Fuzzy Logic and CNN for Social Distance Analysis

Syed Ali Vijdaan Naqvi
Faculty of Computer Science & Eng.
Bachelors of AI
GIK Institute of Eng. Sciences & Tech.
Topi, Khyber Pakhtunkhwa, Pakistan.
alivijdaan@gmail.com

***Abstract*— Widespread isolation has shown us difficult it is to enforce safety protocols among the public and even more difficult to supervise it. The project uses CNN and Fuzzy Logic to identify masks and the state of distance between people. The CNN was able to achieve a high accuracy in identify its target while use of Fuzzy Logic although facing limitations was able to output desired results within those confines. Overall, the system would be able to help in the identification of people who have broken SOP's and efficient supervision of safety protocols.**

***Index Terms*—CNN, Fuzzy Logic, Deep Learning**

I. Introduction

One paragraph on why is it significant to work on this field/topic today.

As machine learning and computer vision experience a boom in its market many applications are being seen, spread across various domains. One such domain is healthcare and public safety. One crucial aspect of public safety was seen just a few years with the spread of a pandemic causing global isolation. The need for enforcing Safety Operations Protocols (SOPs) became all time high and the next challenge faced by the multitudes of institutions around the globe was the adequate and efficient supervision of such SOPs. The paper delves into the use of Convolution Neural Networks and its architecture used to identify correctly worn, incorrectly worn and not worn masks. Furthermore, the paper will discuss how fuzzy logic was used to infer the severity of SOP's being followed – '*severe*', '*moderate*' and '*safe*'.

The selected topic holds immense significance as it showcases the power of synergy between two different approaches. The usage of a combination of soft computing for identification and classification is important as it highlights the importance of different soft computing methods working in tandem to produce better results. The successful implementation of this system can reduce transmission of infectious diseases and promote a healthier society.

Fuzzy logic has high potential in understanding complex systems where analytical solution may not exist or the system is not understood properly but can be observed (Ross et al., 2004). In my opinion, the use of fuzzy logic to measure the distance between moving objects will allow us to create better systems that could react fluidly when interacting with objects in motion. If the fuzzy logic and CNN's are successfully not only will the healthcare domain benefit from it but this research might have far-reaching effects as it may produce results in terms of sports, entertainment, transportation and various other industries.

A. Related Work

The intermingling of fuzzy sets and CNN has already seen some research with researchers like Khan MN [1] and Yalcinkaya [2] already producing impressive results. As shown in Table II, the following authors have published studies related to object detection through CNNs and Fuzzy Logic working in tandem. Khan MN's paper focused on the usage of multiple cameras to detect weed along with fuzzy logic. They set out to improve the robustness of the CNN through multi-sensor based fuzzy fusion algorithm [1]. Yalcinkaya used fuzzy based fuzzy correlation map to produce a fingerprint matrix. This fingerprint matrix could be directly fed to the CNN without the need for medical grade training or preprocessing.

B. Gap Analysis

Currently, the works being produce may not adapt well to changing scenarios such as varying lighting conditions or unexpected obstacles. Furthermore, these systems may not be scalable to a larger crowd [4].

In addition, existing systems may not be compatible with various hardware and software problems. Finally, one of the most important point to take into consideration is that there might be ethical concerns that would need to be addressed and privacy biases needed to be solved. [5]

C. Problem Statement

Following are the main research questions addressed in this study.

1) How to identify people with and without masks
2) How to make sure people are adopting social distancing correctly.

D. Novelty of my work

Although CNNs have been used many times for identification in images, using fuzzy sets on images in tandem to create a proper system of identification of people and the state of social distancing that exists between them is something I believe is unique.

E. Our Solutions

I was able to produce a 95% accuracy on the CNN and produce desirable results using fuzzy logic despite facing numerous limitations.

II. Methodology

A. Dataset

The Face Mask Detection dataset includes around 800 images of people wearing masks, correctly, incorrectly or not at all. Furthermore, the dataset includes the co-ordinates of the bounding boxes in each image. Fig. 1 shows a sample of images included in the dataset. The dataset was then split according to the bounding boxes co-ordinates given in the xml file attached. The total number of images I had to process is mentioned in Fig. 2.

