RECOGNITION OF HUMAN FACES AND TEMPERATURES BASED ON AUTOMATICALLY INTELLIGENT ALGORITHM

Linh T. Nguyen, Duc Long Vu, Minh Tuan Nguyen and Hai-Chau Le
Posts and Telecommunications Institute of Technology, Hanoi, Vietnam

Abstract: The surveillance system plays an essential role in the security field to recognize early illegal access. As a part of such systems, human face recognition provides an effective service to observe and identify people crossing a specific area. Moreover, measurement of body temperature is also mandatory due to Coronavirus pandemic. Therefore, temperature prediction through surveillance systems is the best selection because of its effectiveness, high authentication, and privacy maintenance. In this paper, a novel algorithm is proposed for the recognition of human faces and temperature using deep learning methods and thermal face images. The search-based method is developed for selection of an optimal structure and parameters of the convolutional neural network using the thermal images as the input. In addition, the human temperature is computed directly on the thermal images to assist warning systems in terms of abnormal body temperatures. The proposed application produces a relatively high accuracy of over 98%, which is fitted significantly in practical environments.

Keywords: Thermal image, Deep learning, Human temperature measurement, Face recognition.

I. INTRODUCTION

Nowadays, the growth of human factors in the information age gives an exponential increase to the need for biometric systems for person identification, verification or surveillance systems [1-5]. Among various biometrics, face has been one of the critical traits, which has a variety of applications ranging from health, education, finance and security [1, 2]. Human face recognition and temperature detection have been the most attractive research topic and paid intensive attention from technical experts for many years. Indeed, they are the important elements to construct an effective and reliable surveillance system for early warnings and protection from unauthorized access [6]. In addition, with the outbreak of COVID-19 pandemic and the unbalanced distribution of vaccines around the world, the feasible solution so far is identifying and checking the symptoms at the early stage. Besides, the majority of confirmed cases only has some mild symptoms like fever, cough or fatigue [7]. Therefore, there is a need to measure the body temperature of people, especially the facial temperature, for containing the widespread of the virus.

Various development methods have been implemented for selection of the optimal application related to face recognition and temperature prediction. In general, methods are categorized into two classes in which machine learning and deep learning techniques are employed as the key algorithms of the face recognition. For the machine learning method, the authors of [8] propose a face recognition based on particle swarm optimization in combination with support vector machine (SVM). Furthermore, the principal component analysis PCA) is used for extraction of the feature set, which is then fed into the support vector machine. To increase the final performance of the face detection, a voting algorithm is suggested in [9]. Here, assemble learning methods such as k-nearest neighbours, SVM, random forest, C4.5, random tree, classification and regression tree, naïve bayes, are adopted to detect human faces. An extreme learning machine using a feature combination extracted by the PCA, linear discriminant analysis, and discriminative locality alignment algorithms is proposed in [10-13]. Obviously, the contribution of the feature transformation in aforementioned research is significant in comparison with other studies in terms of performance improvement.

Deep learning techniques have emerged as the alternative method, recently, due to its capability of no human knowledge-based feature extraction, feature selection requirements. Moreover, deep learning techniques offer an attractive characteristic, known as the automatically learning feature, which is absent in machine learning algorithms [14-15]. In addition, existing scientific reports show that the face recognition performance of deep learning algorithms is relatively high compared with that of machine learning methods. Indeed, a simple but effective algorithm is given in [16-18], where the convolutional neural networks (CNN) using the thermal face images is proposed. In [19], the FusionNet including encoder, bridge, decoder and output layer, is shown as the core algorithm for human face recognition using depth and thermal images. The generative adversarial network is applied for

Contact author: Hai – Chau Le
Email: Chaulh@ptit.edu.vn
Manuscript received: 15/11/2021, revised: 12/02/2022, accepted: 19/02/2022.

face recognition in the visible light domain [6]. The rationale behind the use of the visible light domain is that most of the texture and edge information are lost in the thermal images. Therefore, image processing techniques are required to convert thermal images into visible light domain ones for better learning features archived by the generative adversarial network.

In this paper, we propose a simple and effective human face recognition algorithm using CNN and thermal images. Moreover, instead of measuring the entire human body temperature, we consider scanning the facial temperature to restrict impacts of external objects on human body such as clothes and belongings to provide early warnings for security systems with respect to illegal access and people with fever. We first convert the original thermal images into grayscale thermal ones to speed up the process of the CNN. Then different CNN parameters and structures are investigated with a training dataset to figure out an optimal CNN structure and corresponding parameters. On other words, the search-based method is employed for selection of the optimal CNN model. The human temperature is computed directly using the pixel temperatures on the original thermal images.

