International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Dual Shot Face Detecting using Deep Learning

Nishath Ansari, Suresh Dara, Jaala Shruthi

Abstract: ***In the paper, we have used a deep learning technique to identify dual faces i.e. nothing but detecting dual shot faces. As the data is emerging day by day with high dimensionality, recognizing dual faces is a major problem. So wasting time on identifying images is like fiddling around. In order to save time and get absolute accuracy we have implemented a fast preprocessing technique named as Convolutional Neural Network (CNN) along with feature extraction technique which is used to knob the relevant features to detect and identify images/faces. By performing this robust method, our intention is to detect dual images in an efficient way. This technique results in decreased feature cardinality and preserves unique efficiency of the data. The experiment is performed on extensive well liked face detecting benchmark datasets, Wider Face and FDDB. CNN with FE demonstrates the results with superiority and the accuracy was in-depth analyzed by CNN classifier.***

Keywords: Deep Learning, CNN, Feature extraction, face detection, feature subset and Dual Shot.

I. INTRODUCTION

In this emerging population, identifying faces has owned remarkable attention in computer technology in the field of face analysis. Detecting faces can be considered as identifying size and location of any human/animal faces in the format of digital pictures [1]. Detecting faces is the initial steps in face recognition. The process of detecting faces using digital images is quite difficult because of the discrepancy available across human/animal faces for instance skin tone, expressions, features, posture, using glares, hairs along with variability with the camera pixels, weather conditions like light and resolution size of pictures. These days face recognition technique is extensively used for automatic surveillance of videos at malls/roads, authentication for identification, pattern recognition and so on [2].

Till now, the previous work involved in face detection has given a satisfactory outcome in certain conditions for single images; however, this method should be effectively applied on the dual shot faces as well [3]. Therefore, it is substantial to a supplementary inspection for face detection for dual shot faces. Latterly, different works has been done in the area of recognition and detection of faces which gave accurate results, but using CNN alone with feature extraction can be a revolution in the field of detecting dual faces at a time and results in highly efficient and accurate results.

At the moment, deep learning is an evolving research area in machine learning and computer vision [4], whose work has been excellently introduced for reduction of image dimensionality [5], and recognition as well [6]. Deep learning comes from the artificial neural networks concept whose structure comprise of multi-layer perceptron (MLP) of many hidden layers in it [7]. With the help of multi non linear transformation model deep learning algorithms finds excellent features in the data and results in high level accurate objects. A correlation between different images can be detected using deep learning algorithms [8].

In dual shot, detecting faces is the major step in face recognition which comes under a part of object detection. Detecting objects can be widely used in many fields like personal safety, security of people, biometric information, enforcement of law, entertainment, crime safety and so on. Many old methods with discriminative results have solved this issue [9] [10], but is quite difficult to recognize two 300-pixels dual faces then those of recognizing 30 pixels single face. To overcome this problem, we do use a deeper convolution network. So, by including many layers of convolution which may lead to accurate results of detecting faces with discrimination in color, texture, pose, background, moments and so on. When we look on the other side, by digging deeper into layers, there is a possibility of losing spatial information and features. Distribution of pictures in dual shot as frontend and backend is done based on the scores generated [11], whereas frontend size of an image can be around 20-50 pixels and backend face as the negative image can be same as frontend, due to this a lot of work is required to detect tiny objects. This problem has been notified and partially solved with the use of dilation operation and decreasing count of pooling operations; these methods has been widely used in applications of computer vision. A recent face detection technique likes SSD [12] (Single Shot multi box detector), and faster RCNN [13, 14] has given good results in the era of detecting faces. So now it is the turn of detecting dual faces using a fast CNN with the help of feature extraction technique.

Feature extraction comes under the process of reducing dimensionality and extracting relevant features because increase in dimensionality leads to inappropriate results of identifying images [15]. So, by considering feature extraction technique, it will be better to extract or select some features which do posses high relativity and helps in reducing redundancy [16]. By processing lot of computations on raw digital images, it reduces and results in more manageable set of features to identify the pictures.

