Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9315871

Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks

Bickey Kumar shah
Department of CSE
Delhi Technological University
Delhi, India
bickeyshah721@gmail.com

Aman Kumar
Department of CSE
Delhi Technological University
Delhi, India
amandeepnha@gmail.com

Amrit Kumar
Department of CSE
Delhi Technological University
Delhi, India
amritx8@gmail.com

***Abstract*— This paper dealt with the breed classification of dogs. To classify dog breed is a challenging part under a deep convolutional neural network. A set of sample images of a breed of dogs and humans are used to classify and learn the features of the breed. The images are converted to a single label of dimension with image processing. The images of human beings and dogs are considered for breed classification to find the existing percentage of features in humans of dogs and dogs of human. This research work has used principal component analysis to shorten the most similar features into one group to make an easy study of the features into the deep neural networks. And, the facial features are stored in a vector form. This prepared vector will be compared with each feature of the dog into the database and will give the most efficient result. In the proposed experiment, 13233 human images and 8351 dog images are taken into consideration. The images under test are classified as a breed with the minimum weight between test and train images. This paper is based on research work that classifies different dogs breed using CNN. If the image of a dog is supplied then the algorithm will work for finding the breed of dog and features similarity in the breed and if the human image is supplied it determines the facial features existing in a dog of human and vice-versa.**

Keywords*— *Breed, Feature, Images, Neural Network, Principal Component Analysis.

I. INTRODUCTION

Machine learning has been used widely these days. From small prediction models to creating high-end predictive models, machine learning has been used widely in fields ranging from business analytics to healthcare improvement [14-16]. Convolutional Neural networks (CNN) is one of the main achievements of the evolution of Machine learning and deep learning [2]. CNN constitutes the different neurons with different learning biases and weights. In this, every neuron receives some inputs and make some dot product or some mathematical manipulation and gives some outputs as desired. CNN architecture differs from other neural networks. The architecture of CNN helps to reduce the features and make the defined network to increase accuracy [4]. The neuron here is the single layers and different from the previous layers. CNN is carried out in three steps:

1. The convolutional layer of CNN

The images are generally in the form of a large matrix and the pixel intensity ranges from 0 to 256. The image is converted into 2D matrics whose entry has weight and it is used to recognize the features. Generally, a square matrix of size 4*4 or 3*3 will be considered.

2. The pooling layer of CNN

Pooling is a technique that is used to reduce the size of the image when the image is too large [11-12]. The parameters of the image are generally reduced to the size of the image by using the pooling method. Max pooling is the method that is frequently used and it gives the result as same as the original image. Other types of pooling are L2 norm and average pooling. The size of the output image is controlled by three Parameters:

- The number of filters
- Stride
- Zero Padding

3. The output layer of CNN

After passing through the Convolutional layer and Pooling layer the compressed image is obtained with a lot of features in it [13]. The size of the output image is calculated using the formula ((W-F+2P)/S) + 1.
Where W is input volume Size.
F is filter Size.
P is the number of padding applied
S is the number of strides used.
Features are extracted with the help of the convolutional layers. The architecture of CNN can be divided for the dog breed classification as followed:

- Convolutional Layer n°1 with 96 filters
  - Max pooling
  - Relu
- Convolutional Layer n°2 with 256 filters
  - Max pooling
  - Relu
- Convolutional Layer n°3 with 384 filters
  - Max pooling
  - Relu
  - DropOut
- Flatten Layer
- Fully Connected Layer with 500 nodes
  - Relu
  - DropOut
- Fully Connected Layer with n nodes (n = number of breeds)

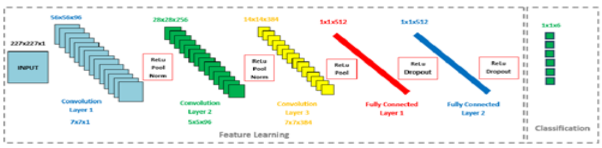

Fig.1 CNN Architecture

We have different mathematical series and terms associated with the convolutional neural networks. For working with Convolutional Neural Networks, the Activation functions in CNN are the basic requirements for non-linear processing [1]. The activation function is used to scale the output values and if required sometimes converts the high value to probability for detection of most probable images [6]. It decides that the upcoming activation states of the neurons that it should be activated or not and it is done by the sum of the bias and weights. Without the activation function model is only the linear model and it won't bring efficiency for the non-linear calculations. the activation functions are as follows:

1. **Sigmoid Function:** The mathematical formula it uses is f(z)=1/(1+exp(-z)). The value ranges from 0 to 1.

