ADRAS: Airborne Disease Risk Assessment System for Closed Environments*

Wilber Rojas[1]✉ ⓘ, Edwin Salcedo[1] ⓘ, and Guillermo Sahonero[1,2] ⓘ

[1] Department of Mechatronics Engineering
Universidad Católica Boliviana "San Pablo", La Paz, Bolivia
{wrojas,esalcedo}@ucb.edu.bo
https://www.ucb.edu.bo/
[2] Institute for Biological and Medical Engineering,
Pontificia Universidad Católica de Chile, Santiago, Chile
gsahonero@uc.cl

Abstract. Airborne diseases are easy to spread in any population. The advent of COVID-19 showed us that we are not prepared to control them. The pandemic has drastically posed challenges to the daily functioning of public and private establishments. In general, while there have been several approaches to reduce the potential risk of spreading the virus, many of them rely on the commitment that people make, which - unfortunately - cannot be constant, for example, wearing a facemask in closed environment at all times or social distancing. In this work, we propose a computer vision system to determine the risk of airborne disease spread in closed environments. We modify and implement the Wells-Riley epidemiological equation. We also evaluate and implement models for facemask and person detection from OpenVino. For mask detection, we applied transfer learning and obtained the best performance for a model based on MobileNetV2. The generated data from several devices is visible in a web platform to monitor multiple areas and locations. Finally, an OAK-D camera and a Jetson device are embedded in a end device meant to monitor a closed environment and send spread risk data continually to the web platform. Our results are promising as we obtained up to 88% of accuracy for the person detection task and up to 57% of mAP for the facemask task. We expect this paper to be beneficial for developing new control measurements and prevetion tools to prevent airborne contagion.

Keywords: Airborne Disease · Risk Assessment · Stereo Vision · Edge Computing.

1 Introduction

In the past few years, society has become more conscious of airborne diseases due to the disruption caused by the COVID-19 pandemic. This aiborne disease,

---

also known as COVID-19, is caused by the SARS-CoV-2 virus and has different symptoms such as fever, breathing difficulties, fatigue, tiredness, cough, and loss of taste and smell [1]. Different from other illnesses transmitted between people, viruses or bacteria of airborne diseases can stay in the air mixed with dust particles and respiratory droplets for longer times [1]. Then, these particles are eventually inhaled by other people and cause a spread of the disease. Airborne diseases comprise Measle, Tuberculosis, Chickenpox, Influenza, Pertussis, SARS-CoV-1, and SARS-CoV-2, among others. Even though the contagious rate and symptoms vary between them, their control and prevention are similar and consist mainly of the installation of isolation rooms for infected people, use of protective clothes such as Personal Protective Equipment (PPE), facemasks, and gloves, better ventilation for closed environments, and stricter practices of sanitation and hygiene. In-depth investigations on SARS-CoV-2 spread proved that the closer or denser a group of people, the higher the risk of airborne disease contagion [1].
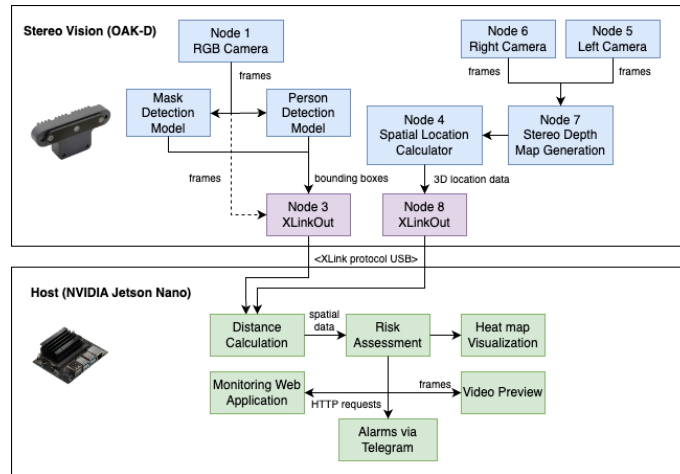

Fig. 1: End device architecture

The current research aims to propose a computer vision-based system to monitor the spread risk of an airborne disease. We considered Coronavirus disease 2019 as a study case for the investigation. So far, it has spread through 591 million people and caused 6.4 million deaths worldwide as of August 19, 2022 according to [2]. This disease has completely disrupted the entire world unexpectedly, consequently, there is a constant research need of new control and prevention methods for new airborne viruses or variants since a rapid increase of infected people can collapse any healthcare system. In the current investigation, we present a real-time hardware and software system that calculates the risk of airborne diseases in one or multiple closed environments. The architecture of the

monitoring end device is illustrated in Figure 1. The model we use for assessing the risk is based on the number of people, compliance with the correct use of masks, and compliance with social distancing and exposure time as variables.

