JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Convolutional Network for Image Recognition Using Gradio

G. Padmini[1], Pacchipulusu Kiran[1],
Sri Gadu Srinivasa Rao[2]

Students, Department of Computer Science Engineering, GITAM Deemed to be University, Visakhapatnam, India-530045[1]
Assistant Professor, Department of Computer Science Engineering, GITAM Deemed to be University, Visakhapatnam, India[2]

Abstract : In today 's world, with ever increasing advancements in the field of technology, various applications and webapps have been developed with many advanced features. In computer vision, image classification is a crucial task. Building better classifiers to determine what object is there in a photograph is a hot topic of research, with applications ranging from driverless vehicles to medical imaging. These models are ideal for Gradio's picture input component. we will build a web demo to classify images using Gradio. We will be able to build the whole web application in Python. Gradio is a package of python that allows users to create simple web apps with just a few lines of code.In this project, we will quickly create image classification model and deploy a web app using Gradio where we can upload new images for class prediction. The model will be a Keras convolutional neural network (CNN) that would be trained on images as the features and their class names as labels.

Keywords: Machine Learning, Decision Tree Classifier, K-Nearest Neighbors (KNN), Deep Learning,VGG16

I. Introduction

Recently, Image categorization has recently grown in popularity among technology developers, owing to the increase in data in various industries such as e-commerce, automotive, healthcare, and gaming. Facebook is the most obvious application of this technology. Facebook can now detect your face with up to 98 percent accuracy with only a few tagged photographs and classify it into your Facebook album. The technology almost matches a human's capacity to classify or recognize images (What is the Working of Image Recognition and How it is Used, 2017).

Deep learning is one of the most popular ways for this technology. Deep learning is classified as Artificial Intelligence since it can act and think like a human. In most cases, the system will be pre-loaded with hundreds, if not thousands, of input data in order to make the 'training' session more efficient and quick. It begins with some type of 'training' using all of the incoming data.

Machine learning is one of the most commonly used systems for picture classification. However, there are still areas of machine learning that may be improved. As a result, image classification will be handled by a deep learning system.

When it comes to Image Classification, Machine Vision has its own context. This technology has the ability to recognize persons, objects, locations, actions, and text in photos. Image classification can be greatly improved by combining artificial intelligence software and machine vision technology.

II. Literature Survey

Naeem T and Park P investigated Neural Network Architecture (NNA) as an image categorization method. The framework is made up of two sets of human eye mimics as well as variation sequence auto-encoding. It involved a lot of complicated photos, but as the study progressed, the algorithm gradually improved the MNIST models. The open source database MNIST will be utilised as the training set. It also runs tests with the Street View House Numbers dataset, with better results because even human eyes can't tell the difference.

C S Krishnamurthi discussed an image categorization system based on a Convolutional Neural Network structure (CNN). By extracting additional face images from the face image data, the training was done in such a way that a balanced number of face and non-facial images were used for training. On the Face Detection Data Set and Benchmark (FDDB), the image classification system uses a biscale CNN with 120 trained data and auto-stage training to achieve an 81.6 percent detection rate with only six false positives, whereas the current state of the art achieves about 80% detection rate with 50 false positives.

H Cecoti employed Decision Tree (DT) as an image categorization technique from [3]. The DT contains many datasets, one for each Hierarchical classifier. It is necessary in order to determine membership in each of the classes. On intermediate phases, the classifier allowed partial rejection of the class. This method has three (3) phases, the first of which is to discover terminal nodes and the second of which is to arrange classes within it. The final option is to partition the nodes. This approach is thought to be very easy to use and has a high efficiency rate.

Abid and Abubakar published in the journal ,and it tackles Support Vector Machine (SVM) active learning, which was a hot topic at the time. It also introduced a novel concept by merging spatial data from a sequential trial method with spectral data. It necessitates three strategies, the first of which is the Euclidean distance. It took some of the training samples from the main section of the dataset and calculated them. The second option employs the Parzen window technique and, finally, incorporates spatial entropy. The outcome revealed that two of the photos have a high resolution in terms of regular efficacy.

