Incept-N: A Convolutional Neural Network Based Classification Approach for Predicting Nationality from Facial Features

Masum Shah Junayed(✉), Afsana Ahsan Jeny, Nafis Neehal, Eshtiak Ahmed, and Syed Akhter Hossain

Daffodil International University, Dhaka 1207, Bangladesh
{junayed15-5008,ahsan15-5278,nafis.cse,eshtiak.cse}@diu.edu.bd,
aktarhossain@daffodilvarsity.edu.bd

Abstract. Nationality of a human being is a well-known identifying characteristic used for every major authentication purpose in every country. Albeit advances in application of Artificial Intelligence and Computer Vision in different aspects, its' contribution to this specific security procedure is yet to be cultivated. With a goal to successfully applying computer vision techniques to predict a human's nationality based on his facial features, we have proposed this novel method and have achieved an average of 93.6% accuracy with very low mis-classification rate.

Keywords: Nationality · Artificial intelligence · Computer vision

1 Introduction

Facial recognition is a complex process in which information and experience are used to set the average face for measuring other faces. The ability to detect face is very important in many aspects of life. It does not help us identify the people near us, but allow us to recognize those people so that we do not know that we can become more aware of possible dangers.

The human face is a very rich inspiration that gives amazing information for social communication adapted to people. For the last several decades, many attempts have been made in the biological, psychological, and cognitive science fields to understand, remember, and remember how humans are brainwashed [11].

The human face is a complex visual pattern that identifies the primary, specific, along with the general in-structure and the unrealistic. With this unique information, we can say that some aspects of the face are not employed in the person's face, but are shared by facial symptoms. These aspects can be used to make both known and unidentified people in general paranormal groups such as district or nationality.

K. C. Santosh and R. S. Hegadi (Eds.): RTIP2R 2018, CCIS 1036, pp. 466–475, 2019.
https://doi.org/10.1007/978-981-13-9184-2_41

With the development of computer technology and digital image processing technology, people start exploring the automatic national identification system through computers, this process is mainly compared to distance, angle, and other characteristics between people, and then people determine the picture and then the people [5]. The oldest method of identifying nationality is to look at living habits, metaphor formations and other characteristics of the people. This classification method is fully artificial, massive, and professionally staffed by professionals who have lots of resources for professional knowledge and experience.

In this paper, we use the technique of learning to review the Inception-V3 [8] model of tensor flow [1] in the datasets (China, Germany, India, Jamaica and Zimbabwe) in 5 countries. We meet an efficient national identity model using a short training time and achieve a high accuracy. The remaining paper is sorted in the following manner: Details of Convolutional Neural Network (CNN) and Inception-v3 [8] model are discussed in Sect. 2. The comparison with other papers is discussed in Sect. 3. Data collection and training are discussed in Sect. 4. Performance analysis is done in Sect. 5. Finally conclusion with some future work scopes is described in Sects. 6 and 7.

2 Background Study

Our work is based on the Inception-v3 [8] model of TensorFlow [1] platform and also used CNN [2].

TensorFlow [1] Like Google's artificial intelligence to learn artificial intelligence, it has received many interesting and presentable introductions for learning across the globe. TensorFlow [1] So far all the deep learning and machine learning programs got the first place. Tensorflow [1] has the advantage of superior facilities and higher facilities, and with the help of TensorFlow [1] researchers, TensorFlow [1] floatation capacity has been improved. To facilitate the use of researchers from various sectors today, TensorFlow [1] has opened many trained models on the official website.

Inception- V3 [8] is one of the TensorFlow [1] training models. This is a reconsideration of the initial framework for computer vision after Inception-V1 [9], Inception-V2 [9] in 2015. Intention-V3 [8] model is trained in ImageNet datasets, contains information that can detect 1000 classes in ImageNet. Inception-V3 [8] consists of two parts: a fully-connected and softmax layer [10] with classification properties and feature extraction part with a convolutional neural network (CNN) [2].

