



Article

Face Alignment in Thermal Infrared Images Using Cascaded Shape Regression

Kent Nagumo , Tomohiro Kobayashi, Kosuke Oiwa and Akio Nozawa *

Graduate School of Science and Engineering, Aoyama Gakuin University, Kanagawa 252-5258, Japan; nagumo@ieee.org (K.N.); kobachi369@gmail.com (T.K.); oiwa@ee.aoyama.ac.jp (K.O.)

* Correspondence: akio@ee.aoyama.ac.jp; Tel.: +81-42-759-6262

Abstract: The evaluation of physiological and psychological states using thermal infrared images is based on the skin temperature of specific regions of interest, such as the nose, mouth, and cheeks. To extract the skin temperature of the region of interest, face alignment in thermal infrared images is necessary. To date, the Active Appearance Model (AAM) has been used for face alignment in thermal infrared images. However, computation using this method is costly, and it has a low real-time performance. Conversely, face alignment of visible images using Cascaded Shape Regression (CSR) has been reported to have high real-time performance. However, no studies have been reported on face alignment in thermal infrared images using CSR. Therefore, the objective of this study was to verify the speed and robustness of face alignment in thermal infrared images using CSR. The results suggest that face alignment using CSR is more robust and computationally faster than AAM.

Keywords: face alignment; thermal infrared image; facial thermal image; cascaded shape regression; real-time measurement; remote measurement



Citation: Nagumo, K.; Kobayashi, T.; Oiwa, K.; Nozawa, A. Face Alignment in Thermal Infrared Images Using Cascaded Shape Regression. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1776. <https://doi.org/10.3390/ijerph18041776>

Academic Editor: Anna Lubkowska

Received: 31 December 2020

Accepted: 9 February 2021

Published: 12 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A method for remotely evaluating physiological and psychological states based on facial skin temperature measured by infrared thermography has attracted considerable interest. Biological information is used in various fields, such as medicine, welfare, and industry. In general, the measurement of biological signals often requires physical restraint, and the measurement itself may cause mental or physical stress to the subject [1]. Conversely, infrared thermography can conduct contactless, non-invasive skin temperature measurements with high sensitivity, accuracy, and reproducibility [2–5]. In a thermal environment that is windless and non-sweat inducing, the main cause of variation in skin temperature is skin blood flow [6]. Since the autonomic nervous system controls skin blood flow as a part of the circulatory system's function to regulate body temperature, skin temperature has been used to assess the activity of the autonomic nervous system [7]. For this assessment, facial thermal infrared images are particularly suitable for application because the face is often exposed and unobstructed by clothing. Many previous studies have been conducted on the estimation of physiological and psychological states based on facial skin temperature distribution [8,9]. For example, studies have been conducted that estimate vital data such as respiratory rate [10] and heart rate [11], sleepiness [12–14], emotions [15–17], mental stress [18,19], and anomaly detection in facial skin temperature distribution [20]. These previous studies used the temperature distribution of the entire face or the temperature of specific Regions of Interest (ROI), such as the nose, mouth, and cheeks for evaluation. Therefore, it is important to automatically detect faces and facial landmarks from thermal infrared images with high accuracy and stability. In recent years, infrared thermography has lowered in price but has bettered in performance. The resolution of thermal infrared images has increased, and it is possible to measure multiple people with a single thermal infrared image. Therefore, in face detection and detection

of facial landmarks in real environments, it is desirable to increase the speed to analyze multiple people at once.

The Active Appearance Model (AAM) [21] is one of the most popular methods for automatically detecting facial landmarks in thermal infrared images. AAM statistically models the changes in face shape and overall facial appearance and aligns the face shape with the model through nonlinear optimization. Kopaczka et al. conducted face alignment in thermal infrared images using AAM based on intensity, Histogram of Oriented Gradients (HOG), and Dense Scale Invariant Feature Transform (DSIFT) features [17,22]. However, in general, AAM is expensive because it solves an exact optimization problem. It also suffers from low robustness to poses, illumination, facial expression changes, and unknown subjects that are not included in the training set [23].

