So, the major emphasis in image processing in a sequence processing is to detect the moving parts. This has various applications. For example, detection and tracking of moving targets and major application is in security surveillance. The other application can be to find the trajectory of a moving target and also monitoring the movement of organ boundaries in medical applications is very important. All these operations can be done by processing video sequences. The image below shows movement detection example. In the first image, some person is moving against a green background. Through image processing techniques, this movement can be identified. In the second processed sequence, the person is moving against a black background. That means, the background from the moving object has been separated.

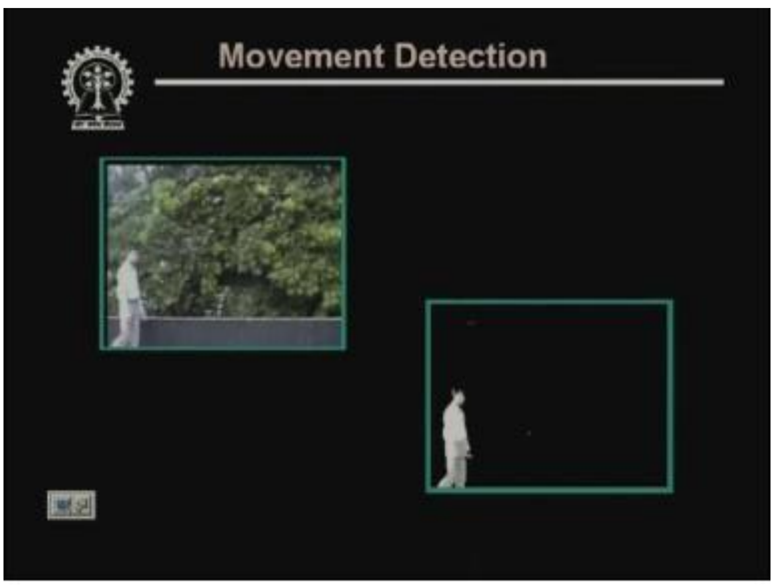

Figure 1.17: Video sequence processing

So these are the some of the applications of digital image processing.

The fundamental steps of digital image processing are shown below.

1) Image acquisition
2) Image Enhancement
3) Image Restoration.
4) Color Image Processing
5) Wavelets and Multi resolution Processing
6) Image Compression
7) Morphological Processing
8) Image Segmentation
9) Representation and description
10) Object Recognition

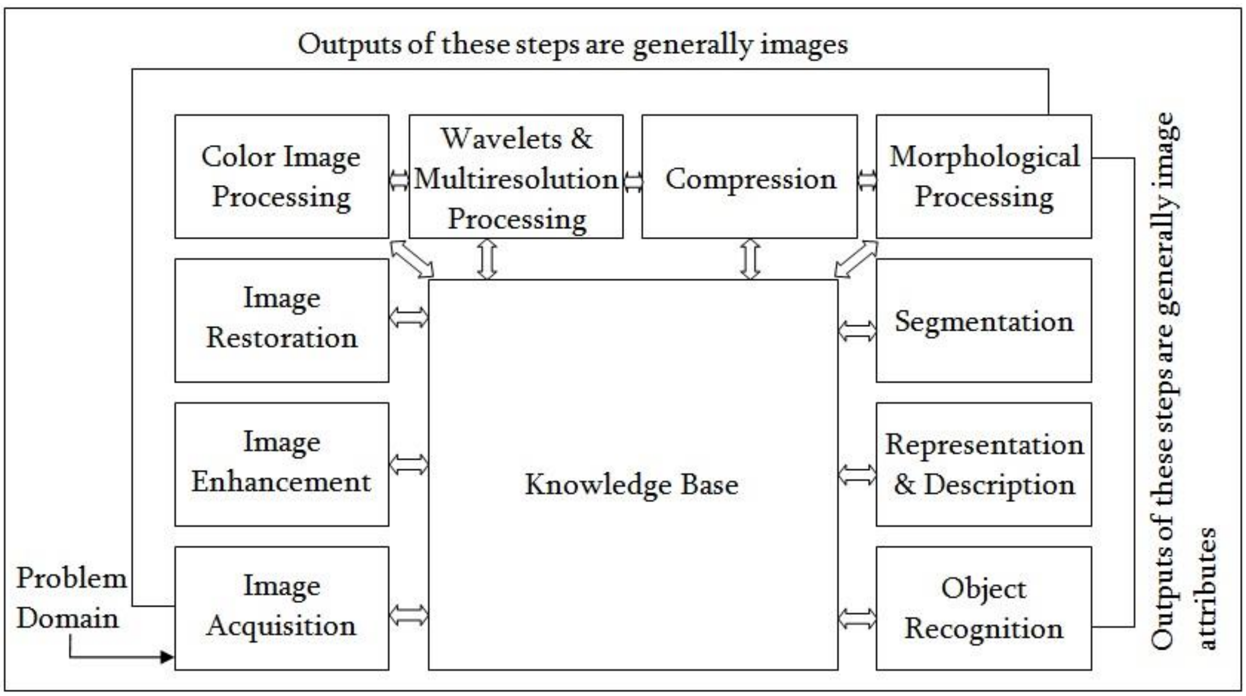

Figure 1.18: Block diagram of Digital Image Processing

Image acquisition: Image acquisition is basically retrieval of an image from a particular source for the purpose of processing it.

Image Enhancement: It is a technique of processing an image to enhance certain features of the image i.e. to improve the quality of the image. There are different image enhancing techniques for various features that is needed to be enhanced. Removal of noise from an image is one of the image enhancement techniques. The main goal of image enhancement is to get a certain quality of image that used for some specific applications. The image enhancement technique fall under two categories i.e. the spatial domain technique where the pixel values of the signal are directly manipulated and the other category is the frequency domain technique where the Fourier transform of the image is found out and the Fourier coefficients are modified and then the inverse Fourier transformation is found out and hence the modified image is received.

Image Restoration: Image restoration becomes important when the image becomes degraded. An image degradation model is found out for image restoration and once the degradation model is found out it becomes very easy to restore the image. In case of image enhancement, image degradation models are not considered but in case of image restoration the image degradation models are very important. A degradation model has a degradation function that is multiplied with the two dimensional image function and then it is added with noise. After this a degraded image is obtained that is passed through low pass filters to restore the original image. The nature of degraded image can be found out by observation, experimentation and some mathematical modeling methods.

Color Image Processing: Color image processing is important in image processing studies even when the information can be obtained from the black and white image itself .The reason is color is very powerful descriptor and using the color information one can extract the objects of interest from an image very easily which is not so easy in case of black and white image. For example if we want to differentiate between petals of a rose and a leaf , it is very difficult to do it in case of a black and white image but if the image is a color image then it is very easy to differentiate. The second motivation to study color image processing is that a human eye can distinguish between thousands of colors and colors shades where as in case of a black and white image one can differentiate number of gray shades. There are two major areas in color processing one is the full color processing and other is the pseudo color processing. In case of full color processing the images that one can acquire from a full color TV camera contains all the colors that one can perceive. In case of pseudo color processing the color is assigned to certain gray level range in the image. For example if a gray level image is considered in the intensity range between 0 and 255, one can subdivide this range into different segments where one can assign different colors to these segments. In case of a black and white image it is possible to distinguish between dozens of gray shades in the image but it is not possible to distinguish between gray shades whose intensity values are near to each other. In such case pseudo color processing can be used where one can extract information from images very easily.

Wavelets and multi resolution processing: Wavelets are rapidly decaying wave like oscillations that has zero mean and finite duration unlike the sinusoids which extend for longer duration.

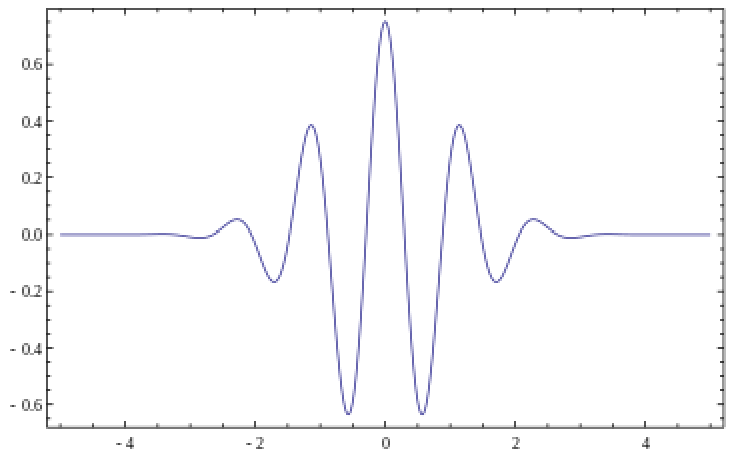

Figure 1.19: Wavelet

Here X axis denotes time in seconds and Y axis denotes the amplitude (can be anything i.e. volts, amperes, watt etc)

Suppose that there is an image with abrupt change in contrast and this abrupt change is represented with a signal. The wavelet is compared with the original parts of the signal and a similarity between the coefficients and wavelet are found and is denoted as c. The wavelet is compared with the remaining parts of the signal and number of c's are collected which are called wavelet coefficients. 'c' is calculated by taking the dot product of original signal vector and wavelet vector basically called correlation. So, wavelets can be used for detecting the minute change in the contrast details. Wavelets are foundation for representing an image in various degrees of resolution. Resolution refers to the number of pixels in the image. A multi resolution representation of an image gives the idea about the extent of details existing at different locations from which the requirements of desired details are chosen. It facilitates efficient compression by exploiting the redundancies across the resolutions.

Image Compression: Image compression is a type of data compression that is applied on digital images to reduce the cost of storage or transmission. Image compression can be lossy and lossless. In the lossy compression method the important features of an image may be lost during its compression while in case of lossless compression the important features of an image will be saved. Lossy images are suitable for photographs where minor loss is acceptable to achieve a reduction in bit rate.

Morphological Processing: Morphology is a term widely used in biology that tells us about the shape and structure of different plants and animals and so on. In image processing application this term morphology is used to discuss the structure and shape of objects. So, the image processing techniques based on structure and shape of objects are classified as morphological operation. This is nothing but mathematical morphology in image processing. Suppose there is an image which contains an object of white color with a black colored background. The goal is to separate the object from the background region. A simple threshold operation can be applied and after applying the threshold, the white pixel will belong to the object region and black pixels will be separated from the white pixels. These black pixels are considered as noise and can be separated by using morphological filters. So morphological operation is a preprocessing operation which is used to remove noise from the image. Morphological operation can be used for shape simplification. Suppose there is an image containing an object of very complicated structure but it is possible to break this structure into number of sub structures where each of the structures will be simple in nature so that it is easy to quantify the shape attributes of that particular structure. In such cases morphological operations will help to simplify the shape of a complicated structure.

Image Segmentation: Image segmentation is basically subdividing an object into constituent parts or objects in the image. The main purpose of sub dividing an image into constituent parts is that one can further analyze each of these constituents in the image once they are further identified. The information that is extracted by segmenting an image is useful for machine language applications.

Suppose an image consists of number of vehicles on a busy road and one is interested to find out the moving pattern of the vehicles. The first level of subdivision will be to extract the road from the image .Once the road is identified further analysis of road is done so that one can identify every single vehicle on road and once the vehicles are identified one can go for vehicle motion analysis. There may certain other information in the image about the residential complex or water bodies but these can be ignored. The only interest is to identify the road and segment it so that each and every vehicle and its moving patterns are identified. The vehicles are not further sub divided in this case and the segmentation process is stopped. Image segmentation is a very difficult task to do but it is very important task. This is useful in machine language application and the application software should have a very robust image segmentation algorithm. There are two approaches for image segmentation for image segmentation. One is the discontinuity based approach and the other is the similarity based approach.

In the discontinuity based approach the subdivision is carried out based on abrupt change in the gray levels of an image. So in the discontinuity based approach the identification of the edges is done. The similarity based approach is slightly different. In this case the grouping of similar pixels is done. The simple approach for similarity based approach is the thresholding operation. Suppose if a gray level image is considered that has the intensity range between 0 to 255 a threshold can be assigned which in between 0 and 255, say 128. So the pixels having intensity values less than 128 will belong to some region and the pixels having values greater than 128 will have will belong to some other region.

The other kind of segmentation in this similarity based approach is the region growing approach. In this case a pixel is considered and the neighborhood of this pixel having similar intensities values is grouped together. So a region is grown in this case by connecting the adjacent pixels having similar intensity values. One more approach in this case is the region splitting and merging. The image is first split into a number of different components following some criteria and after the splitting is done the sub images that are having similar intensity values is merged together. So it can be said, that the processor section of the digital image processing system will perform the operations like image segmentation, enhancement, registration, color image processing etc.

Representation and description: The main aim of this method is to make the computer understand about the segments in an image. In order to do this, proper description of the image segments must be present so that the descriptions present in the image may match against some prior knowledge stored in the computer. So, once such descriptions are generated from the segmented image, these descriptions can be matched with the prior knowledge present in the computer in the form of knowledge base and once the descriptions are matched the computer can immediately infer that the segments belong to some objects.

Object Recognition: Object recognition is a computer vision technique used for identifying objects in images or videos. When humans look at a photograph or a video they can readily spot people, objects, scenes and visual details. The goal is to teach computer to do what naturally comes to humans and gain a level of understanding of what an image contains. Object recognition is a key technology behind driverless cars enabling them to recognize a stop sign or differentiate between a pedestrian and a lamp post. It is also useful in applications such as disease identification in bio imaging. Bio imaging is an imaging technique used to study different tissues and organisms

Knowledge base: This block contains prior knowledge of some portion of an image that is to be further processed after its acquisition.

1.2 Review of Neural Networks and artificial intelligence

The artificial neural networks have matured to a great extent over the past few years. With the advent of very high performance computing, it has assumed a tremendous significance and has got very big application potential in very recent years. The neural network derives its origin from the human brain or the human nervous system which consist of a massively large parallel interconnection of large number of neurons that achieves different tasks like recognition tasks etc in an amazingly small amount of time as compared to today’s very high performance computers. So, this is what inspired the researchers to think of a solution where a computer can be made to mimic the large amount of interconnections and the networking that exists between all the nerve cells can be utilized to do some complex processing tasks which even today’s high performance computers also cannot. Now, whenever one say’s artificial, the question that may immediately come to the mind is that, what can be the natural neural network. One can say the human brain is a natural neural network that is having a highly complex and nonlinear parallel structure that can organize its constituent structural elements. So, the structural constituents of human brain are known as neurons and the neurons are interconnected not in a very simple way, but in a rather complex way and so complex that many of the things are not known yet. Say, if there are large numbers of such neurons or nerve cells that carry out the processing, they will be interconnected typically in a highly complex manner between each other; there will be connections existing from one neuron to the other and that is how a network is realized. As a human brain has got typically billions of nerve cells with trillions of such interconnections existing it can be understood with the help of a recognition task. Whenever someone meets a person who is known to him, he can recognize the name of that person because the person must be his friend. He will be able to tell who that person is whenever they both meet. Now, one may know thousands of people around. If one has to recognize a person whom he has met five years ago at some work place or may be the person was his colleague, then he can immediately recognize the face of that person when he meets him. Now, supposing that this task has been given to a computer. Then to find out how much time the conventional computer would take to accomplish this task efficiently will not be easy for a computer to say because the computer does not know that person. Firstly, it should be taught to the computer that there are photographs of different people. Suppose if a person knows thousand persons then he should be feeding the photographs of thousand people into the computer. Now if the photograph of a person is captured and then fed to the computer, the computer will meticulously perform the recognition task by trying to match the given image with that of the stored photographs after it is trained to perform the recognition task. Then, ultimately at the end of all the computations and comparisons with photographs of thousands of people which is stored in the computer data base, the computer is going to give the results about the person that best resembles the photograph. Now, whether this could take long time or not that would all depend upon how many such images are present in the database. If the database is very large and if a person has to be tracked down from the collection of let say ten thousand images then the task is going to be really complex and is going to take a lot of time. Now if this task has to be done instantly, then the computational capability that exists in the humans is enormously different from the way a computer is doing. If the computation is seen in terms of the processing speed, different types of result will be obtained. The response time of today's silicon IC is expressed in terms of nano seconds. A nano second is time which is 10 to the power minus 9 of a second, whereas if it is compared with the processing speed of a human neuron it is 5 to 6 times slower than that of a typical IC, that is it may take several milliseconds which is 10 to the power minus 3 of a second but still neural processing is faster than computer processing because the network of human neurons in the brain that is massively parallel of the order of something like

10 billion is having approximately 60 trillion interconnections. It is not that easy to mimic the tasks that a human brain does using the electronic components or using computer software because even in the age of parallel computers it cannot be really thought of putting so many processing units and realizing it in a massively parallel scheme. All that one can do within the limitation is interconnect a network of processes and rather than considering the structure of a human brain in totality it can be tried to mimic extremely small part of it in order to do some very specific task i.e. the best thing that can be done using electronic components and software's. The neurons can be made which will surely not be different from the biological neurons. So, the artificial neural networks can be studied. By artificial it inherently means that something that is different from that of the natural or the biological neuron. Now, some of the usefulness and capabilities of artificial neural networks is the exploitation of nonlinearity. Basically, if there is a system to which a set of inputs are given and some output can be expected out of it then in that case the system is said to be linear and the relation between the output and the input can be best described in terms of a simple linear equation. If there are let say 4 inputs and 1 output, then if the output is a linear combination of all the 4 inputs, then naturally the system is linear. Whereas, if the output is written in terms of not only the linear terms but also its higher order terms then in that case the system is no longer linear and it becomes non-linear. A lot of times, for simplicity the linear computational models are considered but by looking at the real life problems, most of the real life problems happen to be highly non-linear in nature. So, for that purpose, non-linear computational units are needed as well the neurons that also happen to be non-linear. So, there is an interconnection of non-linear neurons in the non linear computational units. Therefore, in artificial neural networks, interconnections of non-linear neurons are present and another thing which should be noted is that the nonlinearity is distributed throughout and in fact the very nature of computation one can realize is it's highly distributed in nature. So, the nonlinearity that is used is naturally distributed. The second usefulness is the input output mapping. If some input to a system is provided then in response it is going to get some output. It means that some learning mechanism can be adopted where a teacher has to be involved and what can be done is that, the inputs can be fed and then it can be said what the expected output is going to be. So, in other words it is being specified that for a given input, what is going to be the output or the desired response. Now, it is possible that the computational unit that is present is not able to achieve the actual output that one should get; it may be different from that of the desired output.

