| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 6, June 2020||

The Detection of Selfie Image from Customized Dataset using Deep CNN

Danave Bharati Murlidhar

Asst. Professor, New Arts, Commerce and Science College, Ahmednagar, India

ABSTRACT: To categorize the selfie images from a large number of images is a very tedious task in image processing. Most of the existing researchers have already developed similar kind of applications for image classification. To extract various image features engine train models for classification in supervised learning algorithms. Sometimes few machine learning algorithms still having issues of misclassification of wrong image instances. In this work, we propose selfie image detection from heterogeneous image data set using a deep learning approach. A Convolutional Neural Network (CNN) is a model for learning a synergy function in the Common head and shoulder orientation subspaces, derived from by Local Binary Pattern (LBP) and Oriented Histogram Features of the gradients (HOG), respectively. This synergy was captured through the projection of the above features using the features Correlation. We show that the networks that lead are SIFT-captured spatial key point convolutional activations in the neighbourhood are discriminatory for self-detection. In general, the suggested method aids in capturing present intricacies in the image data and possesses the potential for use in other subtle conditions for image processing aside from self-detection. We study and analyze the performance of the popular CNN (GoogleNet, Alexnet) architectures, used for other images Classification tasks if the task of detecting the selfies on Web Multimedia. Performance of suggested These common architectures is compared to approach on a Ninety thousand images file consisting of roughly equal Amount of selfies as well as non-selfie images. In the experimental analysis, we demonstrate how the proposed system provides better classification accuracy then numerous machine learning algorithms. The system provides around 99% classification accuracy for various data sets.

KEYWORDS: Selfie detection, image classification, DCNN, Deep Learning, Local Binary Pattern, image processing, Canonical Correlation Analysis

I. INTRODUCTION

Detection of selfie images from large data is very tedious problem in image processing. To tacklewe ought to resolve this issue the also there impending questions: What is characteristic of a selfie image? How to doIn recent literature computer vision algorithms builtFor certain activities perform the function of self-identifying as a surrogate?Can we design and track the algorithm to scal it up?For bigger problems and more subtle ones? An attempt was made toMade to address those questions in this work. In most sections ofCases, human beings tend to classify an image asSelfie, by noting the self-portrait taker 's subtle poses.Consider, for example, the selfie in most of selfie image can be easily seen that the strong visual indications for inferring a selfie areConnection between the direction of the shoulder-arm (IndicatedBy the red vector) and the direction of the head gaze (indicatedBy Vector Blue). A short survey consisting roughly ofFifty people, from different backgrounds, were taken toVerify that other people hold a common view, too. It turned its backOut of that most people, 36 to be precise, were considered by hand and head orientation for classifying or grouping the given factor selfish or non-selfish pictures. So, an attempt was estimated head and elbow alignment, made in this respect and our entire approach consists of three stages; followed by synergies or similarity test inTheir direction and eventually their restricted learning understanding model this measure of synergy. In this work we proposed CNN base selfi detection technique using deep learning approach. We have evaluated around 500 test instances and detect the classification using CNN. In result section we demonstrates the proposed system effectiveness and accuracy with some existing systems.

II. LITERATURE SURVEY

Classification is a systematic arrangement based on its features, in groups and categories. Classification of images arose to reduce the gap between computer vision and human vision by training the computer with the data. The grading of the image is achieved by differentiating the image into the prescribed category based on the vision content. Motivation by[1], the study of image classification using deep learning is explored in this paper. The traditional methods used for the classification of images are part and piece of the Artificial Intelligence (AI) area formally called as machine learning. Machine learning consists of an extraction module of features that extracts important features such as edges, textures etc. and a classification module that classifies based on the extracted features.

