A Survey on Image Classification

SIRG
Scientific Innovation Research Group

Lecture 4

Computer Vision: Algorithms and Applications

Supervised by

Dr. Ahmed Elngar
Faculty of Computers and Artificial Intelligence
Beni-Suef University

Our Team

Ahmed Elngar

Founder and Chair of (SIRG) (Egypt)

Omar Mahmoud

Ammar Fathy

Mohamed Shaban

Basma Mostafa

Outline


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG GROUP

Syllabus Queue

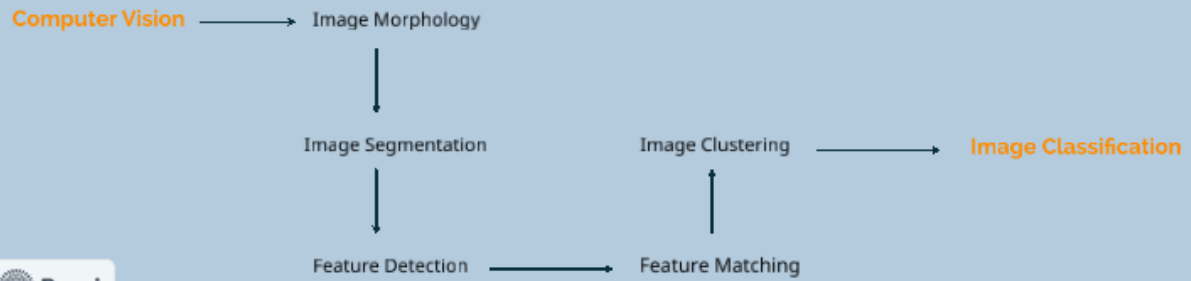
What I'll cover

- Image Classification Concepts
- Image Classification Approach
- Comparison between types
- Image Classification Algorithms
- Classification Process
- Practical part of our topic (Code)

Intro

Intro to Image Classification:

After we went through some stages of the computer vision, we have reached what our topic today is, which Image Classification



What is Image Classification (concepts)

- Classification between the objects is easy task for humans but it has proved to be complex problem for machines.
- The raise of high capacity computers, the availability of high quality and low priced video cameras and has generated an interest in object classification. Image classification refers to the task of extracting information classes from a multi-band **raster** image.

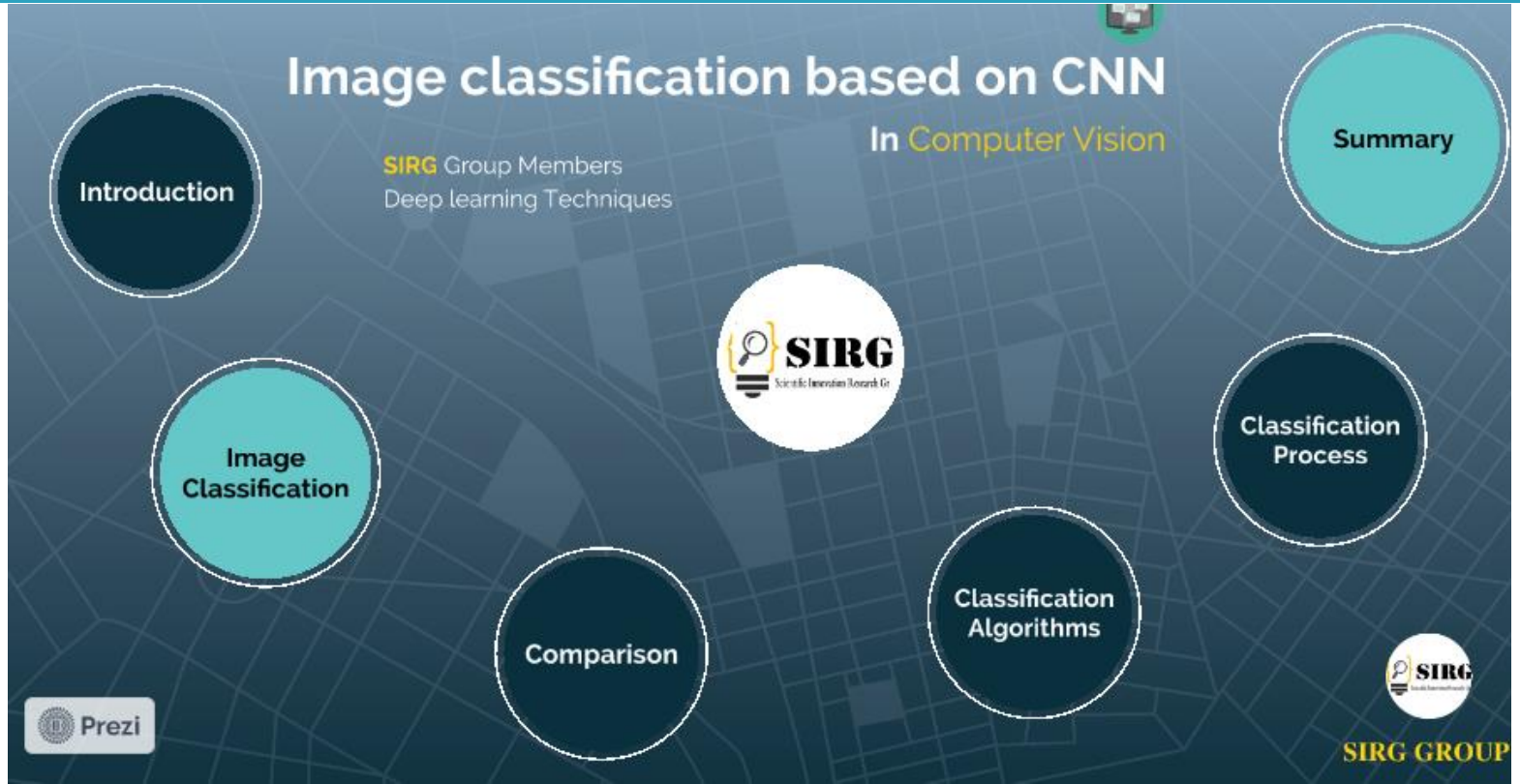


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG GROUP

Classification Approaches

Image Classification

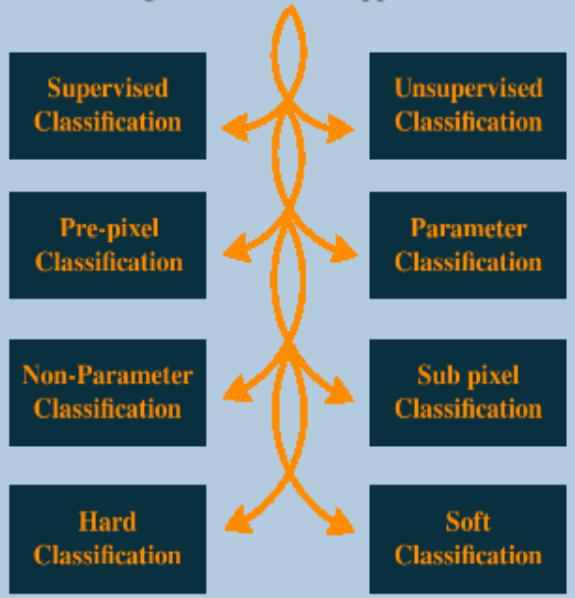
- Image classification refers to the task of extracting information classes from an image. or process of sorting pixels into a finite number of individual classes, or categories of data, based on their spectral response (the measured brightness of a pixel across the image bands, as reflected by the pixel's spectral signature)
- Depending on the interaction between the analyst and the computer during classification, there are two general approaches of pixel-based Image classification: supervised and unsupervised.