Convolutional Neural Networks (CNN) [2] is a class of neural networks that improve effectively in areas such as image classification and recognition. ConvNets have been effective to detect objects, face, and traffic. Generally, three main types of layers are used to create ConvNet architectures:

- Convolutional Layer,
- Pooling Layer and
- Fully-Connected Layer (Fig. 1).

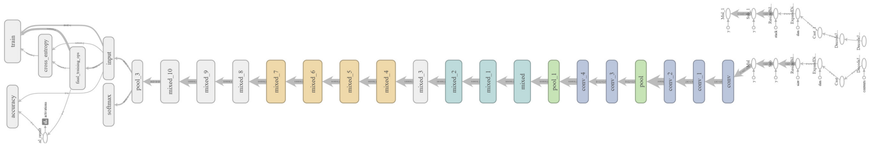

Fig. 1. Main graph of Inception v3 model.

The first layers of a CNN [2] strain (big) features can be accepted and illustrated relatively simply. As a result of the convolution of the neuronal network, the image is divided into emotion, made in local receptor file and compressed synthesis of the size m2 × m3 features map is briefly compressed. In this way, where this map stores the feature in the image and stores how much it is compatible. Therefore, in each file, it is applied to the location of the volume of trained spatial [2].

At each layer, there is a bank of m1 filters. The output characteristics are similar to the depth of the volume on the number of filters applied at one stage. Identifies a particular feature in each location of the input in each file. The output $Y_i^{(l)}$ of layer l consists of $m_1^{(l)}$ feature maps of size $m_2^{(l)} * m_2^{(l)}$. The i^{th} feature map, denoted $Y_i^{(l)}$, is computed as

$$Y_i^{(l)} = B_i^{(l)} + \sum_{(j=1)}^{m_1^{(l-1)}} K_{i,j}^{(l)} * Y_i^{(l-1)}$$

where $B_i^{(l)}$ is a bias matrix and $K_i^{(l)}$, j is the filter of size $2h_1^{(l)} + 1 * 2h_2^{(l)} + 1$ connecting the j^{th} feature map in layer $l - 1$ with i^{th} feature map in layer [2].

Later the layers detect increasingly (smaller) features that are more abstract (and are usually present in many of the larger features detected by earlier layers). The pooling layer l has two hyper parameters, the spatial extent of the filter $F^{(l)}$ and the stride $S^{(l)}$. It takes an input volume of size $m_1^{(l-1)} * m_2^{(l-1)} * m_3^{(l-1)} *$ and provides an output volume of size $m_1^{(l)} * m_2^{(l)} * m_3^{(l)}$ where;

$$m_1^{(l)} = m_1^{(l-1)}$$
$$m_2^{(l)} = (m_2^{(l-1)} - F^{(l)})/S^{(l)}$$
$$m_3^{(l)} = (m_3^{(l-1)} - F^{(l)})/S^{(l)} + 1$$

CNN's last level [2] input data allows a highly-specific classification by combining all the specific features detected by the previous level. It has an image of a specific degree translation, rotation and distortion image. It has made great progress in the field of painting [2].

If l − 1 is a fully connected layer;

$y_i^{(l)} = f(z_i^{(l)})\, with\, z_i^{(l)} = \sum_{(j=1)}^{m_1^{(l-1)}} w_{i,j}^{(l)} * y_i^{(l-1)}$

Otherwise (Fig. 2);

$$y_i^{(l)} = f(z_i^{(l)})\, with\, z_i^{(l)} = \sum_{(j=1)}^{m_1^{(l-1)}} \sum_{(r=1)}^{m_2^{(l-1)}} \sum_{(s=1)}^{m_3^{(l-1)}} W_{i,j,r,s}^{(l)} (Y_i^{(l-1)}) r_l^s$$

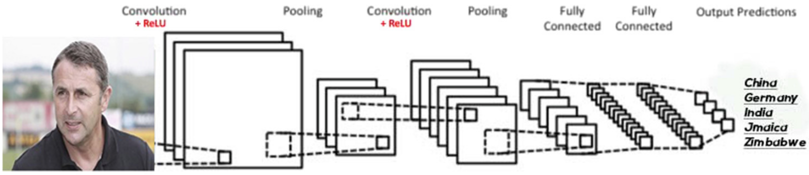

Fig. 2. Structure of Convolutional Neural Network (CNN).