The main limitation of machine learning is that it can only extract certain set of features on images while separating, and can not extract differentiating features from the data training set. This disadvantage is rectified through the use of profound learning[2]. Deep learning (DL) is a sub-field to machine learning which can learn through its own computing method. A deep learning model is used to persistently break down knowledge with a homogeneous framework such as how determinations will be made by a person. To achieve this, deep learning uses multi-algorithm layered structure represented as an artificial neural network (ANN). The architecture of an ANN is simulated with the aid of the human brain's biological neural network. This makes deep learning more able than the standard models of machine learning.

Given these handcrafted features, we intend to learn a multilayer convolutional neural network architecture that will take into account the synergy of these features in the common subspace. A Deep convolutional Neural Network consists of a series of convolutionallayers interwoven by layers of pooling and standardization. Following the convolutionallayers are full-connected layers interleaved with dropout layers to avoid overfitting. Those CNNs help in automated feature learning and classification when trained end-to - end. We use Alexnet architecture[3] in this approach, and use transfer learning by finetuning the model pre-trained on a broad Imagenetdataset[4].

An ablative comparative analysis is conducted with the other unconstrained CNNs to further examine the efficiency of our method. A simple experiment is performed to test this fact, to ensure that the training process is tractable and scalable to different problems. We pass the same test data to the network with the above trained CNNs, but with a slight modification to the input image. Using [5] we detect the face and shoulder, and set those pixel values to zero. A subset of 3000 of these images blocked from the face-shoulder was selected and accuracy was tested. This system can be seen as the unconstrained network shows no significant drop in accuracy (< 10 percent), implying that the networks are learning features that could potentially come from unreliable sources such as background, and are neither tractable nor applicable to subtle image analysis problems. However, a substantial decrease in accuracy from 86.3% to 58.8% of the proposed restricted approach suggests that a separate synergy-based training model is applied and tractable, implying that the features are learned from the accurate and adaptive head and shoulder orientation and can be adjusted to other possible subtle problems. Because SVM 's non-CNN model on synergy function in the previous experimental setup does not produce competitive results, we exclude it in this experimental environment. Tractability is also the key goal of this experimental setup and is more important in evaluating it on strategies based on CNN

Although negatives or non-selfies required for classification could have been compiled by arbitrarily collecting images that do not meet the selfie definition, including abundantly available images such as landscapes, animals and vehicles, the augmented images were chosen with more specificity to present a more demanding and meaningful setting. It consists mainly of images with at least one person performing a task (action dataset), posing for a picture (third person clicked on it) and so on. Most of the non-selfie images were taken from the Imagenet [4] database, the rest were collected manually over time under the hierarchies of people and objects. A novel dataset has thus been compiled which can serve as a good benchmark for testing algorithms.

The initial step of our algorithm is the capture of the head features and the shoulder alignment functions. The [6] method is used in which multi-level Local Binary Pattern (LBP) provides head alignment features and multi-level gradient histogram (HOG) provides head and shoulder alignment features. LBP is used for face and head detection tasks, because it is very successful. But our architecture is general, and also applies to other head apps. Additionally, HOG is measured as complementary to LBP[12], which provides both head and shoulder alignment. In this section, we briefly revisit the method of extraction of features in[6] that we used in our experiments. All images are converted to grayscale and redimensioned to 227 / 227.

Gradient Histogram (HOG)[7] is a local descriptor capturing gradients of the pressure. To obtain the magnitude of the gradient a mask consisting of a vector $[-1, 0, 1]$ is used. The image's resulting gradient magnitudes are divided into blocks of $h = 2l\{l = 1, 2, 3,. \}$, where l is Hierarchy level. In our experiments l is chosen to be 4. In every step of the hierarchy, the gradient magnitudes are voted into 9 bins of $0-180$ degrees of equal size separately from each of those image blocks. Lastly, the resulting histograms are concatenated in order to obtain the final descriptor λ.

As mentioned earlier, handcraft features fail to be learned as standalone descriptors when there is a higher level of séance. A restricted CNN training scheme is then implemented in such a way that all benefits are handcraft features and a deep learning approach is used. In the face verification scenario in[8], this method of constraint learning was used.

