doi: 10.5821/conference-9788419184849.17

CONSTRUCTION OF FACIAL SKIN TEMPERATURE-BASED ANOMALY DETECTION MODEL FOR DAILY FLUCTUATIONS IN HEALTH CONDITIONS

Masahito TAKANO[1a], Kosuke OIWA[a], Akio NOZAWA[a]

[a] *Aoyama Gakuin University, Japan, akio@ee.aoyama.ac.jp*

ABSTRACT

A method for estimating health conditions is required to monitor daily health conditions. Various types of data have been used in healthcare studies; however, imaging data are superior because they allow quick and remote measurements. Thermal face images can be measured safely and economically using infrared thermography. Many physiological and psychological states have been evaluated based on the data from these images. A previous study, using short-term experiments, confirmed that an anomaly detection model constructed using a variational autoencoder enables the detection of anomalous states of thermal face images. A long-term experiment is essential to estimate long-term fluctuating human states, such as health conditions. The purpose of this study is to construct a facial skin temperature-based anomaly detection model for human health conditions. The authors obtained thermal face images with health condition questionnaires for approximately a year. Based on the questionnaire responses, the thermal images in good and poor health conditions were labeled “normal state” and “anomaly state,” respectively. The facial skin temperature-based anomaly detection model for health conditions was constructed using a variational autoencoder with only thermal face images in the normal state. The AUC, which represents anomaly detection performance, was 0.70. In addition, an increasing trend of the performance of the model by learning a wider area of skin temperature was confirmed.

[1] Corresponding author. Post: 5-10-1, Fuchinobe, Chuo-ku, Sagamihara-shi, Kanagawa, 252-5258, Japan. E-mail: takanomasagito.v@gmail.com

Keywords: Thermal face image, Human health monitoring, Healthcare monitoring, Poor health conditions, Deep learning

1 INTRODUCTION

Everyone knows the importance of understanding their own health conditions and dealing with them appropriately. An increasing number of companies are focusing on providing healthcare for their employees. To realize daily health monitoring, a method for estimating the health conditions is required because it is difficult for people to be fully aware of their psychological and physical issues. In healthcare studies, the application of machine learning methods has been increasing, and deep learning methods, in particular, have demonstrated high performance and potential for multiple tasks (Miotto et al., 2018). Deep learning methods can effectively obtain information from complex data, such as biometric data. Various types of data used in these studies include health records (Solares et al., 2020), genomic data (Fakoor et al., 2013), electroencephalogram data (Acharya et al., 2018), and magnetic resonance imaging data (Liu et al., 2014).

Imaging data are superior because they allow quick and remote measurements, and thermal face images are often used in studies to estimate physiological and psychological states. A thermal face image is the skin temperature image of a face and can be measured quickly, remotely, safely, and economically using infrared thermography (Ring, 2014). Skin temperature varies with the amount of skin surface blood and is controlled by the autonomic nervous system. Therefore, it is a reliable tool that estimates physiological and psychological states (Ioannou, Gallese, & Merla, 2014). Based on this information, evaluations of various human states such as stress (Engert et al., 2014), emotions (Ebisch et al., 2012), and drowsiness (Bando, Oiwa, & Nozawa, 2017) have been performed. Studies applying deep learning to thermal face images have estimated some human states such as drunkenness (Koukiou & Anastassopoulos, 2015), exercise-induced fatigue (Lopez, del-Blanco, & Garcia et al., 2017), and drowsiness (Adachi, Oiwa, & Nozawa, 2019). In these studies, estimations were conducted using classification models.

Estimating human health conditions using classification models requires training the model using thermal face images under poor as well as normal health conditions. Generally, obtaining thermal images under poor health conditions is much more challenging than obtaining them under normal health conditions. Anomaly detection methods can solve this problem. A variational autoencoder (VAE) (Kingma & Welling, 2013), known as a deep generative model, enables the construction of anomaly detection models using only the normal data (An & Cho, 2015). Thus, an anomaly detection model for human health conditions is constructed, without the use of thermal images under poor health conditions, using the VAE.