There may be difference between what is actual and what is desired. In this case the modifications of the system parameters can be done by adjusting the parameters. So, if such kind of free parameters are present, then one should be able to adjust the free parameters of the systems such that for a given input or a set of inputs, one can obtain the output that is closest to the desired output. In that case it may not be possible to achieve that immediately. First time a pattern is fed and the system has not encountered that pattern before. So, the difference that exists between the actual output and the desired output is the adjustment of the parameter of the system such that the difference between the actual and the desired output is minimized; and that it may be done several times. So, there is a process of learning and this learning involves a teacher, who says that corresponding to a particular input a particular output should be generated and if it is not, then the teacher will asking to correct i.e. adjust the free parameters which are available in the system so that next time when the input is fed, the output obtained should be closer to that of the desired. So, this is something which is very important and this is what is going to make the neural network remarkably different from the conventional computational unit with a sense that it has got a learning ability. In this case the input output mapping that is being referred to is learning with a teacher where the teacher is definitely monitoring. It should be noted that there may be some situations where the network may have to learn without a teacher a from simple associations. Consider the development process of a child. Now, a child is born with a brain and that brain has got massive interconnection of is neural processing units but the child has to develop him or herself with a process of learning. A child observes so many new things when the world is new to the child and the child learns, finds out many things by himself or herself and that the child is able to go through some process of association. Let us say that a child observes many animals around. Now, the child observes a group of 4 legged animal called cat, a group of 4 legged animal called dog. The child may be making mistakes initially; sometimes the child could get confused between what a dog is and what a cat is. So, he/she may feel little confused but the parents are there as his or her teacher and they correct the child. So, now the child will know that dog has got some specific pattern characteristic and a cat has got some specific characteristic. When he/she observes more number of cats and dogs, it becomes possible for the child to distinguish between a cat and a dog. Now a lot of times the child learns by him/herself through associations, makes mistakes, explores lot of things on his/her own, again makes mistakes and corrects it. Therefore, there are two types of learning i.e. learning with a teacher and learning without a teacher or a sort of auto associations. The third thing about the characteristics of the neural networks is its adaptivity. The neural networks can adapt their free parameters to changes in the surrounding environment. The free parameter basically refers to what is called as the synaptic connection or connection of weights. It is seen that one has to go through the process of learning throughout the life in some sense or the other. The world that was there during our childhood is not the same world that is seen today. There are so many changes, and developments that have taken place in the scientific and technological world. The life style of people has changed all together along with the culture. There are changes everywhere and still everyone are able to cope up with this world. Now, how it is that there are so many changes in the surrounding environment that everyone is able to adapt. That is the capability which the human beings are having and the adaption is done by making some internal adjustments or the adjustments of the free parameters. Another characteristic of the neural networks is that, it not only gives the response, but it can also tell that it is the response with how much confidence level. So, it can be said that the neural network is able to give what is called as evidential response. It is seen that the human beings, a lot of times give the response with some evidentiality in it. Like it can always been said that, ‘I think it is going to happen that way’. The word I think has been associated. Now that a person is hundred percent confident but may be when he say’s ‘I think’ with a good degree of confidence it can be said that it is going to happen. So, it means that there is a confidence associated with the decisions. It is not only a decision, but it is a decision with a confidence measure. Therefore all this are the characteristic of the biological neural network systems. Another very important characteristic which the biological neural network system exhibits is that, its ability for fault tolerance. Now, if one particular nerve cell is mal functioning, one can still carry on with normal activities without any noticeable change. If too many neurons are affected, may be that will have some effect of it, but it is not something that will lead to a catastrophic failure. Unless some fault tolerance has been built in the computer system, the failure of one processing unit could very often lead to a disaster and the entire computer system or the entire network can collapse. This sort of catastrophes can happen, whereas with the biological neural networks, if some neuron malfunctions or if some connections are malfunctioning, all that it leads to is some kind of degradation in the performance which is certainly not a catastrophic failure and that degradation is what is called as graceful degradation. So, graceful degradation in the sense is that, it all depends on how much of fault has taken place.

If the fault is too many then the degree of degradation is large and if the fault is not much then the degree of degradation is small. Therefore, the biological neural network system is highly fault tolerant and it is possible to incorporate this fault tolerance mechanism even in the artificial neural networks also. If some of the capabilities of artificial neural networks are to be listed out then one of them is its VLSI implementation. Using the very large scale integrated circuit, it is possible to integrate a large number of neurons together. Naturally it cannot be thought of integrating ten billion neurons but if it is done then the human brain could have been mimicked completely. The neurons that are existing in a system forming a network can do independent computation and it is dependent upon others but there is large number of parallelism involved within it. An artificial neuron can be made adaptive if it is trained in that way. The evidential response, the fault tolerance, the VLSI implementation are all biologically and neuro biologically motivated. Considering a typical nerve cell that has a cell body, connected with number of axons that are acting like transmission lines by carrying electrical signals. In terms of the electrical characteristic, it can be said that the axons are having a high degree of electrical resistance and offers a large capacitance. All the axons ultimately end up in synaptic terminal and these are basically used for making connections with the other nerve cells which is basically nothing but the output part of the neuron. The synaptic zone is called response zone that will receive inputs from axons and gives the output to other neurons and then the dendrite zone are called receptive zone which is basically referring to the zones from which one can get the inputs from other neurons. As compared to the axons the dendrites are having more number of branches but these are much smaller in length compared to axons. The structure of neuron can be referred to as a cell that can receive inputs from other neurons and these will basically carry the signal to the cell body where processing will be done. All these inputs will be combined in accordance with the strengths of these connections. All the connections are not of the same strength, some connections are very strong and some connections are weak. Now, if the connections happen to be very strong, the signals strength also will be strong i.e. the signal that will be contributed by the input will be much more than the one with a weak synaptic link from where the signal coming will be weaker. So the strength of the synaptic connections or weights will decide what net signal will come to the cell body as the net input. This will ultimately decide what kind of response is going to be transmitted through the synaptic terminals to the other neurons. As mentioned in the previous paragraph about the free parameters it essentially refers to the strengths of the connections. Supposing, initially there is some prior connection strengths that one can take and then accordingly there will be some response that could be different from that of what is desired. If the actual response is different from that of the desired, the internal parameters of the nerve cell have to be adjusted i.e. the connection strengths are to be adjusted. So now connection strengths can be altered and then fed the same inputs and find out that what the output is going to be. This time the output may be closer to that of the desired but still may not be exactly equal to the desired. The neural network will go through another round of synaptic strength modifications and this could go on alternatively till the actual response is close to that of the desired response. It may not be able to achieve exactly, but may be able to achieve close to that. So, the synaptic strengths are basically dictating what the signal strength is going to be and these are basically acting as a free parameter to the biological nerve cell processing. Let the strength of the connections be w_1, w_2 and w_3 and supposing that there are n number of such inputs connected and let this strength be w_n and let the input signal available be that of x_1, x_2 ,x_3 and nth input be x_n. So, here what happens is that the net signal that will be available at this neuron is going to be $x_1 * w_1 + x_2 * w_2 + \ldots\ldots\ldots\ldots + x_n * w_n$, there are n number of such neurons interconnected. Therefore, this will be a linear summation but ultimately linear summation is not an important criteria but the decision whether it is yes or no or a decision which is more quantifiable is more important. In order to arrive at a decision this must be followed up by some non-linear processing unit. There must be some non-linear unit that will follow this summation and effectively the output that will be available by non-linear unit will be given to other neurons. This non linear unit is called synapse. The biological and electrical model of a neuron is shown in figures below.

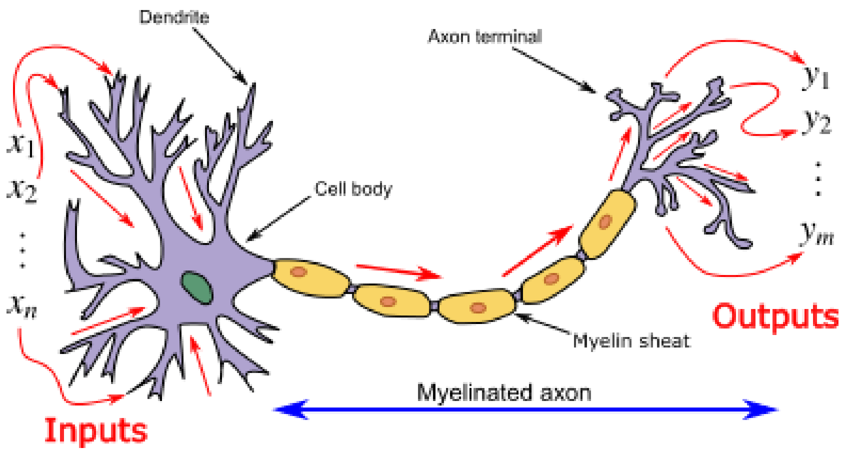

Figure 1.20: Biological model of a neuron

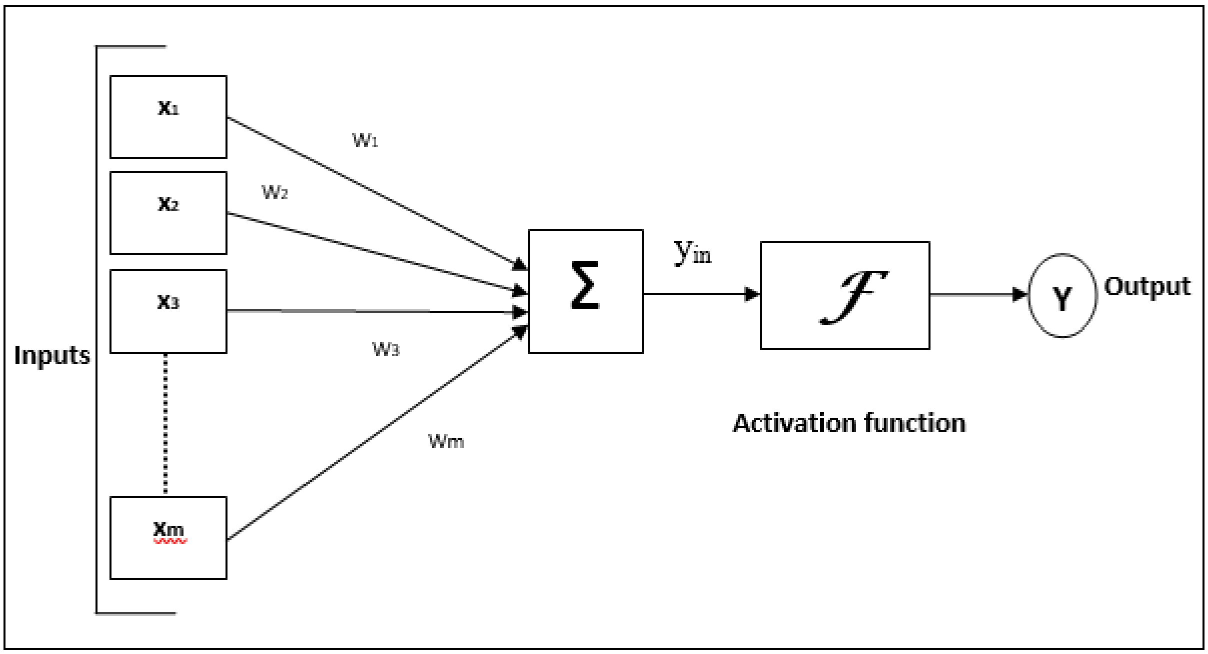

Figure 1.2.1: Electrical model of a biological neural network

The inputs represents Dendrites, summer represents cell body. The output of summer goes through axon and ends at axon terminal called as synapse which is the activation function. The output represents the response of the neuron to the inputs.

Now once the neural network has learnt something, it has to be applied somewhere so that some task becomes easy to accomplish. This can be done by artificial intelligence. Artificial intelligence is an intelligence demonstrated by machines. Artificial intelligence is used to describe the machines that mimic cognitive functions that humans associate with other humans like learning and problem solving. Artificial intelligence can be classified into three different types of system i.e. analytical, human inspired and humanize artificial intelligence. Analytical artificial intelligence has characteristics of cognitive intelligence generating cognitive representation of the world and using learning based on past experience to control future decisions. Human artificial intelligence has elements from cognitive and emotional intelligence. Humanized artificial intelligence shows all type of competencies i.e. cognitive, emotional and social intelligence. The traditional goal of artificial intelligence is reasoning, knowledge representation, planning, learning, natural language processing, perception and the ability to move and manipulate objects. The approaches for these goals include statistical methods, computational intelligence etc. Many tools used in artificial intelligence include versions of search optimization, artificial neural networks and methods based on probability, statistics and economics. Artificial intelligence was founded with the aim of precisely describing human intelligence so that a machine can simulate it. Artificial intelligence is also considered dangerous to humanity because somebody can train a robot to become a terrorist to destruct everything. So, care should be taken to avoid such dangerous situations.

Brief Description of the work done:

In this thesis, a literature survey of thirty four papers was done and out of them, the methods of best ten papers are described. These methods almost gave an accurate result of processing an image.

CHAPTER 2

LITERATURE SURVEY

In this chapter a literature survey of thirty four papers has been done with a brief description. This survey is related to image processing, artificial intelligence and neural networks.

Songcan Chen and Daoqiang Zhan proposed the method of robust image segmentation using fuzzy c means with spatial constraints based on new kernel induced distance measure. In this method clustering of pixels was done to segment the image using fuzzy c means clustering algorithim.[1] to [14]

Wilburn E. Reddick,*Member, IEEE, John O. Glass,Student Member, IEEE, Edwin N. Cook, T. David Elkin, and Russell J.Deaton, Member, IEEE proposed the method of automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. In this method ,the segmentation of brain parenchyma was done using artificial neural networks.[15] to[19]

Son Lam Phung, Member IEEE Abdesselam Bouzerdoum, Sr. Member IEEE and Douglas Chai, Sr. Member IEEE proposed the method of skin segmentation using color pixel classification. In this method three important issues were studied about color pixel classification i.e. color representation, color quantization and classification algorithm.[20] to [24]

S.S. Kumar F. Taheriand M.R. Isla Graduate Student, Professor, Respectively; Dept of Civil Engg, Dalhousie University, Halifax, NS- B3J 1Z1, Canada proposed artificial intelligence and image processing approaches in damage assessment and material evaluation. This paper proposes various various methods of artificial intelligence to detect the defects in a material.[25] to [29].

