A Hybrid Scheme for Detecting Fake Accounts in Facebook

M. Smruthi, N. Harini

Abstract--- A social networking service serves as a platform to build social networks or social relations among people who, share interests, activities, backgrounds, or real life connections. A social network service is generally offered to participants who registers to this site with their unique representation (often a profile) and one's social links. Most social network services are web-based and provide means for users to interact over the Internet. Nevertheless these sites are also constantly preyed by hackers raising various problems related to threats and attacks such as disclosure of information, identity thefts etc. One of the most common ways of performing a large-scale data harvesting attack is the use of fake profiles, where malicious users present themselves in profiles impersonating fictitious or real persons. An attempt has been made in this work to use a hybrid model based on machine learning and skin detection algorithms to detect the existence of fake accounts. The experimentation process clearly brought out the strength of the proposed scheme in terms of detecting fake accounts with high accuracy.

Keywords--- Social Media, Facebook, Privacy, Social Network Analysis, Fake Profiles, Machine Learning, Skin Detection.

I. INTRODUCTION

Online social media are today a standout amongst the well-known Media for conveying, sharing human life and data. It holds network for business, pleasure and for other points in between. Informally speaking, o*nline Social media* are one of the most disruptive communication platforms with high socio-economic value. The media is regularly used by billions of users to interact, and they are key platforms for (among others) content and opinion dissemination, social and professional networking, product recommendations, scouting, alerting, and conducting political campaigns. Online Social networks give people a platform to communicate, share their ideas, to showcase their creative works, to prove their talent and to find new people of same interest. Popular online social Medias include Facebook, Twitter, LinkedIn, etc. Statistical reports reveal that social media have become a daily necessity for most of the human beings particularly for the teenagers.

Along with many advantages, online social media has got many disadvantages too. Some of the major issues with the online social media include privacy breach attacks, viral marketing, and malware attacks. Privacy breach attack is caused when users of online social media exposes their personal information in the online social platforms and then the compromising of these platforms helps attackers to get their private data and can be misused. Online Social Media is used as a platform to spread malicious software. Common ways to spread malware through online social media include indulgence with third parties, interaction among profiles, etc.

Another major issue with the online social media is the existence of fake accounts. A fake account represents the profile of people who guarantees to be somebody else. With 1.44 billion monthly active users and a potentially rich source of information, Facebook is one of the largest online social network in the world. Fake accounts are of two types (i) False accounts and (ii) Duplicate accounts. Duplicate account means that the user creates an extra account while having his own principal account. False accounts are also of two type's (i) Undesirable accounts and (ii) Misclassified accounts. User-misclassified accounts represent the personal profiles created by users for a business, organization, or nonhuman entity such as a pet (Facebooks terms of service permits such entities as a Page rather than a personal profile). On the other hand, undesirable accounts are the user profiles that are intended to be used for purposes that violate Facebooks terms of service, such as spamming.

The rest of the paper is organized as follows: Section 2 includes findings from the literature review done. Section 3 discusses the proposed architecture which gives the detailed explanation of the architecture, Section 4 presents the results obtained in the experimentation process using the proposed model to identify and remove fake accounts and finally section 5 provides an insight on the conclusions and the future scope.

II. LITERATURE REVIEW

Popularity of Online Social Media

Online Social media has met with enormous energy with the new generation first, now the social culture has become part of all the age groups. Initially the utilization of online social networks were limited to corporates and businesses for associating with clients, peers, customers, etc. [3]. Online Social Media offers open door to access information, recordings, augmentation of social gatherings, opportunities for learning, searching and maintaining companions and relatives. Online Social Media is a good method for connecting individuals with common interests [2]. The regular expansion of online social media clients has also resulted in the growth of information available in online social media.

Security Risks in Online Social Media

Online Social network has become a part of society's daily life. The increasing demand of the online social media has also lead to increasing black marks in online social

Manuscript received February 01, 2019

M. Smruthi, Amrita School of Engineering, Amrita Viswa Vidyapeetham, Coimbatore, Tamil Nadu, India. (e-mail: cb.en.p2cse17027@cb.students.amrita.edu)