To solve these problems, Cascaded Shape Regression (CSR) has been proposed for face alignment in visible images [24–27]. In the CSR approach, facial landmark detection is estimated by regression, and the solution is updated multiple times by a multi-stage estimator to detect the facial landmarks. Face alignment using CSR is highly real-time. Ren et al. [26] reported face alignment at more than 3000 FPS speed. Hence, it is expected that CSR can be used for faster face alignment in thermal infrared images. However, no studies have been reported on face alignment in thermal infrared images using CSR. Therefore, the objective of this study was to verify the speed and robustness of face alignment in infrared images using CSR. First, a CSR model was created. Next, we trained and evaluated the CSR on the thermal infrared images acquired in our experiments. The results suggest that face alignment using CSR is more robust and computationally faster than AAMs proposed in the previous study, which is reported in this paper.

2. Cascaded Shape Regression

If x_i, y_i are the x, y coordinates of the i th facial landmark, then the face shape vector represented by the M facial landmarks is $\mathbf{S} = [x_1, y_1, \dots, x_M, y_M]^T$. The cascaded shape regression model is a model with a multi-stage structure estimator with T number of stages, which predicts the face shape $\mathbf{S}^{(t)}$ in a cascaded manner. Given the initial face shape \mathbf{S}_0 and the input image \mathbf{I} , the CSR model is updated by the estimator to find the shape difference fraction $\Delta\mathbf{S}^{(t)}$ and update the solution. At stage t , $\mathbf{S}^{(t)}$ and $\Delta\mathbf{S}^{(t)}$ are regressed as follows:

$$\mathbf{S}^{(t)} = \mathbf{S}^{(t-1)} + \Delta\mathbf{S}^{(t)}, \quad (1)$$

$$\Delta\mathbf{S}^{(t)} = r^{(t)}(\mathbf{I}, \mathbf{S}^{(t-1)}) \quad (2)$$

where $t \in 1, \dots, T$ is the number of estimators corresponding to each stage of the CSR, and $r^{(t)}$ is the estimator. The loss function is represented as follows:

$$\arg \min_{r^{(t)}} \sum_{i=1}^N |\hat{\mathbf{S}}_i - (\mathbf{S}_i^{(t-1)} + r^{(t)}(\mathbf{I}_i, \mathbf{S}_i^{(t-1)}))|^2 \quad (3)$$

where $\hat{\mathbf{S}}^{(t)}$ is the ground's true face shape, and N is the number of images for training. In the CSR, training is performed in such a way that this loss function is minimized. In this study, we estimated facial landmarks using the ensemble of regression tree learning methods used by Vahid et al. [27]. Gradient boosting was used as the training estimator. At each split node of the regression tree, the intensity difference sentence feature of two pixels [25,28] is determined based on the threshold. To train each split node, 400 randomly sampled features were computed.

3. Experiments

3.1. Experimental Methods

Experiments were conducted to acquire thermal infrared images of a face for training a facial landmark detector. Seven subjects (five males and two females) aged 22–24 years participated in the experiment. They were fully informed about the experiment and the

purpose of the study before their participation. All participants signed a consent form. The experimental system is shown in Figure 1. Thermal infrared images were captured using infrared thermography (FLIR A615-model: A615, 45° field of view, FLIR Systems, Oregon). The infrared camera had a resolution of 640×480 pixels and a temperature resolution of less than 0.05 K. Infrared emissivity is the ratio of the thermal radiation from the surface of an object to the radiation from a black body at the same temperature, given by Stefan–Boltzmann’s law. In order to obtain accurate temperature measurements, it is necessary to set the correct infrared emissivity of the surface of an object. In this study, the infrared emissivity of the skin was set to $\varepsilon = 0.98$ [29]. The experimental protocol is shown in Figure 2. Three distances between the subject and infrared thermography were 60 cm, 90 cm, and 120 cm (Figure 3). Each distance consisted of three recording intervals (Small, Large, and Random). As shown in Figure 4, the subjects were asked to turn their heads in nine directions (center, top center, top right, center right, bottom right, bottom center, bottom left, center left, top left) for the Small and Large sections. To evaluate the effect of the angle of face orientation on face alignment, subjects were asked to move their head angles to 20 degrees and 45 degrees in the Small and Large conditions, respectively. To increase the robustness of the face alignment, in the Random section, subjects were asked to move their head in any direction and make any facial expression they wanted. Nothing other than the subject’s body was recorded. The experiment was conducted in the experimental room without convection. Thermal infrared images were taken 15 min after the subjects entered the experimental room for thermal acclimation to the environmental temperature, and the time to take thermal infrared images for each subject was less than 5 min. A total of 609 thermal infrared images were obtained in this experiment. We manually annotated 68 landmarks for the obtained data according to the literature [30] and bounding boxes in the face region.