A custom loss function was modelled in [9] for foreground segmentation, where the CNN outputs were constrained by image level tags. In our approach, CNN loss is modelled as the Less squared loss function in the continuum of HOG and LBP between both the types of networks and the synergistic feature for both the neck and body, including both. Visualization tools and actual outcomes in that such a teaching method is successful by utilizing hand - crafted features, trying to make the model solvable. In addition, this also gives a more generalized approach beyond the problem of self-detection. Finally, the restricted CNN is used as feature pools, and features are extracted at keypoints detected by SIFT from learned convolution maps.

In [10], a total of 85,000 selfie images were collected from selfeed.com using a real-time # selfie update on integral and further manually annotated with various attributes such as age, gender and hair colour. Doing so results in 15,290 images being eliminated. Additionally, pictures of individuals that were either entirely meaningless or general photographs were omitted too. The resulting selfie dataset obtained mainly has single-faced selfies. To further diversify the dataset, additional images were compiled from online and offline resources consisting mainly of multiple-person selfies and also those clicked using equipment such as selfie sticks, apart from the above. Thus the resulting dataset has one chunk of robust data (i.e. positive or selfish), up-to - date images (i.e. positive or selfish) with the current trend of "selfie".

III. RESEARCH METHODOLOGY

Devise a solution for distinguishing self-portrait images with reference to complex features of selfie for high probabilistic accuracy of detection using computer vision techniques and deep learning.We evaluate the proposed system with some synthetic as well as real time image dataset and validate the accuracy of system. This research we design and develop a system which extracts multiple regions of each image to train the data using CNN model. Then we extract various complex features like localization as well as color, texture and shape from input images using proposed CNN. System alsodemonstrates multiple convolutional as auto encode to get best optimize features and improve the classification accuracy.

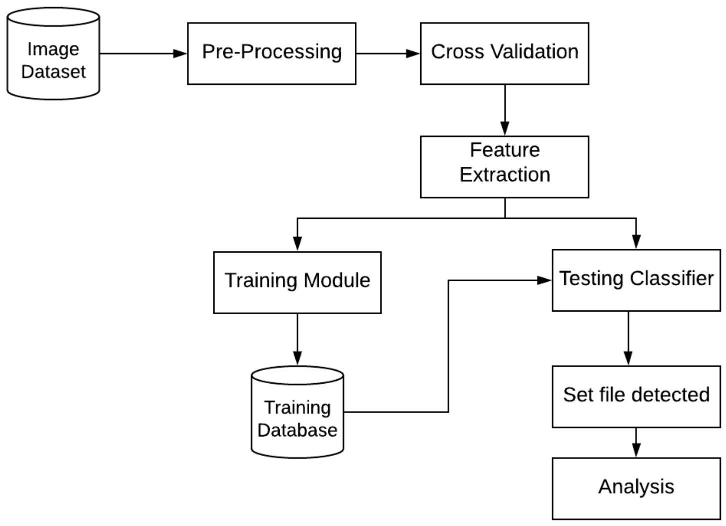

Figure 1 : Proposed system architecture

- System initially deals with large image dataset which contains some selfie as well as some normal images.
- Cross validation has done with 10 fold and 15 fold respectively.
- Apply noise filtration algorithm to remove the existing noise and improve the quality of image, parallel we also normalized the image with fixed dimensions.
- Various features extraction has done using CNN with multiple convolutional layer and train the module.
- Once module has trained system creates Background Knowledge (BK) for entire training dataset.
- For testing dataset feature extraction has similarly done like training module and test all dataset using CNN.
- System work with minimum two convolutional layers during the execution and ReLu as activation function, ReLu provides better accuracy than sigmoid and TanH.