Dr.N. Harini, Amrita School of Engineering, Amrita Viswa Vidyapeetham, Coimbatore, Tamil Nadu, India. (e-mail: n_harini@cb.amrita.edu)

media in the form of presence of fake accounts, viral marketing, privacy breach attack, etc. Privacy breach is caused when the user provide unlimited amount of their personal information in online social media which helps attackers to get the information easily from the online social network and misuse it [4]. Fake account means that somebody is using some identity which doesn't have any relation with any real person [5]. The fake accounts create many problems like bullying other persons, unwantedly trolling, using abusive words, using vulgar images for many illegal activities targeting youth, etc. People make fake profiles also for criticizing a targeted person, spreading wrong information, giving wrong reviews either good or bad for a public figure or a particular brand. After a long period of observance it was noted that teenage group of human beings are badly influenced by the online social networks [2]. They find it really hard for separating bad or good. So for a better result we tried out two approaches namely classifying fake accounts from the real user accounts with the user information data and detecting vulgar images from the user accounts as most of the vulgar images came from the fake accounts.

Machine Learning

According to the Literature many work use supervised machine learning algorithms to detect the fake accounts [5]. Numerous works are present in detecting fake accounts of Twitter, since it is very easy to get the datasets from Twitter. Supervised learning algorithms take dataset as input and with taking one value from the dataset predicts the other dataset values. The dataset used generally are collected after a long period of observance. A total of five machine learning algorithms namely K-nearest Neighbor, Support Vector Machine, Naïve Bayes', Decision Tree and Random Forest algorithms.

Image Processing

Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This procedure is commonly used as a preprocessing step to find regions that potentially have human faces and limbs in images. Skin image recognition is used in a wide range of image processing applications like face recognition, skin disease detection, gesture tracking and human-computer interaction [10]. A skin detection algorithm has been used to detect the vulgar images from the user accounts [11]. A deep learning method has been implemented to detect the image whether it contains human beings [13]. The image that contains human beings are undergone a skin detection where the percentage of skin present in the image will be calculated.

III. PROPOSED SCHEME

Many social networking platforms are available for use although the basic functionality of all the online social media are same. There are many existing solutions for the fake account detection in Twitter. Since the dataset collection is a tedious task in Facebook very less work has been done to detect the fake accounts in it. So the work proposed in this paper discusses results obtained by using Facebook features to detect the fake accounts.

In order to identify an account as real or fake one requires some real life datasets. The experimentation process used dataset collected using the Facebook API graph and also other fields which were collected from our own neighborhood because of strict privacy concerns. After a long period of observance and interaction with various account holders in Facebook, the accounts were identified as fake or legitimate manually.

The identified profiles were went through in detail to understand the features that classified the account as fake or legitimate. For the first part of the proposed work (LHS of the diagram in Figure 1) requires some features of the user profiles to analyze it as a real user account or fake user account. The features that were selected for detection of an account as fake or real were fed to supervised machine learning algorithms namely KNN algorithm, Support Vector Machine, Naïve Bayes' Algorithm, Decision tree and Random Forest and the accuracy were calculated from all these algorithms.

The images collected from the manually identified fake accounts were fed into the skin detection algorithm and the percentage of skin present in each image was calculated and it was seen that the percentage of skin exposed in images collected from fake accounts were higher than that of the real accounts. With the combined approach of the skin detection

