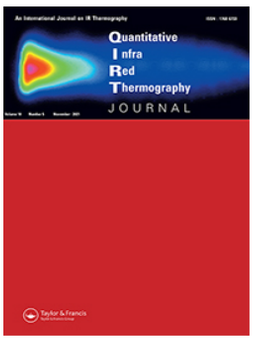

Quantitative InfraRed Thermography Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tqrt20

Automatic body part and pose detection in medical infrared thermal images

Ahmet Özdil & Bülent Yılmaz

To cite this article: Ahmet Özdil & Bülent Yılmaz (2021): Automatic body part and pose detection in medical infrared thermal images, Quantitative InfraRed Thermography Journal, DOI: 10.1080/17686733.2021.1947595

To link to this article: https://doi.org/10.1080/17686733.2021.1947595

QUANTITATIVE INFRARED THERMOGRAPHY JOURNAL
https://doi.org/10.1080/17686733.2021.1947595

Taylor & Francis
Taylor & Francis Group

Automatic body part and pose detection in medical infrared thermal images

Ahmet Özdil [a,b,c] and Bülent Yılmaz [b,c,d]

[a]Faculty of Engineering and Architecture, Computer Engineering Department, Kırşehir Ahi Evran University, Kırşehir, Turkey; [b]Graduate School of Engineering and Sciences, Electrical and Computer Engineering Department, Abdullah Gül University, Kayseri, Turkey; [c]Biomedical Instrumentation and Signal Analysis Laboratory, Abdullah Gül University, Kayseri, Turkey; [d]School of Engineering, Electrical and Electronics Engineering Department, Abdullah Gül University, Kayseri, Turkey

ABSTRACT

Automatisation and standardisation of the diagnosis process in medical infrared thermal imaging (MITI) is crucial because the number of medical experts in this area is highly limited.The current studies generally need manual intervention. One of the manual operations requires physician's determination of the body part and orientation. In this study automatic pose and body part detection on medical thermal images is investigated. The database (957 thermal images - 59 patients) was divided into four classes upper-lower body parts with back-front views. First, histogram equalization (HE) method was applied on the pixels only within the body determined using Otsu'sthresholding approach. Secondly, DarkNet-19 architecture was used for feature extraction, and principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE) approaches for feature selection. Finally, the performances of various machine learning based classification methods were examined. Upper vs. lower body parts and back vs. front of upper body were classified with 100% accuracy, and back vs. front classification of lower body part success rate was 93.38%. This approach will improve the automatisation process of thermal images to group them for comparing one image with the others and to perform queries on the labeled images in a more user-friendly manner.

ARTICLE HISTORY
Received 10 February 2021
Accepted 25 May 2021

KEYWORDS
Medical infrared thermal imaging; pose detection; computer-aided medical diagnosis

Introduction

Human body has many different systems to maintain thermal balance due to its homoeothermic structure. While it generates huge amount of heat, there are different ways to evacuate heat such as perspiration-evaporation, thermal conduction, forced and natural convection, exhalation and infrared (IR) radiation [1]. Core or surface temperature changes of the body indicate certain problems occurring during metabolism. IR thermography is an imaging modality that displays the heat distribution of the target as an image and has emerged to determine certain abnormalities using the temperature changes on the body. The first use of this imaging modality in healthcare was in 1956 by R. Lawson on the breast

CONTACT Ahmet Özdil ahmet.ozdil@ahievran.edu.tr Faculty of Engineering and Architecture, Computer Engineering Department, Kırşehir Ahi Evran University, Kırşehir, Turkey

tumours. Since 1995, there has been an increasing interest for the IR thermography due to the improvements in IR camera technology and image processing algorithms [1]. There are different abnormalities other than the breast cancer that the IR thermography is and can be used, such as the inflammatory arthritis, osteoarthritis, soft tissue rheumatism, tennis elbow, fibromyalgia, complex regional pain syndrome, peripheral circulation and fever [2]. Moreover, the breathing cycle and cardiac pulse wave monitoring can be performed using this technology [3]. In Mercer et al.'s work focusing on autologous transplantation, thermal imaging served as a tool to determine vessels that feed the transplanted tissue and to monitor compliance after transplantation [4]. In several other studies, this technology was used in managing injuries and inflammatory responses on horses or other athletic animals [5–8]. The researchers in this field have used manual or visual analysis of images for the assessment of thermal images; however, improvements in the image processing techniques on such images may open new avenues to apply IR thermal imaging to more complex problems.

On the other hand, the feasibility of using medical IR thermal imaging (MITI) in clinical practice has been a topic of research. The first attempt aims at making this technology available as a pre-diagnosis tool, and currently there are many clinics in the world that have been using MITI as an integral part of their diagnosis process (maybe not the only tool). Automatisation and standardisation of the diagnosis process are crucial because the number of medical experts working with MITI is highly limited.

To give a perspective on the MITI use in the clinics, how the imaging is performed should be explained. It is important to obtain thermal images from patients in a controlled environment, to achieve better results [9]. Patients should rest in a controlled room to settle down before the imaging begins. Precautions should be taken to fix the variables that have influence on thermal conditions. During thermal examination, patients should be disrobed because clothes prevent infrared camera sensing the infrared beams emitted from the body. The position of the camera is adjusted according to the height of the patient. The patient is first asked to stand still facing the camera, then turn almost 90 and/or 180 degrees around his/her own axis, and then raise his/her arms to show the armpits. The camera position is also lowered to facilitate imaging of the legs and feet. From each patient, 15–20 images are obtained for every visit and stored in a database mostly without any specific labelling. From this database, searching for any images with specific disease related diagnostic labels is not possible for now. For the pre-diagnosis phase, the physician should go over the images one by one, adjust colour maps manually and focus on specific regions/parts of the body visually.

Quantitative thermal image analysis emerged recently and needs full collaboration with thermal physiology of body beneath the skin and it will allow objective interpretation of thermogram [10]. Automatic processing of thermal images has been an emerging topic recently in the literature. The studies have been focusing on automatic image thresholding, region-of-interest (ROI) extraction, classification, and remote-sensing methods [11–14]. Automatisation is performed with deep learning (Convolutional Neural Networks, CNN) and conventional machine learning (Support Vector Machines, SVM) methods. However, these types of studies/approaches generally need pre-processing and/or preparation or selection of the images manually. Fully automatic systems need minimal manual intervention. One of the manual operations requires physician's determination of the body part (upper or lower) and orientation of the patient (anterior or

posterior pose) on each thermal image. Before performing image processing and automatic classification for specific diseases, it should be determined what kind of disease/s the image contain/s. In the literature, this is fulfilled mainly by manual selection of a physician. However, there is no study on automatic determination of body part and pose from thermal images.