Manuscript received on March 15, 2020.
Revised Manuscript received on March 24, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
Nishath Ansari*, Computer Science Department, B.V. Raju institute of technology, Narsapur, Telangana, India, Email: nishat.alm@gmail.com
Suresh Dara, Computer Sscience Department, B.V. Raju institute of technology, Narsapur, Telangana, India. Email: darasuresh@live.in
Jaala Shruthi, Computer Science Department, B.V. Raju institute of technology, Narsapur, Telangana, India. Email: jshruthi503@gmail.com

After extracting some features from the original data set, which can be called as feature subset [17]. This feature subset completely and accurately describes the original dataset. So the method of extracting features is used to eliminate redundant features, select relevant ones without losing the importance of the required image features of the dataset [18].

II. BACKGROUND

There are many deep learning algorithms have been noticed for face reorganization and detection problem [19] [20]. Considering this, NN (Neural Networks) [21] is the most popular recycled one for image processing. Neural networks are used by computers to learn about the data by the observational data while Deep learning is considered as a powerful method for processing neural networks [22]. Presently, with the combination of neural networks and deep learning results in the excellent solutions to emerging issues in natural language processing, face detection, image recognition and so on [23]. In this paper, we used CNN (Convolutional Neural Network) for dual face detection with the help of feature extraction technique.

A. An Overview of CNN

At the present era, CNN Convolutional Neural Network is considered as excellent technique for face detection among deep learning approaches [24]. We presented an adequate algorithm for detecting faces and a clear illustration of CNN to results with an accurate computation. We do use 6 CNN's, where 3 are used for detecting faces and 3 are for non faces for binary classification and bounding box calibration which are used for multiple class classification pattern [25]. CNN framework has improved its performance with the use of feature extraction and softmax computation is performed. In this paper, our propose algorithm is composed of two parts, one is CNN and other is feature extraction. Feature extraction is used for selecting relevant and high accurate features whereas CNN is used for classification and saying the accuracy of that images recognized.

III. PROPOSED METHOD FOR CNN WITH FEATURE EXTRACTION

Feature extraction is nothing but extracting features. So to extract features, firstly we need to train our CNN network model. By considering dual shot faces, initially we do train our CNN model with respect to target values with sigmoid/logistic dense layer. The main reason behind training the CNN is to identify the exact density for the model by processing number of forward and backward iterations which results in reducing misclassification cost of the images [26]. In this article, the framework used is keras and backend is developed using tensor flow. Instead of extracting features based on texture, pose and so on, here CNN extracts images based on training raw pixels size and then learns the data from training model, then extracts the features and ultimately results in objects that required [27]. Now with the help of trained model, directly we can extract appropriate features of a dual image at any intermediate dense layer with required dimensions. So, based on training classification mode [28], features extraction which can be considered as feature subset is identified and results in high effective results for all dual shot images. Figure 1. illustrates the overview of our model, i.e., identifying faces with the help of CNN and feature extraction network.

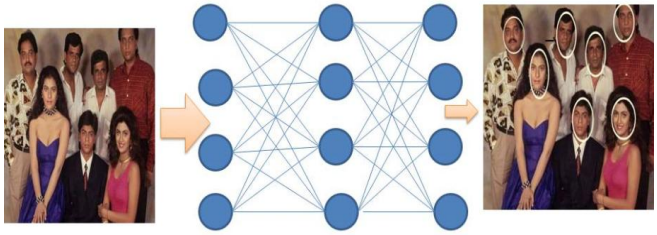

Fig 1. Illustrates overview of process

A.DATASETS

CNN along with feature extraction has been implemented so that we can easily recognize faces. This test can be performed on two datasets, they are:

1. The most popular datasets in image processing are wider face [29], and FDDB [30]. In this paper, we are using these both in dual sense images. Wider face dataset usually consist of publically available images which are dual shot with high possibility of variability in space, texture, pose , brightness, appearance, color and so on [31]. This dataset consists of around 393,703 labeled images/faces bounding boxes in 32,203 pictures. The total number of events constitute of 61 classes. From these training data can be randomly selected as 30% for training, 20% for validation subset and 50% for test. Then the testing can be performed on the dataset. In our method dual images can be trained and subsets are identified and both the validation and testing can be performed. Figure 2 explains about the complete process of detecting images.
2. The other widely used publically available dataset is FDDB (face detection dataset and benchmark) [30]. Our dataset consist of in general 2955 images along with 5381 annotations on face. This is the most challenging dataset with at most complexity of identifying a face. And when it is shot dual, then the complexity increases by 2x. With new dual shot images in the dataset, which is complex when compared to previous FDDB images. In this dataset, 30%, 20%, and 50% data can be used to perform training and testing respectively.

