GAP GYAN
A GLOBAL JOURNAL OF SOCIAL SCIENCES
(ISSN – 2581-5830)
Impact Factor – SJIF – 4.998, IIFS - 4.375
Globally peer-reviewed and open access journal.

FACE MASK DETECTION USING DEEP LEARNING – RETINANET

Mittal R., Pandya B.

Bachelor of Science – Computer Science Department
University of Illinois,
Urbana-Champaign
rahilm3@illinois.edu

Regina, Saskatchewan, CA
bhavikpandya79@gmail.com

Abstract

Technological advancement has provided many facilities and made life easy, but every coin has two sides. As the life of humans has become comfortable but at the same time many new problems are arising. In recent years humans have been fighting fatal diseases like COVID-19. The pandemic has changed life drastically and as an ill effect of it, social distancing and masks has become a societal norm. In the proposed research, RetinaNet is used to identify whether the human face is with mask or without mask.

Keywords: *RetinaNet, Facemask detection, Resnet.*

INTRODUCTION

The world was facing the pandemic situation due to the outbreak of a deadly virus. The pandemic has changed human life drastically. In December 2019 the virus disease had spread and changed every human being's life extremely. After the diseases, the world's eye is on research in medical science that can save human life and as a result vaccination has been identified as one of the important tools but still, to save oneself from this disease certain important parameters have been identified that includes social distancing, use of mask and sanitization. Due to ill effect of the virus, the facemask has become need of an hour. Facemask is useful to reduce the transmission of virus so many countries has made it compulsory to wear Facemask in public areas. To fight against deadly viruses, artificial intelligence (AI) can also be helpful. With use of Deep learning and AI, it is possible to forecast that how speedy it will be spread and what will be the need of medical equipment. It is also useful to identify, what should be rules and regulation for the mass. For example facemask should be compulsory when in a group and maintain social distancing. This research paper identify that the person is without mask or with mask with the use of deep learning technique –RetinaNet. The next section describes the related work in this area.

RELATED WORK

Nowadays in machine learning, computer vision plays important role for different identification includes face, object or even diseases. For object detection RCNN, Faster RCNN, RetinaNet and different versions of YOLO are used[2] YOLO darknet, is the first algorithm that work with face mask detection.[3].Face mask detection algorithm has been already developed for different situations. In a research paper, the detection of face mask has been done on webcam images and different algorithms are compared like YOLO V3-tiny, YOLOv3, Faster-RCNN, YOLOv3-SPP, RetinaNet, Mask RCNN, YOLOv5x, YOLOv5xTTA [3]. In another study to identify that the person is wearing mask or not YOLOv3 algorithm is used. In YOLOv3, the image coordinates are identified by the object detector. It identify the target object features through divide the image in the grid, the cells with high confidence rate are added at one place to generate an output [4]. According to another research, Resnet50 is used for feature extraction and to identify the image with mask or without mask different classification algorithms like Decision Tree, SVM and Ensemble are used and SVM is identified as most efficient classifier [5]. In another study, RGB components identified for upper and lower facial part and with use of SVM classifier it has been tried to identify that the person is with mask or without mask [6].

RESEARCH METHOD

In this research paper RetinaNet algorithm is used for object identification. There are two types of algorithm in object

