# SARS-COV-2-19 Diagnosis Through Integration of Thermal Video and Patient Data

*Manjit Kaur, MBBS[1]; Stephanie Trovato, MD[1]; Lynn A. Fussner, MD[2]; Christina Liscynesky, MD[3]; Bing Zha, PhD[4]; Alper Yilmaz, PhD[4]; Benjamin H. Kaffenberger, MD[1]*

[1]*Department of Dermatology, The Ohio State University Wexner Medical Center, Columbus, OH, USA*

[2]*Department of Internal Medicine, Pulmonary and Critical Care, The Ohio State University Wexner Medical Center, Columbus, OH, USA*

[3]*Department of Infectious Disease, The Ohio State University Wexner Medical Center, Columbus, OH, USA*

[4]*Photogrammetric Computer Vision Lab, The Ohio State University Wexner Medical Center, Columbus, OH, USA*

**Abstract**

**INTRODUCTION** The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic rapidly spread across the world, creating a need for real-time screening of individuals to prevent further transmission. We aimed to develop a mobile thermographic imaging platform, intended to scan an entrance or hallway to identify people with abnormal facial temperatures and/or vital signs. **METHODS** Adult patients who were actively symptomatic with SARS-CoV-2 infection were included along with people with other febrile illnesses, such as pneumonia or viral infection, and healthy/asymptomatic individuals, for comparison. We applied an attention-based model to conduct a thermal image-based diagnosis of the disease. Thermographic images and videos of the study subjects were obtained through a hand-held thermal camera at distances of 5 and 10 feet. Subsequently, the videos were processed and analyzed in the Photogrammetric Computer Vision Laboratory to attempt to categorize individuals with abnormal facial temperatures and vital signs. **RESULTS** Out of a total of 76 participants, 35 patients were positive for SARS-CoV-2 with active infectious syndromes, 21 were diagnosed with other infectious diseases, and the remaining 20 were asymptomatic control subjects. The dataset consists of short videos recorded by the thermal camera to indicate the temperatures of the participants. 53 videos were obtained for SARS-CoV-2 patients and 22 for non-SARS-CoV-2 patients, where 15 frames are sampled from each video at an equal distance, and each frame is resized as 224 pixels. **CONCLUSIONS** This study generated preliminary methods and results aimed at the early identification of febrile illness and can be used as a baseline for future research.

**Keywords:** thermography, diagnosis, fever screening, Coronavirus, infectious disease

*Academic Dermatology* (2025) 3(1):1-8 | https://doi.org/10.18061/ad.v3i1.9672

Published: July 18, 2025.