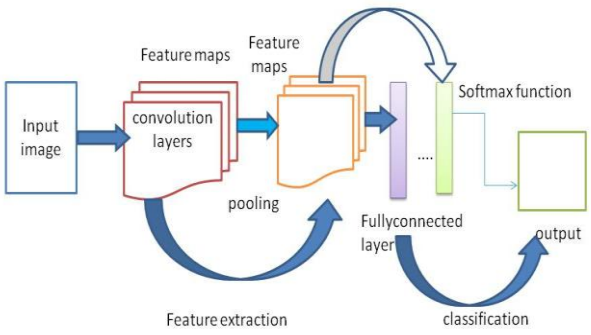

Fig 2. Workflow of CNN with feature extraction

B. CNN based Feature extraction algorithm

1. Input : Model,T, t // T = training dataset, t = test dataset
2. Output: tested dual image with extracted features
3. Start: Initialize neural n/w to create object with sequential class

I. Training classifiers for identifying dual images by adding all layers.
II .Model.add (xConvy), // x = number of filter, y = kernel, Conv = convolution extracts features from i/p dual image and represent pixel values 0&1 for performing operation.
III. Model.add (Relu),// activation funcn for transforming weighted i/p from node into activation of node o/p for that i/p it maps i/p value directly ,o/p value is -ve funcn maps to 0.
IV. Model.add (xPy), //x = pooling size, y= stride, P = pooling layer to detect features in dual images and reduce dimensionality of image.
//adding convolution and pooling layers as many we want for feature extraction.
V. Model.add (Flatten) , //this funcn flattens all feature maps to single column.
VI. Model.add (xFC), // x= number of neurons, FC = fully connected layer flatten matrix into vector combines features to create model.
VII. .Model.add (NC), //NC = number of classes
VIII. Model.add (Softmax), //activation function to classify output as two outcomes.predicts probabilities range 0 to 1 and sums to 1 and performs classification model which has high probability.
IX. Model.compile (optimizer, accuracy measure)
X. Model.fit (T, epochs, validation-split) // epochs use to train cnn, validate dual image through test-set.

4. Stop

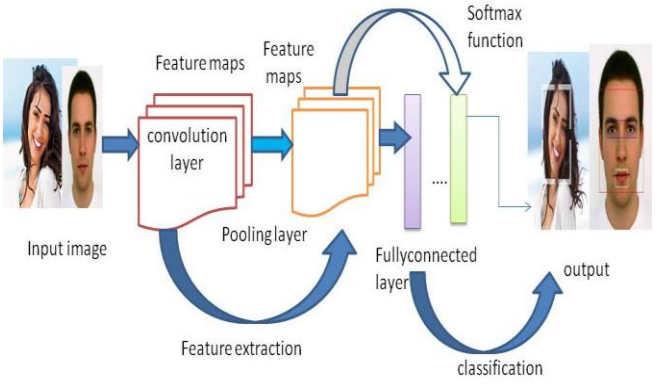

Fig:Feature extraction with CNN

Fig3. Performing classification on images and extracting relevant features.

Figure 3 gives the clear glance of complete workflow and explains each and individual steps clearly and results in relevant and robust features.

IV. EXPERIMENTAL RESULTS

When the input dual images are processed to convolution layer, feature maps, maps the features with the help of pooling layer and extracts relevant and efficient features. Then those features are processed to connected layers and works with softmax function [32], and then by performing classification, our model detects the entire features find the faces in dual shot images.

We used most popular open source dual faces datasets to prove or results. From WIDER FACE data set 30% is used for training the CNN model then 20% is used to validate the data then the feature extraction is performed with the help of soft computing function, then the remaining 50% of dual faces are used for testing the trained model. Our results are classified into three parts, which are easy, medium, and hard [33]. This helps in providing accurate results with different difficulty levels. This strategy produces awesome results of matching faces. While considering precision and recall, average precision around all the three parts of test are 97.1% for easy, 96.2% for medium set of images and 92.6% for hard data. And individual subset of validation is 96.7%, 95.5%, and 91.9% for easy, medium and hard respectively. The accuracy of wider face is improved from 50% to 95% from past to present [34].

