**Artificial Intelligence**

****

**Project Supervisor**

Dr. Talha Waheed

**Project Title**

# Image Classification Using CNN

# Project Report

**Submitted By:**

Samawia Ijaz 2020-CS-124

Wasfa Nasir 2020-CS-145

Aiemen Altaf 2020-CS-148

Department of Computer Science

**University of Engineering and Technology, Lahore**

**Table of Contents**

[Image Classification Using CNN 1](#_Toc121698203)

[Project Report 1](#_Toc121698204)

TOC………………………………………………………………………...……………………………….2

Introduction…………………………………………………………………………………………………3

Problem Statement………………………………………………………………………………………….6

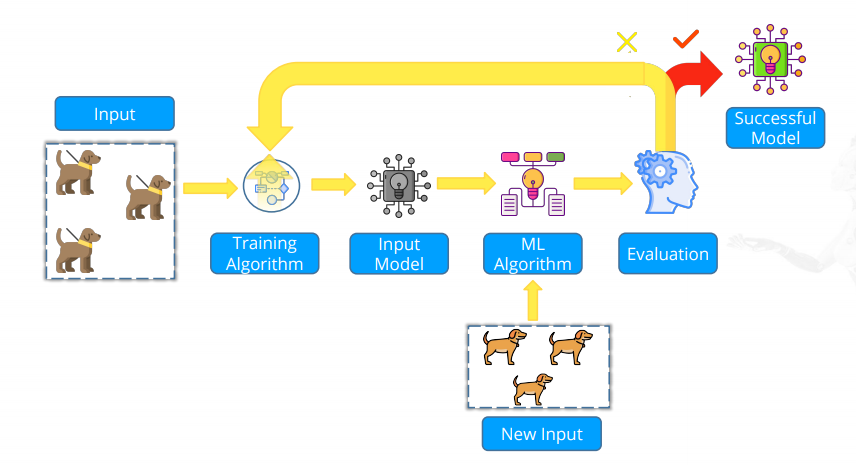
Literature Review…………………………………………………………………………………………...7

Methodology……………………………………………………………………………………………….7

Discussion…………………………………………………………………………………………………12

**Introduction**

**Overview:**

The task of image classification involves the system taking an input image and labelling it appropriately. Today Any business can utilize image categorization to speed up, simplify, and streamline the process. Have you ever wondered how my system can recognize the faces of my family and I, how cars can automatically follow traffic laws, or how people can recognize various animal images? All of this took place after considering image processing. New algorithms are developed as technology advances, and neural networks become more powerful and can handle very huge photos and videos, process them, and finish them with appropriate captions. An image can be classified and identified using a convolutional neural network, a subset of deep learning that deals with processing image and video data by extracting characteristics from them and building a neural network by giving them weights. Every data scientist has CNN on speed dial when dealing with any image or video processing data. It is very simple to use the transfer learning model and alter it using its layers. Image classification systems are typically composed of three crucial components: image preprocessing, image feature extraction, and classifier. For the dataset we will use the Kaggle dataset of cat-vs-dog. In TensorFlow we can now build the Convolutional Neural Network by defining the sequence of each layer. Now after getting the data set, we need to preprocess the data a bit and provide labels to each of the images given there during training the data set. To do so we can see that name of each image of the training data set is either start with “cat” or “dog” so we will use that to our advantage then we use one hot encoder for the machine to understand the labels (cat [1, 0] or dog [0, 1]).

**Fig:1 CNN Modeling and Machine Learning**

**Background:**

To deal with a small number of samples and computation units, traditional machine learning techniques (such as multilayer perceptual machines, support vector machines, etc.) typically use shallow structures. The performance and generalizability of complex classification problems are manifestly inadequate when the target objects have complex meanings. Convolution neural networks (CNNs), which have been developed recently, are widely employed in the field of image processing because they are effective at handling difficulties related to image recognition and classification and have significantly increased the accuracy of many machine learning tasks. It has developed into a strong and versatile deep learning model. More and more image data is being stored online as mobile Internet technology quickly develops. After text, image has grown to be another significant network information carrier. Considering this, it is crucial to employ a computer to intelligently identify and recognize these photos so they can better serve humans. People mostly employ this technique in the early stages of picture categorization and recognition to fulfil a few supplementary needs.

