International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 04 | April - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Gallery Image Classification

Yash Thakkar
Information technology
Shah and Anchor Kutchhi Engineering College, Mumbai, India

Ishani Mody
Information technology
Shah and Anchor Kutchhi Engineering College, Mumbai, India

Hitika Bhandari
Information technology
Shah and Anchor Kutchhi Engineering College, Mumbai, India

Swati Nadkarni
Information technology
Shah and Anchor Kutchhi Engineering College, Mumbai, India

Komal Patil
Information technology
Shah and Anchor Kutchhi Engineering College, Mumbai, India

***Abstract*- The issue of peripheral device space management is a pressing concern for everyday users who deal with massive amounts of media content obtained through social media applications. The content that is shared over the social media application can be important images such as notes, personal pictures like selfies, group pictures, etc., and along with these images, there are many unsolicited images shared that are directly stored in the device. These unsolicited/spam images not only consume storage space but also selecting them from all and then deleting them is a tedious process, sometimes by mistake, a user also deletes important images, which can consume a lot of time. This project aims to specifically detect and extract study notes, personal images & spam images separately and relocate them in three different folders, which will help to segregate the images. This will make searching for a particular image easier and save time and effort.**

Keywords- Unsolicited images, notes images, personal images

I. INTRODUCTION

Over the last decade, Millions and Billions of images have been uploaded and received over the social media applications like WhatsApp, Facebook, etc. These applications are used by almost everyone nowadays. The data that is sent over these social media applications can be personnel images, study notes, or any unsolicited data. Initially, unsolicited data was only available in text format. Many classifiers were developed and trained using machine learning to filter such spam content. However, as time passed the spam textual information was delivered in the form of Images. As a student at the end of the semester, the student's gallery is almost full of study notes. Due to online teaching sessions, most of the notes are received online. The problem of space management of peripheral devices is a pressing concern for everyday users who have to deal with the massive data sent over social media applications. The content that is shared over the social media application can be important images such as notes, personal pictures like selfies, group pictures, etc., and along with these images, there are many unsolicited images shared that are directly stored in the device. These unsolicited/spam images not only consume storage space but also selecting them from all and then deleting them is a tedious process. Keeping all these things in consideration, the proposed system will be an Image classifier, which will classify images that are stored on the user's device. It will specifically detect and extract study notes, personal images, and unsolicited data and store them separately in three different directories. This proposed classifier will help the user with the issues of space management and will also help save the time and effort needed to segregate the images. A convolutional neural network using python's Keras library along with the Tensorflow integration is built to classify the images, which will then segregate the images and dump them into different Directories.

II. LITERATURE SURVEY

Image classification can be done using various classification techniques. In image classification, deep convolutional neural networks recently outperformed humans, and these networks can be used to derive highly predictive [4] features of an image's tags. In this study, the challenges that are faced when implementing a particular image classifier technique and how these techniques are used to identify languages were analyzed. It also gave us an idea about how testing and training are done using different methods.

A. Existing Systems

The Existing system focuses mainly on Image spam filtering using different machine learning techniques. The existing system gives an idea about how images are trained by creating a dataset and testing by using different testing methods for further image classification. Here are a few classifying techniques that are used to classify spam images in existing systems.

- **Support Vector Machine**: When image data is in large quantities or infinite dimension, Support Vector Machines characterize it using hyperplanes or a group of hyperplanes [3]. It is a binary classifier that uses a linear boundary to distinguish two groups. The closest training point in any class has the farthest distance from a single hyperplane, according to a good classification.
- **Artificial Neural Network**: An artificial neural network is a shred of the vast artificial intelligence system..[3] This network mimics how the human nervous system works. The operation of neurons in the human brain is imitated by an artificial neural network.[3]

- **Optical Character Recognition:** Optical Character Recognition is a technique for detecting optical patterns in digital images [1]. That is the process of converting scanned images of hand-written or typewritten text into machine text. It is a three-step procedure [1]. To begin with, the picture is segmented. Then, from a picture, features are extracted. Finally, image recognition takes place.

III. Proposed System

The proposed system Titled " Gallery Image Classifier" focuses on classifying images and relocating them according to their types. The aim of developing this project is to classify Notes images, Personal images, and Spam images. For instance, all the images of 'Study Notes' or all personal images or the unsolicited data sent through social media applications can be grouped and stored in different folders. The images can be grouped as 'Study Notes', 'Spam Images', or 'personal images which will help to segregate the images. This will make searching for a particular image easier and save time and effort.

The dataset is the collection of different types of images. On training the dataset, it will be able to differentiate various kinds of images, to support our instance, a convolutional neural network using python's Keras library along with the Tensorflow integration will be built.

The classification model will try to conclude the input given for training which will help segregate the images and dump them into different folders.

