First International Conference on Computer Engineering
International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
Volume 5 Issue 8, November-December-2020

Face and Food Recognition Techniques using Convolutional Neural Network

Nidhi Shah[1], Vanshika Singh[1], Pranjal Kumavat[1], Samiksha Shirbhate[1], Dr. Meenakshi A. Thalor[1]
[1]Information Technology, B.E AISSMS's Institute of Information Technology, Pune, Maharashtra, India

ABSTRACT

There are different face recognition and food recognition techniques that have been studied and implemented in the past. As the world is moving in a fast pace security is an important aspect of a person's life. Apart from that healthy food has its own benefits. In this paper, various techniques for face recognition and food recognition have been listed along with their methodologies and accuracies. This paper proposes a model which will detect face and food from the given input images using various algorithms and then recognition of face and food item will be done.

Keywords: Face recognition, Food recognition, CNN, EigenFaces, Haar

I. INTRODUCTION

Face recognition is used in some of the high-tech security systems. With the use of videos or images and real time face detection helps in maintaining proper reports and also helps in security management. Maintaining a good diet starts with eating healthy food. So, recognizing the food items we are consuming on daily basis is also important. These processes are made automatic using different algorithms on various datasets. In the past, there have been various techniques for face detection to recognition. These techniques have evolved over a significant timeline. We have discussed different techniques like Eigen Faces, LBPH, CNN for face recognition. Similarly, in food recognition different techniques have been discussed. Each paper has used different datasets and has proposed a model with the highest achievable accuracy. In this paper we have presented a survey on different techniques used for face recognition and food recognition separately. Methodology and conclusions of these papers are explained along with the dataset used. Section 2 explores the literature review of these papers. In Section 3, we have discussed our proposed model and finally in the last section includes conclusion and future scope of face and food recognition.

II. LITERATURE SURVEY

A. Face Recognition Techniques

In [1], face detection and recognition are done on real time basis. Videos are taken as input, where frames are selected using an algorithm for further process. The Haar algorithm is used for face detection. This algorithm performs well for illumination changes. Recognition of faces is done using EigenFaces and Gabor Feature algorithms. EigenFaces algorithm performs well in complex computations and Gabor feature performs well when there are pose changes. The combination is proposed as a robust system. The accuracy of the system in recognizing a face using Eigen face is approximately 60%-65% and using Gabor feature is approximately 50%-60%. This system

will work well under limited conditions and dataset. Real time recognition can be achieved using other methods like CNN.

In [2] as well, the Haar method is used for face detection. A real time attendance system is proposed which detects faces of students using different modes like snap and video. It is built on Raspberry Pi and software module uses Open source computer vision libraries. The face detection here is done using the Haar Cascade method which classifies the input images as "with face" and "without face" images. It is a feature-based method. Recognition of the faces is done using LBPH method. Firstly, LBPs are obtained for the image and histograms are plotted which are then compared with the training dataset. Hence face is recognized whatever the input type is provided. The accuracy of this system is 74%. The accuracy has increased due to the use of EigenFaces for recognition. Another attendance system is proposed in [3],

which uses Haar face detection and PCA and EigenFaces for recognition of faces. The proposed system stores the records of the students in the database. The captured image undergoes the detection process. The process is carried out by checking the landmarks of the face using AdaBoost Algorithm. The recognition is done by Principle Component Analysis with EigenFaces.

In [4], Convolutional Neural Network (CNN) is described for face detection. CNN is used for image classification and object recognition as it works on raw pixel strength of image, which is given as input as flat vector. The collected data is filtered, split and labelled for training. Each image is converted to one fixed size and the model is trained. The ReLu activation function, set dropout proportion to 100% and first order optimization algorithm is used to improve the model efficiency. The accuracy of model is increased from 68.85% to 79.41% after the test set.