https://www.gapgyan.org/

detection. One stage detector algorithm and two stage detector algorithm. Two stage detector algorithms are R-CNN, Fast R-CNN, Faster R-CNN. This algorithm first identify the region and then classify the object. The two stage detector is accurate but bit slower. RetinaNet is one stage detector, that has better performance (accuracy) then two stage detector. There are different one stage detector algorithms are available but the major problem in one stage detector is class imbalance.
RetinaNet is single stage object detection model which is fast and accurate too. The structure of RetinaNet is mainly divided in four parts.
1. Backbone network: The backbone network is using bottom –up approach. It calculates the feature map at different scales. It does not consider the input image size and backbone size. Resnet is used as a backbone of RetinaNet. Resnet (Residual neural network) is collection of many convolutional network and skip connections.
2. Feature pyramid network: it takes single scale image of arbitrary size as input and output proportionally sized feature map at multiple levels. It is not dependent on backbone network. Top –down path way is used.
This architecture combines both, feature map that have small resolution so it is advisable for larger objects, and grid cells have feature map have high resolution so it is better for detecting smaller object.so with combination of top down and bottom up , it does not require extra computation.
3. Class subnet: it is used for object classification. In this network 3*3 convolutional network is used with 256 filters. It is followed by another 3*3 convolutional layer with K*A filters. The output feature map size is W*H*K*A where, W= width of input feature map, H=height of input feature map,
Sigmoid function is used for object classification
4 .Box subnet: it is used for object regression.
RetinaNet uses a feature pyramid network to efficiently detect objects at multiple scales and introduces a new loss, the Focal loss function, to alleviate the problem of the extreme foreground-background class imbalance. [7]

EXPERIMENTAL ANALYSIS

The paper focusses on identification of face mask using RetinaNet. The dataset is created using the VGG Image Annotator for creation of bbox for identification of face mask [1]. Some images are used from Kaggle datasets [2]. The bounding box is generated manually for the face mask and is annoted for each image and labels are encoded for mask and nomask images. The bbox consists of four parameters x, y , width and height of the rectangle box annoted for identifying face mask in the image. All images are of a single face with mask or without mask. Each image has same width and height of 301x400 pixels. The source indicates whether the face image has mask or is without a mask (nomask).

Fig 1 : Facemask images with bounding box

The csv file is created annotation of images as shown in table 1. 800 images are used for training set and have selected a pretrained RetinaNet models.

The batch size =2, steps=100, learning rate =0.0001 and 10 epochs are specified for execution.
The hardware used is CPU model name: Intel(R) Xeon(R) CPU @ 2.20GHz, CPU MHz: 2200.234 CPU-Cores: 2, RAM MemTotal: 18433316 kB and Operating System : Linux 5322ad728f31 5.10.90+

TABLE 1
Dataset representing the bounding box coordinates.

image_id	width	height	bbox	source
m1	301	400	[70, 226, 115, 95]	mask
m2	301	400	[73, 280, 78, 103]	mask
m3	301	400	[148, 90, 47, 57]	mask
m4	301	400	[91, 294, 104, 93]	mask
m5	301	400	[154, 241, 82, 82]	nomask

https://www.gapgyan.org/

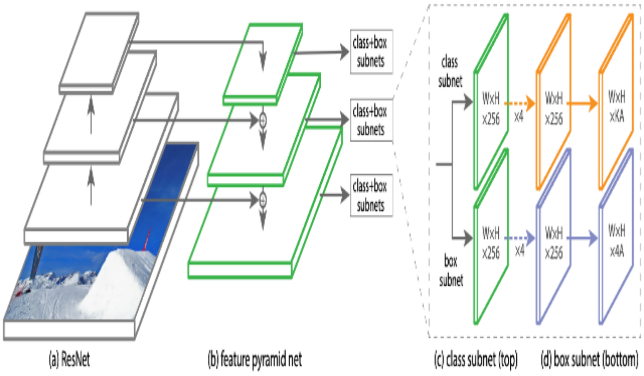


Fig 2 Image describing the RetinaNet architecture [3]

RetinaNet is a single, unified network composed of a backbone network and two task-specific subnetworks. The backbone we are using is resnet50, it is responsible for computing a convolutional feature map over an entire input image. The first subnet performs convolutional object classification on the backbone's output; the second subnet performs convolutional bounding box regression. The pretrained model is used for classification of two classes defined as mask and nomask. The experimental analysis is carried out for different number of images with mask and without mask and the loss, regression loss and classification loss was calculated using the RetinaNet models.