Dataset Description

A novel dataset is to be created that can serve as a good benchmark for testing the algorithm, which contains various kind ofselfie images as well as some non-selfie images for analyzing almost correct accuracy of the algorithm proposed. This dataset can be collected from various sources like Selficity (selficity.com), Selfeed (selfeed.com), Ucf Centre of research in computer vision (crcv.ucf.edu), and Kaggle. Non-selfie images can be collected from Imagenet database, Kaggle, and INRIA Persons Dataset. For real time images it has created manually using smartphones or survey or extraction from social networking cities. University of Florida, Center of Research in Computer Vision Dataset has also used which contains around 46,836 selfie images. It is an annotated with 36 different attributes divided into several categories like:- Gender, Age, Face shape and gestures, Hair color and shape, Accessories, Lightning conditions.

IV. RESULTS AND DISCUSSIONS

After the successfully implementation of entire system we evaluate confusion matrix for overall experimental analysis, below Table 1 and Table 2 demonstrates the confusion matrix evaluation and Figure 2 and 3 shows accuracy of system with existing algorithms.

Table 2 : confusion matrix calculation using proposed system

Class	Selfi	Not-selfie
Selfi	145	2
Not-selfie	2	340

Table 3 : Performance evaluation with NB and SVM

	SVM	**CNN**
Accuracy	0.9892	0.9925
Precision	0.9867	0.9897
Recall	0.9933	0.9963
F-Score	0.9899	0.9929

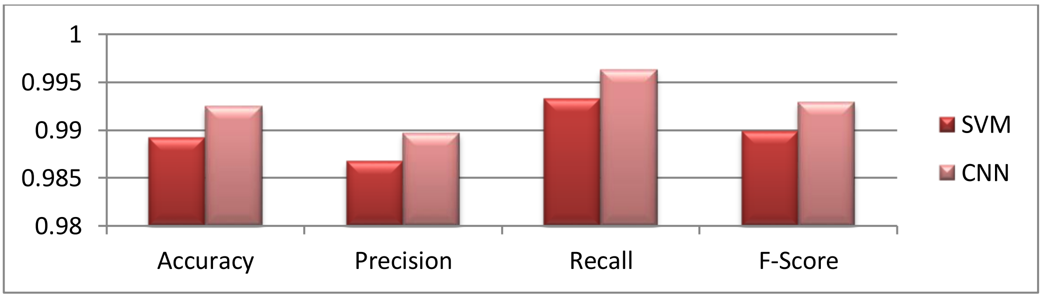

Figure 2 : Performance evaluation using Naïve Bayes and SVM

According to both experiment analyses, SVM shows better classification accuracy than the NB algorithm which is shown in figure 2. Conferring to the above experiment analysis, we can conclude the proposed system produces better accuracy for trust computation in IoT in-service environment. The entire research follows some simulation environmental parameters as well as a combination of machine learning algorithms. Various computation parameters have been used cluster differentiation and id.mi.com using respect to machine learning algorithms.