The rest of the paper is structured as follows. We first review some related works in Section 2. Then, we adapt an epidemiological model by considering several concepts from Wells-Riley estimation model in Section 3. Later, we explain the details of the computer vision models for person and mask detection in Section 4 which is complemented with the description of the distance estimation in Section 5. Moreover, we explain how the monitoring system works in Section 6 and present our results discussing their implications in Section 7. Finally, we conclude the paper in Section 8.

2   Related Works

By the end of 2019, the initial breakthrough of the SARS-CoV-2 virus led to a massive number of deaths and the declaration of a worldwide pandemic. This encouraged many computer vision developers and researchers to collaborate towards the development of new ideas to prevent the virus spread. Since keeping physical distance of at least one meter from others has proved to be one of the most effective measure against SARS-CoV-2 [2], extensive research has been carried out for the development of Visual Social Distance Monitoring Systems (VSDMS) [3]. In addition, other projects complemented the distance measuring idea with face masks detection, or face-hand interaction [4] [5] [6] to control that people comply with measures against a contagion.

The state-of-the-art proposals to measure physical contact among people can be classified as either 2D based or 3D based. The former commonly uses a sequence of methods to recognize people: a) image processing; b) image segmentation; c) shape extraction; d) object recognition. This last step might vary from classical computer vision-methods [7] to deep learning-based methods [8], [4], [9], [5], [6]. Focusing on deep learning methods, researchers usually implement person detectors based on Convolutional Neural Networks (CNNs) such as YOLOv4, Yolov5, MobileNet, SSD, or R-CNN [10]. Then, the investigations regularly calculate the pairwise Euclidean distances among the centroids of the detected bounding boxes [4], [3]. These models are commonly applied to video frames given that monitoring an environment requires continuous surveillance. Consequently, recent proposals combine person detectors with object tracking algorithms such as DeepSORT, SORT, or StrongSort [11][12].

Even though camera-based surveillance systems have been developed up to the point of becoming commercialized solutions [13] [14] [15], this brief literature review let us note that 3D vision has been less explored despite its higher accuracy for distance calculation. To the authors' best knowledge, researchers in [16] have been the first to publish about the application of stereo vision in a VSDMS in April 2022. Not only did they combine stereo and monocular cameras for person detection, but also conclude that stereo vision cameras were superior than monocular cameras. Specifically, they use a Zed M Camera and an MSI

laptop equipped with NVIDIA GTX 1060 3GB GPU, with which they stream videos and obtain depth maps to measure the distance among people. Beyond the published work in indexed venues, developers in [17] and [18] applied stereovision cameras for social distance monitoring in September 2020 and January 2021, respectively.

Beyond social distancing, only a few works explore further implications and contagion risk assessment by leveraging the extracted visual information of the protective measures compliance. For instance, investigators in [12] contributed with a remarkable online infection risk assessment scheme for open environments named DeepSocial, which considers people's moving trajectories and the rate of social distancing violations to calculate the contagion risk. Moving on, researchers in [19] propose BEV-NET to assess social distancing compliance and probability of infection in closed environments using a monocular camera from a top perspective. Furthermore, this investigation proposes the COVID-19 infection risk assessment for each individual present in a scene and a general risk assessment for the complete ambient. In both projects, the closer or denser a group of people is, the higher the risk of contagion [19], [12].

In contrast to COVID-19, the infection risk quantification of other airborne diseases by recognising visually the protective rule's compliance is less investigated, however, their risk monitoring can also be implemented using the approaches proposed for COVID-19. The multiple waves of SARS-CoV-2 have taught us that airborne diseases can severely affect the economy and normal functioning of an entire country, therefore, there is a shortage of research to have better tools that monitor and inform the risk of contagion in an ambient. All these to provide confidence to citizens and users in outdoor and indoor environments.

3   Epidemiological Model

This section describes the definition of the equation for estimating the spread risk of a airborne disease in a closed environment. We know that the classification of risk prediction models can be split into Wells-Riley based and Dose-Response based models [20]. We analyze the first one and modify it so that the implementation is feasible in a computer vision monitoring environment.

Wells developed an equation to estimate the risk of infection in a close environment [21]. Riley proposed an improvement to the equation, considering the ventilation of the room as a parameter [22]. The Wells-Riley equation generalizes the infectivity of pathogens with a new infectious dose unit called *quanta*. A *quanta* is the number of infectious particles required to infect a person [23]. So a *quanta* of influenza would infect the same number of people as a *quanta* of tuberculosis or COVID-19. If the disease is more contagious, the infected person would have a higher *quanta* emission rate.

