

# Introduction of New Technique to Filtering Unwanted Data for Forensic Lab

Bankar Kalyani<sup>1</sup>, Pandore Gayatri<sup>2</sup>, Malkar Shubhangi<sup>3</sup>, Jadhav Manisha<sup>4</sup>  
Prof. Chandgude A.S<sup>5</sup>.

<sup>1</sup>, Student, Department Of Computer Engineering, SND College of Engineering & Research Centre, Yeola, Maharashtra, India

<sup>2</sup>, Student, Department Of Computer Engineering, SND College of Engineering & Research Centre, Yeola, Maharashtra, India

<sup>3</sup>, Student, Department Of Computer Engineering, SND College of Engineering & Research Centre, Yeola, Maharashtra, India

<sup>4</sup>, Student, Department Of Computer Engineering, SND College of Engineering & Research Centre, Yeola, Maharashtra, India

<sup>5</sup>, Lecturer, Department Of Computer Engineering, SND College of Engineering & Research Centre, Yeola, Maharashtra, India

## ABSTRACT

*Now days, people share their views on social media. Many users post various types of post on social media. Some post are good so they can be posted but if the post is not good or some bad words are used then for this type of posting filtering is needed. As it is a social media so the information is being easily flowing to each and every person present on social media. In this paper skin detection algorithm is used for image posting and video posting. For text type of posting stop word removal algorithm is used. By posting image, video, text type of content, system will help to take a certain action on the post or to disallow the certain posting. The system also helps to find out the person who is responsible for certain post.*

**Keyword:** - Key Social networking, frameworks, cloud forensic, social network analysis.

## 1. INTRODUCTION

Online Social Networks is one of the most popular today to help interactive medium to communicate with each other. People share their views about specific thing and disseminate a considerable amount of human life information. For now a day's continuous communications of people, simply exchange of several types of content of post, the content of the post include free text ,image, audio, and video data. According to Face book statistics 1 average client creates 90 pieces of content each month, while more than 30 billion piece of content (web links, news stories, blog posts, notes, photo albums, etc.) are united each month on facebook [1]. A Bayesian based text classifier extracts metadata from the content of the message. The metadata provided by the classifier, together with data extracted from the common graph and users profiles, to implement the filtering, finally message will get filtered [2]. The vast and dynamic character of this information creates the basis for the employment of web content mining strategies aimed to automatically discover useful information dormant within the data. In earlier work, contrast enhancement is detected using the peak artifacts that appear in the digital imagery. However in case of post-processing operation such as JPEG compression, this work fails to detect the contrast enhancement in modified images. So, a new algorithm has been proposed to detect the contrast enhancement not only in uncompressed but also in JPEG compressed images. Algorithm proposed for global contrast enhancement detection in this paper is robust against the post processing operation such as JPEG compression[3].In this paper for image and video, skin

detection techniques is been used. In this method pixel-based skin detection algorithm is used and extracts skin regions from color images. The simplest method for separating skin pixels from non-skin pixels is to give thresholds for each channel in a color space. Examples are combinations of RGB color space. The main goal of the Skin Detection algorithm is to extract all skin areas from an original image. Information filtering systems are planned to classify a flow of dynamically generated information [4]. The main aim of this paper is the design of a system given that customizable content-based message filtering, image filtering and video filtering for Online Social Network. After entering the message on his/her page, the user tries to post a message [5].

## 2. LITERATURE SURVEY

### **[1]"Exploiting Social Networks for the guess of social and civil unrest"**

This paper presents framework which analyses the social media and predict whether the social and civil unrest will occur or not. In this paper the researcher uses the social snapshot framework (Huber's framework) which takes the snapshot of the social media at the given time. In this paper the framework is the combination of monitoring online social media and Digital Cloud Forensics.