In [15], deep learning methods were used to estimate the sleeping position as lateral and supine of the head and body separately. In that study by using a pre-trained CNN architecture called DarkNet-19 92% accuracy level was attained. Different CNN architectures were investigated for semantic segmentation on thermal images in [16], and it was found that the encoder–decoder architectures had better performance than the other architectures did. In [17], Bayesian-based Gaussian Mixture Model (GMM) approach proposed for segmenting and semi-supervised classification of image data; however in [18], a deterministic approach used for data segmentation. In another study [19], exercise-induced fatigue was detected on thermal images using CNN approach with an accuracy over 80%.

Recently in [20], diabetic foot ulcer identification from plantar feet thermal images were evaluated using different machine learning algorithms and k-NN with five neighbours yielded 81.25% accuracy. In [21], two different skin cancer tumours were differentiated using SVM. Another study was carried on predicting haemodynamic shock on thermal images using histogram of oriented gradients (HOG) and random forest classifier [22]. The effect of MITI on diagnosis of cardiovascular disease was studied in [23], and the results were comparable with the standard methods, and even better when combined with standard methods in a computer aided diagnosis system. Moreover, assessment of hypertension was carried out in [24], different pre-processing methods were evaluated, and backpropagation neural network classifier resulted in 89% accuracy.

In this study, automatic body part and pose detection process was investigated on an authentic medical thermal image database that have been developed in collaboration with a clinic in Kayseri, Turkey. Different thresholding and segmentation methods have been evaluated, and binarizing the image according to the threshold determined using Otsu's method performed well in terms on performance, speed and simplicity compared to other more complex methods like active contours region growing and local first-order statistics. Thus, Otsu's thresholding-based ROI extraction and histogram equalisation method within the selected ROI were used to prepare images for the feature extraction phase. Later, features were extracted using DarkNet-19 and t-distributed stochastic neighbour embedding (t-SNE) and principal component analysis (PCA) algorithms were employed to select features from the images. The selected features were employed in the classification of the images for body part (upper/lower) and pose (front/back). As a result, an image was able to labelled with its body part and pose automatically which was a crucial phase in managing unorganised MITI databases. All studies on thermal images need a preliminary step to determine if any image contains the targeted body part and/or pose. Especially the medical centres that investigate the whole body with MITI need an automatic approach to group all the images (hundreds of thousands of them) taken from the patients for the comparison purposes and perform queries on the automatically labelled images.

The aim of this study is that this approach will improve the automatisation process of thermal image processing and help physicians study the thermal images in a more user-friendly manner than it is now.

Materials and methods

Thermal images

In this study, medical infrared thermal images were captured during routine inspection in a special clinic owned by a physician (Dr. M. M. YILMAZ Clinic, Kayseri, TURKEY) using IRIS-XP Infrared Thermography Device (Medicore Co. Ltd., Seoul, South Korea). Iris-XP, is a Class I medical device, that can measure temperatures between 14.5 and 40°C. It has both auto and manual focus and the resolution of the images is 480 (Horizontal) by 640 (Vertical) pixels. The focus range is from 40 cm to infinity, and the system has PACS compatibility. In this clinic, thermal images have been used as a pre-diagnosis tool for the last 10 years. In this study, 957 thermal images from 59 patients (44 male, ages 2–68) consisting of different body parts and poses were used. The images had a fixed size of 480 × 480 pixels. Figure 1 demonstrates eight sample raw images from the database.

This study was approved by the Erciyes University Ethical Council of Clinical Studies, Kayseri, Turkey (2019/524). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

The thermal image set was consisting of 3764 images from 63 patients. However, the image set was not standardised due to working conditions of the clinic, and only 957 images were able to be used for evaluations. Most of the non-standard images were eliminated; however, some images were kept in the image set after we successfully identified the patient as the foreground. For example, in several images, the dress of the patient interfered in the image but was assigned as the background using our pre-processing approach (please see Figure 2).

For the pose and body part detection of this study which included four different classification problems, four different subsets were prepared from the database. The first subset consisted of four classes as upper-back (261 images), upper-front (273

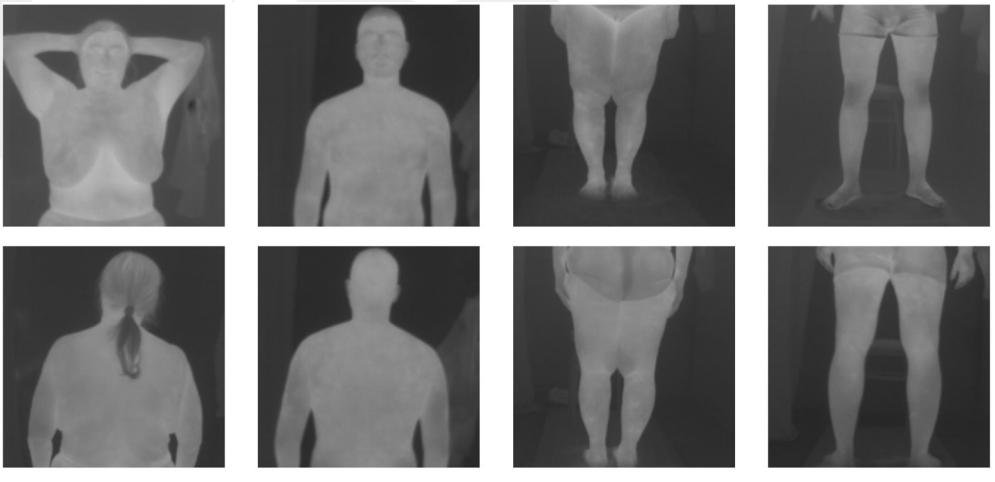

Figure 1. Sample raw thermal images.

