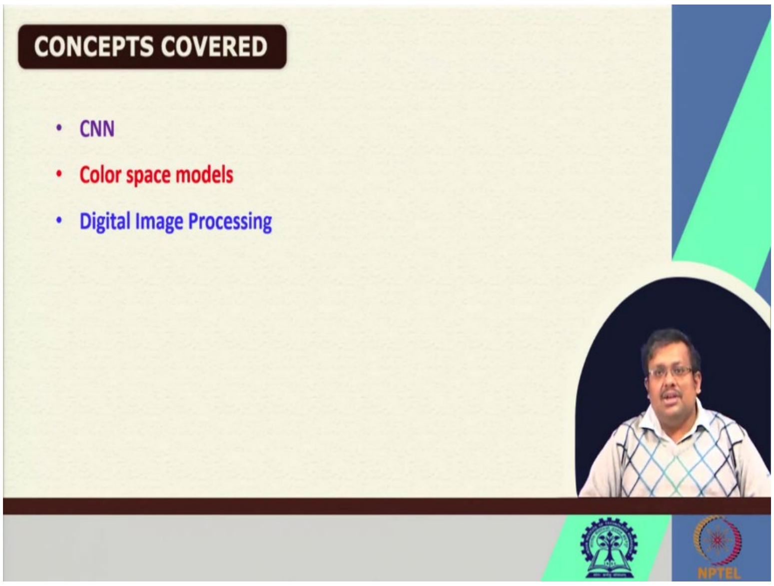
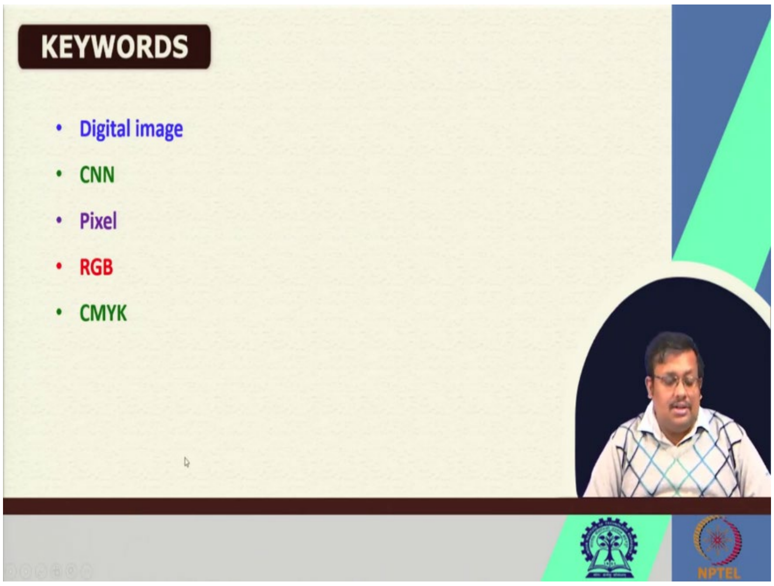
Machine Learning for Soil and Crop Management
Professor Somsubhra Chakraborty
Agricultural and Food Engineering Department
Indian Institute of Technology Kharagpur
Lecture 34
ML and DL for Soil and Crop Image Processing

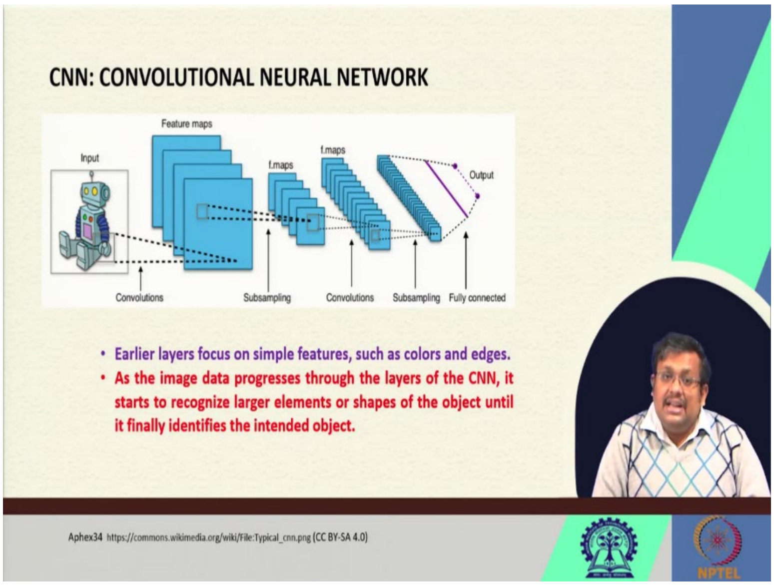
Welcome friends, to this fourth lecture of Week 7. And in this lecture, we are talking about the Machine Learning and Deep Learning for Soil and Crop Image Processing. And in our previous three lectures, we have covered several important regularization, regression methods. We have also discussed the artificial neural networks, and we have discussed their construction, their importance, their advantages and disadvantages. And today, we are going to discuss about the details about convolutional neural network.

In my previous lecture, we have discussed that this convolutional neural network is basically applied for image processing, and you will see that most of the contemporary image processing for soil and crop applications are they are based on the convolutional neural network. Now, before we see those applications, we have to understand what is convolutional neural network and why it is important for processing the images.

(Refer Slide Time: 01:36)







So, in this lecture, we will be talking about this convolutional network neural networks, then we will be also seeing the digital image processing and also different color space models. So, we have to understand all these terms before we see and understand the application of machine learning and images for soil and crop characterization. So, we will start with the convolutional neural network in if you recall, and these are the keywords. Digital image, then convolutional neural network, Pixel, RGB, CMYK these are the some of the keywords for this lecture.

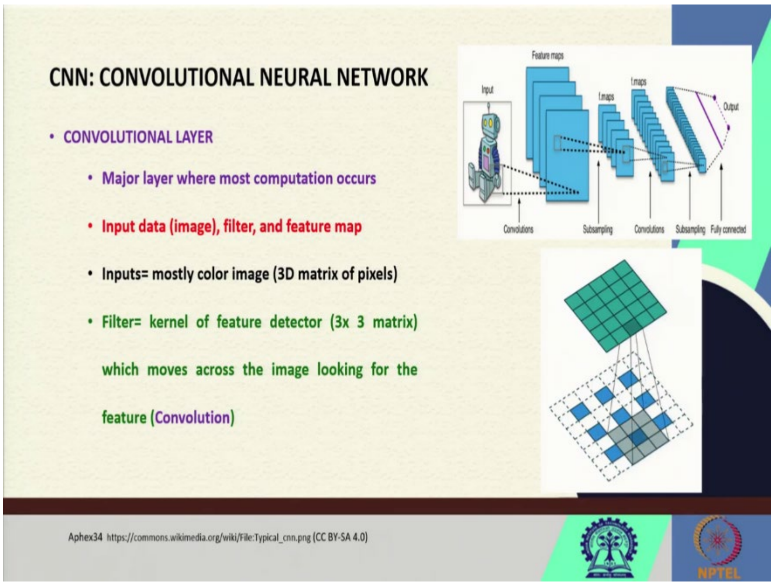
Now, if you see the, if you remember from our last lecture that convolutional neural network basically consists of three layers, one is input layer and or convolutional layer, second is the pooling layer and third one is fully connected layers. So, the convolutional neural network is basically composed of this first convolutional layer, which basically extract the image features from the data from the image features from the original input images, and then the pooling layer will also extract those features.

