Tech Science Press

Affective State Recognition Using Thermal-Based Imaging: A Survey

Mustafa M. M. Al Qudah, Ahmad S. A. Mohamed* and Syaheerah L. Lutfi

School of Computer Sciences, Universiti Sains Malaysia, Penang, 11800, Malaysia
*Corresponding Author: Ahmad S. A. Mohamed. Email: sufril@usm.my
Received: 10 November 2020; Accepted: 20 December 2020

Abstract: The thermal-based imaging technique has recently attracted the attention of researchers who are interested in the recognition of human affects due to its ability to measure the facial transient temperature, which is correlated with human affects and robustness against illumination changes. Therefore, studies have increasingly used the thermal imaging as a potential and supplemental solution to overcome the challenges of visual (RGB) imaging, such as the variation of light conditions and revealing original human affect. Moreover, the thermal-based imaging has shown promising results in the detection of psychophysiological signals, such as pulse rate and respiration rate in a contactless and noninvasive way. This paper presents a brief review on human affects and focuses on the advantages and challenges of the thermal imaging technique. In addition, this paper discusses the stages of thermal-based human affective state recognition, such as dataset type, preprocessing stage, region of interest (ROI), feature descriptors, and classification approaches with a brief performance analysis based on a number of works in the literature. This analysis could help beginners in the thermal imaging and affective recognition domain to explore numerous approaches used by researchers to construct an affective state system based on thermal imaging.

Keywords: Thermal-based imaging; affective state recognition; spontaneous emotion; feature extraction and classification

1 Introduction

For a long time, human affects (emotion, facial expression) have acquired a great attention by researchers. For instance, in the 17th century, John Buwler wrote his book, "Pathomyotomia or Dissection of the Significant Muscles of the Affections of the Mind." He demonstrates the mechanism of muscles on facial expressions [1]. Charles Darwin [2] illustrates the expressions of human and animals by claiming that emotions are innate. Duchenne de Boulogne [3] describes how muscles produce facial expressions. He published facial expression pictures obtained from the electrical stimulation of human muscles. Ekman and Friesen [4] proposed a prototype comprises of six basic human emotions that encompass different cultures. Russel's [5] mapped the categories of emotion into a valance and arousal dimensional model.

Recently, the recognition of human affective state becomes a crucial aspect in different domains and applications due to the sophistication of smart technologies, and growing of the power of processing units

and artificial intelligent techniques. For example, the understanding of humans' current state not only facilitates human to human communication, but can also increase human computer interaction (HCI). Another example is the assessments of patients' needs who are unable to express their emotions, such as people who have autism disorder [6]. In human robot interaction (HRI) [7], such a technique allows robots to understand and distinguish between negative and positive emotions [8]. In security applications, the technique helps identify poker faces (people who have the ability to hide their emotions) [9]. Additionally, it is also useful in stress detection applications to avoid bad consequences, such as cardiovascular diseases, cancer disease [10,11], and suicide phenomenon [12].

Psychological science mainly describes human affects in three main modalities; the categorical model, dimensional model, and appraisal model. In terms of categorical model, human affects can be described as a number of distinctive affective states called basic emotions and various theories have supported this type of emotion description. For example, Mowrer [13] claims that only two unlearned basic emotions exist, which are pleasure and pain. Frijda [14] introduced six basic facial expressions that can be formed by a reading process. Ekman [15] introduced six basic emotions, which can be recognized universally. Conversely, a dimensional model describes human affects in a dimensional way. For example, the study of Russell [5] demonstrates human affects in a 2D-space, which are valance (miserable to gland) and arousal (sleepy to aroused). The third category of human affect is represented as an appraisal-based emotion. In fact, the appraisal approach is a psychological theory, which categorizes human affects by the evaluation (appraisal) of the events, which is a specific reaction. Therefore, emotion could be elicited by appraisal (evaluation) of events [16]. For example, Orton and Clore [17] explain appraisal theories and how cognition shapes influence human emotions.

Human affects are influenced by several factors, and it can be revealed through a set of human channels, such as voice, gesture, pose, gaze direction, and facial expressions [18]. Psychology sciences classify human communication into verbal and nonverbal categories, of which the nonverbal communication comprises signals, such as kinesics, proxemics, haptics, physical appearance, and paralanguage. Kinesics behavior represents gestures, motion, eye gaze, posture, and facial expressions, while paralanguage represents the voice, volume, tone, pitch, and rate of a speaker [19–21]. Facial expressions influence the spoken message about 55%, while the influence of spoken words is only 7% [22]. Facial expressions consist of primitive and nonverbal language, which represent the basic unit of meaning, and it is identical to words in spoken languages [23]. On the other hand, emotions have a significant role in human affective states and represent the stimulations process to conscious or unconscious awareness for events or objects attached to a mental state [24]. Due to the important role of face to convey human affects, Ekman and Friesen [25] introduced an anatomical tool to describe the movement of facial muscles during emotion called facial action coding system (FACS) and the muscles movement is described as an action unit (AU).

In addition to psychological behaviors, past studies have also focused on physiological signals in studying human affects. In fact, the nervous system in human comprises two main parts, which are central nervous system (CNS) and peripheral nervous system (PNS). The PNS also consists of autonomic nervous system [26] and somatic nervous system (SNS) [27]. Considering ANS, it can regulate numerous physiological processes, such as heart rate, breathing rate, digestion, blood pressure, body temperature, and metabolism. Consequently, human behaviors and affects can be reflected from physiological signals [27,28]. The following physiological sensors have been employed by researchers, such as electromyograph (EMG), electrocardiogram (ECG), electroencephalogram (EEG), Galvanic skin response (GSR), and photoplethysmogram (PPG).

In most recent decades, the literature seems to reflect the concern about constructing various types of automatic human affective state recognition modalities. These modalities can be categorized based on the type input source of the model. For example, a large number of studies have focused on visual-based

(RGB imaging) facial emotion recognition, while other studies selected thermal-based (infra-red thermal imaging) facial emotion recognition. Several human affective state recognitions based on voice, gestures, and physiological signals have been obtained. Several reviews and survey studies have been conducted in recent years. For example, the multimodal affective state recognition can be referred to [29–35]. In the physiological-based domain, readers can refer to [27,36–40], for a visual-based facial affective state recognition [41,42], and for thermal-based facial affective state recognition, readers can refer to [43–47]. This study focuses on the thermal-based affective state recognition. The goal of this study was as follows: First, to introduce the important role of thermal imaging in the affective state recognition. Second, to review databases that are used in the thermal affective state recognition and its challenges. Third, to discuss the general stages of thermal-based affective state recognition, such as preprocessing, region of interest (ROI), feature extraction, and classification. The final goal of this survey was to propose a brief performance analysis based on numerous studies to explore the performance of feature descriptors, classification algorithms, and databases, which were used by past studies. The organization of this paper is as follows: Section 2 discusses the thermal-based affective state recognition, while Section 2.1 reviews the types of thermal databases and its challenges. Section 2.2 presents the preprocessing stage based on the literature's perspective, and Section 2.3 discusses the facial region of interest (ROIs) and its correlation with human affects. Section 2.4 discusses numerous feature extraction methods with brief performance analyses. Section 2.5 reviews classification approaches with a brief performance analysis. Finally, Section 3 concludes this study.

