IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 6, 2020, accepted June 23, 2020, date of publication June 30, 2020, date of current version July 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3006004

Implementation of Wavelet Analysis on Thermal Images for Affective States Recognition of Children With Autism Spectrum Disorder

NAZREEN RUSLI[1], SHAHRUL NAIM SIDEK[1], (Senior Member, IEEE), HAZLINA MD YUSOF[1], (Member, IEEE), NOR IZZATI ISHAK[1], MADIHAH KHALID[2], AND AHMAD AIDIL ARAFAT DZULKARNAIN[3]

[1]Department of Mechatronics Engineering, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia
[2]Department of Curriculum and Instruction, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia
[3]Department of Audiology and Speech-Language Pathology, International Islamic University Malaysia–Kuantan, Kuantan 25200, Malaysia

Corresponding author: Shahrul Naim Sidek (snaim@iium.edu.my)

This work was supported by the Transdisciplinary Research Grant Scheme (TRGS19-01-002-0009) and by the Fundamental Research Grant Scheme (FRGS16-030-0529) through the Ministry of Higher Education Malaysia.

ABSTRACT Children with Autism Spectrum Disorder are identified as a group of people who has difficulties in socio-emotional interaction. Most of them lack the proper context in producing social response through facial expression and speech. Since emotion is the key for effective social interaction, it is justifiably vital for them to comprehend the correct emotion expressions and recognitions. Emotion is a type of affective states and can be detected through physical reaction and physiological signals. In general, recognition of affective states from physical reaction such as facial expression and speech for autistic children is often unpredictable. Hence, an alternative method of identifying the affective states through physiological signals is proposed. Though considered non-invasive, most of the current recognition methods require sensors to be patched on to the skin body to measure the signals. This would most likely cause discomfort to the children and mask their ''true'' affective states. The study proposed the use of thermal imaging modality as a passive medium to analyze the physiological signals associated with the affective states nonobtrusively. The study hypothesized that, the impact of cutaneous temperature changes due to the pulsating blood flow in the blood vessels at the frontal face area measured from the modality could have a direct impact to the different affective states of autistic children. A structured experimental setup was designed to measure thermal imaging data generated from different affective state expressions induced using different sets of audio-video stimuli. A wavelet-based technique for pattern detection in time series was deployed to spot the changes measured from the region of interest. In the study, the affective state model for typical developing children aged between 5 and 9 years old was used as the baseline to evaluate the performance of the affective state classifier for autistic children. The results from the classifier showed the efficacy of the technique and accorded good performance of classification accuracy at 88% in identifying the affective states of autistic children. The results were momentous in distinguishing basic affective states and the information could provide a more effective response towards improving social-emotion interaction amongst the autistic children.

INDEX TERMS Autism, affective states, facial skin temperatures, thermoregulation, thermal images, wavelet.

I. INTRODUCTION

For decades, numbers of research have been conducted in the field of affective state recognition, using numerous modalities and exploiting various features from the signal generated by the Autonomous Nervous System (ANS). ANS together with hypothalamus, regulates the blood pressure, breathing, pulse and arousal in response to different emotional features. Seeing the big potentials of affective state recognition system, a systematic review of the literature was performed in the study to identify research gaps in addressing the issue of recognition of affective states using thermal

The associate editor coordinating the review of this manuscript and approving it for publication was Gulistan Raja.

images particularly. In general, the modalities available for affective state detection can be divided into three categories that are behavioral response, physical reaction and physiological signal. The behavioral response is usually measured through the engagement or body language towards the activities whilst physical reaction is presented through facial and speech expressions respectively. To some extent, these two modalities are questionably appropriate to be implemented on the Autism Spectrum Disorder (ASD) children since they have difficulties in expressing their ''true'' affective states physically or behaviorally. Hence, the other perspective of looking at the automatic response inside the body for the features of different affective states was investigated. The chosen modality for recognition of affective states via physiological signals which are directly proportional to the response from the body as the reaction of ANS activities was discussed. The measurement via physiological signals is seemed to be more natural and ''honest'' as it happens automatically and unregulated in our body system. However, in all the measurements of the physiological signals, the in-contact type of sensors is necessary to be patched onto the skin body. As a result, the person who wears the sensor would possibly be aware about its presence and most likely influence his ''true'' affective states response.

Although, there are many ways in analyzing the correlation between affective states and physiological signals in the human body, the study focuses on the frontal facial skin temperature. It is due to the fact that the face is directly exposed to social communication and interaction. Even though, the facial expression is the most rampant means used to identify the affective states [1], yet, these means could still lead to false positive response in recognizing the true affective states due to the fact that facial expression can be faked. On the same note, the use of thermal imaging has been reported in detecting the changes in temperature for masked fearful faces as well as masked happy faces [2].

The target of the study was the individual who cannot reveal and express their affective states verbally or might give inappropriate facial response when induced like the ASD children. Conventional method requires the therapists or caregivers to recognize the affective states of the autistic children through facial expression. The ability of thermal imaging to measure the affective states even in a ''masked'' emotion usually expressed in facial expression form for typical developing (TD) children motivates the study to be extended to investigate the affective states of autistic children. The accurate identification of affective state would be able to help enhance the performance of current rehabilitation and training regimens.

In many cases, the change of affective states in an individual was recognized as either positive or negative emotions through the level of engagements or from a group of emotions displayed. Engagement is epitomized as behavioral feedback meanwhile emotions normally refer to the speech or facial expression. However, in the case of the children with ASD, their speech and facial expression are unpredictable to gauge their emotions. In fact, Gross [3] has identified the impairments in interpreting their facial expressions and it was supported by Sato [4] that found that emotion expressions are less noticeable in children with ASD. These gaps may obviously be related to social interaction deficits. A recent study by Poljac [5] also emphasized on the differences in interpretation of facial expressions and its associated affective states as the relation depended on the extent of autism traits at the individual level.

A. AFFECTIVE STATES RECOGNITION FROM THERMAL IMAGES

Thermal imaging is a passive technique used to create a heat map of objects appearing in a sight without external lighting source. It has been successfully explored and employed in the medical area in which it introduces the use of non-invasive and non-contact physiological sensors to the measurement process. It is a highly adaptable and delicate method to convert infrared light into temperature information, consenting wireless monitoring of the subject and detecting complex regional pain syndrome, inflammatory diseases, Raynaud's phenomenon [6] and cancer [7]. According to [8], pulsating blood flow produces the strongest variation on the temperature signal. Nardelli [9] found that the change in affective states causes variation in the skin temperature. Since the human face composition consists of thermal imprints that reflect the area with different concentration of blood vessels and hence it can be easily spotted using thermal camera. These thermal patterns particularly on facial skin have been used for analysis of correlation between physiological signals and affective states. Correlation between the changes in facial skin temperature and affective states was significantly proven by [10] where the maps of bodily sensation temperature changes were studied in response to the visual stimuli. Supporting the research, Kosonogov *et al.* have proven that thermal imaging provides a reliable tool to detect the changes and hence enables one to differentiate between affective states [11]. Although, they have evaluated the efficiency of thermal imaging in affective states recognition, there were missing gaps in the research. None of the research has ever reported the application of the method to recognize the affective states of ASD children. It can be observed that many studies so far focused on TD individuals.

As to conclude, thermal imaging binds the body's naturally emitted thermal irradiation to allow temperature readings due to the varying pulsation of the blood flow. This method enables cutaneous temperature recordings to be measured noninvasively, and unobtrusively.

B. THERMAL IMAGE ANALYSIS

The use of thermal imaging for affective states analysis has been reported using various methods and algorithms ([12]–[18]). Basu, Routray and Deb [12] have explained the features extracted from histogram analysis and fed to a multi-class support vector machine for classification of the

emotions whilst Bijalwan, Balodhi and Gusain [13] applied eigenface features in Principal Component Analysis and distance classifier. Similar goes to [14] who has studied on the histogram analysis at the specific region of interest on face to detect a deception emotion. A comparative analysis between dermal and thermal was also conducted and proven the capabilities of thermal images to investigate the emotional responses by [11]. Apart from this recent studies, a review of the thermal imaging applications were summarized by [15] to identify the capabilities and successful works done on the similar area of affective states detection using thermal images. Nevertheless, a standard golden procedure for thermal imaging processing has not been reported.

In the field of machine vision, feature extraction is a fundamental topic involving delicate image processing procedures, prior to the extraction of the important information from the input image. The step is regarded as a special form of dimensionality reduction which goal is to reduce computational cost when dealing with real-time applications. The common methods available in feature-based analysis of image processing are statistical method [12] to extract the mean, standard deviation, maximum and minimum values, Grey Level Co-Occurrence Method (GLCM) [19], [20], geometrical method to extract Voronoi tessellation feature and fractal information [21], and signal processing method like Wavelet transform [22], Gabor filters, and curvelets to name a fews. All of these methods manipulating the local variation of images either based on the greyscale intensity or conversion of temperature values in thermal images. Statistical parameters are used to characterize the content of an image by looking at its texture. It can be further grouped into first-order (one pixel), second-order (two pixels) and higher order (three and more pixels). The basic difference is the first order statistics consider a property of an individual pixel and neglecting the neighborhood pixels whereas, a GLCM as an example of second-order statistic generates the information with regards to the neighboring pixel. On the other hand, geometrical analysis usually is used in pattern recognition and it has been extensively discussed in biological and medical applications in a book by [23]. In the meantime, signal processing method is the continuing works of texture analysis from an image into time series data and its counterparts to the first two methods; first-order and second-order statistical analyzes that only focus on the texture of an image. Signal processing can be applied if there are a number of sequences in image of a video file.