Mohamad Awad, Kacem Chehdi, and Ahmad Nasri proposed multicomponent image segmentaition using genetic algorithm and artificial neural networks. Kohonens self organizing feature map method was used to segment the image and a hybrid genetic algorithm was used to detect the main features present in the image.[30] to [34]

Ashish Lakhmani, Shruthi Shukla, Dr. Ambhuj kumar agarwal proposed some of the approaches of artificial intelligence in bio medical image processing. These approaches were used to study, diagnose, enhance , monitor and treat medical disorders [35] to [39]

Datong Chen*, Hem6 Bourlard*, Jean-Philippe Thiran Dalle Molle Institute for Perceptual Artificial Intelligence, Switzerland Signal Processing Laboratory, EPFL, Switzerland @idiap.ch, JP.Thiran@epfl.ch proposed the method of text identification using the support vector machines. The algorithm extracts the text line of the candidate on the basis of edge analysis, base line location and heuristic constraints.[39] to [43].

M. OMRAN and A.P.ENGELBRECHT Department of Computer Science, University of Pretoria, South Africmjomran@engineer.coengel@driesie.cs.up.ac.zaA. SALMAN Computer Engineering Department, Kuwait University, Kuwaitayed@eng.kuniv.edu.kw proposed the method of particle swarm optimization for image clustering.[44] to [48]

John Stoitsis, Ioannis Valavanis, Stavroula G. Mougiakakou, Spyretta Golemat,Alexandra Nikita, Konstantina S. Nikita proposed the computer diagnosis method based on medical image processing and artificial intelligence methods [49] to [53]

Jian Yang, Jing-yu Yang proposed a straight forward projection method of projecting a image vector to image matrix by using principal component analysis. Principal component analysis methods like Eigen faces and fisher faces were used for projection. [54] to 56]

El-Sayed Ahmed El-Dahshan, Tamer Hosny, Abdel-Badeeh M. Salem proposed hybrid intelligent methods for magnetic resonance brain image classification. This was done in three stages namely feature extraction, dimensionality reduction and classification [57] to [61]

MarekR.Ogiela ,RyszardTadeusiewicz proposed artificial intelligence structural imaging methods in visual pattern analysis and medical data understanding. These methods were used for the analysis of morphology of selected organs visualized on medical images.[62] to [66]

RANGACHAR KASTURI AND JUAN ALEMANY proposed the information extraction method from the images of paper based maps. This method proposes the design of a system to automatically extract information from paper based maps.[67] to [71]

D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell Computer Science Division University of California proposed the robust automatic traffic scene analysis in real time. A prototype system was proposed which successfully combined a robust, vision based traffic surveillance system.[72]

Zhengyou Zhang , Rachid Deriche, Olivier Faugeras, Quang-Tuan Luong proposed a robust method for matching two un calibrated images through the recovery of unknown epipolar geometry. This method is based on exploiting the only available geometric constraint namely epipolar constraint. The motion between the image and camera parameters is not known.[73]

Guang-Hai Liu, Zuo-Yong Li, Lei Zhang, Yong Xu proposed a method of image retrieval based on micro phone structure. The micro phone structures are based on edge orientation similarity and underlying colors in micro structures with similar edge detection.[74]

Demetri Terzopoulos, Member, IEEE, and Keith Waters, Member, IEEE proposed analysis and synthesis of facial image sequences using physical and anatomical model. This approach exploits a sophisticated generative model of human face originally developed for realistic facial animation.[75]

JOSEPH O'ROURKE AND NORMAN I. BADLER proposed model based analysis of human motion using constraint propagation. The system is structured as a feedback loop between high and low levels .Predictions are made at semantic level and verifications are made at image level.[76]

J. Jiang, P. Trundle ,J.Ren proposed medical image analysis method using artificial neural networks.[77]

X. Zhuang HIEST University, Department of Electrical Engineering and Computer Science Agiou loannou Theologou 17-23, 15773, Zografou, Athens, Greece proposed the method of edge feature extraction in digital image with ant colony system. In this paper the perceptual graph is proposed to represent the relationship between neighboring pixels.[78]

H Bischof , W Schneider and A.J Pinz proposed multi spectral classification of land sat images using neural networks. The application of three layer back propagation neural network was made for classification of land sat images.[79]

Loris Nanni, Alessandra Lumini, Sheryl Brahnam proposed local binary pattern variants as texture descriptors for medical image analysis. This paper focuses on image based machine learning methods in medical image analysis.[80]

Zhiwei Ye, Zhengbing Hu, Xudong Lai, Hongwei Chen proposed the method of image segmentation using thresholding and swarm intelligence. Two dimensional histogram and fishers methods were used to segment the images.[81]

Yong Rui and Thomas S. Huang proposed the methods of image retrieval.[82]

E.P. Baltsavias proposed the method of object extraction and and revision by image analysis using existing geo data and knowledge. Geo data of buildings and roads were used from which the important geographic contents were extracted.[83]

G. Kuntimad and H. S. Ranganath proposed the perfect image segmentation using pulse coupled neural networks. Conditions for perfect images segmentation were derived.[84]

Ridouane OULHIQ, Saad IBNTAHIR, Marouane SEBGUI, Zouhair GUENNOUN proposed the concept of artificial neural network for finger print recognition framework.[85]

Gonzalo Pajares, Member, IEEE proposed a Hopfield neural network for image change detection. This is an optimization based problem to detect any change in the image [86]

Kiran Talele, Archana Shirsat, Tejal Uplenchwar, Kushal Tuckley proposed a method of facial expression using general regression neural network. This is based on local binary pattern for feature extraction and artificial neural networks.[87]

S. Zhang and E. Salari proposed a method of image de noising using a neural network based non linear filter in wavelet domain. A three layer neural network was trained to properly differentiate between noise and noise less coefficients.[88]

Neeraj Kumar, Rahul Nallamothu, Amit Sethi porposed an image de blurring method using neural networks. This is a learning based method for image de blurring using artificial intelligence and neural networks.[89]

Shiwani Sthapak, Minal Khopade, Chetana Kashid proposed artificial neural network based signature recognition and verification method. Two types of signature verification methods were discussed one is online and other is offline.[90]

Shah Rizam M. S. B., Farah Yasmin A.R., Ahmad Ihsan M. Y., and Shazana K proposed non destructive water melon ripeness detection using image processing and artificial neural networks. The textures of watermelon were captured using digital camera and then image processing was done.[91]

Xueyun Chen, Shiming Xiang, Cheng-Lin Liu, and Chun-Hong Pan proposed a method of vehicle detection in satellite images by hybrid deep convolutional neural networks. Methods of convolutional and hybrid convolutional methods were compared in this paper.[92]

CHAPTER 3

PERFECT IMAGE SEGMENTATION USING PULSE COUPLED NEURAL NETWORK.

Introduction: This chapter deals with the method of image segmentation using pulse coupled neural network.

Method used and its result: A single layer, laterally connected pulse coupled neural network is capable of perfectly segmenting digital images even when there is an overlap in the intensity ranges of the adjacent regions. An inhibition receptive field of a neuron (i.e. the visual field in which a stimulus will modify the firing of the neuron) increases the possibility of perfect image segmentation by effectively compressing the intensity range of each region and therefore reducing the overlap of the intensity ranges.

Pulse coupled neuron model: The figure below shows the model of the pulse coupled neural network.

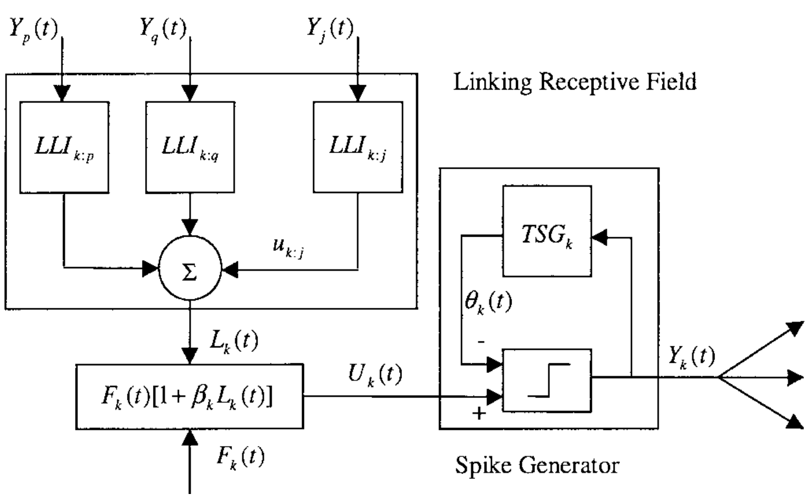

Figure 3.1: Pulse coupled Neuron model

The pulse coupled neural network model can be explained by understanding the following terminologies.

1) **Feeding Receptive field:** The feeding receptive field of the pulse coupled neural network is an external input $F_K(t)$.When the input to the pulse coupled neural network is a static image, the external inputs to the pulse coupled neural network are the intensity of the corresponding pixel.
2) **Internal activity:** The linking input $L_k(t)$ modulates the external input in a nonlinear manner given by the equation below to yield the internal activity of the neuron.

$$U_k(t) = F_K(t)[1 + \beta_k L_k(t)] \ldots\ldots\ldots\ldots\ldots(3.1)$$

where β_k is a positive constant referred as the linking coefficient of the neuron. The linking coefficient provides a simple and effective means of controlling the influence of the net linking input on the neuron's internal activity.

3) **Spike Generator:** The spike Generator consists of an impulse function generator and the threshold signal generator .The threshold signal generator was denoted as TSG_K. Every time the pulse couple neuron pulses, the threshold signal generator was charged to a predetermined value V_k irrespective of the value of the threshold signal just prior to pulsing of neuron. If two successive firings of neurons are assumed to be at time instants t_1 and t_2 the operation of threshold signal generator was described by the following equations.

$$\theta_k(t) = V_k \text{ at } t = t_1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (3.2)$$

$$= V_k e^{-\frac{(t-t_1)}{T_k}} \text{ at } t_1 < t < t_2 \ldots\ldots\ldots\ldots(3.3)$$

$$= V_k \text{ at } t = t_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(3.4)$$

Where T_k is the decay is time constant of threshold signal generator.

Image segmentation using pulse coupled neural network: A reliable and robust image segmentation method is necessary for obtaining meaningful results from any image processing system. Image segmentation is defined as the process of decomposing a given image into disjoint sub images such that the union of all the regions gives the original image. The research related to image segmentation has led to the development of numerous deterministic, statistical, neural network and knowledge-based approaches and algorithms. In spite of the extensive research, reliable image segmentation has proven to be an evasive goal. There are many factors which pose considerable challenge during image segmentation. One of them can be briefly described as the discontinuities in the pixel intensities that may be caused by illumination changes, differing surface reflectance properties, or orientations of the visible surfaces. Often there is not enough contrast between the background and object regions in the image.

Each segmentation algorithm exploits certain characteristics of the image class to accomplish segmentation. An algorithm that is suitable for segmenting a particular type of images may not be suitable for segmenting a other type of images. It may be necessary to use different methods to segment different parts of a given image. Many of the existing techniques may not be suitable for real-time or near real-time implementation. The pulse coupled neural network, by the virtue of its capture phenomenon is capable of producing good segmentation results even when the input images are noisy and of poor contrast. It has been found that even when the intensity ranges of adjacent regions overlap, the pulse coupled neural network can perfectly segment the regions if certain conditions are satisfied. It can be noted that when an image is perfectly segmented, each pixel is correctly assigned to the region it belongs to and for the images that violate the perfect segmentation conditions, it may be possible to achieve perfect or improved segmentation.

The general approach to segment images using pulse coupled neural network is to adjust the parameters of the network so that the neurons corresponding to the pixels of a given region pulse together and the neurons corresponding to the pixels of adjacent regions do not pulse together. It has to be assumed that the image to be segmented consists of regions and is applied as an input to a pulse coupled neural network. The network neurons pulse based on their feeding and linking inputs. The feeding input to a neuron is equivalent to the intensity of its corresponding pixel. Due to the capture phenomenon the neurons associated with each group of spatially connected pixels with similar intensities tend to pulse together. Thus, each set of synchronously pulsing neurons identifies a segment of the image. A segment identified by the pulse coupled neural network may be a region or part of a region or union of several regions and sub regions of the image. Ideally, the goal is to choose the network parameters such that each segment exactly corresponds to a complete region in the image. If such parameters exist, then perfect segmentation is possible.

The pulse coupled neural network used for image processing applications is a single layer two-dimensional array of laterally connected pulse coupled neurons. The number of neurons in the network is equal to the number of pixels in the input image. There exists a one-to-one mapping between the image pixels and network neurons. The key features of segmenting an image with the help of pulsed coupled neural network can be explained as follows.

Let the external input to the neuron be $X_{i,j}$ which is the intensity value of the pixel $P_{i,j}$ .Therefore $F_k(t)$ in equation 3.1 is equal to $X_{i,j}$.

1) All the leaky integrators are identical to one another. Leaky integrators are the type of integrators that takes integral of an input but gradually leaks a small amount of input over a given time. The amplification factor and the decay time constant of every linking leaky integrator in the network are V_l and T_l respectively .
2) Each neuron receives a linking input from its neighbors that is within a certain distance from each other. The combination of all the receptive fields after they are passed through leaky integrators forms linking input. Every neuron corresponds to one pixel. Let $N_{i,j}$ and $N_{p,q}$ for all i , j and p, q be the two neurons and $P_{i,j}$ and $P_{p,q}$ be the two pixel values corresponding to the two neurons. The distance between the two neurons $N_{i,j}$ and $N_{p,q}$ is defined as the Euclidean distance between the pixels $P_{i,j}$ and $P_{p,q}$. The value of the linking coefficient is independent of the neuron. The threshold signal generators of all the neurons are identical to one another.
3) Two terms are defined for better analysis of the pulse coupled neural network. They are natural period and capture range.

   An unlinked neuron with constant feeding input C was considered. A neuron that pulses due to the influence of feeding input alone is said to be pulsing naturally. A naturally pulsing neuron pulses periodically with a time period of T(C) which was computed as $\tau_\theta \ln\left(\frac{V_\theta}{C}\right)$ where τ_θ and V_θ are the decay time constants and the reset value of the threshold signal generator respectively. Now let us consider two mutually linked neuron $N_{i,j}$ and $N_{p,q}$. Assume that $X_{i,j}$ is greater than $X_{p,q}$ and V_θ is greater than maximum possible value of the internal activity of either of the two neurons. At t=0 the neurons $N_{i,j}$ and $N_{p,q}$ are reset and the threshold value is forced to set at zero. Thus the feeding input is applied causing both the neurons to pulse together at t=0. During the first pulsing event the neuron $N_{i,j}$ captures the neuron $N_{p,q}$ at time t= T $(X_{i,j})$ and transmits it. The condition for the neuron $N_{i,j}$ to transmit the neuron $N_{p,q}$ can be given below as

$$X_{p,q}(1 + \beta L_{p,q}(t)) \geq X_{i,j} \ldots\ldots\ldots\ldots(3.3)$$

The intensity range $\left(\frac{X_{i,j}}{1+\beta L_{p,q}(t)}, X_{i,j}\right)$is referred to as the capture range of $N_{p,q}$ with respect to $N_{i,j}$. If $X_{p,q}$ lies within the capture range then $N_{p,q}$ will be captured by $N_{i,j}$.This concept of capture range can be extended to a group of mutually connected neurons also.

Assume that a set of neurons pulse naturally at time t_1.Each neuron which receives a linking input that is high enough to increase the neuron's internal activity above the threshold value is captured by the pulsing neurons. Therefore, the capture range of $N_{i,j}$ with respect to the group of pulsing neurons is given as $\left(\frac{\theta_{i,j}(t_1)}{1+\beta L_{i,j}(t_1)}, \theta_{i,j}(t_1)\right)$ where $L_{i,j}(t_1)$ is the total linking input from all the neurons pulsing at time t_1.If $X_{i,j}$ lies within the capture range then $N_{i,j}$ pulses at t_1 and sends linking inputs to all the neurons that are linked to $N_{i,j}$. The pulsing of $N_{i,j}$ and the transmission of linking inputs to other neurons are assumed to be instantaneous. Since the linking inputs decay fast enough, the group of neurons are not be able to capture $N_{i,j}$ at a later time if it fails to capture it at time t_1 due to current pulsing. A constraint that is to be kept in mind while applying pulse coupled neural network for image segmentation is the proper selection of V_θ.