N Satish Chandra Reddy presented quick image classification by enhancing the Fuzzy Classifiers, based on the article. It was an easy approach to distinguish between the categories of known and unknown. This strategy merely enhances Meta knowledge in areas where local characteristics are most prevalent. It was put to the test with a large amount of image data and compared to the bag-of-features image model. The result was considerably greater classification accuracy because it was a testing method that took a short amount of time and delivered results that were 30 percent faster than the prior one. The summary of relevant categorization system works is shown in Table.

Research No	Name/Year	Title of project	Purpose	Method Used	Result
Research 1	Gregor, Danihelka, Graves, Rezende, & Wierstra (2015)	DRAW: A Recurrent Neural Network for Image Generation	➢ Train neural network for image classification ➢ Trained complex images with MNIST models	Artificial Neural Network (ANN)	Classification improved even naked eye cannot distinguish it with main data
Research 2	Rastegari, Ordonez, Redmon, & Farhadi (2016)	XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks	➢ Balanced number of face images and non-face images are used for training ➢ Employing the bi-scale CNN 120 trained with the auto-stage training	Convolutional Neural Network	Current state of the art achieves about 80% detection rate with 50 false positives.
Research 3	Kamavisdar, Saluja, & Agrawal (2013)	A Survey On Image Classification Application Techniques	➢ Multiple dataset that being located under each of Hierarchical classifier ➢ Rejection of the class on the intermediary stage	Decision Tree	Considered very simple and high rate of efficiency
Research 4	Pasolli, Melgani, Tuia, Pacifici, & Emery (2014)	SVM Active Learning Approach for Image Classification Using Spatial Information	➢ Combining spatial information from sequential process of trial process with spectral	Support Vector Machine (SVM)	Two (2) of the images have high resolution in terms of effectiveness of regularly.
Research 5	Korytkowski, Rutkowski, & Scherer (2016)	Fast Image Classication by Boosting Fuzzy Classifiers	➢ Simply boosting Meta knowledge where local characteristic can be mostly found	Fuzzy Classifiers	Testing process give short period of time where it produce 30% shorter compared to the previous one.

III. Proposed Methodology

3.1 Dataset

ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. The project has been instrumental in advancing computer vision and deep learning research. The data is available for free to researchers for non-commercial use.

According to the WordNet hierarchy, the ImageNet dataset comprises 14,197,122 annotated images. The dataset has been utilized in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an image classification and object identification benchmark, since 2010.

3.2 Algorithms

K-Nearest Neighbor:

The number of nearest neighbours to new input has to be classified by symbol k. It can be used as both classification and regression. It is a non parametric algorithm, lazy learning algorithm.

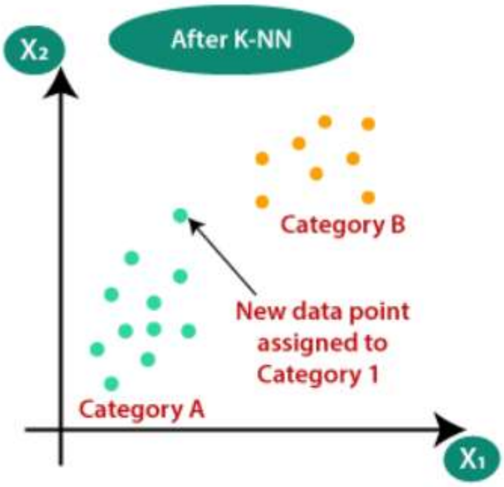


Fig 3.3.3 K-Nearest Neighbor

Decision Tree Classifier:

It is a tree-structured classification, where branches, leaf nodes are present for classification and regression. Branches are decision nodes, leaf nodes are the output of the decisions. There are various terminologies such as root node, branches or sub trees, pruning, parent node, child node. Information gain, entropy of each node is calculated. The node which has the highest information gain is will be split first.