### **[2]" A System to filter preventable Messages from Online Social Network User Walls"**

In this paper, we have presented a system to filter undesired messages from online server network walls. This work is the first step of a wider project. The early encouraging results we have obtained on the categorization process prompt us to continue with other work that will aim to improve the quality of classification. The second task involves the learning phase. Since the underlying domain is dynamically shifting, the gathering of pre classified data may not be representative in the longer term. The present batch learning approach, based on the preliminary collection of the entire set of labeled data from experts.

### **[3]"Text Mining Facebook condition Updates for Sentiment Classification"**

In this research paper, researcher focuses on the usage of text mining for sentiment classification. Illustration is performed on Tunisian user's statuses on Facebook posts during the Arabic Spring era. The aim was to extract useful information, about user's sentiments and behaviors during this sensitive and significant period. For that purpose, they propose a method based on Support Vector Machine (SVM) and Nave Bayes. They also create a emotion lexicon, based on the emoticons, interjections and acronyms, from extracted statuses updates.

### **[4] "Revealing picture Forgery through Image Manipulation Detection"**

In this paper, contrast enhancement is detected using the peak/gap artifacts that appear in the digital images. However, In case of post-processing operation such as JPEG compression, this work fails to detect the contrast enhancement in modified images. So, a new algorithm has been proposed to detect the contrast enhancement not only in uncompressed but also in JPEG compressed images. Algorithm proposed for global contrast enhancement detection in this paper is robust against the post processing operation such as JPEG compression.

### **[5] "SkinSheriff: A Machine Learning Solution for Searching Explicit Images"**

In this paper, we presented skin sheriff, a novel approach to detect pornography in arbitrary images. We evaluated the best performing skin detection algorithms and compared them to our approach for image and video skin filtering algorithm is used and also histogram is match.

### **[6]"Filtering Undesired Messages from Online Social Network:A Content Based Filtering Approach"**

This paper proposes a system that creates a content-based message filtering service for Online Social Networks (OSNs). Our system allows Online Social Network peoples to have a direct handle on the messages that

are posted on their walls. This is done by a rule-based system, that allows a user to customize the filtering criteria, which is to be applied to their walls, and a mechanism based classifier which can automatically create association labels for the support of our content-based filtering mechanism.

#### **[7]"Digital Image Alteration Detection using Advance Processing"**

In this paper, at the start, discusses finding of basic improvement and histogram equalization that is frequently applied on forged images and some challenge of copy-move forgery finding. After, put forward a process for image confirmation; this will beat the face up to of copy-move forgery finding should be planned. In the forged images, the image alteration may be made on the full image (global) or to the particular portions (local) of the image.

#### **[8]"Image Enhancement by Histogram Specification"**

In this paper, appearance and uniqueness of the histogram plays a most important role in searching the value of an image. Histogram Specification (HS) is an image enrichment method, where the far above the ground resolution images known as goal Images, histogram of the these images are coordinated with each other, then most excellent image will be chosen with high PSNR value to enhance the input image .

#### **[9]"Automatic Text Classification: A Technical Review"**

This paper gives details of the general method for automatic text categorization which includes steps such as pre-processing characteristic choice using a variety of statistical or semantic approaches, and modeling using proper machine learning methods. This paper also discusses some of the most important issue concerned in automatic text categorization such as selling with shapeless text, behavior huge amount of attributes, exploratory achievement of purely statistical pre-processing methods for text categorization v/s semantic and natural language processing based methods. Automatic text classification involves assigning a text document to a set of pre-defined classes automatically, using a machine learning methods. The categorization is usually done on the basis of significant words or *features* extracted from the text document.