Selection of V_θ : Let X_{min} and X_{max} be the minimum and maximum values of feeding inputs to the network. All the neurons in the network pulse at t=0. Later the neurons for which the feeding input is equal to X_{max} pulse naturally at t=T(X_{max}).Those neurons whose feeding inputs lie within their capture ranges are also forced to pulse at t= T(X_{max}). This process of natural pulsing and capture continues indefinitely. By the time t = T(X_{min}) all the neurons in the network pulse at least once. By the proper selection of value of V_θ, each neuron can be restricted to pulse exactly once during the interval [0,

$T(X_{min}]$. If the network is reset at $t = T(X_{min})$ the network will repeat the pulsing pattern of the interval $[0, T(X_{min})]$ in the interval $[T(X_{min}), 2T(X_{min})]$.Thus if the network periodically resets at times $t = nT(X_{min})$ where n is the positive integer the network will produce periodic output. The time between two successive resetting of the network is referred to as a pulsing cycle. The determination of minimum value of V_θ that guarantees the pulsing of neurons exactly once during a pulsing cycle is explained below.

Consider the sequence of events that takes place during the first pulsing cycle.

1) All the neurons pulse at $t = 0$. The outputs of all the threshold signal generators are charged to V_θ and start to decay exponentially.
2) The neurons corresponding to the highest intensity (X_{max}) pixels pulse naturally at $t = T(X_{max})$ and capture all neurons whose feeding inputs are within their respective capture ranges.
3) The neurons corresponding to the lowest intensity (X_{min}) pixels which have not pulsed at an earlier time due to the capture phenomenon pulse naturally at $T(X_{min})$

   At $t = T(X_{min})$, the value of threshold signal of a neuron $N_{i,j}$ which pulsed at $t = T(X_{max})$ is given as $\theta_{i,j}(T(X_{min})) = V_\theta e^{-\frac{(T(X_{min}-T(X_{max}))}{\tau_\theta}}$. Let $U_{i,j}(T(X_{min}))$ be the maximum value of $\theta_{i,j}(T(X_{min}))$. If the value of V_θ is selected such that $\theta_{i,j}(T(X_{min}))$ is greater than the maximum value $U_{i,j}(T(X_{min}))$ then each neuron is guaranteed to pulse exactly once during a pulsing cycle. Therefore,

   $V_\theta e^{-\frac{(T(X_{min}-T(X_{max}))}{\tau_\theta}} > U_{i,j}(T(X_{min}))$. Since $T(C) = \tau_\theta \ln\left(\frac{V_\theta}{C}\right)$ the inequality reduces to $V_\theta\left(\frac{X_{min}}{X_{max}}\right) > X_{max}(1 + \beta L_{i,j}(T(X_{min}))$. The value of the linking input $L_{i,j}(T(X_{min}))$ depends on the linking receptive field radius r and the number of neurons pulsing at $T(X_{min})$ that are linked to $N_{i,j}$ .For a given value of radius r the linking input attains a maximum value L_{max} when all the neurons that are linked to $N_{i,j}$ pulse at $T(X_{min})$.

   Conditions for perfect image segmentation: Consider an image consisting of two regions i.e. object and background. Let the object pixels be denoted as R and background pixels be denoted as B. Let (X_{Rmin}, X_{Rmax}) and (X_{Bmin}, X_{Bmax}) be the intensity ranges of the object and the background pixels respectively. It was assumed that $X_{Rmax} > X_{Bmax}$ and $X_{Bmax} > X_{Rmin}$. Since the intensity ranges of R and B overlap, thresholding techniques cannot produce perfect segmentation. Therefore the above image is applied as input to the pulse coupled neural network and the following events have occurred in the given order during each pulsing cycle.

1) The object neurons with the intensity X_{Rmax} pulse naturally at t = $T(X_{Rmax})$.

2) All object neurons that lie within their respective capture ranges were captured. For example $N_{i,j}$ is captured by the pulsing neurons if the following inequality is true:

$$X_{i,j}\left(1+\beta L_{i,j}\left(T(X_{Rmax})\right)\right) \geq X_{Rmax} \quad \text{……………. (3.4)}$$

3) Every background neuron $N_{p,q}$ for which the following inequality is not true was also captured by the pulsing neurons:

$$X_{p,q}\left(1+\beta L_{p,q}\left(T(X_{Rmax})\right)\right) < X_{Rmax} \quad \text{…………………(3.5)}$$

4) The object neurons that are not captured at t = $T(X_{Rmax})$ pulse in several groups after $T(X_{Rmax})$ and before $T(X_{Bmin})$.The number of groups and the exact time at which each group pulses were determined by the intensity distribution of the image, β and other network parameters.

5) At t = $T(X_{Bmax})$ the neurons corresponding to the background pixels with the intensity X_{Bmax} that were not captured by object neurons so far pulse naturally.

6) Every neuron $N_{m,n}$ for which the following inequality is true, was captured by the pulsing neurons:

$$X_{m,n}\left(1+\beta L_{m,n}\left(T(X_{Bmax})\right)\right) \geq X_{Bmax} \quad \text{…………………..(3.6)}$$

7) The remaining background neurons organize themselves into several groups and pulse after t = $T(X_{Bmax})$ and before t = $T(X_{Bmin})$.

8) If (3.3) is true for every object neuron and (3.5) and (3.6) are true for every background neuron, all the object neurons pulse together at t = $T(X_{Rmax})$ and all the background neurons pulse together at t = $T(X_{Bmax})$ thus segmenting the image perfectly.

Conclusions: The pulse couple neural networks have proven to be a good processing element for image processing applications. Perfect segmentation conditions were established which were not established for other methods also.

CHAPTER 4

NEURAL NETWORK BASED IMAGE DE BLURRING, FINGER PRINT RECOGNITION AND SIGNATURE RECOGNITION.

Introduction: This chapter explains the method of image de blurring, finger print recognition and signature recognition using back propagation neural network.

Method used for image de blurring and its results: Image de blurring is a particular type of image degradation which makes an image unfocussed as it has adverse effect on the high frequency components in the image. Blurring may be caused due to many reasons like blur due to movement of subject i.e. motion blur, improper holding of camera, dirty lens, wrong focusing etc. Many de blurring algorithms were applied but these algorithms required prior knowledge of the blur kernel which was difficult to obtain. Therefore a method for de blurring the image using neural networks was proposed that assumed no prior knowledge of the blur kernel. The original image was modeled as a Markov random field and the blurred image as degraded version of Markov field.

Markov Random field: Let $\{x_1, x_2, x_3 \dots \dots x_n\}$ be the set of n random variables. A Markov Random field is a probability distribution over the variables x_1 to x_n defined by an undirected graph in which the nodes correspond to variables x_n. The probability p has the form

$$P(x_1, x_2, x_3 \dots \dots . x_n) = \frac{1}{Y} \prod_{c \in C} \emptyset_c (x_c) \dots\dots\dots\dots\dots (4.1)$$

Where C denotes the set of cliques i.e. fully connected sub graphs of the original graph, $\emptyset_c$ is a nonnegative function over the variables in a clique, Y is a partition function which is a normalizing constant that ensures that the distribution sums to 1.

Blur Modeling: A statistical model was developed to find out the relationship between the original image and the blurred image. Let I be the original image, J be the blurred image, H be the blur kernel. The blurring process was represented in terms of convolution of I with H as

$$J = I*H+n \dots\dots\dots\dots (4.2)$$

n is the noise that was added during the blurring process. No assumptions were made about the nature of blur kernel. This statistical relationship between the blurred and original image was modeled using the Markov random field. A random change was made in the image based upon the current value of the pixel and its immediate neighborhood which makes the image easy to analyze using the markov random field. This was done by characterizing the mutual influence among the pixel intensities for a given neighborhood using the conditional probabilities. A 3x3 neighborhood for a pixel located at any point (i , j) was considered as shown in figure below.

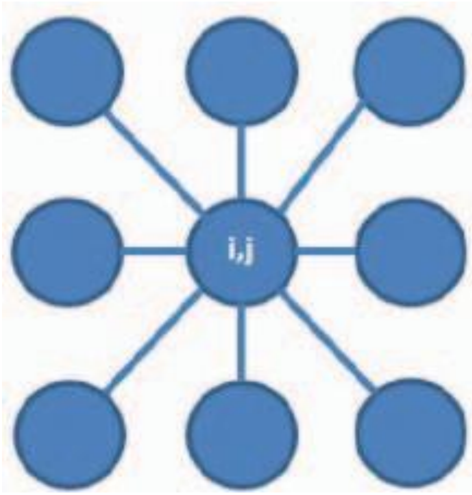

Figure 4.1: Markov random field relationship between pixel at location i, j with its neighborhood.

Let Z be a random field over a given NxN original image. From the Hammersly-Clifford theorem, P(Z=z) = $\frac{1}{Z_p} e^{-U(z,\theta)}$ where z is the realization of Z, Z_p is a partition function given as $\sum_z e^{-U(z,\theta)}$, θ is the parameter that defines the Markov random field model i.e. it gives the information about the interaction among the pixels, U(Z,θ) is the energy function given by $\sum_{c\in C} V_c\,(Z,\theta)$ which is the potential function associated with the clique c, and C is a set of all the cliques. A clique is any fully connected subset of a graph. In order to differentiate between original image and blurred image using markov random field, a parameter called Gibbs energy function was defined separately for original image and blurred image. For a 3×3 neighborhood of a pixel in original image at location (i, j), the Gibbs energy function was given as follows

$$U(Z,\theta) = \sum\sum a[(z_{ij} - z_{ij+1})^2 + (z_{ij} - z_{ij-1})^2] + b\left[(z_{ij} - z_{i-1j})^2 + (z_{ij} - z_{i+1j})^2\right] + c\left[(z_{ij} - z_{i-1j+1})^2 + (z_{ij} - z_{i+1j-1})^2\right] + d[(z_{ij} - z_{i-1j-1})^2 + (z_{ij} - z_{i+1j+1})^2] \text{........(4.3)}$$

The parameter set θ consists of unknown parameter a,b,c,d which was estimated by learning the markov random field model depending upon the choice of clique potentials. Blurred image corresponding to given original image was modeled using Gibbs energy function but with different parameter set as for the blurred image we have a degraded markov random field. For 3×3 neighborhood of blurred pixel at location (i, j), Gibbs energy function was represented as shown in equation below

$$U(Z,\theta) = \sum\sum a_{ij}[(z_{ij} - z_{ij+1})^2 + (z_{ij} - z_{ij-1})^2] + b_{ij}\left[(z_{ij} - z_{i-1j})^2 + (z_{ij} - z_{i+1j})^2\right] + c_{ij}\left[(z_{ij} - z_{i-1j+1})^2 + (z_{ij} - z_{i+1j-1})^2\right] + d_{ij}[(z_{ij} - z_{i-1j-1})^2 + (z_{ij} - z_{i+1j+1})^2] \ldots\ldots(4.4)$$

A functional mapping was established between the markov random field of original image and the degraded markov field of the blurred image by using a neural network. A 3x3 patch was picked up from the blurred image and given to the input and the center pixel of the original patch was given to the output. Training samples were generated by shifting the center pixel of a 3x3 patch by one pixel in either direction for both blurred patch and the original patch. Then the desired mapping was captured by using a small number of sigmoid basis functions. This mapping gives the information about the de blurred image. This experiment was performed on a three layer neural network using the back propagation algorithm for weight adaption. This neural network had one output node with an activation function and twelve hidden nodes with sigmoid activation functions and input node dependent on size of neighborhood. This experiment used 3x3 patches and hence nine input node. Therefore a computationally tedious task of de blurring an image was overcome by using a neural network that showed better de blurring capability

Method used for finger print recognition and its results: Finger print is one of the most used biometrics for identification of a person that rely on image processing and classification algorithm. A framework that enables finger print detection using the set of image preprocessing algorithm and an artificial neural network for feature extraction and classification has been proposed. An Artificial neural network was configured and trained to adapt to the variations of fingerprints images of one's person especially when fingerprint images were affected with noise. The extraction of true details was done if it was a bad quality input image and then the image was divided into sixty four zones and modeled with a vector and was given as a input to the neural network.

The framework used for finger print recognition contains three stages. They are (A) preprocessing, (B) feature extraction (C) classification using the artificial neural network

(A) Preprocessing: Noise and poor quality image are the main causes of the finger print identification error. To improve the finger print detection results preprocessing of image was used. It is a 5 step process namely normalization, enhancement, binarization, filtering and thinning.

1) Normalization: The purpose of normalization was to bring the gray level values in the finger print image within the desired range of values. The normalized image was calculated using the equation given below.

$$N(x, y) = M_o + \sqrt{\frac{VAR_0 * (I(x,y) - M)^2}{VAR}} \text{ if } I(x, y) > 0$$

$$\ldots\ldots(4.5)$$

$$= M_o - \sqrt{\frac{VAR_0 * (I(x,y) - M)^2}{VAR}} \text{ , otherwise}$$

Where N(x, y) is the normalized version of the original image I(x, y). M_0 and VAR_0 are the desired mean and variance values respectively. While M and VAR are the real mean and variance values of the original image respectively.

2) Fingerprint enhancement by Fourier transform: The input image was divided into blocks of mxn pixels and then the Fourier transform was applied to each block according to the equation shown below.

$$f(u, v) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} F(x, y) e^{-j2\pi\left(\frac{ux}{m} + \frac{vy}{n}\right)} \ldots\ldots..(4.6)$$

To obtain the enhanced block the Fourier transform of the block was multiplied by its own magnitude certain number of times according to the equation given below

$$g(u, v) = F^{-1}(F(x, y) * F(x, y)^k) \ldots\ldots\ldots..(4.7)$$

The higher the value of k in the formula, the more the image improves the appearance of edges and filling up of small holes in the edges. However having too high value of k will result in the false joining of edges. The inverse furrier transform was computed using the equation below

$$F(x, y) = \frac{1}{mn}\sum_{u=0}^{m-1}\sum_{v=0}^{n-1} f(u, v)e^{j2\pi\left(\frac{ux}{m}+\frac{vy}{n}\right)} \text{……….(4.8)}$$

The advantage of Fourier transform was to connect some falsely broken points on edges and to decrease the background noise like false connection between edges.

3) **Binarization:** The gray nature of the finger print image is more related to the image quality and doesn't have considerable effect on person's identification. The 256 level of gray was transformed into black and white with no loss of information. For this purpose a global threshold was defined for the gray image and a comparison between the local gray value of each pixel and the threshold was made. The pixel value turns to 0 (Black) if it is less than the threshold and to 1 (white) if it is greater than threshold. By the end of this process all the pixel values within the image were either 0 or 1 and the image was converted to binary format. The value of 1 corresponds to the information in the image and the value of 0 corresponds to the background.

$$B(x, y) = 1 \text{ if } F(x, y) < T$$

$$\text{……….(4.9)}$$

$$= 0 \text{ otherwise}$$

Where T is the threshold gray value of F(x, y)

4) **Gabor Filter:** The Gabor filter was used to remove noise and to preserve the true edges. The Gabor filter is described by the following equation.

$$g = ke^{-\frac{d}{2}M(x,y,f,\theta)} \text{ …… (4.10)}$$

$$\text{where } k = \frac{1}{2\pi s_x s_y}, \text{………….. (4.11)}$$

$$d = \left(\frac{x}{s_x}\right)^2 + \left(\frac{y}{s_y}\right)^2 \text{ ……… (4.12),}$$

$$M(x, y, f, \theta) = \cos\left(2\pi f(x\cos\theta + y\sin\theta)\right) \text{…….(4.13)}$$

Where s_x and s_y are variance along x and y direction, f is the frequency of sinusoid and θ is the orientation of the gabor filter.

5) **Thinning:** The objective of thinning was to eliminate the redundant pixel in the edges

(B) Feature extraction: Feature extraction involves 2 steps one is minutiae extraction and profile building.

Minutiae Extraction: For the minutiae extraction, cross number method was used which generally includes the bifurcation. The bifurcations and edge endings were extracted from the thinned image by scanning the local neighborhood of each edge pixel in the image using a 3×3 window. The value of the cross number C_n was calculated using equation.

$$C_n = \frac{1}{2}\sum_{1}^{8}(p_{i-1} - p_i) \qquad \text{.................(4.14)}$$

If $C_n = 1$, p is the end point and if $C_n = 3$ point p is a bifurcation.

Profile Building: To build the profile corresponding to the image, the fingerprint image was divided to a grid of sixty four zones and then the number of the bifurcations in each zone was calculated which gives the system a basic fault tolerance to image rotation and translation. Fault tolerance is the property of the system to continue to work properly even if there is failure in the components of the system. Each neuron of the input layer corresponds to number of bifurcations in the equivalent zone.