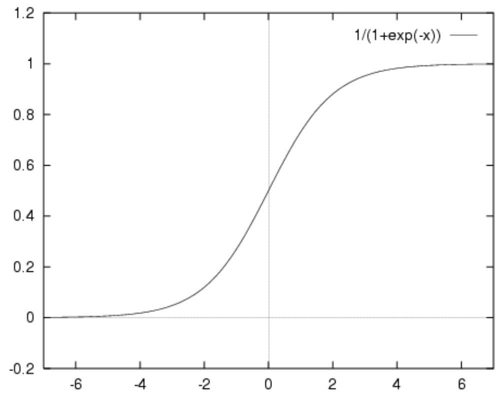

Fig.2 Sigmoid Function

2. **Tanh (Hyperbolic Function):** The value of this ranges from -1 to 1. The mathematical formula used for this function is f(z)=(1-exp(-2z))/(1+exp(-2z)).

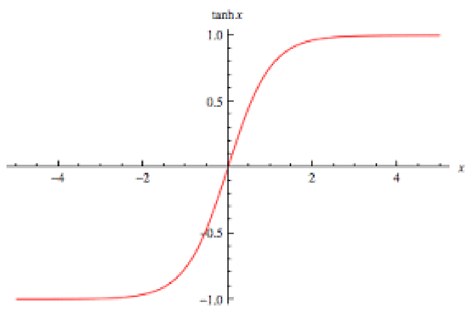

Fig.3 Tanh function

3. **ReLU**- Rectified linear Units:The mathematical formula for this is R(z)= maximum(0,z) i.e if z<0, R(z)=0 and if z>0 R(z)=a.

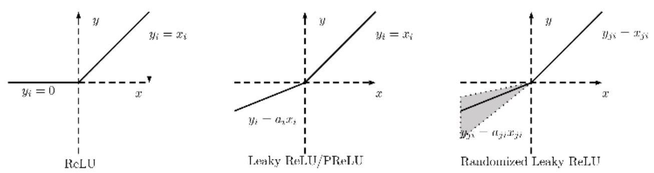

Fig.4 ReLU Function

II. METHODOLOGY

Computer vision is used for the image processing work to convert the images into black and white images [3]. In this we revolve around the CNN algorithm. The data is preprocessed and after that classified into the group of training data, testing data, and validation data. Training data is used for finding patterns of the images using deep neural networks. for the adjustment of error in the weights, Feedforward and backpropagation work are carried out over the networks.

A. *Transfer Learning*

It is a popular approach to deep learning. Pre-trained models are used at the starting points and in these cases Transfer learning is considered to be best [7-9]. To solve any type of problem we generally need a model of a similar type and transfer learning plays important role in it. To make a fine model through the transfer learning we modify the learning rates, go through the features extraction, use the architecture of the pre-trained model, and train some layers and freeze others [8]. It is an optimization technique. The learning rates from the pre-trained model should be kept in mind while performing transfer learning.

B. *Import Datasets*

We need to load the required datasets and all the required library that is important for the computation work. We will import all the required libraries and then perform the processing parts using TensorFlow. We will use the TensorFlow framework of version 2.3.0.

C. *Resize and Normalize the data*

All the images do not have the same shapes and for preparing the model we need to resize all the images into one dimension of square shape. We normalize our dataset by dividing all the pixels by 255.

