Detection of textual and faces for images with clustering and duplicates

Kesavan T [1],Madhan balaji M [2], Maheswari M[3]

1 *Student, Computer science and engineering, Anand Institute of Higher Technology, Chennai, India.*

2 *Student, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India.*

3 *Assistant Professor, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India.*

ABSTRACT

As cameras integrated with the mobile phones the image files are growing exponentially and there is a need of proper management of them. This can be effectively done by using various image-processing techniques. The existing system does the management of images for their account holders by providing feature like face clustering where the clusters are made for each individuals. These accounts can be logged into N number of devices and sometimes people forget logging off. The unknown user of the device can get access to the account. The paper does the management of images on offline and also provides features like face clustering, Duplicate or near duplicate image detection and identification of textual images.

Keywords: Image-processing, Face clustering, Duplicate detection, Textual identification.

1. INTRODUCTION

An image file is a sequence of pixels of certain size each of which have a pixel value based on the RGB(red,green,blue) combination. Each pixel possess a color and contribute to the overall image. A single pixel cannot be used in representing image but it helps in contributing the overall image. The maximum number of pixels that can be placed within the resolution in a sequence defines the quality of the image. An image can be considered as a two dimensional or three dimensional matrix based on the number of image channels. For black and white image, there are only two channels, black and white therefore two-dimensional matrix. For a color image three channels are used known as RCB therefore three-dimensional matrix is used to represent the image. The matrix values are the pixels and for a three dimensional matrix the RGB values are combined to form a pixel. OpenCV-python which is library of python bindings designed for the computer vision problems. OpenCV-Python makes use of numpy arrays which does support for numerical operations in a highly optimized way. These numpy values can be a real or complex numbers represented by a finite number of bits. The image is first get converted into a numpy array with N dimension based on the channels the image possess. These images can then be processed and displayed on a high-resolution monitor. For display, the image is stored in a rapid-access memory buffer memory, which refreshed the monitor at a rate of 25 frames per second to produce a visually continuous display.

LITERATURE SURVEY

The detection of face coordinates on any images is the primary task for the face clustering task which is extracting the bounding region of a face on an image. This can be achieved by a deep study of face detection using OpenCv [2]. This is followed by extracting the features of face such as width of mouth, width of eyes,

pupil and comparing with the record set and returns the closest record. That record is return along with the label where for

that label the attendance system marks as present [4]. Clustering the face images according to identity includes two applications one is grouping a collection of face images when no external labelling is done for the images and indexing for efficient large scale face retrieval. The clustering problem is composed into two parts, the representation and similarity metric for face images and choice of the partition algorithm. The paper propose a representation which has been shown to perform very well in image clustering algorithm. . It uses an *k*-NN variant of ConPac , which has a linear time complexity given a *k*-NN graph, suitable for large datasets [1]. The system about near duplicate images which contains extensive bounces of matching image that is not present in other images (i.e. non-duplicate images). These image fragments which are present only in a few images are the distinctive topographies that differentiate analogous images with disparate images in a modest way. The Idiosyncratic topographies are used to compare the images for the recognition of duplicates or near-duplicates. By equating the topographies instead of entire images, in the outcome of duplicates or near-duplicates in a huge assortment of images [6]. Textual image grouping comes with two important task text detection and text recognition. Text recognition follows a sequence which starts with capturing the image that has text which is the primary requirement of the processing, Then followed by detecting the signboard which can be 3D, Neon, Creative signboard etc. Having different layout with different shapes and angles. Text detection is the next phase where the text region on the sign board is identified. Text detection can be done with EAST known as Efficient and Accurate Scene Text Detector [5] which uses a simple yet powerful pipeline for accurate text detection on images. This is an effective method for identifying the text region on the image after on the region the text recognition can be made. This includes identification of text region on a blurred, having broken ligature, variation in size, colour, resolution, shape texture, background geometry of text , lighting problem and contrast with background. After finding the region the text extraction is done which is also known as localisation [3].

3. EXISTING SYSTEM

The existing system provide a free cloud service for their account holders with certain amount of storage for each account. The user device has to be synchronized with the cloud. Now the person can log into N number of system and can view, download or modify the image from there. The system minimize the image management efforts by providing features like face clustering where the images are made into clusters of each individuals and location based clustering where based on the location the image was capture the clustering is made. Whenever person forget to log off from any one of the devices an unknown can get access to all the image file with that account.