Mechanism

Image Classification Approaches

- there are eight general approaches of pixel-based Image classification: and we will take deeply about supervised and unsupervised classification

Image Classification Approaches



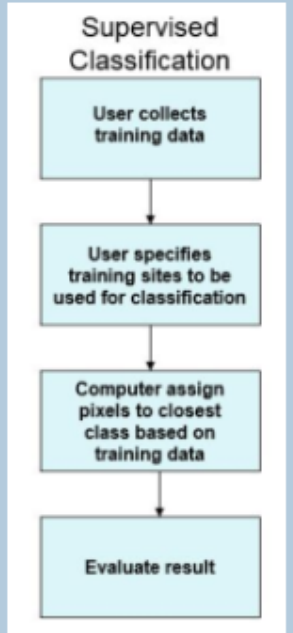
Supervised Classification

Unsupervised Classification

Supervised Classification :

- Supervised classification uses image pixels representing regions of known, homogenous surface composition (training areas) or uses the spectral signatures obtained from training samples to classify an image

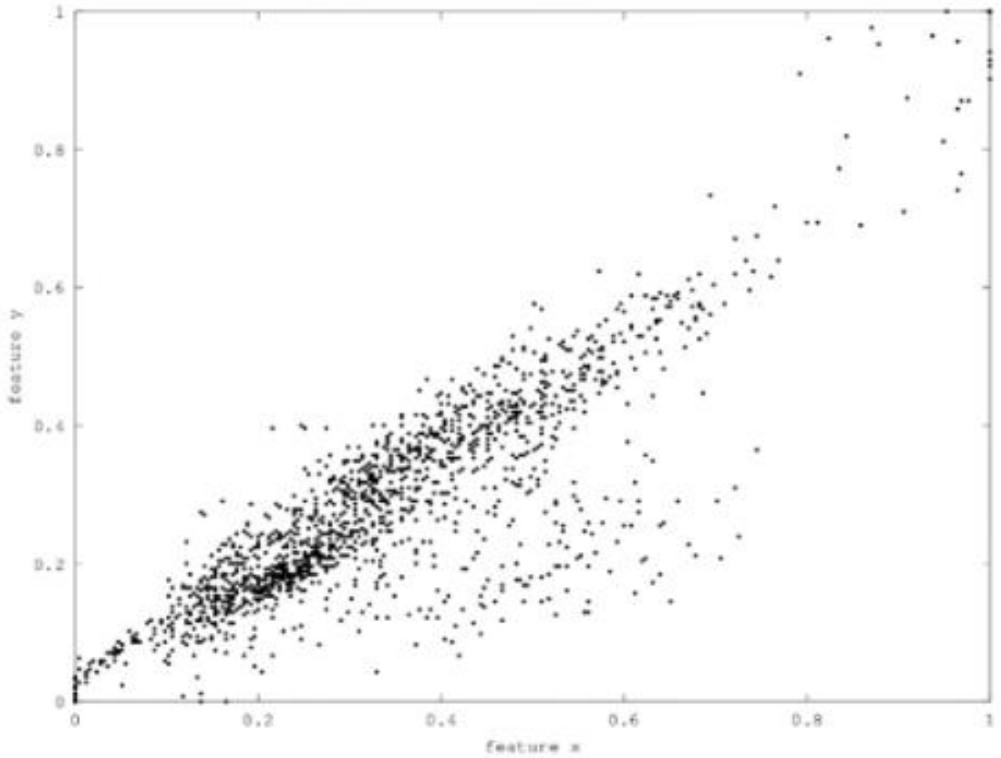
- Divide image into **user classes** defined by user (grass, corn ,water)
- User provides **training data** for each class



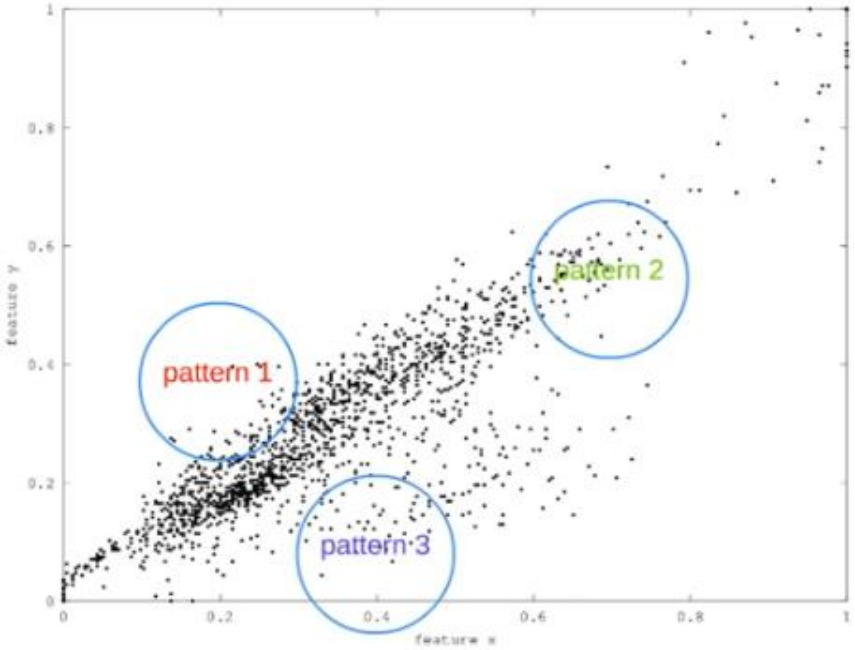
In ENVI there are four different classification algorithms you can choose from in the supervised classification procedure. There are as follows:

1- Maximum Likelihood
2- Minimum Distance
3- Mahalanobis Distance
4- Spectral Angle Mapper

Supervised Classification



Supervised Classification



Priori information Needed

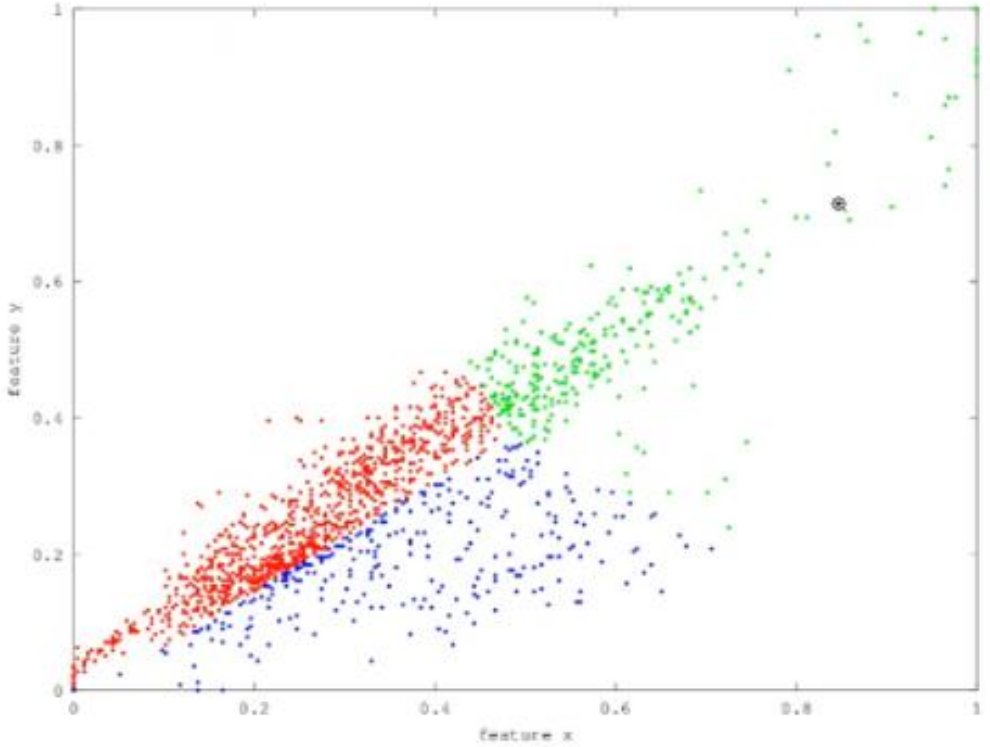
Training classification 1 .