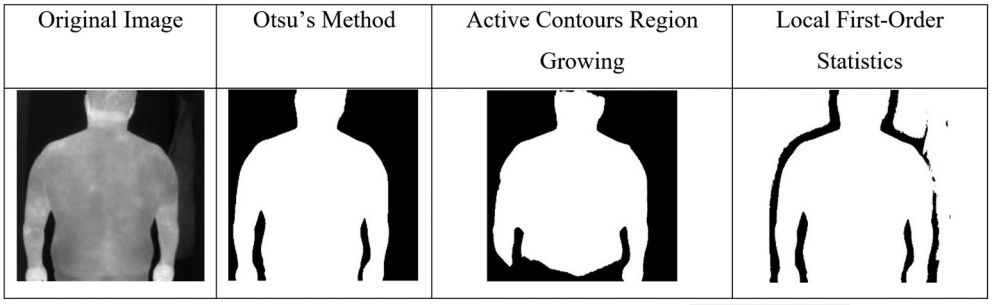

Figure 2. Evaluation of segmentation methods.

images), lower-back (210 images), and lower-front (213 images), which includes all classes in the entire image dataset. The second subset consisted of upper (534 images) and lower (423 images) parts of the body for the second classification problem which is called upper versus lower. The third subset included 210 posterior/back and 213 anterior/front poses for the third classification problem, which is called lower back versus lower front. The last subset was comprised of 261 back and 273 front images from the upper parts of the body for the fourth classification problem, which is called upper back versus upper front.

Pre-processing of thermal images

A thermal image consists of thermal values ranging from the lowest room temperature to the highest body temperature. The informative portion of the temperature range exists on the body temperature level, which makes lower temperatures uninformative. The pixels corresponding to non-body parts of the scene decrease the success of the histogram equalisation (HE) approach, and the resulting image becomes a low contrast one, which makes the success of the classification process also low. For this reason, a different approach was prepared that the HE method should be used to equalise the histogram of the pixels only within the ROI in order to improve the feature selection procedure. HE method was used to adjust the contrast by manipulating the histogram of the image [25]. Otsu's approach [26], which is a well-known thresholding method for grey level images, was used for discriminating foreground (patient) from the background (non-body parts). Different thresholding and segmentation methods have been evaluated and binarizing the image according to the threshold determined using Otsu's method was good on performance, speed and easy to apply aspects compared to other methods. After determining a proper threshold for each image separately with Otsu's method, the images were binarized with the calculated threshold to generate the background mask. The binarized images served as masks for the body parts, which is referred to as the ROI, and the histogram equalisation was performed on the mask of each image one-by-one.

Feature extraction and selection

Convolutional Neural Networks (CNNs) have recently become a popular approach for image processing. In this study, the pre-trained DarkNet-19 architecture was used for

Table 1. DarkNet-19 architecture [27].

Type	Filters	Size/Stride	Output
Convolutional	32	3 × 3	224 × 224
Maxpool		2x2/2	112 × 112
Convolutional	64	3 × 3	112 × 112
Maxpool		2x2/2	56 × 56
Convolutional	128	3 × 3	56 × 56
Convolutional	64	1 × 1	56 × 56
Convolutional	128	3 × 3	56 × 56
Maxpool		2x2/2	28 × 28
Convolutional	256	3 × 3	28 × 28
Convolutional	128	1 × 1	28 × 28
Convolutional	256	3 × 3	28 × 28
Maxpool		2x2/2	14 × 14
Convolutional	512	3 × 3	14 × 14
Convolutional	256	1 × 1	14 × 14
Convolutional	512	3 × 3	14 × 14
Convolutional	256	1 × 1	14 × 14
Convolutional	512	3 × 3	14 × 14
Maxpool		2x2/2	7 × 7
Convolutional	1024	3 × 3	7 × 7
Convolutional	512	1 × 1	7 × 7
Convolutional	1024	3 × 3	7 × 7
Convolutional	512	1 × 1	7 × 7
Convolutional	1024	3 × 3	7 × 7
Convolutional	1000	1 × 1	7 × 7
Avgpool		Global	1000
Softmax			

feature extraction from the thermal images. The Darknet-19, which was developed by Redmon et al. [27], consists of a CNN architecture with 19 layers as depicted in detail in Table 1.

1024 features were extracted using 'avg1' layer of DarkNet-19 from each image. After the feature extraction phase, feature selection was applied to decrease the number of features. Principal component analysis (PCA) [28] and t-distributed stochastic neighbour embedding (t-SNE) [29] algorithms were applied for this purpose. PCA is a method to represent data, which captures most variable dimensions of data and removes low ones to reduce the data dimension. PCA was selected for testing due to its low cost in computation [28]. In contrast, t-SNE embeds high-dimensional points to lower dimensions by considering similarities between neighbours [29] and calculates generally 2–3 features from the available feature set. Different number of features were selected using PCA and t-SNE and the effect of the number of features was investigated for each selection approach and classification methodology.

Pose detection-classification

Different standard classification methods were applied on selected feature sets for four different image-subsets separately. The Waikato Environment for Knowledge Analysis (Weka) [30] and MATLAB were the software tools that were used for testing different classification approaches. MATLAB's 'Deep Learning Toolbox' was used during CNN architecture evaluations. Features were extracted with DarkNet19 pretrained network available in that toolbox. After the feature extraction and selection phases, the classification methods were evaluated on Weka except for the support vector machines (SVM)

method, which was performed on MATLAB. 59 classification methods were applied to feature sets which were selected by PCA and t-SNE methods. These feature sets contained up to 956 features according to the number of samples in the datasets. For the first, second, third and fourth data sets 956, 956, 534 and 422 different ARFF files are prepared, respectively, for this purpose. The selected methods with best training performances among 59 methods were used in this study were Random Forest (RaF), Rotation Forest (RoF), Sequential Minimal Optimisation (SMO), Bayesian Logistic Regression (BLR), Instance-Based 1 Nearest Neighbour (IB1), Reduced Error Pruning Tree (REPTree) and Support Vector Machines (SVM). These methods were the best among the tested methods available in Weka and MATLAB in terms of classification accuracy. The details of these approaches can be found in [31]. The classification accuracy metrics used were the accuracy, true positive rate, false positive rate, precision, recall, and F-measure.