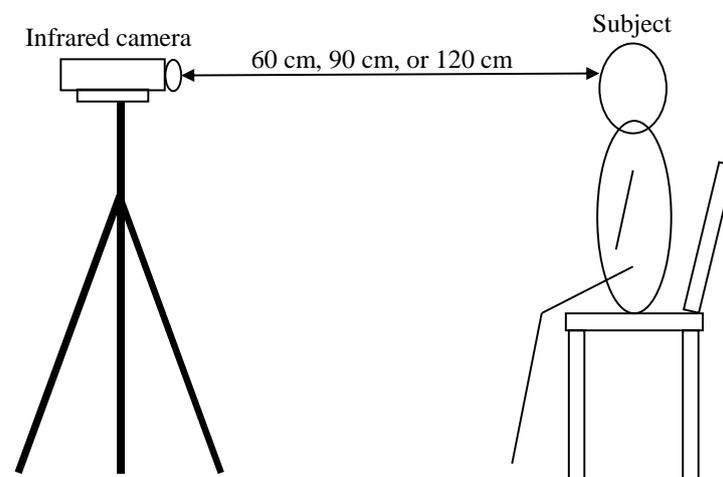


Figure 1. Experimental system. The distance between the subject and infrared camera was set at three different distances: 60 cm, 90 cm, or 120 cm.

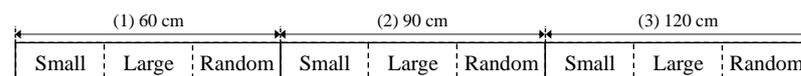


Figure 2. Experimental protocol.



Figure 3. Examples of thermal infrared images acquired under three different distance conditions. From left to right: 60 cm, 90 cm, and 120 cm.

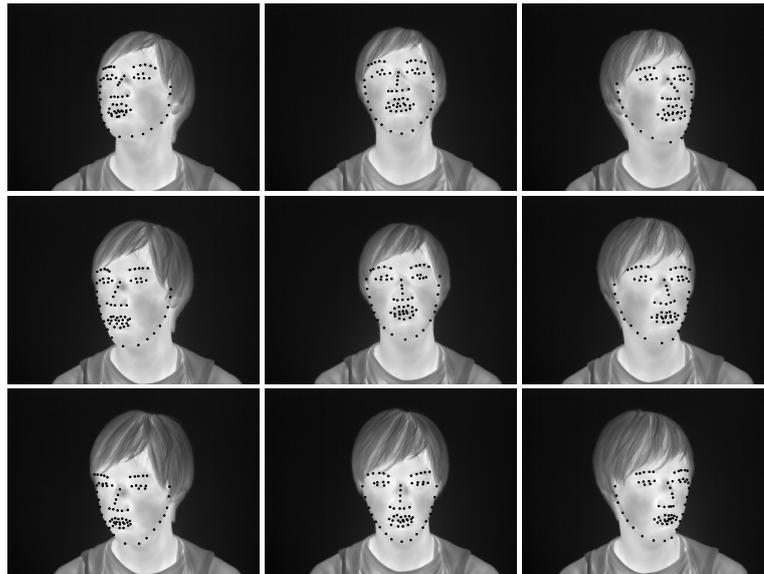


Figure 4. Examples of thermal infrared images oriented in nine different directions.