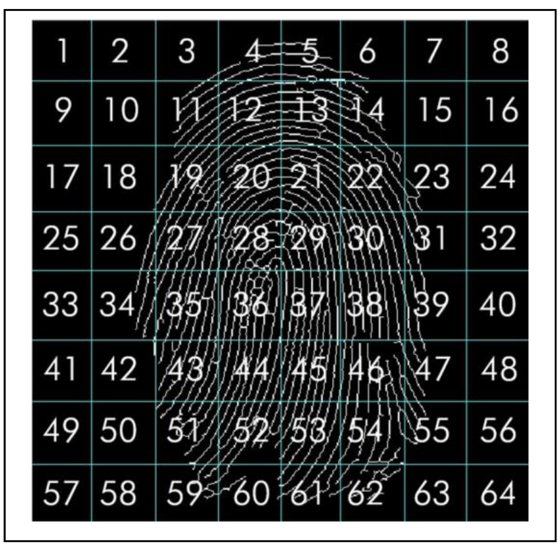

Figure 4.2: Profile building

(c) Classification: The objective of this step was to train an artificial neural network to perform classification of different fingerprint images as belonging to their owners. When the learning was achieved, the algorithm was capable for identification. The feed forward back propagation neural network was used because it is one of the most popular neural networks. The term feed forward describes how this neural network works. In a feed forward neural network, neurons are only connected forward. Each neuron of a layer was connected to all the other neurons of the next layer, but there are no connections backward. The second term back propagation describes how this type of neural networks is trained. An iterative weight-adjusting scheme is used to propagate backward the error term by modifying the weights for all the connections in the neural network structure.

$(U_1, U_2, U_3 \ldots \ldots \ldots U_{I+1})$ are the input vectors, $U_{I+1} = 1$ was the bias term,

$(S_1, S_2, S_3, \ldots \ldots S_K)$ is the vector of output for classification.

I: number of units in the input layer

J: number of units in the hidden layer

K: number of units in the output layer.

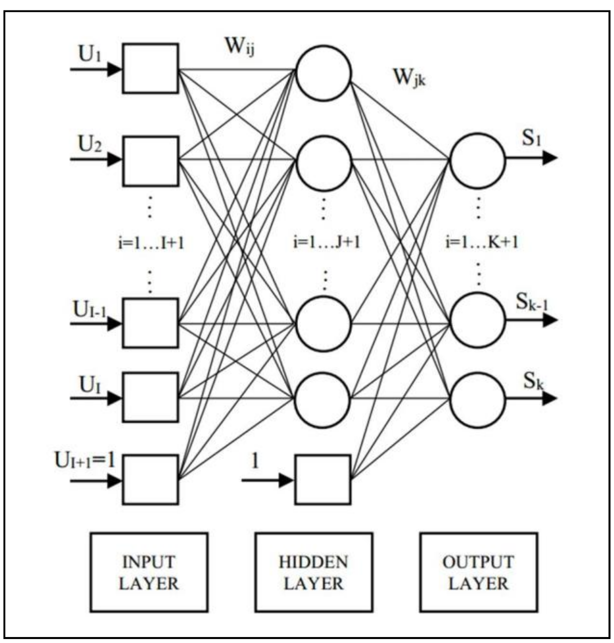

Figure 4.3: Artificial Neural Network for performing classification

Method used for signature recognition and its results:

Human signature has proven to be the most important for access and is socially accepted and extensively used means for authentication. This method of signature verification system is an offline system of signature verification which extracts features that characterizes each input signature. The approach starts by scanning the signature image using peripheral devices and then modifying their quality through image enhancement followed by feature extraction and neural network and finally it verified whether the signature is genuine or morphed. It involves two steps which is preprocessing and feature extraction.

Step1: Preprocessing

The signature image given at the input was scanned in gray code with a normal scanner using sufficient resolution using equation given below.

$$\text{Gray colour} = (0.299 * \text{Red}) + (0.5876 * \text{Green}) + (0.114 * \text{Blue})$$

The steps involved in preprocessing stage were 1) Scaling (2) Noise Reduction (3) Background elimination (4) Signature Normalization (5) Thinning.

1) Scaling:

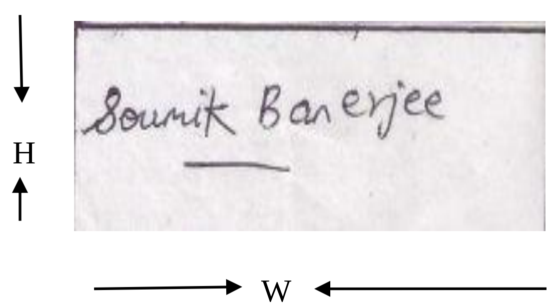

Figure 4.4: Signature Image Recognition

Image is transformed to uniform 100x100 pixels image by using equations

$$X_{new} = \frac{Xold*100}{H} \text{ where}$$

$$X_{old} = \text{Original X Coordinate}$$

$$X_{new} = \text{New X cordinate}$$

$$Y_{new} = \frac{Y_{old} * 100}{W} \quad \text{, where}$$

$$Y_{old} = \text{Original Y Co ordinate,}$$

$$Y_{new} = \text{New Y Co ordinate}$$

2) Noise Reduction: Median Filter was extensively used for smoothing and restoring images corrupted by noise. Usually mean filter is used but in this case median filter is used as it has the ability to suppress impulse noise while preserving edges.

3) Background Elimination: Thresholding is used to separate the signature from the background. A case of dark object on light background has been considered and hence a Threshold value T was entitled as brightness threshold and was suitably chosen and applied to the image .After thresholding the image

Pixels of signature = Binary 1,

Other pixels = Binary 0.

Mathematically, it F(x,y) is the image and

if $f(x,y) > T$ - it represents background.

if $f(x,y) < T$ – it represents object.

(4) **Signature Normalization:**

Height and width of signature fluctuate from person to person and occasionally even the same person may exercise different sizes of signature. So it becomes necessary to normalize the signature i.e. all the signatures will have to have similar dimensions. Normalization was done using following equations.

$$\text{Xnew} = \left(\frac{\text{Xold} - \text{Xmin}}{\text{Xmax} - \text{Xmin}}\right) * \text{M}$$

$$\text{Ynew} = \left(\frac{\text{Yold} - \text{Ymin}}{\text{Ymax} - \text{Ymin}}\right) * M$$

Where X_{new}, Y_{new} are Coordinates of normalized signature, X_{old}, Y_{old} are Coordinates of Original Signature.

M – width/height meant for Normalized Signature.

5) **Thinning:**

Thinning process was used to eliminate thickness differences of pen. It was used to reduce original image into most compact representation.

Step 2: Feature Extraction

A feature vector is used to uniquely characterize a candidate signature. These features are extracted as follows.

Global features: Global features give information about the shape like signature area, signature height to width ratio, slope and slope direction skewness of signature etc.

Slope of the signature is calculated by the use of thinned image obtained during preprocessing. A 3x3 sliding window is used for slope calculation.

Width to Height ratio = $\frac{\text{Xmax}-\text{Xmin}}{\text{Ymax}-\text{Ymin}}$

Skewness is the measure of symmetry .It allows us to determine how curved is each segment of the signature.

Summary:

Input = Signature image

Output = Conformation from the system whether the signature is genuine or false.

1) The signature image whose authenticity has to be found out was taken from the database..
2) The signature image was enhanced by preprocessing.
3) Various features of the signature image were extracted.
4) The feature vector of the signature image was created by combining the extracted features from the preprocessed image.
5) The feature vector of the signature image was normalized for further processing.
6) This normalized feature vector was applied to the neural network for training purpose.

7) Steps from 1 to 6 were repeated to train neural network to test signature.
8) The pattern matching with the test dataset present in the hidden layer of neural network was performed.
9) The Classification of the datasets was done.
10) The authenticity of the signature i.e. whether the signature is genuine or morphed was found out by the outcome produced by the output layer of the neural network.

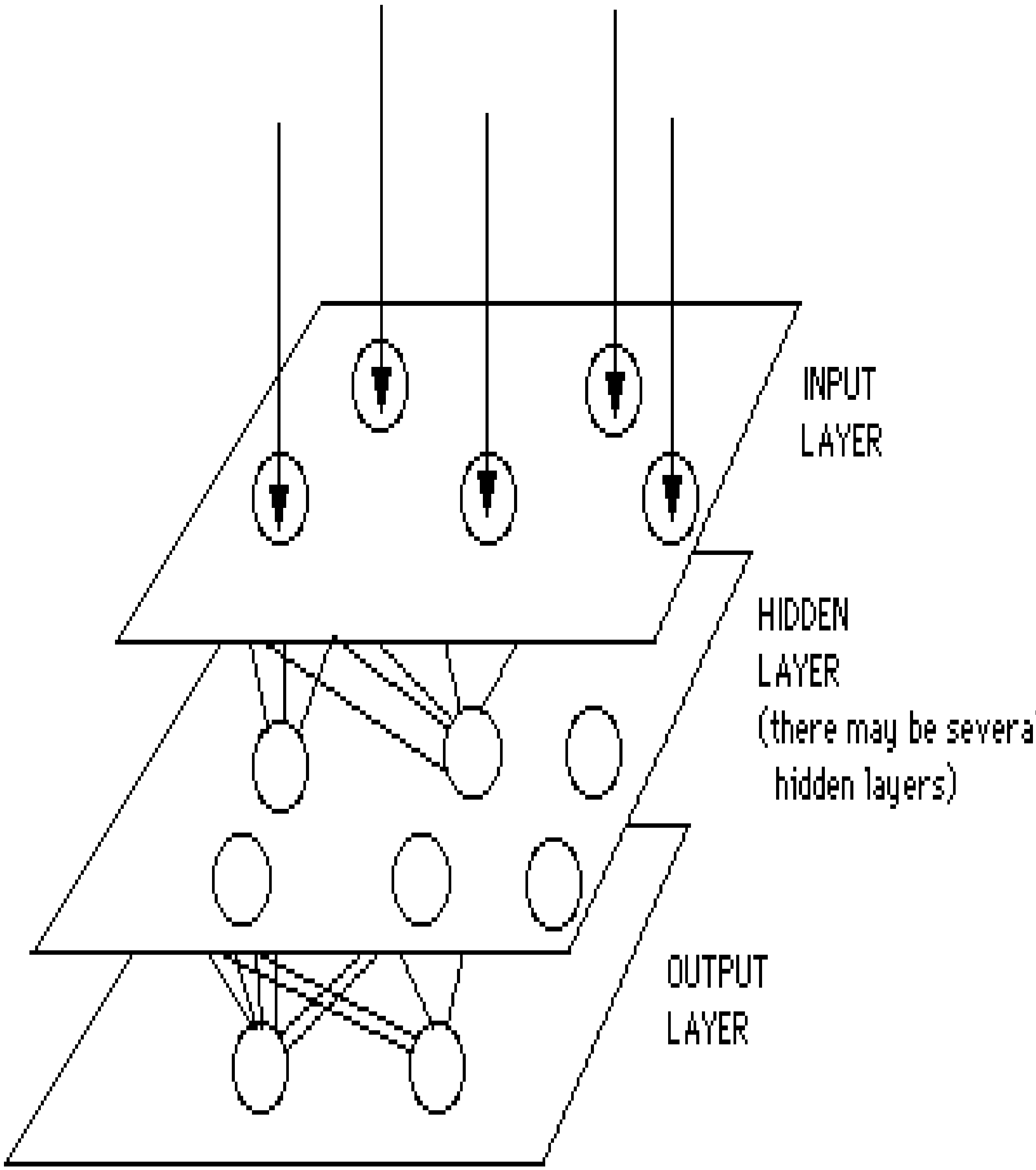

Figure 4.5: Three layer neural network for image segmentation

Conclusion: 19 different signatures which are both genuine and morphed were taken of 30 individuals and kept in the database. After relating a feature vector of test signature if the output neuron generates a value close to +1 test signature was declared as genuine and if it generates value close to -1 it was declared as morphed.

In this case, 3 Constraints were defined for signature recognition.

False acceptance range: $\frac{\textbf{Number of morphed signatures accepted}}{\textbf{Number of morphed signatures tested}} * \mathbf{100}$

False rejection range: $\frac{\textbf{Number of original signature rejected}}{\textbf{Number of original signature tested}} * \mathbf{100}$

Correct classification range: $\frac{\textbf{Number of samples correctly recognized}}{\textbf{No of samples tested.}}$***100**

Therefore this method can be a reliable method for signature verification as the constraints can be classified into false acceptance range, false rejection range and correct classification range.

In the image de blurring method the dependency of pixel intensity on its neighborhood exists even for blurred image .In order to de blur the image a mapping was established between original Markov random field and degraded Markov random field. Learning the Markov random field is a tedious task and this was done easily with the help of neural networks **.**This is a good method for de blurring the image.

In the finger print recognition method the   number of hidden layers was chosen to be 1 and any number between 2 and 120, so that the performance of the neural network could be evaluated according to the number of units in the hidden layer. Each neuron of the neural network has an activation function which specifies the output of a neuron to a given input. Different activation functions such as hyperbolic sigmoid activation function, Elliot activation function and logistic sigmoid activation function were used for testing. These activation functions gave 81% accuracy of identifying correct finger print of a person. Therefore this method can be a reliable method but had the accuracy been more than 90% it would have been more reliable.

CHAPTER 5

VEHICLE DETECTION IN SATELLITE IMAGES BY HYBRID DEEP CONVOLUTIONAL NEURAL NETWORKS

Introduction: This chapter explains the method of detecting satellite image by hybrid deep convolutional neural networks.

Method used and results: Detecting small objects such as vehicles in satellite images is a difficult problem. Many features have been used to improve the performance of object detection but mostly in simple environment such as those on roads. No satisfactory accuracy has been achieved in complex environments .To solve this problem a deep convolutional neural network was used that can learn rich features from the training data automatically and has achieved a good performance in many image classification databases. Though the deep neural network has shown vigorousness to distortion, it only extracts the features of the same scale and hence is insufficient to tolerate large-scale variance of object. Therefore, a Hybrid deep convolutional neural network was introduced by dividing the maps of the last convolutional layer and the max-pooling layer of deep neural networks into multiple blocks of variable receptive field sizes or max-pooling field sizes to enable the hybrid deep neural network to extract variable-scale features. Max pooling is a sample based discretization process whose objective is to down sample an image by reducing its dimensionality. The deep convolutional neural network as a feature learning architecture has yielded superior performance in many object recognition tasks. Deep neural networks use convolutional layers and max pooling layers. A convolution layer is the core building block of convolutional neural network which consists of a set of learnable filters which is used to perform convolution between the input data and the learning filters. The extracted features are combined by the hidden layer and output layer for the purpose of classification. The layers of the deep neural networks are divided into two parts one is the feature extractor which extracts features systematically using convolutional layers and max pooling layers and other is the multi layer perceptron classifier which classifies the data by extracted features. A multilayer perceptron is a class of feed forward neural network which consists of three layers of node i.e. input layer, hidden layer, output layer. It uses a supervised learning algorithm called back propagation algorithm. Each convolutional layer of deep neural network generates feature map on a local receptive field in the maps of the preceding layer. Feature map is the output received by convolving the image with a particular filter.

Structure of Deep Neural Network: The layers of deep neural network are divided into two parts one is the feature extractor and the other is the multi layer perceptron classifier. For convenience all the convolutional layers were assumed to have same number of maps. All the convolutional and max pooling layers consist of feature extractor of deep neural network. A max pooling layer is a layer in convolutional neural network which progressively reduces the dimensions of the image to reduce the amount of parameters and computation of the network. The last layer of max pooling layer outputs the extracted features to the multi layer perceptron classifier. The output value of the multi layer perceptron denotes the brightness of the pixels of the output image which signifies the likelihood of vehicle detection. The Tanh function was used as the kernel function for all the nodes of the deep neural network. The convolutional layer map was determined by the filters sliding on the previous layers pixel by pixel. The max-pooling layer maps were determined on the non overlapped max-pooling fields sliding over the previous convolutional layer.