2 Thermal-Based Affective State Recognition

Recently, studies have increasingly adopted the thermal imaging in the affective state recognition due to several factors. In the past, the technology of thermal camera was not feasible because of its low resolution, high cost, heavy weight, and the need to control environment for a stable ambient temperature [48]. The advancement of thermal devices produced new categories of portable and flexible thermal sensors in lightweight, low cost price, and with high resolutions, like mobile thermal sensors. Therefore, the sophisticated thermal sensors motivate researchers to explore thermal imaging inside laboratories and in real world environments for several application, such as human stress recognition [49], and physiological monitoring, such as respiration rate and heart rate [50]. Moreover, the Covid-19 pandemic has increased the popularity of thermal sensors to detect people's temperatures in a contactless and noninvasive way. Another factor that reveals the importance of thermal imaging is the nature of human affect. From the psychophysiological perspective, ANS is responsible for regulating human physiological signals, such as heart rate, respiration rate, blood perfusion, and body temperature during human affect. Hence, thermal imaging becomes a potential solution to measure the facial transient temperature [51]. Past studies have also used thermal imaging to detect another vital signals, which are correlated with human affect, such as respiration rate and pulse rate and it could overcome the challenges of contact-based and invasive physiological sensors [52–57]. When comparing between thermal-based imaging with visual-based (RGB) imaging, studies have reported several advantages of thermal images over visual. For example, visual-based methods reported challenges to distinguish between the original and fake human affective state, especially for those people who are skillful in disguising their emotions [58], visual-based systems are sensitive to an illumination change [59–61], the variety of human skin color, facial shape, texture, ethical backgrounds, cultural differences, and eyes could influence the accuracy of the affective state recognition rate [59,62]. Visual-based recognition approaches have used posed databases due to the lack and difficulties in constructing a spontaneous database, which means that it reduces the efficiency of recognition of the spontaneous affective state [63–65]. Moreover, visual-based imagining approaches provide an inaccurate recognition accuracy in uncontrolled environments [66]. On the other hand, thermal-based imaging is robust against light conditions, which means that it can be used even in full

dark environments. Thermal-based imaging can capture a temperature variation influenced by human affects, therefore, it can be adopted to differentiate between spontaneous and fake emotions. For example, studies have shown that the variation of temperature between the left and right side of the face and a temperature change in periorbital and nasal facial regions are correlated to human affective state [67], also, the variation of blood flow in periorbital region provides an opportunity to measure the instantaneous and sustained stress conditions [68]. Past studies have also reported challenges obtained from thermal imaging, such as facial occlusion of glass opacity in individuals who wear eyeglasses. Consequently, the occlusion creates a challenge to read facial heat pattern. Head pose is another challenge that could influence recognition accuracy in thermal-based affective state recognition. Several studies proposed a solution to solve occlusion and head pose challenges, such as in the following studies [58,69–72]. Past studies have also used the thermal-based imaging as a complementary tool to overcome challenges of visual-based imaging and to enhance human affective state's recognition accuracy by the fusing of features from both types of images, such as [69,73–77]. The thermal-based affective state recognition comprises numerous stages, such as dataset selection, preprocessing, region of interest (ROIs) selection, feature extraction (descriptors), and classification process. The section that follows reviews the abovementioned stages of affective state recognition and presents a brief performance analysis based on a number of works in the literature as demonstrated in Tab. 1

Table 1: Literature work of thermal-based affective state recognition, R/A is recognition accuracy

Author	Dataset	ROI	Emotion	Descriptors	Classifier	R/A
[58]	KTFE	● Forehead ● Eyes ● Cheeks ● Nose ● Mouth	● Anger ● Fear ● Happy ● Sad ● Normal	● Moment invariants ● Histogram statistics	● Multiclass-SVM	87.5
[60]	KTFE	● Refer to paper	● Angry ● Disgust ● Fear ● Happy ● Sad ● Surprise ● Neutral	● Statistical	● PCA ● EMC ● PCA-EMC	90.42 89.99 85.94
[78]	Their own dataset	● Whole face	● Angry ● Disgust ● Fear ● Happy ● Sad ● Surprise	● HOG	● modified Hausdorff distance (MHD)	66.8
[79]	Their own dataset	● Supraorbital ● Periorbital ● Nose ● Mouth	● Angry ● Disgust ● Surprise	● GLCM	● SVM :Fine Gaussian ● Tree Complex ● Weighted KNN	88.5 61.4 99.1

Table 1 (continued).

Author	Dataset	ROI	Emotion	Descriptors	Classifier	R/A
[80]	Dataset from [80]	• Whole face	• Neutral • Happiness • Sadness • surprise	• DSIFT	• SVM	75.4
[81]	Their own dataset	• Forehead • Periorbital • Lower cheeks • Below bottom lip	• High Stress • Neutral • Low stress	• Statistical	• LDA	89.3
[82]	USTC-NVIE	• Facial AUs	• Angry • Disgust • Fear • Happy • Sad • Surprise	• FCTH	• KNN	34
[83]	Their own dataset	• Face • Neck • shoulders	• Affective vs Neutral • Positive Vs Negative Vs neutral • Positive Vs Negative	• Statistical	• Ada boost	90 54 50
[84]	USTC-NVIE	• Forehead • Left cheek • Right cheek	• Embarrassment • Surprise • Happiness • Sadness • Disgust • Anger or Upset • Pain • Fear or Nervous	• Histogram of SIFT flow bag of words • Facial temperature changes	• SVM	91
[85]	NVIE	• Forehead • Eyes • Mouth	• Low Valance and High Valance	• Thermal statistic features • 2D-DCT • GLCM • DBM features	• DBM • SVM	62.9
[86]	Their own dataset	• Forehead • Periorbital • Nose • Mouth	• Happy • Anger	• PCA	• SVM	77.5
[87]	Their own dataset	• Supraorbital • Periorbital • Maxillary /Nose • Mouth	• Angry • Disgust • Fear • Happy • Sad • Surprise	• 2DWT • GLCM • LBP	• SVM	98.6

(Continued)

Table 1 (continued).