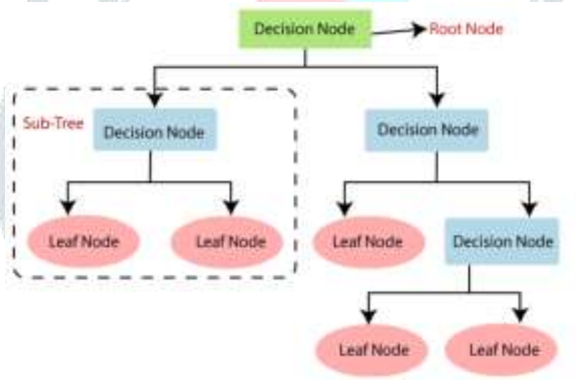
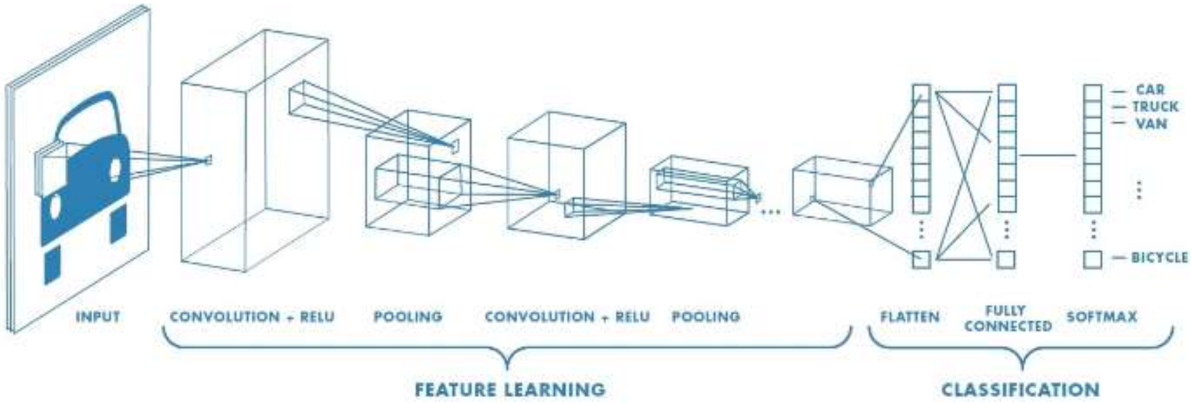


Fig 3.3.4 Decision tree Classifier

Convolutional Neural Network:

A CNN or Convolutional Neural Network is a deep learning neural network designed to analyze structured arrays of data-like representations. When one thinks about neural networks, one usually thinks of matrix multiplications, but this isn't the case with ConvNet. It employs a technique known as Convolution. Convolution is a mathematical operation on two functions that yields a third function that explains how the shape of one is changed by the other.

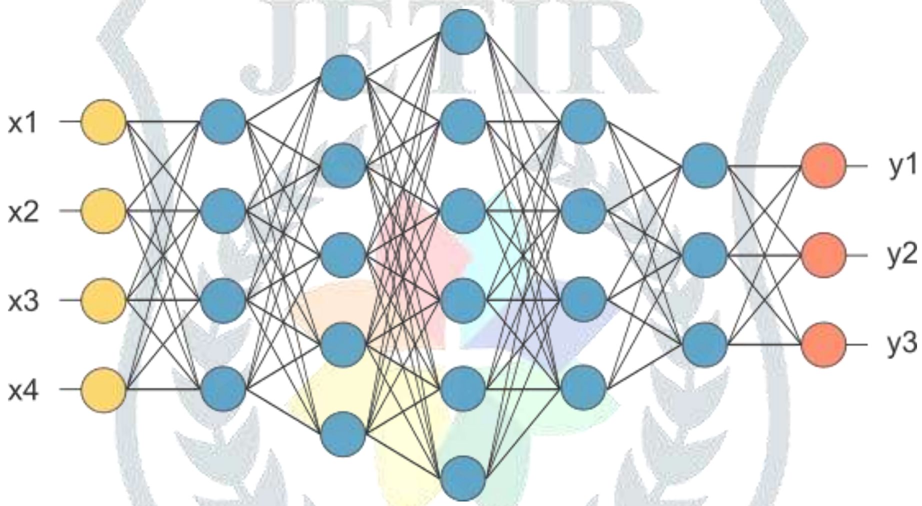
CNN's are excellent at detecting unique features in input images, such as lines, gradient circles, and even eyes and faces. Because of this feature, convolutional neural networks are effective in computer vision. CNN does not require pre-processing and can run straight on an under-done image. Feedforward neural network with up to 20 layers is known as a Convolutional Neural Network. Convolutional Neural Network strength stems from a layer known as the convolutional layer. CNN is made up of multiple convolutional layers stacked on top of each other, capable of identifying more complex structures.