And then finally, there will be a feedforward a fully connected layer. This fully connected layer will be producing the output. So, we are going to discuss these in details. So, earlier layers, so the first layer generally focus on. So, the first layer in this total construction is basically focus on the simple features such as colors of the image, then ages of the image. Now, as that image data progresses through the layers of convolutional neural network it starts to recognize larger elements or shapes of the objects until it finally identifies the intended object.

Generally, the CNN is useful also for identification of several objects in the image and they are also used for classification. So, for identifying it is a kind of a hierarchical approach, we will see that when the image features process or progress from one layer to the subsequent layer it is a kind of a hierarchical process where we are joining all the bits together in the upper layer, we are going to get the final feature or final picture of the and based on that we are going to classify.

Now, so as we, as this image data progresses through the layers of the convolutional neural network, it starts to recognize larger elements or shapes of the object until it finally identifies the intended objects.

(Refer Slide Time: 04:46)



So, let us see, what are the importance of these three layers. So, first layer is called the convolution layer. So, first layer is called the convolutional layer and this convolutional layer is basically composed of one or more filters. Now, these are the major layers where most computation occurs. So, here the input data is image and also it is also composed of the filter and feature map.

So, all these three consist or are the features are the component of this convolutional layer. So, again the convolutional layer is the major layer where most of the computation work and it is composed of the input data or image and then filter and the feature map. So, where are the what are the inputs, inputs are the mostly color image let us consider this in color image which is basically 3D matrix of pixel. What are the pixels? What is the 3D matrix?

Basically, remembered that any pixel is denoted by x, y and their color brightness or color intensity. So, there are three features x, y and z. So, x and y are basically the length and the width of the pixel and z basically is the value of the color or the intensity of the color. So, you know, that the image any image can be considered as a 3D matrix of pixels. Pixels are the smallest feature in an image.

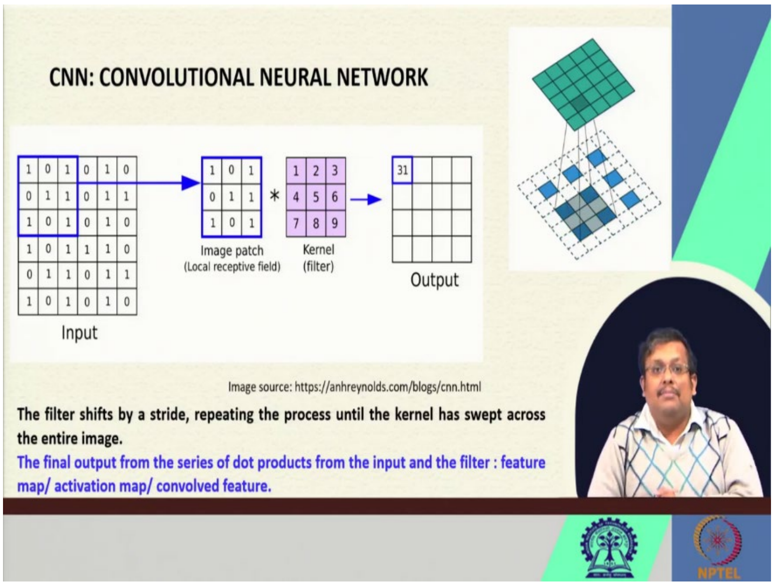
So, here if you see this is an original image. So, if we the smallest feature in this image is called the pixel or the full name is picture elements. So, we will be discussing this in detail in our upcoming slides. So, these inputs are the color image and then the filter. What are the filters? Filters are basically are known also known as kernel or feature detector. Generally they are 3 by 3 matrix, but the size varies which moves across the image looking for the feature, and this process is known as convolution.

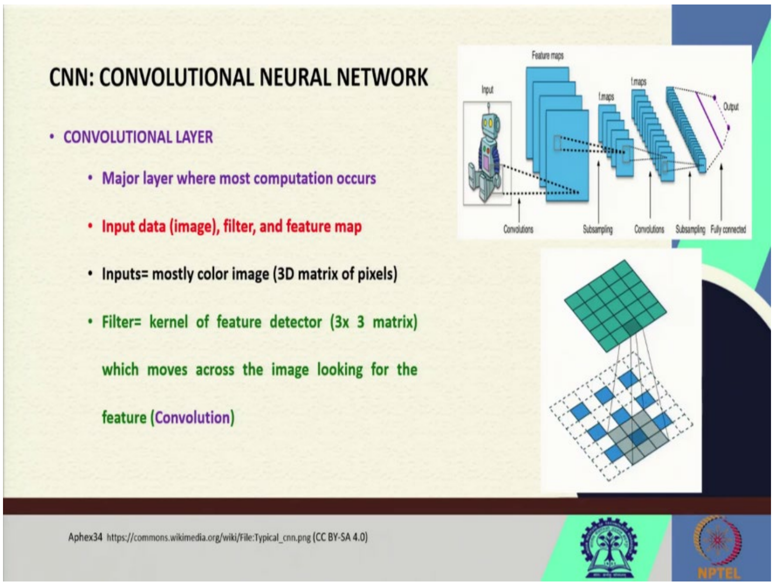
So, here you can see that a 3 by 3 matrix of feature detector or kernel is moving through the receptive part of the image not it is covering the whole image at a glance, but it is moving through the image. And as it is moving through the image, they are computing the dot products of these image pixels and the weights which are given in this 3 by 3 feature detector or kernel.

So, ultimately the idea is when this filter is moving throughout the image it will basically compute in a dot compute the dot products of the weights which are assigned to these kernel as with the values of the inputs and then they will output will be a final quantitative value. So, these quantitative values will be considered. So, these quantitative values will be considered for all these pixels, so this will be known as the feature map.

So, this whole process of moving this kernel for looking or identifying the feature. So, when this kernel moves, it tries to identify a feature based on this dot product and ultimately computation of the final output and this process is known as convolution process.