Author	Dataset	ROI	Emotion	Descriptors	Classifier	R/A
[88]	KTFE	• Whole face	• Angry	• Statistical	• PCA	73.96
			• Happy		• EMC	
			• Fear		• PCA-EMC	
			• Sad			66.45
			• Neutral			70.34
[89]	Their own dataset	• Left and right forehead	• Disgust	• statistical	• LDA	85
		• Left and right periorbital	• Fear			
		• Tip of nose	• Happiness			
		• Cheeks	• Sadness			
		• Left and right perinasal	• Surprise			
		• Left and right chin				
[90]	Their own dataset	• Facial thermal feature points on facial muscle map (forehead, Around eyes, cheeks, around mouth, chin)	• Neutral	• PCA	• LDA	83.3
			• intentional happy and sad	• Statistical features		
			• Neutral			67.2
			• intentional happy, disgust, surprise, angry			
			• Neutral			57.1
			• Six pretended basic emotions			
			• Neutral			72
			• Evoked happy, sad, disgust, angry			

2.1 Thermal Dataset

Thermal databases can be divided based on the affective type into three main categories, such as posed, induced, and spontaneous databases [91]. In posed databases, in past studies, participants were asked to act different types of emotions to collect posed affective state. Therefore, this type of emotion does not reflect true human affects. To elicit human affects, the participants were exposed to stimuli to induce their affects, while in the spontaneous type, the participants were not aware of the video recording of their acts that captured their images. Hence, the process of the construction of spontaneous databases is very difficult [63]. More importantly, thermal-based affective state recognition systems require databases that comprise different factors, which play an important role on recognition accuracy, such as aligned and cropped faces, illumination variation, head poses, and occlusion [92]. Due to the importance of thermal imaging in recognition of human affective state, several methods have been proposed to construct thermal-based databases. From the study of Wang et al. [92], a natural visible and infrared facial expression (NVIE) database was constructed. This database comprises both thermal and visible images acquired simultaneously from 215 participants. Also, the database was manually annotated and contained landmarks for spontaneous and posed expressions. In addition, facial images with and without eyeglasses were captured under different illumination conditions. For validation and assessment purposes, the authors have applied PCA, LDA, and AAM algorithms for visible expression recognition and PCA, and LDA for infrared expression recognition from the results of the thermal classification algorithms and analysis of correlation between facial temperature and emotions. The authors reported that NVIE database provides suitable features for the thermal affective state recognition process.

Kopaczka et al. [70] constructed a thermal database, which includes a number of features that facilitate the affective state recognition in thermal imaging as follows: Firstly, the database comprises spontaneous and posed expressions, and secondly, for occlusion challenge the database encompasses faces with and without spectacles to solve occlusion challenges, then, the authors captured faces with different positions to allow algorithms to deal with the head movement problems. In addition, the database consists of 68 manually annotated facial landmarks and numerous AUs for thermal recognition. Finally, the authors evaluated their database by applying the different types of features and classification algorithms, such as HOG, LBP, SIFT, SVM, KNN, BDT, LDA, NB, and RF classifiers. Their study reported the ability of the database to train machine learning algorithms. Nguyen et al. [93] constructed the KTFE facial expression database. This database focuses on natural and spontaneous expressions in thermal and visible videos. The authors claim that their database solves a number of problems found in other facial expression databases, such as time interval between stimuli and participants' response. More importantly, thermal databases are very few compared to visual databases [91], and numerous challenges could influence the quality of data. For example, the number of annotated facial images is small, which limits the use of machine learning algorithms and the type of expressions in these databases are posed expressions [92,93]. Consequently, the rarity of thermal databases could limit the validation process and comparison between several thermal based affective state methods.

Due to the important role of spontaneous affective state, this paper conducted a brief analysis to show the percentage of spontaneous database compared to posed databases. The analysis relied on the study of Ordun et al. [94] as this study collected 17 thermal emotion databases. The percentage of abovementioned dataset types in their collected databases is as follows: Spontaneous database 17.6 %, posed database 58.8 %, spontaneous with posed database 11.8 %, and unknown type of databases is 11.8 %. This analysis may assert the challenge of collecting spontaneous human affects. As listed in Tab. 1, 50 % of the studies conducted their own experiment (handcraft) to collect induced affective state dataset rather than posed affects. The percentage of studies that used KTFE and UTC-NVIE dataset is 18.75 %. From this analysis, it can be concluded that past studies tended to use their own handcrafted dataset more than the public dataset and the reason could be linked to the rarity of spontaneous and induced databases and not all databases are available for public. For example, the percentage of databases, which require a permission from the owners to use and the ones listed in Ordun et al.'s [94] study is 52.9 %, while 11.7 % of databases is publicly available and 35.2 % of the listed databases have no information about publicity of use. Another important issue that could be helpful in selecting an appropriate database is by analyzing the performance of recognition affective state with respect to the type of dataset. Therefore, Fig. 1 demonstrates the average recognition accuracy results with respect to the dataset type. From Fig. 1, it can be noticed that the average recognition accuracy results obtained by using KTFE, dataset developed based on the literature (handcraft), Ref. [80] database, and UTC-NVIE database are 80.65%, 75.98%, 75.40%, and 62.63%, respectively.

2.2 Preprocessing Stage

Thermal-based images are different in structure than visual-based images due to its thermal temperature. Therefore, several characteristics of geometric, appearance, and texture of thermal images need different preprocessing methods for image enhancement, noise reduction, and for facial extraction. In the pre-processing stage, past studies have used several methods to enhance thermal images and to extract facial region. For example, Kopaczka et al. [70] applied HOG with SVM for face detection. Merhof et al. [71] implemented unsharped mask for image enhancement. The study used lowpass gaussian filter and subtracted the filtered image from the original image. Liu and Yin [84] proposed a unified model, which comprises a mixture of trees with shard pools of parts for face detection taken from [95] to allocate a facial region for the first frame, then, they used the first frame to calculate the head motion. In

Wang et al. [85], they applied the Otsu thresholding algorithm to generate a binary image. Then, they calculated vertical and horizontal curves from the binary image and detected a face boundary from the largest gradient of projection curve. To improve the image contrast, Latif et al. [87] implemented the contrast limited histogram equalization (CLHE) on thermal images. Mohd et al. [59] applied the Viola-Jones boosting algorithm with Haar-like features to detect a facial region. The study applied a bilateral filter for noise reduction and facial edge preservation. Wan et al. [88] proposed a temperature space method to distinguish facial region from image background. Trujillo et al. [96] applied the bi-modal thresholding method to locate facial boundaries. Kolli et al. [78] proposed a color-based detection, region growing with morphological operation for face detection. Goulart et al. [89] extracted a facial region form thermal image by using the median and gaussian filters with a binary filter to convert the facial region into a pure black and white. Khan et al. [90] applied the median smoothing filter for blurring and noise reduction and sobel filter for edge detection. Cruz-Albarran et al. [97] proposed a thresholding value to extract a facial region from an image background. Mostafa et al. [72] proposed the tracking ROI method, of which their method is composed of an adaptive particle filter tracker to track the facial ROI. From these studies, it can be noticed that different methods have been proposed for thermal image enhancement and for facial extraction.
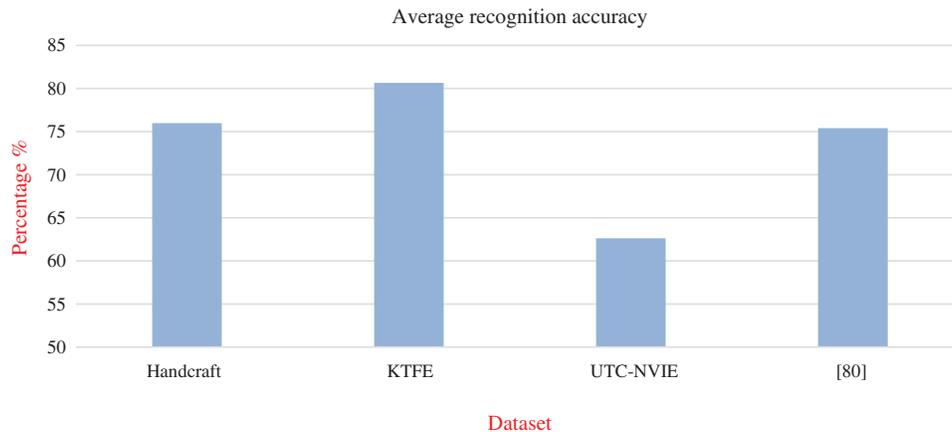
Figure 1: Average recognition accuracy results with respect to database type