Training classification 1 .

Training classification 1 .

Unclassified .

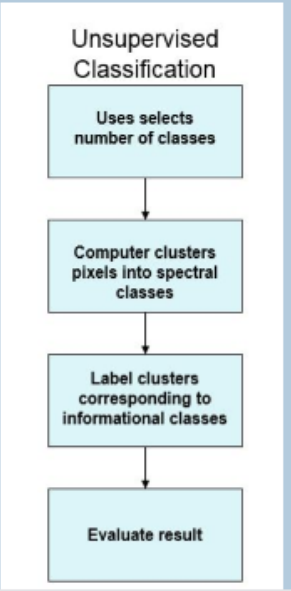
Supervised Classification



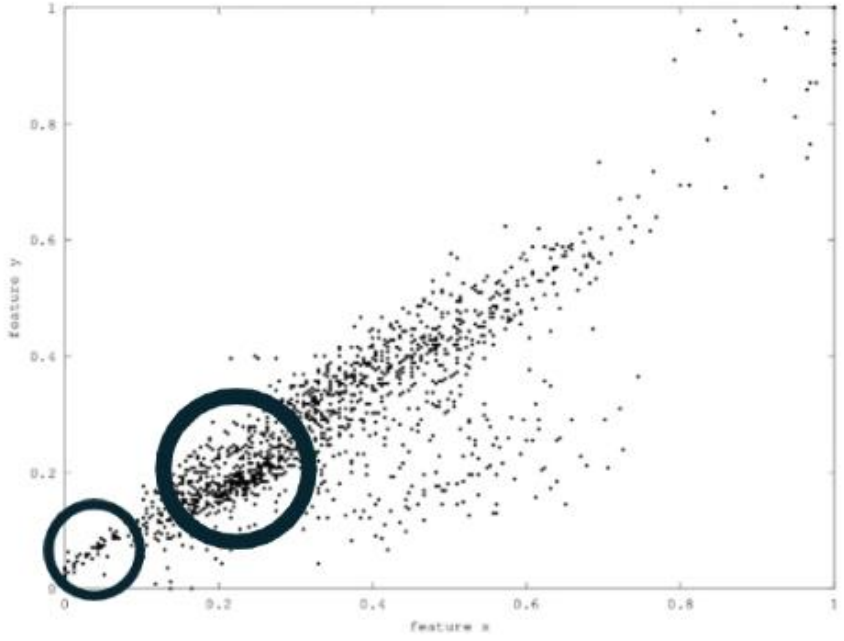
Unsupervised Classification :

- Unsupervised classification is a method which examines a large number of unknown pixels and divides into a number of classed based on natural groupings present in the image values. unlike supervised classification, unsupervised classification does not require analyst-specified training data. The basic premise is that values within a given cover type should be close together in the measurement space (i.e. have similar gray levels), whereas data in different classes should be comparatively well separated

- Divide image into **spectral classes** defined by user (grass, corn ,water)
- No **training data** is provided as based on clustering



Unsupervised Classification



Priori information Unavailable

How Image Classification Works

Image classification is a supervised learning problem: define a set of target classes (objects to identify in images), and train a model to recognize them using labeled example photos. Early computer vision models relied on raw pixel data as the input to the model. However, as shown in Figure 2, raw pixel data alone doesn't provide a sufficiently stable representation to encompass the myriad variations of an object as captured in an image. The position of the object, background behind the object, ambient lighting, camera angle, and camera focus all can produce fluctuation in raw pixel data; these differences are significant enough that they cannot be corrected for by taking weighted averages of pixel RGB values.

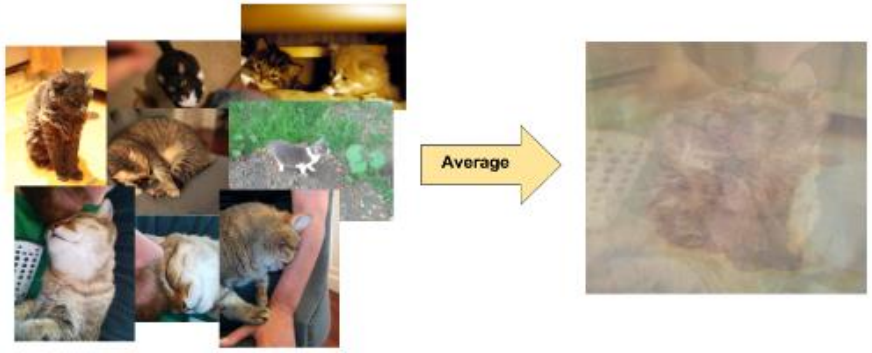



Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG
SIRG GROUP

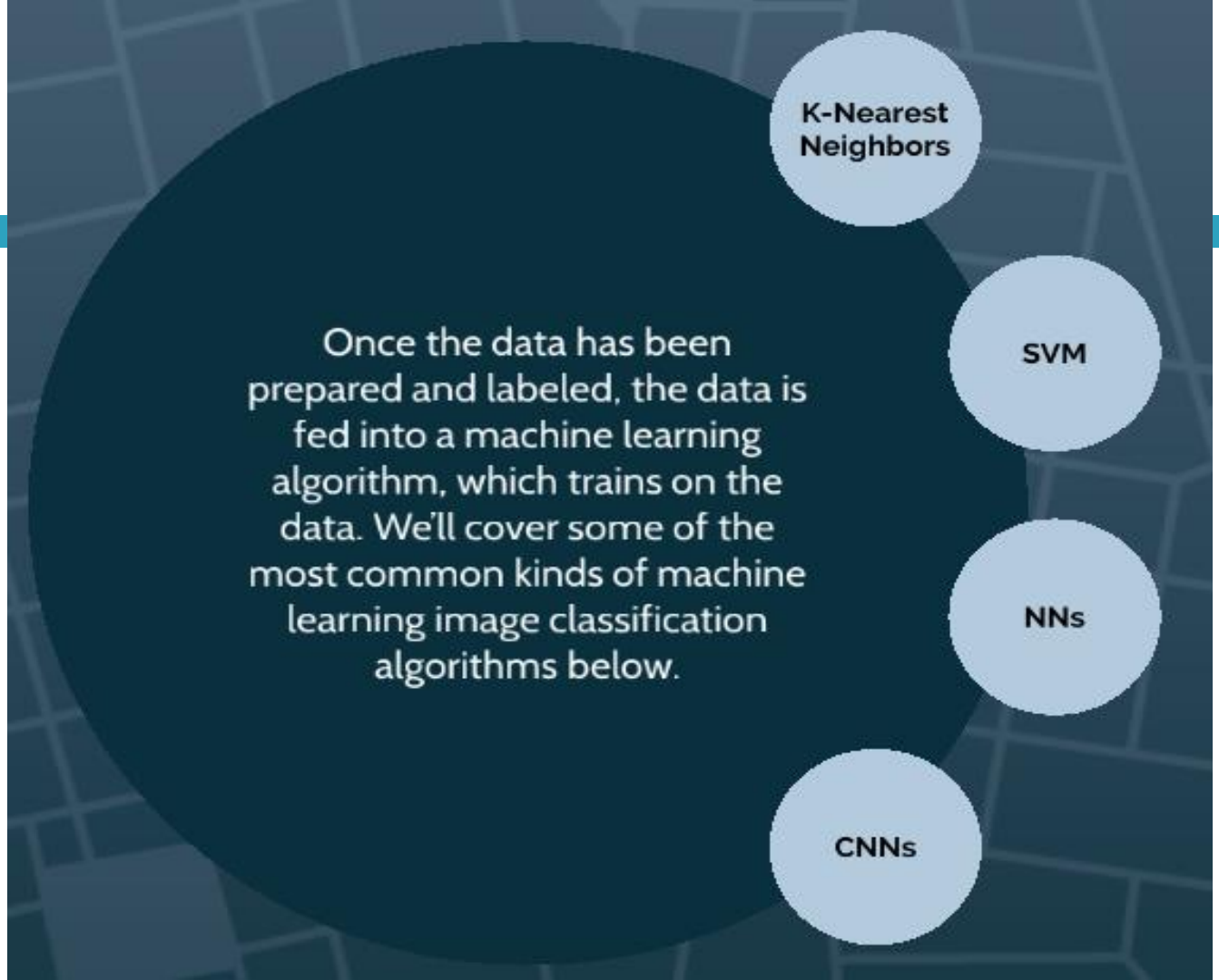
Unsupervised vs. Supervised Classification