In addition to wider face we do use another famous data set called FDDB. FDDB is particularly used for identifying exact annotations for regions of face. As we know changing the positions and format may result in decreased detection accuracy. So to say our model works efficiently, we need to check the model for different annotations, and then we can say this model is accurate and efficient for face detection. So in this model, we have taken 60% data for training and 40% for testing the accuracy of each feature in the dual images. The accuracy obtained with CNN and FE is 88%. Our model enjoys more effective and efficient way to amalgamate different networks, which helps us to use simple backend model and result high performance accuracy.

V. CONCLUSION

Detecting dual faces is a fundamental issue in emerging vision function. To solve the problem of identifying dual shot faces, in this work, we proposed a robust approached named CNN algorithm alone with feature extraction amalgamated into it to detect dual faces and boost predominant performance. By understanding drawbacks of previous methods, we proposed this method to overcome those drawbacks and results in more relevant, efficient and effective recognition of dual faces. Our experiment has been evaluated on the most admired datasets. Overall, we would like to conclude that with the help of this technique detecting dual faces can be easier and an efficient way without consuming much time. With this approach we demonstrate the superiority of our technique over the state of the arts. This deep learning technique provides robustness and effectiveness for faces. This innovative idea can be further applied to other image detection issues

REFERENCES

1. Hjelmås, Erik, and Boon Kee Low. "Face detection: A survey." *Computer vision and image understanding* 83, no. 3 (2001): 236-274.
2. Ban, Yuseok, Sang-Ki Kim, Sooyeon Kim, Kar-Ann Toh, and Sangyoun Lee. "Face detection based on skin color likelihood." *Pattern Recognition* 47, no. 4 (2014): 1573-1585.
3. Li, Jian, Yabiao Wang, Changan Wang, Ying Tai, Jianjun Qian, Jian Yang, Chengjie Wang, Jilin Li, and Feiyue Huang. "DSFD: dual shot face detector." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5060-5069. 2019.
4. Zhang, Cha, and Zhengyou Zhang. "A survey of recent advances in face detection." (2010).
5. Sarikaya, Ruhi, Geoffrey E. Hinton, and Bhuvana Ramabhadran. "Deep belief nets for natural language call-routing." In *2011 IEEE International conference on acoustics, speech and signal processing (ICASSP)*, pp. 5680-5683. IEEE, 2011.
6. Zhong, Sheng-hua, Yan Liu, and Yang Liu. "Bilinear deep learning for image classification." In *Proceedings of the 19th ACM international conference on Multimedia*, pp. 343-352. 2011.
7. Pal, Sankar K., and Sushmita Mitra. "Multilayer perceptron, fuzzy sets, classifiaction." (1992).
8. Ciregan, Dan, Ueli Meier, and Jürgen Schmidhuber. "Multi-column deep neural networks for image classification." In *2012 IEEE conference on computer vision and pattern recognition*, pp. 3642-3649. IEEE, 2012.
9. Yang, Shuo, Ping Luo, Chen-Change Loy, and Xiaoou Tang. "From facial parts responses to face detection: A deep learning approach." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3676-3684. 2015.
10. Yang, M-H., Narendra Ahuja, and David Kriegman. "Face recognition using kernel eigenfaces." In *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, vol. 1, pp. 37-40. IEEE, 2000.
11. Liang, Rui, Qingxin Zhu, Honglei Wei, and Shujiao Liao. "A video shot boundary detection approach based on CNN feature." In *2017 IEEE International Symposium on Multimedia (ISM)*, pp. 489-494. IEEE, 2017.
12. Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *European conference on computer vision*, pp. 21-37. Springer, Cham, 2016.
13. Jiang, Huaizu, and Erik Learned-Miller. "Face detection with the faster R-CNN." In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 650-657. IEEE, 2017.
14. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.
15. Nixon, Mark, and Alberto Aguado. *Feature extraction and image processing for computer vision*. Academic Press, 2019.
16. Hong, Zi-Quan. "Algebraic feature extraction of image for recognition." *Pattern recognition* 24, no. 3 (1991): 211-219.