Results

Figure 2 shows a visual comparison of different thresholding and segmentation methods that were evaluated in this study. Otsu's method performed significantly better in terms on performance, speed and simplicity compared to other more complex methods like active contours region growing and local first-order statistics. As depicted in the figure, Otsu's method was more successful in distinguishing the patient and the background than the other two approaches.

Figure 3 demonstrates the outcomes of Otsu's thresholding approach for three sample images to differentiate the pixels corresponding to body parts from the ones corresponding to non-body parts (background). On the left column original images are presented and, in this column, although the room temperature was controlled, the clothes of the patient could be seen as part of the background, therefore a proper method is needed to successfully segment body parts in the image. There were several images like this one in our image set because the imaging procedure was not fully standardised due to working conditions of the clinic. The masks obtained using Otsu's thresholding method indicated as the ROIs on the images (middle panel) were used to apply on the original image whose output can be seen on the right most column, that is, only the intensity values corresponding to the pixels inside the mask were kept as they are and the intensities on the remaining pixels (background) were set to zero to be depicted as black.

Figure 4 shows the outcome of applying histogram equalisation (HE) method on the original or pre-processed images. The figure demonstrates the superior performance of applying the HE method only for the intensity values corresponding to the binary mask obtained using Otsu's thresholding approach. On the left column, the sample raw thermal images can be seen. The second column demonstrates the results of applying the HE method on the original image without any masks. It is obvious that the objects on the background may negatively affect the feature extraction process. On the third column, the HE method was applied on the whole masked image (please see right-most column of Figure 4); however, the temperature/intensity details over the body parts were lost. On the right-most column, the HE method was applied only within the ROI, which did not include the background pixels. It was observed that the temperature details on the body parts were sharpened and more obvious than the other two cases. This step was critical in terms of feature quality extracted from the

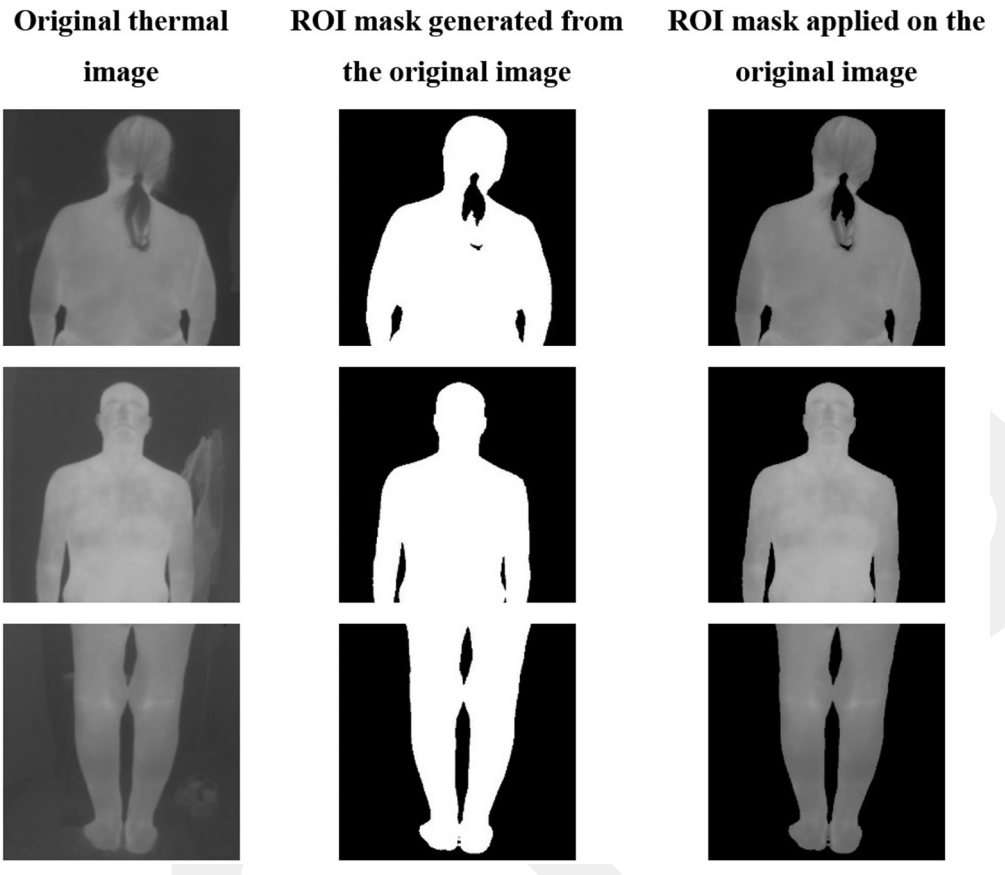

Figure 3. Application of ROI mask generated using Otsu's thresholding method on thermal images.

images. Thus, it was decided to perform histogram equalisation only on the body parts determined by the Otsu's thresholding method.

Medical thermal image analysis aims to explore the slightly thermal differences on body surface, even less than a half degree can be meaningful. However, there is about 10-degree Celsius difference between body temperature and room temperature, which must be excluded while processing thermal images. Since the background temperature values were not included in the histogram equalisation process, the resulting images were better in terms of being able to differentiate the slight changes on the temperature values on the body when HE was applied on the pixels only within the mask or when HE applied on the whole image. This was the case not only for these sample images but also for many other thermal images in the database.

The first image dataset was prepared for multi-class classification problem to classify four classes into: lower-back, lower-front, upper-
back and upper-front. The best correctly classification rate observed during experiments was 96.55% with Sequential Minimal Optimisation SMO classifier using 103 features selected by PCA.