TensorFlow [1] is available in the detailed tutorials. Repeat the final level of inspection for new sections using transfer learning. Learning movements is a new ma-chine learning process that can use existing knowledge from an environment and can answer another new problem, but it is related to the old problem. Measuring with the traditional Neural Network, it requires only fewer data for model training and high accuracy with minimal training time [5,7].

3 Literature Review

Inception v3 [8] model used in many experiments. Among them:

In 2017, Xiaoling Xia and Cui Xu used the transfer learning technique to retrain the Inception-v3 [8] model of TensorFlow [1] on the datasets of flower [11,13] of Oxford-I7 and Oxford-102 for Flower Classification. The accuracy of the classification of the model was 95% on the dataset of Oxford-I7 flower and 94% on the dataset of Oxford-I7 flower [5].

In 2017, Alwyn Mathew, Jimson Mathewa, Mahesh Govindb, Asif Mooppanb from bVuelogix Technologies Pvt Ltd. used Google's TensorFlow [1] deep learning a framework to train, validate and test the network for Intrusion Detection and the accuracy was 95.3%. But the proposed network is found to be harder to train due to vanishing gradient [3] and degradation problems [3].

In 2017, Brady Kieffer, Morteza Babaie Shivam Kalra, and H. R. Tizhoosh used CNN and Inception v3 [8] model for Histopathology Image Classification [6]. All experiments are done on Kimia Path24 dataset and the accuracy was 56.98% [6]. In 2017, Xiao-Ling Xia, Cui Xu, Bing Nan worked for Facial Expression Recognition based on the Inception-v3 [8] model of TensorFlow [1] platform. They used CK+ dataset [15] and selected 1004 images of facial expression. Their accuracy was 97% but it wasn't based on dynamic sequences [7].

In 2016, Bat-Erdene and Ganbat worked on Effective Computer Model for Recognizing Nationality from Frontal Image [4]. They used SVM, AAM, ASM and the accuracy was 86.4%. Their experiment was worked manually and images must be the frontal face image that has smooth lighting and does not have any rotation angle.

Our experiment is based on the Inception-v3 [8] model of TensorFlow [1] platform for Nationality Recognition based on facial features with Deep Learning. Nobody did it before. This is the first approach from us. It works automatically and has rotation angle and translations in the picture.

4 Methodology

Now in this section, we have described the following part is as follows: at first we make a flowchart [12] of our experiment; second, we provide a simple description on the dataset; third, we give about the data pre-processing; then, we describe the model installation; eventually, we discuss about the train model.

A flowchart [12] is a diagram that represents a workflow or process. Flowchart [12] connects boxes with the steps and arrows of the box and shows their order. Flowcharts [12] is used in a process analysis, design or management. The following diagrammatic representations illustrate the solution model of our system (Fig. 3).

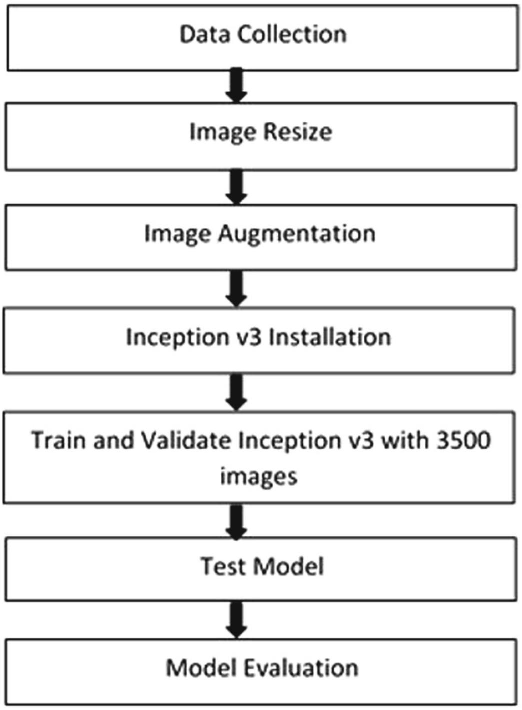

Fig. 3. Flowchart of the system model.