17. Sun, Zehang, George Bebis, and Ronald Miller. "Object detection using feature subset selection." *Pattern recognition* 37, no. 11 (2004): 2165-2176.
18. McAllister, Patrick, Huiru Zheng, Raymond Bond, and Anne Moorhead. "Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets." *Computers in biology and medicine* 95 (2018): 217-233.
19. Zou, Xuan, Josef Kittler, and Kieron Messer. "Illumination invariant face recognition: A survey." In *2007 first IEEE international conference on biometrics: theory, applications, and systems*, pp. 1-8. IEEE, 2007.
20. Hancock, Peter JB, Vicki Bruce, and A. Mike Burton. "Recognition of unfamiliar faces." *Trends in cognitive sciences* 4, no. 9 (2000): 330-337.
21. Beale, Hagan Demuth, Howard B. Demuth, and M. T. Hagan. "Neural network design." *Pws, Boston* (1996).
22. Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.
23. Collobert, Ronan, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. "Natural language processing (almost) from scratch." *Journal of machine learning research* 12, no. Aug (2011): 2493-2537.
24. Li, Haoxiang, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. "A convolutional neural network cascade for face detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5325-5334. 2015.
25. Shin, Hoo-Chang, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning." *IEEE transactions on medical imaging* 35, no. 5 (2016): 1285-1298.
26. Chen, Yushi, Hanlu Jiang, Chunyang Li, Xiuping Jia, and Pedram Ghamisi. "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks." *IEEE Transactions on Geoscience and Remote Sensing* 54, no. 10 (2016): 6232-6251.
27. Chen, Yushi, Hanlu Jiang, Chunyang Li, Xiuping Jia, and Pedram Ghamisi. "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks." *IEEE Transactions on Geoscience and Remote Sensing* 54, no. 10 (2016): 6232-6251.
28. Sheppard, Jeremy M., and Warren B. Young. "Agility literature review: Classifications, training and testing." *Journal of sports sciences* 24, no. 9 (2006): 919-932.
29. Yang, Shuo, Ping Luo, Chen-Change Loy, and Xiaoou Tang. "Wider face: A face detection benchmark." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5525-5533. 2016.
30. Jiang, Huaizu, and Erik Learned-Miller. "Face detection with the faster R-CNN." In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 650-657. IEEE, 2017.Wang, Kai, Yuan Dong, Hongliang Bai, Yu Zhao, and Kun Hu. "Use fast R-CNN and cascade structure for face detection." In *2016 Visual Communications and Image Processing (VCIP)*, pp. 1-4. IEEE, 2016.
31. Bouchard, Guillaume. "Efficient bounds for the softmax function, applications to inference in hybrid models." In *Presentation at the Workshop for Approximate Bayesian Inference in Continuous/Hybrid Systems at NIPS-07*. 2007.
32. Chauhan, Aajad, Baizel Kurian Varghese, LM Abdul Rahman, Vishweswar Mohapatra, and Tapas Badal. "WIDER Face Challenge using Multi-Task Cascading Neural Network." In *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, pp. 188-192. IEEE, 2019.
33. Qin, Hongwei, Junjie Yan, Xiu Li, and Xiaolin Hu. "Joint training of cascaded CNN for face detection." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3456-3465. 2016.

AUTHORS PROFILE

Nishath Ansari, she is currently working as an Assistant Professor in the Department of CSE in BVRIT, Narsapur, India. She completed her M.Tech from B.V. Raju institute of technology, India 2019. At early age, she started her major interest in research-oriented problems and started publishing excellent research articles. Her major interest lies in Machine Learning, Deep Learning, Soft Computing, Bioinformatics, Data Science and IoT.

Suresh Dara, he is currently working as a Professor in the Department of CSE in BVRIT, Narsapur, India. He did his Ph.D. from the Indian Institute of Technology (ISM) Dhanbad, India in 2015 and M.Tech from Andhra University in 2008. He has published many good research articles in SCI, Scopus Indexed Journals. Also, he has been serving as a reviewer for many reputed SCI Journals. Now his research interests are Machine Learning, Deep Learning, Soft Computing, Bioinformatics and Data Science.

Jaala Shruthi, she is pursuing her M.Tech in B.V. Raju institute of technology, Narsapu, India. Her major interest lies in Machine learning and deep learning.