Vs

- Unsupervised: bulk of analyst's work comes after the classification process
- Supervised: bulk of analyst's work comes before the classification process

	Supervised	Unsupervised
Advantages	▪ No require prior knowledge of the image area ▪ Human error is minimized ▪ produce unique spectral classes ▪ Relatively fast and easy to perform	▪ Generates informational classes representing features on the ground ▪ Training areas are reusable (assuming they do not change)
Disadvantages	▪ Spectral classes do not represent features on the ground ▪ Does not consider spatial relationships in the data ▪ Can be very time consuming to interpret spectral classes ▪ Spectral properties vary over time, across images	▪ Information classes may not match spectral classes ▪ (e.g., a supervised classification of "forest" may mask the unique spectral properties of pine and oak stands that comprise that forest) ▪ Homogeneity of information classes varies ▪ Difficulty and cost of selecting training sites ▪ Training areas may not encompass unique spectral classes


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG GROUP


Once the data has been prepared and labeled, the data is fed into a machine learning algorithm, which trains on the data. We'll cover some of the most common kinds of machine learning image classification algorithms below.
K-Nearest Neighbors
SVM
NNs
CNNs

K-Nearest Neighbors

- K-Nearest Neighbors is a classification algorithm that examines the closest training examples and looks at their labels to ascertain the most probable label for a given test example. When it comes to image classification using KNN, the feature vectors and labels of the training images are stored and just the feature vector is passed into the algorithm during testing. The training and testing feature vectors are then compared against each other for similarity.

- KNN-based classification algorithms are extremely simple and they deal with multiple classes quite easily. However, KNN calculates similarity based on all features equally. This means that it can be prone to misclassification when provided with images where only a subset of the features is important for the classification of the image.

KNN - Different names

-K-nearest Neighbors
-memory -based reasoning
-Example -based reasoning
-instance-Based Learning
-Lazy learning

What is the KNN?!

-its a powerful classification algorithm
-it stores all available cases and classifies new cases based on similarity measures such as distance function (1-Euclidean 2-Manhattan (City Block) 3-Chessboard)

$$\text{dist}((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2}$$

$$\text{distance} = |x1 - x2| + |v1 - v2|$$

$$D_{Chess} = \max(|x_2 - x_1|, |y_2 - y_1|)$$

-a non parametric lazy learning algorithm

What KNN require ?

1- Feature space(training data)
2- Distance metric To compute distance between records
3- The value of k The number of nearest neighbors to retrieve from which to get majority class

To classify an unknown record:

1- Compute distance to other training records
2- Identify k nearest neighbors
3- Use class labels of nearest neighbors to determine the class label of unknown record

Mechanism of the KNN algorithm

- KNN can be used for both classification problems and predictive regression. This algorithm goes through stages to predict the image classification.

Step 1: load the training set as well as the test set.
Step 2: choose the value of k, which is the closest data point. K can be any integer.
Step 3: For each new point, do the following:

- Calculate the distance between test data and each class of training data.
- Sort the distance values in ascending order.
- Next, the upper K rows of the sorted array will be selected.
- Now, it will assign a category to the test point based on the most frequent category of those rows.

Step 4: End.