Our main contribution of this work is to apply the optimization method for identification of the optimal CNN structure and parameters. The remainder of the paper is structured as follows. Section II is to introduce the dataset we use in our research. Section III explains the research methodology and our research approach and Section IV is devoted to describe simulation results and provide discussions. Finally, Section V concludes the paper.

II. DATA

The dataset used for this work is the Terravic Facial Infrared Database, which is available on OTCBVS Benchmark Dataset Collection [20]. This is a collection of paired thermal and visible ear images that were acquired in different illumination conditions ranging between 2 and 10700 lux. The dataset contains variations in facial poses such as front, left, right and environments like indoor and outdoor. Accessories such as glasses and hats also are included in this dataset. There are 18 individuals with a total of 22,784 thermal images. However, the number of thermal images is different between individuals.

Therefore, in this research, we select only 227 thermal images from each person, which are then divided into 107 thermal images of the training set, 50 thermal images of test set, and 70 thermal images of validation set. As a result, only 4096 thermal images of 18 individuals are considered for construction of the CNN model.

III. METHOD

We use the CNN as the core algorithm for human face recognition using thermal images. Figure 1 shows the method development procedures to address the optimal CNN model.

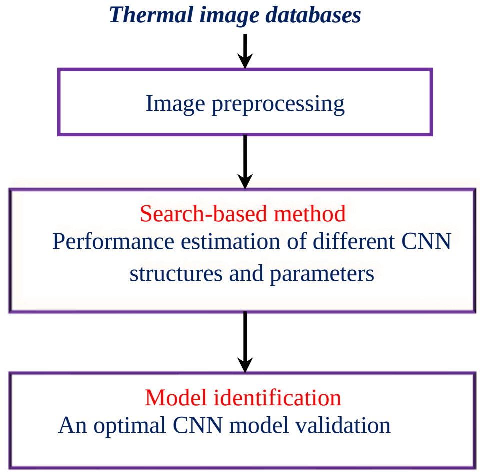


Figure 1. Method diagram.

A. CNN

The CNN includes different layers, which are input, convolutional, rectified linear unit, max-pooling layer, fully connected layers, and output layer [21-23].

Input layer (IP): This layer defines the dimension of the thermal images.

Convolutional layer (CL): The Conv layer implements feature detection at different positions on the thermal images. The filter is the key element of this layer and known as the line-shaped object scanning over entire the thermal images. The output is called as the Feature map which provides us information like the corners and edges.

Rectified linear unit layer (RL): A nonlinear activation function is activated on this layer to map nonlinearity into the data. Consequently, the learning process is converged faster.

Max-pooling layer (MP): For reduction of computational complexity and costs, subsampling the feature maps of the preceding Conv is implemented on this layer by computing the average or maximum responses in small overlapping neighborhoods.

BatchNormalization (BN): This layer consists of normalizing activation vectors from hidden layers that can stabilize the learning process and reduce the number of training epoch.

Flatten layer (FL): This layer makes the multidimensional input one-dimensional.

Dense layer (DL): This layer has deep connection, meaning that each neuron in dense layer receives input from all neurons of its previous layer.

B. Search-based method

Firstly, we define a block, which contains three component layers that are convolutional layer, rectified linear unit layer and BatchNormalization respectively (denoted as CL-RL-BN). A network segment contains differently consecutive blocks followed by a max-pooling layer. We investigate the CNN structure with 3 network segments, in which each network segment includes 1 or 2 or 3 consecutive CL-RL-BN blocks. Three considered CNN structures are called as CNN1, CNN2 and CNN3 respectively. Figures 2, 3 and 4 respectively show 3 CNN structures, which are performance estimation for selection of an optimal model.

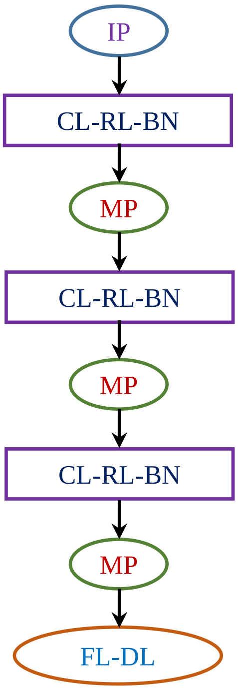


Figure 2. CNN1 structure with 3 network segments based on single CL-RL-BN blocks.