Generally, features extraction from thermal images can be divided into two domains namely: imaging features and temperature features [24]. Imaging features are defined as features processed directly on the image intensity values, meanwhile temperature features are the conversion values from intensity to temperature readings from each pixel in the thermal image. Yasunari [25] reported the use of 2-dimensional discrete cosine transformation (2D-DCT) to transform the greyscale values as imaging feature of the facial image into frequency components, and applied the extracted features in the expression recognition systems. In support of the work was Basu, Routray and Deb [12], who have applied the histogram features to classify the affective states. This feature falls under imaging features in which the greyscale intensity values were taken into measurement. The features were used in the classification of four affective states namely the anger, fear, happy and sad with the accuracy of 81.95%. Conversely, for the temperature features, [26] has exploited the variances in the thermal intensity values recorded at thermally significant locations on human faces as the features to distinguish pretended and evoked affective expressions. In the similar context, [11] has extracted the time, frequency and time-frequency features derived from temperature readings in thermal images to classify the natural responses of subject- levels of arousal and valence induced by series of pictures from the International Affective Picture System (IAPS) and others. [27] examined the discriminative features based on the temperature distribution of the periorbital region for detecting deceptions. According to their work, a total of 492 responses were gathered from twenty-five subjects consisting of 249 lies and 243 truth responses. Throughout the data collections stage, the right and left sides of periorbital region were continuously tracked, and the images were captured at the rate of 25 frames per second. The developed classifier model attempted to classify the truths or the lies by using two approaches that were between-person approach and within-person approach. In another interesting work by [24], they proposed a method for recognizing affective states based on the thermal images. They used facial temperature data from the thermal images to classify different emotions (happy, fear and disgust). The sources of the thermal images were taken from the Natural Visible and Infrared facial Expression (USTC-NVIE) database. In each identified ROI, there were six features extracted namely the mean, mean of absolute values, mean of positive and negative values individually, temperature difference histogram features (TDHFs) and variance, sampled at rate of two frames per second. In the similar study of deception detection, [28] has also applied temperature features on segmented face and overall areas of the thermal images. The temperature features consist of mean, individual minimum and maximum temperature, standard deviation and range of temperature for facial and overall areas correspondingly. In most of the experiments, they have done the analysis based on the first order statistical features such as computation of mean, standard deviation, variance, minimum and maximum and variance in either intensity or heat signature of pixels within a ROI [15]. However, there was also a study done on the second-order statistical features called as Gray Level Co-Occurrence Matrix (GLCM). [29] has started to choose GLCM for region descriptors computation of the thermal images to distinguish the expressions of surprise, happiness and anger. GLCM was first introduced by [30] and it is amongst the widely used texture feature extraction method in computer vision field. The GLCM characterizes the second-order statistic of an image by computing how

often pairs of pixels with specific values and in a specified spatial relationship occur in an image.

In our preliminary study, [19] and [20], only the GLCM has been successfully applied in the affective states recognition of the adults with the accuracy of 86%. Thus, a combination of statistical method and signal processing method was proposed in this study where the texture of a single image was first analyzed. The mean intensity, area of ROI and GLCM parameters were later chosen as thermal features. Among the extracted thermal features, the mean intensity values were further scrutinized into using signal processing method. The mean features were translated into continuous signal and Wavelet Transform (WT) was applied.

WT was introduced as the solution to overcome the limitation of Fourier Transform (FT) and Short-time Fourier Transform (STFT) since FT can only analyze the frequency components of the signal but without information on the time-domain of the particular frequency rises meanwhile STFT applies the sliding window and is able to give data in both time and frequency domains but yet the length of the window limits the resolution of the frequency. This is due to the huge window measure in STFT recovers the frequency resolution and yet supposition of stationary data inside the window might be undermined while picking little window size may prompt poor frequency resolution. Thus, WT was preferred in order to locate the frequency of interest in the time domain. Theoretically, WT are divided into three types namely, Discrete Wavelet Transform (DWT), Wavelet Packet Transform (WPT) and Continuous Wavelet Transform (CWT). The DWT can be efficiently realized by decomposing the signal into approximation (low frequency) and detail (high frequency) coefficients and it operates at the integer having the power of two. On the other hand, WPT, after the first level, both the detail and approximation are decomposed into further level. Meanwhile, for the CWT, it can operate at any scale. The later type is more competent and reliable due to its capabilities to keep all the information without the need of down-sampling. The CWT results in Continuous Wavelet Coefficients (CWC), which illustrates how well a wavelet function correlates with a specific signal. If the signal has a major frequency component corresponding to a particular scale, then the wavelet at this scale is similar to the signal at the location where this frequency component occurs, regardless of its amplitudes and phases.

In this study, CWT was used to locate the affective states event in the time series data. The outcome of the convolution between signals was established as the basis for comparing the nominated mother wavelet functions. Selection of the most similar mother wavelet function has been a challenge for the application of wavelet transform in signal processing, as different mother wavelet applied on to the same signal may produce different results. The mother wavelet function is the main base of wavelet transforms that would authorize similarities of correlated coefficients across multiple signals. The more alike the mother wavelet function is to the wavelet coefficients across signals, the more accurate the signal of interest can be secluded and recognized; hence, the selection of a mother wavelet function is of supreme implication. Heretofore, there is no standard or general method to select a mother wavelet. A study reported in [31] has determined the most suitable and similar wavelet function to the biosignals: forearm electromyographic (EMG), electroencephalographic (EEG), and vaginal pulse amplitude (VPA). The results showed that a periodic behavior can be extracted from those three biological signals using Morlet function thus opted in this study. Furthermore, a rational reason to use the Morlet wavelet function was due to the thermoregulation phenomena in human's body.

During the expression of specific emotions or affective states, a change in facial temperature would appear due to the thermoregulation mechanism in the body. Thermoregulation mechanism is a special, natural, body control system to maintain the temperature of a body at a relatively constant hypothalamic setpoint [32]. Since the variation of the transient change of hypothalamic setpoint is of non-stationary nature, the signal is suitable for wavelet-based analysis. Of the same note, early study has been conducted to find the correlation between skin blood flow and wavelet by [33] and the results were able to differentiate the control mechanisms of blood flow in response to a variety of stimuli. The method used has been validated by comparison with laser Doppler imaginary effect amongst 20 healthy subjects whereby, cutaneous blood flow values, simultaneously computed by thermal IR imagery and measured by laser Doppler imaging. The results depicted a linear correlation ($R = 0.85$, Pearson Product Moment Correlation). Therefore, it is possible to transform raw thermal image series in cutaneous blood flow image series. This method has been applied in psychophysiology study of emotion assessment [15] and deception detection [34].

The mother wavelet should be carefully selected to better estimate and capture the transient in the de-noised thermal signals. To the extent, the Morlet Wavelet Function that was used as a mother wavelet has the basic characteristics of individual body thermoregulation for both transient and steady-state shape of the signals. It is used to detect any abrupt changes or transient event in the temperature signals to be correlated with the salient change in the affective states. The outcome of the correlation between signal and wavelet basis function has been established as the basis for comparing selected mother wavelet functions. In this research study, Morlet wavelet was chosen as the mother wavelet.

II. RESEARCH DESIGN

A well-planned experimental setup and systematic flow of experiment was designed for the collection of thermal imaging data from TD children (control group) and children with ASD. According to [35], it was possible to infer subjects' affective states from information collected from different types of resources, such as audio, video, self-report tools and interaction analysis.

A. SUBJECT SELECTION

The cross-sectional approach was adopted where two groups of subjects were recruited namely the TD children and the ASD children. The range of age of the male and female population was set between 5 to 9 years old from both groups. However, 86% of the selected subjects from ASD children were male aged between 5 to 6 years. A random sampling method was applied to select the subjects from TD children group. Meanwhile a purposive sampling method was used to select the ASD children from the category of advance level based on behavioral analysis score which required the subjects to be able to adequately follow the given instructions, such as to sit down on the chair and face the screen for stimulus. The required number of subjects recruited in the experiment was calculated from a priori analyze in G*power software where it is an open software used to calculate statistical power and sample size. A reference value of mean and standard deviation taken from Nicolini's work [36], and the outcomes with error probability at 0.20 and Power at 0.80 is tabulated in Table 1.

TABLE 1. Parameter in G*Power analysis.

Term	Descriptions
Analysis	A priori: Compute required sample size
	Tails(s) = One
Input	α err prob = 0.20
	Power(1-β err prob) = 0.80
	Sample size group 1 = 23
Output	Sampe size group 2 = 23
	Total Sample size = 46
	Actual Power = 0.8058150

As can be seen in the Table 1, 23 subjects were required as the minimum number from each group of TD and ASD subjects accordingly. In the study, a total of 57 subjects (34 TD Children and 23 ASD children) were recruited in which ethical clearance (IREC465) was secured in prior.

B. STIMULI SELECTION

In order to elicit the specific affective states, three distinct sets of audio-visual stimuli corresponding to happy, fear and sad basic affective states were administered to the subject via a monitor screen. Many research in the field of physiology explicitly specializing in emotion and attention have adopted still images from the standard database, the International Affective Picture System (IAPS) in their studies. The images from the IAPS database are systematically categorized and rated to meet the standard set by the psychologists. However, from observation, the ASD children seemed to face with difficulties in understanding and responding to the still images from IAPS appropriately and it was supported by [37]. As a result, two sets of audio-visual based stimuli were developed for TD and ASD children groups, respectively. The audio-visual stimuli were composed of five set of IAPS still images for each affective state to be induced from TD group. Meanwhile, the stimuli used for ASD group were further personalized based on the feedback obtained from the caregivers. The final forms of stimuli with proper scenario of situations and sounds were adopted in the study post consultation with psychiatrist, therapist and special needs teachers. Both sets of stimuli were composed of three standard segments namely the baseline (T1) (Pre-segment), stimuli for induction (T2) (During-segment), and rest (T3) (Post-segment) respectively. The T2 segment was designed to induce the intended affective states and composed of either series of IAPS still-images or dynamic images with respected scenario. Whereas, the segments T1 and T3 were composed of a black dot at the center of a white screen and full black screen accordingly. The goal was to ascertain engage state and relax state respectively. The timeframe of the stimuli spanned at the intervals of one to five minutes. A quick break of two to three minutes between the stimuli was given. In order to avoid misreading or mixing of affective states in the data sequence, the stimuli were arranged from positive to negative valence as depicted in the Fig.1 for Circumplex Model (i.e.1-happy, 2-fear and 3-sad).