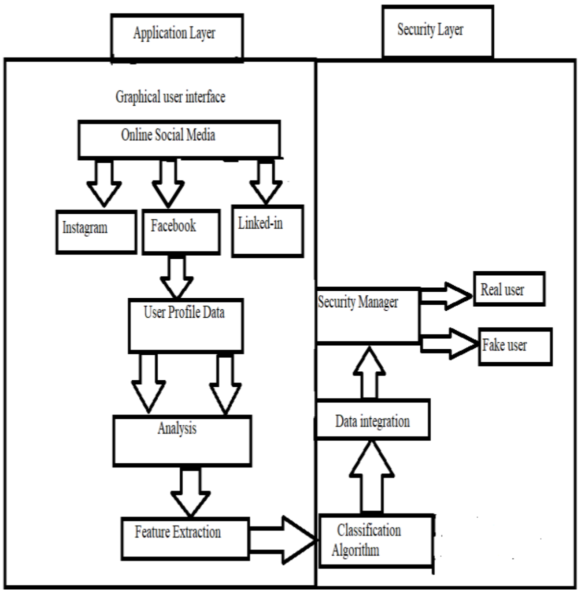


Fig. 1: Architecture of Proposed System

IV. RESULTS AND DISCUSSIONS

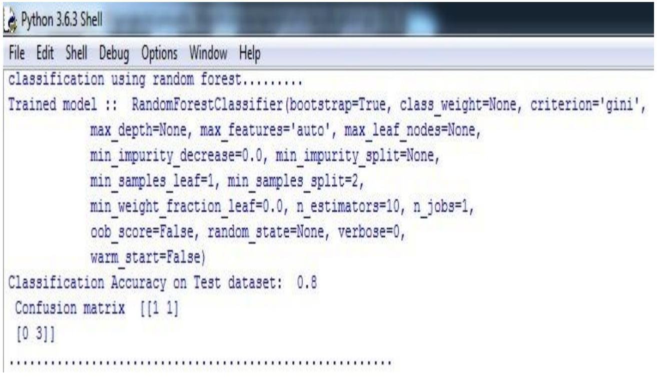
This section describes of how the evaluation was done to detect an account as real or fake. All the five supervised machine learning algorithms were applied to the mixed datasets of fake accounts and real accounts. A total of 400 datasets were used which comprises of 200 manually identified fake accounts and 200 real accounts which is from our own social neighborhood. The accuracy of different supervised machine learning algorithms were calculated.

Along with this a skin detection algorithm was used and the percentage of skin exposed in each image collected from the fake accounts and real accounts were calculated. In supervised machine learning algorithms a 10 fold cross validation were applied. The dataset contained manually identified 13 features of user accounts which included:

Table 1: Features extracted for Experimentation

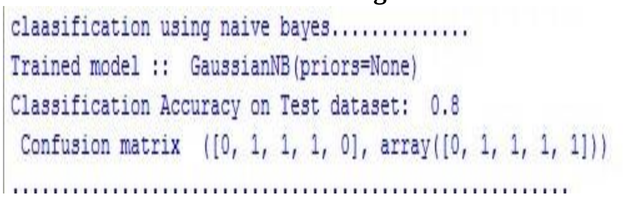
Name	Description	Explanation	Measured way
Post Count	Average number of posts created by the user in his own timeline	Fake accounts are expected to have low count of posts in their accounts	Number of post can be calculated from their timeline
Comment Count	Average number of comments received for all the posts.	Fake accounts are expected to share and post unwanted materials which results in lower comments.	Comment count is collected from the user's timeline.
Google image match	The display pictures matching with the google images count.	Fake accounts usually use the images downloaded from the google	The profile picture and the google image match was done.
Presence in other social media	Presence of user in other social medias were checked	Fake accounts are created by fake identities, so it is expected no presence of fake accounts in same name in other social media.	The user profile names were checked in other social media namely Instagram, Linked-in, Twitter, etc.
Memory Count	Average number of the memories posted or tagged in posts.	Fake accounts are expected to have low or no memory posts shared by them or tagged.	The number of memory posts were counter from user timeline.
Usage of Safeguard	The privacy settings for the images were noted.	Fake accounts don't use their identity and are expected not to use their own images, so relatively the protection usage by fake accounts were low	The privacy settings were checked for the images from the user timeline.
Check-in's	The places visited by the users were marked as checked-in	Fake users use fake identities and doesn't want to relieve their current location	The number of check-in were collected from the user profile.
Tagged-Post	The average number of posts that tagged them by other users	The number of tagged posts were comparatively less for a fake user	The average number of tagged posts were taken from user timeline.
Reviews	The average number of reviews given by a user	The fake users are expected to give very less reviews comparatively to real users	The average number of reviews given were noted from user timeline
Self-Information	The information about a particular user given by themselves	Fake profiles come with fake identities so the information provided them very less comparative to real accounts	The self - information provided by the user were noted down by checking out the user accounts
Friendsversary	The post celebrating the anniversary of when they became friends	Fake profiles doesn't care about celebrating friendships.	The number of friendsversary posts are taken out from users' timeline.
Followers	The number of followers following a particular user	The fake profiles will be having low followers count.	The number of followers are taken from their timeline.
Events	The number of events that person has attended in his life	The fake profiles are expected not add any events they attended.	The number of events attended were collected from users' timeline.

When the five supervised machine learning algorithms were applied, an accuracy of almost 80% were obtained with decision tree classifier which is the highest among all the classifiers and the rest of the classifiers gave almost 60-80% accuracy with an error rate of 20%.