The Wells-Riley equation is defined by Equation 1:

$$P_i = 1 - exp\left(\frac{I * q * p * t}{Q}\right) \qquad (1)$$

Where, P_i is the probability of infection, I is the number of infected people, q is the quanta emission rate, p is the pulmonary ventilation rate of a person, t is the exposure time, and Q is the ventilation rate of the room [23]. For this investigation, each of these parameters were obtained as follows:

– Number of infected people (I):
  Wells-Riley estimates the risk of infection based on a certain number of infected people [23], however, we cannot know this information with entire certainty. We calculate the probability of infected people based on the population percentage of cases in a region:

$$p_c = \frac{c}{P_r} \qquad (2)$$

Where, p_c is the percentage of cases, c is the number of cases in a region, and P_r is the total population of the region. If we multiply the percentage of cases by the number of people in a room, we obtain the probability of infected people in that room. So, our value for I is:

$$I = N * p_c \qquad (3)$$

Where N is the number of people.

– Quanta emission rate (q):
  The quanta generation rate is the only parameter that contains the infectivity of the virus [23] [11], so each pathogen has its own value of q. Mikszewski et al. [24], analyzed the quanta generation rate for the most common airborne diseases, including SARS-CoV-1, SARS-CoV-2, MERS, Tuberculosis, and Influenza. These values are constants in the monitoring system and are described in Table 1.

Table 1: Quanta emission rate values $[q * h^{-1}]$ [24].

Pathogen	Resting, oral breathing	Standing, Speaking	Light activity, speaking loudly
SARS-CoV-1	0.008	0.042	0.71
MERS	0.011	0.056	0.96
Tuberculosis (On Treatment)	0.020	0.098	1.70
Influenza	0.035	0.17	3.00
Coxsackievirus	0.062	0.31	5.20
Rhinovirus	0.210	1.00	18.00
SARS-CoV-2	0.550	2.70	46.00
Tuberculosis (Untreated)	0.62	3.1	52.00
Adenovirus	0.780	3.90	66.0
Measles	3.100	15.00	260.00

Li et al. determined that the viral load, therefore also the quanta, is almost the same between presymptomatic, asymptomatic, and symptomatic subjects [25]. In the case of the advent of a new airborne disease or actual disease variant, the value of q should be calculated using Equation 4:

$$q = c_v * c_i * p * v_d \tag{4}$$

where, c_v is the viral load, c_i is a conversion factor between a quanta and the infectious dose, p is the inhalation rate, and v_d is the volume of droplets expelled by a person [26].

– Pulmonary ventilation rate (p):
Adams [27], conducted a study where he empirically determined the average person inhalation rate for different activities. Table 2 shows the values obtained from this study, which are used in Equation 1.

Table 2: Inhalation rate values [27].

Activity	Inhalation Rate [m3 h-1]
laying down	0.49
stand	0.54
very light exercise	0.72
light exercise	1.38
moderate exercise	2.35
heavy exercise	3.30

– Ventilation rate of the room (Q):
It refers to the ACH (Air Changes per Hour) value of a close environment. To obtain this parameter, first we determined the air flow rate.

$$A_{FR} = w * A_v \tag{5}$$

In Equation 5, A_{FR} is the air flow rate, w is the window area, and A_v is the air velocity. Finally, the ACH is the Air flow rate divided by room volume (V_R)[28].

$$ACH = \frac{A_{FR}}{V_R} \tag{6}$$

In case of having artificial ventilation, the ACH value can be obtained from the specifications of the machine, and should be added to the natural ventilation calculated in the equation 6.

– Exposure time (t):
This parameter refers to the time that "I" number of infected people will remain in the closed environment. Note that the Wells-Riley formula requires the total exposure time as a parameter. And, in this research, the objective is to implement it in a monitoring system, so the number of infected people for a specific time is variable and will be calculated in real time.

We can represent the Wells-Riley equation with the use of integrals for the time variation:

$$P_i = 1 - exp\left(\int_0^\infty \frac{I * q * p}{Q} dt\right) \qquad (7)$$

– Facemask detection:
  Wells-Riley does not consider if people wear facemasks, which is required for estimating the spread of a disease in a closed environment. Therefore, we implement an additional parameter to the equation considering that we will monitor the presence of facemasks on the detected people's faces.

  It is well-known that the worst type of facemasks are made of cloth, so we consider it as the default type of facemask used by everybody detected in a scene. It is worth noting that the use of cloth facemask reduces the contagion risk of an airborne disease by half [29]. From the Wells-Riley equation, we know that if we double the ventilation rate, we also halve the risk of spread. Consequently, the final equation is:

$$P_i = 1 - exp\left(\int_0^\infty \frac{I * q * p}{Q * (1 + M)} dt\right) \qquad (8)$$

  Where M is the percentage, in range $[0, 1]$, of people wearing a mask in a closed environment. To conclude the equation requires setting the ventilation of the room (Q), the pulmonary ventilation rate (p) and the Quanta generation rate (q) as constants before monitoring. The exposure time (t), the estimated number of infected people (I), and the percentage of masks (M) will be calculated in real time by computer vision. Equation 8 is the Risk Assessment block in Figure 1.