C. EXPERIMENTAL PROTOCOL

A proper data collection process during the experiment is paramount to generate dataset with the highest accuracy and consistency. In the experiment, a FLIR thermal camera model T420 sitting on a pan-tilt system to track the face was used to capture thermal facial images of the subjects. The camera was set with emissivity value of 0.98 for temperature reading of human skin. Throughout the data collection session, a set of standard protocols was observed and described as follows:

1) The experiment was prepared in a controlled room with temperature set to 24°C. The room was free from ambient noise and clear from foreign distraction. The subject should be healthy and with the presence of therapist or caregiver and researcher during data collection.
2) The subject was required to seat comfortably in front of a monitor screen with the distance of 1.5m apart between the subject and the thermal camera.
3) The subject was ensured to secure the forehead area from any occlusion like hat or scarf. Glass made material should also be cleared from the image frame, which could represent opaque in thermal spectrum. Hence, the subject with eye glass was excluded in the study.
4) The experiment was conducted by administering stimuli via a monitor screen and at the same time the subject's frontal facial was recorded using thermal camera.
5) At the end of each session, the subject was requested to fill-up the questionnaire related to the elicited affective state and should rest for 2-3 minutes before the elicitation of the next affective state. This step was only applicable to TD group of children as it was sure that the children could render response appropriately. In the case of ASD children, the experiment was executed upon the subject readiness and the therapist would assist in responding to the questionnaire based on his

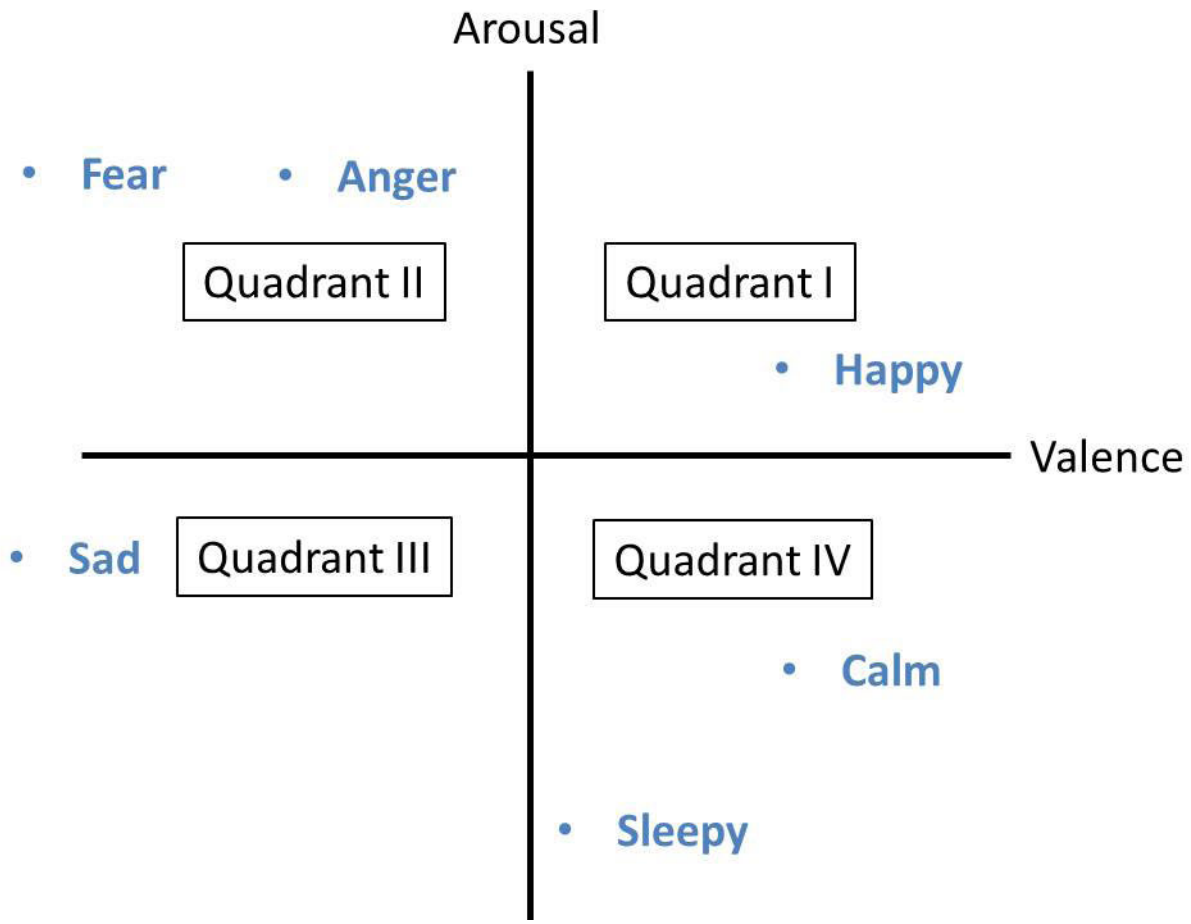

FIGURE 1. Russell's Circumplex Model.

personal assessment of the subject. The process might completed in the same day or continue in the following day subjected to the child readiness and advice from the therapist.

The experiment was conducted in a single trial for every subject from both groups of children to avoid bias. However, for the ASD children, the subjects would have a few prior sessions following similar experimental protocol but using different set of stimuli. This was to ensure the children to be familiar and comfortable with the new environment so he/she could place more focus in doing the actual experiment.

III. THERMAL DATA ANALYSIS

In a Machine Learning, a proper selection of features is important to avoid unnecessary complexity in the model and it is known that the algorithms have the tendency to be affected by noisy data. Noise should be reduced as much as possible in order to improve the efficiency of the algorithm. Thus, detail analyses were executed for every taken step to produce decent thermal features with minimal noise. The section expounds the equitable findings in the pre-processing thermal images and feature extractions and selections.

A. SEGMENTATION OF REGION OF INTEREST (ROI)

The selection of supraorbital and periorbital regions as the ROIs in the study stemmed from the fact that the two ROIs were found to be highly independent and significant in the affective state model analysis [38]. A matching template technique was implemented to automatically locate the ROIs for a complete series of thermal images for a subject. Since the process was done in an offline mode, every recorded video was initially converted into raw frontal facial thermal images. Then, template was taken from a manual cropped image of periorbital region from the first image of each subject. A threshold value of 150 was applied onto template to convert to a binary image as shown in Fig.2.

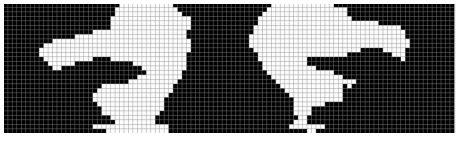

FIGURE 2. Template in binary.

The matching process started from the top left corner of the image and it computed the difference for each pixel in the frame size (3×3) of template and the process continued to shift over the entire image. The absolute differences were

FIGURE 3. Output of ROI cropped images.

summed up and stored. The minimum of the summation of absolute differences over entire process was selected as the best matched image. Fig.3 displays results of cropped image in frames over time. Supraorbital was automatically cropped at the same time during the segmentation process of periorbital.

B. THERMAL FEATURES

Strong correlated features were required to give better performances in the accuracy of the affective states classifier. In this study, three major steps were considered in order to confirm good results of the developed classifier from the TD children which were:

1) Time detection of the affective state change
2) Selection of the best thermal features
3) Selection of quality datasets

1) TIME DETECTION OF AFFECTIVE STATE CHANGE

A dedicated algorithm was designed to automatically locate the event where the induced affective states transpired within the measured signals. It was stemmed from the fact that the continuous difference in content of the stimuli used was causing the response to vary by time. As such, a CWT technique was deployed to analyze the signals. CWT was used to locate the affective states event in the time domain for thermal image video data. The outcome of the convolution between signals was established as the basis for comparing the nominated mother wavelet functions. The ability to locate the instantaneous event in the scalogram resulted in the important information about the specific timing and duration of the induced affective states within the signals. This helped in the analysis of the signals as such the identified window of the signals was further analyzed statistically to extract the required features.

The occurrence of similar frequency pattern of non-sinusoidal signal at specific time was most vigorous to identify the affective states change reflected in the de-noise thermal signals. The wavelet scalogram was used to identify signals with specific frequency components. It is a 3D grey color representation of wavelet coefficients, where the horizontal axis represents the time, t, the vertical represents the scale at evenly distributed frequency, a and the z-axis displays the color scale of square magnitude of the wavelet coefficients on an intensity graph. The scaling factor of the mother wavelet was heuristically changed and tested to supreme match with the time of occurrence. The Analytic Wavelet Transform (AWT) was used in the LabVIEW as it is a special case of the CWT with Morlet wavelet. The output of the AWT is the CWT coefficients, which reflects the wavelet energy spectrum and can be viewed by plotting its scalogram. The scalogram defines the time-frequency behavior of the signal and can be used to locate the events of affective states. The scale value of AWT (i.e. a) set in the LabVIEW took the form of 2^M where M ranges between 0 and 8. The squared magnitude of CWT coefficient showed in the scalogram was jointly representing the signal in terms of time and frequency at a constant scale. Large scales corresponded to low frequencies, and small scales corresponded to high frequencies.

2) SELECTION OF THE BEST THERMAL FEATURES

The selection of thermal features is crucial as it may affect the performance and computational cost of the classifier. A proper selection of best thermal features was then required. Three signals sampled at 3 Hz namely de-noise temperature signals, the setpoints values, and de-noise thermal signals minus setpoints values were generated from mean intensities in the ROI. The statistical parameters such as mean, μ, standard deviation σ, median η and frequency of the highest amplitude point in the signals were extracted to form the new feature set totalling to 12 number of features (4 features x 3 signals) and added to the original features extracted from the ROI thermal image that were blood vessels area, GLCM contrast, GLCM correlation and GLCM homogeneity. This has been detailed in [19], [20]. A complete set of 20 thermal features were presented in Table 2.

In order to select the best features from the list, a statistical analysis was employed based on ANOVA test. Kruskal-Wallis H test and Post-Hoc Test were executed to result in the best fifteen thermal features (two were rejected from Kruskal-Wallis Test that were Stdev_Diff and Mean_area and three from Post-Hoc Test namely Freq_DTS, Stdev_SP and Freq_WC). The final 15 features from each ROI thermal image were then combined independently to form the complete input feature dataset. The result from the combination of these features promised a better performance as explained in the following section.