3.2. Analysis Methods

The acquired images were flipped to the left and right for data augmentation. As a result, 1218 images were created. To perform k-fold cross-validation ($k = 7$) using CSR, we split the data of six subjects into training data and the data of the remaining subjects into test data. All subjects' data were used as test data. Unless otherwise specified, all experiments were run with the following fixed parameter settings: the number of stages in the cascade $T = 10$, tree depth $F = 4$, number of weak regressors $K = 500$, and a random pair of pixels $P = 400$ used as the difference feature between two points. The average coordinates of the facial landmarks in the training data were used as the initial shape. The Normalized Point to Point Error (NPPE) introduced by Zhu et al. [31] was used as a method to evaluate the estimation accuracy of the face alignment. The $NPPE_i$ of each ith image is the following equation:

$$NPPE_i = N_i \sqrt{\frac{\sum_{n=1}^N [(x_{n,r} - x_{n,g})^2 + (y_{n,r} - y_{n,g})^2]}{2N}}, \quad (4)$$

$$N_i = \frac{1}{\frac{1}{2}(w_i + h_i)} \quad (5)$$

where $x_{n,r}$ and $y_{n,r}$ are the coordinates of the estimated facial landmarks, $x_{n,g}$ and $y_{n,g}$ are the coordinates of the correct facial landmarks, N is the number of facial landmarks, w_i is the width of the face, h_i is the height of the face, and N_i is the reciprocal of the mean of w_i and h_i . To compare the estimation accuracy of CSR models, we performed Intensity, DSIFT, and HOG-based AAM methods that were effective in aligning faces in thermal infrared images in previous studies [17,22]. Marciniak et al. [32] reported that the accuracy

of face recognition in visible images is lower when the number of pixels in the face region is small. To evaluate the effect of the number of pixels in the face region on face alignment, the number of pixels per face width was calculated. To evaluate the computation time, we measured the frames per second (FPS) of the face alignment of the test data for each method. The specifications of the evaluation PC in this experiment were Intel Core i7-8700 CPU and 16GB RAM. Only one CPU core was used. The program was implemented in C++ and Python.

4. Results

Table 1 shows the minimum, maximum, and mean values of the facial skin temperature and the ambient temperature. The ambient temperature was almost the same for all subjects in the experiment. Figure 5 shows the percentage of test images satisfying a given NPPE evaluated with CSR and Intensity, DSIFT, and HOG-based AAM. It is probably due to the problem that AAM is less robust to unknown subjects that are not part of the training set [23]. The CSR method has the highest number of images below 0.05, which is an acceptable accuracy value for NPPE [33]. Figure 6 shows examples of NPPE for face alignment. From Figure 6, it can be confirmed that the accuracy of face alignment becomes worse when the NPPE is greater than 0.5. The CSR method reached a higher total accuracy value. Figure 7 shows the mean value of NPPI of the test images for each method. Figure 8 shows examples of face alignment using each method. The mean NPPEs by CSR and conventional AAM were almost equal. Conversely, the variation of NPPEs was the smallest for CSR. From Figure 8, it can be confirmed that the accuracy of the face alignment of AAM becomes worse when the face is not looking the front. This suggests that face alignment by CSR is more robust than the AAM method and can be applied to face alignment for more varieties of images.

Table 2 shows the FPS of each method: the FPS of the CSR model is over 80, which is more than ten times larger than the AAM methods. The FPS of CSR was the largest, and the FPS of Intensity, DSIFT, and HOG-based AAM were smaller in that order. DSIFT and HOG-based AAMs are considered to have taken more time than Intensity-based AAM because of the calculation of DSIFT and HOG features. It is suggested that face alignment in thermal infrared images using CSR is highly real-time.

Table 1. Mean values \pm SD of the minimum, maximum, and mean facial skin temperature. The number of thermal infrared images for each subject was 87.