4 Person and Mask Detection

In the proposed system, two CNN-based object detection models localize person and facemask appearances in the frames coming from the OAK-D camera. The collected datasets, the object detection models, and most importantly, the used metrics are presented in the following subsections.

4.1 Dataset and Preprocessing

We collected the dataset of person and mask instances separately through three different means: web scrapping, video processing, and public datasets. First, we developed a Python script to download images from Google images. The terms used to find people images were: "pedestrians", "people in room", and "meeting". Once we obtained 377 images, we needed to review their quality due to some unrelated images downloaded by mistake. Second, we used the Computer

Vision Annotation Tool (CVAT) [30] to obtain image samples of people and face-masks independently by labelling video frames. This let us get 1,084 instances of people and 337 instances of facemasks. Finally, we obtained 756 images by combining public datasets [31]. All these subsets were sorted to create two sub-sets: one with instances of people without facemasks, and another with images with facemask instances. Since the subsets were small, we needed to implement data augmentation techniques, such as horizontal flip, brightness change, and grayscale, by using the Roboflow platform [32].

Table 3: Datasets

Subset	Collected	Augmented	Total
Person Instances	1084	2277	3361
Facemask Instances	337	606	943

4.2   Object Detection

First, instead of creating and training object detection models from scratch, the person detection models that we tried were based on pre-trained architectures provided by Intel OpenVino [33]. Specifically, we tested the models 106, 200, 201, 202, 203, 301, 302, and 303 to achieve a good performance with our collected dataset. In addition, it is worth mentioning that OpenVino models can be easily deployed on OAK-D devices by using the MiryiadX Blob Converter [34], which made us decide to use them. Secondly, for facemask detection, we applied transfer learning to re-train the models Mobilenet v2, which had been previously trained with bigger datasets. It is important to note that several developers who worked with OAK-D devices defined Mobilenetv2-based models as the best object detectors to deploy on them [33].

4.3   Metrics

Intersection Over Union: The main tool to evaluate the human and facemask detection models with respect to each localized bounding box was Intersection Over Union (IoU). This metric, also known as the Jacquard Index, measures the overlap area between the ground-truth bounding boxes and the predicted bounding boxes, and ranges between 0 and 1. For object detection tasks, it is recommended to set a confidence threshold to filter good-quality detected bounding boxes. In the current project, we used 50% as threshold confidence because it gave us good experimental results.

Confusion Matrix: We calculated a confusion matrix per image where the pixels under the area of the detected objects were classified as 1; otherwise, they were classified as 0. Consequently, both predicted and ground truth masks let us quantify the well and wrong classified pixels as True Positive (TP), True

Negative (TN), False Positive (FP), or False Negative (FN). Then, the matrix let us calculate the Matthew's Correlation Coefficient, F1 score, Sensitivity (Recall), Specificity, Accuracy, Negative Predictive Value (NPV), False Positive Rate (FPR), and False Negative Rate (FNR) according to Equations 18, 17, 10, 11, 9, 12, 13, and 14, respectively. Among them, Precision shows the percentage of correct predictions among all the positive-predicted pixels in the images while Sensitivity describes the percentage of actual positives that were identified correctly.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP+FP+TN+FN}} \quad (9) \qquad \text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \quad (10)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (11) \qquad \text{NPV} = \frac{\text{TN}}{\text{TN+FN}} \quad (12)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP+TN}} \quad (13) \qquad \text{FNR} = \frac{\text{FN}}{\text{FN+TP}} \quad (14)$$

Mean Average Precision: Commonly, precision and sensitivity (recall) can be plotted against each other to obtain the precision-recall curve, and the Average Precision will be the area under this curve. This metric is defined in Equation 15, where r represents recall, p represents precision as a function of r. Therefore, $p(r)$ means "precision at recall r".

$$AP = \int_0^1 p(r) \ dr \tag{15}$$

Given that the person and facemask detection models were applied separately, Mean Average Precision (mAP) in our case is the same as Average Precision (AP). However, it is worth mentioning that mAP is the mean of Average Precisions of all individual classes for multi-class detection tasks and should be calculated as in Equation 16.

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{16}$$

Here, mAP is Mean Average Precision, N is the number of class labels, and AP_i is the Average Precision for the ith class. We evaluated and calculated mean average precision for different IoU thresholds: mAP@50% IoU, mAP@75% IoU, mAP@50%:5%:95% IoU. However, we finally kept 50% as confidence threshold.