B. Overall Workflow

Fig. 4 shows the workflow of the entire system explained in this study. The entire workflow starts off with the image being processed through the Convolution Neural Net. It is only after the CNN the has verified that there exists more than one person in the image that we move on to the next step of our workflow. If the CNN has determined that there only exists one person then the output would be regarded as the final result. When moving onto our next stage of the pipeline – three crucial pieces of information are to extracted as our fuzzy logic rules are based on these variables. These variables include:

- The distance between bounding boxes
- The size of the bounding boxes
- The local density of bounding boxes

It is through these three variable that we can infer the state of SOP severity as shown in Fig. 4

C. Experimental Settings

As mentioned in Table II, the CNN was trained on 50 epochs with a minimum learning rate of 0.00001. I used early stopping and learning rate reduction on plateaus on every 8 epochs. This was to stop it the model from overfitting and save computational power. One extra hyper-parameter which was introduced to stop overfitting was dropout. Fig. 3 shows the entire architecture of the CNN which was developed through its ONNX file. The architecture has three convolutional layers with max pooling layers to extract features from the input image. It also includes two dropout and dense layers for regularization and classification in one of the three categories. The architecture includes ReLU activations and same paddings to preserve spatial information. The final layer uses a softmax function to produce a probability for each class.

The total amount of rules set for the fuzzy logic were 27. Each antecedent had 3 linguistic variables while the consequent also had 3 linguistic variables. The rules are mentioned in Table III.

III. Results

Our first problem statement included identifying people with masks. This problem statement was to be overcome through the CNN model. After training the model through several images the model could successfully identify which bounding box belongs which category. The progress through the training and the testing via validation set is shown in Fig. 5. In Fig. 5 (a) we can see the training progress and the validation progress was very similar with both their ending points being the same as well as their general trend. This means that our model did not encounter any overfitting or under fitting and can produce accurate results on new data as well. In Fig. 5 (b) we see that the model produces an accuracy of around 95%. Other than accuracy, we also checked through different metrics. These metrics are shown in Table IV.

The second problem statement dealt with enforcing proper social distancing among the masses. To do this we required a system that could give meaning to linguistic variables such as *'far'* or *'close'*. By using fuzzy logics, we were able to produce a change of states as two people or objects changed the distance between each other. Fig. 6 shows the result on a testing image. The figure shows 4 images regarding a scene where a man and woman slowly move closer to each other. The first column shows the initial position of the two people. Below the images of the couple are images showing one of the variables that influence the social distance state – *'localized bounding box density'.*

By dividing an image into grids and counting the number of bounding boxes per grid we were able to find the number of people in a specific area. Using the localized bounding box density, distance between the bounding boxes of the couple and the sizes of their bounding boxes we were able to get the membership function of SOP severity. The results show that initially the state of the couple (as shown by the black line) is in the safe zone. During the second image the black line moves closer to the 'moderate' zone. The third image is where the couple has entered the 'moderate' zone. This zone although not too dangerous still holds a sense of caution as the subjects might get too close to each other. The last scene shows the couple hugging and the membership function shows their state as severe. The testing hence proves that the fuzzy logic can produce desirable results.

TABLE I

LITERATURE REVIEW TABLE SHOWING THE CONTRIBUTIONS OF VARIOUS AUTHORS FOR OBJECT DETECTION

Author Name	Applied On		Fuzzy Rules	Dataset Used	Results
	Convolution Neural Network	Fuzzy Logic			
Khan MN et la. [1]	✓	✓	20 Rules		Confidence Score of 88.6%
Yalcinkaya et la. [2]	✓	✓	81 Rules		80% Accuracy when AlexNet was combined with fuzzy logic
Glukhov et la. [3]	✓	✓		VOTT	72.1% Accuracy

Fig. 1. Image showing some sample images present in the dataset

TABLE II

CONFIGURATION TABLE SHOWING THE NETWORK CONFIGURATION OF CCN USED IN THIS STUDY. THE TABLE SHOWS THE VARIOUS CONFIGURATION SETTINGS USED FOR CCN.