Masaki et al. confirmed that an anomaly detection model constructed using the VAE enabled the separation of two states of thermal face images (Masaki et al., 2021). Two states of the thermal face images were obtained in a short-term experiment; one state was obtained when the subject was just set in a chair, and the other state was obtained when the subject raised their blood pressure by breath-holding. Takano et al. performed anomaly detection considering diurnal variations in thermal face images (Takano et al., 2022). Thermal face images in normal and anomalous states were obtained every hour in a 16-hour experiment using the method described by Masaki et al. In these studies, thermal face images were obtained during the experiment in a day, and high blood pressure caused by breath-holding was defined as an anomalous state. However, obtaining thermal face images through a long-term experiment is essential to estimate long-term fluctuating human states, such as health conditions. Additionally, defining poor health conditions as anomalous states is needed to evaluate an anomaly detection model for health conditions. The purpose of this study is to construct a facial skin temperature-based anomaly detection model for human health conditions. To achieve this, thermal face images were obtained for approximately one year, and an anomaly detection model was constructed. Finally, the performance of the constructed model was evaluated using thermal images under good and poor health conditions.

2 METHODS

2.1 Data collection

2.1.1 Experiment systems

Thermal face images with health conditions questionnaires were obtained for approximately a year (September 2020 to November 2021). The subjects constructed of 33 healthy adult males aged 21–51 years. The measurement environment is illustrated in Figure 1. An infrared thermography camera (Boson, FLIR Systems, Wilsonville, OR, USA) was placed at 1.0 m in front of the subject. A tablet (iPad, Apple, Cupertino, CA, USA) was placed near the camera to complete the questionnaire. The measurement was conducted at a room temperature range of 25.3–30.9 °C at the subject's discretion. The temperature resolution of the camera was 0.1 °C. The size of the thermal image was 256 × 320 pixels. An example of a measured thermal images is shown in Figure 2.

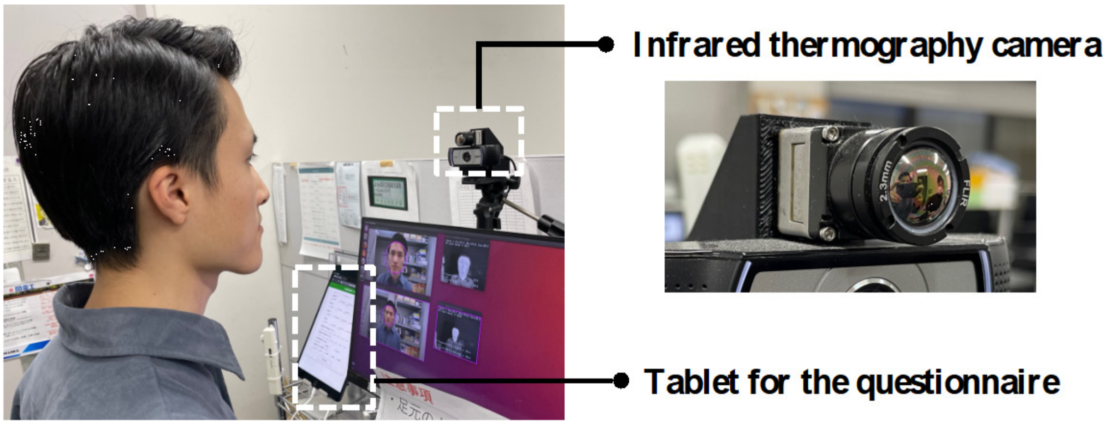

Figure 1. Measurement environment

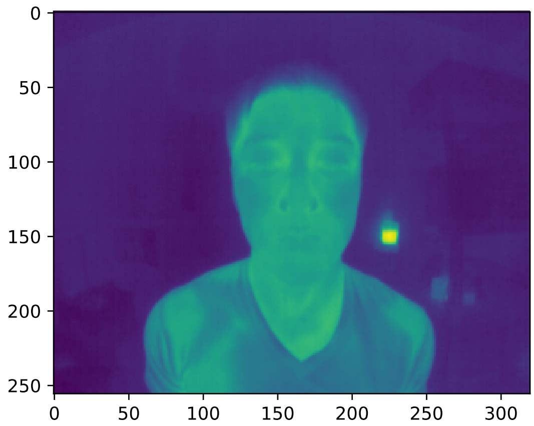

Figure 2. Example of measured thermal image

The subjects were asked to complete the questionnaire immediately after measuring their thermal images. The screen for answering the questionnaire is shown in Figure 3. The questionnaire consisted of four questions regarding health conditions, and the participants choose only one answer for each question.