Then, multi-class classification problem was changed into binary-class classification by increasing classification steps to improve classification accuracy. For this purpose, the first

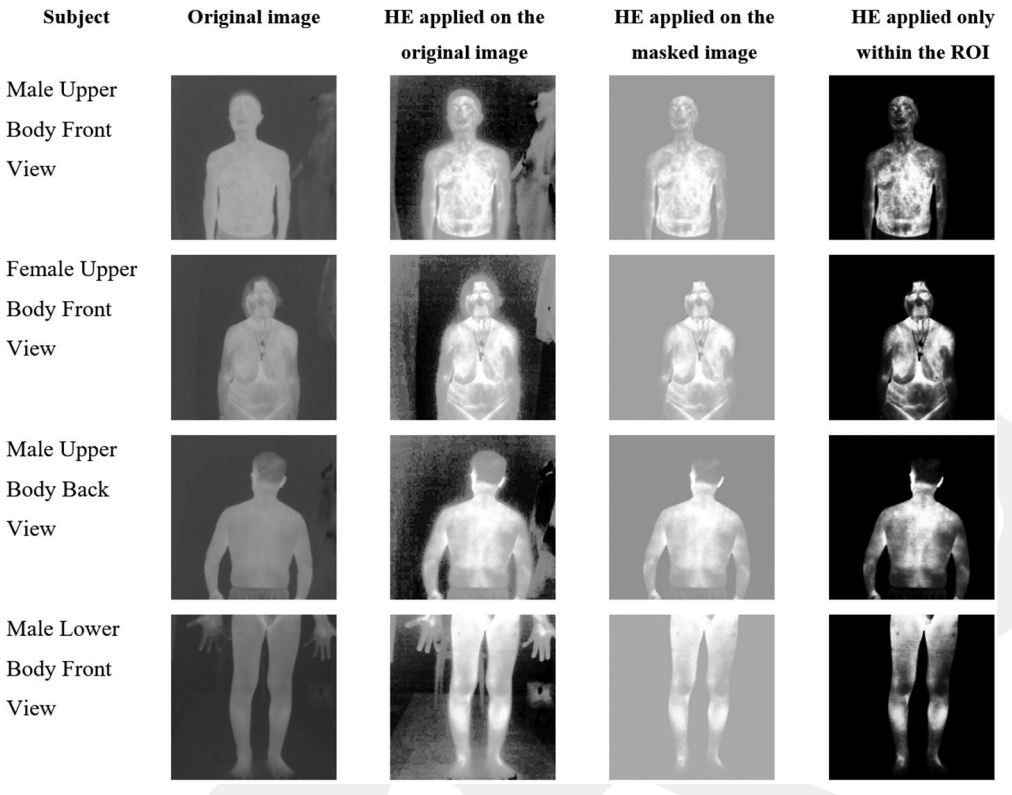

Figure 4. The results of histogram equalisation (HE) method with and without focusing on the body parts as the ROI The first column: original image, second column: HE applied on the whole image, third column: HE applied on the whole masked image, and the fourth column: HE applied only within the ROI.

dataset was divided into two classes as upper and lower body parts for the first step of classification, and for the second step the resulting two datasets were divided into front and back classes separately. In the end, four classes were acquired. Selected features in the second image subset were successful in separating the upper and lower body parts. However, the selected features on the third image subset could not differentiate the back and front poses of the lower part of the body. For the fourth image subset, the discrimination success increased as good as the second case. The feature selection results using two most features on the first, second, third and fourth image subsets using PCA and t-SNE can be seen as scatter plots in Figure 5. PCA selected the two most principal components and t-SNE calculated two features.

The results depicted in this figure clearly show that t-SNE outperforms PCA in terms of separating these two features and further evaluations were conducted to find out more-than-two-feature bundles can perform better. During evaluations, all four datasets features reached up to the number of images in the subsets, and 59 different classifiers were applied on generated feature sets using WEKA. For example, for the first dataset 956 different incremental ARFF files were prepared using 956 components selected by PCA. These 956 files consisted of the features as the following: the first ARFF file contains 1

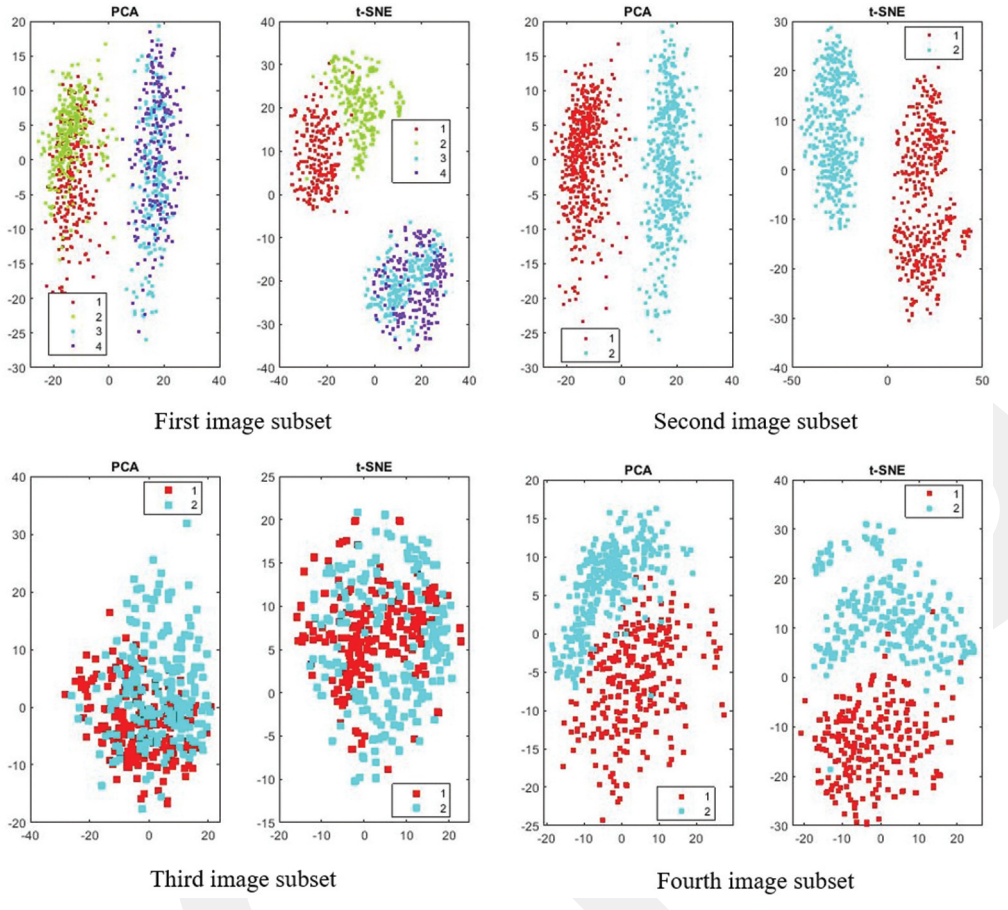

Figure 5. PCA and t-SNE results on the first, second, third and fourth image subsets.

feature, the second ARFF file contains 2 features and 956th ARFF file contains 956 features. For the second, third and fourth datasets 956, 422 and 533 different ARFF files were generated, respectively. These evaluations demonstrated that increased number of features perform better classification results on both PCA and t-SNE than low number of features do (Table 2).