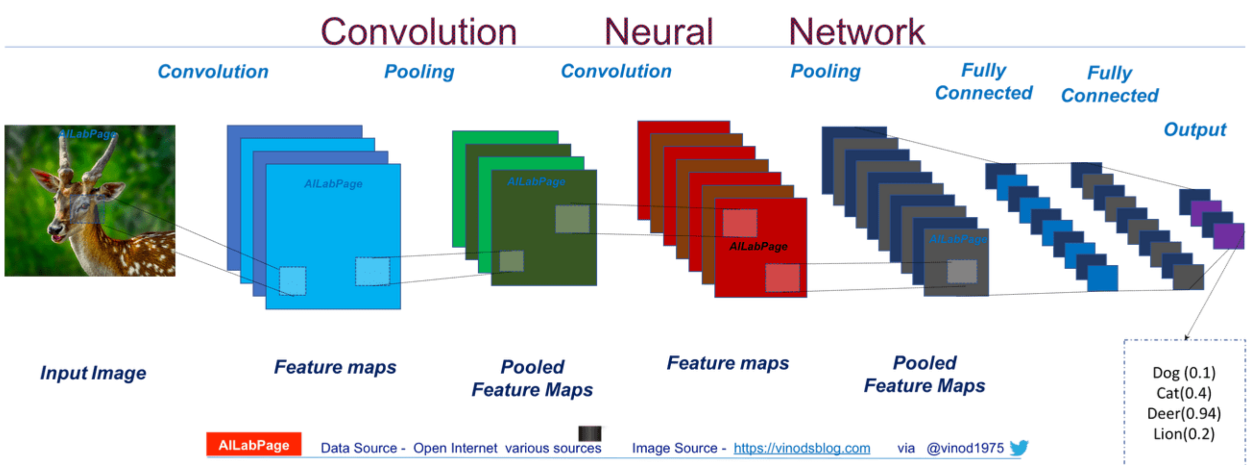

Figure 5.1: Deep Convolutional Neural Network

Two definitions were explained for understanding the deep neural networks

1) **Source area**: The source area of a pixel is a part that includes all the relevant pixels of the input image.
2) **Feature scale**: Feature scale of a pixel is the largest size of the possible source areas. In deep neural network all the pixels of a particular max pooling layer was passed through 3 different filters present in 3 convolutional layers and it was seen that the deep neural network extract features of the same scale.

Structure of Hybrid deep neural network: A set of convolutional layers and max pooling layers were divided into different blocks with each block having a different size. The pixels of a particular pooling layer in the blocks were passed through 3 different filters in convolutional layer and it was found that the hybrid deep neural network extracts features of 2 scales.

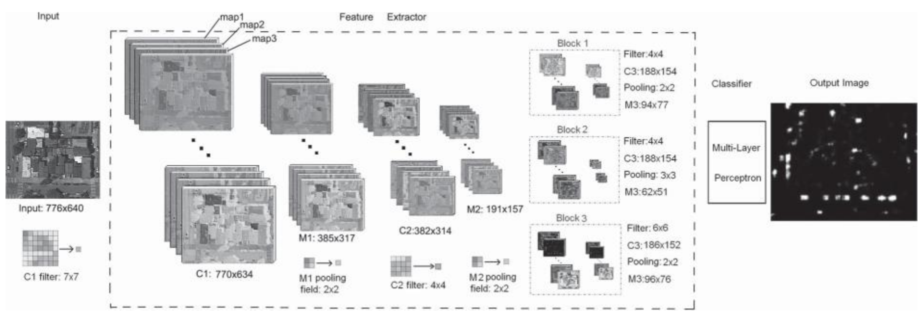

Figure 5.2: Hybrid Deep Convolutional neural network

Implementation on location of vehicle: Image gradient is a directional change in the intensity of the image. It is one of the fundamental building blocks in image processing. It is used for gradual blend of colors also called color progression. In case of vehicle detection, the edge detection was the necessary condition and a gradient computing norm was presented which was based on RGB channels. For vehicles of black color and the vehicles sheltered with tree, the gradient method was able to generate dim edges of black car and failed to separate the sheltered car from the textures of the tree. Therefore, a multiple gradient method was used where borders of dim black cars were enhanced and tree textures were erased. A sliding window technique was also important for vehicle detection. The usual sliding window technique slides the windows at a fixed sliding step which didn't ensure the coverage of entire vehicle exactly. Therefore, the window was slided to its geometric center which is the arithmetic mean of all the pixel values in the image. If the initial window covered minor parts of the image, sliding once was not enough. Thus the window was enlarged to get the main part of the vehicle and was slided again to its new geometric center. The sliding windows were collected for 3 gradient images and the repetitive windows were filtered and the remaining unfiltered image was sent to the hybrid deep neural network for classification.

Conclusion: 99.7% vehicles were located exactly. This can be a reliable method for vehicle detection as it is able to locate almost 100% of the vehicles.

CHAPTER 6

FACIAL EXPRESSION RECOGNITION USING THE GENERAL REGRESSION NEURAL NETWORK.

Introduction: This chapter explains the method of recognition using general regression neural networks

Method used and its results: Facial expression using the local binary pattern for feature extraction and artificial neural network for feature classification was presented. Facial feature vectors were obtained using the local binary pattern algorithm by considering blocks of 4 different sizes i.e. 256x256, 128x128, 64x64 and 32x32.The multiclass classification of the image into 6 basic universal expressions was done using artificial neural networks. The 6 expressions were happy, surprise, unhappy, disgust, fear and anger. The general regression algorithm was also used for classification purpose as this method doesn't require iterative training procedure.

Local Binary Pattern: It is an efficient operator for texture representation of image pattern and a simple tool for detection of features and is vigorous to the illumination variations in an image. The pixels of the image was labeled by the binary operator by comparing the center pixel value with the 3x3 neighborhood of each pixel to form a 8 bit binary number which was then converted to decimal value. For a given pixel at (x, y) the local binary pattern was obtained using the equation given below

$$\text{LBP}(x, y) = \sum_{p=0}^{7} S(g_p - g_c) * 2^p, \quad S(x) = 1 \text{ for } x \geq 0 \ldots\ldots\ldots (6.1)$$

$$= 0 \text{ for } x < 0$$

where, g_c = gray value of center pixel

g_p = gray value of neighbouring pixel of g_c.

The input image block of 3x3 size was considered. The value of the central pixel is 7 and is surrounded by the 8 pixels in all eight directions. The Local binary pattern converts all 9 pixel values in to a single value. This was done by comparing the pixel value of every neighboring pixel with the central pixel value that is the intensity value at that point. The pixel values were usually considered in the grayscale. The pixel with the value greater than or equal to the central pixel was assigned as 1 and lower value was assigned as 0 since only binary values of 0 and 1 can be assigned to the pixels. Now these 8 pixel values surrounding the center pixel forms 1 byte. This resultant byte then was turned into a decimal number. As long as binary values are present, any pixel block can be encoded into a byte. Converting this byte into a decimal value yields a single decimal value for the block. Thus using the values obtained for each of the block in the complete image, the feature vectors obtained. These feature vectors acts as the input for the classification process. The feature vector reduces the entire image data into a single vector ready for classification. This feature vector gives the information about how efficiently the face expressions can be identified. The efficiency of the algorithm was tested for four different blocks of sizes 256x256, 128x128, 64x64 and 32x32

Facial Expression Recognition system: The facial expression system consists of two main phases namely training phase and testing phase. The training phase consists of known input facial samples that are processed digitally using preprocessing tools and techniques. The second phase is the testing of any unknown facial expression. The unknown face expression is cropped and the extractions of facial features were done. The facial features were classified into 6 basic emotional states of a person i.e. happy, unhappy, surprise, disgust, fear and angry. The preprocessing of input image was used to reduce noise and eliminate illumination variations. It involves conversion of colored image into a gray scale image and then resizing it for fast processing.

General Regression Neural Network: Every neural network works on the principle that it requires training data to train itself and then classify unknown data into known classes based on the training. The network is then able to predict the output for the unknown test data and classify the new test images. The General Regression Neural Network consists of 4 layers i.e. input layer, pattern layer, summation layer and output layer. The output estimate was obtained by calculating the Euclidean distance between the training and test data. The input layer acts as the feed to the next layers. The pattern layers were used for calculating the Euclidean distance and the activation function. The activation function decides how much weight the training sample contributes. Summation layer consists of the numerator and denominator part fed to the final output layer. The basic principle on which this network works is the joint probability estimate of the input and output given below

$$f(x,y) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} * \frac{1}{n} \sum e^{\frac{(x-x_i)^T (x-x_i)}{2\sigma^2}} \quad e^{\frac{(y-y_i)^2}{2\sigma^2}} \text{.........(6.2)}$$

Where n is the number of observed samples, σ is the spread parameter, x_i is the ith training vector, y_i is the corresponding output value. The conditional mean of y given x is given by

$$\mathrm{E}\left(\frac{\mathrm{Y}}{\mathrm{X}}\right) = \frac{\int_{-\infty}^{+\infty} \mathrm{Y}\, \mathrm{f(x,y)dy}}{\int_{-\infty}^{+\infty} \mathrm{f(x,y)dy}} \text{.........(6.3)}$$

From (1) and (2) we get

$$\mathrm{Y(x)} = \frac{\sum \mathrm{Y_i} e^{-\left(\frac{d_i^2}{2\sigma^2}\right)}}{\sum e^{-\left(\frac{d_i^2}{2\sigma^2}\right)}} \text{......... (6.4)}$$

Where $d_i^2 = (x - x_i)^2 (x - x_i)$ the Euclidean is distance between x and x_i.....(6.5)

$e^{-\left(\frac{d_i^2}{2\sigma^2}\right)}$ is the activation function

From (6.5) it can be inferred that if the Euclidean distance has small value the activation function will return a large value and if the Euclidean distance has large value the activation function function will return a small value. The only free parameter in the network is the spread constant. The training procedure involves calculation of spread parameter so that the error is minimized. The advantage of of general regression neural network over other neural network is it can be trained faster and it is easy to determine only one parameter that is the spread constant. The maximum efficiency of classification was seen for 64x64 block among the other 3 blocks.

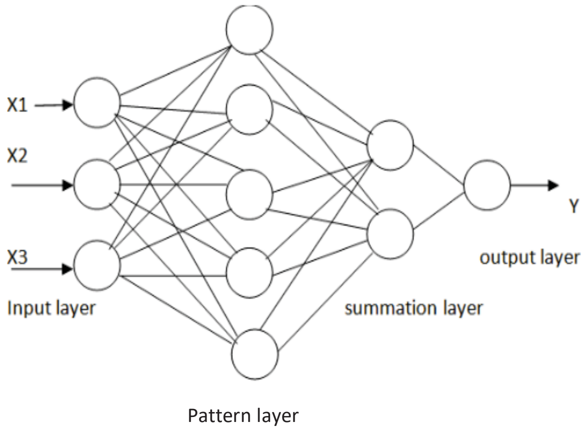

Figure 6.1: Artificial neural networks for facial expression recognition

Conclusion: Local binary pattern was obtained for different block dimensions and for all the block the efficiency of recognition of face was more than 90%. Therefore this method can be the reliable method for face recognition.

CHAPTER 7

A HOPFIELD NEURAL NETWORK FOR IMAGE CHANGE DETECTION

Introduction: This chapter explains the method of image change detection using Hopfield neural network

Method used and its results:

Image change detection: The change in the image was seen for two images by obtaining a difference image formed after subtracting pixel by pixel in both images. Let the two images be I_1 and I_2.There were four major detecting schemes used for detecting the difference between two images of the same scene taken at different times.

1) **Temporal difference model:** A difference image D was computed at pixel location (x, y) by subtracting the corresponding intensity values of the images. The changed image O(x , y) was generated according to the following decision rule

$$O(x,y) = \begin{cases} 1, & \text{if } \; D(x,y) > T \text{ change} \\ 0, & \text{otherwise} \qquad\qquad \text{no change} . \end{cases}$$

   The threshold T was chosen empirically. This method was very sensitive to changes in illumination.

2) **Significance and hypothesis test models**: The decision to whether or not a change has occurred at the given pixel location (x, y) corresponded to choosing one of the two hypothesis i.e. the null hypothesis $\mathbf{H_0}$ or the alternative hypothesis $\mathbf{H_1}$.Null hypothesis is the hypothesis where there is no significant difference between two specified populations due to experimental error. The alternative hypothesis is opposite to null hypothesis i.e. there is a difference between two specified populations due to some variation superimposed on the population. Let the image pair be $\mathbf{I_1}$ and $\mathbf{I_2}$ whose difference image found out as D was viewed as a random vector. Both the hypothesis was characterized by modeling the observations of difference image D(x, y) assuming the gray level

differences to assume Gaussian distributions with variances $\boldsymbol{\sigma_0^2}$ **and** $\boldsymbol{\sigma_1^2}$ for $\mathbf{H_0}$ **and** $\mathbf{H_1}$ respectively.

3) **Vector and shading models**: This method is also known as linear dependence detector and is based on linear algebra. Each pixel at location (x, y) of images can be represented by one vector of n neighboring pixels $\mathbf{x = x_1, x_2, x_3 \dots \dots x_n}$. Therefore whether or not the same pixel is changed in different frames with $\mathbf{y = y_1, y_2, y_3 \dots \dots y_n}$ can be validated if both representative vectors are linearly dependent or independent. The shading model is closely related to vector model. It generally compares the ratio of image intensities to a threshold value determined empirically.

4) **Clustering Models**: Every background pixel of an image can be divided into clusters and these clusters can be modeled as Gaussian distributions about mean values that vary from cluster to cluster and this abnormality have pixel values that deviate significantly from the distribution of the cluster. The clustering was carried out by vector quantization technique. Vector quantization technique is a process of modeling the probability distribution function using vectors. Let there be L levels to which the pixel values of the image can be assigned. The levels are chosen in such a way so that the mean square quantization error is minimized. The mean square quantization error (MSQE) is given by the equation below

$$\mathrm{MSQE} = \sum_{l=0}^{L-1} \left[\int_{x_l}^{x_{l+1}} (x - x_l)^2 p(x) dx \right]$$

Where p(x) is the probability distribution over the intensity values of the image

Hopfield Neural Network: The Hopfield neural network is a recurrent neural network whose output nodes are connected to the input nodes i.e. it contains feedback paths. This means that if an input is applied to the network the output produced will be fed back to the input so that the input is modified. The output is again recalculated and the process is repeated again and again until the outputs become constant. When the output becomes constant the network is said to have achieved stability. The connection weights between the nodes of the network were considered in matrix form. If this matrix is a symmetric matrix then the network is said to be stable.

Let us consider a special case of Hopfield neural network with a symmetric matrix. The input to some ith node comes from 2 sources one is the external source and the other is the inputs from the other nodes. Let $\mathbf{u_i}$ be the total input for some node i which is given as shown below

$$u_i = \sum_{j \neq i} T_{ij} V_j + \theta_i$$

Where $\mathbf{V_j}$ represents output of the jth node, $\mathbf{T_{ij}}$ represents the connection weights between the nodes i and j and $\boldsymbol{\theta_i}$ is the external input bias value. There are two types of Hopfield neural networks one is analog where the states of neurons are allowed to continuously vary between -1 and +1 and other is discrete where the states of neuron are restricted to the binary values 0 and 1.The drawback of discrete neural network is the oscillation of network between different binary states and setting down of the network at many stable states. The analog neural network performs better than the discrete neural network as it has the ability of not getting stuck at some local minima. For analog Hopfield networks the total input to the node can be converted into an output value by sigmoid mono activation function. The dynamics of a node can be defined as shown below

$$\frac{du_i}{dt} = -\frac{u_i}{R_i} + \sum_{j \neq i} T_{ij} V_j + \theta_i \quad \text{where } V_j = g(u_j)\ \forall j$$

Where g($\mathbf{u_j}$) is the sigmoid activation function and $\mathbf{R_i}$ is the time constant. The sigmoid activation function was chosen as the hyperbolic tangent function Tanh(x).

Image change detection by Hopfield neural network: The problem of image detection by Hopfield neural network was to label each pixel as changed or unchanged. With such a purpose the output image was considered as a network of nodes where each node was associated to a pixel location in the difference image i.e. the number of nodes is exactly the number of pixels of the incoming image. Also each node was characterized by its state value $\mathbf{V_i}$ ranging between -1 and +1.After the optimization process by the neural network, the stability of the neural network was reached and the nodes achieved its final state value. This final value will indicate unchanged for -1 and maximum change for +1 and the intermediate values gives the strength of the change. The following steps were carried out for detecting the change in the image.

Summary of change detection procedure:

1) Each node i was created for the pixel location (x, y) from the difference image. Each node was loaded with state values $\mathbf{V_i}$.
2) For each node i the value of $\mathbf{V_i(t)}$ was calculated and updated using Runge Kutta method
3) Each updated $\mathbf{V_i}$ indicates changed or unchanged pixel locations.