Contact author: dr.manjitk@yahoo.com

## INTRODUCTION

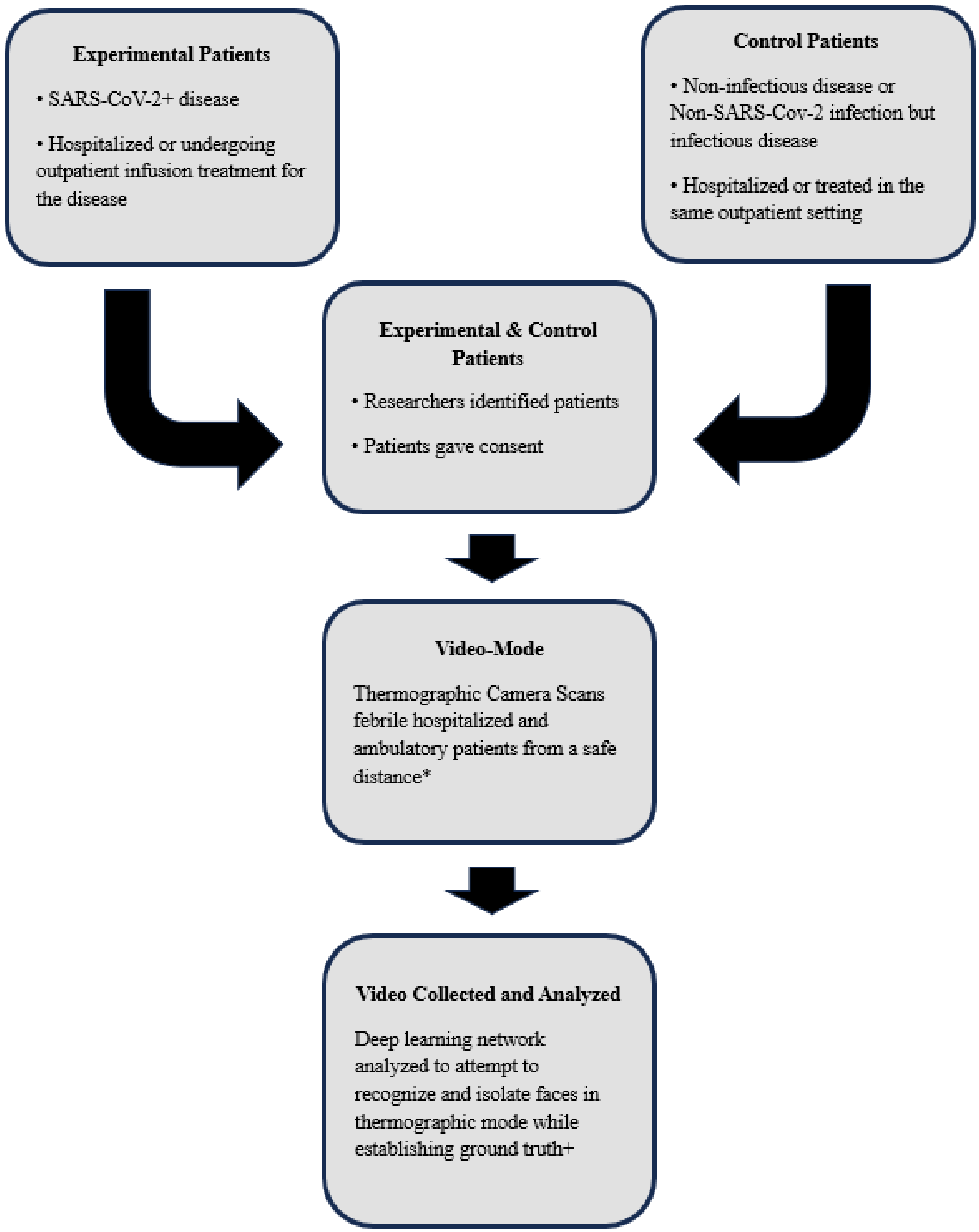
The rapid spread of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic created an urgent need for screening of individuals to limit transmission. The use of thermographic imaging could be helpful for identifying symptomatic individuals earlier and thereby decrease person-to-person transmission in public settings such as hospitals, classrooms, and airports, although evidence for its use was limited at the time of deployment.

Over the last few years, attention-based neural networks1 have revolutionized the field of natural language processing (NLP) due to their excellent capabilities of capturing long-range dependencies, as well as their ability to train scalability. Video recognition and NLP share a high-level similarity with respect to their sequential characteristics. Thus, one would expect the long-range attention models from NLP to be effective for video analytics. Other studies exist which discuss activity recognition problems from video analysis.2-4 In this study, the goal was to train a neural network to identify whether a patient has been infected with SARS-CoV-2 through the use of thermal video detection of variations in body temperature, respiratory rate, and heart rate.

## METHODS

The primary objective of this study was to develop a thermographic imaging platform that can scan a wide-angle entrance or hallway to identify individuals with abnormal facial temperatures and/or vital signs. The secondary objective was to explore the technical feasibility of identifying characteristics related to SARS-CoV-2 by noninvasive thermographic imaging and artificial intelligence. Approval was obtained from the Ohio State University, Institutional Review Board Ethics Committee (OSU IRB protocol no. 2020H0241). Eligible participants were adult patients aged ≥18 years with or without SARS-CoV-2 infections who were able to consent and who were able to expose their skin for images. Subjects from the healthy control group served as the basis for identifying patients who were not infectious. The method of recruitment of study participants for both the groups, screening and video collection, is outlined in Figure 1.