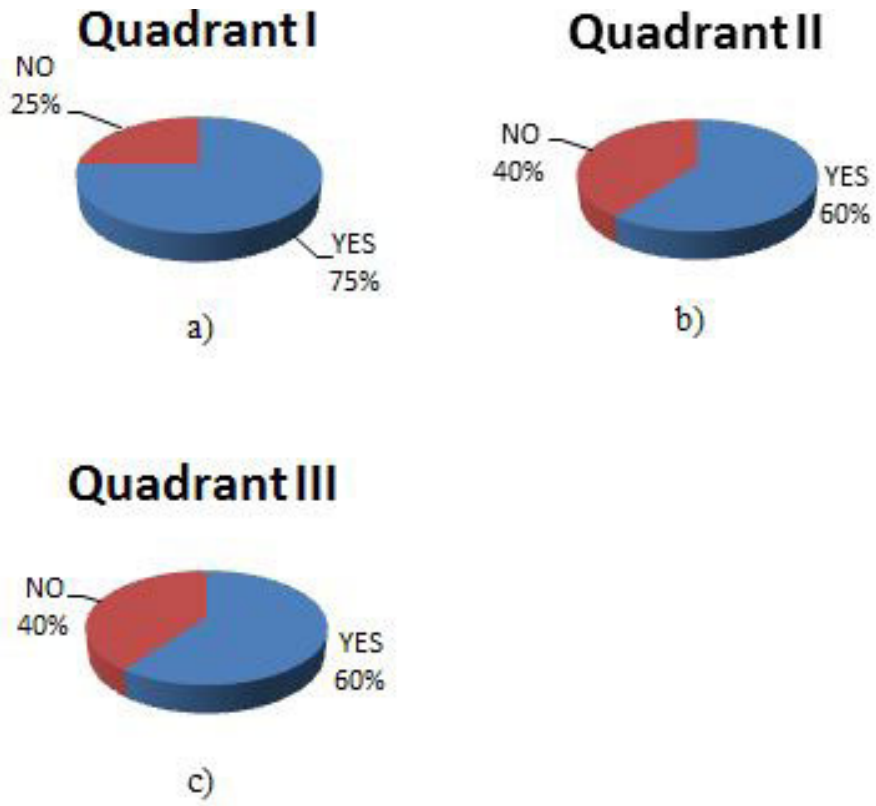

FIGURE 4. Summary of the responses of TD children from the questionnaires (i.e DEQ) for three quadrants (i.e. "YES" signifies the same affective state as the targeted state and "NO" for otherwise.)

TABLE 2. A complete list of thermal features.

List of Thermal Features	Computed Statistical Parameters	Denotations
De-noise thermal signals (DTS)	Mean	Mean_DTS
	Standard Deviation	Stdev_DTS
	Frequency of a signal with maximum amplitude	Freq_DTS
	Median	Median_DTS
Set-point values (SP)	Mean	Mean_SP
	Standard Deviation	Stdev_SP
	Frequency of a signal with maximum amplitude	Freq_SP
	Median	Median_SP
Difference (DIFF)	Mean	Mean_Diff
	Standard Deviation	Stdev_Diff
	Frequency of a signal with maximum amplitude	Freq_Diff
	Median	Median_Diff
Blood Vessel's Area	Mean	Mean_area
	Standard Deviation	Stdev_area
GLCM Homogeneity (HMG)	Mean	Mean_HMG
	Standard Deviation	Stdev_HMG
GLCM Contrast (CRST)	Mean	Mean_CRST
	Standard Deviation	Stdev_CRST
GLCM Correlation (CORR)	Mean	Mean_CORR
	Standard Deviation	Stdev_CORR

3) SELECTION OF QUALITY DATASETS

A high quality dataset should contribute to the optimum accuracy in the classification process. Since the reference model was formed based on the datasets from TD children, the results from questionnaires were paramount to verify the model. Therefore, a properly developed set of questionnaires was designed. The approach was supported by [40] who reported that children were better informants than parents or teachers when it concerned their own emotions. A set of questionnaires was developed by using a new tool for measuring state Self-Reported emotions used by [41]. They conducted various studies against the new tool, Discrete Emotions Questionnaires (DEQs) to measure eight distinct affective states: anger, anxiety, disgust, desire, fear, happiness, relaxation and sadness. The ANOVA results of the research have revealed the subscales of the DEQ in properly detecting the manipulated affective states. Their studies showed that the DEQs were sensitive to several different manipulations of affective states such as nonfictional recall, guided pictures, and graphic stimuli to name a few. In our study, the application of DEQ to find out the response of the subjects towards the audio-video stimuli has resulted in the percentages shown in Fig. 4.

As can be seen from the pie chart, 60% from the total responses matched the expected affective state of Quadrant II and Quadrant III (i.e answering ''YES'') whereas 75% of the responses matched the Quadrant I. Thus, only the datasets from the ''YES'' group were chosen to be included in the development of the reference model so to form the best classifier.

IV. RESULTS

In the research study, there were two sample groups involved TD and ASD children, correspondingly. The ability to regulate affective states between the groups was assumed to be asymmetrical due to the neurodevelopmental difference

that affects the affective states regulation. An exploratory and separate analysis was executed to examine the performance accuracy by varying the scaling factors a and the results were presented.

A. ANALYSIS OF WAVELET SCALES

The recorded thermal video from TD children formed the basis to pinpoint the changes in facial expression for different affective states. Fear affective state was the only one chosen as a guideline to confirm the identified location as it is a universal emotion or affective state with typical and unique facial expression by Ekman [39]. In the thermal images, the fear face can be observed with known prescribed description such as a dropped open jaw and the lips stretched horizontally backwards as depicted in Fig.5b.

Therefore, a visual observation was made to spot the similarity in time for the corresponding fear expression with the output of CWT coefficients. The red color shown in the scalogram was the maximum energy generated while purple color represents the least amount of energy. The energy was computed such that the center frequency of the wavelets at the analyzed scales evenly sample the frequency range from 0 to sampling rate/2. The central frequency of a wavelet is inversely proportional to the scale. The resulting scalogram is a kind of joint time-frequency representation with an adaptive time-frequency resolution. For example, at the chosen scale a=256, the energy was computed at frequency range from 0 to 0.002 Hz. As can be seen in the scalogram (Fig. 6e), the maximum of energy was clearly charted from 24 to 34 seconds and yet, in the study, the energy was taken from green color and onwards. The chosen range of energy was mapped accordingly with the fear facial expression of the child that was a dropped open jaw and lips stretched backwards. The comparison between scaling factors were also conducted on the same child to ensure the validity of the chosen scaling factor.

As noted in the Fig. 6b, c and d, three signals were produced and analyzed namely de-noise temperature signals, the setpoint values and the difference between de-noise thermal signal and setpoints values. The thermal signals extracted from mean intensities in a series of thermal images was formed then filtered using a Morlet wavelet filter to generate a clean and clear pattern. After that, a hypothalamic setpoint and differences between hypothalamic setpoints and de-noise thermal signals were derived. The hypothalamic setpoints were computed from the average of temperature readings over three seconds period. The period was the fastest time to detect any change in affective state [42]. The inter-threshold range is the range of hypothalamic setpoints within which no thermoregulatory mechanism is required

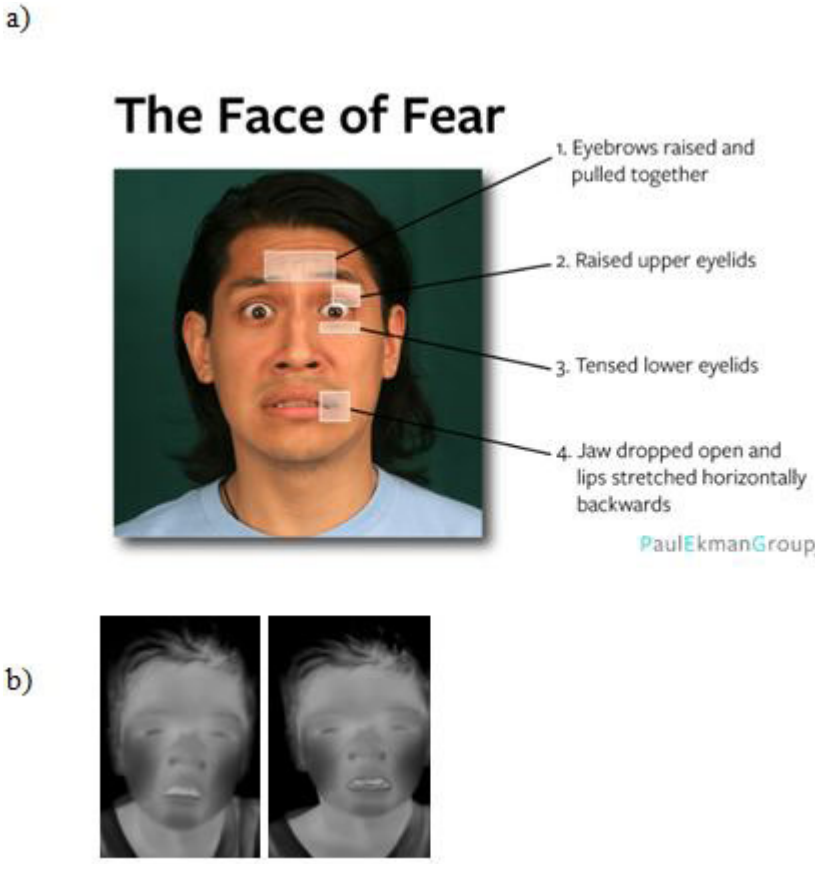

FIGURE 5. a) Fear facial expression by Paul Ekman Group b) Thermal images of TD children for fear facial expression.