D. *Preprocessings Images, CNN and Models*

We perform some more data pre-processing and features the reduction of large data using the principal components analysis and then proceed for the efficient model prediction works.

```
# Downloading the pre-trained face detector
import os
if not os.path.isdir('haarcascades'):
    os.mkdir('haarcascades')
!wget -cq -O haarcascades/haarcascade_frontalfa
clear_output()
print("Downloaded Successfully")

Downloaded Successfully
```

Fig. 5. Pre-trained face detector algorithm

We download the pre-trained face detector with the computational transfer learning algorithms taking the learning rates into considerations. It helps to detect the face that whether it belongs to the human or Dog and then after it again find the similarities to predict the breed.

```
human_files_short = human_files[:100]
dog_files_short = dog_files[:100]

homosapiens, canis = 0, 0
out = display(progress(0, 100), display_id=True)
for il, (human, dog) in enumerate(zip(human_files_short, dog_files_short)):
  out.update(progress(il+1, 100))
  if face_detector(human):
    homosapiens += 1
  if face_detector(dog):
    canis += 1

print("Human detection accuracy: {}%".format(100.* float(homosapiens) / len(human_files_short)))
print("Dogs detected as human: {}%".format(100. * float(canis) / len(dog_files_short)))

Human detection accuracy: 99.0%
Dogs detected as human: 11.0%
```

Fig. 6 Algorithm to find Human face.

This algorithm finds the number of the face detected i.e. it specifies whether the algorithm can detect or not that the facial accuracy of the human and dogs. This is for human face detection and it gives the accuracy of 99% when input images of humans are given and for the same when dog image is input it gives 11%. Since the algorithm is designed for humans it implies that our accuracy is correct. And, similarly, we have the algorithm for dog face detection which gives the accuracy of 100% when a dog face is an input and 0% when the face of a human is input. The same algorithm over different computations helps to classify breeds and further we give the image input of humans that finds the prediction of the dog breed facial match of a human face. We transform the training, validation, and testing set. We define the CNN architecture and then convolutional layers with increasing numbers of filters in each layer and max-pooling after 3 layers of value 2 [10]. We perform the non-linear task and dropout value of the layer to be 2. We give the feedforward algorithm by adding the sequence of a convolutional and max-pooling layer. We perform the ReLU activation function over the first and second hidden layers. We predict the loss function and make the system more optimized with the minimum occurrence of loss. We initialize the variable for training and testing loss predictor and also trackers for the minimum loss and preparation of the models efficiently. We train the model and clear the gradients of all optimized variables. We then work on validating the model. We update the average validation loss and calculate the batch loss. After this, we train the model and transform the training, validation, and testing sets. Further computations are made with the learning rates and image classifications.

III. EXPERIMENTAL RESULT

We performed the Convolutional Neural network experiment with a lot of images of humans and dogs. We come across the result that the total number of the dog images sample was classified into 103 dog breed images. We were able to classify the breed of the dog and finally, we detected the probability of features of dog breed that match the human features of a particular sample. The experiment was performed over a lot of images and data, distinctly some of the obtained results are: -

Fig. 7 sample Image1

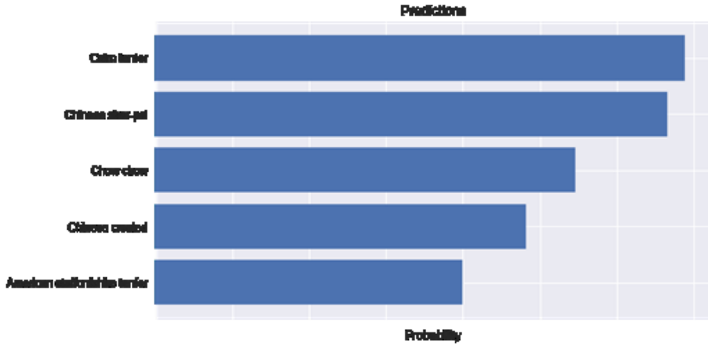

Fig. 7.1 sample Image1 result

In fig. 7 It is shown that when a human sample is given as an input, the output result shows that the human sample image has a shadow of five different dog breeds in it. The maximum probability that the human seems like is 'Cairn terrier' breed, after that the probability goes max to 'Chinese shar-pei' and so on. It signifies that the facial expression of the human in the sample matches maximum with the cairn terrier and least with the American Staffordshire terrier.