4.1 Dataset

There are many countries in this world with many people. There is a similarity in the outlook of human faces. For recognizing nationality, we have used 600 images of five countries for our experiment. They are China, Germany, India, Jamaica, and Zimbabwe.

4.2 Data Preprocessiong

Image promotion is a very important step to promote the effect of image classification. The Convention Neural Network's [2] learning process follows our activity monitoring and guidance on machine learning, so we need to label the data in the image pre-processing step. Then we have resized the data and also augmented (Rotate +30, Rotate −30, Translation, Lighting and Flip). Finally, we have found 3600 images for training (Fig. 4).

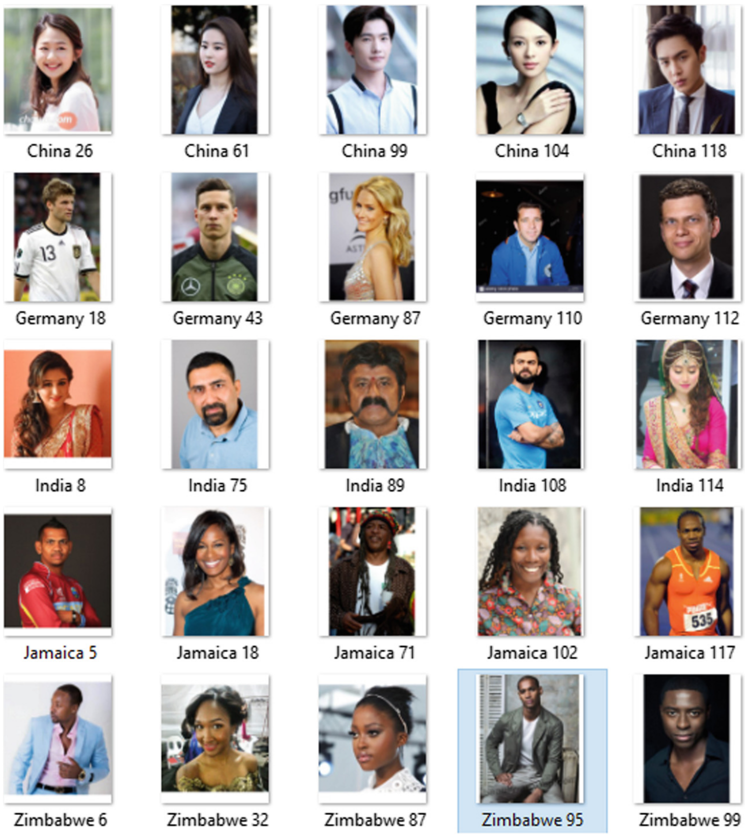

Fig. 4. The example of our dataset.

4.3 Model Installation

This test is based on the Tensorflow [1] platform's Inception-V3 [3] model. Processor 2 GHz Intel i3, Memory 4 GB 1600 MHz DDR 3, System Type: 64-bit OS, X-64 based processor.

After all, we've downloaded Tensorflow [1]. Then we've downloaded the Inception V3 [8] model. We also used the transfer learning process, which had previously set up the level parameter and removed the final level of the inception V3 [8] model, then re-marked a final layer.

4.4 Train Model

At this step, we have to keep the parameters of the previous level, then the last layer will be removed and we need to input dataset to restore the next level.

By the backing promotion algorithm, the cross-entropy spend function is used to synthesize weight parameters by calculating the last level of the model and counting the output error of the texture level, and the label vector used for the given test category is used [5,7].