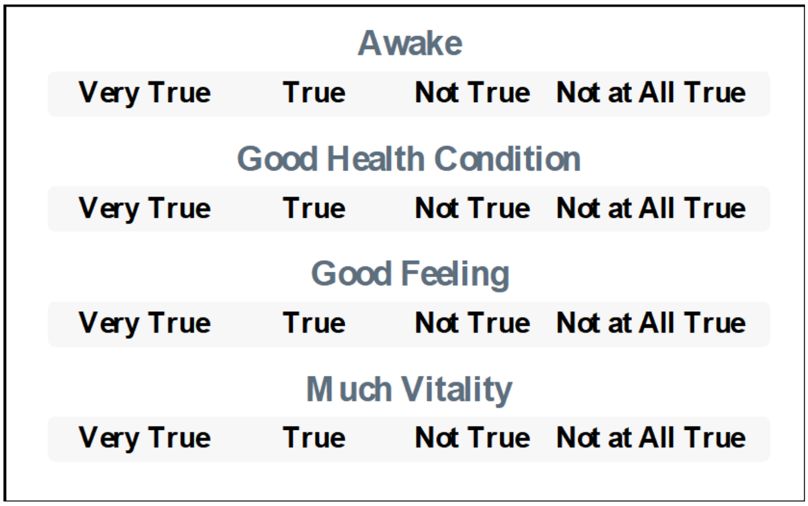

Figure 3. Screen for answering the questionnaire

2.1.2 Definitions of Normal and Anomaly

The obtained thermal face images were labeled as “normal state” and “anomaly state” based on the questionnaire answers. If all answers were “Very True,” the thermal image was labeled as “normal state,” indicating good health conditions. If at least one answer was “Not at All True,” the thermal image was labeled as “anomaly state,” indicating poor health conditions. Among obtained 612 thermal images, 175 were labeled as “normal state,” and 66 were labeled as “anomaly state.” 371 thermal images not labeled as “normal state” or “anomaly state” were not used in this study.

2.2 Model construction

2.2.1 Overview of VAE framework

VAE is an autoencoder-type network, as shown in Figure 4. The input and output data have the same dimensions, and the latent variables have fewer dimensions than the input data. The VAE comprises an encoder, which transforms the input data into the latent variables, and a decoder, which reconstructs the input data from the latent variables. The difference between the input and the reconstructed input is called the reconstruction error. Training is performed to minimize the reconstruction error. Through this process, the VAE learns the probability distribution of latent variables to reconstruct the input data. An anomaly detection model is constructed under the framework of semi-supervised learning using only the normal data for training the VAE. The reconstruction error calculated by inputting data into the trained VAE indicates the anomaly degree of the data.