2.3 Region of Interest (ROI)

A facial temperature is influenced by human affects as an autonomic nervous system (ANS) responses to human affects, like emotions or mental trigger physiological changes, such as an increasing blood flow inside the vascular system. Hence, the human body temperature is regulated through the propagation of internal body heat to human skin. The variation of facial temperatures could be measured by using the IR-thermal [98]. Another factor that influences facial temperature is the facial muscles contraction, which also produces heat that propagates to facial skin [99]. However, the literature has mainly focused on numerous facial areas to measure human affects, such as forehead, eyes, mouth, periorbital, tip of nose, maxillary, cheeks, and chin [60,81,82,84,87,100]. Several studies have observed the variation of temperature in the facial ROI with respect to human emotion type. For example, the study of Ioannou et al. [47] reported that the temperature of the nose, forehead, and maxillary reduces in stress and fear affective state, and the periorbital temperature increases in the anxiety affective state. Their study demonstrates the temperature variation in facial ROI as shown in Tab. 2. Moreover, Jian et al. [67] have shown that cheeks and eyes have positively correlated with human emotion. Furthermore, Cruz-Albarran et al. [97] reported that the temperature of the nose and maxillary is reduced in fear, joy, anger, disgust, and sadness emotions, while the cheeks' temperature increases in disgust and sadness affects. Also, Ioannou et al. [47] reported the

decrease in the forehead's temperature when in fear and sadness, and increases in an anger affective state. Tab. 2 demonstrates the temperature changes in facial ROIs within human affects, taken from [47].

Table 2: Temperature variation of facial ROIs during human affective state, '+' temperature increased, '-' temperature decreased, 'NA' not available, this table adapted from [47]

| Facial ROI	Stress	Fear	Startle	Sexual Arousal	Anxiety	Pain	Guilt
Forehead	+/−	−	NA	+	+	−	NA
Periorbital	NA	NA	+	+	+	NA	NA
Supraorbital	NA	NA	+	NA	+	NA	NA
Cheeks	NA	NA	−	NA	NA	NA	NA
Nose	−	NA	NA	NA	NA	NA	NA
Maxillary	−	−	−	NA	NA	−	−
Lips/mouth	NA	NA	NA	+	NA	NA	NA

2.4 Feature Extraction (Descriptors)

Features play an important role in the recognition human affects, especially in thermal images, which have a different texture, appearance, and shape than the RGB images. Past studies have used several types of thermal descriptors in their works. The performance of affective state recognition relies on the efficiency of facial features from either the whole face or facial parts (ROI). Tab. 1 in this survey demonstrates numerous types of feature extraction. From Tab. 1, it can be noticed that more than 50% of the studies have used statistical features. For example, Cross et al. [81] selected the mean value of temperature points, while Pérez-Rosas et al. [83] used an average temperature, overall minimum and maximum temperature, standard deviation, and standard deviation of minimum and maximum temperature. Nguyen et al. [69,88] applied the covariance matrix over facial temperature points. Liu and Yin [84] applied the histogram of SIFT features with the variation of facial temperature. Wang et al. [85] selected a number of statistical features in addition to 2D-DCT, GLCM, and DBM features. Basu et al. [58] applied the moment invariant with histogram statistic features, while Goulart et al. [89] applied numerous statistical features, such as mean emissivity, variance of emissivity, mean value of rows and columns, median of rows and columns, and another seven statistical features. Similarly, the study of Khan et al. [90] applied a number of statistical features, such as mean of thermal intensity values, covariance, and eigenvectors. Another important feature has also been paid attention to by researchers that is the gray level co-occurrence matrix (GLCM) [79,85,87]. GLCM is used to describe the texture information in thermal images and co-occurrence distributions, which represent the distance and angles of spatial relationship over the image. GLCM composes multiple directions during the analysis, such as vertical direction (0°), vertical direction (90°), and diagonal direction (45°) and (135°). Histogram of oriented gradient (HOG) is another facial descriptor used in thermal-based affective state recognition [78]. The HOG descriptor counts the occurrence of gradient orientation from a local region in an image, by dividing the image into a number of sub-regions called cells and computes the histogram of gradient direction for each cell. Local Binary Pattern (LBP) is applied in thermal-based affective state recognition [87]. This descriptor thresholds the neighborhood of pixels image by using 3×3 filter where the center of the filter is the pixel image, while the others represent pixel neighborhoods, and the threshold value is represented in the form of binary numbers. LBP operators have been updated by extending the pixel neighborhoods as a circular neighborhood to utilize the features from different scales [101]. Different kinds of LBP have been established, for example, LBP for spatial domain, LBP for spatiotemporal

domain, and LBP for face description. Scale-invariant feature transform (SIFT) and Dense SIFT have also been selected as facial descriptors in thermal-based affective state recognition [84,80]. The general idea of the SIFT descriptor is to describe the local features in an image based on Euclidian distance. Principle Component Analysis (PCA) is a statistical approach selected by past studies for dimensionality reduction [86,90] by converting a large set of variables to a small set, while preserving the information that exist in the large set. Principal Component Analysis (PCA) is another important descriptor. This descriptor transforms numerous possible correlated variables to small uncorrelated variables (principal components), in a facial expression recognition domain. PCA extracts the facial features from a set of images, which are global facial images and varies from mean face (eigenface). However, another several feature descriptors have also been utilized by past studies. For example, 2D-DCT was used with other descriptors in [85]. 2DWT applied by [87] and Fuzzy Color and texture histogram (FCTH) were used in [82]. This paper analyzed the performance of facial descriptors as follows: The first analysis was the maximum recognition accuracy result reported in Tab. 1, which is 98.6% obtained by combining 2DWT, GLCM, and LBP features as reported in [87]. The next analysis represents the maximum feature selected by past studies. As shown in Tab. 1, about 50% of past studies used statistical features and reported an overall average recognition accuracy result of 75.91%. The reason of selecting statistical features could be due to the nature of thermal imaging, which motivated past studies to measure the variation of facial skin temperature rather than the RGB images, which have a better facial texture and shape.