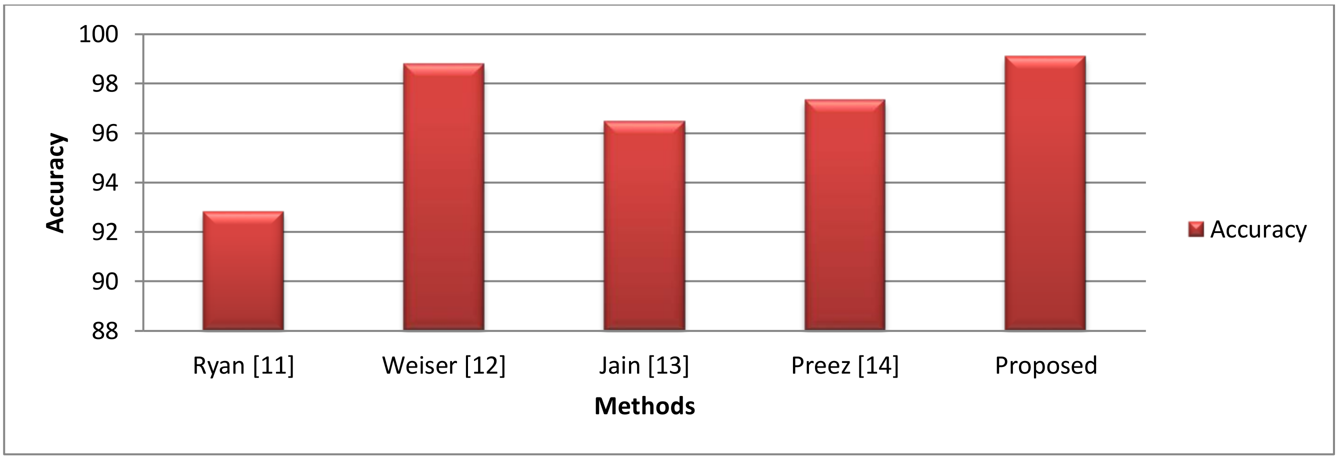

Figure 3 :Selfie detection accuracy of proposed system vs. various existing approaches

V. CONCLUSION AND FUTURE WORK

Systems results show that deep learning does provide promising results with a performance comparable to some methods using handcrafted features on emotion classification task, and also a few methods using deep learning for sentiment analysis. The important features of selfie image can be collected extensively using Image Processing , Computer Vision methodologies and can be classified, when mapped with deep learning techniques for training paradigm for obtaining features which are discriminative for selfie detection, enhances detection with more precise features.Large scale implementation of selfie detection algorithms which contains more than two features and merged with multiple other techniques can be the major part of future scope aegregation of groupies from selfies can the other important part.To evaluate the similar system on large heterogeneous dataset including some deep features will be the interesting task for future work.

REFERENCES

[1] https://in.mathworks.com/matlabcentral/fileexchange/59133 - neural -network -toolbox -tm--model -for -alexnet -network
[2] H. Lee, R. Grosse, R. Ranganath, and A.Y. Ng. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th Annual International Conference on Machine Learning, pages 609–616. ACM, 2009
[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information processing systems, 2012, pp. 1097–1105
[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in Computer Vision and Pattern Recognition, 2009. CVPR 2009.IEEE Conference on. IEEE, 2009, pp. 248–255
[5] X. Zhu and D. Ramanan, “Face detection, pose estimation, and landmark localization in the wild,” in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 2879–2886.
[6] C. Zeng and H. Ma, “Robust head-shoulder detection by pca-based multilevel hog-lbp detector for people counting,” in Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010, pp. 2069–2072.
[7] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), vol. 1. IEEE, 2005, pp. 886–893
[8] M. S. Sarfraz and R. Stiefelhagen, “Deep perceptual mapping for thermal to visible face recognition,” arXiv preprint arXiv:1507.02879, 2015.
[9] D. Pathak, P. Krahenbuhl, and T. Darrell, “Constrained convolutional neural networks for weakly supervised segmentation,” in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1796–1804.
[10] M. M. Kalayeh, M. Seifu, W. LaLanne, and M. Shah, “How to take a good selfie?” in Proceedings of the 23rd Annual ACM Conference on Multimedia Conference. ACM, 2015, pp. 923–926
[11] Weiser E. B., \#Me: Narcissism and its facets as predictors of selfie-posting frequency. In: 2015, Personality and Individual Differences, Elsevier. p: 477 - 481. Doi: 10.1016/j.paid.2015.07.007.
[12] Ryan T., \&Xenos S, Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. In: 2011, Computers in Human Behavior, Elsevier. p: 1658 - 1664. Doi: 10.1016/j.chb.2011.02.004.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 6, June 2020||

[13] Jain M. J. \&Mavani K. J., A comprehensive study of worldwide selfie-related accidental mortality: a growing problem of the modern society. In: 2017, International Journal of Injury Control and Safety Promotion, Taylor \& Francis. p: 544 - 549. Doi: 10.1080/17457300.2016.1278240.

[14] Preez A. D., Sublime selfies: To witness death. In: 2017, European Journal of Cultural Studies, Sage. p: 1 - 17. Doi: 10.1177/1367549417718210.