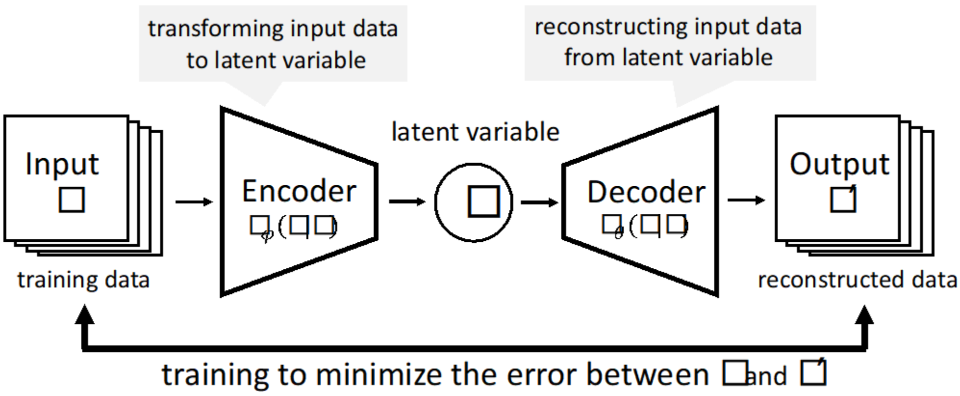

Figure 4. Overview of VAE

2.2.2 Model construction using VAE

The anomaly detection model was constructed using 109 thermal face images in the normal state. Model performance evaluation was conducted using 66 thermal face images in both, the anomaly state and as well as the normal state; these images were not used for model construction. The thermal face image was cut out from the thermal image based on 68 facial landmarks. Facial landmarks were obtained using a method proposed by Nagumo et al. (Nagumo et al., 2021). The examples of thermal image with 68 facial landmarks and cut-out image are shown in Figure 5; the red plots are facial landmarks, and the black square frame represents the cut-out area

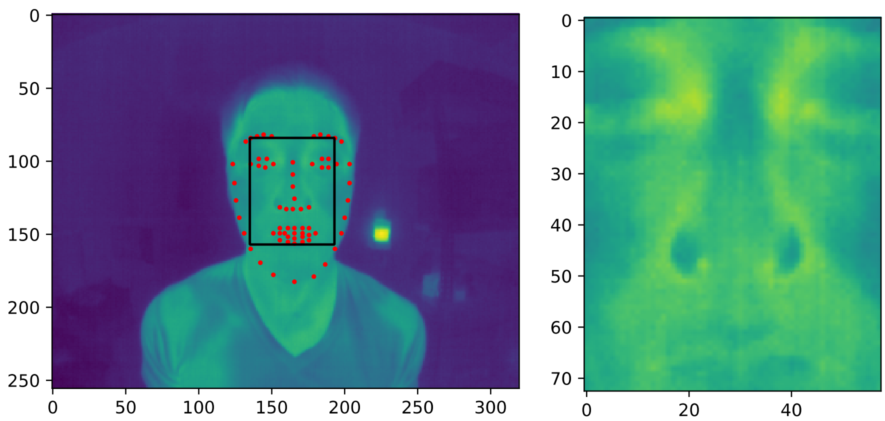

Figure 5. Examples of thermal image with facial landmarks and cut-out thermal face image

Each thermal face image was resized to 66 × 66 pixels. The mean and standard deviation of the pixel values were set to 0 and 1, respectively, using z-score normalization. The VAE was trained using 100,000 patches that were randomly cut from 109 resized thermal face images in the normal state. The patch size was set to $S_p \times S_p$ pixels, and S_p was selected from $S_p \in \{8, 16, 32, 64\}$. The dimensions of the latent variables N_z were selected from $N_z \in \{2, 8, 16, 32, 64\}$. Nine models were constructed for each condition, considering the variations in model performance.

2.3 Anomaly detection

Anomaly detection was performed by plotting the thermal face images based on the mean unregularized score (Matsubara et al, 2020). The unregularized score M_{VAE} was calculated by applying a thermal face image for evaluation to a trained VAE. The equation for the unregularized score is as follows.

$$M_{VAE} = \sum_{i=1}^{N_x} \frac{1}{2} \frac{\left(\boldsymbol{\mu}_{x_i} - \boldsymbol{x}_i\right)^2}{\boldsymbol{\sigma}_{x_i}{}^2} \Bigg|_{z=\mu_z} \tag{1}$$

The above equation is directly related to the reconstruction error. N_x and $\boldsymbol{x}_i$ represent the number of pixels in the image and pixel value, respectively. $\boldsymbol{\mu}_x$ and $\boldsymbol{\sigma}_x$ are a pair of mean and variance vectors for the conditional probability of the latent variables $\boldsymbol{z}$. The mean unregularized score was calculated by dividing the unregularized score by the number of pixels.

3 RESULTS AND DISCUSSION

An example of thermal face images plotted for model evaluation is shown in Figure 6.