**Motivation:**

Since the inception of computer vision, finding effective internal representations of picture objects and features has been the major objective. As a result, numerous tools have been developed to handle photos. Many of these are built on the convolution mathematical technique. Convolution is still the key step in image processing, even when neural networks are included. Convolutional Neural Networks eventually incorporate the benefits of conventional Neural Networks and go one step farther to handle two-dimensional input. As a result, two-dimensional filters' training parameters are their components. Each level also includes non-linearity and rectification processing of the data. Deep learning ultimately produces numerous trainable phases, resulting in a hierarchically structured internal representation. It found out that such a depiction is quite potent, particularly for visuals.

Graphical user interface, diagram, application

Description automatically generated

**Fig:2 Supervised Learning**

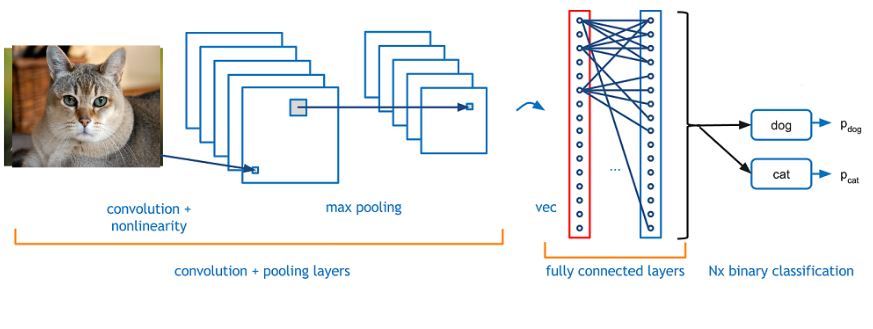
**Objectives:**

The objective of image classification is to identify and portray, as a unique gray level (or color), the features occurring in an image in terms of the object or type of land cover these features represent on the ground. Image classification is perhaps the most important part of digital image analysis. While convolutional neural networks have been around for less than a decade, other pattern identification techniques and neural networks have been around for the previous 50 years.

**Resilience to changes and picture distortion:** The use of CNN for detection makes it resistant to distortions like form changes brought on by camera lenses, various lighting conditions, variations in positions, the existence of partial occlusions, horizontal and vertical shifts, etc.

**Fewer memory requirements:** A hidden layer with 1000 features and a 32x32 input image would need an order of 106 coefficients, which would consume a lot of memory. In the convolutional layer, the same coefficients are used across multiple places in space, hence the memory required is dramatically decreased.

**Easier and better training:** Since there are many less parameters in a CNN, training time is also considerably reduced. A normal neural network would have more parameters, which would increase the amount of noise added during training. As a result, a normal neural network's performance will never match that of a CNN.

****

**Fig 3: How image classification is done using CNN**

**Scope:**

While imaging was associated with security functions and surveillance missions, the term has grown to represent something larger in recent years. Thanks to advancements in science and technology, image classification is now an integral part of AI systems. It forms the basis for other computer vision problems. Image classification applications are used in many areas, such as medical imaging, object identification in satellite images, traffic control systems, and more.

**Problem Statement:**

The main challenge in image classification are the large number of images, the high dimensionality of the data, and the lack of labeled data. Images can be very large, containing many pixels. The data in each image may be high-dimensional, with many different features. Given a set of images and prior knowledge about the content of the images find the correct Symantec label for the pixels in the images.

**Literature Review:**

The most traditional and widely used deep learning framework is the multilayer convolutional neural network (CNN).

Previous research and innovation:

* **Newman** offered a new reconstruction approach based on convolutional neural networks, and they show its benefits in terms of performance and speed.
* Three approaches were covered by **Wang**, namely the CNN model with pretraining or fine-tuning and the hybrid approach. The last category employs a patch-based feature extraction method, whereas the first two executive images are sent to the network only once. The survey reviews numerous earlier works in various categories and represents a turning point in contemporary case retrieval.
* **Huang** introduced a depth learning framework based on convolutional neural network (CNN) and Naive Bayes data fusion technique (referred to as NB-CNN). To improve the system's overall performance and robustness, a new data fusion approach is simultaneously suggested to aggregate the information retrieved from each video frame.
* **R. Gupta and Bhavsar** suggested extending the architecture of any convolutional neural network (CNN) to improve omnidirectional picture recognition using conventional 2D significant prediction (ODI). It is demonstrated from beginning to end that each step in the pipeline they describe is intended to increase the accuracy of the salient map built over the ground live data.
* For applications involving object detection and classification, the invariance retrieved by **Zhang**—such as the specificity of uniform samples and the invariance of rotation invariance—is crucial. The invariance of features, such as rotation invariance, is the subject of current research.
* **Al-Saffar** presented a framework for CNN classification that is region-based, pluralistic, and may encode representations with semantic context to produce useful features.
* **Chaib** began by outlining the emergence and advancement of convolution neural networks and deep learning, as well as summarizing the fundamental model architecture, convolution feature extraction, and pool operation of convolution neural networks.