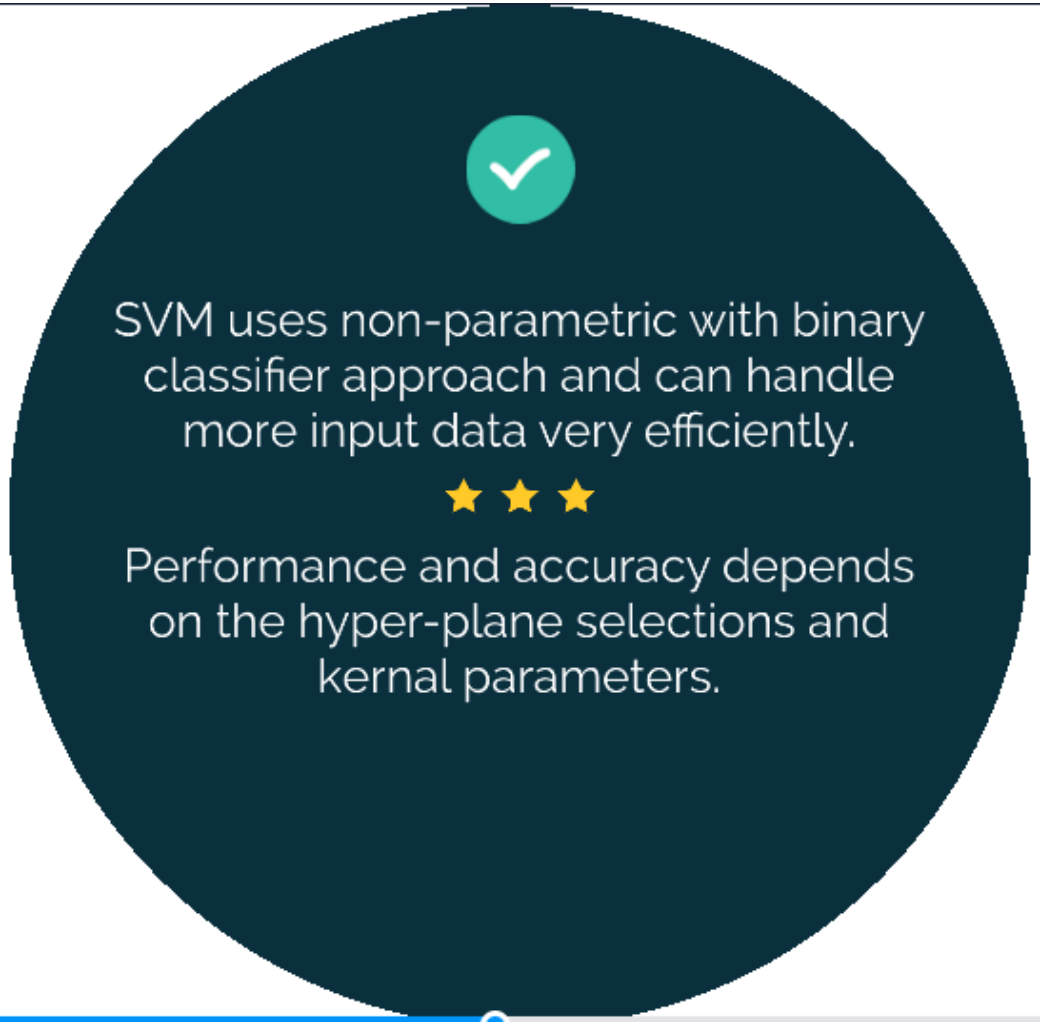
Support Vector Machine

- Support Vector Machines are a classification method that places points in space and then draws dividing lines between the points, placing objects in different classes depending on which side of the dividing plane the points fall on.

- Support Vector Machines are capable of doing nonlinear classification through the use of a technique known as the kernel trick. While SVM classifiers are often very accurate, a substantial drawback to SVM classifiers is that they tend to be limited by both size and speed, with speed suffering as size increases

- SVM build a hyper plane or set of hyper planes in a high or infinite dimentional space used for classification .

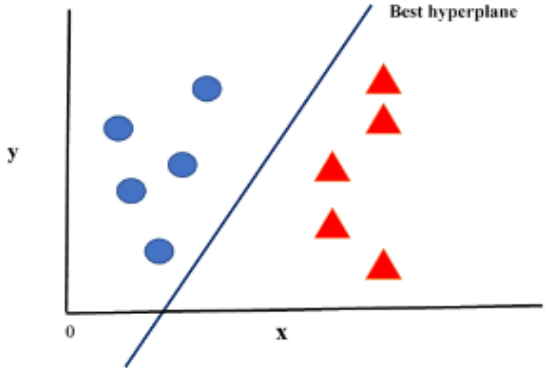
- Good separation is achived by the hyper plan that has the largest distance to the nearest trainig data point of any class, generally larger the margin lower the genralization error of the classifier.


SVM uses non-parametric with binary classifier approach and can handle more input data very efficiently.
Performance and accuracy depends on the hyper-plane selections and kernal parameters.

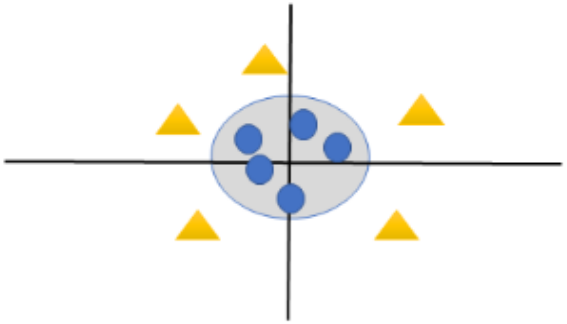
Mechanism of the SVM algorithm

The main idea of the SVM is to separate two classes from each other, either by line or by kernel

- There are two types of data and SVM is divided into linear and non-linear data :



Linear data



Non-linear data

Mechanism of the SVM algorithm

- Linear Data

Fig 13.Not all hyperplanes are created equal

The formula of hyperplane:

$$ax+c=0$$

But the maximum hyperplane:

$$x = c + \sum_{j=1}^{k} \alpha j * yj * a(j) - a$$

where c and αj are learned parameters, k is the number of support vectors, j is a support vector instance, yj is the class value of a particular training instance of vector t, and a(j) is the vector of support vectors.

Mechanism of the SVM algorithm

- Non-linear Data

The use of kernel trick in the non-linear data which it doesn't separate by line. Note that the kernel trick is not actually part of SVM. It can be used with other linear classifiers such as logistic regression. Support vector machine is only concerned with finding resolution boundaries. It can be classified in this figure

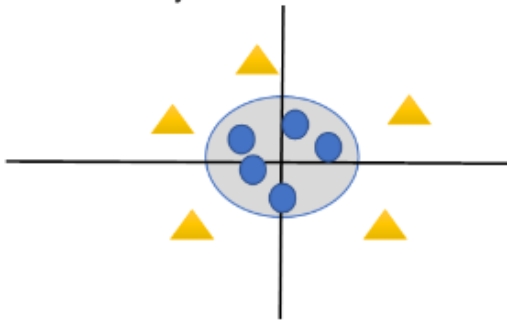


Fig 14.classification of non-linear data using kernel

- Examples of SVM Kernels:

  - Polynomial kernel:

$$k(xi, xj) = (xi.xj + 1)^d$$

where d is the degree of the polynomial.

  - **Gaussian kernel:**

$$k(x, y) = exp(-\frac{||x - y||^2}{2\sigma^2})$$

  - **Gaussian radial basis function (RBF):**

$$k(x, y) = exp\left(-\gamma||xi - xj||^2\right), \gamma > 0$$

Sometimes parametrized using:

$$\gamma = \frac{1}{2\sigma^2}$$

Mechanism of the SVM algorithm

- Non-linear Data

- **Linear splines kernel in one-dimension:**

$$k(x,y) = 1 + xy + xy * min(x,y) - \frac{x+y}{2} min(x,y)^2 + \frac{1}{3} min(x,y)^3$$

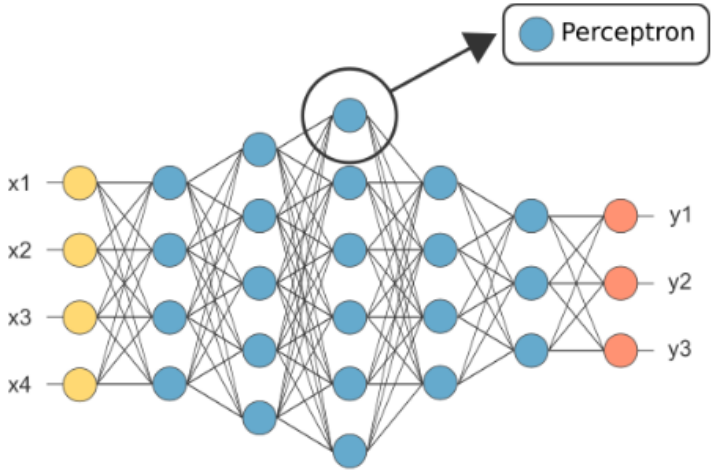
- **ANOVA radial basis kernel:**

$$k(x,y) = \sum_{k=1}^{n} exp(-\sigma(x^k - y^k)^2)^d$$

Multi-Layer Perceptrons (Neural Nets)

A Neural Network is a multil-layer networks based on a collection of connected units or nodes called perceptrons, a perceptron that receives a signal can process it and then signal additional perceptrons connected to it.

Each single perceptron has a weight and bias that define the output of each perceptron based on the input it receives.



Multi-Layer Perceptrons (Neural Nets)

- Multi-layer perceptrons, also called neural network models, are machine learning algorithms inspired by the human brain. Multilayer perceptrons are composed of various layers that are joined together with each other, much like neurons in the human brain are linked together.

- Neural networks make assumptions about how the input features are related to the data's classes and these assumptions are adjusted over the course of training. Simple neural network models like the multi-layer perceptron are capable of learning non-linear relationships, and as a result, they can be much more accurate than other models. However, MLP models suffer from some notable issues like the presence of non-convex loss functions.

