# Neural network based face mask detection using thermal imaging sensors

C. Elwinger, L. Elwinger, P. Arnold, N. Schmitz
*Department of Engineering, Pforzheim University*
*mailto:norbert.schmitz@hs-pforzheim.de*

The data processing of cameras with the help of artificial intelligence has arrived in areas of our daily routine. A trained artificial intelligence can precisely identify persons to unlock secured data. Today, it is already state of the art to recognize faces precisely with the help of cameras and to derive actions from them. Since the corona pandemic is omnipresent, there are many ideas to be explored. Medical masks will probably protect against infection and support the fight against the pandemic. For this reason, masks are very often mandatory in public buildings and transportation. Since the wearing of protective masks is often perceived as restrictive and annoying, many organizations, such as an university or a grocery store, are forced to hire additional inspectors. With a mechatronic system that recognizes masks, it would be possible to reduce the additional work and increase the safety of the staff.

## 1 Introduction

Nowadays, intelligent cameras are already fully integrated into our everyday life. This area is already very well researched and has yielded very good results. The general level of knowledge about image recognition from infrared (IR) camera data, however, is much lower. The possibilities are already partially applied in industrial manufacturing processes, but there is still a lot of unused potential. The main task of the present research project is to investigate the data generation of two different IR cameras and the conceptual design and implementation of an artificial intelligence. By combining these components, the question "Is mask detection by AI feasible using IR data?" should be answered. The system should detect whether the scanned person is wearing a protective mouth-nose mask or not.

## 2 State of the art

### *2.1 Related work*

The research and development of mask detectors received a boost from the corona pandemic. Mainly color images are used for this purpose, like the SSDMNV2 [1]. A big benefit here is the very large amount of training data available. The SSDMNV2 is a system that uses a color camera to detect facial masks with an accuracy of 0.9264 and an F1 value of 0.93. In the past, only a few facial recognition systems have been developed using thermal imaging cameras. However, the basic function of facial recognition is possible [2]. The image data is further processed to provide information on exact temperature values. The system achieves a total mean absolute error and a root mean square error compared to an industrial instrument of 0.375℃ and 0.439℃, respectively.

### *2.2 Outcome*

The analysis of the above-mentioned publications provided insights into the quality and quantity of the data set. Fine-tuning pre-trained models or the use of data augmentation can redue the problem of small datasets. Both sources have shown that the detection with a SSMD or YOLO can be useful.

## 3 Proposed methodology

In the context of this work, two approaches are considered. One is object classification and the other is object localisation with YOLO. For both concepts, data sets were created, neural networks trained and the results evaluated. The research question could already be answered with the first approach "object classification". In the further course, the network for object localisation was implemented in a demonstrator.

## 4 Data set

***Fig. 1*** *Example images from the dataset with mask (top row) and without mask (bottom row)*

There are still no open source data sets for mask detection using an IR camera. Therefore, data sets for object classification and object localisation had to be created within the scope of this work. The captured

DGaO Proceedings 2022 – http://www.dgao-proceedings.de – ISSN: 1614-8436 – urn:nbn:de:0287-2022-B009-9
submitted: 29.07.2022 – published: 05.09.2022

datasets fulfil the basic requirements of sharp and clear images, sufficient training data, high variation of geometry and position as well as the balanced appearance of classes. A data set was created for each of the two approaches. Both data sets contain two classes, people with mask and people without mask. The above figure shows a small section of the data sets. The final data set for the demonstrator consists of 1134 images for training and 251 images for validation.The data for validation represents 18% of the total number of images.

## 5 Object localization

In the field of real-time methods, the YOLO approach has become very important. YOLO is very fast, because detection is considered a regression problem and does not require a complex pipeline [3]. In addition, YOLO achieves twice the average accuracy compared to other real-time systems. By using the open source repository (https://github.com/ultralytics/yolov5) it was possible to use a pre-trained YOLO network. The pre-trained model is called YOLOv5s and has 7.3 million parameters. The existing model was re-trained and fine-tuned with the self-generated data.

### *5.1 Training results*

After the last training session, a precision of 0.9 has been achieved with no tendency to increase further. The confusion matrix shown below represents a comparison between the labels, the model prediction, and the actual labels that the model should predict. Both the strengths and weaknesses of the model can be seen here. The system achieves a very high accuracy value, but also has many background errors. However, background errors can be reduced by adjusting thresholds in the application.

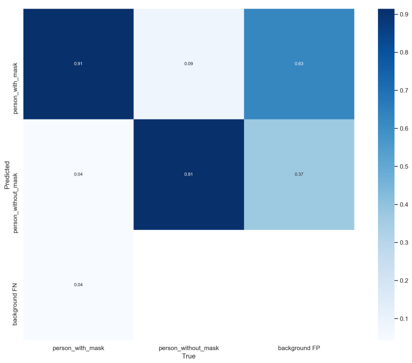


***Fig. 2*** *Confusion matrix of the final training results*

## 6 Summary and conclusion

The research question ”Is it possible to implement a mask recognition using a thermal imaging camera” could be answered with two different approaches. Both the first approach ”using a CNN” and the second approach ”using YOLO” can answer the research question positively. Within the project it was possible to generate image data from a Lepton 2.5 and to process them afterwards with different neural networks. By combining the developed subsystems, it was possible to implement a selected approach in a demonstrator. The system can test several people at the same time for wearing a face mask and display the result via a graphical interface as depicted below. Persons wearing a mask are marked with a green label, persons not wearing a mask with a red label.

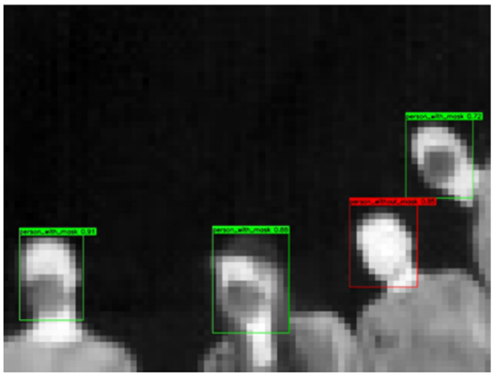

***Fig. 3*** *Real-time capable demonstrator with detection of persons with and without mask*

The existing potential of thermal cameras is not yet fully exploited today and is still rather rarely used. The use of thermal cameras makes it possible to solve problems in a new way. The resulting system works well and reliably in the first place. The research question could be answered positively by the results. However, there is a lot of potential to improve the system. Points such as environment dependency or the susceptibility to background errors could be worked on further in the future.

## References


[1] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and J. Hemanth, “SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2,” Sustain Cities Soc **66**, 102692 (2020).

[2] J.-W. Lin, M.-H. Lu, and Y.-H. Lin, “A Thermal Camera Based Continuous Body Temperature Measurement System,” in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 1681–1687 (2019).

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788 (2016).