(Refer Slide Time: 08:56)





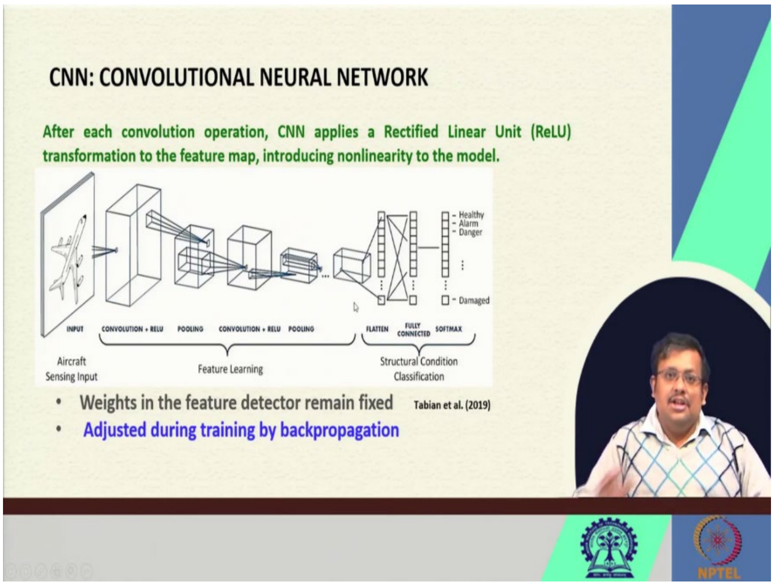
So, here you can see suppose, this is an input and it is an input image and this is the 3 by 3 kernel which we are moving. So, once we are moving these 3 by 3 kernels. So, suppose, these image patch is there, so suppose, this is an image patch, which is 3 by 3 and the kernel or filter which is supposed this shaded area is the kernel which is moving here. So, you can see. So, this shaded area is having a different weights, and also this is the image patch.

So, if we give the dot product so, you can see here 1 is to 1 plus 0 sorry, 1 multiplied by 1 plus 0 dot 2 plus 1 dot 3 plus 4 dot 0 plus 1 dot 5 plus 6 dot 1 plus 7 dot 1 plus 8 dot 0 plus 9 dot 1 you will get the final value of 31. So, this 31 is the final output. So, that is why when this kernel moves through the whole image it will calculate in this session this final output, which will be a multiplication or dot product of the image patch as well as the kernel weight. So, ultimately this will be the output.

So, the filter shifts by a stride. So, as you can see the filters or kernel is moving the repeating the process until the kernel swept across the entire image. So, in this way it moves from one patch of the image to another patch of the image and it will cover the whole image. So, the final output from the series of dot products from the inputs final outputs is from the series of dot products from the inputs and this kernel filters. So, they are known as feature space or feature map or activation map or convolved feature. So, this is the output which is known as the feature map.

So, in this construction you can see this is the feature map this is the input and convolutional neural. A convolution layer decides in between. So, I hope you are clear now, how this convolution process goes on and these what is the function of this convolutional layer.

(Refer Slide Time: 11:17)

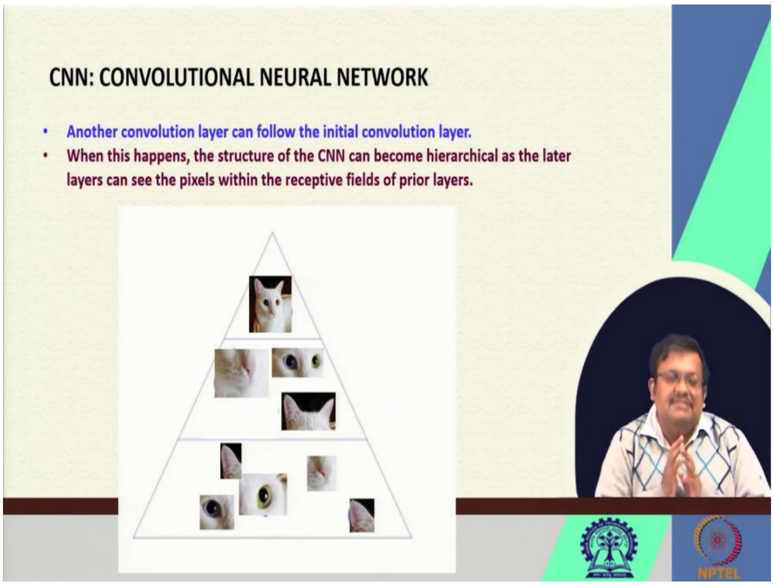


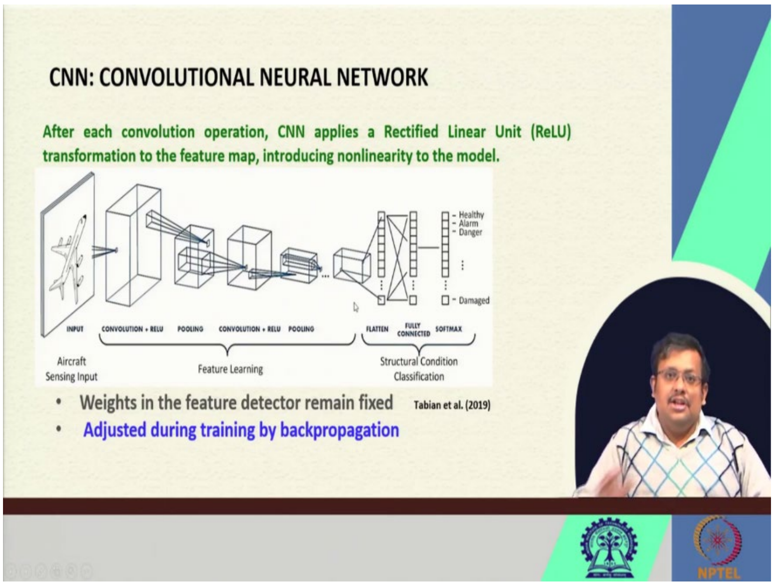
So, if you see, so, after each convolution operation, the CNN applies a rectified linear unit, which is also known as RELU transformation to the feature map introducing the non-linearity to the model. In the previous lecture, I told you that in the real world this feature generally related to each other in a non-linear way. So, to add this, this non-linear feature, after each of this convolution operation, this CNN applies a rectified or ReLU transformation to this feature map.

So, once this feature map is being calculated then it applies these non-linear transformation to this feature map introducing this non-linearity to the model. So, weights in this feature detectors remains fixed, but and also that they are adjusted during the training process. Remember, where we are using this kernel or feature the filters, these the weights of these filters or kernel are basically fixed, but they are adjusted during the training process by backpropagation.

I have already told you what is backpropagation in our previous lecture. So, remember how it happens. So, you can see here this is the input image and from this input image we are getting the convolved convolution. After the convolution we are getting the feature map and that feature map is also being we have also introduced the ReLU transformation to introduce the non-linearity to mimic the real world situation. So, this is the first step.