Verbal consent was obtained from all study participants to reduce the risk of exposure via the implement sharing necessary for written consent. The patients were instructed to breathe normally during the evaluation. Vital signs were collected before the thermography study. Patients were on anti-pyretics as needed; these medicines were not given as part of the study protocol. Rather, these patients were already taking anti-pyretics at their time of recruitment.




* *For consistency, images were taken from a distance and used as ground truth.*

\+ *For the screening system.*

Figure 1. *Infographic illustration of screening, recruitment of study participants, and video collection methods.*

After consent was given, the clinical investigators entered the patient room with a hand-held thermographic video and collected 2-5 minutes of video of the patient with a black body temperature set at 37°C in the background. The camera focused on the patient's exposed face, while the patient was breathing at a regular cadence for several minutes, and the image included a black body within the frame set at 37°C. Thermographic image examples from varying distances are shown in Figure 2.

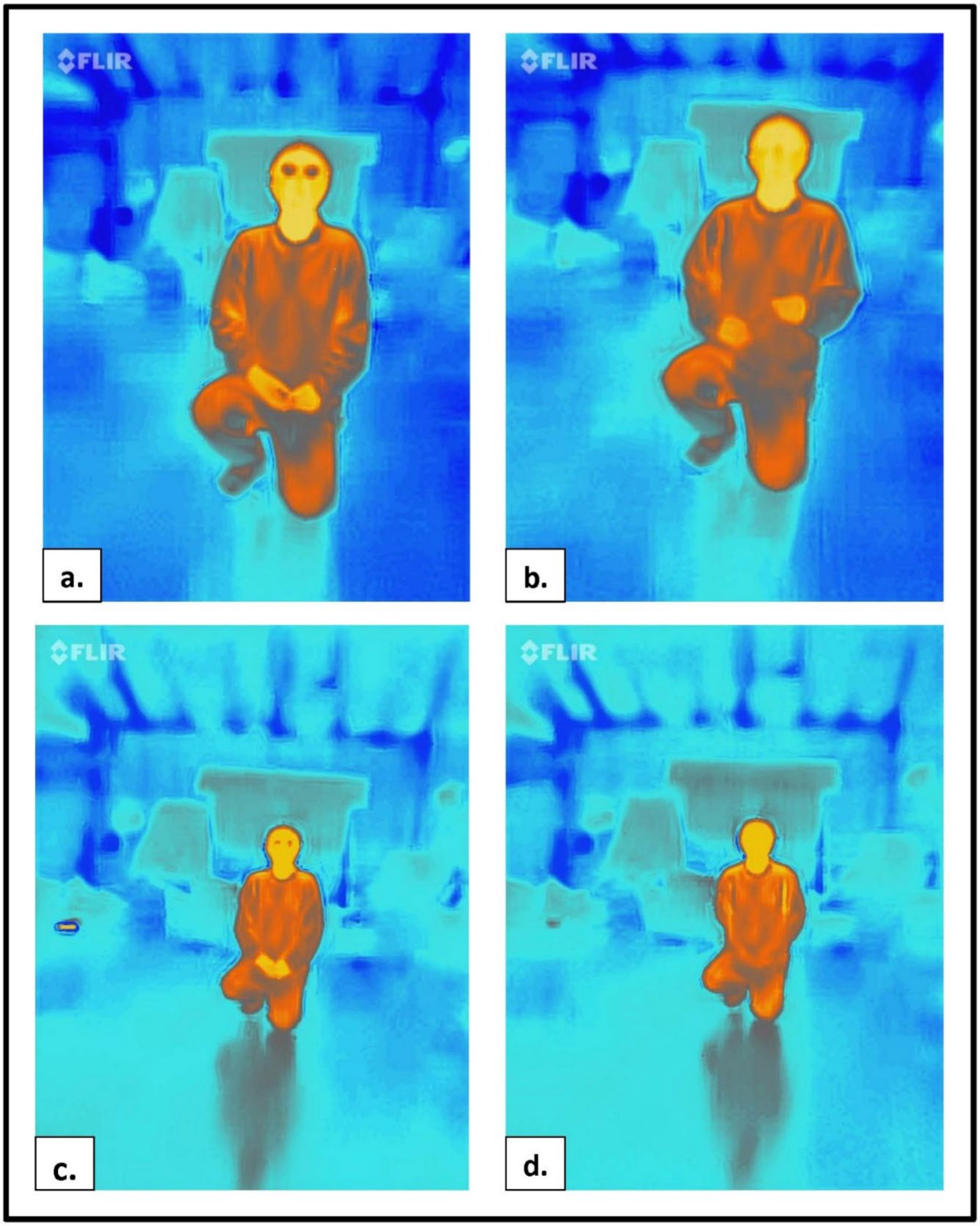



Figure 2. *Thermographic image samples in a patient assessed from varying distances with and without glasses: a. 5 ft with glasses; b. 5 ft without glasses; c.10 ft with glasses; d. 10 ft without glasses.*

In this work, we applied an attention-based model[5] to carry out a thermal image-based diagnosis of the disease. The fundamental idea of an attention-based modeling method is to learn a scoring function that assigns a different weight to each given datapoint. Specifically, the network takes as input a clip $X \in R^{T \times H \times W \times 3}$ consisting of $T$ RGB frames of size $H \times W$ sampled from the original video. A stack of the RGB representations from consecutive frames was used as an input. The output of the network is a binary value either 0 or 1, indicating non-SARS-CoV-2 or SARS-CoV-2.