The loss derived using Resnet50 backbone is shown in the table 2.
TABLE 2
Loss generated for different number of epochs using Resnet50.

epoch	loss	regression loss	classification loss
1	2.0479	1.4849	0.5630
2	1.4710	1.1838	0.2872
3	1.3587	1.1354	0.2233
4	1.2881	1.0440	0.2441
5	1.3276	1.0648	0.2628
6	1.2571	1.0463	0.2108
7	1.3450	1.0857	0.2593
8	1.1804	1.0109	0.1695
9	1.1565	0.9705	0.1860
10	1.1826	0.9898	0.1927

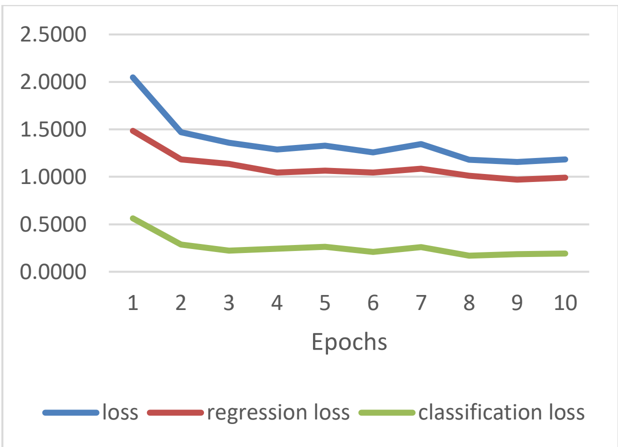


Fig 3 Graph depicting loss generated for different epochs using Resnet50

The loss derived using Resnet101 backbone is shown in the table 3

TABLE 3
Loss generated for different number of epochs using Resnet101.

GRAND ACADEMIC PORTAL
RESEARCH JOURNALS
GAP GYAN
A GLOBAL JOURNAL OF SOCIAL SCIENCES
(ISSN – 2581-5830)
Impact Factor – SJIF – 4.998, IIFS - 4.375
Globally peer-reviewed and open access journal.

epoch	loss	regression loss	classification loss
1	4.5633	2.2124	2.3508
2	2.1694	1.6858	0.4836
3	1.9634	1.5142	0.4492
4	1.9634	1.4273	0.3876
5	1.9634	1.2863	0.3497
6	1.7017	1.3177	0.3840
7	1.6087	1.2741	0.3346
8	1.5134	1.2064	0.3070
9	1.5232	1.2059	0.3173
10	1.5188	1.2276	0.2912

Fig 4 Graph depicting loss generated for different epochs using Resnet101

The loss derived using Resnet152 backbone is shown in the table 4.
TABLE 4
Loss generated for different number of epochs using Resnet152.

epoch	loss	regression loss	classification loss
1	5.0707	2.1371	2.9336
2	2.1938	1.7023	0.4915
3	1.945	1.4903	0.4547
4	1.8719	1.4735	0.3984
5	1.7796	1.3995	0.3802
6	1.7636	1.3503	0.4133
7	1.7613	1.3911	0.3702
8	1.7035	1.3442	0.3612
9	1.6529	1.3303	0.3226
10	1.5627	1.2464	0.3163