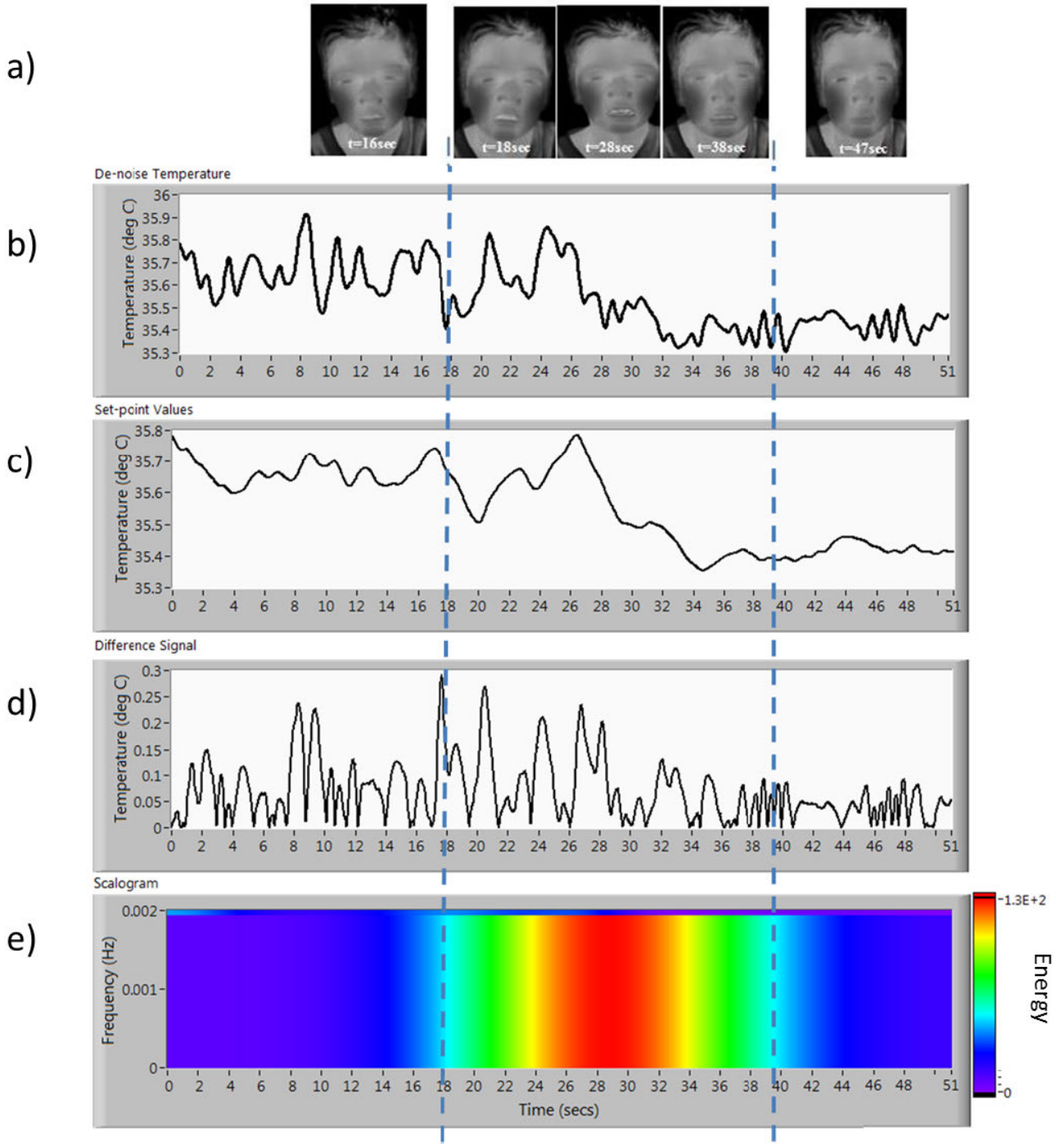

FIGURE 6. a) The recorded fear affective state response in thermal image for TD child #3. b) The thermal readings in ° C c)The computed average setpoints for every 3 secs d) The difference between filtered signals and average Setpoint e) A scalogram where it shows the energy projected from the convolution of thermal signals and Morlet wavelet at scale a=256.

or activated. A ground study by [43] has highlighted that the inter-threshold range with no activation of autonomic effector was about 0.3°C. When a person was induced with specific emotion, the hypothalamus in the brain sensed the changes in the temperature above the inter-threshold value, and reacted upon it by activating the thyroid gland so to produce hormones to control the phenomenon of vasoconstriction and vasodilation of blood vessels and hence affect the blood flow. As observed in Fig. 6c, the hypothalamic setpoints seemed to change to a new setpoint value within the selected range of energy as the indication of autonomic response. To emphasize the findings, the heuristical observations were done to select the best scale where the scales were tested to have varied value with the power of two as such $a = 2, 4, 16, 32, 64, 128, 256$ and the results were shown in Fig. 7.

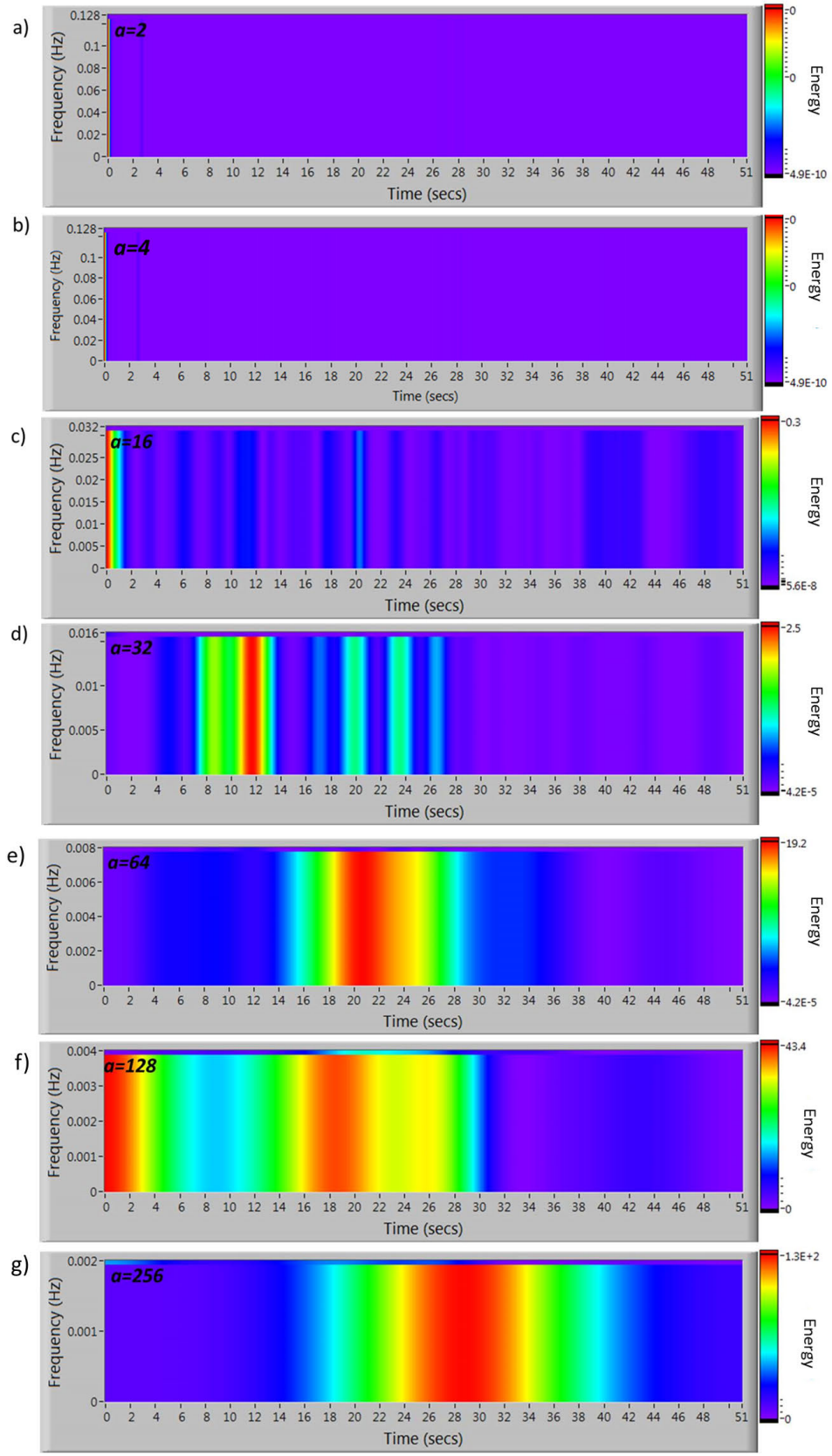

FIGURE 7. The scalogram output for TD children at different scaling factors, a at a)2 b)4 c)16 d)32 e)64 f)128 and g)256.

The results have remarked the effects of the location of the event occurrence within the session. The smaller value of a would miss any event triggered meanwhile a large scale would lead to a wide coverage of detection thus mismatched the facial expression. The best scaling factor of 256 was finally chosen to locate the occurrence of affective states for all TD children. To note, some of the subjects produced fear expression in multiple events. The significant findings was critical and has proven as evidence to further deploy the AWT in the affective states classifier development process.

B. PERFORMANCE OF CLASSIFIER

Three classes of affective states from Russell's Circumplex model [44] were investigated to represent three different quadrants in the model namely Quadrant I, Quadrant II and Quadrant III as shown in Fig.1. The specific affective states were distinguished by the degree of valence and arousal measurements in the respective quadrants. In this research, the type of affective states under study were happy as it is the only state having positive valence, anger or fear which are located at the negative valence axis with high degree of arousal and sad which is a negative valence affective state with low arousal.

A k-NN classifier was selected due to its low computational cost but relatively high performance, and applicable simplexes approach which worked fine in the study. The relatively small number of data (57 subjects) from the self-collected data of TD and ASD subjects and low dimensionality features also contributed to the selection of k-NN classifier. The k-NN classifier is a non-parametric method which works by determining a cluster of k objects in training dataset which is the nearest to the test object. In this work, a heuristically analysis was done to determine the best k neighbors in k-NN classifier by using a five fold Cross-Validation (CV) technique in which 80% were used for training and the remaining 20% used for testing. The cross-validation was done independently between TD and ASD datasets. The best results were obtained at k=5.

In the research study, a cross-validation was chosen to statistically evaluate the performance of the machine learning models on a limited data sample. The k-fold cross-validation technique is an ideal method because the technique increases the chances to train on multiple train-test separations hence, gives a better indication on the performance towards the unseen data.

1) TD CHILDREN

The selection of scaling factor plays an important role to detect occurence of the induced affective state. However, the improper scaling might possibly lead to a different location of event that should match the occurence of affective states and thus, reducing the performance accuracy of the classifier. Since there were three different classes of affective states under study, a number of heuristic trials was conducted using different value of scaling factor a from a=32 to a=256. The results from the TD children were tabulated in Table 3

TABLE 3. The matched occurrence of affective states for different scaling factors for TD subjects in percent when algorithm was run automatically.

Affective States	*a=32*	*a=64*	*a=128*	*a=256*
Quadrant I (Happy)	0%	0%	16%	84%
Quadrant II (Fear)	0%	4%	18%	78%
Quadrant III (Sad)	0%	0%	22%	78%

where it can be seen that the highest matched occurrence of affective states in percent happened at a=256.

Subsequently, a cross-validation of the classifier was performed using the best a value (i.e. 256) and the result showed in Table 4 was promising with the average accuracy of 88% in classifying the three affective states.

TABLE 4. The percentage in accuracy of k-NN classifier among TD children.