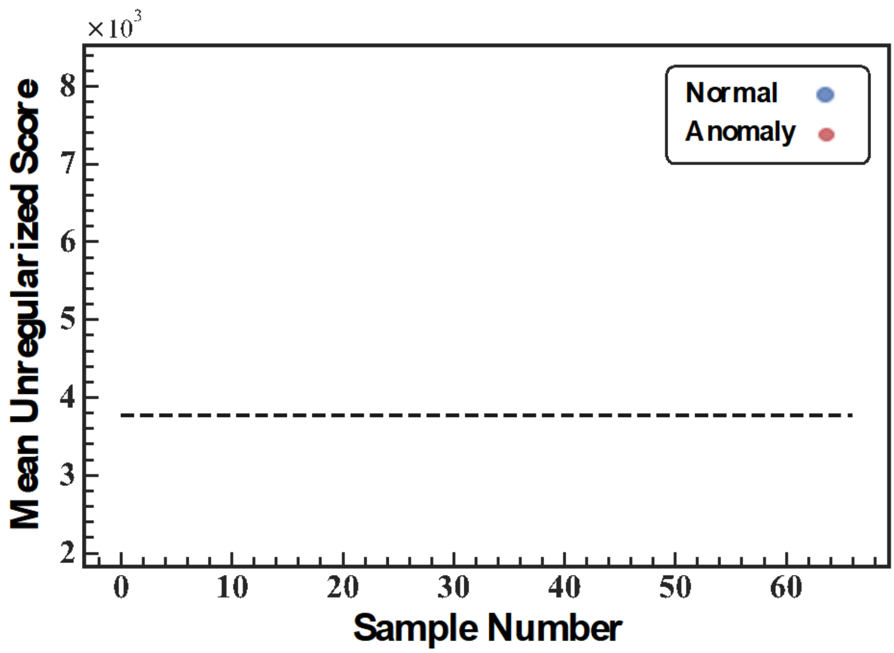

Figure 6. Example of thermal face images plotted for evaluation

Plotting was performed by the VAE trained under the conditions of $S_p = 64$ and $N_z = 64$. The red plots are thermal face images in the anomaly state, and the blue plots are thermal face images in the normal state; these images were not used for model construction. The black dotted line represents the threshold for anomaly detection. A mean unregularized score at which the sum of the true positive rate and the true negative rate is maximum was set as the threshold. The true positive rate is the proportion of correct predictions in the anomaly class, and the true negative rate is the proportion of correct predictions in the normal class. The equations for the true positive and true negative rates are as follows.

$$True\ Positive\ Rate = \frac{(True\ Positive)}{(True\ Positive\ + False\ Nagative)} \tag{2}$$

$$True\ Negative\ Rate = \frac{(True\ Negative)}{(True\ Negative\ + False\ Positive)} \tag{3}$$

In this example, the true positive and true negative rates were 0.67 and 0.68, respectively. This result indicates that facial skin temperature-based anomaly detection models enable the detection of poor health conditions. The receiver operating characteristic (ROC) curve of the model under the conditions of 64 for S_p and 64 for N_z is shown in Figure 7.

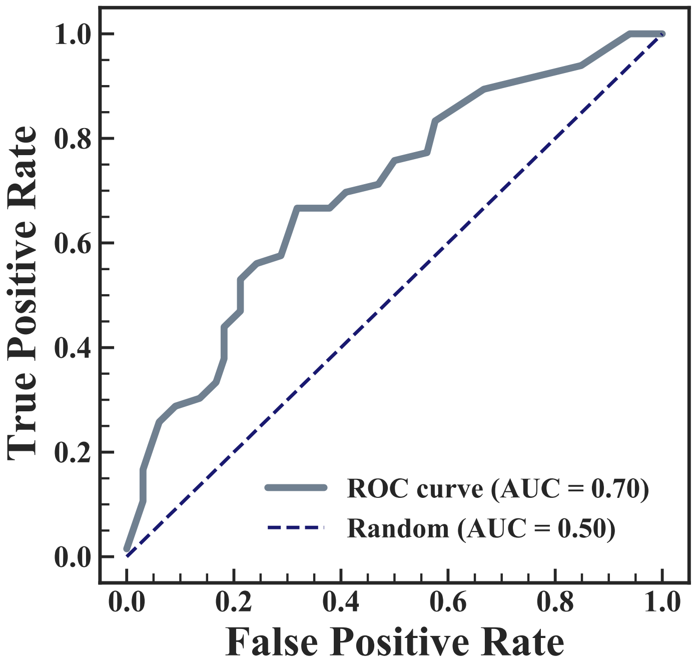

Figure 7. Example of a ROC curve

The ROC curve is a two-dimensional graph that plots the true positive rate on the y-axis and false positive rate on the x-axis. The false positive rate is the proportion of incorrect predictions in the normal class. The equations for the false positive rate are as follows.