In [5], CNN technique on FPGA is used for face recognition. Generally, CNN is implemented using CPU and GPU. This paper proposes real time face recognition by using FPGA. The accuracy of the network is determined by the testing unit. The system built on FPGA consists of 6 CNN layers. [5] Fig. 1 shows the CNN layers. There are 4 hidden layers where there are 2 convolutional layers and 2 pooling layer. The activation function determines the sigmoid function. This function determines the error rate of the network. The maximum error seen in the network after training is of 0.001. The system claims to be 99.25% better than other systems built on CPU, GPU and even other algorithms like PCA and LBP. Network is trained and tested on face datasets having a range of different expressions, lights and other factors. The system shows good robustness. The only drawback of this is that there isn't much research done on face recognition using FPGA. It is also difficult to code a CNN algorithm on a FPGA as there is yet no mature research in this field. The hardware of this system can be a limiting factor.

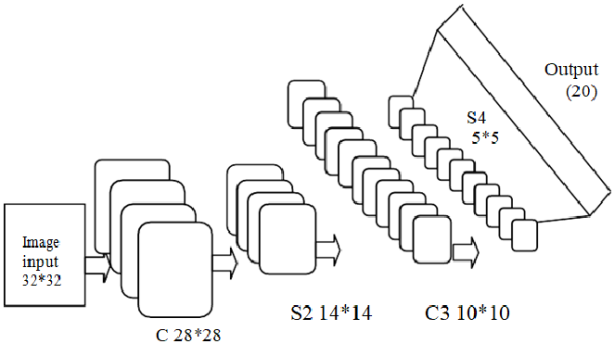

Figure 1: CNN of Face Recognition on FPGA

In [6] specifically focuses on face recognition when there are some obstacles in capturing the face. As in real time getting a clear image of the face and recognition can be difficult. This paper mainly focuses on recognition of face when it is partially visible. CNN is used for extracting different features from an occluded face image. It is done by weighted neurons which shred on a weighted plane. This decreases the computational complexity of the model. The recognition of the face is done by calculating the

Euclidean distance. Euclidean distance is calculated of the input image and the referenced image. It should be less than 1 if it is the same person. The triplet loss function is used in the paper for determining the degree of inconsistency. In this the dataset had 30% occlusion over the faces. The model proposed in this paper has a recognition rate of 98%.

Multiple layer model based real time face detection [7] works on CPU. Different ways are used are used to cover the face for training the model for real-time detection. This system works on CPU using different hardware. Multiple Images are detected using CNN architecture which uses shallow layers that produce light variations so the system can work as close to real- time. This system has built a fast face detection system. They have achieved this using Average Precision for measuring accuracy of the system which is 90.84%.This can be further improved by using pre-processing.

B. **Food Recognition Techniques**

In [8] deep convolution neural network (DCNN) is proposed to classify food images. Food image analysis has three steps: 1. image segmentation 2.food recognition 3.quantity analysis. In food analysis colour, intensity, texture, etc are important features. The model proposed consists of 54 layers.3 datasets have been used for training and testing of the model. In this approach the network is pre- trained using Inception v3.Feature extraction is done using this Inception v3. To improve the performance of this type of network increasing its size, depth layers, width makes a huge difference. The model shows 92.3% accuracy in food detection. The only drawback of this model is that even after using pre trained model the computational resource requirement is quite high.

This paper [9] uses ensemble network containing GoogleNet, ResNet, and Alexnet for food image detection. The model of this CNN has been trained and tested on 2 types of databases. One is the ETH-101 and the other is Indian food database. The feature extraction is done with the help of external pre-trained model. AlexNet and GoogleNet are used in the max pooling layer of the CNN module and ResNet is used in the last output layer. The accuracy of this model is 73% for Indian food database which in comparison to AlexNet, ResNet and GoogleNet individually is quite high. The hardware used in this experiment include NVIDEA processor with system requirement of 128GB of ram. the parameters used in the sub-network is different which adds to the complexity as reshaping the parameters according the sub-network increases load. As the architecture of each sub-network is different, it could be challenge to build such a network. The building of such a networking module could be difficult to code and debugging might be a challenge in future.