F1 Score: Both precision and recall were used to calculate the F1 Score metric provided in Equation 17. The benefit of this metric is that it considers the number of prediction errors that the model makes and also the type of errors that are made.

$$F1Score = \frac{2 * TP}{2 * TP + FP + FN} \tag{17}$$

Matthews Correlation Coefficient: Finally, the Matthews Correlation Coefficient (MCC) was used as a measure of the quality of binary classifications of bounding boxes. This is defined in Equation 18.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \tag{18}$$

5   Distance Estimation

To calculate the distance among people, we need to estimate the relative position of the objects detected with respect of the camera. Stereo vision helped us estimate these distances and positions by obtaining a three-dimensional view of a scene through the OAK-D camera and its binocular vision. Stereo vision can be applied to calculate the depth of an object by making use of the angle of convergence.
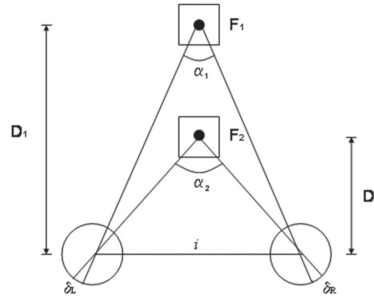

Fig. 2: Stereo vision definition, where α is the convergence angle, D is the distance between the camera and the detected object, and i is the distance between cameras. F_1 and F_2 are different objects, which have different distances, and therefore, different angles [16].

As shown in the Figure 2, the convergence angle α_2 is in the middle of the two monocular cameras that capture an object located on the front [16]. Also, this angle varies depending on the distance of the object detected, this allows a precise approximation of its real distance. We used the OAK-D device built-in functions to calculate this distance. Specifically, the function used was "Spatial Location Calculator" [35]. In order to control social distancing, we calculated the Euclidean distance between each person, defined by Equation 19, where d is

the distance between person p_1 and person p_2, and x, y, z refers to the positions in the three dimensional plane.

$$d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \qquad (19)$$

6 Monitoring System

Combining the result of Wells-Riley equation and the positions of the people determined by the stereo camera, we can analyze the sectors with the highest risk of infection in the monitoring area. We implement a Gaussian analysis to distribute the concentration of infectious particles. The Fig. 3 shows an example analysis of a room monitoring COVID-19. This analysis can only be displayed in square or rectangular shapes. And the unit of the axes is centimeters.
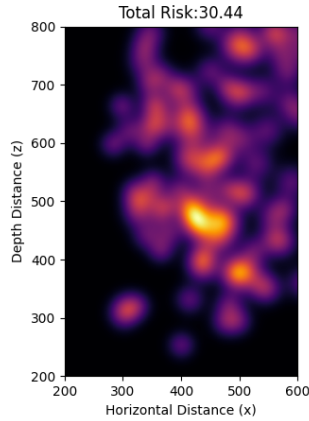

Fig. 3: Gaussian analysis of quanta concentration

In Fig. 3, a heatmap shows the sectors where people remained for the longest time. Additionally, we send the data to an online monitoring system shown in Fig. 4. This system was developed as a web application and was deployed on a cloud server, which collected the airborne spread risk calculated from multiple end devices. If the risk in an environment exceeded a threshold, the system sent notifications via Telegram to alert the authorities. Specifically, Fig. 4 (a) shows an example of the interface in which a bank can register multiple areas for real-time surveillance. The system was developed to monitor multiple areas by collecting data from end in several public places. For instance, the interface shown in Fig. 4 (b) was the detail view that helped us monitor a specific closed environment every 30 seconds.
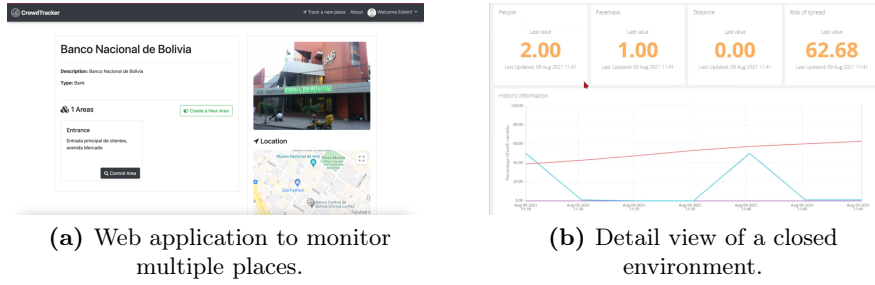
(a) Web application to monitor multiple places.

(b) Detail view of a closed environment.

Fig. 4: Online Monitoring System.

About the end device, this was designed using SolidWorks, assembled with 3D printing and installed to monitor a closed environment continually. Internally, the device consisted of a Jetson Nano computer and an OAK-D camera. The final design is shown in Fig 5. The device has a screen that shows the risk of spread, as well as, activates an alert when the calculated risk exceeds a safety threshold.