Main components of ANN

1-interconnection (input layer,output layer,hidden layer).
2-activation Function.

two ways of ANN:

1-Feed-forward networks

2-back-propagation networks

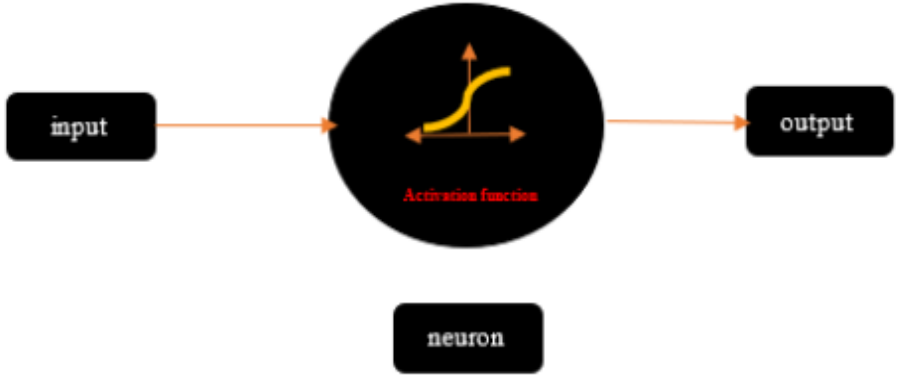


Fig 20. Activation functions.

there are 3 types of activation function :

1-Binary step function.

$$f(x) = \begin{cases} 0 \text{ for } x < 0 \\ 1 \text{ for } x \geq 1 \end{cases}$$

2-linear activation function.

$$V=cX$$

,c is a constant

there are 3 types of activation function contd :

3-Non-linear activation function.

seven Common Nonlinear Activation Functions and How to Choose an Activation Function:

- Sigmoid / Logistic.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (13)$$

- TanH / Hyperbolic Tangent

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^{x}\, e^{-x}}{e^{x} + e^{-x}} \quad (14)$$

- ReLU (Rectified Linear Unit)

$$y = \max(x, 0) \quad (15)$$

the function equlal x if x>=0 and 0 other wise.

- Leaky ReLU

f(x)=1(x<0)(αx)+1(x>=0)(x) where α is a small constant. (16)

- Parametric ReLU

$$f(y) = y \text{ if } y \geq 0$$
$$f(y) = ay \text{ if } y \leq 0 \quad (17)$$

- SoftMax

$$s(xi) = \frac{e^{xi}}{\sum_{j=1}^{n} e^{xj}} \quad (18)$$

- Swish

$$\sigma(x) = \frac{x}{1 + e^{-x}} \quad (19)$$

Convolutional Neural Network in Keras (NNs).

- The most commonly used image classification algorithm in recent times is the Convolutional Neural Network (CNNs). CNNs are customized versions of neural networks that combine the multilayer neural networks with specialized layers that are capable of extracting the features most important and relevant to the classification of an object. CNNs can automatically discover, generate, and learn features of images.

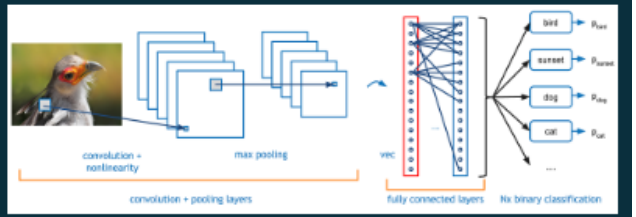
- This greatly reduces the need to manually label and segment images to prepare them for machine learning algorithms. They also have an advantage over MLP networks because they can deal with non-convex loss functions.

- Convolutional Neural Networks get their name from the fact that they create "convolutions". CNNs operate by taking a filter and sliding it over an image.

Convolutional Neural Network

Convolutional neural networks (CNNs) were proposed by Yann LeCun in 1988, and they are a special architecture of neural networks. The most popular use of this architecture is image classification.

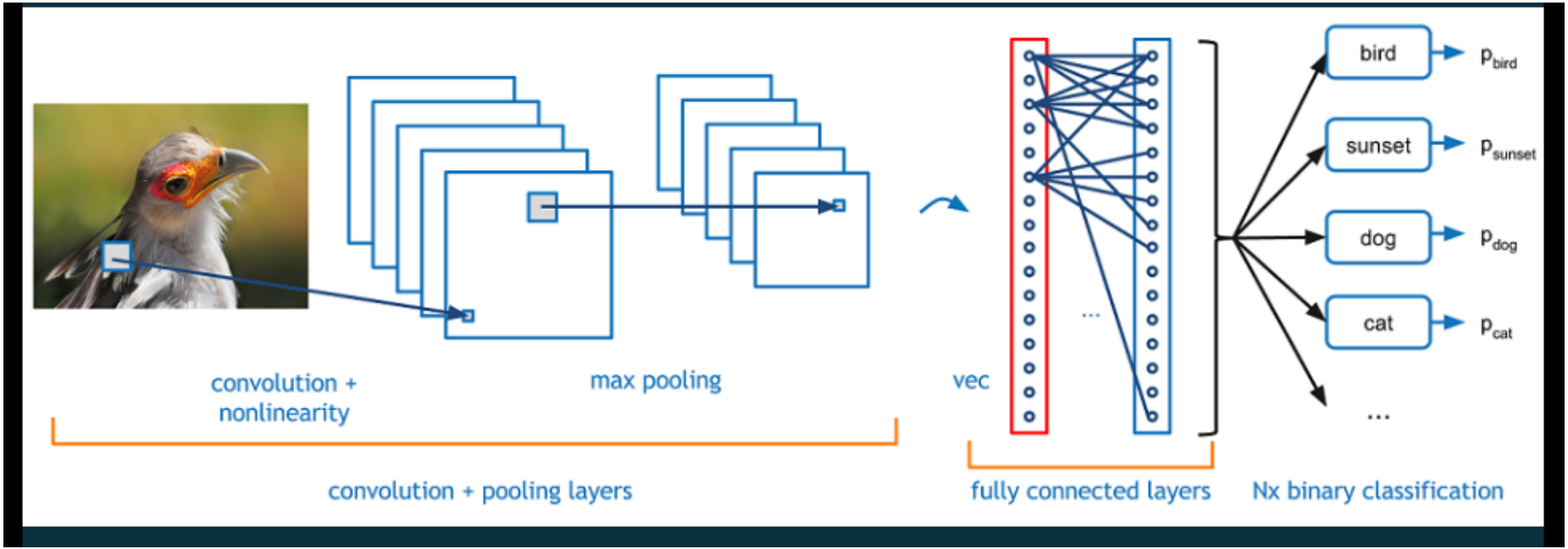
CNNs has two main processes: convolution and sampling.



Convolution Process: The purpose of this operation is to identify edges, boundaries, curves and simple colors on the image.

Sampling Process: Its purpose is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network.

Prezi


bird
sunset
dog
cat
...
p_{bird}
p_{sunset}
p_{dog}
p_{cat}
convolution + nonlinearity
max pooling
vec
convolution + pooling layers
fully connected layers
Nx binary classification

architecture of CNN

1-input layer.
2-convolution layer.
3-pooling layer.
4-fully connection layer.