Diagram

Description automatically generated

**Fig 4: History recalls**

**Methodology**

1. **Data collection:**

The basic principles of data collection include keeping things as simple as possible; planning the entire process of data selection, collection, analysis and use from the start; and ensuring that any data collected is valid, reliable, and credible. Data scientists and machine learning enthusiasts can connect online at Kaggle. Users of Kaggle can work together, access and share datasets, use notebooks with GPU integration, and compete with other data scientists to solve data science problems. We have used cats and dog’s dataset from Kaggle.

1. **Data understanding:**

Quality data is fundamental to any data science engagement. To gain actionable insights, the appropriate data must be sourced and cleansed.

* **Data cleaning**: The data may contain a lot of unnecessary and missing information. Data cleansing is completed to handle this portion. It entails dealing with erroneous data, noisy data, etc.

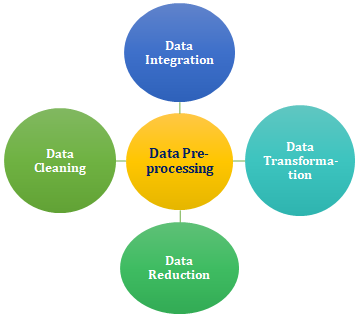
1. Missing Data: This circumstance occurs when there are gaps in the data. It can be dealt with in several ways. Among them are:
2. Ignore the tuples:
3. fill in the missing values.
4. Noisy Data: Noisy data is meaningless data that can’t be interpreted by machines. It can be generated due to faulty data collection, data entry errors etc.
5. **Binning Method:** This method works on sorted data in order to smooth it. The whole data is divided into segments of equal size and then various methods are performed to complete the task. Each segment is handled separately.
6. **Regression:** Here data can be made smooth by fitting it to a regression function. The regression used may be linear (having one independent variable) or multiple (having multiple independent variables).
7. **Clustering:**   
   This approach groups the similar data in a cluster. The outliers may be undetected, or they will fall outside the clusters.

* **Data transformation:** This procedure is used to change the data into a format that is suited for the mining process. This entails the following:
  + 1. **Normalization** is the process of scaling data values to fit inside a predetermined range (-1.0 to 1.0 or 0.0 to 1.0)
    2. **Selection of Attributes:** To aid the mining process, new attributes are created from the existing set of attributes in this technique.
    3. **Discretization:** This process substitutes interval levels or conceptual levels for the raw values of a numerical attribute.
    4. **Concept Hierarchy Generation:** In this step, qualities are raised in the hierarchy from a lower level to a higher level. As an illustration, the attribute "city" can be changed to "country."
* **Data Reduction:** Since data mining is a technique that is used to handle huge amounts of data. While working with huge volumes of data, analysis became harder in such cases. To get rid of this, we use data reduction technique. It aims to increase storage efficiency and reduce data storage and analysis costs. The various steps to data reduction are:

1. **Data Cube Aggregation:** Aggregation operation is applied to data for the construction of the data cube.
2. **Attribute Subset Selection:** The highly relevant attributes should be used; the rest can be discarded. For performing attribute selection, one can use the level of significance and p- value of the attribute. The attribute having p-value greater than significance level can be discarded.
3. **Numerosity Reduction:** This enables us to store the model of data instead of whole data, for example: Regression Models.
4. **Dimensionality Reduction:** This reduces the size of data by encoding mechanisms. It can be lossy or lossless. If after reconstruction from compressed data, original data can be retrieved, such reduction are called lossless reduction else it is called lossy reduction.

* **Data Integration:** A data preprocessing technique that combines data from multiple heterogeneous data sources into a coherent data store and provides a unified view of the data. These sources may include multiple data cubes, databases, or flat files.

1. **Tight Coupling:** Here, a data warehouse is treated as an information retrieval component. In this coupling, data is combined from different sources into a single physical location through the process of ETL – Extraction, Transformation, and Loading.
2. **Loose Coupling:**  Here, an interface is if takes the query from the user, transforms it in a way the source database can understand, and then sends the query directly to the source databases to obtain the result. And the data only remains in the actual source databases.