#### **[10]"forensic estimation and reconstruction of a contrast enhancement mapping"**

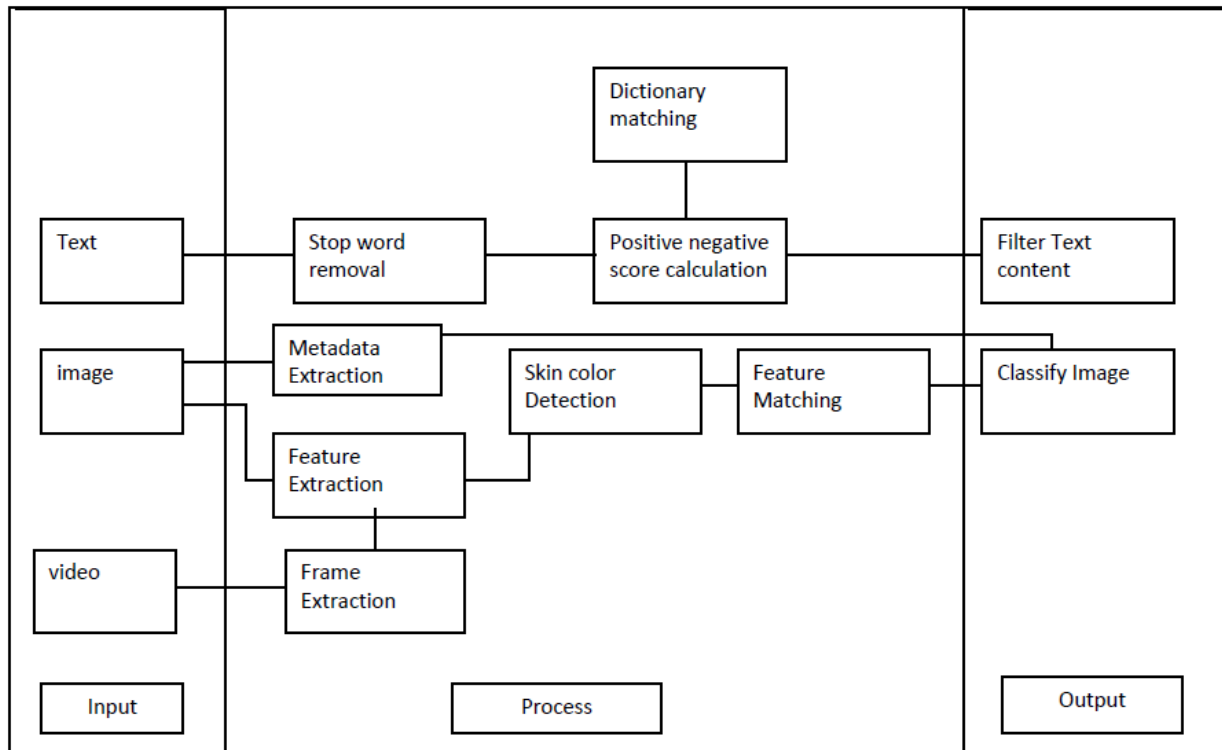
In this paper we recommend an iterative algorithm to together estimate any arbitrary contrast enhancement mapping used to adjust an image as well as the pixel value histogram of the image earlier than contrast enhancement. To do this, we use a probabilistic model of an image's pixel value histogram to determine which histogram entries are most likely to correspond to contrast enhancement artifacts. Experimental outputs are presented to express the valuable of our proposed method. Previous work has dealt with the forensic recognition of contrast enhancement in digital pictures.

### **3. PROPOSED SYSTEM**

Now a day, communication is increasing on social media .Users use more social network for communication purpose. In the current proposed system we are implementing the dictionary for text, on ly on the basis of contents of the users to compare with dictionary. For increasing the accuracy of the system, Corpus -based approach can be used along with the Dictionary based approach .Every individual person uses its own language for the purpose of communication. There exits certain words in the posts that gives us the complete meaning of that post. Those words are referred as keywords, which gives complete information about those posts. The words are compared with the dictionary and then those are the filtered according to keyword which are matched to dictionary of keywords in the dataset. In extra, it is possible to use texture, structural irregularities, images signatures or even trainable classifiers for pixel-based detection we determine to use such complex methods only later in the classification step, where they can actually act as features. The main contribution of this paper is the design of a system providing customizable content-based message filtering for online social network. As we have pointed out in the introduction, to the best of our knowledge, we are proposing such kind of application for online social network. Also for image and video type of posting, skin detection technique has been a key to get different recognitions like face recognition, human motion detection, image prediction, etc. Skin detection plays an vital role in tracking people, filtering out adult web images,