The paper [10] proposes a method for food image classification and furthermore how to determine its calories. Food recognition is done by 2 techniques namely, Graph cut segmentation and deep learning neural network. From these two methods CNN has proven to be more efficient comparatively in recognizing the food and hence providing its calorie count. The application in the paper calculates probability of the input image against the labelled data in the database hence predicting the food in the input image. Calorie measurement of the recognized food item is done by 2 methods.

1) Finger Based Calorie Measurement.
2) Calorie Measurement using Distance Estimation.

The application shows more than 95% accuracy in identifying the food and showing perfect calorie count. The database used to train and test the model contained images of only single food items.

The CNN approach proposed in paper [11] is built on a 5-layer network architecture. The first 4 layers are convolutional-pooling layers and last layer is a fully connected layer. The steps in recognizing the food

given in the input image are as follows: Feature Extraction, Shift and distortion invariance and Classification. This Fig. 2 [10] shows the system architecture used. Different layers of the CNN perform these steps and the output is provided. UEC-FOOD100 and an open source database are used in training and testing of model. The accuracy on single food item is 80.8% and 60% on multi- food item datasets.

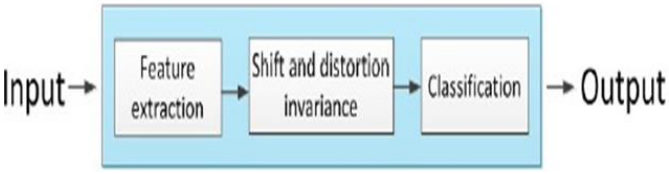

Figure 2: System Architecture

III. PROPOSED MODEL

In reference to above literature survey, CNN has performed well under all the different scenarios. We propose a system which will detect face and food from the given input images using algorithms and further recognize the face and food item. We will be training the system under supervised manner. The feature extraction and further classification will be done by CNN.

Face detection is done using MT-CNN algorithm which detects the faces in the given input image and crops it to 224 * 224 size array. This array is input to extended model of VGG-Face model which performs recognition of the faces and marks attendance of the person accordingly. The VGG- Face model is used after altering its output layer with respect to our dataset. The process of using the pre-trained model as a feature extractor for a new model is called transfer learning. Only the output layer is fine-tuned according to our dataset.

Food Recognition is done using VGG-16 model. This model is mainly used for object detection and classification of objects in a photograph. We will be using this model after fine tuning its output dense layers. We have created our own dataset of purely Indian dishes. Some of the images are taken from the internet and some are taken by us of some dishes which are common in our household. The models will be given input of the images which are uploaded on a website. The final output of these models will also be shown on the website itself. The output will also be stored in the database for future reference. Fig.3 show the simple architecture of proposed system.