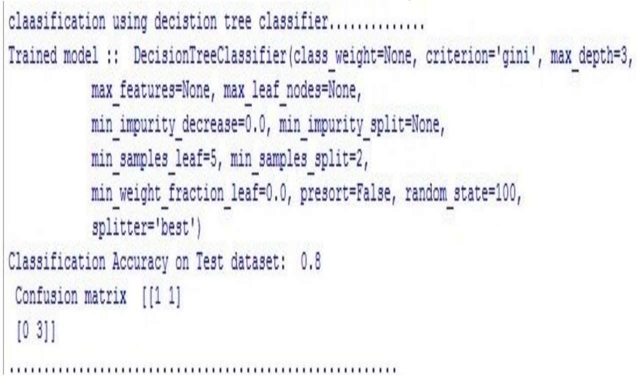
```
classification using random forest.........
Trained model ::  RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
            max_depth=None, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=10, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
Classification Accuracy on Test dataset:  0.8
 Confusion matrix  [[1 1]
 [0 3]]
.........................................................
```

Fig. 2: Accuracy calculation of fake accounts using Random Forest Algorithm



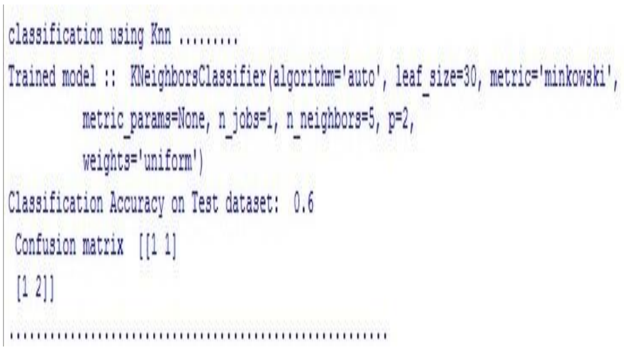
```
claasification using naive bayes..............
Trained model ::  GaussianNB(priors=None)
Classification Accuracy on Test dataset:  0.8
 Confusion matrix  ([0, 1, 1, 1, 0], array([0, 1, 1, 1, 1]))
.........................................................
```

Fig. 3: Accuracy calculation of fake account detection using Naïve Bayes' Algorithm



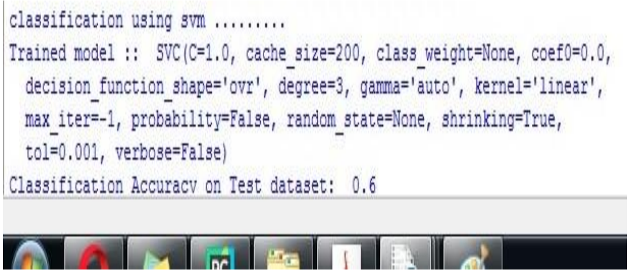
```
claasification using decistion tree classifier..............
Trained model ::  DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=3,
            max_features=None, max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=5, min_samples_split=2,
            min_weight_fraction_leaf=0.0, presort=False, random_state=100,
            splitter='best')
Classification Accuracy on Test dataset:  0.8
 Confusion matrix  [[1 1]
 [0 3]]
.........................................................
```

Fig. 4: Accuracy calculation of detection of Fake accounts using Decision tree classifier



```
classification using Knn .........
Trained model :: KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
           metric_params=None, n_jobs=1, n_neighbors=5, p=2,
           weights='uniform')
Classification Accuracy on Test dataset:  0.6
 Confusion matrix  [[1 1]
 [1 2]]
```

Fig. 5: Accuracy calculation of detection of Fake accounts using KNN algorithm



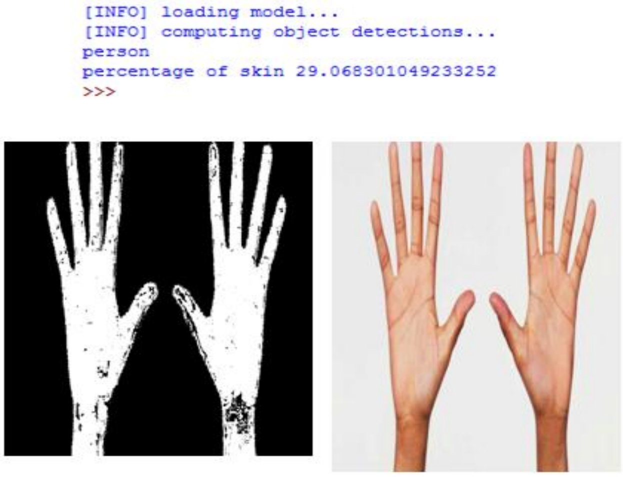
```
classification using svm .........
Trained model ::  SVC(C=1.0, cache_size=200, class_weight=None, coef0=0.0,
  decision_function_shape='ovr', degree=3, gamma='auto', kernel='linear',
  max_iter=-1, probability=False, random_state=None, shrinking=True,
  tol=0.001, verbose=False)
Classification Accuracy on Test dataset:  0.6
```

Fig. 6: Accuracy Calculation of detection of fake account using SVM classifier.