Fig. 8 sample image2

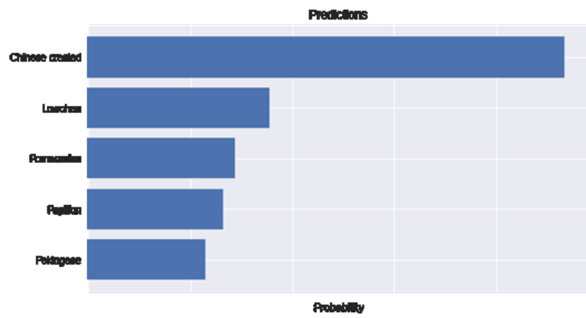

Fig.8.1 sample image2 result

In fig. 8, It is similar to that of Fig. 7, that when a human sample is given as an input, the output result shows that the human sample image has a shadow of the top five different dog breeds in it. The maximum probability that the human seems like is a Chinese crested dog breed, after that the probability goes max to Lowchen and so on. It signifies that the facial expression of the human in the sample matches maximum with the Chinese crested and least with the Pekingese.

Fig. 9 sample image3

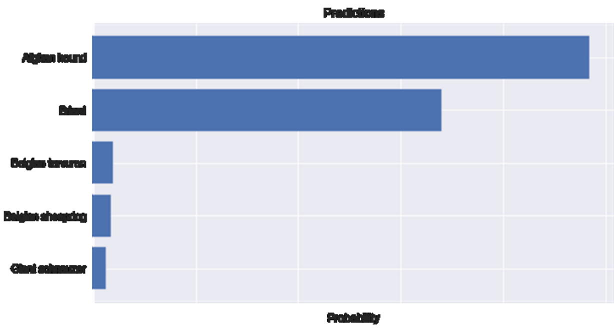

Fig. 9.1 sample image3 result

In fig. 9 It is similar to that of Fig. 7, that when a human sample is given as an input, the output result shows that the human sample image has a shadow of the top five different dog breeds in it. The maximum probability that the human seems like is Afgan hound dog breed, after that the probability goes max to briard and so on. The facial expression of human has negligible matching with Belgian Tervuren and Giants Schnauzer. It signifies that the facial expression of the human in the sample matches maximum with the Chinese crested and least with the Giants Schnauzer.

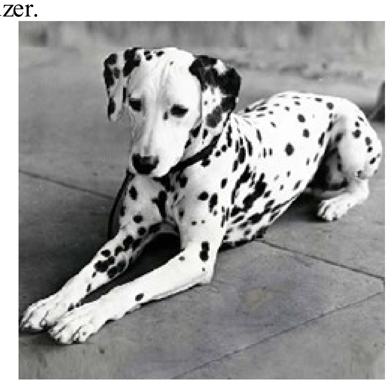

Fig. 10 sample image4

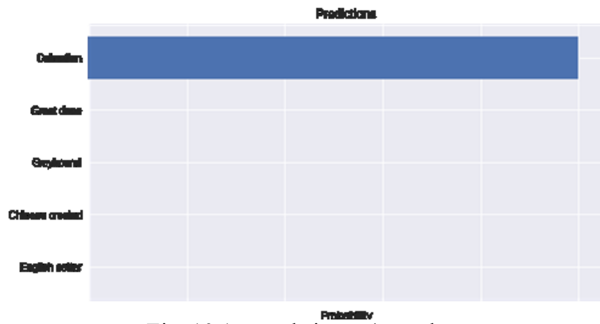

Fig. 10.1 sample image4 result

In Fig. 10 the sample images of the dog give full probability prediction with Dalmation breed, the sample dog is really of dalmation breed that is why it has shown full prediction accuracy, this sample verifies that the working model is almost accurate. This sample has not given any errors or features match with other dog breed samples.