Fully Connected Layer:

The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like neural network.

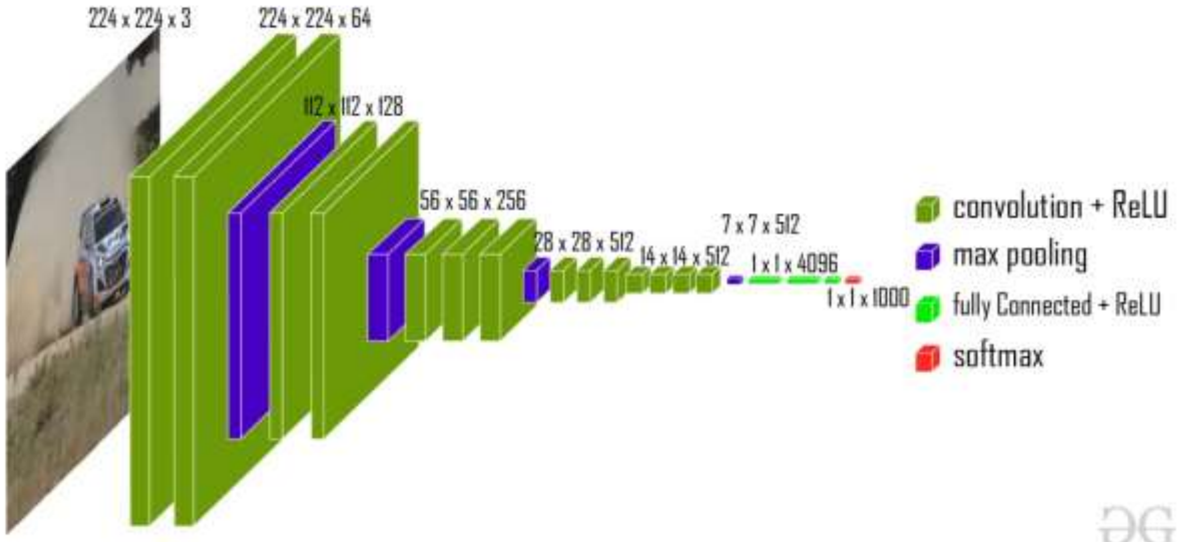
In the above diagram, feature map matrix will be converted as vector (x1, x2, x3, …). With the fully connected layers, we combined these features together to create a model. Finally, we have an activation function such as softmax or sigmoid to classify the outputs as cat, dog, car, truck etc.,



VGG16 Training Model:

It is a pre-trained model which was created by Visual Geometry Group at University of Oxford, specializing in building very deep convolutional networks for large scale visual recognition. It is trained on a huge dataset (ImageNet) with a lot of diverse image categories. The model learned a robust hierarchy of features which are spatial, rotation, and translation invariant. Thus, this gives an advantage of using pre-trained models as effective feature extractors for new images to solve diverse and complex computer vision tasks.[6]

It is a 16 layer (convolution and fully connected) network built on the ImageNet database , specially build for Image recognition and classification.