2.5 Classification

Robust classification method is crucial in thermal-based affective state recognition in order to maintain a high accuracy of affective prediction. Consequently, several classifiers have been used for this purpose. As demonstrated in Tab. 1, it can be noticed that the SVM classifier has been used widely in past studies [58,61,79,80,84–87]. The second classifier focused by past studies is LDA [81,89]. In addition, KNN is another important classifier [82,102]. PCA, EMC, and PCA-EMC have been used by [60,88,90]. AdaBoost was also selected by [83]. Deep Boltzmann machine (DBN) was used in [85]. The Modified Hausdorff distance (MHD) was applied in [78]. Tree complex classifier was implemented in [79]. This paper analyzed the performance of numerous classification algorithms. The maximum accuracy results is 99.1% reported in [79]. The classifier used in this study was weighted KNN. Moreover, the average accuracy result reported from all literatures is 76.69%. Fig. 2 demonstrates the analysis of classifier recurrent with its average recognition accuracy. From this figure it can be noticed that the classifier that was more frequently used in past studies is SVM with 28% of recurrent and reported higher average recognition accuracy results, which is 83.05%.
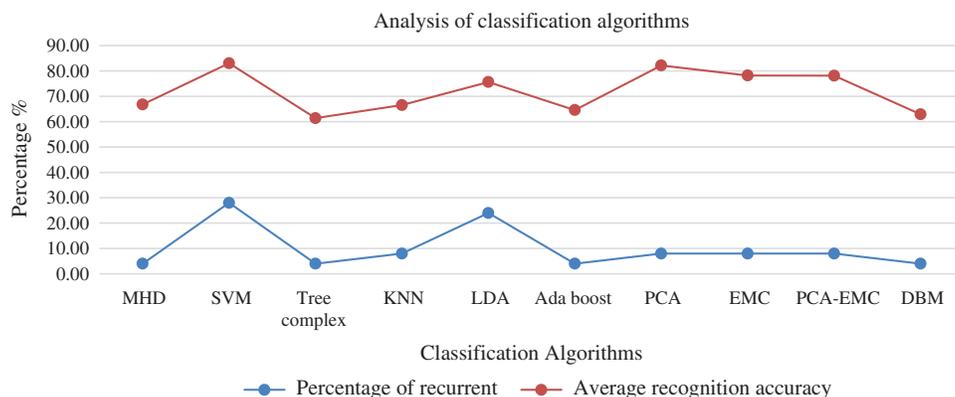
Figure 2: Classification algorithm analysis, percentage of recurrent and average recognition accuracy

3 Conclusion

This paper discusses the importance of thermal-based imaging in affective state recognition. Thermal images have important characteristics that could help researchers to overcome well-known challenges of other techniques like visual-based affective state recognition and physiological-based affective state recognition. For example, thermal images are robust against illumination changes and it captures the facial transient temperature, which reflects human inner affects. Moreover, past studies have used thermal images to measure important vital signals, such as pulse rate and respiration rate in a contactless and noninvasive way. Therefore, researchers have recently paid their attention to the thermal-based imaging technique as a potential and supplemental solution to the challenges of affective state recognition. This paper also discusses general stages of affective state recognition based on thermal imaging and conducted a brief performance analysis. Based on the analysis of this paper, 50% of past studies have conducted their own dataset and the reason could be due to the rarity of spontaneous and induced databases, and 52.9% of the public databases require the permission to use. This paper also reported that 50% of past studies applied statistical features with an overall recognition accuracy of 75.91%. In the classification stage, the analysis of this paper shows that SVM is the classifier that have been more frequently used by past studies with 28% recurrent of use and an overall average recognition accuracy of 83.05%.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report the findings of the present study.

References

[1]   S. Greenblatt, "Toward a universal language of motion: Reflections on a seventeenth-century muscle man," *LiNQ (Literature in North Queensland)*, vol. 21, no. 2, pp. 56–62, 1994.

[2]   C. Darwin, "Introduction to The First Edition," in *The Expression of Emotions in Man and Animals*, 3rd edition, London, England: John Murray, pp. 7, 1872.

[3]   G. B. Duchenne, G. B. D. de Boulogne , R. A. Cuthbertson, A. S. R. Manstead and K. Oatley, "A review of previous work on muscle action in facial expression," in *The Mechanism of Human Facial Expression*. New York, USA: Cambridge University Press, pp. 14–16, 1990, https://books.google.com.sa/books?id=a9tjQC7xbNMC.

[4]   P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, no. 2, pp. 124–129, 1971.

[5]   J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161, 1980.

[6]   C.H. Chen, I.J. Lee and L.Y. Lin, "Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders," *Research in Developmental Disabilities*, vol. 36, pp. 396–403, 2015.

[7]   E. Fendri, R. R. Boukhriss and M. Hammami, "Fusion of thermal infrared and visible spectra for robust moving object detection," *Pattern Analysis and Applications*, vol. 20, no. 4, pp. 907–926, 2017.

[8]   L. Desideri, C. Ottaviani, M. Malavasi, R. di Marzio and P. Bonifacci, "Emotional processes in human-robot interaction during brief cognitive testing," *Computers in Human Behavior*, vol. 90, no. 1, pp. 331–342, 2019.

[9]   M. S. Lee, Y. R. Cho, Y. K. Lee, D. S. Pae, M. T. Lim *et al.,* "PPG and EMG based emotion recognition using convolutional neural network," in *ICINCO 2019—Proc. of the 16th Int. Conf. on Informatics in Control, Automation and Robotics*, Prague, Czech Republic, pp. 595–600, 2019.

[10] G. E. Miller, S. Cohen and A. K. Ritchey, "Chronic psychological stress and the regulation of pro-inflammatory cytokines: A glucocorticoid-resistance model," *Health Psychology*, vol. 21, no. 6, pp. 531–541, 2002.

[11] L. Vitetta, B. Anton, F. Cortizo and A. Sali, "Mind-body medicine: Stress and its impact on overall health and longevity," *Annals of the New York Academy of Sciences*, vol. 1057, no. 1, pp. 492–505, 2005.

[12] J. L. Gradus, "Posttraumatic stress disorder and death from suicide," *Current Psychiatry Reports*, vol. 20, no. 11, pp. S7, 2018.

[13] O. H. Mowrer, "Introduction: Historical review and perespective," in *Learning Theory and Behavior*. Hoboken, NJ: John Wiley & Sons Inc, pp. 9, 1960.

[14] N. H. Frijda, "Emotional behavior," in *The Emotions*. Cambridge, UK: Cambridge University Press, pp. 9–109, 1986, https://books.google.com.sa/books?id=QkNuuVf-pBMC.