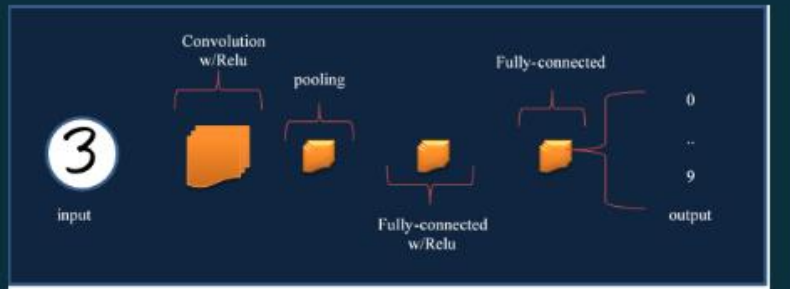


Fig 25. An simple CNN architecture, comprised of just five layers

The Convolution layers are used to reduce the complexity of the model through the optimization of its output. These are improved through three hyper parameters, the depth, the stride, and the setting zero-padding.

$$((I-F)+2P)/(S+1)$$

I --> refers to the input volume size (height x weight x depth).
F --> refers to the respective field size.
P--> refers to the amount of zero padding set.
S--> refers to the stride.

architecture of CNN contd.

2-Pooling / sub-sampling layer

Pooling layers used to gradually reduce the dimensionality of the representation, and thus further reduce the number of parameters and the computational complexity of the model.

$$Y_{vug} = \max_{(p,q) \in Rug} (a_{vpq})$$

-Relu: A Relu performs the function y = max(x,0), so the size of input and output in this layer are similar.

-Continuous trigger (non-linear) function:

The non-linear layer implements element by element in every feature. A continuous non-linear function can be a hyperbolic tangent, absolute of hyperbolic tangent, or sigmoid.

A continuous non-linear function :

1-Sigmoid:

$$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$$

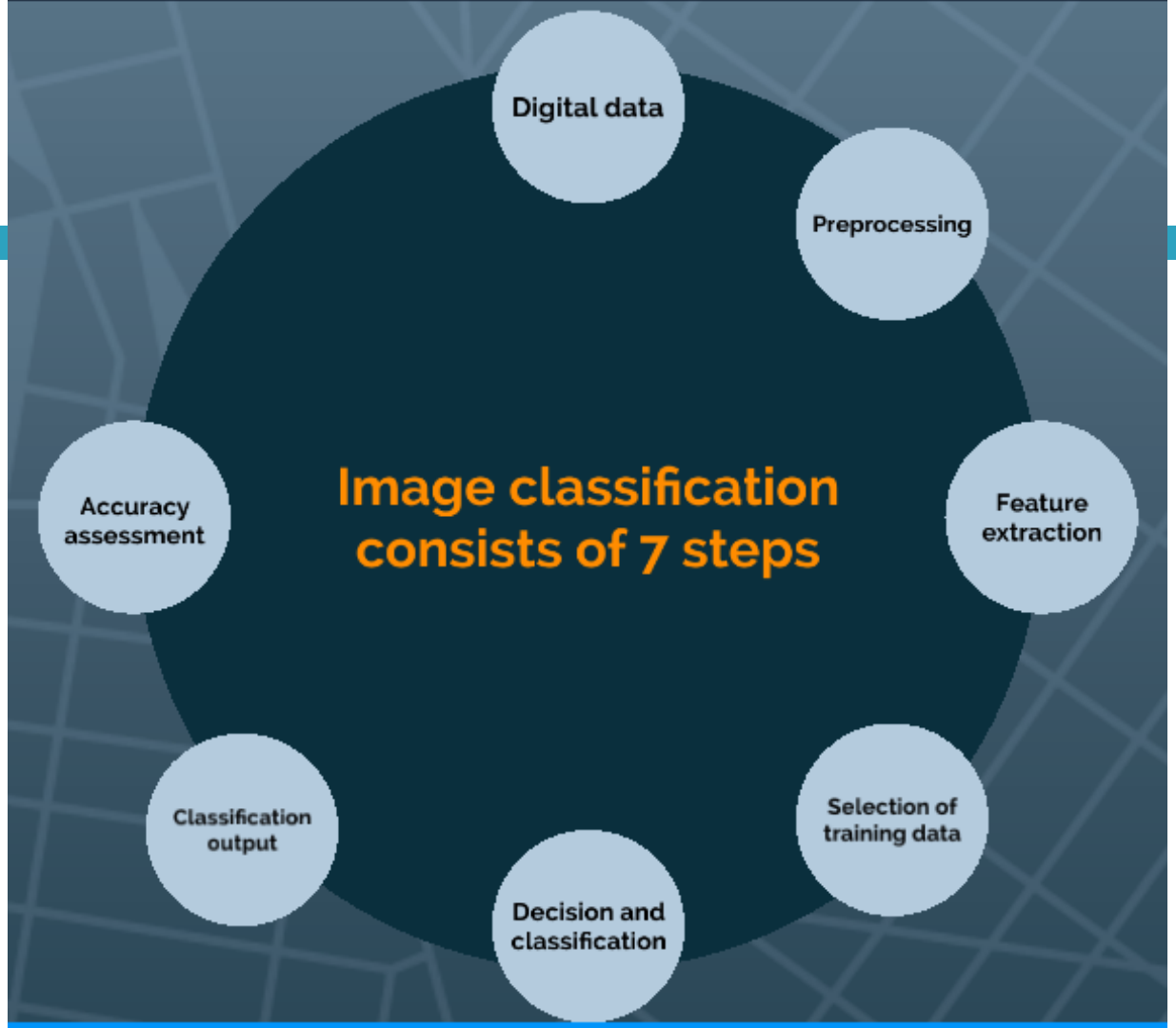
2- hyperbolic tangent

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^{x}\, e^{-x}}{e^{x} + e^{-x}} = \frac{e^{2x}\, 1}{e^{2x} + 1}$$

or

$$f(x) = tanh(x) = 2\sigma(2x) - 1$$


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Image Classification
Comparison
Classification Algorithms
Classification Process
Summary
SIRG
Prezi
SIRG GROUP

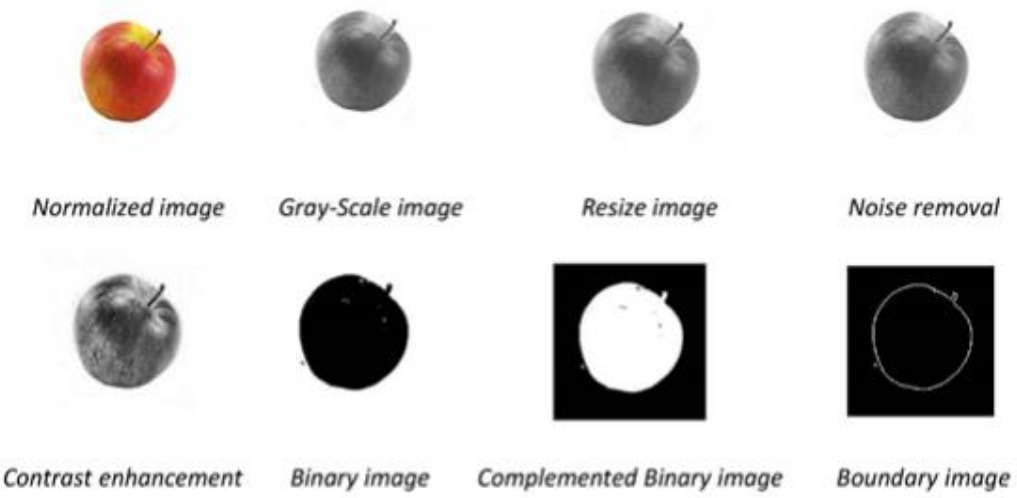

Digital data
Preprocessing
Feature extraction
Selection of training data
Decision and classification
Classification output
Accuracy assessment
Image classification consists of 7 steps

1. *Digital data*

- An image is captured by using digital camera or any mobile phone camera.