(Refer Slide Time: 13:06)





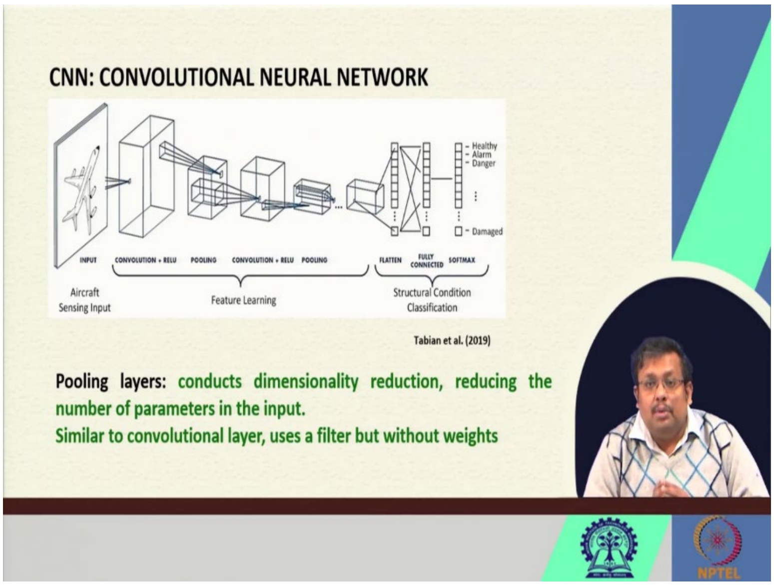
So, you see here another convolutional neural network layer can follow the initial convolutional neural network layer. So, you can see this convolution neural convolutional layer is followed by another convolution layer that followed by another convolutional neural network layer. So, it repeats multiple times, until we get this fully connected layer.

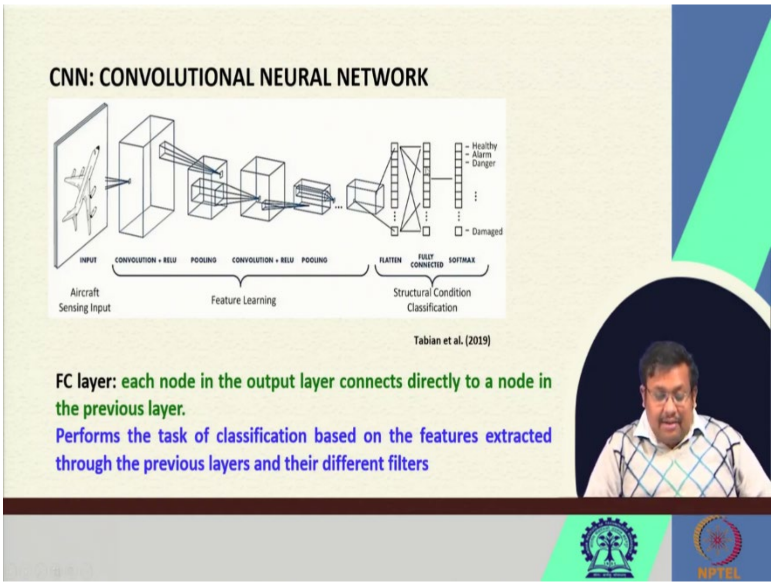
So, when this happens, the structure of this convolutional neural layer can become hierarchical as the later layers can see the pixel within the receptive fields of the prior layers. So, basically, you can consider this is a hierarchical identification. As this image moves through the features of this image moves through different layers ultimately, at the end, these convolutional neural network will identify the object.

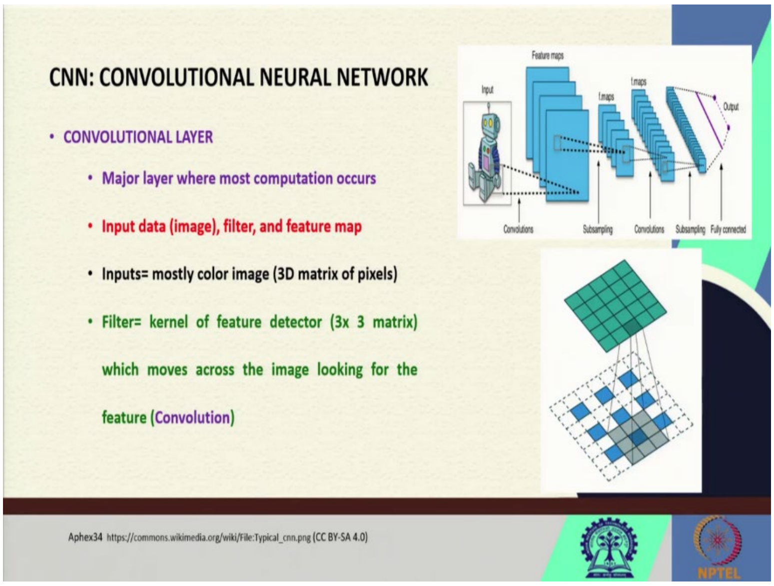
For example, in the lower layers in the initial layers, it can identify different parts of the image of a cat. So, here you can see the image of a cat they have different parts. So, in the lower layer, they are identifying these, in the subsequent layer they are joining together and in the upper layers they will identify the final feature and they will identify the object in this case it is a cat picture.

So, when this happens, the structure of the convolutional neural network can become hierarchical as the later layers can see the pixels within the receptive field of the prior layers. So, this is how they can identify these images or objects in an image.

(Refer Slide Time: 14:46)







So, we have seen the convolutional layer, we have also known what is ReLU function. Now, let us see what is the second important component of CNN, that is pooling layers. So, pooling layers are basically used for dimensionality reduction. When you have huge amount of data, you need to reduce the numbers of the parameters of the inputs, so that you can easily deal with the data.

So, it is a similar to convolution layer, but it uses a filter, but without any weight. Remember, in case of convolutional filter it uses some predefined weights, which are being calculated during the training process through backpropagation. However, in case of pooling layers, they do not have any weight assigned. So, the third layer or the final layer is called a fully connected layer as you can see in this picture. So, this fully connected layer here, these each node in the output layer connects directly to a node in the previous layer.

So, here you can see they are being fully connected. This picture is also very helpful. So, you can see the ultimately the fully connected layer individual they will be these features will be considered as individual nodes and they will be connected from to each other and this is called the fully connected layers.

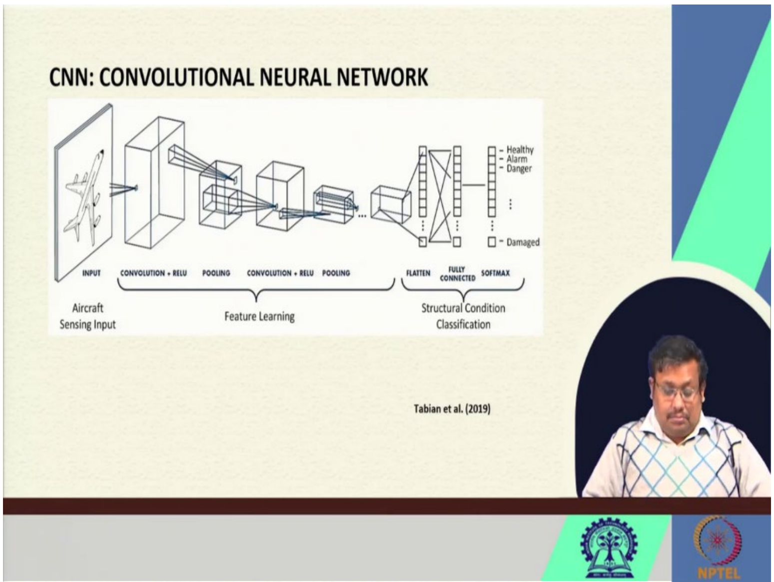
So, this fully connected layers performs the tasks of classification based on the features, which are being extracted to the previous layers and their different filters. So, suppose we have started with the image of inputs and then we have undergone through feature learning and this feature learning process is nothing but the combination of convolutional layer, as well as the pooling layers.