[15] P. Ekman, "Universals and cultural differences in facial expressions of emotion," in *Nebraska Sym. on Motivation*, J. Cole, (ed.) 19. Lincoln, NE: Univercity of Nebraska Press, pp. 209–282, 1971.

[16] K. R. Scherer, A. Schorr and T. Johnstone, "Appraisal theory," in *Appraisal Processing in Emotion*. Madison Avenue, New York, USA: Oxford University Press, 2001, https://books.google.com.sa/books?id=fmtnDAAAQBAJ,

[17] A. Ortony and G. Clore, "Appraisal theories: How cognition shapes affect into emotion," in *Handbook of Emotions*, M. Lewis, J. M. Haviland-Jones, L. F. Barrett, (ed.) 3rd edition, New York, USA: Guilford Press, pp. 628–642, 2008.

[18] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259–275, 2003.

[19] J. A. Devito, "Silence and paralanguage as communication," *ETC: A Review of General Semantics*, vol. 74, no. 3–4, pp. 482–487, 2017.

[20] M. L. Knapp, J. A. Hall and T. G. Horgan, "Nonverbal communication: Basic perspectives," in *Nonverbal Communication in Human Interaction*, 8th edition, Boston, USA: Wadsworth Cengage Learning, pp. 3–19, 2013.

[21] S. Johar, "Emotion, affect and personality in speech: The bias of language and paralanguage." New York, USA: Springer International Publishing, 1–6, 2016. [Online]. Available: http://gen.lib.rus.ec/book/index.php?md5=79cb2f8fcb7c48eaa270d6b1ab364612.

[22] N. N. Khatri, Z. H. Shah and S. A. Patel, "Facial expression recognition: A survey," *Int. Journal of Computer Science and Information Technologies (IJCSIT)*, vol. 5, no. 1, pp. 149–152, 2014.

[23] C. Smith and H. Scott, "A Componential approach to the meaning of facial expressions," in *The Psychology of Facial Expression (Studies in Emotion and Social Interaction)*, G. Mandler, J. Russell and J. Fernández-Dols, (ed.) Cambridge, UK: Cambridge University Press, pp. 229–254, 1997.

[24] J. F. Kihlstrom, S. Mulvaney, B. A. Tobias and I. P. Tobis, "The emotional unconscious," in *Cognition and Emotion*, J. F. Kihlstrom, G. H. Bower, J. P. Forgas and P. M. Niedenthal, (eds.) New York, USA: Oxford University Press, pp. 30–86, 2000, [Online]. Available: https://psycnet.apa.org/record/2001-00519-002.

[25] E. Friesen and P. Ekman, "Facial action coding system: A technique for the measurement of facial movement," *Consulting Psychologists Press* vol. 3, Palo Alto, 1978.

[26] B. Khalid, A. M. Khan, M. U. Akram and S. Batool, " Person detection by fusion of visible and thermal images using convolutional neural network," in *Int. Conf. on Communication, Computing and Digital Systems (C-CODE)*. Islamabad, Pakistan, 143–148, 2019.

[27] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li *et al.,* "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, no. 7, pp. 2074, 2018.

[28] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological Psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[29] N. Samadiani, G. Huang, B. Cai, W. Luo, C.H. Chi *et al.,* "A review on automatic facial expression recognition systems assisted by multimodal sensor data," *Sensors (Basel)*, vol. 19, no. 8, pp. 1863, 2019.

[30] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Computer Vision and Image Understanding*, vol. 108, no. 1–2, pp. 116–134, 2007.

[31] D. Mehta, M. F. H. Siddiqui and A. Y. Javaid, "Facial emotion recognition: A survey and real-world user experiences in mixed reality," *Sensors*, vol. 18, no. 2, pp. 416, 2018.

[32] H. Gunes, B. Schuller, M. Pantic and R. Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey," in *International Conference on Automatic Face and Gesture Recognition*, Santa Barbara, CA, USA, pp. 827–834, 2011.

[33] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee et al., "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proc. of the 6th Int. Conf. on Multimodal Interfaces (ICMI04)*, Pennsylvania, USA: State College, pp. 205–211, 2004.

[34] C. A. Corneanu, M. O. Simón, J. F. Cohn and S. E. Guerrero, "Survey on RGB, 3D, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 8, pp. 1548–1568, 2016.

[35] N. J. Shoumy, L. M. Ang, K. P. Seng, D. M. Rahaman and T. Zia, "Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals," *Journal of Network and Computer Applications*, vol. 149, no. 1, pp. 102447, 2020.

[36] A. Chunawale and D. Bedekar, "Human emotion recognition using physiological signals: A survey," in *2nd Int. Conf. on Communication and Information Processing (ICCIP-2020)*. Tokyo, Japan, p. 9, 2020.

[37] S. Saganowski, A. Dutkowiak, A. Dziadek, M. Dzieżyc, J. Komoszyńska et al., "Emotion recognition using wearables: A systematic literature review-work-in-progress," in *IEEE Int. Conf. on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 1–6, 2020.

[38] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Computer Methods and Programs in Biomedicine*, vol. 161, no. 1, pp. 1–13, 2018.

[39] S. Z. Bong, M. Murugappan and S. Yaacob, "Methods and approaches on inferring human emotional stress changes through physiological signals: A review," *Int. Journal of Medical Engineering and Informatics*, vol. 5, no. 2, pp. 152–162, 2013.

[40] S. Jerritta, M. Murugappan, R. Nagarajan and K. Wan, "Physiological signals based human emotion recognition: A review," in *IEEE 7th Int. Colloquium on Signal Processing and its Applications*, pp. 410–415, 2011.

[41] B. C. Ko, "A brief review of facial emotion recognition based on visual information," *Sensors*, vol. 18, no. 2, pp. 401, 2018.

[42] I. M. Revina and W. S. Emmanuel, "A survey on human face expression recognition techniques," *Journal of King Saud University-Computer and Information Sciences*, pp. 1–33, 2018.

[43] C. Filippini, D. Perpetuini, D. Cardone, A. M. Chiarelli and A. Merla, "Thermal infrared imaging-based affective computing and its application to facilitate human robot interaction: A review," *Applied Sciences*, vol. 10, no. 8, pp. 2924, 2020.

[44] B. Yang, X. Li, Y. Hou, A. Meier, X. Cheng et al., "Non-invasive (non-contact) measurements of human thermal physiology signals and thermal comfort/discomfort poses-a review," *Energy and Buildings*, vol. 224, no. 1, pp. 110261, 2020.

[45] C. N. Kumar and G. Shivakumar, "A survey on human emotion analysis using thermal imaging and physiological variables," *Int. Journal of Current Engineering and Scientific Research (IJCESR)*, vol. 4, no. 4, pp. 122–126, 2017.