Network Configuration	
Epochs	50
Learning rate	0.00001
Batch size	8
Optimizer	Adam
Weight decay	None
Dropout	0.3
Samples in training set	3256
Samples in validation set	244

TABLE IV

METRIC TABLE SHOWING THE RESULTS OF THE CNN

METRICS	RESULT
ACCURACY	0.950
RECALL	0.949
PRECISION	0.957
AREA UNDER CURVE	0.995

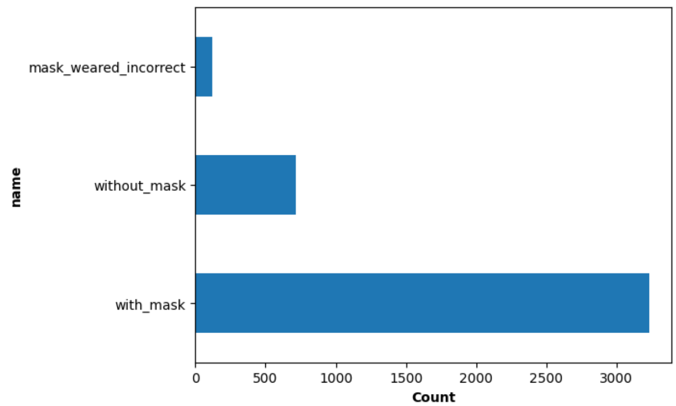

Fig. 2. Image showing the total number of bounding boxes belonging to each category – ‘masks worn incorrectly’, ‘without mask’ and ‘with mask’

IV. DISCUSSION

The results of the system show that we were able to produce the desirable results by using CNN and Fuzzy Logic in tandem to enforce social policing. However, these results come with limitations. Firstly, CNNs are unable to identify the location of different objects. Therefore, there comes the need to use region based object detection algorithms such as R-CNN, Fast R-CNN and Mask R-CNN [6]. The sole reason why we were able to directly feed our image into the fuzzy logic module was because our dataset came pre-defined with bounding box co-ordinates. When we talk in terms of scalability using region based object detection algorithms will become a must.

Secondly, finding the distance between 2D bounding boxes mean that we stay confined in a 2D spectrum. To find the distance in a crowd of people one of the most important variables is *depth*. A number of methods have been proposed for 3D object detection [7], one of these methods include the pseudo-LiDAR-based method which transforms the input image into dense artificial point clouds using existing depth estimation algorithms [8][9]. Semi-supervised learning also focuses on training models that excel in 2D and 3D object detection [10 – 12]. Lastly, for the system implemented in this study, it is necessary for the images to be in a fixed dimension. Changing dimensions could mean you have the probability to go out of bounds of the crisp values defined. Hence, the system lacks in terms of scalability.

A. Future Directions

In my opinion although this project faced severe scalability limitations it still has potential. With more dynamic fuzzy sets such as *scalable fuzzy* [13] which propose to extend the conventional fuzzy logic approach from linguistic variables to all numbers. As we explore more dynamic and scalable fuzzy sets I believe we would be looking forward to better results when paired up with CNNs and object detection algorithms.

V. Conclusion

In a nutshell, the system being implemented in this study holds immense potential in the machine learning world. Its potential to track 3D objects and classify them holds merit enough to prove that it can breach into other domains of transportation, entertainment, sports and other industries and actually thrive. Yes, this system faces limitations, these limitations however, are worth exploring. With our CNN giving an accurate result of 95% and the fuzzy logic giving desirable results hold true to its word that this topic is worth the work.

References

[1] Khan MN, Al Hasan M and Anwar S (2021) Improving the Robustness of Object Detection Through a Multi-Camera–Based Fusion Algorithm Using Fuzzy Logic. *Front. Artif. Intell.* 4:638951. doi: 10.3389/frai.2021.638951

[2] Fikret Yalcinkaya* , Ali Erbas (2021) Convolutional Neural Network and Fuzzy Logic-based Hybrid Melanoma Diagnosis System doi: http://dx.doi.org/10.5755/j02.eie.28843

[3] Dmitry Glukhov, Rykhard Bohush, Juho Mäkiö and Tatjana Hlukhava (2017) A Joint Application of Fuzzy Logic Approximation and a Deep Learning Neural Network to Build Fish Concentration Maps Based on Sonar Data. https://ceur-ws.org/Vol-2353/paper11.pdf

[4] Vishnu Kumar TV, John A, Vighnesh M, Jagannath M. Social distance monitoring system using deep learning and entry control system for commercial application. Mater Today Proc. 2022;62:4605-4611. doi: 10.1016/j.matpr.2022.03.077. Epub 2022 Mar 11. PMID: 35291397; PMCID: PMC8914536.