Fig. 5: Monitoring Device

7   Results and Discussion

In the validation our Wells-Riley model, we collected data from similar studies and used them as input to the equation 8. The results, in Table 4, show that our equation obtains a value close to that reported in these studies, with an average error of 4.29%. These values were expected because the equation was essentially unchanged, and the structure was only slightly modified to allow for different input values. We can conclude that the small differences between the results are due to the fact that each study proposes new parameters to consider.

Table 4: Evaluation of the modified Wells Riley equation

Input						Output		Reference
I	M	q	p	t (min)	Q	P_i	Expected P_i	
1	1	10	0.54	60	12.00	9.88	10.00	F. Velarde et al. [36]
1	1	30	0.54	4	3.86	2.76	2.50	Z. Liu et al. [37]
1	1	100	0.36	60	15.00	21.34	21.20	Y. Guo et al. [38]
1	1	14	0.35	60	0.56	1.44	1.40	Z. Wang et al. [39]
1	1	14	0.35	420	0.56	10.30	9.70	Z. Wang et al. [39]

Regarding person detection model, we evaluated several pretrained models provided by OpenVino on its Github repository [33]. These tests were performed with the total subset named "Person Instances", described in Table 3. As mentioned above, we mainly used these models because they were already trained with bigger datasets for person detection, in consequence, the obtained results were better than implementing detection models from scratch. The resulting metrics are shown in Table 5, which shows the models' code in the first row. Note that all models in [33] have the term "person-detection-0" as the prefix name and a unique code as an identifier. In addition, each model has three versions that vary in their weights' precision: "float point 16" (FP16), "float point 16 - int 8" (FP16-INT8), and "float point 32" (FP32). We focused on testing the FP16 version of the models for better memory usage when deployed in the OAK-D device. Based on the evaluation results, the model person-detection-0106 obtained the best results according to the MCC, IoU and Accuracy metrics. The final evaluation results are presented in Table 5, where the row "Complexity" defines the number of computational operations to pass a frame through the model, and the row "Size" defines the memory footprint needed for each model.

Table 5: Evaluation results for the tested models for person detection [33]. Complexity row is in GFLOPS. Size row is in MParams. IoU, Accuracy, mAP, F1 Score, Sensitivity, Specificity, NPV, FPR, and FNR are in the [0;1] interval. MCC is in the [-1;1] interval.

Metric/Model	106	200	201	202	203	301	302	303
Complexity	404.264	0.786	1.768	3.143	6.519	79318.216	370.208	24.758
Size	71.565	1.817	1.817	1.817	2.394	55.557	51.164	3.630
IoU	0.829	0.605	0.605	0.605	0.716	0.632	0.657	0.657
mAP	0.737	0.021	0.021	0.021	0.545	0.335	0.370	0.376
MCC	0.630	-0.001	-0.001	-0.001	0.366	0.131	0.172	0.186
Accuracy	0.884	0.706	0.706	0.706	0.816	0.749	0.770	0.768
F1 Score	0.884	0.706	0.706	0.706	0.816	0.749	0.770	0.769
Sensitivity	0.683	0.001	0.001	0.001	0.510	0.385	0.371	0.428
Specificity	0.962	0.999	0.999	0.999	0.864	0.742	0.801	0.757
NPV	0.869	0.706	0.706	0.706	0.817	0.796	0.788	0.801
FPR	0.045	0.001	0.001	0.001	0.135	0.258	0.197	0.242
FNR	0.316	0.999	0.999	0.999	0.490	0.615	0.629	0.572

Finally, for mask detection, we trained a transfer learning-based model with Tensorflow using the subset "Facemask Instances". This task was performed using the training instructions for the OAK-D camera provided by Luxonis [40], and the selected base model was MobileNetV2. From training, we obtained 57.98% of mAP result with IoU 0.50. Finally, the facemask model and the best person detection models were deployed on the OAK-D camera. A sample of detection results using both models can be seen in Figure 6.
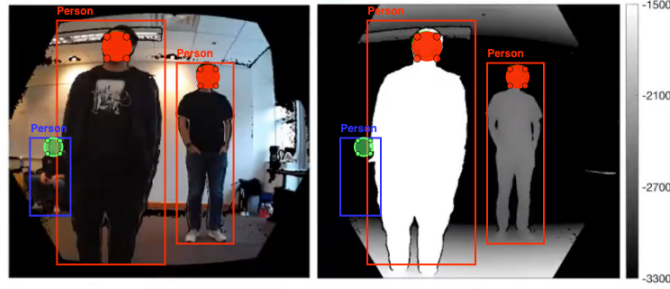