Subject	Facial Skin Temperature/ $^{\circ}$ C			Ambient Temperature/ $^{\circ}$ C
	Minimum	Maximum	Mean	
A	30.64 \pm 0.22	34.27 \pm 0.15	32.50 \pm 0.13	24.42 \pm 0.07
B	29.23 \pm 0.28	33.76 \pm 0.12	31.89 \pm 0.22	24.38 \pm 0.37
C	29.99 \pm 0.20	33.67 \pm 0.17	31.85 \pm 0.17	24.46 \pm 0.14
D	30.81 \pm 0.17	34.13 \pm 0.17	32.64 \pm 0.19	24.02 \pm 0.06
E	31.42 \pm 0.16	34.18 \pm 0.09	32.98 \pm 0.09	24.33 \pm 0.06
F	30.32 \pm 0.33	34.05 \pm 0.27	32.27 \pm 0.25	24.44 \pm 0.29
G	31.20 \pm 0.32	34.15 \pm 0.17	32.80 \pm 0.22	24.11 \pm 0.10

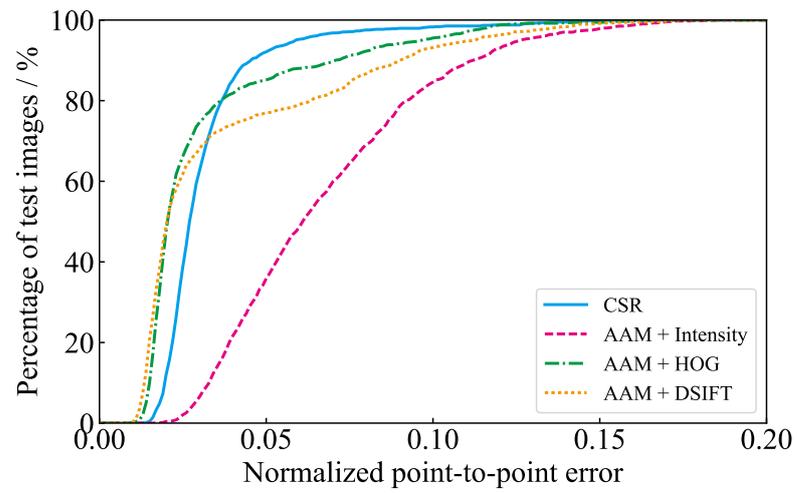


Figure 5. The percentage of test images satisfying a given Normalized Point to Point Error (NPPE) for each method.

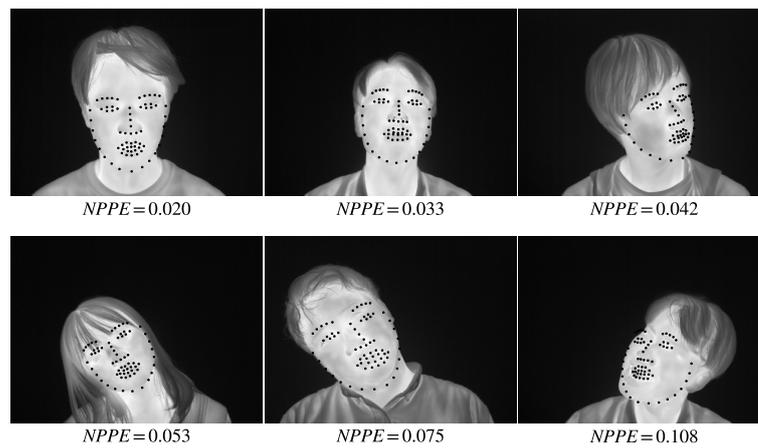


Figure 6. Examples of NPPE for each face alignment.

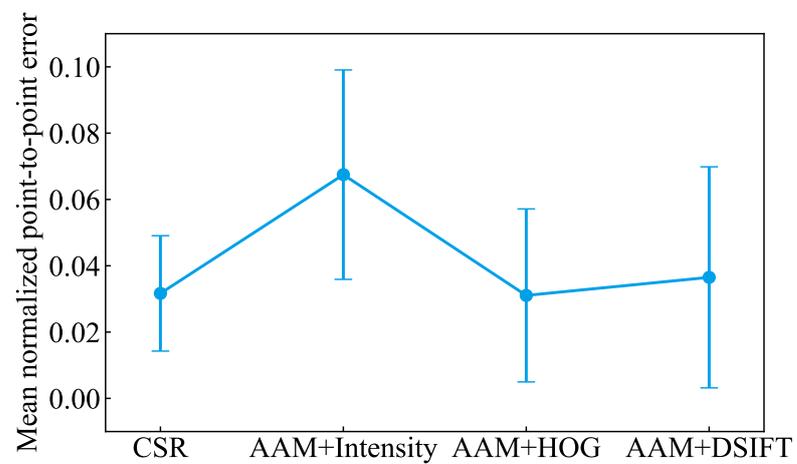


Figure 7. The mean value of NPPE of the test images for each method. Error bars represent standard deviations.