We used a time-space former (TimeSformer) approach. The default hyperparameter set from the repository was used, and the complete set of hyperparameters are available on GitHub. For statistical analysis, a chi-square test was performed for categorical variables in the data, while a two-tailed t-test was used for continuous variables to compare the difference in the means of two groups.

## RESULTS

The dataset consisted of short videos recorded by the thermal camera to indicate the temperature of the patient. Each video had a varying length. The original dataset was imbalanced and contained 88 videos of SARS-CoV-2 patients and 22 of non-SARS-CoV-2 patients. The dataset contained several incomplete video recordings. Therefore, we selected 53 videos for SARS-CoV-2 patients and 22 for non-SARS-CoV-2 patients, where 15 frames were sampled from each video at an equal distance apart, and each frame was resized as 224 pixels.

Tables 1a-1e. *Demographic breakdown for all 76 study participants.*

Table 1a. *Patient age (in years, mean± standard deviation).*

| SARS-CoV-2 (35 Patients) | Control Group (41 Patients) | Statistical Tests |
|---|---|---|
| 49.97±15.05 | 46±17.65 | t-value = 1.0<br>p-value = 0.23 |

Table 1b. *Patient sex.*

| Sex | Total (76 Patients) | SARS-CoV-2 (35 Patients) | Control Group (41 Patients) | Statistical Tests |
|---|---|---|---|---|
| Males | 39 (51.32%) | 19 (54.28%) | 20 (48.78%) | X2 = 0.22<br>p-value = 0.63 |
| Females | 37 (48.68%) | 16 (45.71%) | 21 (51.21%) | |

Table 1c. *Patient locations.*

| Location | Total (76 Patients) | SARS-CoV-2 (35 Patients) | Control Group (41 Patients) | Statistical Tests |
|---|---|---|---|---|
| CRC | 38 | 12 (34.28%) | 26 (63.41%) | X2 = 5.35<br>p-value = 0.02∗ |
| Off-site | 36 | 21 (60%) | 15 (36.58%) | |

Key: *CRC- clinical research center; ∗ statistically significant*

Table 1d. *Patients vaccinated against SARS-CoV-2.*

| Total (76 Patients) | SARS-CoV-2 (35 Patients) | Control Group (41 Patients) | Statistical Tests |
|---|---|---|---|
| 25 (32.89%) | 13 (37.14%) | 12 (29.26%) | $X^2 = 0.53$<br>p-value = 0.46 |