The above five figures give the accuracy of detecting fake accounts using the supervised machine learning algorithms. With the new features of Facebook added and used for the detection of fake accounts an increase in the accuracy of detection was noted. The accuracy has increased up to 80% which is shown in the figure 2, 3 and 4 when the three algorithms namely random forest algorithm, Naïve Bayes' and decision tree classifier were applied. The accuracy has gone up to 80% with an error rate of 20%. With the support vector machine classifier and the KNN algorithm the accuracy has been observed is 60%. Even though the accuracy rate is low the false positive rate is also very low with 10%. This shows that the new features added in the Facebook when considered for the detection of fake accounts has given a higher result than before.

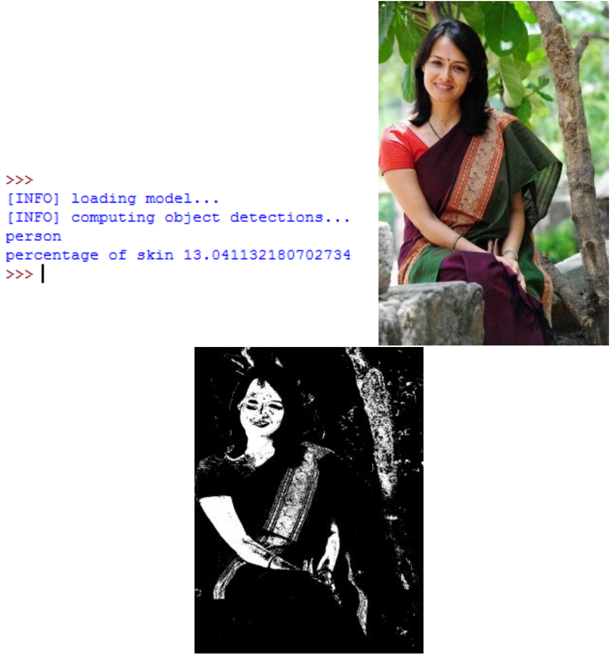
For the skin detection algorithm, a deep learning algorithm was first implemented to identify the image contained human or not. When the image detects human being in it, then the skin detection algorithm is applied and it calculates the percentage of skin present in it. A skin color matching algorithm is applied to calculate the percentage of skin in the image. Most of the images collected from the identified fake accounts were having a high percentage of skin in it greater than 13%. The images taken from the manually identified real accounts had less percentage of skin exposed in their images.

The Fig (7) shows the skin percentage calculated for a naked hands. Since the pixels will be same for a naked hand and a full nude human picture, the result is shown using a naked hand picture where the skin percentage calculated is 29%. The Fig (8) shows the skin percentage calculated for the image from a real account and the result acquired is 13%. Thirteen percentage can be set as a threshold for the normal skin percentage exposed in an image. For all the images first the image was tested to identify whether it contain human beings or not, if and only if the image contained human being the algorithm will calculate the skin percentage exposed in the image. Figure nine is an example for the algorithm where the image contained image of a dog and the algorithm din calculate the skin percentage in it where as it calculated for the Fig (7) and Fig (8).



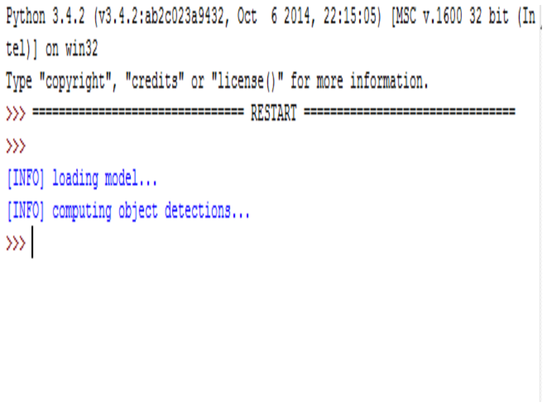
```
[INFO] loading model...
[INFO] computing object detections...
person
percentage of skin 29.068301049233252
>>>
```