Affective States	Accuracy (%)			Average Acuracy (%)
	Quadrant I (Happy)	Quadrant II (Fear)	Quadrant III (Sad)	
Quadrant I (Happy)	88	8	4	88
Quadrant II (Fear)	5	88	7	
Quadrant III (Sad)	5	6	89	

With respect to the other studies on affective state recognition based on thermal images, the performance of classification model proposed in the study was on par. For example, Basu, Routray and Deb [12] has applied the histogram features to classify affective states in adults. The features were then the imaging features in which the greyscale intensity values were taken into measurement. The features were used in a classification of four affective states anger, fear, happy and sad with the accuracy of 81.95%. In a similar study, [13] reported an excellent performance with accuracy of 97.00% for a distance classifier using principal component analysis (i.e. eigen-faces for feature extraction) on adults facial thermal images. However, the study was focused on the facial expression in the thermal images. Meanwhile, in [16], the authors presented a new marker of affective arousal in facial thermal imaging where the ROI was focused on the tip of the nose region on adults. The results accorded the accuracy at 89.90% for a different set of affective states (happiness, disgust, fear and sadness). Furthermore, a Deep Boltzmann Machine has also been used in the affective states classification as reported in [17] where the implementation of thermal databases of facial expressions, includes 38 adult subjects. They obtained the accuracy of 62.90% for classification of negative and positive valence affective states. In the latest study found in [18], the authors leveraged on the emissivity variation values of the thermal camera to classify five affective states namely happiness, surprise, disgust, fear and sadness. The results accorded a mean accuracy of 85.25%.

2) ASD CHILDREN

The same procedure and algorithms in data analysis to develop the affective states classifier for TD children were extended to the ASD children datasets. The results were tabulated in Table 5.

TABLE 5. The percentage in accuracy of *k*-NN classifier among ASD children at *a*=256.

Affective States	Accuracy (%)			Average Acuracy (%)
	Quadrant I (Happy)	Quadrant II (Fear)	Quadrant III (Sad)	
Quadrant I (Happy)	88	4	8	83
Quadrant II (Fear)	11	72	17	
Quadrant III (Sad)	8	4	88	

It was interesting to note that the results were sporadic and decreased. Thus, further investigation was conducted to scrutinize the procedure and analysis. As reported by Paula [45], the power spectrum of EEG signals collected from the ASD children when subjected to visual stimulus of happy, neutral, and angry faces revealed that there was an increase in power in the higher frequencies region of the signals recorded from frontal, occipital, and center-parietal areas when compared to control group. Thus, there was a need to look into the empirical evidence in the approach as there was high probability that ASD children might emit different pattern of signals alike TD children. Hence, the same procedure was repeated for ASD children but now with focused attention given to the thermal video images by manually pinpoint the resemblance in the timing of matched affective states with the actual facial expression. As happy was generally recognizable affective state of the ASD children [46], thus, the state became the reference state for the case of ASD. The chosen subject (ASD #4) was selected by therapist. He has been diagnosed with mild autism and had attended a few sessions of behavioral therapy. In addition, he has been taught with few types of affective states and ways to react in appropriate manners. It is important to note that the stimuli was individually designed and suggested by the parents in which the subject favorite audio-video clip was presented to induce the happy response. The algorithm was then executed to identify the best scaling factor, *a* and the results were shown in the scalogram of Fig.8.

When matched to the actual expression in thermal image as shown in Fig. 9 it was noted that the best matched occurrence of affective state happened when the scaling factor, *a* was set to 32.

The smaller value of scaling factor, *a* signified a smaller magnitude of time and it was inversely proportional to frequency which correlated with the work of Paula [45]. The result also indicated a unique observation in the data of ASD children in which the higher scaling factor, *a* at 256 has also resulted in a good match similar to TD children. This might suggest, the wide spectrum nature of the ASD, and as a result a specialized algorithm was developed to find the best matched occurrence for ASD children based on the best two scaling factors, *a* of 32 and 256 as shown in Table 6.

TABLE 6. The matched occurrence of affective states for different scaling factors for ASD subjects in percent when the algorithm was run automatically.

Affective States	*a=32*	*a=64*	*a=128*	*a=256*
Quadrant I (Happy)	38%	0%	0%	62%
Quadrant II (Fear)	50%	0%	0%	50%
Quadrant III (Sad)	46%	0%	0%	54%

Therefore, to exclusively focus into the development of affective state classifier for ASD children, the decision to split the classifier into two separate models was done. It was expected to increase the performance accuracy of the classifier. The datasets of ASD children were based on the response verified by the therapist as the DEQ approach was not applicable to ASD children. The detailed breakdown of the performance accuracy for the unique classification model of ASD children with fifteen extracted thermal features at two-option of scaling factors, *a* was tabulated in Table 7.

TABLE 7. The percentage in accuracy of k-NN classifier among ASD children at *a* fixated value of *a* at 32 and 256 and the switched *a* technique.

Affective States	Accuracy (%)		
	a=32	*a=256*	*a=32 and a=256*
Quadrant I (Happy)	84	88	86
Quadrant II (Fear)	78	72	80
Quadrant III (Sad)	82	88	93
Average	**82**	**83**	**88**

The average accuracy was noted at 88% which is remarkably comparable to the classifier of TD children at 88%. The affective state that accorded the highest classification accuracy was sad (Quadrant III), then happy (Quadrant I) followed by state fear (Quadrant II).

The relatively lower performance accuracy to recognize the affective state fear might stem to the fact that the ASD children were born with impairment in limbic area especially amygdala that contributed to the deficiency in most emotions. In particular, the amygdala plays a vital role in processing negative emotions with high arousal such as fear, thus, it was only speculative that it might also affect the overall performance of the ASD children affective state classifier. This is consistent with supposition that the ASD children have deficiency in the affective state regulation especially when associated with impairment of amygdala [47]. The degree in valence correlates mainly with the activation of orbitofrontal cortex, whereas arousal correlates with activation of the amygdala. The ASD children with impaired amygdala were found to have more deficits in processing static negative emotions than positive emotions [48]. Likewise, the amygdala deficit might

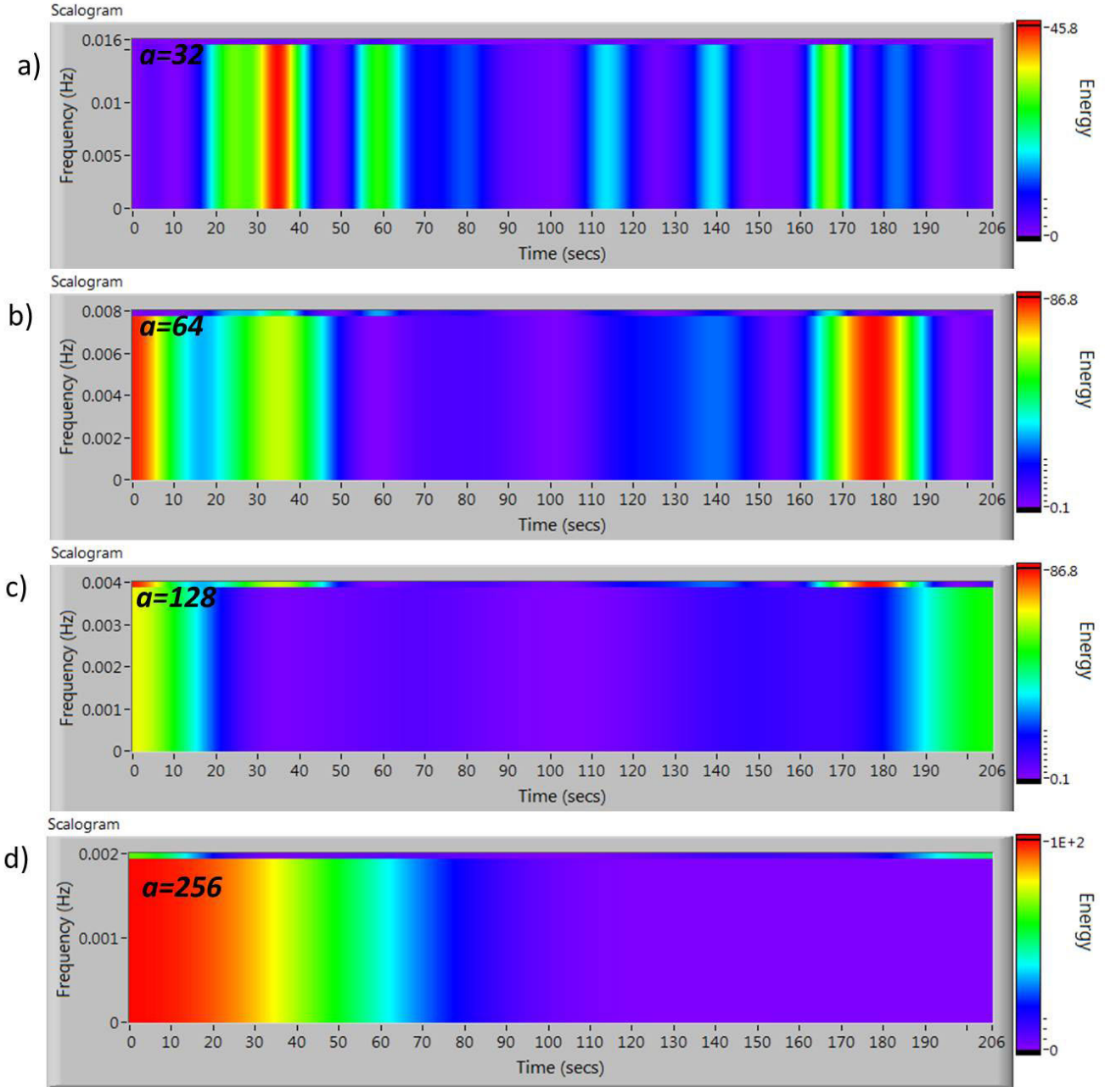

FIGURE 8. The scalogram output for ASD children at varying scaling factors, *a* at a) 32 b) 64 c) 128 d) 256.

lead the ASD children to abnormal fear response such he/she may either show too little or too much fear or none, when compared to TD control children.

C. QUESTIONNAIRES ANALYSIS

In many research [49]–[51] and from the consultation of expert therapists, guided questionnaire responded by the caregivers, done in a very controlled and structured environment was the best option to gage the affective state of the autistic children. In this research the questionnaire was used to verify the classifier and help guided the development of classifier in the prior stage based on the response obtained from the control group (the TD children who answered the questionnaire themselves). Otherwise similar signal processing technique were applied on the thermal signals from both group of subjects.