Fig. 6: Object Detection Inferences

8   Conclusions

In this paper, we proposed the implementation of epidemiological models in computer vision systems, with the aim of reducing the spread of airborne diseases. We adapted the Wells-Riley equation to be able to calculate the risk of infection in real time. We use object detection models to determine the number of people in an environment, as well as the number of people wearing masks. We implement stereo-cameras and Gaussian mathematical models to obtain a Heatmap of the sections with the highest risk of infection. All information is sent to an online monitoring system, with which multiple environments can be monitored. Our results, although preliminary, are promising and contribute by bridging the gap between computer vision, health care and spread modeling. As future work, the facemask detector requires to be more robust. Another important limitation is that our percentage of infected (I) is an estimate and it cannot be guaranteed that it is the real value of infected in the environment. An oversize can be added to this value in order to prevent false security values. Something to consider, is that no ethical protocol is needed. Although the system monitors people, no private data or video recording is saved, only general data, such as the number of people and how many have a mask.

References

1. S. E. Park, "Epidemiology, virology, and clinical features of severe acute respiratory syndrome -coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19)," *Clinical and experimental pediatrics*, vol. 63, no. 4, pp. 119–124, apr 2020. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32252141

2. "WHO: Coronavirus (COVID-19) Statistics." [Online]. Available: https://covid19.who.int/

3. Y. Himeur, S. Al-Maadeed, N. Almaadeed, K. Abualsaud, A. Mohamed, T. Khattab, and O. Elharrouss, "Deep visual social distancing monitoring to combat COVID-19: A comprehensive survey," *Sustainable Cities and Society*, vol. 85, p. 104064, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670722003821

4. I. Ahmed, M. Ahmad, J. J. P. C. Rodrigues, G. Jeon, and S. Din, "A deep learning-based social distance monitoring framework for COVID-19," *Sustainable Cities and Society*, vol. 65, p. 102571, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670720307897

5. M. Razavi, H. Alikhani, V. Janfaza, B. Sadeghi, and E. Alikhani, "An automatic system to monitor the physical distance and face mask wearing of construction workers in covid-19 pandemic," *CoRR*, vol. abs/2101.0, 2021. [Online]. Available: https://arxiv.org/abs/2101.01373

6. F. I. Eyiokur, H. K. Ekenel, and A. Waibel, "A Computer Vision System to Help Prevent the Transmission of COVID-19," *undefined*, 2021.

7. N. Petrovic and D. Kocić, *IoT-based System for COVID-19 Indoor Safety Monitoring*, sep 2020.

8. S. Degadwala, D. Vyas, H. Dave, and A. Mahajan, "Visual Social Distance Alert System Using Computer Vision Deep Learning," *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020*, pp. 1512–1516, nov 2020.

9. D. Yang, E. Yurtsever, V. Renganathan, K. A. Redmill, and Ü. Özgüner, "A Vision-Based Social Distancing and Critical Density Detection System for COVID-19," *Sensors 2021, Vol. 21, Page 4608*, vol. 21, no. 13, p. 4608, jul 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/13/4608/htmhttps://www.mdpi.com/1424-8220/21/13/4608

10. O. Karaman, A. Alhudhaif, and K. Polat, "Development of smart camera systems based on artificial intelligence network for social distance detection to fight against COVID-19," *Applied Soft Computing*, vol. 110, p. 107610, oct 2021.

11. J. Li and Z. Wu, "The application of yolov4 and a new pedestrian clustering algorithm to implement social distance monitoring during the covid-19 pandemic," *Journal of Physics: Conference Series*, vol. 1865, no. 4, apr 2021. [Online]. Available: https://doi.org/10.1088/1742-6596/1865/4/042019

12. M. Rezaei and M. Azarmi, "DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic," *Applied Sciences*, vol. 10, no. 21, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/21/7514

13. V. Kamalasanan and M. Sester, "Living with Rules: An AR Approach," *Adjunct Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2020*, pp. 213–216, nov 2020.

14. "Delta variant: 8 things you should know | Coronavirus | UC Davis Health." [Online]. Available: https://health.ucdavis.edu/coronavirus/covid-19-information/delta-variant.html

15. "Covid-19 vaccine tracker: View vaccinations by country." [Online]. Available: https://edition.cnn.com/interactive/2021/health/global-covid-vaccinations/

16. M. L. M.D., N. Varble, B. T. M.D., S. Xu, and B. J. W. M.D., "Camera-based distance detection and contact tracing to monitor potential spread of COVID-19," in *Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*, C. R. Mello-Thoms and S. Taylor-Phillips, Eds., vol. 12035, International Society for Optics and Photonics. SPIE, 2022, p. 120351D. [Online]. Available: https://doi.org/10.1117/12.2612846

17. ibaiGorordo, "Ibaigorordo/social-distance-feedback-for-the-blind: A social distancing feedback system for the blind using the oak-d camera." [Online]. Available: https://github.com/ibaiGorordo/Social-Distance-Feedback-For-The-Blind

18. R. Kanjee, "Social distance detection system — using raspberry pi and opencv ai kit." [Online]. Available: https://medium.com/augmented-startups/social-distance-detection-system-using-raspberry-pi-and-opencv-ai-kit-97fd68ff8dd4

19. Z. Dai, Y. Jiang, Y. Li, B. Liu, A. B. Chan, and N. Vasconcelos, "Bev-net: Assessing social distancing compliance by joint people localization and geometric reasoning," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct 2021, pp. 5381–5391.