Fig. 11 sample image5

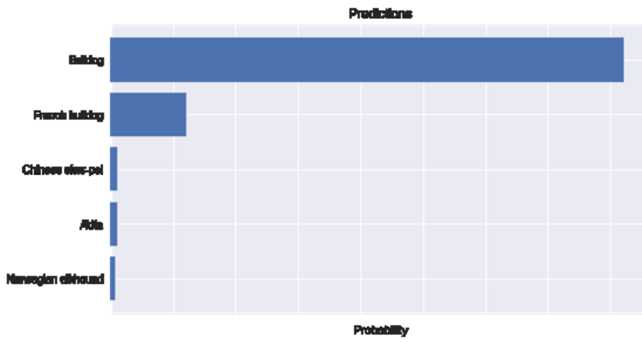

Fig. 11.1 sample image5 result

Fig. 12. Sample image6

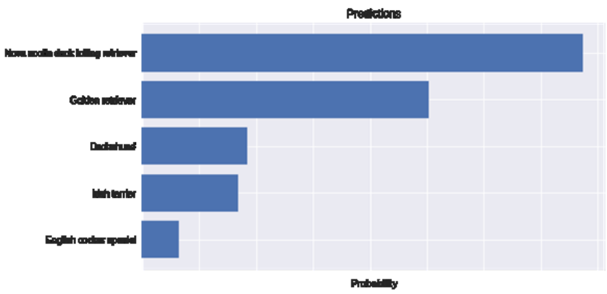

Fig. 12.1 Sample image6 result

We have similar cases in Fig.11 and Fig.12 as that of fig.10. This is the breed identifying for the given dog image. In fig. 9 it is a real Bulldog image and it has shown almost the greatest accuracy with bulldog breed and almost negligible match with the rest of the breeds. In Fig.10 the image is of a toy dog and It has shown a similar breed property as humans in the above image. All of the samples have a different result and almost accurate results on practical ideas and theory. It signifies that the models are nearly accurate. The results of the data truly depend on the type of data that we use. For the Deep learning analysis, we need millions of data for getting more and more accurate results. We have obtained the following model accuracy with the epoch.

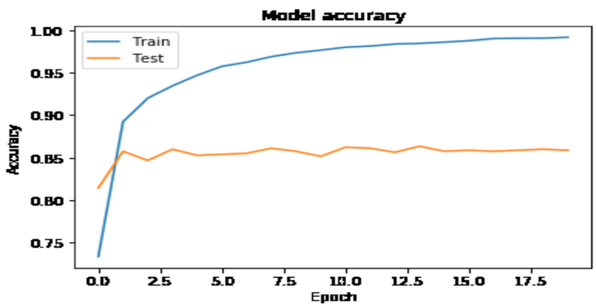

Fig13. Accuracy Vs Epoch

We have also bear some model loss that were significantly optimized with the learning rate and loss variable detection over every epoch. A snap of model loss and epoch is as:

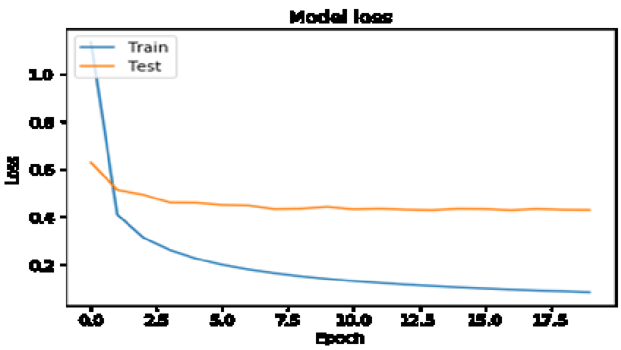

Fig. 14 Model loss VS Epoch

This paper presents that CNN is a good approach for the dog breed classification work. The experiment shows that the CNN algorithm has an accuracy of 95.5%. This signifies that CNN can be used for dog breed classification.

IV. CONCLUSION AND FUTURE SCOPE

Convolutional Neural network seems to be efficient for the dog breed classification. This learning has shown that multiple algorithms can be processed with different parameters for efficient learning and classification of dog breed. Newer Technology and better computation power are developed and more study work can be done during the course time for better classification. The algorithm can be made further efficient with the deep learning models, and the breed classification can be performed further with deep neural networks by structuring similar algorithms and better computation methods.

ACKNOWLEDGMENT

The Authors are indebted to Central Library, Delhi Technological University, New Delhi for providing the access to all the materials that were needed for the study.

REFERENCES

[1] P. Road, "Face Recognition Based Dog Breed Classification Using Coarse-to-Fine Concept and PCA Massinee Chanvichitkul , Pinit Kumhom , and Kosin Chamnongthai Department of Electronics and Telecommunication Engineering Faculty of Engineering King Mongkut's Univers," *System*, pp. 5–9, 2007.