4. PROPOSED SYSTEM

The proposed system aims at reducing the management effort and also minimize storage. This is achieved by using three techniques namely Face clustering, Duplicate or near duplicate detection and Text recognition. The system uses all the above technique in a flow and as a result the images are organized as clusters of faces and textual images. The flow starts with finding the textual images and separating them from the dataset. The remaining dataset is passed to the face clustering and followed by duplicate identification on the clusters.

4.1 Textual image grouping:

This is the first process, In this step the system gets the dataset from any local storage location of the device. This part does the management of textual images like meme type images, images with Quotes on it and softcopy of student notes. These images serves no purpose after some period of time, But these images will be left in the storage for a long time unnecessarily. The system detects the text bounding areas present in the image using OpenCV EAST model and based on the number of text or word, the text recognition is used. For images which consist more word it is considered as text image and saved in the destination location and for images with minimum words the text recognition is used to get what text in it and the system decides whether it is a textual image or normal image.

Fig 1. Text detection and text extraction from images

4.2 Face clustering:

The dataset is the remaining set of images in the system after textual image grouping is made for which the management has to be made. The clustering task begins by converting the images as a numpy arrays. Then the arrays is passed to OpenCV face detector which is pre-trained deep learning model, helps in providing the bounding coordinates of the face in the image. Then we obtain the description of the image as an 128D vector. This is done for all the images in the dataset as a result we have a set of descriptions. These descriptions are passed to the Chinese whisper clustering algorithm which performs clustering and return a set of labels of each cluster. Using the labels the images are saved as separate clusters. Fig 2. Is the result of image after face clustering.

Fig 2. Face clustering on dataset.

4.3 Duplicate Identification:

This process is done only after the clustering step as this part work on the result of the clustering process. The dataset for this process is the resultant set of clusters from face clustering. Same image captured many times, same image downloads from any social media or same image on separate location of same device still resides on the storage as it is. The clustering results contains these duplicates also. So this process does the identification of duplicate or near duplicate image on each clusters separately. The resulting duplicate images are grouped so that the user can select the one which he/she thinks the best. This is done by extracting the features by SIFT and SURF algorithms on the duo of images and comparing the features of the images matching the similar one.

The duplicate image files need not be with similar file size, it can be identical images but with different size due to some compression technique followed while sharing them on a network. The system detect these duplicates also(Fig 3). The below example is set of identical images on which the right one is the compressed one. Another sample set are dissimilar image set (Fig 4).

Fig 3. Duplicate image set

Fig 4. Non-duplicate image set

5. CONCLUSION AND FUTURE ENCHANCEMENTS

An optimized automated computer based method is proposed and validated for management of images file. The dataset from any specific location is taken and the automation begins with textual image identification followed by face clustering process which takes the remaining dataset and performs the clustering, then the cluster result are stored on destination location. The duplicate identification performs duplicate image identification present among the cluster resultant.

In future the system will be capable of grouping location based clustering where the clustering will be made based on the location the image was captured. Also on face clustering , images with common person in any cluster will be maintained as numpy arrays and can be viewed dynamically only when the person wants result of clusters. Further Text detection from images that are not good in quality will also be capable of being processed.

6. REFERENCES

[1] Yichun Shi;Charles Otto;Anil K Jain(2018) Face Clustering Representation And Pairwise Constraints International Conference of Electionics ,Communication And Aerospace Technology(ICECA).

[2] Kruti Goyal; Kartikey Agarwal; Rishi Kumar(2017).Face Detection and Tracking Using OpenCv. International Conference of Electronics, Communication and Aerospace Technology (ICECA).

[3] Partik Madhukar Manwatkar Shashank H Yadav (2015).Text Recongition From Images International Conference on Innovations In Information And Communications Systems (ICIIECS).

[4] Adrian Rhesa Septian Siswanto;Anto Satriyo Nugroho; Maulahikmah Galnium(2014). Implementation Of Face Regonition Algorithm For Biometrics Based Time Attendance System.International Conference On ICT For Smart Society (ICISS).

[5] Xinyu Zhou,cong Yao He Wen Yuzhi Wang,Shuchang zhou Weiran He ,Jaiajun Liang(2017) East:An Efficient And Accurate Scene Text Detector. Computer Science: Computer Vision And Pattern Recognition.

[6] V.A Naruana ; SreevaniGaddameedhi; Vijaya Kumar Koppula; K. Srujan Raju (2018). Framework for Proficient Proof Of Identity Of Duplicate And Near-Duplicate Images And Image Distances Using High-Disguisable Image Fragment . Fifth International Conference On Parallel, Distributed And Grid Computing (PDGC).