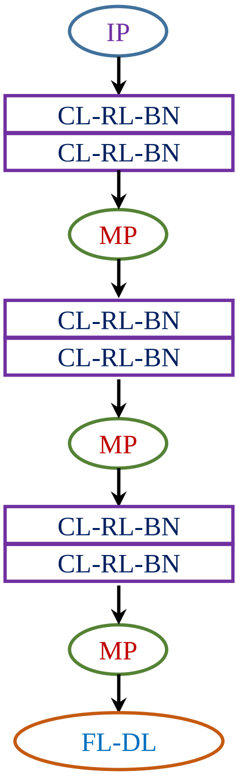


Figure 3. CNN2 structure with 3 network segments based on two consecutive CL-RL-BN blocks.

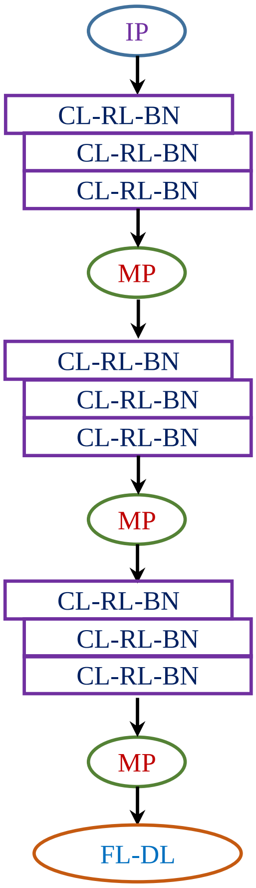


Figure 4. CNN3 structure with 3 network segments based on three consecutive CL-RL-BN blocks.

C. Temperature prediction

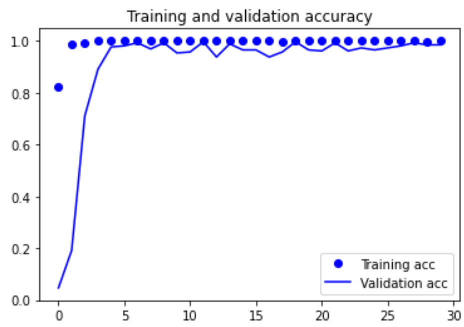
The binary thermal images are converted into temperature using standard equations [24-25], which is frequently used in infrared thermography.

IV. SIMULATION RESULTS AND DISCUSSION

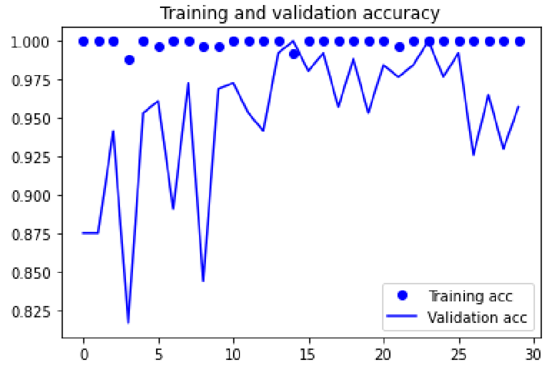
We use a performance parameter for model estimation, which is accuracy (Acc) to measure the person who is identified correctly.

A. Search-based method

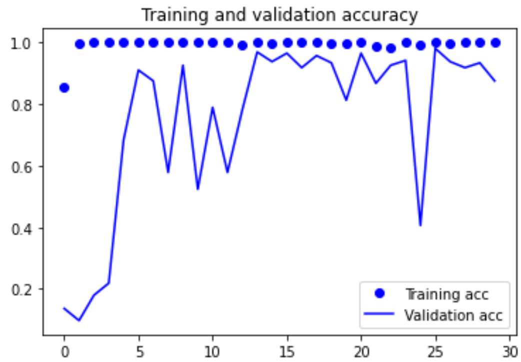
All the CNN structures are trained with 30 epoch and 128 steps-per-epoch on training dataset and then estimated on validation dataset to figure out the best number of epochs. Figure 5 shows the training and validation Acc of three CNN models.



a. CNN1



b. CNN2



c. CNN3

Figure 5. Accuracy of three CNN models on training and validation datasets.