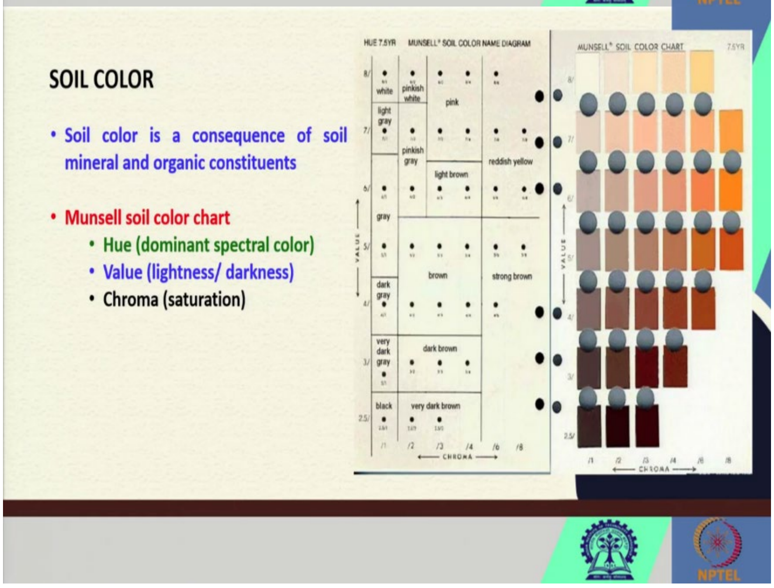
So, in the convolution layer, we have the filters, the filters are having the predefined weights calculated by during the training by backpropagation, and use and these moving filters give, the final output, which is basically the series of dot products also known as the featured map. This feature map is also introduced with some amount of non-linearity by these ReLU layer or transformation, then it moves to the pooling layer.

Ultimately, the importance of the pooling layer is to remove or dimensionality reduction to reduce the number of inputs, ultimately, we will be getting the final layer which is called a fully connected layer, which will go for the actual classification based on these softmax criteria.

Now, this is if you see the three major conditions, three major feature, three major sets of this whole convolutional neural network you can see there are input images and also feature learning and finally there will be structural condition classification. So, this is how these convolutional neural network works, and this convolutional neural network is very, very helpful for object identification and prediction based on the images.

(Refer Slide Time: 18:14)





So, now, we have seen that is, this is called a convolutional neural network in a, we can identify different features, we can identify different objects. So, let us start with the application of this color and the image features for soil property identification. So, far we have learned different types of machine learning and deep learning approaches. Remember these artificial neural network or convolutional neural network, they are very, very powerful deep learning methods and there are multiple application for both soil and crop which we are going to see.

Now remember that in case of soil, soil color is a consequence of soil mineral and organic constituents. So, based on the variation of organic matter or based on the variation of the textural features, you can see the, you can see the color of the soil varies.

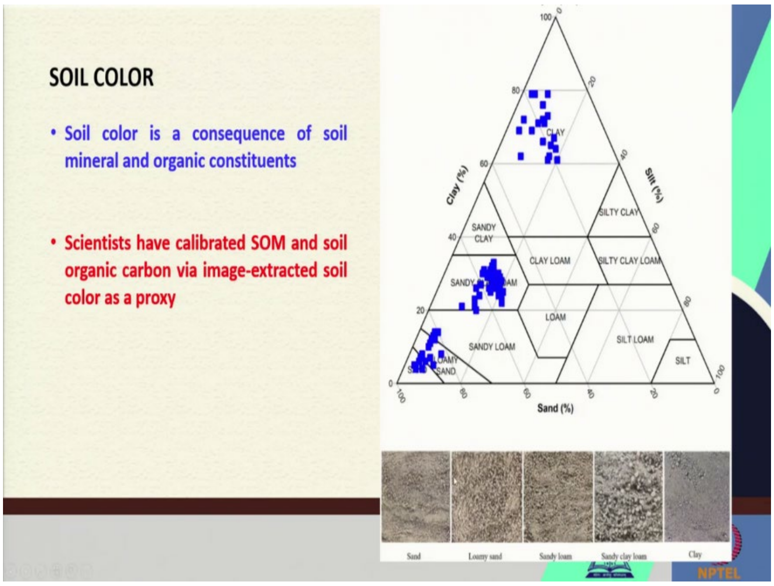
So, generally how we define the soil color? The soil color is being defined by these Munsell cells soil color chart. So, this Munsell this is a Munsell soil colored chart as you can see, in this Munsell soil colored chart, there are three major components one is called hue, which is determined by the which is basically indicating the dominant spectral color, then the value which shows the brightness that means lightness or darkness and finally, the Chroma also known as the saturation.

So, you can see here, here in this case the hue is the dominant spectral color and then values generally the lower values means, it is dark higher values means, it is lighter in color. And as you can see the increasing the purity from right to left where the colors are getting more and more saturated and more so, this saturation is the value of the Chroma is less when there is high saturation, and where it is low, there will be higher values of the Chroma.

So, based on these we generally define the soil color. Suppose, the hue is 7.5 are in this case and suppose we are taking on soil samples and we see that the soil color matches with these color chip, with this color chip. So, here we can see, the value is 5 and the Chroma is 1, so we define the soil color as 7.5 YR 5 slash 1. So, this is how we define the soil color.

Remember here the wire stands for yellow red, so there are several hues in this soil color book also. So, this is how qualitatively we define the soil color and scientist have calibrated the soil organic matter and soil organic carbon via image extracted soil color as a proxy.

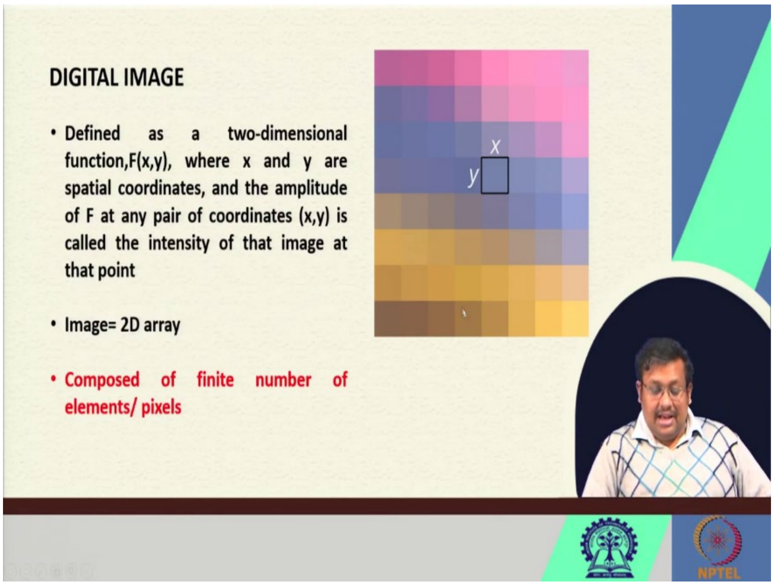
(Refer Slide Time: 21:40)