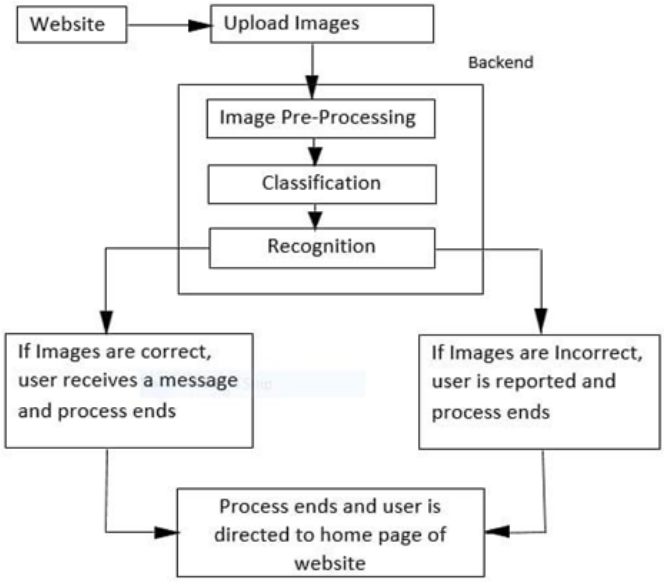

Figure 3: Simple Architecture of Proposed System

IV. CONCLUSION

In this review it is seen that CNN shows the best accuracy among the algorithms, compared for both face and food recognition. It can be implemented on medium database efficiently. It works efficiently despite the occlusion, pose, expression and light change in face recognition. Similarly, for food recognition colour, texture and multi-food items are recognized with high accuracy. As a whole these will help in making the society safer to place to live in and even healthier.

V. REFERENCES

[1]. S. V. Tathe ,A. S. Narote,S. P. Narote, " Face Detection and Recognition in Videos", IEEE. Bangalore Section, Institute of Electrical and Electronics Engineers. India Council, & Institute of Electrical and Electronics Engineers. (n.d.). 2016 IEEE Annual India Conference (INDICON) : 16-18 Dec. 2016.

[2]. V, T., Vellore Institute of Technology, Institute of Electrical and Electronics Engineers. Madras Section, IEEE Communications Society., & Institute of Electrical and Electronics Engineers. (n.d.). Conference proceedings, International Conference on Vision Towards Emerging Trends In Communication and Networking (ViTECoN 2019) : 30-31, March 2019, Vellore, Tamilnadu, India.

[3]. Shreyak Sawhney, Karan Kacker,Samyak Jain, Shailendra Narayan Singh, Rakesh Garg,"Real-Time Smart Attendance System using Face Recognition Techniques" 9th International Conference On Cloud Computing, Data Science and Engineering : Confluence 2019 : 10-11 January 2019, Uttar Pradesh, India.

[4]. Di Wang , Ding Wang , Hongzhi Yu , Guanyu Li " Face Recognition System Based On CNN" 978-1-7281-9837- 8/20/$31.00 ©2020 IEEE

[5]. Qu, X., Wei, T., Peng, C., & Du, P. (2018). A Fast Face Recognition System Based on Deep Learning. Proceedings- 2018 11th International Symposium on Computational Intelligence and Design, ISCID 2018, 1, 289–292. https://doi.org/10.1109/ISCID.2018.00072.

[6]. Gui Wu1, Jun Tao, Xun Xu " Occluded Face Recognition Based on the Deep Learning "IEEE 31st Chinese Control and Decision Conference (2019 CCDC) : 3-5 June, 2019, Nanchang, China.

[7]. M. D. Putro, Wahyono and K. -H. Jo, "Multiple Layered Deep Learning Based Real-time Face Detection," 2019 5th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia, 2019, pp. 1-5. doi: 10.1109/ICST47872.2019.9166172

[8]. Hassannejad, H., Matrella, G., Ciampolini, P., De Munari, I., Mordonini, M., & Cagnoni, S. (2016). Food image recognition using very deep convolutional networks. MADiMa 2016 - Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management, Co-Located with ACM Multimedia 2016, 41–49. https://doi.org/10.1145/2986035.2986042

[9]. Pandey, P., Deepthi, A., Mandal, B., & Puhan, N. B. (2017). FoodNet: Recognizing Foods Using Ensemble of Deep Networks. IEEE Signal Processing Letters, 24(12), 1758– 1762. https://doi.org/10.1109/LSP.2017.2758862

[10]. Pouladzadeh, P., Kuhad, P., Peddi, V. B., Yassine, A., & Shirmohammadi, S. (n.d.). Food Calorie Measurement Using Deep Learning Neural Network.

[11]. Zhang, W., Zhao, D., Gong, W., Li, Z., Lu, Q., & Yang, S. (2015). Food Image Recognition with Convolutional Neural Networks. 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), 690–693. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom- IoP.2015.139