3.4 Workflow of the system

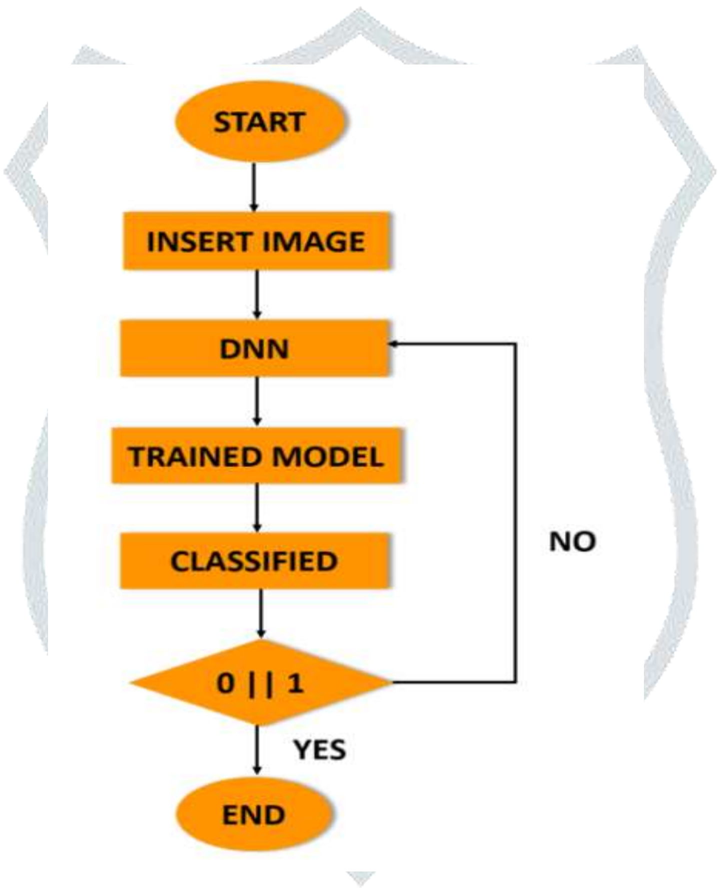


Fig 3.4.1 Workflow of the proposed methodology

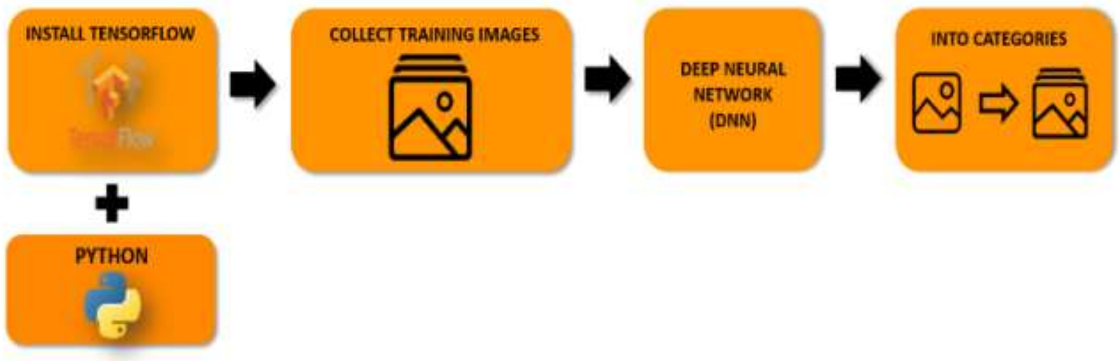


Fig 3.4.2 Dataflow of the proposed methodology

3.5 Performance Metrics

The performance metrics in this research work are cross validation with k=5, precision, recall, f1-score, support, confusion matrix. Confusion matrix consists of True positive, True Negative, False Positive, False Negative.

Accuracy - Ratio of correctly predicted observation to the total observations.
Accuracy = (True Positive + True Negative)/(True Positive + False Positive + False Negative + True Negative)

Precision - Ratio of correctly predicted positive observations to the total predicted positive observations

Precision = True Positive/ (True Positive + False Positive)

Recall (Sensitivity) - (False negatives) ratio of correctly predicted positive observations to the all observations in actual class - yes.

Recall = True Positive / (True Positive + False Negative)

F1 Score- It is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false

F1-score = 2(Recall*Precision) / (Recall + Precision)

Confusion Matrix: To determine the performance of actual class and predicted class