2. *Preprocessing*

- Improvement of the image data. Normalized image Contrast enhancement Gray-Scale image Binary image Resize image Complemented Binary image Noise removal Boundary image



3. *Feature extraction:*

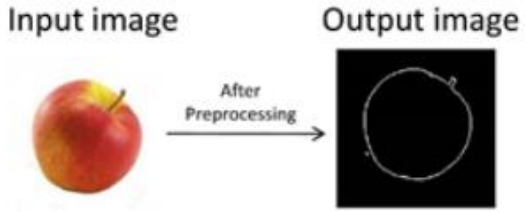
- The process of measuring or calculating or detecting the features from the image samples.

The most common two types of feature extraction are:-

- Geometric feature extraction
- Color feature extraction

4. *Selection of training data:*

- Selection of the particular attribute which best describes the pattern.



Train Data Set

Name	Major	Minor	Area	Perimeter	Red #	Yellow #
'AP.jpg'	[175.4774]	[169.6791]	[23328]	[604.9848]	[1267]	[621]

5. *Decision and classification:*

- Categorizes detected objects into predefined classes by using suitable method that compares the image patterns with the target patterns.

6. *Classification output:*

- Selection of the particular attribute which best describes the pattern.

Sampled Image						
Area or Size	1200000	900000	800000	1100000	600000	700000
Colour Intensity	80	90	91	82	100	95
Grading	Grade A	Grade B	Grade B	Grade A	Grade C	Grade C

6. *Classification output:*


```
[ ] #get the model predictions
    predictions=model.predict(np.array([resized_image]))
    #show the predictions
    predictions
```

```
array([[2.6671082e-04, 2.6750495e-05, 4.7759484e-03, 9.7983164e-01,
        4.7273710e-04, 1.3412381e-02, 5.5663224e-04, 5.1465066e-04,
        1.8441120e-04, 3.8224585e-05]], dtype=float32)
```

```
[ ] #Sort the Predictions from least to greatest
    list_index=[0,1,2,3,4,5,6,7,8,9]
    x=predictions
    for i in range(10):
        for j in range(10):
            if x[0][list_index[i]]:
                temp=list_index[i]
                list_index[i]=list_index[j]
                list_index[j]=temp

    #show sorted labels in order
    print (list_index)
```

```
[2, 3, 4, 5, 6, 7, 8, 9, 0, 1]
```

```
#printn the first 5 predictions
for i in range(10):
  if (round(predictions[0][list_index][i]*100,2)>50):
     print (classification[list_index[i]],':',round(predictions[0][list_index][i]*100,2),'%')
```

```
cat : 97.98 %
```

7. *Accuracy assessment:*

- An accuracy assessment is realized to identify possible sources of errors and as an indicator used in comparisons.


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG GROUP

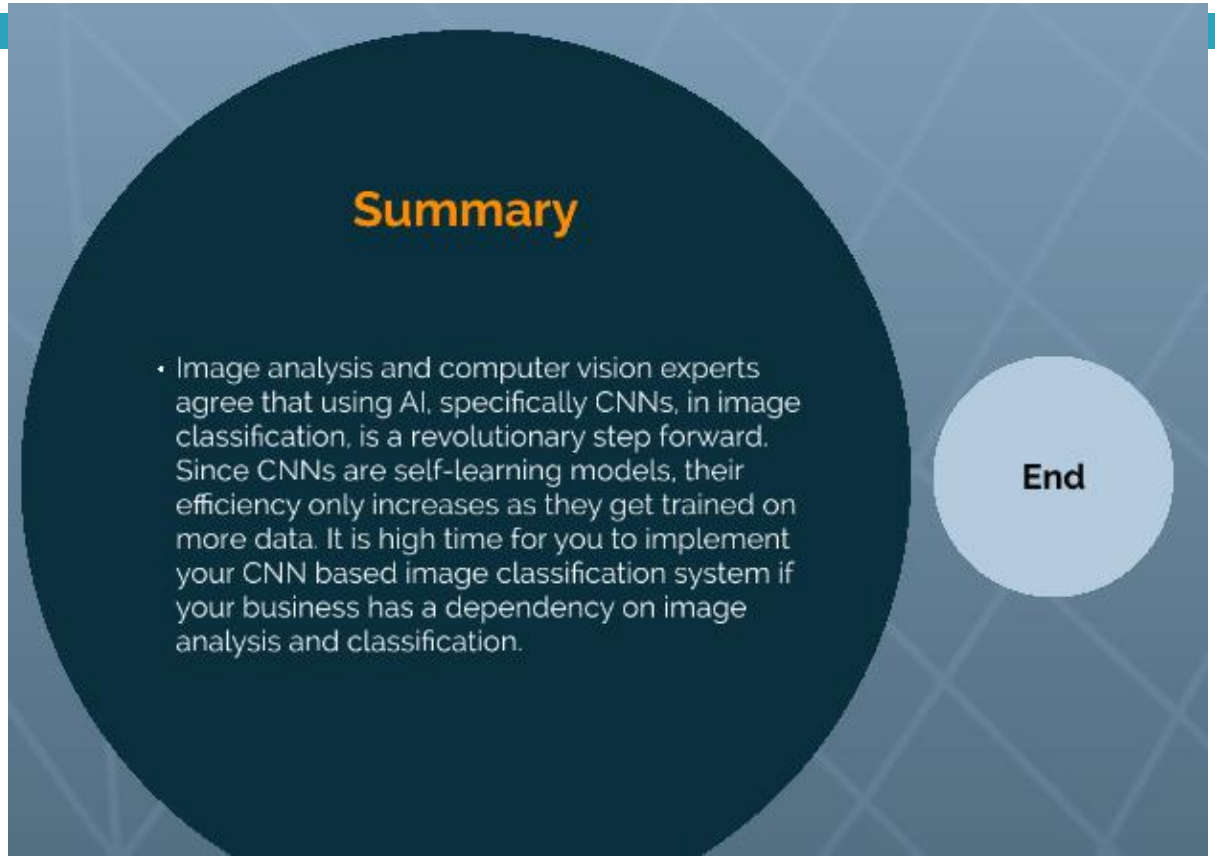

Summary
• Image analysis and computer vision experts agree that using AI, specifically CNNs, in image classification, is a revolutionary step forward. Since CNNs are self-learning models, their efficiency only increases as they get trained on more data. It is high time for you to implement your CNN based image classification system if your business has a dependency on image analysis and classification.
End


Image classification based on CNN
In Computer Vision
SIRG Group Members
Deep learning Techniques
Introduction
Summary
SIRG
Image Classification
Classification Process
Comparison
Classification Algorithms
Prezi
SIRG GROUP

Thanks and Acknowledgement

Thank you

Dr. Ahmed A. Elngar (Ph.D)
Assistant Professor
Faculty of Computers & Artificial Intelligence
Beni-Suef University, Beni Suef City, office box # (62511),, Egypt
Founder and Chair of the Scientific Innovation Research Group (SIRG)
Managing Editor in Journal of CyberSecurity and Information Management (JCIM)
Email: elngar_7@yahoo.co.uk
Email: ahmedelnagar@fcis.bsu.edu.eg
Mobile : (+2)01007400752
www.sirg.club