Table 1e. *Vaccinated patients by number of doses.*

| One Dose | Two Doses |
|---|---|
| 18 (72% of 25 vaccinated patients) | 7 (28% of 25 vaccinated patients) |

Out of the 76 total participants enrolled in this study, 39 were male and 37 were female. Thirty-five (46%) of participants were in the study group (M: F; 19:16), and 41 were in the control group (M:F 20:21). 38 patients were seen in the clinical research center and 36 were seen at an off-site location; this difference was statistically significant ($\chi 2$ 5.35; p= 0.02). The average age of patients in the study group was 49.97±15.05 years, and the average age in the control group was 46±17.65 (t=1.0; p =0.23). 35 out of 76 study participants were diagnosed with SARS-CoV-2 infection. In the study group, 13 out of 35 (37.14%) patients were vaccinated against COVID, as were 12 of 41 (29.26%) in the control group ($\chi 2$ 0.53; p 0.46).

Table 2. *Vital signs of study participants.*

| Vital Sign | SARS-CoV-2 (35 Patients) | Control Group (41 Patients) | Statistical Tests |
|---|---|---|---|
| Temperature ºF (mean ± SD) | 98±0.63 | 97.9±0.38 | t-value = 0.21<br>p-value = 0.82 |
| Respiratory Rate/min (mean ±SD) | 17.17±2.39 | 15.9±1.43 | t-value = 2.74<br>p-value = 0.07 |
| Heart Rate/min (mean ±SD) | 78.2±15.92 | 75±10.98 | t-value = 1.0<br>p-value = 0.31 |
| Oxygen Saturation % (mean ± SD) | 96.5±2.18 | 97±2.35 | t-value = 1.18<br>p-value = 0.23 |

Key: *SD-standard deviation.*

Table 3. *Severity of SARS-CoV-2 disease in study participants.*

| | |
|---|---|
| **Mild** | 25 (71.42%) |
| **Moderate** | 5 (14.28%) |
| **Severe** | 5 (14.28%) |
| **Oxygen requirement (on room air)** | 11 (31.42% |

Out of the 41 patients in the control group, 21 were diagnosed with infectious diseases other than SARS-CoV-2, such as pneumonia or viral diseases, while the remaining 20 were asymptomatic controls. Among those infected with SARS-CoV-2, 25 out of 35 (71.42%) had a mild disease, while 5 patients each had moderate and severe disease. The mean temperature of the patients with SARS-CoV-2 infections was 98±0.63 ºF versus 97.9±0.38 ºF in the control group (t 0.21; p= 0.82). The average heart rate was 78.2±15.92 beats per minute in the SARS-CoV-2 group and 75±10.9 in the control group (t 1.0; p= 0.31). The average respiratory rate was 17.17±2.39 breaths per minute in the SARS-CoV-2 group and 15.9±1.43 in the control group (t 2.74; p= 0.07). The average oxygen saturation was 96.5±2.18 % in the SARS-CoV-2 group and 97±235 % in the control group (t 1.18; p= 0.23). The differences between the means in the SARS-CoV-2 and control groups were not statistically significant. 11 of the 35 people (31.42%) with SARS-CoV-2 infection had an oxygen requirement.

We trained the attention-based network as a binary classification problem and observed that the loss function was not converging to a solution. This fact indicates that the neural network is not learning well from the given videos to identify SARS-CoV-2 infection. The developed code is available at: https://github.com/OSUPCVLab/InfraredCovid.