The second image subset was classified with 100% accuracy since upper and lower parts of the body have significant differences. Lower part has two legs, but upper part has a head, two arms and the body, which makes significant difference and automatic classification achieved a great success. However, the success of the back versus front classification of lower and upper parts separately decreased, since both parts have high similarity between front and back scenes. Although upper part separation on fourth dataset, success was as good as the second dataset, lower part separation should have been increased by further processing techniques.

Figure 6 demonstrates the resemblance of lower part back view in the first row, and front view in the second row. One can infer that the degraded performance was caused by this resemblance. The high resemblance here can be reduced by using other pre-

Table 2. Performance metrics for different classification methods on different image subsets.

Image subset	Classification method	Correctly classified %	TP rate	FP rate	Precision	Recall	F-measure	Feature Set
1	RaF	93	0.93	0.02	0.93	0.93	0.93	t-SNE 5
1	SMO	96.55	0.97	0.01	0.97	0.97	0.97	PCA 103
1	SVM	83.5	0.83	0.17	0.86	0.83	0.83	
2	IB1	100	1	0	1	1	1	t-SNE 2
2	REPTree	100	1	0	1	1	1	PCA 2
2	RoF	100	1	0	1	1	1	t-SNE 2
2	SVM	100	1	0	1	1	1	
3	BLR	93.38	0.93	0.07	0.94	0.93	0.93	PCA 60
3	SMO	93.14	0.93	0.07	0.93	0.93	0.93	PCA 55
3	IB1	82.98	0.83	0.17	0.83	0.83	0.83	t-SNE 2
3	SVM	81	0.81	0.19	0.86	0.81	0.80	
4	BLR	100	1	0	1	1	1	PCA 104
4	IB1	97.19	0.97	0.03	0.97	0.97	0.97	t-SNE 2
4	SMO	100	1	0	1	1	1	PCA 110
4	SVM	85	0.85	0.16	0.86	0.84	0.84	

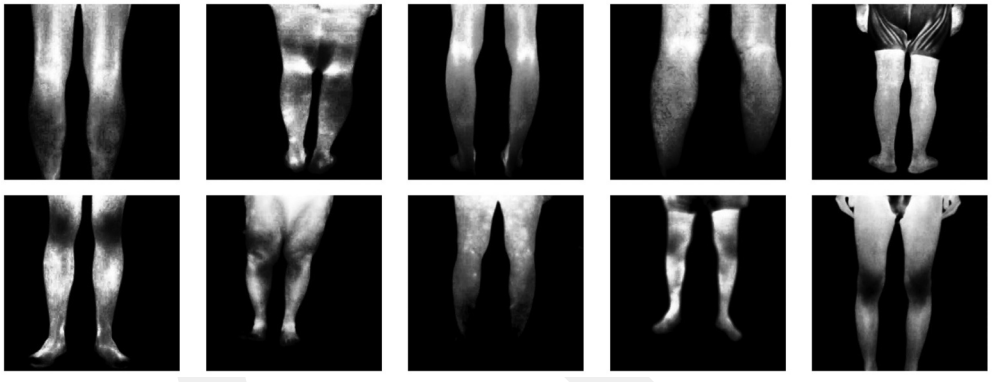

Figure 6. A visual explanation of the resemblance of the back and front views of the lower part of the body. Lower part **back view** is shown on the first row, and lower part **front view** on the second row.

processing approaches or by eliminating non-standard images from the image set in a future study.

Table 2 summarises the results of automatic classification with different methods for abovementioned classification problems. The first column demonstrates the data set number used for classification. Since four different data sets were prepared from the thermal images set, the results of three or four methods showing the highest performances were reported here even though the performances of many other methods were investigated. Image subset 1 includes all the images divided into four classes. Image subset 2 also includes all images, but this time images were divided into upper vs. lower classes. Image subsets 3 and 4 include upper and lower images respectively, and in these subsets, images were divided into front versus back classes. The second column indicates the best performing classification methods for each dataset indicated on the first column. The other columns present the associated method's performance metrics such as the classification accuracy, true positive (TP) rate, false positive (FP) rate, precision, recall, and F-measure. The right-most column includes the feature selection method (PCA or t-SNE) and the number of selected features. SMO-based classification method performed

remarkable well when compared to other methods for all image subsets. One may argue that since the SMO divides the problem into smaller sub-problems and is used to solve quadratic functions it is better suited for the datasets in this study.

In Figure 7, the performance comparison of evaluated methods can be seen. The first dataset has relatively low success rate due to the multi-class classification problem it contains. The second and fourth datasets were well suited for the classification methods and gained 100% performance. However, the performance on the third dataset was not as good as the other datasets, but it had a promising performance, which was over 90% accuracy. Different processing methods will be considered for further studies.

DISCUSSION and CONCLUSIONS

MITI is gaining an increased attention in medicine. Since it is a cheap, harmless/non-invasive, and non-destructive approach for medical examination, it has a potential to be employed more frequently as a pre-diagnostic imaging modality. Since the number and skills of experts using MITI in clinical practice are highly limited, automatic schemes might be used to close this gap. The first step for the development of computer-aided diagnosis tool is to determine,