The performance of the independent affective state classifier of ASD children was validated against the analysis of input response from therapist-report questionnaires. A Leave One-Out Cross Validation (LOCV) technique was executed to verify the ASD classifier where data from a child was completely removed from the training dataset and solely used as a testing dataset. The process was reiterated with another set of unique data until all the data has been assigned as a

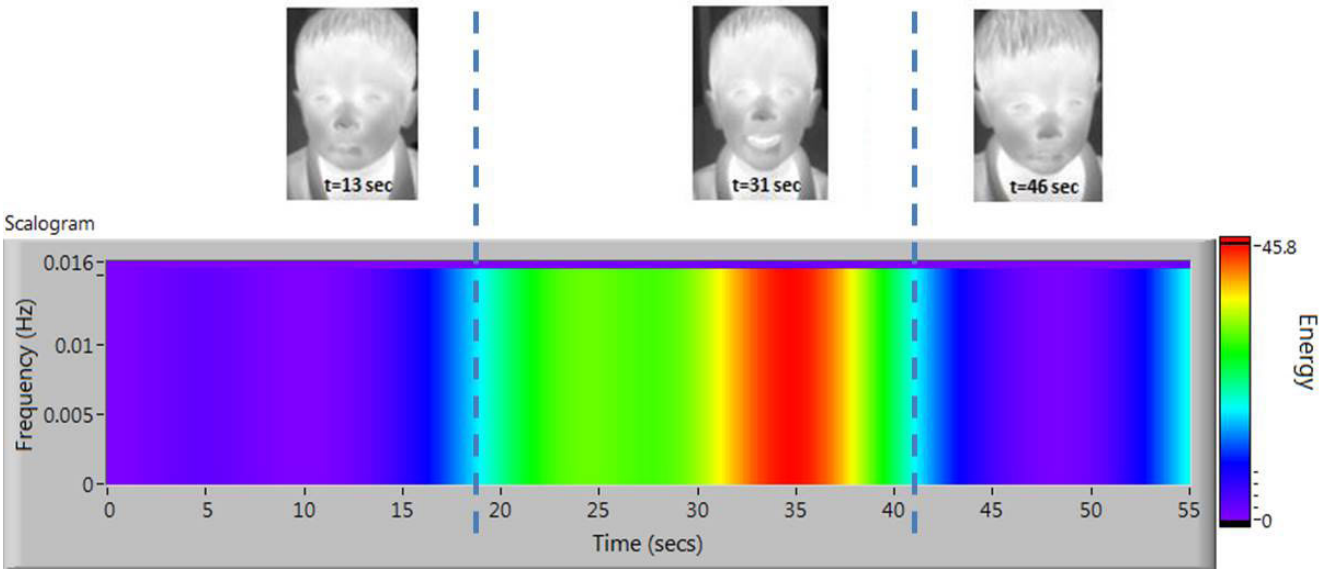

FIGURE 9. The recorded thermal response and results from AWT analysis with scaling factor at a=32.

TABLE 8. Kappa statistics of the therapist reading and classifier's output of the affective state recognition of ASD children.

Affective States	Inter-raters Therapist vs. ASD Thermal Classifier		
	Kappa, κ	Standard Error	p-value
Quadrant I (Happy)	0.831	0.082	<0.005
Quadrant II (Fear)	0.560	0.075	<0.005
Quadrant III (Sad)	0.560	0.075	<0.005

testing dataset. The therapist is an experienced occupational therapist with a qualification in Bachelor (Hons) in Occupational Therapy, UiTM Puncak Alam. She has been involved in the special needs children (ASD) early intervention training for more than 5 years. She responded the questionnaires based on her experiences and observations on the children's daily behavior. The inter-rater analysis was then executed to find the reliability of therapist rating of the affective states and the classifier's output. Weighted kappa (κ) with linear weights [52] was run to determine the agreement between all the actual affective states and the output from affective state classifier for ASD. The results were tabularized in Table 8.

There was a statistically significant agreement between the therapist-prediction and classifier-prediction, κ=0.560, $p < 0.005$ for affective states in Quadrant II and III respectively. The strength of agreement was classified as medium to good according to [53] and good according to [54]. On the other hand, a nearly perfect agreement was found between therapist-prediction and classifier prediction where κ=0.831, $p < 0.005$ for an affective state in Quadrant I. The result explained that, there was an excellent agreement between the classifier's affective states of ASD children and therapist reading for happy. This shows that the happy is the easiest distinguishable and noticeable affective state. This explanation was in accordance to the previous meta analysis conducted by Mirko [46] where they summarized the individuals with ASD to have difficulties in recognizing five basic affective states or emotions namely anger, disgust, fear, sad and surprise. The only state that was mostly recognized was happy. However, there were still some inconsistent reports that suggested on the sturdiness of determination of distinct affective states in ASD especially associating to negative affective states. The states fear and sad were suspected to post a challenge in visual observations by therapist. Furthermore, there were also researchers that considered the assessments of the behavior by the therapist may need to be tailored to a more specific question. For instance, for sad state, it was better to be correlated with a measure of empathy in daily life and experience, rather than droopiness in life. The observations might sometimes tend to relate to the therapist personal experiences as well.

V. CONCLUSION

This study has presented the application of the Wavelet Transform to instantaneous detect the occurrence of affective states during experiments with high accuracy. The method contributed to the development of affective state classifier to classify three distinct emotions induced by the audio-video stimuli. The agreement between both raters (therapist and independent affective state classifier of ASD children) verifies the model and justifies the needs of the thermal imaging modalities and the use of non-invasive tool in recognizing the affective states especially the negative valence affective states. The outcome of the research work is a thermal based classifier for affective state model of autistic children. It is envisaged that the classifier could be used to increase the

efficacy and efficiency of the rehabilitation or training regimens. The therapist will be more aware of the ''true'' affective states of the autistic children during the session and allows them to react and address the children accordingly.

VI. LIMITATIONS

The study has successfully developed an affective state classifier for three basic emotions. More affective states based on arousal-valence model should be investigated to generalize the result in the future study.

ACKNOWLEDGMENT

The study presented was carried out at NASOM and IDEAS Autism Centers. Our deepest appreciation also goes to special-needs teachers, therapist at the centers and psychiatrist Associate Professor Dr. S.K. Pillai that have involved in making this project possible. This research received ethical approval from IIUM Research Ethics Community (IREC no. 465). Written informed consent was gathered from all the subjects for publication of the research.

REFERENCES

[1] C. F. Benitez-Quiroz, R. B. Wilbur, and A. M. Martinez, ''The not face: A grammaticalization of facial expressions of emotion,'' *Cognition*, vol. 150, pp. 77–84, May 2016.

[2] M. Kanat, M. Heinrichs, R. Schwarzwald, and G. Domes, ''Oxytocin attenuates neural reactivity to masked threat cues from the eyes,'' *Neuropsychopharmacology*, vol. 40, no. 2, pp. 287–295, Jan. 2015.

[3] L. Giombini, *Handbook Emotion Regulation*. Tokyo, New Delhi: Taylor & Francis, 2015.

[4] W. Sato, S. Uono, and M. Toichi, ''Atypical recognition of dynamic changes in facial expressions in autism spectrum disorders,'' *Res. Autism Spectr. Disorders*, vol. 7, no. 7, pp. 906–912, Jul. 2013.

[5] E. Poljac, V. Hoofs, M. M. Princen, and E. Poljac, ''Understanding behavioural rigidity in autism spectrum conditions: The role of intentional control,'' *J. Autism Develop. Disorders*, vol. 47, no. 3, pp. 714–727, Mar. 2017.

[6] E. F. J. Ring and K. Ammer, ''Infrared thermal imaging in medicine,'' *Physiol. Meas.*, vol. 33, no. 3, pp. R33–R46, Mar. 2012.

[7] S. Mambou, P. Maresova, O. Krejcar, A. Selamat, and K. Kuca, ''Breast cancer detection using infrared thermal imaging and a deep learning model,'' *Sensors*, vol. 18, no. 9, p. 2799, Aug. 2018.

[8] A. Saxena, E. Y. K. Ng, and V. Raman, ''Thermographic venous blood flow characterization with external cooling stimulation,'' *Infr. Phys. Technol.*, vol. 90, pp. 8–19, May 2018.

[9] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, ''Recognizing emotions induced by affective sounds through heart rate variability,'' *IEEE Trans. Affect. Comput.*, vol. 6, no. 4, pp. 385–394, Oct. 2015.

[10] E. Salazar-López, E. Domínguez, R. V. Juárez, L. F. J. de, A. Meins, O. Iborra, G. Gálvez, M. A. Rodríguez-Artacho, and E. Gómez-Milán, ''Arousal, emotional valence, empathy, facial temperature, subjective experience, thermography,'' *J. Consciousness Cognit.*, vol. 34, pp. 149–162, Oct. 2015.

[11] E. S. Kosonogov, J. M. Nandrino, and S. Henrique, ''Facial thermal variations: A new marker of emotional arousal,'' *J. PLoS ONE*, vol. 12, no. 9, 2017, Art. no. e0183592.

[12] A. Basu, A. Routray, and A. K. Deb, ''Human emotion recognition from facial thermal image using histogram based features and multi-class support vector machine,'' in *Proc. Annu. IEEE India Conf. (INDICON)*, 2015, pp. 1–5.

[13] V. Bijalwan, M. Balodhi, and A. Gusain, ''Human emotion recognition using thermal image processing and eigenfaces,'' *Int. J. Eng. Sci. Res.*, vol. 1, pp. 34–40, May 2015.

[14] M. S. Panasiti, D. Cardone, E. F. Pavone, A. Mancini, A. Merla, and S. M. Aglioti, ''Thermal signatures of voluntary deception in ecological conditions,'' *Sci. Rep.*, vol. 6, no. 1, Dec. 2016, Art. no. 35174.

[15] D. Cardone and A. Merla, ''New frontiers for applications of thermal infrared imaging devices: Computational psychopshysiology in the neurosciences,'' *Sensors*, vol. 17, no. 5, p. 1042, May 2017.

[16] 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,'' *Infr. Phys. Technol.*, vol. 81, pp. 250–261, Mar. 2017.

[17] S. Wang, M. He, Z. Gao, S. He, and Q. Ji, ''Emotion recognition from thermal infrared images using deep Boltzmann machine,'' *Frontiers Comput. Sci.*, vol. 8, no. 4, pp. 609–618, Aug. 2014.

[18] C. Goulart, C. Valadão, D. Delisle-Rodriguez, E. Caldeira, and T. Bastos, ''Emotion analysis in children through facial emissivity of infrared thermal imaging,'' *PLoS ONE*, vol. 14, no. 3, Mar. 2019, Art. no. e0212928.

[19] M. H. Abd Latif, H. Md. Yusof, S. N. Sidek, and N. Rusli, ''Thermal imaging based affective state recognition,'' in *Proc. IEEE Int. Symp. Robot. Intell. Sensors (IRIS)*, Oct. 2015, pp. 214–219.