Table 2 presents the accuracy of three CNN models on training and validation sets after 30 epochs.

TABLE 2. PERFORMANCE OF 3 CNN STRUCTURES ON DIFFERENT DATASETS

	CNN1	CNN2	CNN3
Training set	100%	100%	100%
Validation set	98.4%	94.1%	87.5%

As shown in Figure 5, the validation Acc of the CNN1 remains stable and better than that of the CNN2 and CNN3 through 30 epochs. With a small dataset, the CNN2 and CNN3 that have overwhelming capacity will overfit the training dataset and cannot perform accurately against unseen data.. Especially, the Ac of the CNN1 is similar from about epochs of 25 to 30. On the other hand, the figures for the CNN2 and CNN3 are erratic over the whole period of training and validation. Hence, we select the CNN1 model and an epoch number of 30 as the final face recognition algorithm. We further test the performance of the selected algorithm on the test set. The Ac of 98.4% implies that our proposed algorithm is effective and potential for application in practical environments.

B. Temperature prediction

As mentioned in [24], for simplicity, the flux-linear signal is related to scene temperature by the Planck curve and the conversion is typically approximated by the following equation,

$$S = \frac{R}{e^{\frac{B}{T_k}} - F} + O \qquad (1)$$

where S denotes the output signal from the thermal camera, R is the thermal camera responsivity, T_k is the target's absolute temperature in the unit of Kelvin, F and O are parameters generated during calibration.

The conversion from flux to temperature is performed using the inverse of Equation (1). Hence, we can express the equation (1) as,

$$T_k = \frac{B}{ln(\frac{R}{S-O} + F)} \qquad (2)$$

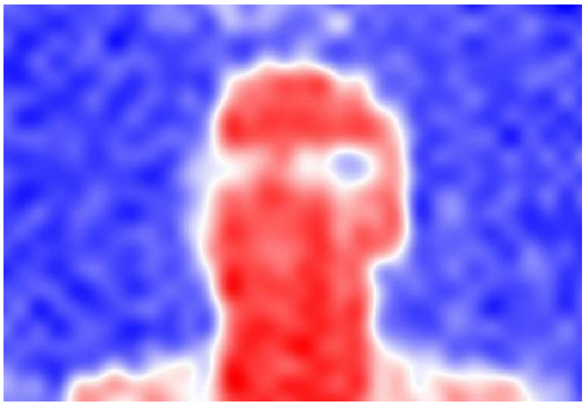


Figure 6. An example of thermal image for temperature prediction.

Based on the temperature of each pixel in images, we calculate the body temperature by using the average temperature of human face area. Figure 6 shows an example of the original thermal image on which the temperature of human face is computed. A person in this image has a normal temperature, which is often lower than 38-degree Celsius.

C. Discussion

The database used for this work includes only 18 individuals, which a relatively small number of classes, which possibly reduce the final detection performance. Another significant characteristic is that the number of thermal images of individuals is different. On other words, unbalanced data is available in this dataset, which make us difficult to use entire database for construction of our algorithm. Therefore, we only use a part of database to make the number of thermal images is equal for all individuals.

The facial occlusion issue is a serious challenging, which maybe cause an unstable validation Acc though epochs as shown in CNN2 and CNN3 of Figure 3. Obviously, the CNN1 structure is the better solution for such issue compared to the others. However, we only investigate three CNN structures in this work, which leave certainly a room for further research.

V. CONCLUSIONS

Correct detection of human face and body temperature are important at the first place, which allows to provide early warnings in terms of security and infection prevention during COVID-19 pandemic. In this paper, we have proposed a noble human face recognition using deep learning method, which can be applied for the advanced security systems. The algorithm is constructed carefully using search-based method to select the optimal CNN structure and its parameters. The proposed algorithm uses the thermal images as the input, which are also employed for calculation of human temperatures. The high-performance result shows that our proposed algorithm is effective and can be applied for practical security systems.

REFERENCES

[1] Z. Zhao, P. Zheng, S. Xu and X. Wu, "Object Detection With Deep Learning: A Review," in IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[2] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj and L. Song, "SphereFace: Deep Hypersphere Embedding for Face Recognition," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6738-6746, doi: 10.1109/CVPR.2017.713.

[3] M. Taskiran, N. Kahraman, and C. E. Erdem, "Face recognition: Past, present and future (a review)," *Digit. Signal Process.*, vol. 106, p. 102809, Nov. 2020, doi: 10.1016/J.DSP.2020.102809.