## DISCUSSION

Despite our intentions and prospective collection, the neural network was not adapting from the given videos to identify SARS-CoV-2. Given that the attention-based neural network has been demonstrated as a powerful model in video classification problems,[6,7] we strongly believe that the attention-based neural network can be applied to solve this thermal image sequence classification if a balanced, larger, and better-quality training dataset is possible. However, this approach was found to be a consistently ineffective screening system in multiple different settings during the SARS-CoV-2 pandemic.[4,8,9] Studies have reported varied efficacy of thermography cameras for fever screening in other infectious diseases.[3,4]

We determined that the training result was not as consistent as expected. Our analysis indicated that the results were due to both the quality and the quantity of the dataset. Specifically, three important factors resulted in a reduction in performance:

- Camera shake during data acquisition.
- Camera internal compensation for thermal values for visualization.
- No difference in recorded vital signs between control and infected patients.

From the recorded videos, we observed that the camera was shaking and moving, which caused the area (and thereby, the temperature) around the patient to change. Furthermore, the scale of the captured face-image changed with camera movement.

Furthermore, a more ideal video dataset would show higher temperatures for SARS-CoV-2 patients versus the normal temperature for non-SARS-CoV-2 patients to better differentiate them. However, as noted in Table 2, even our most severe SARS-CoV-2 patients were often not febrile when they were being treated and enrolled. This most likely was a consequence of enrolling ambulatory patients and not excluding those who were on anti-pyretics.

## CONCLUSIONS

Despite the initial excitement around this method as a screening process to detect infectious diseases, even in a controlled hospital room with the use of a ground truth temperature device, we were unable to adapt a neural network to differentiate SARS-CoV-2 from control patients. The method introduced and the results obtained are preliminary; however, this study can be used as a baseline for future research. This dataset can be enlarged with potentially-healthy people. An initial training on only healthy people using an auto encoder to learn the feature maps of normal heat distribution, followed by a combined healthy/sick training set, can be used to perform a secondary training (transfer learning) to improve the accuracy for future research.

**Acknowledgements:** We are grateful to Dr. Ronald Xu, PhD, Associate Professor, Department of Biomedical Engineering, The Ohio State University, for his contributions to the study design and manuscript editing.

## REFERENCES

1. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS; 2017:6000-6010. https://doi.org/10.48550/arXiv.1706.03762
2. Qu Y, Meng Y, Fan H, et al. Low-cost thermal imaging with machine learning for non-invasive diagnosis and therapeutic monitoring of pneumonia. *Infrared Physics & Technology.* 2022;123:104201. https://doi.org/10.1016/j.infrared.2022.104201
3. Mercer JB, Ring EFJ. Fever screening and infrared thermal imaging: Concerns and guidelines. *Thermology International.* 2009;19(3):67-69.
4. Khaksari K, Nguyen T, Hill B, et al. Review of the efficacy of infrared thermography for screening infectious diseases with applications to COVID-19. *Journal of Medical Imaging.* 2021;8(S1):010901-010901. https://doi.org/10.1117/1.JMI.8.S1.010901
5. Bertasius G, Wang H, Torresani L. Is space-time attention all you need for video understanding?. Paper presented at: 38th International Conference on Machine Learning; February 9, 2021. https://doi.org/10.48550/arXiv.2102.05095
6. Girdhar R, Carreira J, Doersch C, et al. Video action transformer network. In: *Proceedings of the Institute of Electrical and Electronics Engineers/Computer Vision Foundation Conference on Computer Vision and Pattern Recognition(CVPR)*. IEEE/CVF; 2019:244-253. https://doi.org/10.1109/CVPR.2019.00033
7. Liu Z, Ning J, Cao Y, et al. Video swin transformer. In: *Proceedings of the Institute of Electrical and Electronics Engineers/Computer Vision Foundation Conference on Computer Vision and Pattern Recognition*. IEEE/CVF; 2022:3202-3211. https://doi.org/10.1109/CVPR52688.2022.00320
8. Fiscal MRC, Treviño V, Treviño LJR, et al. COVID-19 classification using thermal images. *Journal of Biomedical Optics.* 2022;27(5):056003-056003. https://doi.org/10.1117/1.JBO.27.5.056003
9. Quilty BJ, Clifford S, Flasche S, et al. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Eurosurveillance.* 2020;25(5):2000080. https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000080