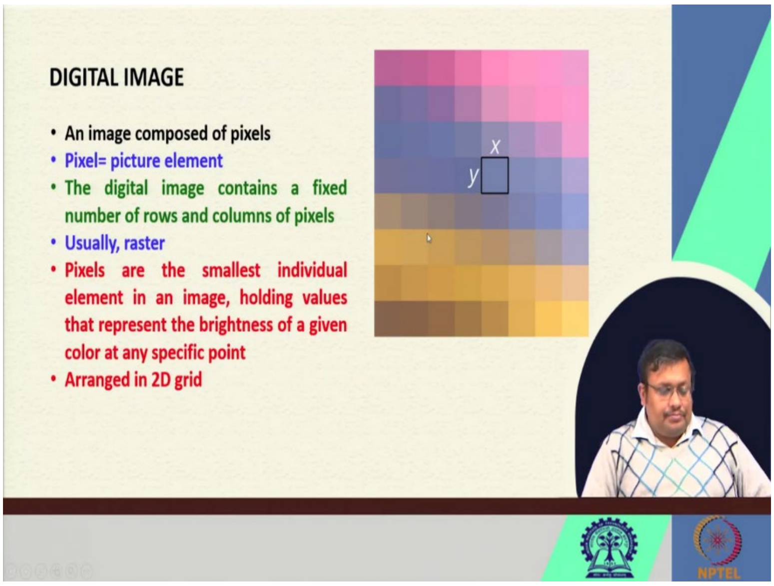


As you can see here one example here depending on the variation of the soil texture, it is clay, this is sandy clay loam, this is sandy loam, this is loamy sand, this is sand. So, based on the variation of this clay and sand, so here clays dominant, clay is darker in nature, however, sand is lighter in nature. So, you can clearly see the difference in this in these images.

So, basically these images are indicating the changes in the soil texture and it is possible to predict the soil texture, and also in similar way we can also predict the soil organic carbon or soil organic matter. I will show you some examples, but you see that color of the soil the major importance of this slide just to show you that color tells us so many things. So, using this color as a proxy we can classify any soil as a sandy soil or clay soil.

(Refer Slide Time: 22:51)





So, before we go for any type of further discussion on soil image analysis or digital image analysis, let us first discuss what is digital image and what is digital image processing? So, digital image is basically defined as a two-dimensional function which is denoted by Fxy. So, where x and y are the spatial coordinates and the amplitude of F are at any pair of the coordinate is called the intensity of that image at the point.

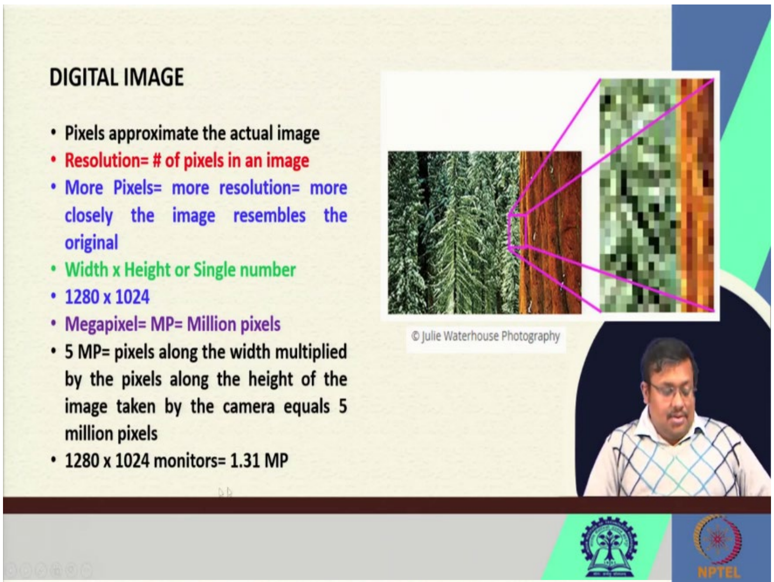
So, you can see here it is defined as the two-dimensional function. So, you can see here these x and y are the spatial coordinates and so these pair of coordinates and it has the any value, which is known as the intensity of that image of that point. So, image is of course, a 2D two-dimensional array and consists of finite numbers of elements or pixels.

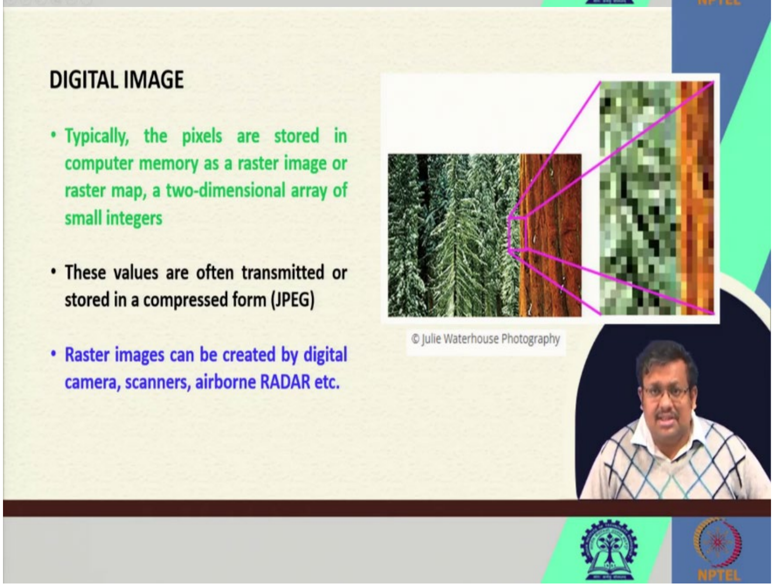
So, here these are individual pixels, I hope you all have seen these types of pixels when you try to zoom an image. So, you try to zoom in one image. So, an image is always composed of multiple pixels, pixels is the short form of picture element and the digital image contains a fixed number of rows and columns of pixels.

And you can see these are continuous and these continuous type of data is known as the raster data. We will discuss more about these raster in our digital soil mapping lectures, in our in the coming weeks. So, remember these most of the time these digital image are generally in the raster format, and pixels are the, what is the definition of the pixel. So, pixels are the smallest individual element in any image holding values that represents the brightness and or of a given color at any specific point.

So, of course, not only their coordinates are important, but also their brightness is also important. So, based on this changing brightness also we can assign some values. So, remember they are as arranged in obviously, in that 2D grids or two-dimensional grids.

(Refer Slide Time: 25:33)





Now, remember that a pixels approximates the actual image, and more number of pixel to define an area of the image or in other words, if an image is having more number of pixels, then their resolution is also increasing. So, resolution depends on the number of pixels in an image. More pixels means more resolution more closely the image resembles the original. So, generally you can define the resolution by width multiplied by height or a single number.