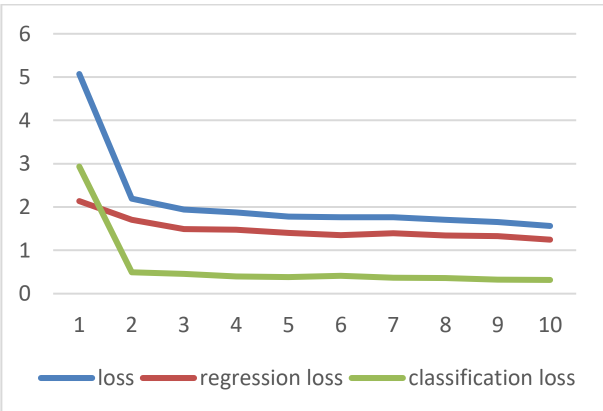
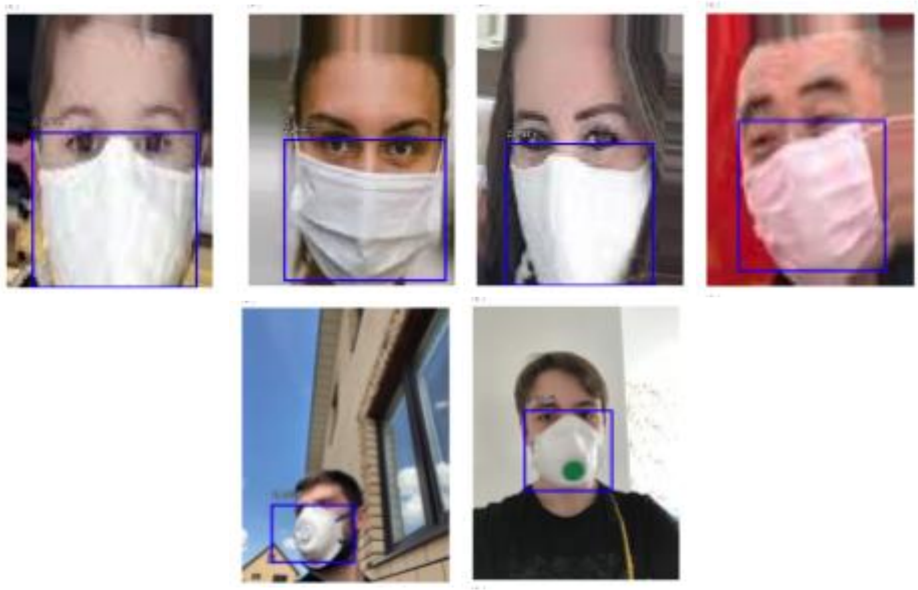


Fig 5 Graph depicting loss generated for different epochs using Resnet152

GAP GYAN
A GLOBAL JOURNAL OF SOCIAL SCIENCES
(ISSN – 2581-5830)
Impact Factor – SJIF – 4.998, IIFS - 4.375
Globally peer-reviewed and open access journal.

The predictions are carried for resnet50, resnet101 and resnet152, the face mask is accurately identified with the threshold value of 0.9.



CONCLUSION

This research paper used the deep learning technique – Retinanet to identify whether the person is without a mask or with mask. The experiments were conducted for various backbones like ResNet50, ResNet101, and ResNet152. The loss is calculated for each backbone, and predictions were generated for each back-bone. Our approach is simple but effective for identifying the facemask with the one-stage detector RetinaNet and produces experimental results by applying different backbones. From the results, we can conclude that for images with similar sizes, RetinaNet with the ResNet50 backbone gives the same predictions with lower loss, regression loss, and classification loss.

REFERENCES

[1] X. Liu and S. Zhang, "Covid-19: Face masks and human-to-human transmission," Influenza and Other Respiratory Viruses, 2020.

[2] P. Wu , H. Li , N. Zeng , F. Li, "FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public", Image and Vision Computing,vol.117,p.104341,jan.2022,doi:10.1016/j.imavis.2021.104341

[3] E. Adhikarla and B. D. Davison, "Face Mask Detection on Real-World Webcam Images," Proceedings of the Conference on Information Technology for Social Good, Sep. 2021, doi: 10.1145/3462203.3475903.

[4] M. R. Bhuiyan, S. A. Khushbu, and M. S. Islam, "A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3," IEEE Xplore, Jul. 01, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9225384

[5] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," Measurement, vol. 167, p. 108288, Jan. 2021, doi: 10.1016/j.measurement.2020.108288.

[6] I. Buciu, "Color quotient based mask detection," IEEE Xplore, Nov. 01, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9301079

[7] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2018, doi: 10.1109/tpami.2018.2858826.

https://www.gapgyan.org/