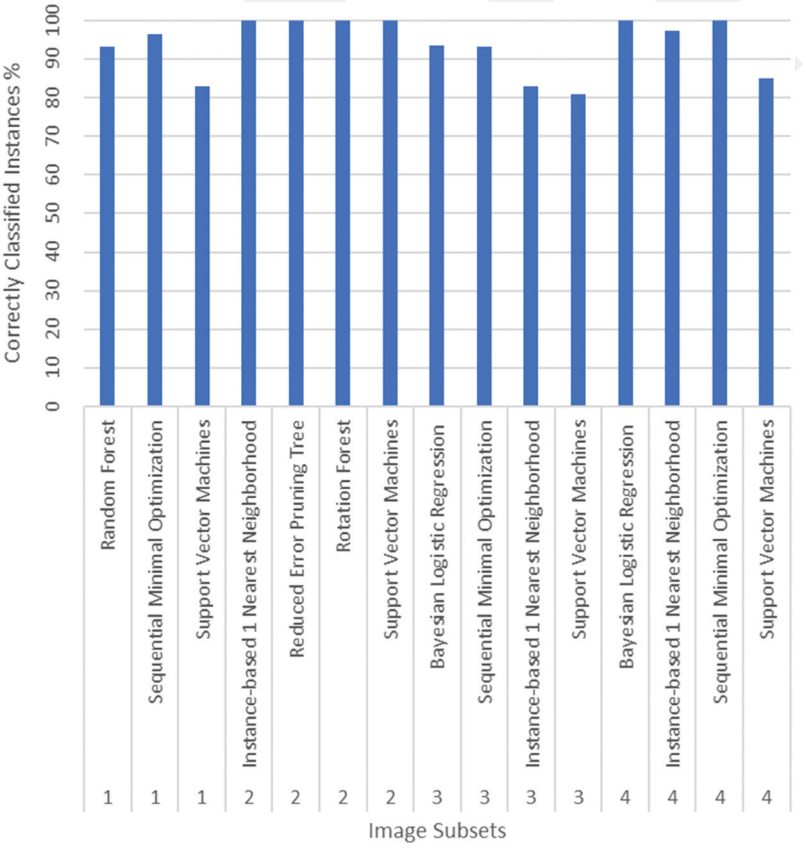

Figure 7. Summary of the classification accuracies for the best-performing approaches on 4 different image subsets.

which part of the body any thermal image contains for further processes such as searching similar cases, creating subsets, performing comparisons between patients or before and after treatment images, and labelling properly. However, in the literature, automatic detection of different body parts on thermal images automatically has not been studied so far.

It is envisioned that one possible area of body part detection might be the proper labelling of huge datasets for academic use. As expected, upper and lower body parts could be classified with great success and classification of back versus front of upper body part had same success rate (100%). The clinics employing MITI would need an automatic approach to group thousands of images taken from the patients that is stored in their databases for performing queries on the automatically labelled images and help the physician to compare a case with the other cases in the database.

There are many approaches that can be used to extract features from images. These approaches can be grouped as follows: Low-level (edge, corner, blob, ridge detection and scale-invariant feature transform-SIFT), curvature, shape-based, deformable and active contours [25]. Another group of approaches include texture analysis such as grey-level co-occurrence, grey-level run length, contrast, entropy, homogeneity, etc [25]. A relatively new set of techniques to extract features from the images is the deep neural networks (NNs) such as convolutional neural networks (CNNs) [32]. CNNs have also been used as the feature extractors. The filter slides along the image and generates a feature map. This map becomes the input of the subsequent layer. This is followed by other layers (pooling layers, fully connected layers, etc.). This whole process lets textural and spectral features to be revealed and exploited. Furthermore, there are numerous studies in the literature in which CNNs were used as the feature extractors [33–35], and in this study, it was decided to use a CNN based feature extraction approach to be investigated on the medical thermal images. DarkNet-19 is a CNN that is employed as the backbone of YOLOv2.

One of the limitations of this study was the relatively low number of images for processing and classification. The clinic is in the process of collecting more thermal images for different diseases coming from the patients visiting the clinic. Another limitation of the study was that there were non-standard images in the image set. The posture of patients was not similar due to age variance, especially it was hard for children and elderly patients to stand on the right position. This may be overcome by making thermal camera move 360 degrees around the patient where patient stands still, but this is out of the scope of this study.

Current and future research efforts in our laboratory are aiming and will aim at studying different aspects of MITI for it to become an accepted pre-diagnosis tool for certain diseases and conditions in clinical practice. Many researchers believe that this technology has this potential and there are many clinics and experts around the world that employ MITI in their diagnosis (or pre-diagnosis) process. It is obvious that the automatisation and standardisation of the diagnosis is highly necessary. Along these lines several parallel studies are underway aiming at the automatic detection of certain diseases using conventional machine learning and deep learning approaches from medical thermal images. Possible research directions include the detection of pain centres from the dorsal view of the upper body, or finding the varicose vein extension on the legs using thermal images.

Acknowledgments

This study was approved by the Erciyes University Ethical Council of Clinical Studies, Kayseri, Turkey (2019/524) and was carried out in collaboration with Dr. M.M. YILMAZ Clinic in Kayseri, Turkey, and we thank Dr. Mustafa Mücahit Yılmaz for his sincere help. We would like to thank Assoc. Prof. Zafer Aydın from Abdullah Gul University for his constructive and competent criticism on the scientific content of this work.

The dataset used in this study will be provided upon request.

Disclosure statement

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Notes on contributors

Ahmet Özdil obtained his bachelor's degree from Yıldız Technical University, İstanbul, Turkey in 2008 and master's degree from Melikşah University, Kayseri, Turkey in 2015. He is currently pursuing a PhD in Electrical and Computer Engineering at Abdullah Gül University, Kayseri, Turkey specializing in medical infrared thermal image processing and classification. He has got 10 years of research experience in the field of Computer Science. His research interests include medical image processing, computer vision and biometric image processing & recognition, signal processing.

Bülent Yılmaz is a Professor of Electrical-Electronics Engineering at Abdullah Gül University (AGU), Kayseri, Turkey. He has got over 20 years of research experience in the fields of Biomedical Image and Signal Processing and Machine Learning. Prof. Yilmaz received his BS and MS degrees from Electrical-Electronics Engineering at the Middle East Technical University (METU), Ankara, Turkey in 1997 and 1999 respectively. He got his PhD degree from Bioengineering Department of the University of Utah, Salt Lake City, Utah, USA in 2004. His current research interests are biomedical signal and image processing, especially brain-computer interfaces and automatic disease detection. He is the principal investigator at the Biomedical Instrumentation and Signal Analysis (BISA) Laboratory at AGU. He is currently serving as the vice rector of AGU.

ORCID

Ahmet Özdil http://orcid.org/0000-0002-6651-1968
Bülent Yılmaz http://orcid.org/0000-0003-2954-1217

References

[1] Jiang LJ, Ng EY, Yeo AC, Wu S, Pan F, Yau WY, Chen JH, Yang Y. A perspective on medical infrared imaging. J Med Eng Technol. 2005 Nov-Dec;29(6):257–67. doi:10.1080/03091900512331333158. PMID:16287675.