**Fig 5: Data understanding**

1. **Data preparation:**

Data preparation is the process of preparing raw data so that it is suitable for further processing and analysis. Key steps include collecting, cleaning, and labeling raw data into a form suitable for machine learning (ML) algorithms and then exploring and visualizing the data.

1. **Data visualization:**

A graphical representation of information and data. By using visual elements data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Steps in data visualization are:

* Develop your research question.
* Get or create your data.
* Clean your data.
* Choose a chart type.
* Prepare data and Create Chart.

1. **Design models:**

A convolutional neural network, or CNN, is a deep learning neural network sketched for processing structured arrays of data such as portrayals. CNN are very satisfactory at picking up on design in the input image circles, or even eyes and faces. We are using CNN (convolutional neural network) model in this project. There are five different layers in CNN:

* Input layer
* Convo layer (Convo + ReLU)
* Pooling layer
* Fully connected (FC) layer
* SoftMax/logistic layer and Output layer

Diagram

Description automatically generated

**Fig:6 layers in CNN**

1. **Performance analysis:**

Performance analysis seeks to explore how the innovative application of technology may help performance. Practitioners aim to provide systematic and objective feedback to athletes and coaches in order to understand, accelerate, and develop performance, including optimizing existing techniques and learning new skills. It may also be used to analyze the strengths and weaknesses of competitors.

The main components of performance analysis include tactical and technical evaluation, analysis of movement, and statistical compilation. These components are often facilitated using technology including computer software and video technology.

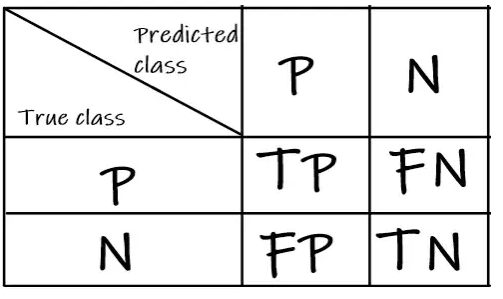
1. **Accuracy**: This is a very popular metric that tells you out of all the predictions what percentage is correct.

Accuracy = Total correct predictions / Total predictions

1. **Precision: The fraction of relevant instances among the retrieved instances**[3]. It is also known as the positive predictive value. It is the ratio of true positives to the predicted positives in the set i.e., sum of true positives and false positives.

Precision = TP/ (TP + FP)

1. **Visualization:** In machine learning for classification, a confusion matrix (also called as error matrix) is a table that allows visualization of classification performance. Each row represents the instances in the actual class and the columns represent the instances in the predicted class or the other way.



**Fig 7: confusion matrix**

**Discussion**

Limitations and challenges:

As a result, when we feed an image classification system an image from a particular category, the system will label the image with the appropriate label. For instance, let's say we provide the "image classification system" the image shown below. The image ought to be given the label "dog" by the system.

There are following main challenges in image classification:

* **Intra-Class Variation:**

Intra-Class Variation is the variation between the images of the same class. The example of intra-class variation is having chairs of multiple types in our dataset. Chair can be “office chair”, “comfy chair”, “dining table chair”, etc. So, our image classification system should be able to solve the problem of intra-class variation.



* **Scale Variation:**

This problem is very common in image classification. Scale variation is having image of same object with multiple size. Picture below is showing scale variation of same object \_\_ spoon, but they are all different sizes of spoons.



* **View-Point Variation:**

We have viewpoint variation, where an object can be oriented/rotated in multiple dimensions with respect to how the object is photographed and captured in image. No matter the angle in which we capture the image of chair, it’s still a chair.

A picture containing dining table

Description automatically generated

* **Occlusion:**

There are a lot of objects which we want to classify in image that cannot be viewed completely. There a large part is hidden behind other objects. Given is the image of cat but notice how it is resting underneath the cover, occluded from our view. It means it is not completely visible, but our image classification system should be able to detect and classify it as cat.

A cat lying on a blanket

Description automatically generated with medium confidence

* **Illumination:**

Our image classification system should be able to handle the variation in illumination. So, when we give any picture of the same object with different brightness levels (Illumination) to our images classification system, the system should be able to assign them the same label.



* **Background Clutter:**

It means that there are a lot of objects in the image and for observer it is very tough to find the object. These images are very “noisy”. It’s not easy to pick out objects. This is because the image classification system has learnt the patterns in the images during the training process. So, our images classification system should be able to solve the problem of background clutter as well.