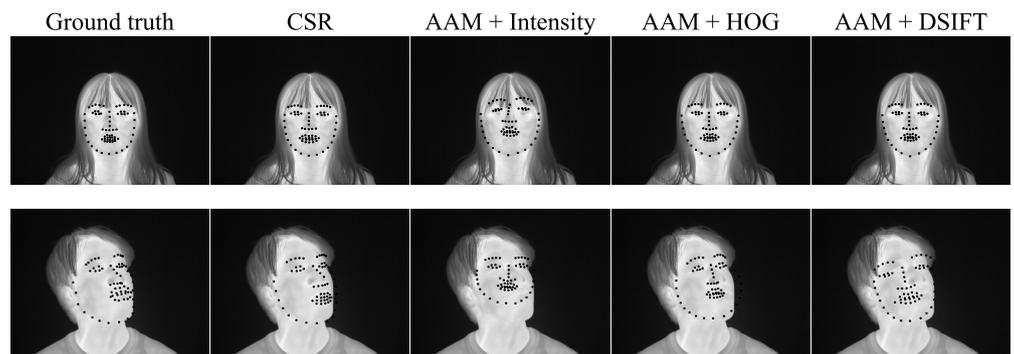


Figure 8. Examples of face alignment using Cascaded Shape Regression (CSR) and Intensity, Dense Scale Invariant Feature Transform (DSIFT), and Histogram of Oriented Gradients (HOG)-based Active Appearance Model (AAM).

Table 2. The frames per second (FPS) of each method.

Method	FPS
CSR	83.3
AAM + Intensity	5.56
AAM + HOG	0.35
AAM + DSIFT	0.64

Figure 9 shows the results for each cascade stage for tree depth = 3, 4, 5, and 10 and Figure 10 shows examples of facial alignment using for tree depth = 3, 4, 5, and 10. When the tree depth was 4, the accuracy of face alignment was the highest. When the tree depth was 5 or 10, the model features were large, and overfitting to the training data occurred, resulting in small accuracy. When the tree depth was 3, the model features were small and under-fitted to the training data, resulting in small accuracy.

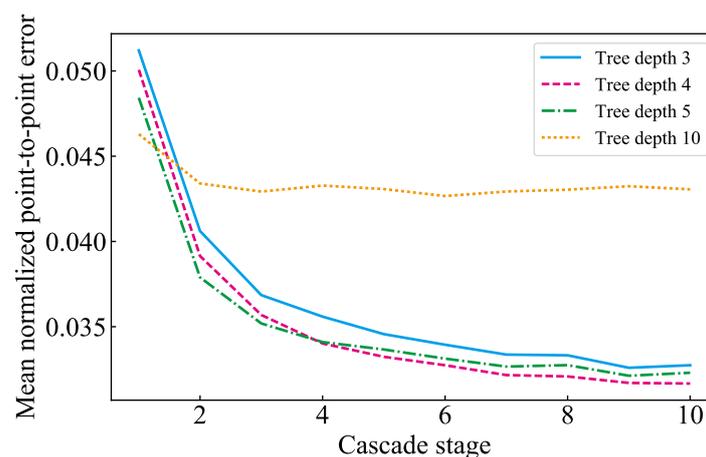


Figure 9. The mean NPPE using CSR for each cascade stage for tree depth = 3, 4, 5, and 10.

Table 3 shows the mean number of pixels and mean NPPE of the face’s width for each distance between the infrared thermography and the subject and Figure 11 shows examples of face alignment for each distance. From Table 3 and Figure 11, the accuracy of the fitting did not decrease with distance. In this experimental condition, differences in distance to infrared thermography and the number of pixels of the face in the image did not affect the face alignment estimates’ accuracy. In thermal measurements, one meter is known to be an excellent standard to assure stable consistency [34]. It is suggested that the face alignment can be done with high accuracy when the distance between the thermography

and the person is between 60 and 120 cm. This satisfies the length of 1 m, which is the right standard length for thermal measurements.