[2] A. Ghimire, S. Thapa, A. K. Jha, A. Kumar, A. Kumar, and S. Adhikari, "AI and IoT Solutions for Tackling COVID-19 Pandemic", in 2020 International Conference on Electronics, Communication and Aerospace Technology, 2020: IEEE

[3] P. Borwarnginn, K. Thongkanchorn, S. Kanchanapreechakorn, and W. Kusakunniran, "Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning," *2019 11th Int. Conf. Inf. Technol. Electr. Eng. ICITEE 2019*, vol. 7, pp. 3–7, 2019, doi: 10.1109/ICITEED.2019.8929955.

[4] P. Prasong and K. Chamnongthai, "Face-recognition-based dog-breed classification using size and position of each local part, and PCA," *2012 9th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2012*, pp. 3–7, 2012, doi: 10.1109/ECTICon.2012.6254212.

[5] A. K. Jha, S. Adhikari, S. Thapa, A. Kumar, A. Kumar, and S. Mishra, "Evaluation of Factors Affecting Compressive Strength of Concrete using Machine Learning" in 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), 2020: IEEE, pp. 70-74, doi: 10.1109/ACCTHPA49271.2020.9213199.

[6] J. Krause, T. Gebru, J. Deng, L. J. Li, and F. F. Li, "Learning features and parts for fine-grained recognition," *Proc. - Int. Conf. Pattern Recognit.*, pp. 26–33, 2014, doi: 10.1109/ICPR.2014.15.

[7] A. Ghimire, S. Thapa, A. K. Jha, S. Adhikari, and A. Kumar, "Accelerating Business Growth with Big Data and Artificial Intelligence", in 2020 Fourth International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020: IEEE, pp. 441-448, doi: 10.1109/I-SMAC49090.2020.9243318.

[8] C. D. Casuat, A. S. Alon, J. L. Dioses, R. M. Dellosa, and R. A. Diamante, "RFID controlled 'GG' pieces ranking detection with watch-dog enable," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 1 Special Issue 3, pp. 274–279, 2020, doi: 10.30534/ijatcse/2020/4191.32020.

[9] R. O. Sinnott, F. Wu, and W. Chen, "A Mobile Application for Dog Breed Detection and Recognition Based on Deep Learning," *Proc. - 5th IEEE/ACM Int. Conf. Big Data Comput. Appl. Technol. BDCAT 2018*, pp. 87–96, 2019, doi: 10.1109/BDCAT.2018.00019.

[10] J. J. Cooper *et al.*, "Necrotizing meningoencephalitis in atypical dog breeds: A case series and literature review," *J. Vet. Intern.*

Med., vol. 28, no. 1, pp. 198–203, 2014, doi: 10.1111/jvim.12233.

[11] C. Marín Navas, F. J. Navas González, V. Castillo López, L. Payeras Capellà, M. Gómez Fernández, and J. V. Delgado Bermejo, "Impact of breeding for coat and spotting patterns on the population structure and genetic diversity of an islander endangered dog breed," *Res. Vet. Sci.*, vol. 131, no. April, pp. 117–130, 2020, doi: 10.1016/j.rvsc.2020.04.003.

[12] R. Kumar, M. Sharma, K. Dhawale, and G. Singal, "Identification of Dog Breeds Using Deep Learning," *Proc. 2019 IEEE 9th Int. Conf. Adv. Comput. IACC 2019*, pp. 193–198, 2019, doi: 10.1109/IACC48062.2019.8971604.

[13] M. V. S. Rishita and T. A. Harris, "Dog breed classifier using convolutional neural networks," 2018 Int. Conf. Networking, Embed. Wirel. Syst. ICNEWS 2018 - Proc., 2018, doi: 10.1109/ICNEWS.2018.8903980.

[14] S. Thapa, P. Singh, D. K. Jain, N. Bharill, A. Gupta, and M. Prasad, "Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease", in 2020 International Joint Conference on Neural Networks (IJCNN), 2020: IEEE, pp. 1-8, doi: 10.1109/IJCNN48605.2020.9207359.

[15] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020: IEEE, pp. 975-978, doi: 10.1109/ICESC48915.2020.9155833.

[16] S. Thapa, S. Adhikari, A. Ghimire, and A. Aditya, "Feature Selection Based Twin-Support Vector Machine for the diagnosis of Parkinson's Disease", in 2020 8th R10 Humanitarian Technology Conference (R10-HTC), 2020: IEEE