[46] G. Ponsi, M. S. Panasiti, G. Rizza and S. M. Aglioti, "Thermal facial reactivity patterns predict social categorization bias triggered by unconscious and conscious emotional stimuli," *Proc. of the Royal Society B: Biological Sciences*, vol. 284, no. 1861, pp. 20170908, 2017.

[47] S. Ioannou, V. Gallese and A. Merla, "Thermal infrared imaging in psychophysiology: Potentialities and limits," *Psychophysiology*, vol. 51, no. 10, pp. 951–963, 2014.

[48] M. van Gastel, S. Stuijk and G. de Haan, "Robust respiration detection from remote photoplethysmography," *Biomedical Optics Express*, vol. 7, no. 12, pp. 4941–4957, 2016.

[49] Y. Cho, "Automated mental stress recognition through mobile thermal imaging," in *Seventh Int. Conf. on Affective Computing and Intelligent Interaction (ACII)*. San Antonio, TX, USA, 596–600, 2017.

[50] X. He, R. Goubran and F. Knoefel, "IR night vision video-based estimation of heart and respiration rates," in *IEEE Sensors Applications Symposium (SAS)*, pp. 1–5, 2017.

[51] A. Merla, L. Di Donato, P. Rossini and G. Romani, "Emotion detection through functional infrared imaging: preliminary results," *Biomedizinische Technik*, vol. 48, no. 2, pp. 284–286, 2004.

[52] A. H. Alkali, R. Saatchi, H. Elphick and D. Burke, "Thermal image processing for real-time non-contact respiration rate monitoring," *IET Circuits, Devices & Systems*, vol. 11, no. 2, pp. 142–148, 2017.

[53] Y. Nakayama, G. Sun, S. Abe and T. Matsui, "Non-contact measurement of respiratory and heart rates using a CMOS camera-equipped infrared camera for prompt infection screening at airport quarantine stations," in *IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. Shenzhen, China, 1–4, 2015.

[54] S. B. Park, G. Kim, H. J. Baek, J. H. Han and J. H. Kim, "Remote pulse rate measurement from near-infrared videos," *IEEE Signal Processing Letters*, vol. 25, no. 8, pp. 1271–1275, 2018.

[55] M. Yang, Q. Liu, T. Turner and Y. Wu, " Vital sign estimation from passive thermal video," in *IEEE Conf. on Computer Vision and Pattern Recognition*. Anchorage, AK, USA, 1–8, 2008.

[56] J. Fei and I. Pavlidis, "Thermistor at a distance: Unobtrusive measurement of breathing," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 988–998, 2009.

[57] J. N. Murthy, J. van Jaarsveld, J. Fei, I. Pavlidis, R. I. Harrykissoon *et al.,* "Thermal infrared imaging: A novel method to monitor airflow during polysomnography," *Sleep*, vol. 32, no. 11, pp. 1521–1527, 2009.

[58] A. Basu, A. Routray, S. Shit and A. K. Deb, "Human emotion recognition from facial thermal image based on fused statistical feature and multi-class SVM," in *Annual IEEE India Conf. (INDICON)*. New Delhi, India, 1–5, 2015.

[59] M. N. H. Mohd, M. Kashima, K. Sato and M. Watanabe, "Mental stress recognition based on non-invasive and non-contact measurement from stereo thermal and visible sensors," *Int. Journal of Affective Engineering*, vol. 14, no. 1, pp. 9–17, 2015.

[60] T. Nguyen, K. Tran and H. Nguyen, "Towards thermal region of interest for human emotion estimation," in *10th Int. Conf. on Knowledge and Systems Engineering (KSE), Ho Chi Minh City.* Vietnam, 152–157, 2018.

[61] S. Wang, B. Pan, H. Chen and Q. Ji, "Thermal augmented expression recognition," *IEEE Transactions on Cybernetics*, vol. 48, no. 7, pp. 2203–2214, 2018.

[62] X. Yan, T. J. Andrews, R. Jenkins and A. W. Young, "Cross-cultural differences and similarities underlying other-race effects for facial identity and expression," *Quarterly Journal of Experimental Psychology*, vol. 69, no. 7, pp. 1247–1254, 2018.

[63] S. Wang, Z. Liu, Z. Wang, G. Wu, P. Shen *et al.,* "Analyses of a multimodal spontaneous facial expression database," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 34–46, 2013.

[64] M. K. Bhowmik, P. Saha, A. Singha, D. Bhattacharjee and P. Dutta, "Enhancement of robustness of face recognition system through reduced Gaussianity in Log-ICA," *Expert Systems with Applications*, vol. 116, no. 1, pp. 96–107, 2019.

[65] X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale *et al.,* "Bp4d-spontaneous: A high-resolution spontaneous 3d dynamic facial expression database," *Image and Vision Computing*, vol. 32, no. 10, pp. 692–706, 2014.

[66] C.-H. Chu and S.-M. Peng, "Implementation of face recognition for screen unlocking on mobile device," in *Proc. of the 23rd ACM Int. Conf. on Multimedia*, Brisbane, Australia, pp. 1027–1030, 2015.

[67] B.-L. Jian, C. L. Chen, M. W. Huang and H. T. Yau, "Emotion-specific facial activation maps based on infrared thermal image sequences," *IEEE Access*, vol. 7, pp. 48046–48052, 2019.

[68] V. Elanthendral, R. Rekha and M. Rameshkumar, "Thermal imaging for facial expression-fatigue detection," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 2, no. XII, pp. 14–17, 2014.

[69] H. Nguyen, F. Chen, K. Kotani and B. Le, "Fusion of visible images and thermal image sequences for automated facial emotion estimation," *Journal of Mobile Multimedia*, vol. 10, no. 3-4, pp. 294–308, 2014.

[70] M. Kopaczka, R. Kolk and D. Merhof, "A fully annotated thermal face database and its application for thermal facial expression recognition," in *IEEE Int. Instrumentation and Measurement Technology Conf. (I2MTC)*. Houston, TX, USA, 1–6, 2018.

[71] D. Merhof, K. Acar and M. Kopaczka, "Robust facial landmark detection and face tracking in thermal infrared images using active appearance models," in *Int. Conf. on Computer Vision Theory and Applications*. Rome, Italy, 150–158, 2016.

[72] E. Mostafa, A. Farag, A. Shalaby, A. Ali, T. Gault *et al.,* "Long term facial parts tracking in thermal imaging for uncooperative emotion recognition," in *IEEE Sixth Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*. Arlington, VA, USA, pp. 1–6, 2013.

[73] S. Wang, S. He, Y. Wu, M. He and Q. Ji, "Fusion of visible and thermal images for facial expression recognition," *Frontiers of Computer Science*, vol. 8, no. 2, pp. 232–242, 2014.

[74] X. Shi, S. Wang and Y. Zhu, "Expression recognition from visible images with the help of thermal images," in *Proc. of the 5th ACM on Int. Conf. on Multimedia Retrieval*, Shanghai, China, pp. 563–566, 2015.