[20] M. H. Abd Latif, H. Md. Yusof, S. N. Sidek, and N. Rusli, ''Implementation of GLCM features in thermal imaging for human affective state detection,'' *Procedia Comput. Sci.*, vol. 76, pp. 308–315, 2015.

[21] Q.-H. Zhao, X.-L. Li, Y. Li, and X.-M. Zhao, ''A fuzzy clustering image segmentation algorithm based on hidden Markov random field models and Voronoi tessellation,'' *Pattern Recognit. Lett.*, vol. 85, pp. 49–55, Jan. 2017.

[22] V. Lyashenko, R. Matarneh, and Z. Deineko, ''Using the properties of wavelet coefficients of time series for image analysis and processing,'' in *Proc. SCEP*, 2016.

[23] Y. Ogata and K. Katsura, ''Maximum likelihood estimates of the fractal dimension for random spatial patterns,'' *Biometrika*, vol. 78, no. 3, pp. 463–474, 1991.

[24] S. Wang, P. Shen, and Z. Liu, ''Facial expression recognition from infrared thermal images using temperature difference by voting,'' in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Intell. Syst.*, Oct. 2012, pp. 94–98.

[25] Y. Koda, Y. Yoshitomi, M. Nakano, and M. Tabuse, ''A facial expression recognition for a speaker of a phoneme of vowel using thermal image processing and a speech recognition system,'' in *Proc. 18th IEEE Int. Symp. Robot Human Interact. Commun.*, Sep. 2009, pp. 182–186.

[26] 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 Trans. Appl. Perception*, vol. 6, no. 1, pp. 1–22, Feb. 2009.

[27] B. A. Rajoub and R. Zwiggelaar, ''Thermal facial analysis for deception detection,'' *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 1015–1023, Jun. 2014.

[28] B. Diana, M. Elia, V. Zurloni, A. Elia, A. Maisto, and S. Pelosi, ''Multimodal deception detection,'' in *Proc. ACM Workshop Multimodal Deception Detection*, 2015, pp. 419–453.

[29] B. Hernández, G. Olague, R. Hammoud, L. Trujillo, and E. Romero, ''Visual learning of texture descriptors for facial expression recognition in thermal imagery,'' *Comput. Vis. Image Understand.*, vol. 106, nos. 2–3, pp. 258–269, May 2007.

[30] R. M. Haralick, ''Statistical and structural approaches to texture,'' *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.

[31] J. Rafiee, M. A. Rafiee, N. Prause, and M. P. Schoen, ''Wavelet basis functions in biomedical signal processing,'' *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6190–6201, May 2011.

[32] N. Charkoudian, ''Mechanisms and modifiers of reflex induced cutaneous vasodilation and vasoconstriction in humans,'' *J. Appl. Physiol.*, vol. 109, no. 4, pp. 1221–1228, Oct. 2010.

[33] M. J. Geyer, Y.-K. Jan, D. M. Brienza, and M. L. Boninger, ''Using wavelet analysis to characterize the thermoregulatory mechanisms of sacral skin blood flow,'' *J. Rehabil. Res. Develop.*, vol. 41, no. 6, p. 797, 2004.

[34] M. Dcosta, D. Shastri, R. Vilalta, J. K. Burgoon, and I. Pavlidis, ''Perinasal indicators of deceptive behavior,'' in *Proc. 11th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG)*, May 2015, pp. 1–8.

[35] M. Ez-zaouia and E. Lavoué, ''EMODA: A tutor oriented multimodal and contextual emotional dashboard,'' in *Proc. 7th Int. Learn. Analytics Knowl. Conf.*, Mar. 2017, pp. 429–438.

[36] Y. Nicolini, B. Manini, E. De Stefani, G. Coudé, D. Cardone, A. Barbot, C. Bertolini, C. Zannoni, M. Belluardo, A. Zangrandi, B. Bianchi, A. Merla, and P. Ferrari, ''Autonomic responses to emotional stimuli in children affected by facial palsy: The case of moebius syndrome,'' *J. Neural Plasticity*, pp. 1–13, 2019.

[37] A. Saitovitch, A. Bargiacchi, N. Chabane, A. Phillipe, F. Brunelle, N. Boddaert, Y. Samson, and M. Zilbovicius, ''Studying gaze abnormalities in autism: Which type of stimulus to use?'' *Open J. Psychiatry*, vol. 3, pp. 32–38, 2013.

[38] N. Rusli, S. N. Sidek, H. M. Yusof, and N. I. Ishak, "Mean of correlation method for optimization of affective states detection in children," *IEEE Access*, vol. 6, pp. 68487–68497, 2018.
[39] P. Ekman. (2004). *What is Fear What Causes Fear Paul Ekman Group*. [Online]. Available: https://www.paulekman.com/universal-emotions/what-is-fear/
[40] P. DiBartolo, G. Marten, and E. Amie, "Who is best at predicting children's anxiety in response to a social evaluative task?: A comparison of child, parent, and teacher reports," *J. Anxiety Disorders*, vol. 20, no. 5, pp. 630–645, 2006.
[41] C. Harmon-Jones, B. Bastian, and E. Harmon-Jones, "The discrete emotions questionnaire: A new tool for measuring state self-reported emotions," *PLoS ONE*, vol. 11, no. 8, Aug. 2016, Art. no. e0159915.
[42] Y. Fu and C. Frasson, "Detecting thermal emotional profile," in *Proc. PhyCS*, 2016, pp. 142–151.
[43] K. J. Collins, *Regulation of Body Temperature* (Care of the Critically Ill Patient), J. Tinker and W. M. Zapol, Eds. London, U.K.: Springer, 1992.
[44] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980.
[45] C. A. R. Paula, C. Reategui, B. K. D. S. Costa, C. Q. D. Fonseca, L. D. Silva, E. Morya, and F. L. Brasil, "High-frequency EEG variations in children with autism spectrum disorder during human faces visualization," *BioMed Res. Int.*, vol. 2017, pp. 1–11, Oct. 2017.
[46] M. Uljarevic and A. Hamilton, "Recognition of emotions in autism: A formal meta-analysis," *J. Autism Develop. Disorders*, vol. 43, no. 7, pp. 1517–1526, Jul. 2013.
[47] P. Vuilleumier, M. P. Richardson, J. L. Armony, J. Driver, and R. J. Dolan, "Distant influences of amygdala lesion on visual cortical activation during emotional face processing," *Nature Neurosci.*, vol. 7, no. 11, pp. 1271–1278, Nov. 2004.
[48] R. C. Leung, E. W. Pang, J. A. Brian, and M. J. Taylor, "Happy and angry faces elicit atypical neural activation in children with autism spectrum disorder," *Biol. Psychiatry, Cognit. Neurosci. Neuroimaging*, vol. 4, no. 12, pp. 1021–1030, Dec. 2019.
[49] A. Koáakowska, A. Landowska, A. Anzulewicz, and K. Sobota, "Automatic recognition of therapy progress among children with autism," *Sci. Rep.*, vol. 7, no. 1, Dec. 2017, Art. no. 13863.
[50] C. Wong, "Play and joint attention of children with autism in the preschool special education classroom," *J. Autism Develop. Disorders*, vol. 42, no. 10, p. 2152–2161, 2012.
[51] H. Pickard, C. Hirsch, E. Simonoff, and F. Happé, "Exploring the cognitive, emotional and sensory correlates of social anxiety in autistic and neurotypical adolescents," *J. Child Psychol. Psychiatry*, Mar. 2020.
[52] Cicchetti, Domenic V and Allison, Truet, "A new procedure for assessing reliability of scoring EEG sleep recordings," in *Amer. J. EEG Technol.*, vol. 11, no. 3, pp. 10–101, 1971.
[53] J. Landis, K. Richard, and G. Gary, "The measurement of observer agreement for categorical data," *J. Biometrics*, pp. 159–174, Mar. 1977.
[54] J. L. Fleiss, "Measuring nominal scale agreement among many raters," *Psychol. Bull.*, vol. 76, no. 5, p. 378, 1971.

NAZREEN RUSLI received the bachelor's degree in communications engineering from International Islamic University Malaysia, in 2005, and the Masters of Professional Engineers degree in electrical engineering from The University of Sydney. She is keen in signal and image processing, real-time computing, and intelligent control related research.

SHAHRUL NAIM SIDEK (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees from Vanderbilt University, USA, in 1998 and 2008, respectively. He is currently a Senior Staff with the Department of Mechatronics Engineering, IIUM. He also serves as an Associate Professor of mechatronics engineering at the Department of Mechatronics Engineering, International Islamic University Malaysia. His research interests include human centered elctromechanical systems, human–robot interation, and affective state computing.

HAZLINA MD YUSOF (Member, IEEE) received the B.Eng. degree (Hons.) in mechatronics and the M.Eng. degree in electrical from Universiti Teknologi Malaysia, in 1998 and 2000, respectively, and the Ph.D. degree in electrical and electronics engineering from Loughborough University, U.K., in 2012. Her research interests are in the areas of control systems design and applications concentrating on vehicle suspensions, sensors and actuators, and engineering education for upper and lower secondary. She is also actively involved with research on thermal imaging for human robotic interaction focusing on learning rehabilitation.

NOR IZZATI ISHAK was born in Kedah, Malaysia, in 1993. She received the bachelor's degree in mechatronics engineering from International Islamic University Malaysia, in 2017, and the master's degree from International Islamic University Malaysia. Her current research interests include biomechatronics and robot interactions.

MADIHAH KHALID received the Ph.D. degree in mathematics education from the Curtin University of Technology, Western Australia, in 2004. She is currently with the Department of Curriculum and Instruction, International Islamic University Malaysia (IIUM). She regularly publishes papers in refereed journals, conference proceedings, as well as book. Her areas of interests include curriculum and instruction, education for special needs, mathematics education, teacher education, assessment and evaluation, and lesson study.

AHMAD AIDIL ARAFAT DZULKARNAIN received the bachelor's degree (Hons.) in audiology from Universiti Kebangsaan Malaysia (UKM), in 2003, and the Doctor of Philosophy (Ph.D.) degree in audiology from the University of Queensland, Australia, in 2008. He has been a Deputy Dean of Postgraduate and Research of Kulliyyah of Allied Health Sciences, since 2016. His research interests include auditory-electro and neurophsyiology, clinical audiology, neuroaudiology, and signal processing.