Similarly, you can see some time people are using these 1280 by 1024 pixels as a resolution 1028 multiplied by 1024 pixels. So, that means, that image has 1280 pixels in the width and

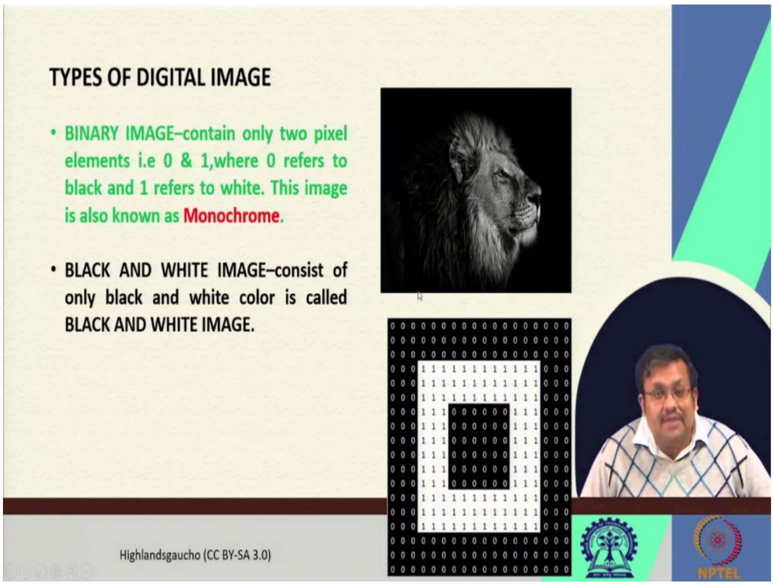
1024 pixels length wise or height wise. So, this is one way of showing the resolution. Another way is showing through megapixel. So, these megapixels are known as the 1 megapixel is basically 1 million pixels.

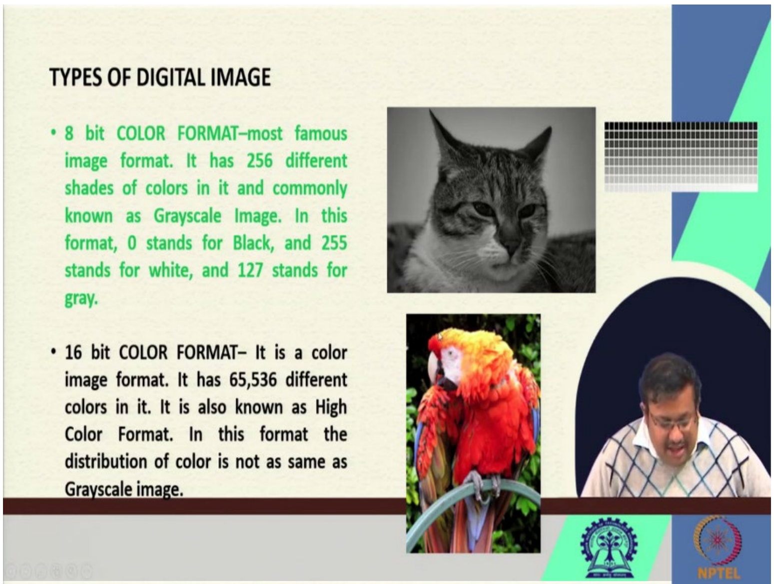
So, 5 megapixels when we talk about 5 megapixel camera, 10 megapixel camera, so, this 5 megapixel denote that pixels along the width multiplied by the pixels along the height of the image taken by this camera equals to 5 million pixels. So, similarly, if we, for this 1280 by 1024 resolution also we can see in the monitors we can see if we multiply it with it, this will be 1.31 megapixels. So, this is how these image resolution is being calculated.

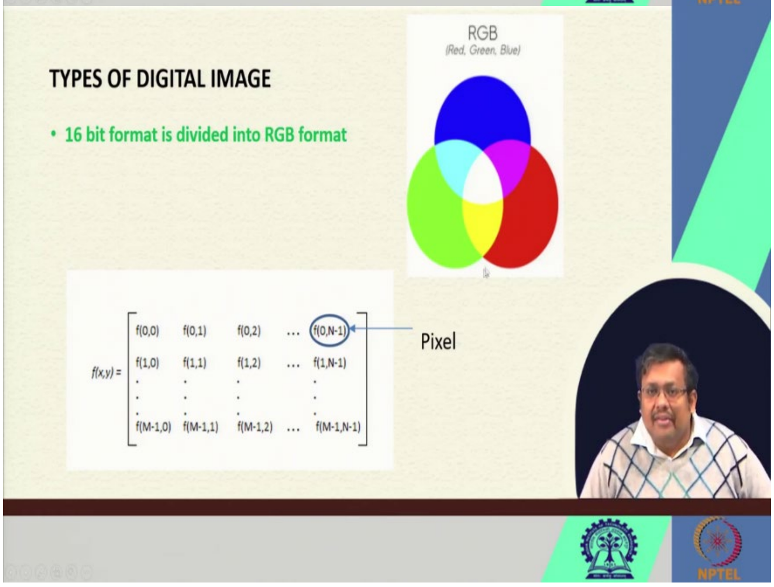
So, typically the pixels are stored in computer memory, as raster image or raster map a two-dimensional array of small integers, and these values are often transmitted or stored as compressed form. So, generally, they are quite large in volume, so generally they are stored in a compressed form one of the compressed form is this JPEG image, you all know about JPEG image, which is Joint Photographic Experts Group, the full name.

So, the raster images can be created by different devices like cameras, scanners, airborne radar, etc. So, these digital images are very much important for digital image processing because they are the one on which we do these digital image processing.

(Refer Slide Time: 28:26)







Now, let us see what are the types of the digital image. There are some binary image binary image are consist of two pixel element that is either 0 or 1. As you can see, this is the binary image where 0 refers to the black and 1 refers to white. So, this image is also known as monochrome.

And the second another image type is black and white image which is consist of only black and white color, and it is also known as the black and white image. Also, there is 8 bit color format, which is the most famous image format, and it has 256 different shades of color in it added commonly known as the grayscale image. As you can see it is a Grayscale image in this format 0 stands with a black pure black and 127 stands for gray and 255 stands for pure white.

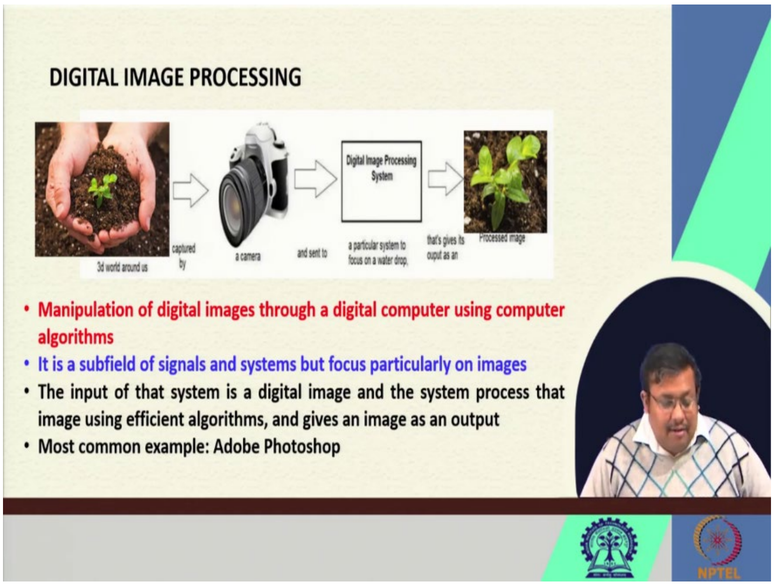
So, and in between there are different grades differentiates so based on that they can assume these pixels can assume any value between 0 to 255. And based on their values, they will be color coded, and when we combine to see them together, we will get this type of final image.