[5] Lee, HS., Shin, BS. Potato Detection and Segmentation Based on Mask R-CNN. J. Biosyst. Eng. 45, 233–238 (2020). https://doi.org/10.1007/s42853-020-00063-w

[6] S. Sood, H. Singh, M. Malarvel and R. Ahuja, "Significance and Limitations of Deep Neural Networks for Image Classification and Object Detection," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2021, pp. 1453-1460, doi: 10.1109/ICOSEC51865.2021.9591759

[7] Yang, Lei & Zhang, Xinyu & Li, Jun & Wang, Li & Zhu, Minghan & Zhang, Chuang & Liu, Huaping. (2023). Mix-Teaching: A Simple, Unified and Effective Semi-Supervised Learning Framework for Monocular 3D Object Detection. IEEE Transactions on Circuits and Systems for Video Technology. PP. 1-1. 10.1109/TCSVT.2023.3270728.

[8] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, “Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving,” in CVPR, 2019, pp. 8445–8453

[9] M. Ding, Y. Huo, H. Yi, Z. Wang, J. Shi, Z. Lu, and P. Luo, “Learning depth-guided convolutions for monocular 3d object detection,” in CVPRW, 2020, pp. 4306–4315

[10] H. Wang, Y. Cong, O. Litany, Y. Gao, and L. J. Guibas, “3dioumatch: Leveraging iou prediction for semi-supervised 3d object detection,” in CVPR, 2021

[11] N. Zhao, T.-S. Chua, and G. H. Lee, “Sess: Self-ensembling semisupervised 3d object detection,” in CVPR, 2020, pp. 11 076–11 084

[12] J. Yin, J. Fang, D. Zhou, L. Zhang, C.-Z. Xu, J. Shen, and W. Wang, “Semi-supervised 3d object detection with proficient teachers,” in ECCV, 2022, pp. 727–743.

[13] Mirbeygi Moghaddam, Amirreza & Kinsner, Witold & Sepehri, Nariman. (2022). Scalable Fuzzy Systems. 10.11159/cist22.132.

TABLE III

FUZZY RULE TABLE THAT SHOWS ALL FUZZY RULES THROUGH WHICH WE INFER SOP SEVERITY

INDEX	RULE	OUTPUT
1.	distance['close'] & bounding_box_size['large'] & local_box_density['high']	sop_severity['severe']
2.	distance['close'] & bounding_box_size['large'] & local_box_density['medium']	sop_severity['severe']
3.	distance['close'] & bounding_box_size['large'] & local_box_density['low']	sop_severity['moderate']
4.	distance['close'] & bounding_box_size['average'] & local_box_density['high']	sop_severity['severe']
5.	distance['close'] & bounding_box_size['average'] & local_box_density['medium']	sop_severity['severe']
6.	distance['close'] & bounding_box_size['average'] & local_box_density['low']	sop_severity['severe']
7.	distance['close'] & bounding_box_size['small'] & local_box_density['high']	sop_severity['severe']
8.	distance['close'] & bounding_box_size['small'] & local_box_density['medium']	sop_severity['severe']
9.	distance['close'] & bounding_box_size['small'] & local_box_density['low']	sop_severity['severe']
10.	distance['moderate'] & bounding_box_size['large'] & local_box_density['high']	sop_severity['severe']
11.	distance['moderate'] & bounding_box_size['large'] & local_box_density['medium']	sop_severity['severe']
12.	distance['moderate'] & bounding_box_size['large'] & local_box_density['low']	sop_severity['moderate']
13.	distance['moderate'] & bounding_box_size['average'] & local_box_density['high']	sop_severity['severe']
14.	distance['moderate'] & bounding_box_size['average'] & local_box_density['medium']	sop_severity['severe']
15.	distance['moderate'] & bounding_box_size['average'] & local_box_density['low']	sop_severity['moderate']
16.	distance['moderate'] & bounding_box_size['small'] & local_box_density['high']	sop_severity['severe']
17.	distance['moderate'] & bounding_box_size['small'] & local_box_density['medium']	sop_severity['moderate']
18.	distance['moderate'] & bounding_box_size['small'] & local_box_density['low']	sop_severity['safe']
19.	distance['far'] & bounding_box_size['large'] & local_box_density['high']	sop_severity['severe']
20.	distance['far'] & bounding_box_size['large'] & local_box_density['medium']	sop_severity['moderate']
21.	distance['far'] & bounding_box_size['large'] & local_box_density['low']	sop_severity['safe']
22.	distance['far'] & bounding_box_size['average'] & local_box_density['high']	sop_severity['severe']
23.	distance['far'] & bounding_box_size['average'] & local_box_density['medium']	sop_severity['safe']
24.	distance['far'] & bounding_box_size['average'] & local_box_density['low']	sop_severity['safe']
25.	distance['far'] & bounding_box_size['small'] & local_box_density['high']	sop_severity['severe']
26.	distance['far'] & bounding_box_size['small'] & local_box_density['medium']	sop_severity['safe']
27.	distance['far'] & bounding_box_size['small'] & local_box_density['low']	sop_severity['safe']