Conclusion: Mapping of data was done perfectly as compared to the classical manual method. This method has proven to be robust against noise compared to other methods. So, the Hopfield method can be used for real time surveillance tasks.

CHAPTER 8

IMAGE SEGMENTATION USING PARTICLE SWARM INTELLIGENCE

Introduction: This chapter explains the method of segmenting an image using particle swarm optimization problem

Method Used and its results:

Particle swarm optimization problem

Suppose there are group of birds searching for food in an area. There is only one piece of food available. All the birds do not know where the food is exactly but they know at what distance the food is available. The best strategy to acquire the food is to follow the bird that is nearest to the food. The above scenario can be made analogous to a particle swarm optimization problem where particle is analogous to a bird. Each particle has a fitness value evaluated by the fitness function to be optimized. Each fitness function has velocities which direct the flying of the particles.

Let D – Dimension of the search space

M- No. of particles in D Dimension space.

Let i_{th} particle in the search space be a vector X = $(X_{i1}, X_{i2}, X_{i3}, \ldots \ldots \ldots . . X_{id})$ where

I = 1,2,3…….M in the D dimension search space.

Now movement of each particle is analyzed which is influenced by its best known position. For every iteration the best known position is found out which is the best possible solution.

Two best solutions is seen one is local best value which is

$P_{best1} = (P_{i1}, P_{i2}, P_{i3}, \ldots \ldots \ldots . P_{id})$ ….……………….(8.1)

Another best value is the globally best value among the population

$G_{best2} = (P_{g1}, P_{g2,} P_{g3}, \ldots \ldots \ldots \ldots . . P_{gd})$……………(8.2)

After finding out the two best values of velocities and positions, the current velocity and position value of the particle was updated as follows:

$$V_{id}(t+1) = wV_{id}(t) + k_1\big(P_{id}(t) - X_{id}(t)\big) + K_2(P_{gd}(t) - X_{id}(t))\ldots\ldots(8.3)$$

$$X_{id}(t+1) = X_{id}(t) + V_{id}(t+1) \ldots\ldots\ldots\ldots (8.4)$$

Where V = Velocity of the particle, X= Position of the particle, t = time of previous best position, t+1 = time of next best position.

Fisher's two dimensional image segmentation algorithm

Fisher's linear discriminant is a method that is used to project high-dimensional data onto a line and performs classification in this one-dimensional space**.** It **a**ims to search for a linear transform

that reduces the dimension of a given n-dimension statistical model, consisting of K classes to d ($d \leq n$) dimensions. The transform should be such that a maximum amount of discrimination information is preserved in the lower dimensional model and it tries to preserve distance of already well separated classes, which may result in a large overlap or even occlusion of neighboring classes. The Fisher's criterion is defined as

$$J(w) = \frac{|M_1 - M_2|^2}{S_1^2 + S_2^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8.5)$$

Where m is the mean, s is the variance and the subscripts denote the two classes.

The Threshold pair to which the position of the particles is to be mapped to segment the image was found out by maximizing the above equation.

Particle swarm optimization in Image segmentation

1) Input image to be segmented is having levels [0 to 255] of maximum N iterations. These levels [0 to 255] represent particles which are generated at t=0.
2) Position of the particles is mapped to threshold pair using 2-D Fisher's discriminant method.
3) The best position of the particle was found out which was the local best value
4) The global best position of the particle was found out
5) The velocity and position of the particles were updated regularly.

6) If the particles generated were >N then position of the particles was mapped in the best threshold pair and the image was segmented
7) If the particles generated were <N then the algorithm was started from step 1.

Conclusion: This is the good method of segmenting an image but it gives rise to exponential increment of computation time.

CHAPTER 9

NON DESTRUCTIVE WATERMELON RIPENESS DETECTION USING ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS.

Introduction: This chapter explains the method of determining the ripeness of water melon using artificial intelligence and neural networks

Method used and its results: The process to determine the ripeness of watermelon ripeness involved 5 steps.

1) Data collection: The watermelon samples were collected and the samples were divided into 3 categories i.e. ripe, under ripe and over ripe.90 samples were collected of which 45 samples were used as the training set, 15 samples were used as the validation set and 30 samples were used as the testing set.

As the collection of data was very small for training the artificial neural network a virtual training patterns were generated by adding Gaussian noise to the true training points. Thus we get 360 data points all together of which 215 points were used for training, 35 samples for validation and 110 samples for testing. Hence artificial neural network was trained using the new data set.

2) **Image Acquisition:** The water melon images were collected using a digital camera under standard and controlled environment in the image capturing studio room.

3) **Preprocessing:** The Y Cb Cr color space was used for image segmentation and processing. The Y component stores the luminance information while the Cb, Cr component stores the chrominance information. The Cb component represents the difference between the blue component and the reference value. Cr represents the difference between red component and the reference value. The Y Cb Cr component are generated as follows

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$Cb = -0.1687 * R - 0.3312 * G + 0.5 * B$$

$$Cr = 0.5 * R - 0.4186 * G - 0.0813 * B$$

Where R, G, B are non linear Red, Green and Blue components. Each of the components has a nominal range from 0 to 1 where 0 represents the minimum intensity and 1 represents the maximum intensity. The resulting Y value will have nominal range from 0 to 1 and chrominance values will have the nominal range from -0.5 to 0.5.

The color in watermelon image is segmented into 3 dimensional color map and then converted into YCbCr color space. The Y component is completely eliminated because the color of the image will not be affected by the brightness of the surroundings. Therefore image contains only component colors namely Cb and Cr

Figure 9.1: Image obtained before preprocessing

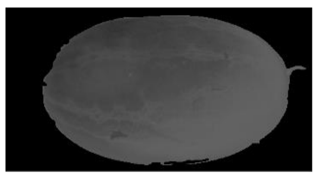

Figure 9.2: Image obtained after preprocessing

4) **Feature Extraction**: The Cb Cr color feature was extracted after completing the pre processing stage. The sum of each color chrominance (Cb and Cr) was computed for every pixel area to determine the mean value.

Artificial neural network model: Artificial neural network is a problem solving tool that has become an alternating modeling method to some physical and non physical systems with scientific or mathematical basis. The ANN acquires knowledge from its environment through a learning process and synaptic weight connections are used to store the acquired knowledge. The common type of ANN is called the multi-layer perceptron (MLP). It consists of three layers of units .A layer of input unit is connected to a layer of hidden unit which is connected to a layer of output units shows an example of three layer MLP architecture. This method employed supervised learning where the target values for the outputs were presented to the network by updating the weights of the network. Supervised learning attempts to match the output of the network to values that have already been defined. After training network, verification is applied in which only the input values were presented to the network so that the success of the training could be established. The input for the ANN model will be the mean value of the two color components namely Cb and Cr.

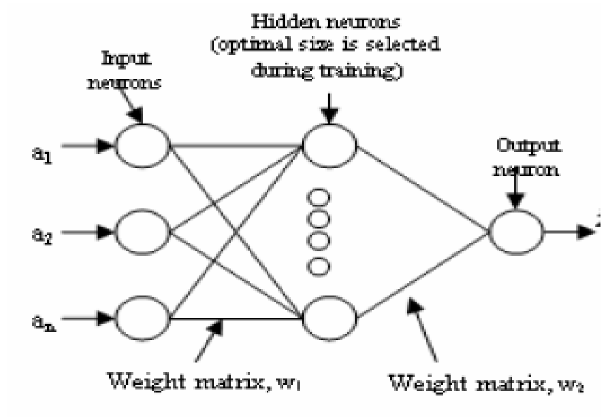

Figure 9.3: Artificial neural networks for water melon ripeness detection

Conclusion: The output will be the ripeness stage of the watermelon with an accuracy of 86.51%.. If it were a little more efficient i.e. at least 90% then would have been considered as a reliable method.

CHAPTER 10

IMAGE DENOISING USING A NEURAL NETWORK BASED NON-LINEAR FILTER IN WAVELET DOMAIN.

Introduction: This chapter explains the image de noising method using a neural network based nonlinear filter in wavelet domain.

Method used and its results: Images are often corrupted as a result of various factors that can occur during acquisition and transmission processes. Image de noising is aimed at removing or reducing the noise so that a good-quality image can be obtained for various applications. In this method of image de noising a noisy image was first wavelet transformed into four sub bands and then a trained layered neural network was applied to each sub band to generate noise-removed wavelet coefficients from their noisy ones. The de noised image was thereafter obtained by applying the inverse transform on the noise-removed wavelet coefficients. Compared with other methods performed in wavelet domain, it requires no a priori knowledge about the noise and needs only one level of signal decomposition to obtain very good de noising results. Wavelet transform provides excellent properties for signal and image processing. One of the non-linear de noising methods performed in the wavelet transform domain was wavelet thresholding. This method removes the noise in an image by removing the wavelet coefficients that are too noisy and preserving or shrinking the coefficients that contain important image signals. The success of the method depends heavily on the choice of the threshold parameters. A layered neural network was properly designed and trained to explore the learning capability of the neural network to learn the correlation between the noise-free wavelet coefficients and their noisy observations. After the training process the network is applied to the noisy coefficients to produce noise-reduced output values. Unlike the wavelet thresholding methods which usually require three or more levels of wavelet decomposition and need the accurate estimate of the involved noise to obtain good de noising results, this method needs only one level of wavelet decomposition and can adapt itself to the various noise environments by learning.

Let y be the corrupted signal which can be modeled as shown below

$$y = x + v \quad \text{………………….(10.1)}$$

Where x is the original uncorrupted signal and v is the density function of noise and is independent of x. The noise signal used to corrupt the signal was assumed to be an additive noise. In the wavelet transform domain equation (1) can be expressed as shown below

$$Y = Wy \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10.2)$$

Where Y is the wavelet transformed version of y and W is the orthogonal wavelet transform operator. Since W is orthogonal and v is independent of x, Y can be further expanded as

$$Y = Wx+Wv = X+V \ldots\ldots\ldots..(10.3)$$

Where X is the contribution to the wavelet coefficients related to the original signal x and V is the contribution to the wavelet coefficients related to the noise v. The de noising problem in the wavelet transform domain is to obtain the estimate X' of X from Y. The de noised signal in spatial domain is simply the inverse transformed version of x' of X' i.e.

$$x' = W^{-1}X' \ldots\ldots\ldots\ldots\ldots (10.4)$$

Where W^{-1} is the inverse wavelet transform operator. The layered neural network was used to learn the correlation coefficients corresponding to signal and therefore used to remove the coefficients corrupted due to the noise.

The Layered Neural Network filter: The Layered neural network filter is a three-layer neural network with inputs derived from an NxN neighborhood of the transformed image and appropriately selected neuron activation functions. The network takes Y_p and ΔY_k as the inputs, where Y_p is the wavelet transform coefficient under consideration, which is the center of a N x N processing window and $\Delta Y_k = Y_k - Y_p$ is the difference value between Y_p and the coefficient Y_k where k = 0,1,2,………..$N^2 - 1$, k$\neq$ p of other points in the NxN window. The figure 2 shows the processing window with a size of 5x5 pixels where Y_{12} is the center of the window and $\Delta Y_k = Y_k - Y_{12}$ where k = 0, 1, 2, 3…………24 and K≠12

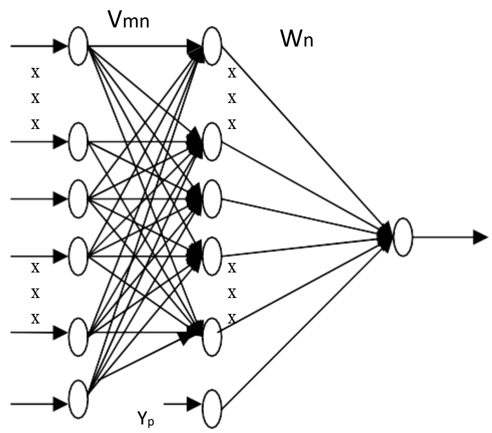

Figure 10.1: Neural Network Structure

Y_0	Y_1	Y_2	Y_3	Y_4
Y_5	Y_6	Y_7	Y_8	Y_9
Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}
Y_{15}	Y_{16}	Y_{17}	Y_{18}	Y_{19}
Y_{20}	Y_{21}	Y_{22}	Y_{23}	Y_{24}

Figure 10.2: a 5x5 processing window

The output of the neural network was formulated as shown below

$$X'_p = f_o\left(\sum_{n=0}^{M} W_n\, f_h\left(\sum_{m=0}^{N^2} Y_m\, v_{mn}\right)\right) \ldots\ldots (10.5)$$

Where X'_p is the output of the layered neural network corresponding to the central point in the processing window, Y_m is the mth input layer, M is the number of neurons in the hidden layer,w_n is the connecting weight between the nth neuron in the hidden layer and the neuron in the output layer, v_{mn} is the weight of the connection between the m-th neuron in the input layer and the n-th neuron in the hidden layer. The activation functions of the neurons in hidden and output layers are

$$f_h\,(x) = xe^{-\gamma x^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10.6)$$ and

$$f_o(x) = \frac{2}{1 + e^{-\lambda x}} - 1 \ldots\ldots\ldots\ldots\ldots..(10.7)$$

Where $\gamma > 0$ and $\lambda > 0$ are the function coefficients. The selection of f_h (x) shown in equation (6) provides the network with a good ability to learn the correlation of the wavelet coefficients. A four layered neural network each having the same structure as shown in Figure 1 was used in this method. A noisy image was decomposed into four sub bands using a wavelet transform. Each of the layered neural network was trained using one of the four sub bands of the decomposed image. After the training process the four layered neural networks were applied to the corresponding sub band of the wavelet transformed noisy image.

Conclusion: The outputs of the network were noise-removed coefficients and the de noised image was obtained by performing an inverse wavelet transform on the coefficients. This is the reliable method of de noising an image as the visual quality was not degraded after finding out wavelet transform and inverse wavelet transform coefficients.

CHAPTER 11

CONCLUSION

The Methods related to image processing using artificial intelligence and neural networks were discussed. Image segmentation is a very tedious task to do, but using pulse coupled neural networks, an image can be perfectly segmented without any hassle. If an Image of a fruit is there, then it is difficult to judge whether the fruit is ripe or not, but with the help of neural networks it can be easily said whether the fruit is ripe or not. Image segmentation can also be done almost accurately using search optimization problem. If there are two images present and if it is told to detect the difference between two images then it will be very difficult because mapping of each pixel has to be done one by one which is a tedious task. The Hopfield neural network can detect any change in the image in a easy way without taking any time. If the satellite image of a particular region is taken and a vehicle has to be detected in that image then manually it is going to be a very difficult task. By using deep convolutional neural networks, this particular work can be done very easily. Similarly, Signature recognition, finger print recognition, facial recognition, image de blurring, image de noising is also very easy to do using neural networks.

BIBLIOGRAPHY

[1] J. C. Bezdek, L. O. Hall, and L. P. Clarke,“Review of MR image seg-mentation techniques using pattern recognition,”Med. Phys.
, vol. 20,pp. 1033–1048, 1993.

[2] D. L. Pham, C. Y. Xu, and J. L. Prince,“A survey of current methodsin medical image segmentation,”
Annu. Rev. Biomed. Eng., vol. 2, pp.315–337, 2000.

[3] W. M. Wells, W. E. LGrimson, R. Kikinis, and S. R. Arrdrige,“Ada-tive segmentation of MRI data,”IEEE Trans. Med. Imag., vol. 15, pp.429–442, Aug. 1996.

[4] J. C. Bezdek,Pattern Recognition With Fuzzy Objective Function Algo-rithmsNew York: Plenum, 1981.

[5] D. L. Pham and J. L. Prince,“An adaptive fuzzy C-means algorithmfor image segmentation in the presence of intensity inhomogeneities,”Pattern Recognit. Lett., vol. 20,pp.57–68, 1999.

[6] Y. A. Tolias and S. M. Panas,“On applying spatial constraints in fuzzyimage clustering using a fuzzy rule-based system,”
IEEE Signal Pro-cessing Lett., vol. 5, pp. 245–247, Oct. 1998.

[7]“Image segmentation by a fuzzy clustering algorithm using adap-tive spatially constrained membership functions,”IEEE Trans. Syst.,Man, Cybern. A, vol. 28, pp. 359–
369, May 1998.

[8] A. W. C. Liew, S. H. Leung, and W. H. Lau,“Fuzzy image clusteringincorporating spatial continuity,”Inst. Elec. Eng. Vis. Image SignalProcess, vol. 147, pp. 185–192, 2000.