Third another one is 16-bit color format. So, it is a color image format, as you can see here, it has 65,536 different colors in it. It is also known as the high color format. In this format, the distribution of the color is not same as the grayscale image.

So, if you see the type of the digital image of course, the 16-bit format is further divided into RGB. Any color image can be divided into RGB data. So, you can see mathematically in the matrix form we can define the color as an image as this. It is a combination of pixels. So, here these are individual pixels, individual pixels have some values. So, this is basically the mathematical matrix notation of a digital image.

Now, remember any color image or system-based image can be divided into RGB format. They are called additive color models, because, if we mix these RGB one by one together, we can get other colors also. So, that is why it is called the additive color model.

(Refer Slide Time: 30:56)



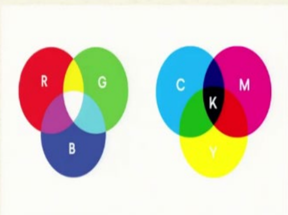
DIGITAL IMAGE PROCESSING

- In the above figure, an image has been captured by a camera and has been sent to a digital system to remove all the other details, and just focus on the plant and soil by zooming it in such a way that the quality of the image remains the same.

STEPS OF DIGITAL IMAGE PROCESSING

1. Import the image via image acquisition tools/devices
2. Analyze and manipulate the image
3. Produce altered image or extract image features

DIFFERENT COLOR MODELS



RGB color mode is best for digital work, while CMYK is used for print products

Now, what is digital image processing? This digital image processing is basically the manipulation of digital images through a digital computer using computer algorithm. So, here you can see here, we are taking the 3D, we are actually using a camera to capture the image of the real 3D world around us, and saying to this digital image processing system to finally get the processed image. So, it is a subfield of the signals and systems but focus particularly on images.

So, input of the system is a digital image and the system processes that image using different computer algorithms and gives the image an output. So, this is an output image and these output images produced by different types of computer algorithms.

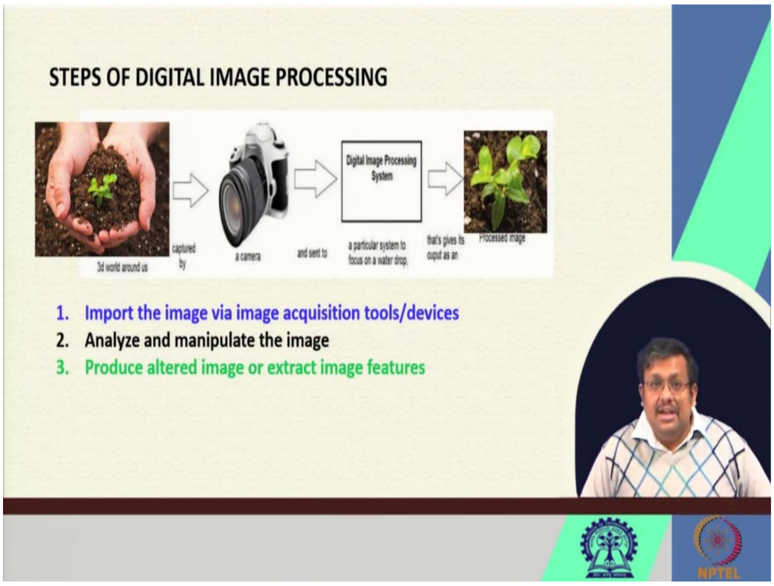
Most common example of this digital image processing is Adobe Photoshop. So, you all have used this Adobe Photoshop and these Adobe Photoshop use this digital image processing technology for their performance.

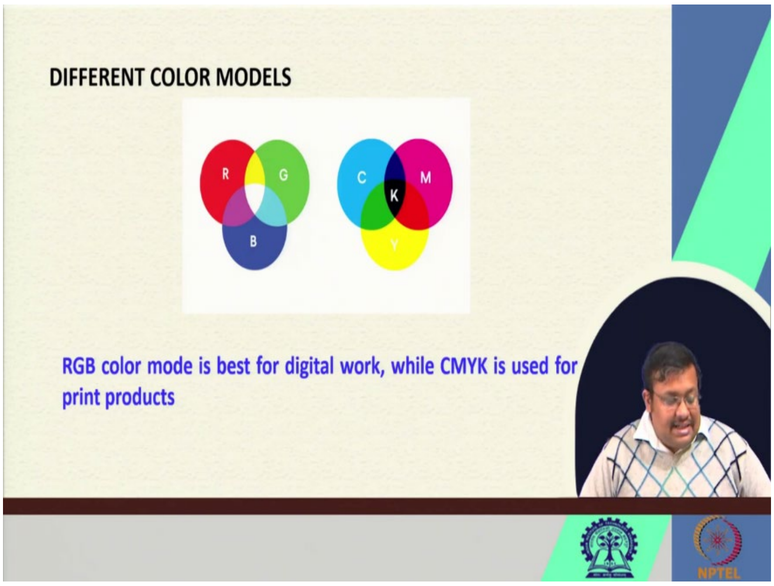
Now, in the above figure you can see an image has been captured by an camera and has been sent to a digital system to remove all the other details and just focus on the plant and soil by zooming it in such a way that the quality of the image remains same. So, this is one example of digital image processing.

Now, what are the steps of the digital image processing, there are three major steps one is importing the image using different tools. Secondly, analyzing the or and manipulating the image and thirdly producing altered image or extracted image features. So, these three are the major steps of a digital image processing.

So, our time is limited. So let us wrap up our lecture here. In our next lecture we will start from here and we will see what are the major steps of digital image processing, and we will also see the other color models like RGB models. We have already discussed. We will discuss the CMYK models, HSB models and so on so forth.

(Refer Slide Time: 33:20)





So, let us wrap up our lecture here. And this is the reference for this lecture. Tabian et al in 2019 and let us wrap up our lecture here. In the next lecture we will start from where we left, we will start with the phases of digital image processing and then we will discuss different color models, and then, I will show you how these digital image processing is helpful for identification and prediction of different features in soil and crop.

So, I hope that you have gained some good knowledges in this lecture. You have of course, this digital image processing it is an entire discipline and it requires years and years of experience, there are so many terms. Of course, it is not possible to cover all these in a single lecture, but I think I am being able to give you some basic overview of convolutional neural network, artificial neural network as well as some basics of digital images.

We will be discussing more about this digital image in our coming lectures and their processing and their application for soil and crop images. Thank you, let us meet in our next lecture.