[4] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[5] I. Masi, Y. Wu, T. Hassner and P. Natarajan, "Deep Face Recognition: A Survey," 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, pp. 471-478, doi: 10.1109/SIBGRAPI.2018.00067.

[6] TV-GAN: Generative Adversarial Network Based Thermal to Visible Face Recognition Emmanouilidou, M. The Internet of Healthcare Things: A European Perspective and a Review of Ethical Concerns. International Journal of Business, Human and Social Sciences, 13(5), 675-679, (2019).

[7] Logu, K., Devi, T., Deepa, N., Kumar, S.R. and Gayathri, N. (2021). Real-Time Mild and Moderate COVID-19 Human Body Temperature Detection Using Artificial Intelligence. In AI and IoT-Based Intelligent Automation in Robotics (eds A.K. Dubey, A. Kumar, S.R. Kumar, N. Gayathri and P. Das). https://doi.org/10.1002/9781119711230.ch11.

[8] M. Hasan, S. N. H. S. Abdullah and Z. A. Othman, "Face recognition based on opposition particle swarm optimization and support vector machine," 2013 IEEE International Conference on Signal and Image Processing Applications, 2013, pp. 417-424, doi: 10.1109/ICSIPA.2013.6708043.

[9] Z. B. Lahaw, D. Essaidani and H. Seddik, "Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), 2018, pp. 1-5, doi: 10.1109/TSP.2018.8441452.

[10] F. Mahmud, M. T. Khatun, S. T. Zuhori, S. Afroge, M. Aktar and B. Pal, "Face recognition using Principle Component Analysis and Linear Discriminant Analysis," 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 2015, pp. 1-4, doi: 10.1109/ICEEICT.2015.7307518.

[11] Weiwei Zong and Guang-Bin Huang, "Face recognition based on extreme learning machine" 2011 Neurocomputing, vol. 74, no. 16, pp. 2541-2551, doi: 10.1016/j.neucom.2010.12.041.

[12] X. Ning, W. Li, B. Tang and H. He, "BULDP: Biomimetic Uncorrelated Locality Discriminant Projection for Feature Extraction in Face Recognition," in IEEE Transactions on Image Processing, vol. 27, no. 5, pp. 2575-2586, May 2018, doi: 10.1109/TIP.2018.2806229.

[13] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8.

[14] Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," in IEEE Access, vol. 6, pp. 35365-35381, 2018, doi: 10.1109/ACCESS.2018.2836950.

[15] P. P. Shinde and S. Shah, "A Review of Machine Learning and Deep Learning Applications," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697857.

[16] Zhan Wu, Min Peng and Tong Chen, "Thermal face recognition using convolutional neural network," 2016 International Conference on Optoelectronics and Image Processing (ICOIP), 2016, pp. 6-9, doi: 10.1109/OPTIP.2016.7528489.

[17] Głowacka, Natalia, and Jacek Rumiński. "Face with Mask Detection in Thermal Images Using Deep Neural Networks." Sensors (Basel, Switzerland) vol. 21,19 6387. 24 Sep. 2021, doi:10.3390/s21196387.

[18] Z. Wang, G. Horng, T. Hsu, C. Chen and G. Jong, "A Novel Facial Thermal Feature Extraction Method for Non-Contact Healthcare System," in IEEE Access, vol. 8, pp. 86545-86553, 2020, doi: 10.1109/ACCESS.2020.2992908.

[19] Litvin, A., Nasrollahi, K., Escalera, S. et al. A novel deep network architecture for reconstructing RGB facial images from thermal for face recognition. Multimed Tools Appl 78, 25259–25271 (2019). https://doi.org/10.1007/s11042-019-7667-4.

[20] S. M. Z. S. Z. Ariffin, N. Jamil and P. N. M. A. Rahman, "DIAST variability illuminated thermal and visible ear images datasets," 2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2016, pp. 191-195, doi: 10.1109/SPA.2016.7763611.

[21] G. Hu et al., "When Face Recognition Meets with Deep Learning: An Evaluation of Convolutional Neural Networks for Face Recognition," 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), 2015, pp. 384-392, doi: 10.1109/ICCVW.2015.58..