[9] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Mo-riarty,“A modified fuzzy C-means algorithm for bias field estimationand segmentation of MRI data,”
IEEE Trans. Med. Imaging, vol. 21, pp.193–199, Mar. 2002.

[10] D. L. Pham,“Fuzzy clustering with spatial constraints,”inIEEE Proc.Int. Conf. Image Processing, New York, Aug. 2002, pp. II-65–II-68.

[11] K. R. Muller and S. Mikaet al.“An introduction to kernel-basedlearning algorithms,”IEEE Trans. Neural Networks, vol. 12, pp.181–202, Mar. 2001.

[12] N. Cristianini and J. S. Taylor,An Introduction to SVM’s and OtherKernel-Based Learning Methods. Cambridge, U.K.: Cambridge Univ.Press, 2000.

[13] V. N. Vapnik,Statistical Learning TheoryNew York: Wiley, 1998.

[14] B. Scholkopf, Support Vector Learn.

[15] R. L. Galloway, Jr., R. J. Maciunas, and A. L. Failinger, “Factorsaffecting perceived tumor volumes in magnetic resonance imaging,”Ann. Biomed. Eng., vol. 21, pp. 367–375, 1993.

[16] P. A. Narayana, E. F. Jackson, and J. S. Wolinsky, “Inter- and intraoper-ator variability in the quantitative volumetric measurements of multiplesclerosis lesions,”J. Magn.Reson.Imag.,(Supp.), vol. 3, pp. 111–112,1993. (Abstract)

[17] T. Taxt, A. Lundervold, B. Fuglaas, H. Lien, and V. Abeler, “Multispec-tral analysis of uterine corpus tumors in magnetic resonance imaging,”Magn. Reson. in Med., vol. 23, pp. 5576, 1992.

[18] R. P. Velthuizen, L. P. Clarke, S. Phuphanich, L. O. Hall, A. M. Bensaid,J. A. Arrington, H. Greenberg, and M. L. Silbiger, “Unsupervisedmeasurement of brain tumor volume on MR images,”J. Magn. Reson.Imag., vol. 5, pp. 594–605, 1995.

[19] S. C. Carpentieri, R. K. Mulhern, S. Douglas, S. Hanna, and D.L. Fairclough, “Behavioral resiliency among children surviving braintumors: A longitudinal study,”J. Clin. Child Psychol., vol. 22, pp.236–246, 1993.

[20] R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, “Face Detection in ColorImages,” IEEE Trans. Pattern Analysis and Machine Intelligence,vol. 24, no. 5, pp. 696-707, May 2002.

[21] J. Yang and A. Waibel, “A Real-Time Face Tracker,”Proc. IEEE Workshop Applications of Computer Vision,pp. 142-147, Dec. 1996.

[22] X. Zhu, J. Yang, and A. Waibel, “Segmenting Hands of Arbitrary Color,” Proc. IEEE Int’l Conf. Automatic Face and Gesture Recognition,pp. 446-453, Mar. 2000.

[23] M.J. Jones and J.M. Rehg, “Statistical Color Models with Application to Skin Detection,”Int’l J. Computer Vision,vol. 46, no. 1, pp. 81-96, Jan. 2002.

[24] J. Brand and J. Mason, “A Comparative Assessment of Three Approaches to Pixel-Level Human Skin Detection,”Proc. IEEE Int’l Conf. Pattern Recogni-tion, vol. 1, pp. 1056-1059, Sept. 2000.

[25] “NDT Resource Center” http://www.ndt-ed.org index_flash.htm.

[26] Hagan, M.A., H.B. Demuth, M.H. Beale. (2003): Neural Network Design, Brooks Cole, ISBN: 0-9717321-0-8.
[27] L. Zadeh. (1987): Fuzzy Sets and Applications: Selected Papers by L.A. Zadeh, ed. R.R. Yager et al, John Wiley, New York.
[28] Kumar, S. and F. Taheri, F. (2004): "Neuro-Fuzzy Approaches for FRP Oil and Gas Pipeline Condition Assessment", American Society of Mechanical Engineers, Pressure Vessels and Piping Division (publication), V490, Storage Tank Integrity and Materials Evaluation, p271275.
[29] "Image processing toolbox user's guide" (2005) The Math Works, Natick, Massachusetts, USA.
[30] W. Pratt,Digital Image Processing, 2nd ed. New York: Wiley, 1991.
[31] S. Perkins, J. Theiler, S. Brumby, N. Harvey, R. Porter, J. Szymansk, and J. Bloch,"GENIE: A hybrid genetic algorithm for feature classification in multi-spectral images," in Proc. SPIE 4120 Appl. and Sci. Neural Netw.,Fuzzy Syst. and Evol. Comput. III, 2000, pp. 52–62.
[32] P. Zhang, B. Verma, and K. Kumar, "Neural vs statistical classifier in conjunction with genetic algorithm feature selection in digital mammog-raphy," inProc. IEEE Congr. Evol. Comput., Canberra, Australia, 2003,pp. 634–639.
[33] Q. Chen, C. Zhou, J. Luoet al., "Fast segmentation of high-resolution satellite images using watershed transform combined with an efficient region merging approach,"Lecture Notes Comput. Sci., vol. 33, no. 22,pp. 621–630, 2004.
[34] P. Pina and T. Barata, "Classification by mathematical morphology," inProc. IEEE Int. Geosci. and Remote Sens. Symp., 2003, pp. 3516–3518.
[35] International Conference on VLSI, Communication & Instrumentation (ICVCI) 2011 Proceedings published by International Journal of Computer Applications® (IJCA) – A Review Of Medical Image Classification Technique
[36] IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 5, No 1, September 2013 ISSN (Print):1694-0814 | ISSN (Online): 1694-0784 .
[37] IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727 Volume 15, Issue 3 (Nov. - Dec. 2013), PP 71-78 www.iosrjournals.org - A Survey of Image Segmentation based on Artificial Intelligence and Evolutionary Approach

[38] G Anil Kumar, Int.J.Computer Technology & Applications,Vol 5 (3),851-860 IJCTA | May-June 2014 ISSN:2229-6093 www.ijcta.com - Analysis of Medical Image Processing and its Applications in Healthcare Industry
http://www.ijcta.com/documents/volumes/vol5issue3/ijcta2014050308
[39] Asachi" din Iaúi Tomul LIX (LXIII), Fasc. 2, 2013 Secēia AUTOMATICĂú
CALCULATOARE MEDICAL BULETINUL INSTITUTULUI POLITEHNIC DIN IAùI Publicat de Universitatea Tehnică„Gheorghe IMAGE PROCESSING BY MEANS OF SOMEARTIFICIAL INTELLIGENCE METHODS BY HARITON COSTIN and CRISTIAN ROTARIU http://www.ace.tuiasi.ro/users/103/023-041_2_Costin_.pdf
[40] M. Bokser, "Omnidocument technologies", Roc. IEEE,
[41] S. V. Rice, F. R. Jenkins, and T. A. Nartker. "OCR accuracy: UNLV's fifth annual test",INFORM, 10(8), September 1996.
[42] L. O Gorman and R. Kasturi, "Document Image Analysis", IEEE Computer Society Press, Los Alamitos, 1995.
[43] J. Ohya, A. Shio, and S. Aksmatsu, "Recognition characters in scene images. IEEE Trans. Pattern Analysis and Machine Intelligence", 16(2):214--220, 1994. [SI V. Wu, R. Manmatha, and E. M. Riseman, "Finding text in images", In Proc,ACM Int. Conf. Digital Libraries, 1997.
[44] 1. N. Alldrin, A. Smith and D. Turnbull, Clustering with EM and K-means, unpublished manuscript, http://louis.ucsd.edu/ nalldrin/research/wo3/cse253/project1.pdf.
[45] M. Antonie, O. Zaiane and A. Coman, Application of data mining techniques for medical image classification, in Proc. 2nd Int. Workshop on Multimedia Mining ,in San Francisco, USA, 2001, pp. 94–101.
[46] J. Bala, J. Juang, H. Vafaie, K. DeJong and W. Wechsler, Hybrid learning using genetic algorithms and decision trees for pattern classification,Int. Joint Conf. Arti-ficial Intelligence, Montreal, August, 1995, pp. 719–724.
[47]. G. Ball and D. Hall, A clustering technique for summarizing multivariate data,Behav. Sci.12(1967) 153–155.
[48]. J. Bezdek, A convergence theorem for the fuzzy ISODATA clustering algorithms, IEEE Trans. Patt. Anal. Mach. Intell.2 (1980) 1–8.

[49] E.S. Berner, M.J. Ball, Clinical decision support systems: theory and practice, Springer, Berlin, 1998.
[50] R. Mousa, Q. Munib, Expert Syst. Appl. 28 (2005) 713.
[51] D.R. Chen, R.F. Chang, Y.L. Huang, Radiology 213 (1999) 407.
[52] A. Reeves, W. Kostis, Radiologic Clinics N. Amer. 38 (3) (2000) 497.
[53] T.-K. Yin, N.-T. Chiu, IEEE T. Bio.-Eng. 51 (7) (2004) 1286.
[54] M. Turk, A. Pentland, Face recognition using Eigenfaces,Proceedings o fthe IEEE Conference on Computer Vision andPattern Recognition, 1991, pp. 586 –591.
[55] P.N. Belhumeur, J.P. Hespanha, D.J. Kriengman, Eigenfaces vs.Fisherfaces: recognition using class speci<c linear projection,IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711-720.
[56] K. Liu, Y.-Q. Cheng, J.-Y. Yang, et al., Algebraic feature extraction for image recognition based on an optimal discriminant criterion, Pattern Recognition 26 (6) (1993) 903–91.
[57] L.M. Fletcher-Heath, L.O. Hall, D.B. Goldgof, F.R. Murtagh, Automatic segmentation of non-enhancing brain tumors in magnetic resonance images,Artif.Intell. Med. 21 (2001) 43–63.
[58] S. Chaplot, L.M. Patnaik, N.R. Jagannathan, Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network, Biomed. Signal Process. Control 1 (2006) 86–92.
[59] F. Gorunescu, Data mining techniques in computer-aided diagnosis: Non-invasive cancer detection, PWASET 25 (2007) 427–430.
[60] S. Kara, F. Dirgenali, A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks,ExpertSyst. Appl. 32 (2007) 632–640.
[61] M. Maitra, A. Chatterjee, Hybrid multiresolution Slantlet transform and fuzzy c-means clustering approach for normal-pathological brain MR imagesegregation, Med. Eng.Phys.(2007), doi:10.1016/j.medengphy.2007.06.009
[62] M.G. Khan, Heart Disease Diagnosis and Therapy, Williams& Wilkins, Baltimore, 1996.
[63] F.A. Burgener, S.P. Meyers, R.K. Tan, W. Zaunbauer,Differential Diagnosis in Magnetic Resonance Imaging,Thieme, Stuttgart, 2002.
[64] S.E. Silvus, Ch.A. Rohrmann, H.J. Ansel, Text andAtlas of Endoscopic Retrograde Cholangiopancreatography,Igaku-Shain, New York, 1995.

[65] M.R. Ogiela, R. Tadeusiewicz, Semantic-oriented syntactic algorithms for content recognition and understanding of images in medical data bases,in:Proceedings of the 2001 IEEE International Conference on Multimedia and Expo-ICME,Tokyo, 2001, pp. 621–624.

[66] R. Tadeusiewicz, M. Flasinski, Pattern Recognition, PolishScientic Publisher, Warsaw, 1991.

[67] N. J. W. Throwers, Maps and Man: An Examination of Cartography Englewood Cliffs, NJ: Pren- 121 .

[68] Proc. Ini. Symp. Spatial Data Handling. Zurich, Switzerland, 1984,

[69] L. E. Starr. "Computer-assisted cartography research and develop- 141

[70] D. M. McKeown. Jr., "Maps," Dep. Comput. Sci., Carnegie-Mel- in Relation to Culiure and Civilization. tice-Hall. 1972. pp. 10-15. and Second Ini. Symp. Sparial Dura Handling, Seattle, WA, 1986. ment report," Int. Cartographic Ass., July 1984. Ion Univ., Pittsburgh, PA, Rep. CMU-CS-83-117, 1983.

[71] W. C. Mahoney, "Defence mapping agency (DMA) overview of mapping, charting and geodesy (MCIG) applications of digital image pattern recognition," in Techniques and Applications of Image Un- derstanding, Proc. SPIE, vol. 281, 1981, pp. 11-23.

[72] Andersen et al. 891 S. Andersen, K. Olesen, F. V. Jensen, F. Jensen, HUGIN* a Shell for Building Bayesian Belief Universities for Expert Systems, in Proc. of the 10th Int'l Joint Conference on Arti- ficial Intelligence, 1989.

[73] S. Ayer, P S&meter and J. Bigiln, Segmentation of moving objects by robust motion parameter estimation over multiple frames, in: J.-O. Eklundh, ed., Proceedings Third European Conference on Computer Vision II, Stockholm, Sweden (1994) 316-327.

[74] Y. Liu, D. Zhang, G. Lu, W.Y. Ma, A survey of content-based image retrieval with high-level semantics, Pattern Recognition 40 (11) (2007) 262–282.

[75] K. Aizawa, H. Harashima, and T. Saito, "Model-based analysis syn- thesis image coding (MBASIC) system for a person's face," Signal Processing: Image Commun., vol. 1, no. 2, pp. 139-152, 1989.

[76] N. I. Badler, J. O'Rourke, and H. Tolzis, "A human body modelling system for motion studies," Proc. IEEE,vol. 67, pp. 1397-1403,Oct. 1979.

[77] Ge J, Sahiner B, Hadjiiski LM, et al. Computer aided detection of clusters of micro calcifications on full field digital mammograms. Medical Physics 2006 ; 33(8):2975–88.

[78] Bonabeau E., Dorigo M. and Theraulaz G., “Swarm Intelligence, From Namnl to Artificial Systems,” Oxford University Press, Oxfoxd, 1999.

[79] V.V Anshelerish, A.V Amirikian, A.V Lukashin, and M.D Frank Komentskii “ on the ability of neural networks to perform generalization by induction”, Biol Cybernetics, vol .61 .pp 125-128, 1989.

[80] Glory E, Murphy R. Automated subcellular location determination and high throughput microscopy. Dev Cell 2007;12:7–16.

[81] D.Karaboga, “An idea based on bee swarm for numerical optimization.”Tech. Rep. TR-06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

[82] J. R. Smith and S.-F. Chang, An image and video search engine for the world-wide web, inProc. SPIE Storage and Retrieval for Image and Video Databases, 1997.

[83] Aas, K., Solberg, A., Koren, H., Solberg, R., 1997. Semi-automatic revision of forest maps combining spectral, textural, laser altim-eter and GIS data. Proc. 3rd Int. Airborne Rem. Sens. Conf. andExh., Copenhagen, Denmark, pp. 405–411.

[84] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, “Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex,” in Neural Comput., vol. 2, pp. 293–307, 1990.

[85] Mohamed S, Nyongesa H,”Automatic fingerprint classification system using fuzzy neural techniques”. In: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, Washington, DC, April, 2002.

[86] T.. Aach and A. Kaup,“Bayesian algorithms for adaptive change detec-tion in image sequences using Markov random fields,”Signal Process.Image Commun., vol. 7, pp. 147–160, 1995.

[87] kiran, Talele, and Tuckley Kushal,"Facial Expression Classification Using Support Vector Machine Based On Bidirectional Local Binary Pattern Histogram Feature Descriptor",17th IEEE International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD),2016.

[88] M. Malfait and D. Roose, “Wavelet Based Image De noising Using a Markov Random Field a Priori Model,” IEEE Trans. Image Processing, vol. 6, No. 4, pp.549-565, 1997.

[89] D. Kundur and D. Hatzinakos, “Blind image deconvolution,” Signal Processing Magazine, IEEE, vol. 13, no. 3, pp. 43–64, May 1996.

[90] Hwang, H., and Haddad, R.A. “Adaptive Median Filters: new Algorithm and Results”, Transactions on Image processing, Vol. 4, No. 4 pp.449-505, April 1995.

[91] R. B. Paolo Gay, "Innovative Techniques for Fruit Color Grading," presented at Innovative Techniques for Fruit Color Grading, American Society of Agricultural and Biological Engineers, St. Joseph, Michigan,2002 ASAE Annual Meeting, 2002.

[92] H. Grabner, T. Nguyen, B. Gruber, and H. Bischof, “On-line boosting-based car detection from aerial images,”J. Photogramm. Remote Sens.,vol. 63, no. 3, pp. 382–396, May 2008.

.