IV. Method

For the proposed system, Python's Keras Library along with the Tensorflow integration is used to build a neural network. Tensorflow is an open-source platform for machine learning. It has many libraries and is flexible which allows users to create neural networks with many layers. On the other hand, Keras is the high-level Application-programming interface of Tensorflow. In Figure 1, there are four phases through which the image goes for image classification.

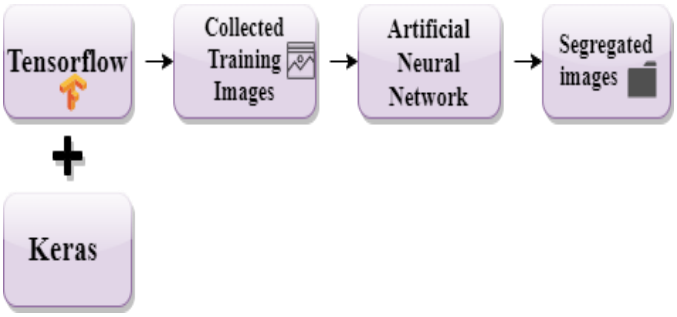

Fig. 1. Image classification. (Block diagram)

Tensorflow and Keras integration is applied on the collected trained images and then the process will continue to collect inputs by applying an artificial neural network lastly; it will classify the image given for testing into the specified group.

V. Training Dataset and Results

Dataset for the proposed system mainly consists of three different directories/classes namely 'Spam', notes' and 'personal' with a total of up to 3012 images. From the dataset, 2416 images are used for training, and 603 images are used for validation.

TABLE I. Number of Images in Dataset

No.	Directories/ classes in the dataset	No of images
1	Notes class	1113
2	Spam class	1092
3	Personal class	807
Total		3012

The proposed classifier intends to classify the images on the user's device and then relocate them to the specified folder. If the image given for testing is a group photo, the user here already knows that it is a personal picture but the model should classify it accordingly and then store the image in the personal folder of the user's device. Thousands of images are collected and trained the images so that the model will classify the images.

Let's test few different images to check whether the classifier is classifying the particular image correctly or not.

Fig. 2. Testing Image 1 (Personal image)

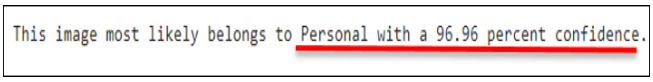

Fig. 3. Testing results (Image 1)

Fig. 4. Testing Image 2 (Notes image)

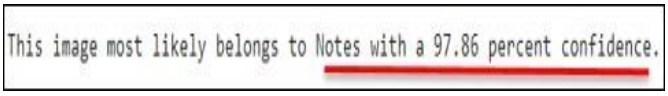

Fig. 5. Testing results (Image 2)

Fig. 6. Testing Image 3 (Spam image)

Fig. 7. Testing results (Image 3)

The images given for the testing were correctly classified with good accuracy. It is shown that the accuracy of the personal image (Fig 2.) is 96.96% when it was simulated ad tested. Similarly, the accuracy for the Notes image (Fig 4.) and the accuracy for Spam mage (Fig 6.) is 97.87 % and 78.09% respectively

VI. Conclusion

Machine learning, specifically Artificial neural networks, is a growing trend in software development that has gone mainstream in recent years. Image Classification is one of the supervised learning problems in Deep learning. There are many image classification models built and although these existing models are considered successful they mainly focus on classifying unsolicited data. The proposed system not only classifies unsolicited/spam data but also classifies study notes and personal images. The model along with classifying will also segregate the images in three different folders. The images gallery on the device is relocated into different folders so that finding a particular image does not become a tedious job for the user.

References

[1] Hamad, Karez & Kaya, Mehmet. (2016), "A Detailed Analysis of Optical Character Recognition Technology.", International Journal of Applied Mathematics, Electronics, and Computers. 4. 244-244. 10.18100/ijamec.270374

[2] K.Karthick, K.B.Ravindrakumar, R.Francis, S.Ilankannan, "Steps Involved in Text Recognition and Recent Research in OCR; A Study", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1, May 2019

[3] Abhishek Rungta, Bhawna Arya, G. Usha, "Image Spam Filtering Using Machine Learning Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2S4, July 2019.

[4] Mourya Gupta Vakacharla, Bharath Kumar N.G, Chilukuri Vishal Reddy, "Identification of Spam Images Using Image Processing and Machine Learning ", International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 5, Issue 10, October 2018, ISSN (Online) 2394-2320.

[5] Namrata Musale, Dishant Mohite, Avinash Monde, Ramesh Holkar, 2019, "Image Tagging and Classification", International Journal Of Engineering Research & Technology (IJERT) Volume 08, Issue 03 March – 2019.

[6] Mohd Azlan Abu, Nurul Hazirah Indra, Abdul Halim Abd Rahman, Nor Amalia Sapiee and Izanoordina Ahmad, "A study on Image Classification based on Deep Learning and Tensorflow", International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 4 (2019), pp. 563-569.