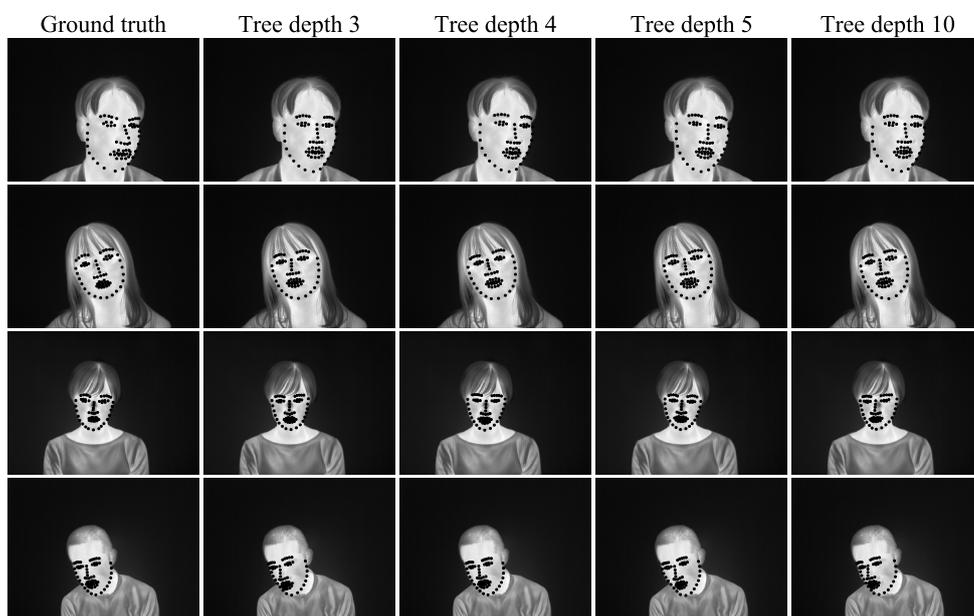


Figure 10. Examples of facial alignment using CSR for each cascade stage for tree depth = 3, 4, 5, and 10.

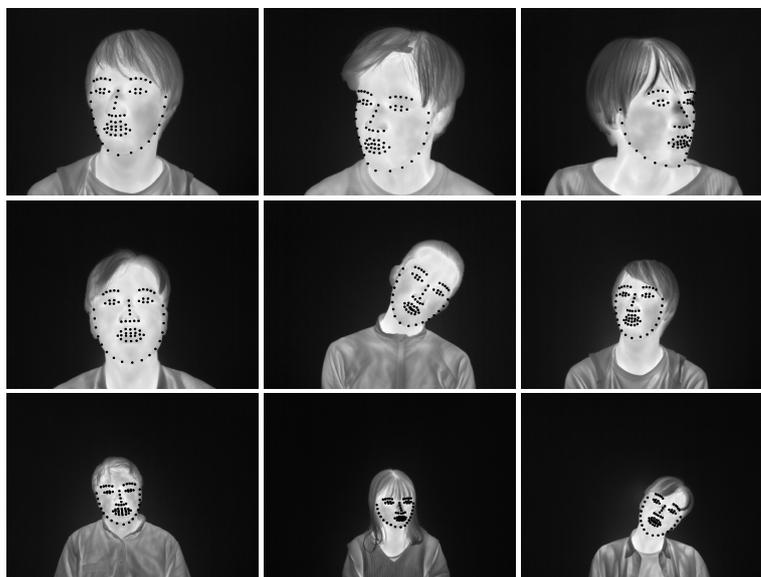


Figure 11. Examples of face alignment for each distance. The distances from the top are 60 cm, 90 cm, and 120 cm.

Table 3. NPPE using CSR for each distance.