[2] Ring EFJ, Ammer K. Infrared thermal imaging in medicine. Physiol Meas. 2012 Mar;33(3):R33–R46.

[3] Andras S, Hana Skala K, Simeon G. Infrared thermography and image analysis for biomedical use. Period Biol. 2011;113(4):385–392.

[4] Mercer J, De Weerd L, Miland Å, et al. Pre-, intra-, and postoperative use of dynamic infrared thermography (DIRT) provides valuable information on skin perfusion in perforator flaps used in reconstructive surgery. Proc Inframation. 2010;2010(1889):313–320.

[5] Purohit RC, McCoy MD. Thermography in the diagnosis of inflammatory processes in the horse. Am J Vet Res. 1980 Aug;41(8):1167–1174.

[6] Turner TA. Diagnostic thermography. Vet Clin North Am Equine Pract. 2001 Apr;17(1):95–114.
[7] Eddy A, Van Hoogmoed L, Snyder J. The role of thermography in the management of equine lameness. Vet J. 2001 Nov;162(3):172–181.
[8] Holmes LC, Gaughan EM, Gorondy DA, et al. The effect of perineural anesthesia on infrared thermographic images of the forelimb digits of normal horses. Can Vet J. 2003 May;44 (5):392–396.
[9] Bagavathiappan S, Jayakumar T, Raj B, et al. Infrared thermal imaging for detection of peripheral vascular disorders. J Med Phys. 2009 Jan.;34(1):43. .
[10] Ring FJ. Pioneering progress in infrared imaging in medicine. Quant Infrared Thermogr J. 2014;11(1):57–65.
[11] Singh J, Arora AS. Automated approaches for ROIs extraction in medical thermography: a review and future directions. Multimed Tools Appl. 2020;79(21–22):15273–15296.
[12] Manda MP, Kim HS. A fast image thresholding algorithm for infrared images based on histogram approximation and circuit theory. Algorithms. 2020 Aug.;13(9):207.
[13] Sahraoui Y, Korichi A, Kerrache CA, et al. Remote sensing to control respiratory viral diseases outbreaks using Internet of Vehicles. Trans Emerg Telecommun Technol. 2020 Sep;e4118. https://onlinelibrary.wiley.com/doi/full/10.1002/ett.4118
[14] Aarthy SL, Prabu S. Classification of breast cancer based on thermal image using support vector machine. Int J Bioinform Res Appl. 2019;15(1):51–67.
[15] Akbarian S, Delfi G, Zhu K, et al. Automated non-contact detection of head and body positions during sleep. IEEE Access. 2019;7:72826–72834.
[16] Kakileti ST, Dalmia A, Manjunath G. Exploring deep learning networks for tumour segmentation in infrared images. Quant Infrared Thermogr J. 2020 Jul.;17(3):153–168.
[17] Gogebakan M, Erol H. A new semi-supervised classification method based on mixture model clustering for classification of multispectral data. J Indian Soc Remote Sens. 2018;46 (8):1323–1331.
[18] Gogebakan M, Erol H. Mixture model clustering using variable data segmentation and model selection: a case study of genetic algorithm. Math Lett. 2019;5(2):23.
[19] Lopez MB, Del-Blanco CR, Garcia N. Detecting exercise-induced fatigue using thermal imaging and deep learning, ***Proceedings*** *7th* ***International Conference*** *on* ***Image Processing Theory, Tools*** *and* ***Applications*** *IPTA 2017*, Montreal, Canada; 2018 Jan 1–6; 2018.
[20] Vardasca R, Magalhaes C, Seixas A, et al. Diabetic foot monitoring using dynamic thermography and AI classifiers. Quant Infrared Thermogr J. 2019. DOI:10.21611/qirt.2019.027
[21] Magalhaes C, Vardasca R, Rebelo M, et al. Distinguishing melanocytic nevi from melanomas using static and dynamic infrared thermal imaging. J Eur Acad Dermatol Venereol. 2019;33 (9):1700–1705.
[22] Nagori A, Dhingra LS, Bhatnagar A, et al. Predicting hemodynamic shock from thermal images using machine learning. Sci Rep. 2019;9(1):1–9.
[23] Thiruvengadam J, Mariamichael A. Model-based computer-aided method for diagnosis of cardiovascular disease using IR thermogram. Biomed Res. 2019;30(1):95–101.
[24] Thiruvengadam J, Mariamichael A. A preliminary study for the assessment of hypertension using static and dynamic IR thermograms. Biomed Tech. 2018;63(2):197–206.
[25] Gonzalez RC, Woods RE. *Digital image processing (Second Edition)*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall; 2002.
[26] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. 1979;9(1):62–66.
[27] Redmon J, Farhadi A, YOLO9000: better, faster, stronger, ***Proceedings*** *30th* ***IEEE Conference*** *on* ***Computer*** *Vision and* ***Pattern Recognition***, *CVPR 2017*, Honolulu, HI, USA; 2017 Jan; 2017, p. 6517–6525.
[28] Karamizadeh S, Abdullah SM, Manaf AA, et al. An overview of principal component analysis. J Signal Inf Process. 2013;4(3):173–175.
[29] van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9 (11):2579–2605.

[30] Hall M, Frank E, Holmes G, et al. The WEKA data mining software. ACM SIGKDD Explor Newsl. 2009 Nov.;11(1):10–18.
[31] U. of Waikato. Class hierarchy (weka-dev 3.9.5 API). [cited 2021 Jan 4]. https://weka.sourceforge.io/doc.dev/overview-tree.html
[32] Chollet F. Deep learning with python. Vol. 361. New York: Manning; 2018.
[33] Li Q, Cai W, Wang X, et al., Medical image classification with convolutional neural network, *2014 13th International Conference on Control Automation Robotics and Vision, ICARCV 2014,* Singapore; 2014 Dec; 2014, p. 844–848.
[34] Yang J, Wang W, Lin G, et al. Infrared thermal imaging-based crack detection using deep learning. IEEE Access. 2019;7:182060–182077.
[35] Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and the future. In: Dey N, Ashour AS, Borra S, editors. Classification in BioApps: automation of Decision Making. Cham: Springer International Publishing; 2018. p. 323–350.