$$Folse\ Positive\ Rate = \frac{(False\ Negative)}{(True\ Positive\ + False\ Nagative)} \tag{2}$$

The area under the ROC curve (AUC) is a single scalar value that represents the expected classification performance of a model (Fawcett, 2006). In this example, the AUC of the model was 0.70. The maximum AUC for each condition is shown in Figure 8.

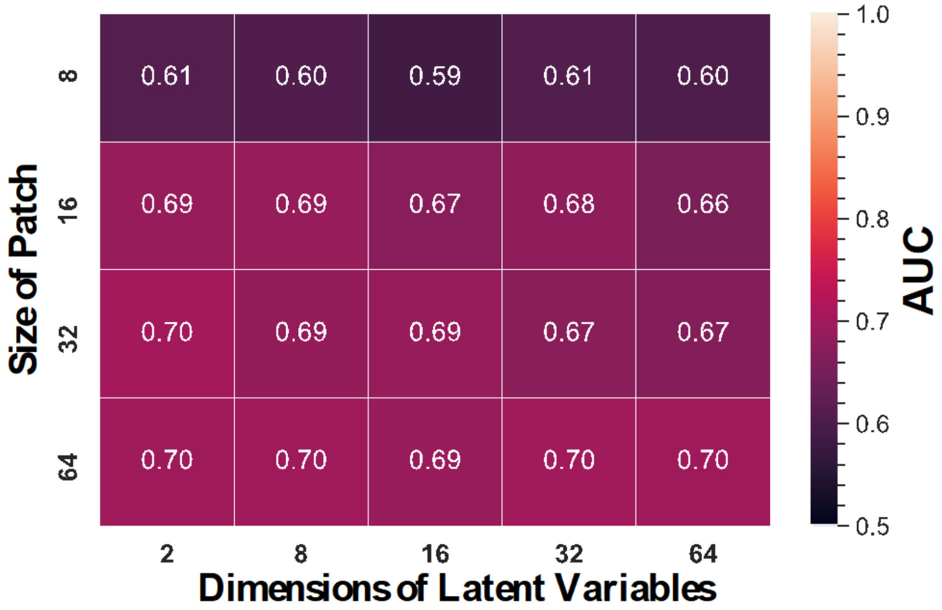

Figure 8. Maximum AUC for each model construction conditions

The overall maximum AUC was 0.70, and the AUC was higher for conditions with a larger patch size. This result indicates that learning a wider area of skin temperature improves the classification performance of the model. Oiwa et al. concluded that long-term variability of facial hue information around the periorbital region could be related to health conditions (Oiwa et al., 2021). This suggests a relationship between a wide range of facial hue information and health conditions. The authors believe that facial skin temperature is also related to health conditions because facial skin temperature also varies with the amount of skin surface blood, similar to facial hue information.

4 CONCLUSION

The purpose of this study was to construct a facial skin temperature-based anomaly detection model for human health conditions. The authors obtained thermal face images for approximately one year and constructed an anomaly detection model. Finally, the performance of the constructed model was evaluated using thermal face images under good and poor health conditions. Consequently, thermal face images in poor health conditions were detected using the proposed detection model. The AUC of the model with the highest accuracy was 0.70. The AUC was improved by learning a wider area of skin temperature. However, continuous data collection is required in the future to improve accuracy. Moreover, investigations on the influence of ambient temperature and construction of models using other information, such as facial hues, are needed to gain better insights.

REFERENCES

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adeli, H., & Subha, D. P. (2018). Automated EEG-based screening of depression using deep convolutional neural network. Computer methods and programs in biomedicine, 161, 103-113.