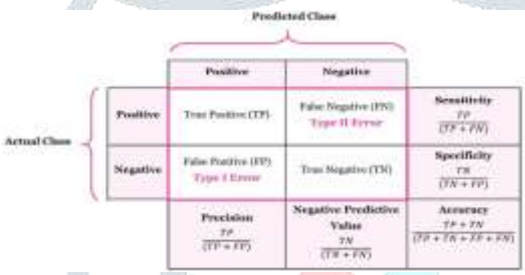
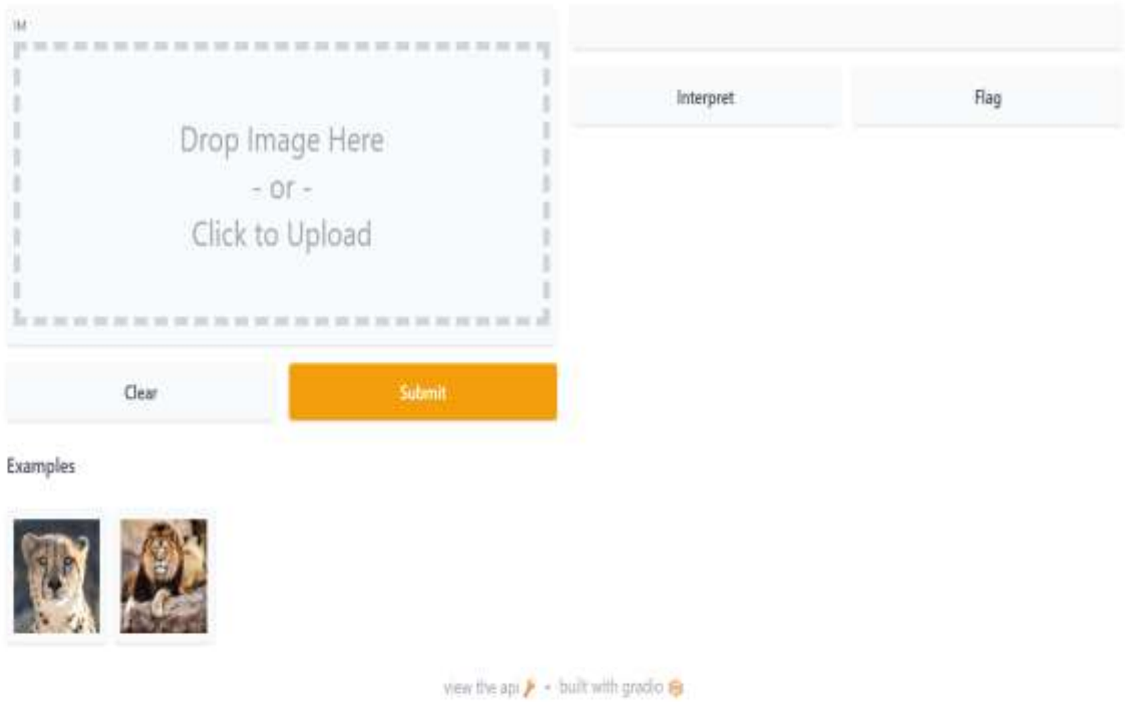
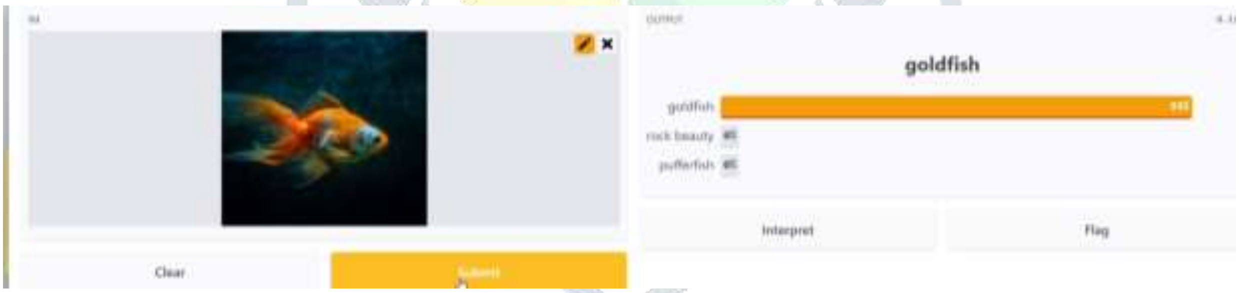
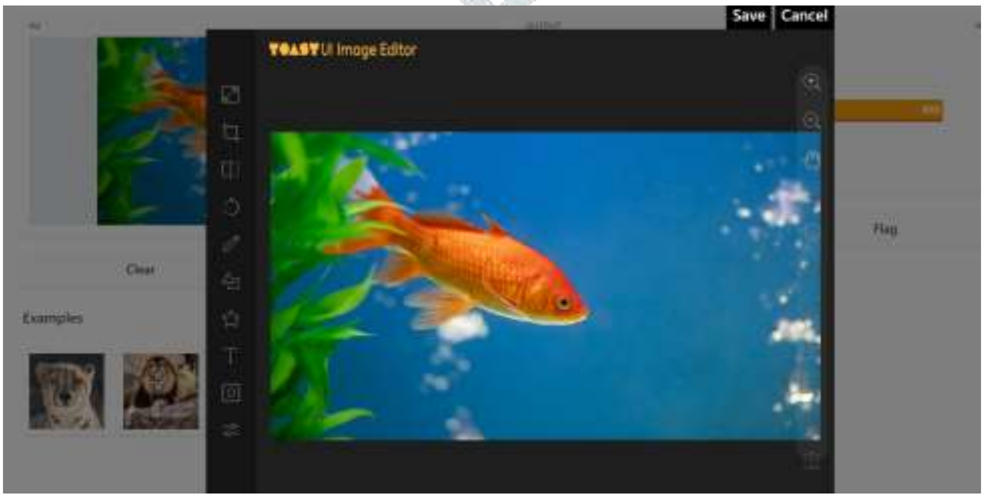