For final accuracy, we have also created Confusion Matrix. From Confusion Matrix, we have calculated Precision, Recall, Accuracy, and F1-Score. And finally, we have determined Macro Average Accuracy of our experiment.

Here is the Confusion Matrix of our model. From the following Confusion matrix of Table 1, we can say that our model has given a very high number of True Positive values.

Table 1. Confusion matrix

	China	Germany	India	Jamaica	Zimbabwe
China	18	1	0	1	0
Germany	0	19	1	0	0
India	1	0	16	1	2
Jamaica	0	1	0	16	3
Zimbabwe	1	1	0	3	15

5 Result Analysis

Figures 5 and 6 show the variation of accuracy and the cross-entropy which are based on our training dataset. The orange line represents the training set while the blue line represents the validation set.

Table 2 shows two statistical descriptions. For our dataset, training acupressure can reach 95% and validity accuracy can be maintained at 89%–90%.

Figure 7 shows the precision, withdrawal, accuracy and F1 scores graph of China, Gar-On, India, Jamaica, and Zimbabwe and shows macro-average accuracy, withdrawal, accuracy, and F1 scores [14,15].

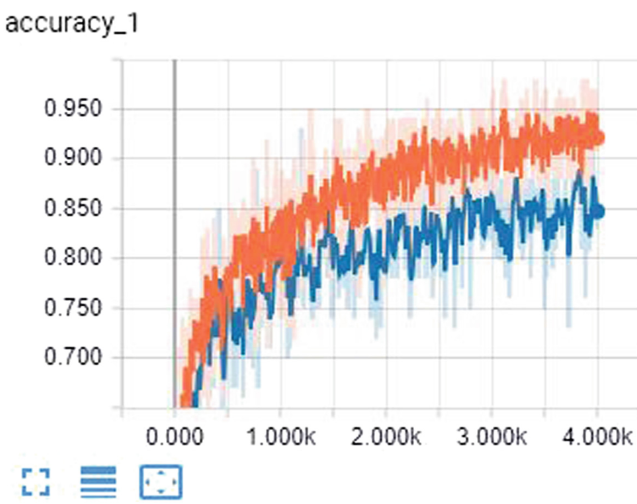

Fig. 5. The accuracy graph of Inception-V3 model.

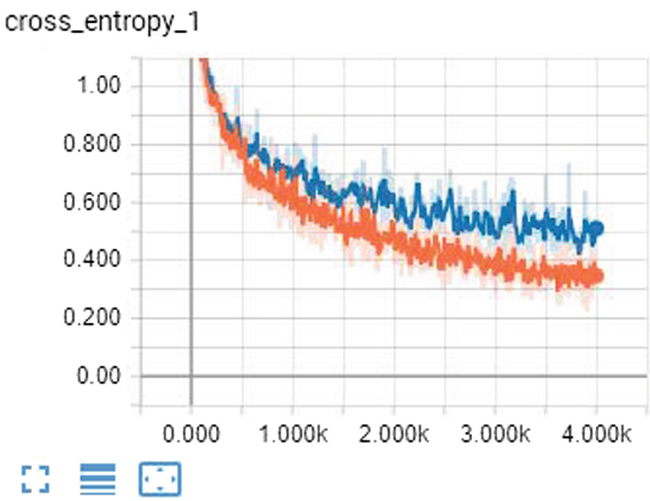

Fig. 6. The cross-entropy graph of Inception-V3 model.

Table 2. Description of the two figures.