editing photo etc. More number of image processing models have been done practical for skin detection. The main goal of the Skin Detection component is to extract all skin areas from an image. We label all pixels separately and mark them in a binary image also for videos and images. Also in our system for video and image process we can see that from where the photo has been taken and from which mobile or computer that image has been send. It is also clear that whether the photo is edited or original.

#### 4. SYSTEM ARCHITECTURE



**Fig: System Architecture for filtering data**

Input:-

For our proposed system we will take three types of input text, image, audio and video.

Process:-

For text type of filtering, stop word removal algorithm is used with the help of stop word removal the common characters like . , \* + - / etc are removed from the post. For text we will maintain a dictionary as a dataset for detecting positive and negative words. The message of positive words and negative words matching will be displayed. After comparing the words with dictionary, if the word will go above the threshold value then the particular message will not be allow to post on social media.

For image filtering skin detection algorithm is used. This algorithm will give how much percentage of skin is open .If it will go above threshold value then that particular image will not be displayed on post. In extra metadata of image is also been extracted. Just like the image is taken from which cell phone, what is the version of it, whether the image is edited or no etc. After metadata extraction, if the image is in good level then it will be posted as it is on the social media.

For video type of filtering, frames of the video will be generated in the form of image and then again skin detection algorithm will be applied. If the video contains images like gun then for such type of video histogram matching algorithm will be used.

Output:-

At the end, a filtered text, image and video will be posted.

## 5. CONCLUSION

In this paper, we will proposed a system to filter the messages of type text, image and video from online social network. The flexibility of the system helps all users to be secure from unwanted post on social media. This paper presents the detection of bad or unwanted images, video and text type of posting. We evaluated the best performing skin detection algorithms and compared to previous algorithm .The skin detection algorithm is used for image to robust against the previous algorithm. So, the proposed algorithm overcomes the limitations of previous approaches. From this paper, if post includes text, image or video type and if the post will not harm anyone's emotion or ego then and then only the post will be posted on media..In this way, this system is useful to detect the person who is always posting unwanted images, text and video type of posting. So that we can stop his / hers posting on social media permanently.

## REFERENCES

- [1] Benkhelifa, Elhadj. "Exploiting Social Networks for the prediction of Social and Civil Unrest:A Cloud based Framework." International Conference (2014): 565-572.
- [2] Vanetti, Marco. "A System to Filter Unwanted Messages from OSN Walls." IEEE (2013): 285-297.
- [3] Jalel Akaichi. "Text mining facebook updates for sentiment classification." IEEE (2014).
- [4] Charpe, Jayshri. "Revealing Image Forgery through Image Manipulation Detection." IEEE (2015): 723-727.
- [5] Platzter, Christian. "Sheriff: A Machine Learning Solution for Detecting Explicit Images." International Journal (2014): 1-11.
- [6] Vashistha, Dhruv. "Filtering Undesired Messages from Online Social Network:A Content Based Filtering Approach." International Journal (2014): 2046-2049.
- [7] Patil, Mahesh Mahipati. "Digital Image Alteration Detection using Advance Processing." International Journal (2015): 18-21.
- [8] Gayakwad, Shweta B. "Image Enhancement by Histogram Specification." JATIT Journal (2014): 2347 - 2812.
- [9] Dalal, Mita K. "Automatic Text Classification: A Technical Review." International Journal (2011): 37-40.
- [10] C., Matthew. "forensic estimation and reconstruction of a contrast enhancement mapping." IEEE (2013): 1698-1701.