Adachi, H., Oiwa, K., & Nozawa, A. (2019). Drowsiness level modeling based on facial skin temperature distribution using a convolutional neural network. IEEJ Transactions on Electrical and Electronic Engineering, 14(6), 870-876.

An, J., & Cho, S. (2015). Variational autoencoder based anomaly detection using reconstruction probability. Special Lecture on IE, 2(1), 1-18.

Bando, S., Oiwa, K., & Nozawa, A. (2017). Evaluation of dynamics of forehead skin temperature under induced drowsiness. IEEJ Transactions on Electrical and Electronic Engineering, 12, S104-S109.

Ebisch, S. J., Aureli, T., Bafunno, D., Cardone, D., Romani, G. L., & Merla, A. (2012). Mother and child in synchrony: thermal facial imprints of autonomic contagion. Biological psychology, 89(1), 123-129.

Engert, V., Merla, A., Grant, J. A., Cardone, D., Tusche, A., & Singer, T. (2014). Exploring the use of thermal infrared imaging in human stress research. PloS one, 9(3), e90782.

Fakoor, R., Ladhak, F., Nazi, A., & Huber, M. (2013, June). Using deep learning to enhance cancer diagnosis and classification. In Proceedings of the international conference on machine learning (Vol. 28, pp. 3937-3949). ACM, New York, USA.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), 861-874.

Ioannou, S., Gallese, V., & Merla, A. (2014). Thermal infrared imaging in psychophysiology: potentialities and limits. Psychophysiology, 51(10), 951-963.

Iwashita, Y., Nagumo, K., Oiwa, K., & Nozawa, A. (2021). Estimation of resting blood pressure using facial thermal images by separating acute stress variations. Artificial Life and Robotics, 26(4), 473-480.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

Koukiou, G., & Anastassopoulos, V. (2015). Neural networks for identifying drunk persons using thermal infrared imagery. Forensic Science International, 252, 69-76.

Liu, S., Liu, S., Cai, W., Pujol, S., Kikinis, R., & Feng, D. (2014, April). Early diagnosis of Alzheimer's disease with deep learning. In 2014 IEEE 11th international symposium on biomedical imaging (ISBI) (pp. 1015-1018). IEEE.

Lopez, M. B., del-Blanco, C. R., & Garcia, N. (2017, November). Detecting exercise-induced fatigue using thermal imaging and deep learning. In 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.

Masaki, A., Nagumo, K., Lamsal, B., Oiwa, K., & Nozawa, A. (2021). Anomaly detection in facial skin temperature using variational autoencoder. Artificial Life and Robotics, 26(1), 122-128.

Matsubara, T., Sato, K., Hama, K., Tachibana, R., & Uehara, K. (2020). Deep generative model using unregularized score for anomaly detection with heterogeneous complexity. IEEE Transactions on Cybernetics.

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), 1236-1246.

Nagumo, K., Kobayashi, T., Oiwa, K., & Nozawa, A. (2021). Face alignment in thermal infrared images using cascaded shape regression. International Journal of Environmental Research and Public Health, 18(4), 1776.

Oiwa, K., Urakami, K., Lamsal, B., Nagumo, K., & Nozawa, A. (2021). Relationship Between Long-Term Variability of Facial Hue Information in Physiological and Psychological ROIs and Health Condition. IEEE Access, 9, 145554-145562.

Ring, F. J. (2014). Pioneering progress in infrared imaging in medicine. Quantitative InfraRed Thermography Journal, 11(1), 57-65.

Solares, J. R. A., Raimondi, F. E. D., Zhu, Y., Rahimian, F., Canoy, D., Tran, J., ... & Salimi-Khorshidi, G. (2020). Deep learning for electronic health records: A comparative review of multiple deep neural architectures. Journal of biomedical informatics, 101, 103337.

Takano, M., Iwashita, Y., Nagumo, K., Kobayashi, T., Oiwa, K., & Nozawa, A. (2022). An optimization of anomaly detection algorithms considering diurnal variation in facial skin temperature using VAE: Proceedings of the 27th International Symposium on Artificial Life and Robotics (AROB 27th 2022), pp.311-315 (in Japan)