20. R. Mittal, C. Meneveau, and W. Wu, "A mathematical framework for estimating risk of airborne transmission of covid-19 with application to face mask use and social distancing," *Physics of Fluids*, vol. 32, p. 101903, 10 2020.

21. W. Wells, "Airborne contagion and air hygiene. an ecological study of droplet infections." *Cambridge MA*, 1955.

22. E. Riley, G. Murphy, and R. Riley, "Airborne spread of measles in a suburban elementary school," *American Journal of Epidemiology*, vol. 107, 1978.

23. G. N. S. To and C. Y. H. Chao, "Review and comparison between the wells-riley and dose-response approaches to risk assessment of infectious respiratory diseases," *Indoor Air*, vol. 20, 2 2010.

24. A. Mikszewski, L. Stabile, G. Buonanno, and L. Morawska, "The airborne contagiousness of respiratory viruses: A comparative analysis and implications for mitigation," *Geoscience Frontiers*, p. 101285, 8 2021.

25. J. Li, Z. Cheng, Y. Zhang, N. Mao, S. Guo, Q. Wang, L. Zhao, and E. Long, "Evaluation of infection risk for sars-cov-2 transmission on university campuses," *Science and Technology for the Built Environment*, vol. 27, pp. 1165–1180, 10 2021.

26. G. Buonanno, L. Morawska, and L. Stabile, "Quantitative assessment of the risk of airborne transmission of sars-cov-2 infection: Prospective and retrospective applications," *Environment International*, vol. 145, p. 106112, 12 2020.

27. W. Adams, "Measurement of breathing rate and volume in routinely performed daily activities," 1993.

28. R. Teppner, B. Langensteiner, W. Meile, G. Brenn, and S. Kerschbaumer, "Air change rates driven by the flow around and through a building storey with fully open or tilted windows: An experimental and numerical study," *Energy and Buildings*, vol. 80, pp. 570–583, 9 2014.

29. S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, and A. B. Gumel, "To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic," *Infectious Disease Modelling*, vol. 5, 2020.

30. "Computer Vision Annotation Tool," 2022. [Online]. Available: https://cvat.org/

31. GotG, "How to train an object detector using mobilenet ssd v2," 2020. [Online]. Available: https://github.com/GotG/test_object_detection_demo/tree/master/data/medmask_voc

32. "Roboflow," 2022. [Online]. Available: https://roboflow.com/

33. "Overview of OpenVINO™ Toolkit Intel's Pre-Trained Models," 2022. [Online]. Available: https://docs.openvino.ai/latest/omz_models_group_intel.html

34. Luxonis, "Luxonis mytiadx blob converter." [Online]. Available: http://blobconverter.luxonis.com

35. Luxonis, "Spatial location calculator." [Online]. Available: https://docs.luxonis.com/projects/api/en/latest/components/nodes/spatial_location_calculator/

36. F. Velarde, R. Rub´, R. Mamani-Paco, and M. Andrade-Flores, "Estimation of the probability of contagion of covid-19 by aerosols in closed environments: Applications to cases in the city of la paz, bolivia," vol. 37, pp. 22–30, 2020. [Online]. Available: http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1562-38232020000200004&lng=es&nrm=iso

37. Z. Liu, W. Zhuang, X. Hu, Z. Zhao, R. Rong, J. Li, N. Li, and W. Ding, "Potential infection risk assessment of improper bioaerosol experiment operation in one bsl-3 laboratory based on the improved wells-riley method," *Building and Environment*, vol. 201, p. 107974, 8 2021.

38. Y. Guo, H. Qian, Z. Sun, J. Cao, F. Liu, X. Luo, R. Ling, L. B. Weschler, J. Mo, and Y. Zhang, "Assessing and controlling infection risk with wells-riley model and spatial flow impact factor (sfif)," *Sustainable Cities and Society*, vol. 67, p. 102719, 4 2021.

39. Z. Wang, E. R. Galea, A. Grandison, J. Ewer, and F. Jia, "A coupled computational fluid dynamics and wells-riley model to predict covid-19 infection probability for passengers on long-distance trains," *Safety Science*, vol. 147, p. 105572, 3 2022.

40. GotG, "How to train an object detector using mobilenet ssd v2," 2020. [Online]. Available: https://github.com/GotG/test_object_detection_demo