Fig 3.5.1 Confusion Matrix Performance Metric

IV. RESULTS AND DISCUSSION







.

4.1 Decision Tree Classifier

For Decision Tree Classifier, Accuracy was 86.32%, Precision was 86%, Recall was 86%.

4.2 K-Nearest Neighbours (KNN)

For KNN, Accuracy was 84.61%, Precision was 85%, Recall was 84%

4.3 Convolutional Neural Network Using Transfer Learning (CNN):

For CNN, Accuracy was 91.11%, Precision was 93%, Recall was 89%

```
Performance Summary of Sequential Neural Network on test data:
              precision    recall  f1-score   support

           0       0.89      0.75      0.82      3723
           1       0.79      0.91      0.84      3761

    accuracy                           0.83      7484
   macro avg       0.84      0.83      0.83      7484
weighted avg       0.84      0.83      0.83      7484

[[2808  915]
 [ 353 3408]]
```

4.7 Model Evaluation

Each model is compared with their accuracy, precision and recall.

Classifier	Accuracy	Precision	Recall	Time for output
K-NN	86.32 %	0.86	0.86	2.7s
Decision Tree	84.61 %	0.85	0.84	1.5s
CNN	91.11 %	0.93	0.89	0.1s

Fig 4.7.1 Model Evaluation of each model

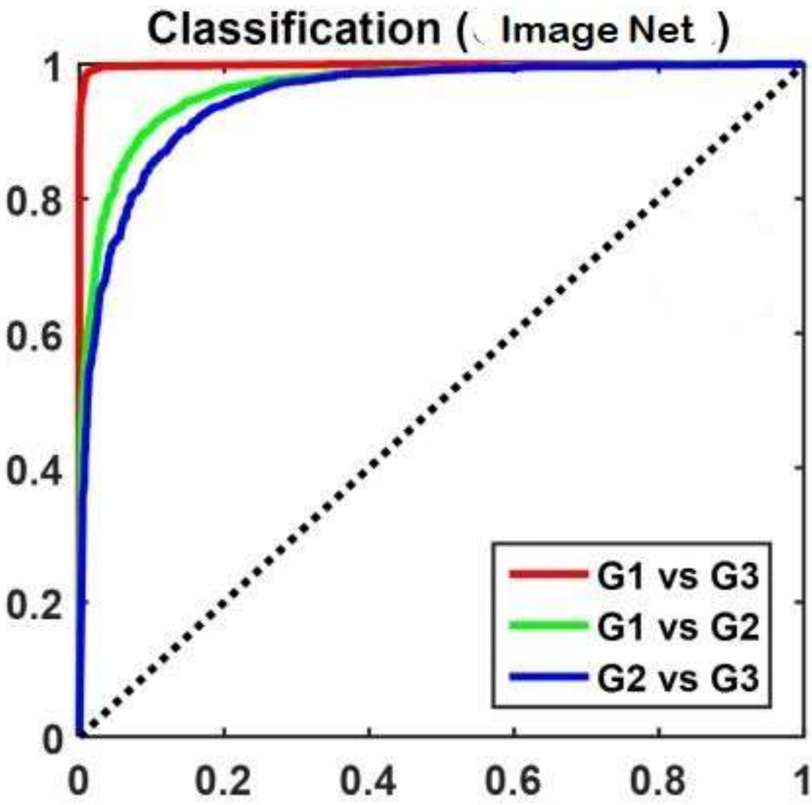


Fig 4.7.2 RoC

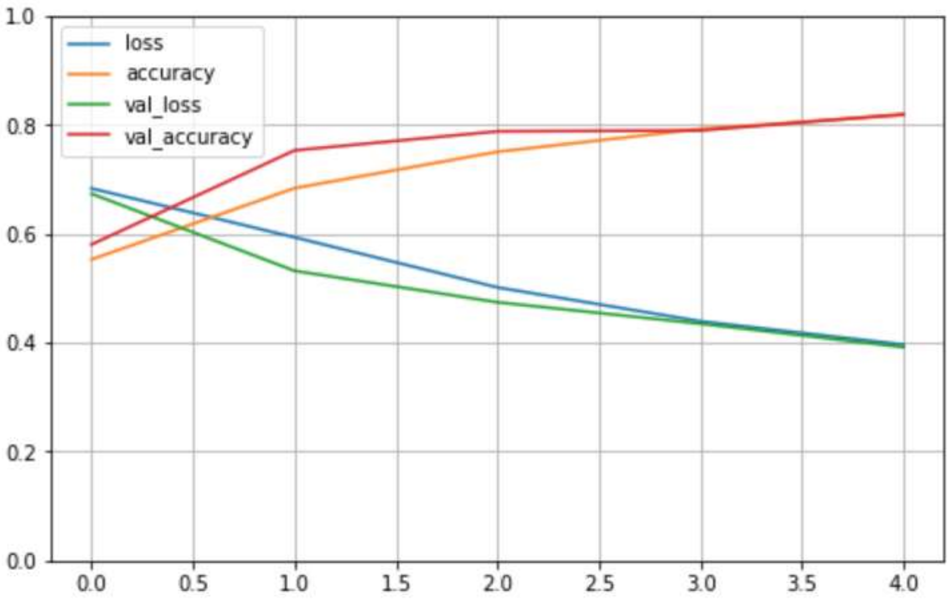


Fig.: Training Vs Validation

V. Conclusion

The main purpose of this paper is to classify images using Machine learning model and deploy the model using Gradio Package, Using the GradioML tool, a CNN image classification model was developed and deployed as a short example. Gradio, on the other hand, can be utilised for a wide range of applications, including regression/numerical prediction and natural language processing. Gradio is a tool that can serve as a quick and efficient solution for delivering