[75] M. F. H. Siddiqui and A. Y. Javaid, "A multimodal facial emotion recognition framework through the fusion of speech with visible and infrared images," *Multimodal Technologies and Interaction*, vol. 4, no. 3, pp. 46, 2020.

[76] S. Wang and S. He, "Spontaneous facial expression recognition by fusing thermal infrared and visible images," in *Intelligent Autonomous Systems 12. Advances in Intelligent Systems and Computing*. vol. 194. Heidelberg, Berlin, Germany: Springer, pp. 263–272, 2013.

[77] S. Nayak, S. K. Panda and S. Uttarkabat, "A non-contact framework based on thermal and visual imaging for classification of affective states during HCI," in *4th Int. Conf. on Trends in Electronics and Informatics (ICOEI)(48184)*. Tirunelveli, India, 653–660, 2020.

[78] A. Kolli, A. Fasih, F. Al Machot and K. Kyamakya, " Non-intrusive car driver's emotion recognition using thermal camera," in *International Workshop on Nonlinear Dynamics and Synchronization, INDS. Proc. of the Joint INDS'11 & ISTET'11*, Klagenfurt, Austria, pp. 1–5, 2011.

[79] M. Latif, S. Sidek, N. Rusli and S. Fatai, "Emotion detection from thermal facial imprint based on GLCM features," *ARPN J. Eng. Appl. Sci*, vol. 11, no. 1, pp. 345–350, 2016.

[80] M. Kopaczka, R. Kolk and D. Merhof, "A fully annotated thermal face database and its application for thermal facial expression recognition," in *IEEE Int. Instrumentation and Measurement Technology Conf. (I2MTC)*. Houston, TX, USA, 1–6, 2018.

[81] C. B. Cross, J. A. Skipper and D. T. Petkie, "Thermal imaging to detect physiological indicators of stress in humans," *Proc. SPIE Thermosense: Thermal Infrared Applications XXXV*, vol. 8705, pp. 87050I–1-8705I-15, 2013.

[82] R. Carrapiço, A. Mourao, J. Magalhaes and S. Cavaco, "A comparison of thermal image descriptors for face analysis," in *23rd European Signal Processing Conf. (EUSIPCO)*. Nice, France, 829–833, 2015.

[83] V. Pérez-Rosas, A. Narvaez, M. Burzo and R. Mihalcea, "Thermal imaging for affect detection," in *PETRA '13: The 6th Int. Conf. on Pervasive Technologies Related to Assistive Environments*. Rhodes, Greece, 1–4, 2013.

[84] P. Liu and L. Yin, "Spontaneous facial expression analysis based on temperature changes and head motions," in *11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG)*. Ljubljana, Slovenia, pp. 1–6, 2015.

[85] S. Wang, M. He, Z. Gao, S. He and Q. Ji, "Emotion recognition from thermal infrared images using deep Boltzmann machine," *Frontiers of Computer Science*, vol. 8, no. 4, pp. 609–618, 2014.

[86] L. Boccanfuso, Q. Wang, I. Leite, B. Li, C. Torres *et al.,* "A thermal emotion classifier for improved human-robot interaction," in *25th IEEE Int. Sym. on Robot and Human Interactive Communication (RO-MAN)*. New York, NY, USA, 718–723, 2016.

[87] M. Latif, M. H. Yusof, S. Sidek and N. Rusli, "Texture descriptors based affective states recognition-frontal face thermal image," in *IEEE EMBS Conf. on Biomedical Engineering and Sciences (IECBES)*. Kuala Lumpur, Malaysia, 80–85, 2016.

[88] H. Nguyen, K. Kotani, F. Chen and B. Le, "Estimation of human emotions using thermal facial information," in *Fifth Int. Conf. on Graphic and Image Processing (ICGIP2013) Proc*. vol. 9069. Hong Kong, China, pp. 90690O-1–90690O-5, 2014.

[89] C. Goulart, C. Valadão, D. Delisle-Rodriguez, E. Caldeira, T. Bastos *et al.,* "Emotion analysis in children through facial emissivity of infrared thermal imaging," *PloS One*, vol. 14, no. 3, pp. e0212928, 2019.

[90] M. M. Khan, R. D. Ward and M. Ingleby, "Classifying pretended and evoked facial expressions of positive and negative affective states using infrared measurement of skin temperature," *ACM Transactions on Applied Perception (TAP)*, vol. 6, no. 1, pp. 1–22, 2009.

[91] R. E. Haamer, E. Rusadze, I. Lüsi, T. Ahmed, S. Escalera *et al.,* "Review on emotion recognition databases," in *Human-Robot Interaction-Theory and Application*, 1st edition, London, UK: IntechOpen, pp. 39, 2017.

[92] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu *et al.,* "A natural visible and infrared facial expression database for expression recognition and emotion inference," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 682–691, 2010.

[93] H. Nguyen, K. Kotani, F. Chen and B. Le, "A thermal facial emotion database and its analysis, in: Pacific-Rim Sym. on Image and Video Technology, Berlin, Heidelberg: Springer, 397–408, 2013.

[94] C. Ordun, E. Raff and S. Purushotham, "The use of AI for thermal emotion recognition: A review of problems and limitations in standard design and data," pp. 1–13, 2020. [Online]. Available: https://arxiv.org/pdf/2009.10589.pdf.

[95] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *IEEE Conf. on Computer Vision and Pattern Recognition*. Providence, RI, USA, 2879–2886, 2012.

[96] L. Trujillo, G. Olague, R. Hammoud and B. Hernandez, "Automatic feature localization in thermal images for facial expression recognition," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*. San Diego, CA, USA, p. 14, 2005.

[97] I. A. Cruz-Albarran, J. P. Benitez-Rangel, R. A. Osornio-Rios and L. A. Morales-Hernandez, "Human emotions detection based on a smart-thermal system of thermographic images," *Infrared Physics & Technology*, vol. 81, no. 1, pp. 250–261, 2017.

[98] J. Stemberger, R. S. Allison and T. Schnell, "Thermal imaging as a way to classify cognitive workload," in *Canadian Conf. on Computer and Robot Vision*. Ottawa, Canada, 231–238, 2010.

[99] S. Jarlier, D. Grandjean, S. Delplanque, K. N'Diaye, I. Cayeux *et al.,* "Thermal analysis of facial muscles contractions," *IEEE Transactions on Affective Computing*, vol. 2, no. 1, pp. 2–9, 2011.

[100] H. Rivera, C. Goulart, A. Favarato, C. Valadão, E. Caldeira *et al.,* Development of an automatic expression recognition system based on facial action coding system, Simpósio Brasileiro de Automação Inteligente (SBAI2017), Porto Alegre, Brazil, pp. 615–620, 2017.

[101] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[102] K. Mallat and J. L. Dugelay, "A benchmark database of visible and thermal paired face images across multiple variations," in *Int. Conf. of the Biometrics Special Interest Group (BIOSIG)*. Darmstadt, Germany, pp. 1–5, 2018.