[22] C. Ding and D. Tao, "Robust Face Recognition via Multimodal Deep Face Representation," in IEEE Transactions on Multimedia, vol. 17, no. 11, pp. 2049-2058, Nov. 2015, doi: 10.1109/TMM.2015.2477042.

[23] Srivastava Y., Murali V., Dubey S.R. (2020) A Performance Evaluation of Loss Functions for Deep Face Recognition. In: Babu R.V., Prasanna M., Namboodiri V.P. (eds) Computer Vision, Pattern Recognition, Image Processing, and Graphics. NCVPRIPG 2019. Communications in Computer and Information Science, vol 1249. Springer, Singapore. https://doi.org/10.1007/978-981-15-8697-2_30.

[24] FLIR. Lepton radiometry application note. Datasheet, 2014.

[25] Julia, Converting DNs to temperature in FLIR cameras. Accessed on: January 27, 2022. Available: https://juliageographer.wordpress.com/2017/08/30/converting-dns-to-temperature-in-flir-cameras/

NHẬN DIỆN KHUÔN MẶT VÀ NHIỆT ĐỘ DỰA TRÊN THUẬT TOÁN THÔNG MINH TỰ ĐỘNG

Tóm tắt: Hệ thống giám sát đóng một vai trò thiết yếu trong lĩnh vực an ninh để nhận biết sớm các truy cập bất hợp pháp. Là một phần quan trọng của hệ thống giám sát tự động, tính năng nhận diện khuôn mặt người cung cấp một dịch vụ hiệu quả để quan sát những người băng qua một khu vực cụ thể. Hơn nữa, việc đo nhiệt độ cơ thể cũng là bắt buộc do đại dịch COVID-19. Do đó, dự đoán nhiệt độ thông qua hệ thống giám sát là lựa chọn tốt nhất do tính hiệu quả, xác thực cao và khả năng duy trì sự riêng tư. Trong bài báo này, một thuật toán mới được đề xuất giúp nhận dạng khuôn mặt người và nhiệt độ bằng cách sử dụng phương pháp học sâu và hình ảnh khuôn mặt nhiệt. Phương pháp dựa trên kĩ thuật tìm kiếm được phát triển để lựa chọn cấu trúc tối ưu và các tham số của mạng nơ ron tích chập bằng cách sử dụng các hình ảnh nhiệt làm đầu vào. Ngoài ra, nhiệt độ của con người được tính toán trực tiếp trên các hình ảnh nhiệt để hỗ trợ hệ thống cảnh báo về nhiệt độ cơ thể bất thường. Ứng dụng được đề xuất tạo ra độ chính xác tương đối cao trên 98%, phù hợp đáng kể trong môi trường thực tế.

Từ khoá: Ảnh nhiệt, Học sâu, Đo thân nhiệt, Nhận diện khuôn mặt.

Linh T. Nguyen is a B.E. student in Electronics and Telecommunications Engineering of Posts and Telecommunications Institute of Technology (PTIT) of Vietnam. His research interests include IoT, software-defined networks and Deep learning.

Duc Long Vu is currently a B.E. student in Electronics and Telecommunications Engineering of Posts and Telecommunications Institute of Technology (PTIT) of Vietnam. His research interests include cloud computing and machine learning.

Minh Tuan Nguyen received the B.E. degree in Electronics and Telecommunications Engineering from Posts and Telecommunications Institute of Technology (PTIT) of Vietnam in 2003, and the M.Eng. from Hanoi University of Science and Technology in 2007 and D.Eng. degrees in Electrical Engineering and Computer Science from Nagoya University of Japan in 2018, respectively. He is currently with PTIT. His research interests include Biomedical signal processing; Wireless signal processing; Signal decomposition; Biomedical application design; Optimization algorithm; Machine learning; Deep learning; Intrusion detection system; Internet of things; Internet of everything; Big data; Next generation of anti-jamming satellite communication; Underwater communication.

Hai-Chau Le received the B.E. degree in Electronics and Telecommunications Engineering from Posts and Telecommunications Institute of Technology (PTIT) of Vietnam in 2003, and the M.Eng. and D.Eng. degrees in Electrical Engineering and Computer Science from Nagoya University of Japan in 2009 and 2012, respectively. From 2012 to 2015, he was a researcher in Nagoya University of Japan and in University of California, Davis, USA. He is currently a lecturer in Telecommunications Faculty at PTIT. His research interests include optical technologies, network design and optimization and future network technologies. He is an IEEE member.