	Index	Performance
Dataset	The accuracy of the training set	95%
	The accuracy of the validation set	89%-90%
	The cross-entropy of the training set	0.24
	The cross-entropy of the validation set	0.41

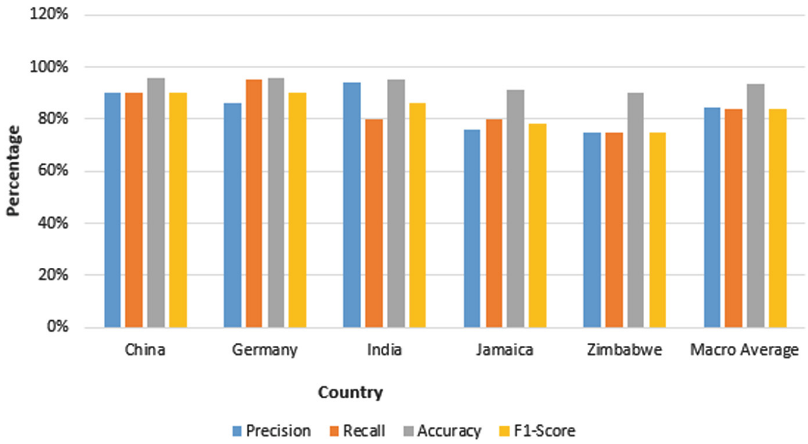

Fig. 7. Precision, Recall, Accuracy and F1-Score graph.

Table 3. The accuracy of five countries and final accuracy.

Country	Accuracy
China	96%
Germany	96%
India	95%
Jamaica	91%
Zimbabwe	90%
Macro average	93.6%

Table 3 shows the accuracy of five countries from the graph. For our dataset, the accuracy of China is 96%, Germany is 96%, India is 95%, Jamaica is 91%, Zimbabwe is 90% and the final accuracy is 93.6%.

6 Future Work

Google creates the TensorFlow [1] platform of the Inception-V3 [3] model and we used it. So our future work is to read and develop a more effective model so that we can use that model and enhance our accuracy.

7 Conclusion

Based on the Tensor Flow [1] platform Inception-V3 [3] model, we use transfer learning technology to identify the nationalities of five countries based on our datasets. And we get 93.6% accuracy of the model. We hope that our work will be useful for the future.

References

1. Abadi, M., Agarwal, A., et al.: TensorFlow: large-scale machine learning on heterogeneous distributed systems. CoRR abs/1603.04467 (2016)
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks, 03–06 December 2012
3. Mathew, A., Mathew, J., Govind, M., Mooppanb, A.: An improved transfer learning approach for intrusion detection. In: 7th International Conference on Advances in Computing & Communications, ICACC 2017, Cochin, India, 22–24 August 2017 (2017)
4. Bat-Erdene, B., Ganbat, T.: Effective computer model for recognizing nationality from frontal image (2016)
5. Xia, X., Xu, C.: Inception-v3 for flower classification. In: 2017 2nd International Conference on Image, Vision and Computing (2017)
6. Kieffer, B., Babaie, M., Kalra, S., Tizhoosh, H.R.: Convolutional neural networks for histopathology image classification: training vs. using pre-trained networks. arXiv:1710.05726v1 [cs.CV] 11 October 2017
7. Xia, X.-L., Xu, C., Nan, B.: Facial expression recognition based on TensorFlow platform. In: ITM Web of Conferences, vol. 12, p. 01005 (2017)
8. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. arXiv: 1512.00567v1 [cs.CV] 2 December 2015
9. https://datascience.stackexchange.com/questions/15328/what-is-thedifference-between-inception-v2-and-inception-v3
10. https://codelabs.developers.google.com/codelabs/tensorflow-forpoets/#0
11. Lu, X., Jain, A.K.: Ethnicity identification from face images. In: Proceedings of SPIE, vol. 5404. Biometric Technology for Human Identification (2004). https://doi.org/10.1117/12.542847
12. Charntaweekhun, K., Wangsiripitak, S.: Visual programming using flowchart. In: 2006 International Symposium on Communications and Information Technologies (2006)
13. https://datascience.stackexchange.com/questions/15989/microaverage-vs-macro-average-performance-in-a-multiclass-classification
14. Salmon, B.P., Kleynhans, W., Schwegmann, C.P., Olivier, J.C.: Proper comparison among methods using a confusion matrix. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2015)
15. Sawat, D.D., Hegadi, R.S.: Unconstrained face detection: a deep learning and machine learning combined approach. CSI Trans. ICT **5**(2), 195–199 (2017)