Fig. 7: Skin percentage computed for nude hands



```
>>>
[INFO] loading model...
[INFO] computing object detections...
person
percentage of skin 13.041132180702734
>>> |
```

Fig. 8: Skin percentage calculated for image from real account



```
Python 3.4.2 (v3.4.2:ab2c023a9432, Oct  6 2014, 22:15:05) [MSC v.1600 32 bit (In
tel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>> ============================== RESTART ==============================
>>>
[INFO] loading model...
[INFO] computing object detections...
>>> |
```

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Fig. 9: Tried Calculating Skin percentage of a dog

V. CONCLUSION AND FUTURE SCOPE

Fake accounts are being continuously evolving in online social media. Therefore, it is very essential to invent new methods to detect Fake profiles in online social media. So the real time Facebook dataset were required to detect the fake accounts and vulgar images in Facebook. For the detection of Fake accounts the user timeline information namely post-count, comment-count, etc. were used and for the vulgar image detection the images from the user time line and the display picture of the users were taken out. The performance were evaluated using the supervised machine learning algorithms and the highest 80%accuracy were obtained and the maximum percentage of skin exposed were calculated from the images collected from the fake accounts. For the future scope, a more complex algorithm for the skin detection can be implemented. The natural language processing techniques can be implemented to detect fake accounts more accurately. The new features will be certainly introduced by the Facebook, and these features can also be included while analyzing the fake accounts.

REFRENCES

1. Wani, Suheel Yousuf, Ahmad Wani, Mudasir and Ahmad Sofi, Muzafar. *Why Fake Profiles: A study of Anomalous users in different categories of Online Social Networks*. International Journal of Engineering, Technology, Science and Research Vol-4 (September 2017).
2. Bhardwaj, Akashdeep, Goundar, Sam and Avasthi, Vinay. *Impact of Social Networking on Indian Youth-A Survey*. International Journal of Electronics and Telecommunications Vol-7(September 2017).
3. Persia and D. D'Auria, "*A Survey of Online Social Networks: Challenges and Opportunities,*" 2017 IEEE International Conference on Information Reuse and Integration *(IRI)*, San Diego, CA, 2017, pp. 614-620.
4. H. Gao, J. Hu, T. Huang, J. Wang and Y. Chen, *"Security Issues in Online Social Networks,"* in IEEE Internet Computing, vol. 15, no. 4, pp. 56-63, July-Aug. 2011.
5. A. Gupta and R. Kaushal, *"Towards detecting fake user accounts in facebook,"* 2017 ISEA Asia Security and Privacy (ISEASP), Surat, 2017, pp. 1-6.
6. MR, Neethu.; HARINI, N. " *Safe sonet: a framework for building trustworthy relationships*". International Journal of Engineering & Technology, [S.l.], v. 7, n. 2.26, p. 57-62, may 2018. ISSN 2227-524X.
7. E. Van Der Walt and J. Eloff, *"Using Machine Learning to Detect Fake Identities: Bots vs Humans,"* in IEEE Access, vol. 6, pp. 6540-6549, 2018.
8. Neethu M.R. and Harini N. "*Securing Image Posts in Social Networking Sites".* In: Hemanth D., Smys S. (eds) Computational Vision and Bio Inspired Computing. Lecture Notes in Computational Vision and Biomechanics, vol 28. Springer, Cham (2018)
9. M. Priyadharshini, V and Valarmathi, A. "Breast and nipple line localization for adult image identification in online social networks". 10.1109/ICETETS.2016.7603006 (February 2016).
10. Y. Lei, W. Xiaoyu, L. Hui, Z. Dewei and Z. Jun, *"An algorithm of skin detection based on texture,"* 2011 4th International Congress on Image and Signal Processing, Shanghai, 2011, pp. 1822-1825.
11. Tan, Wei Ren, Chee Seng Chan, PratheepanYogarajah, and Joan Condell.: "A fusion approach for efficient human skin detection", Industrial Informatics, IEEE Transactions on 8, no. 1,138-147(2012)
12. Patil, Prajakta M., and Y. M. Patil, "Robust Skin Colour Detection and Tracking Algorithm", International Journal of Engineering Research and Technology Vol. 1. No.8 (October-2012), ISSN: 2278- 0181 (2012).
13. X. Zhou, W. Gong, W. Fu and F. Du, "Application of deep learning in object detection," *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, Wuhan, 2017, pp. 631-634.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*