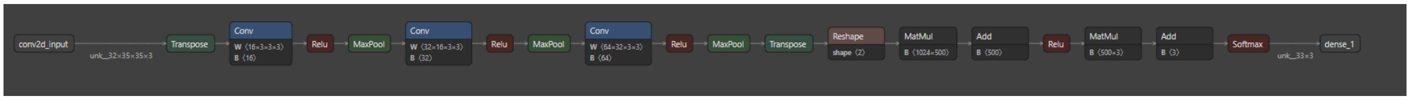

Fig. 3 The image shows the architecture of the CNN used through its ONNX file via the netron.app website.

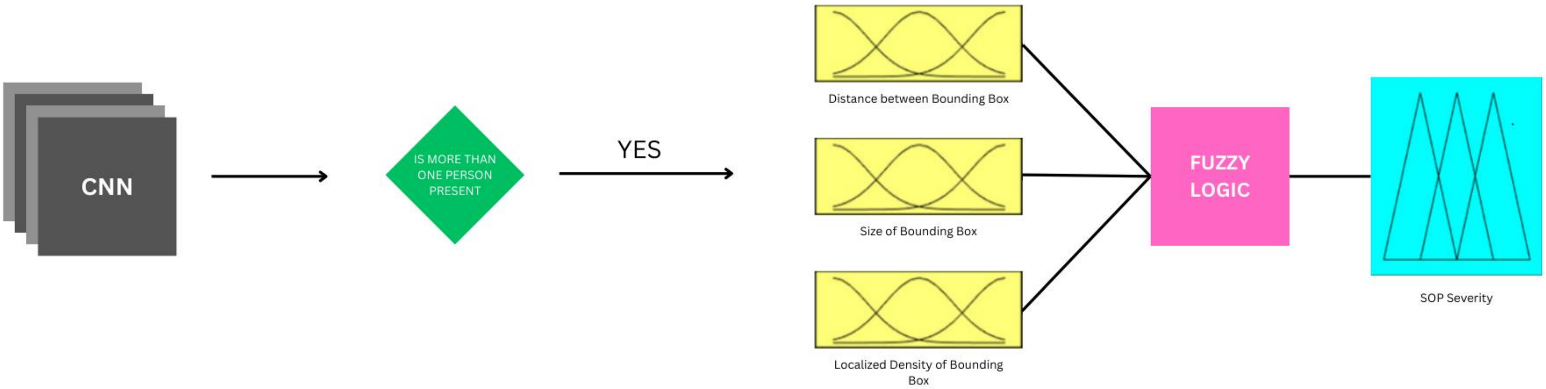

Fig. 4. The image shows the pipeline of the entire workflow. The pipeline starts with the Convolution Neural Network and ends with the membership function of our desired output – in this case is the SOP severity.

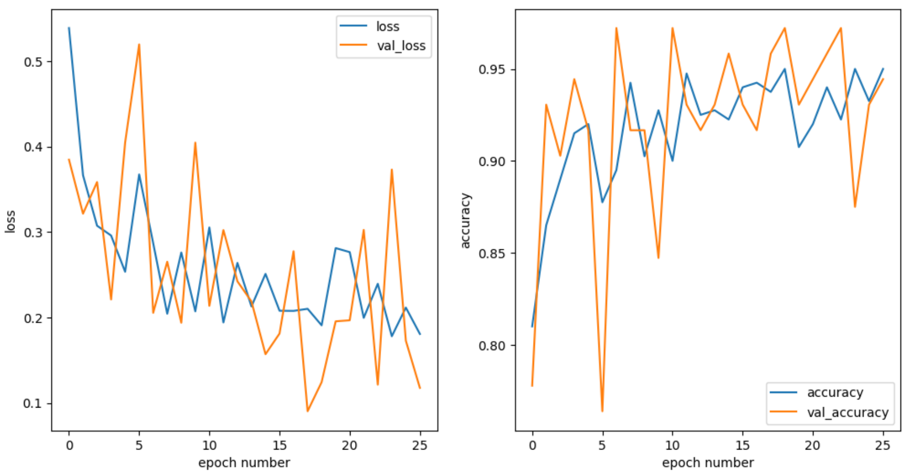

Fig. 5. The image shows the Loss and Accuracy of the CNN and the comparison between the training data and the validation data.