a quick machine learning demo, with a variety of interface options that are ideal for these jobs. Gradio allows us to Deploy a model in a functional user interface with just a few lines of code. Web app successfully classified the images provided in testing. web application we deployed can run in local at in localhost URL : http://127.0.0.1:7860/ and run publically at URL : https://14280.gradio.app .

REFERENCES

[1] Naeem, T., Qamar, S., & Park, P. (2021). User Interface Application for Cancer Classification using Histopathology Images using Gradio. Journal of the Korean Society of Systems Engineering , 17 (2), 91–97. https://doi.org/10.14248/JKOSSE.2021.117.2.091

[2] C. S. Krishnamurthi and I. M, "Sign Language Translator Using Deep Learning Techniques," 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2021, pp. 1-5, doi: 10.1109/ICECCT52121.2021.9616795.

[3] H. Cecotti and G. Prasad, "Single-trial detection of realistic images with magnetoencephalography," 2015 International Joint Conference on Neural Networks (IJCNN), 2015, pp. 1-6, doi: 10.1109/IJCNN.2015.7280778.

[4] Abid, Abubakar, et al. "Gradio: Hassle-free sharing and testing of ml models in the wild." arXiv preprint arXiv:1906.02569 (2019).

[5] N. Satish Chandra Reddy, Song Shue Nee, Lim Zhi Min and Chew Xin Ying, "Classification and Feature Selection Approaches by Machine Learning Techniques: Heart Disease Prediction", International Journal of Innovative Computing, 2019

[6] Bukhsh, Zaharah A., Andrej Anžlin, and Irina Stipanović. "BiNet: Bridge Visual Inspection Dataset and Approach for Damage Detection by gradio." International Conference of the European Association on Quality Control of Bridges and Structures. Springer, Cham, 2021.