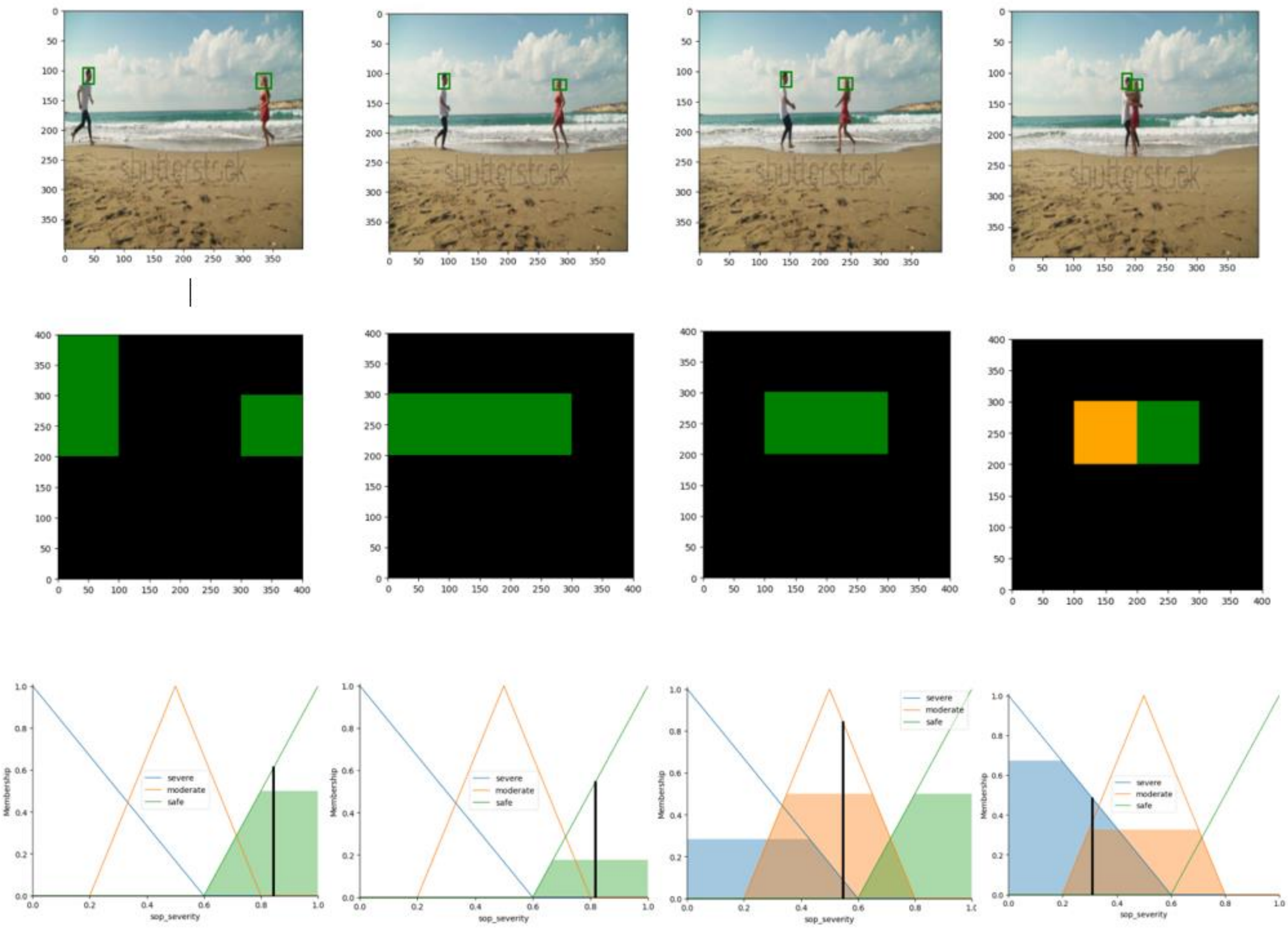

Fig. 6. Shows the initial images with their respective bounding boxes. The second layer shows the localized bounding box density. Lastly the third layer shows the membership function of SOP severity. As the man and woman in the initial picture get closer you can see the black line moving from the 'safe' zone (green) to the 'moderate' zone (orange) and finally at the 'severe' zone shaded in blue.