Distance/cm	Pixels Per Face Width	NPPE/%
60	184 ± 20	3.25 ± 1.99
90	126 ± 13	3.02 ± 1.71
120	97 ± 10	3.23 ± 1.47

5. Conclusions

As mentioned in the introduction, the objective of this study was to conduct face alignment in thermal infrared images using CSR. CSR is more robust than AAM in face alignment in facial thermal images and can be applied to various types of images. The FPS of face alignment using CSR is more than 80, and it can detect facial landmarks at a high speed. Therefore, facial landmark detection by CSR may be useful for real-world applications. However, the limitation of this study is the small sample size of 609 thermal infrared images, and we have not dealt with thermal infrared images in the wild. In the future, we plan to conduct studies using thermal infrared images of many more varieties and conditions.

Author Contributions: Conceptualization, K.N., T.K., K.O. and A.N.; methodology, K.N., T.K., K.O. and A.N.; software, K.N. and T.K.; validation, K.N. and T.K.; writing—original draft preparation, K.N.; writing—review and editing, K.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the ethics review board.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AAM	Active Appearance Model
CSR	Cascaded Shape Regression
ROI	Regions of Interest
HOG	Histogram of Oriented Gradients
DSIFT	Dense Scale Invariant Feature Transform
NPPE	Normalized Point to Point Error
FPS	Frames Per Second

References

1. Hall, T.; Lie, D.Y.; Nguyen, T.Q.; Mayeda, J.C.; Lie, P.E.; Lopez, J.; Banister, R.E. Non-contact sensor for long-term continuous vital signs monitoring: A review on intelligent phased-array Doppler sensor design. *Sensors* **2017**, *17*, 2632. [[CrossRef](#)]
2. van den Heuvel, C.J.; Ferguson, S.A.; Dawson, D.; Gilbert, S.S. Comparison of digital infrared thermal imaging (DITI) with contact thermometry: Pilot data from a sleep research laboratory. *Physiol. Meas.* **2003**, *24*, 717. [[CrossRef](#)] [[PubMed](#)]
3. Ahmadi, N.; McQuilkin, G.L.; Akhtar, M.W.; Hajsadeghi, F.; Kleis, S.J.; Hecht, H.; Naghvi, M.; Budoff, M. Reproducibility and variability of digital thermal monitoring of vascular reactivity. *Clin. Physiol. Funct. Imaging* **2011**, *31*, 422–428. [[CrossRef](#)] [[PubMed](#)]
4. Zaproudina, N.; Varmavuo, V.; Airaksinen, O.; Närhi, M. Reproducibility of infrared thermography measurements in healthy individuals. *Physiol. Meas.* **2008**, *29*, 515. [[CrossRef](#)]
5. McCoy, M.; Campbell, I.; Stone, P.; Fedorchuk, C.; Wijayawardana, S.; Easley, K. Intra-examiner and inter-examiner reproducibility of paraspinal thermography. *PLoS ONE* **2011**, *6*, e16535. [[CrossRef](#)]
6. Nilsson, A.L. Blood flow, temperature, and heat loss of skin exposed to local radiative and convective cooling. *J. Investig. Dermatol.* **1987**, *88*, 586–593. [[CrossRef](#)]
7. Ioannou, S.; Ebisch, S.; Aureli, T.; Bafunno, D.; Ioannides, H.A.; Cardone, D.; Manini, B.; Romani, G.L.; Galese, V.; Merla, A. The autonomic signature of guilt in children: A thermal infrared imaging study. *PLoS ONE* **2013**, *8*, e79440. [[CrossRef](#)]
8. Ioannou, S.; Gallese, V.; Merla, A. Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology* **2014**, *51*, 951–963. [[CrossRef](#)]
9. Fernández-Cuevas, I.; Marins, J.C.B.; Lastras, J.A.; Carmona, P.M.G.; Cano, S.P.; García-Concepción, M.Á.; Sillero-Quintana, M. Classification of factors influencing the use of infrared thermography in humans: A review. *Phys. Technol.* **2015**, *71*, 28–55. [[CrossRef](#)]

10. Lewis, G.F.; Gatto, R.G.; Porges, S.W. A novel method for extracting respiration rate and relative tidal volume from infrared thermography. *Psychophysiology* **2011**, *48*, 877–887. [[CrossRef](#)] [[PubMed](#)]
11. Gault, T.; Farag, A. A fully automatic method to extract the heart rate from thermal video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Portland, OR, USA, 23–28 June 2013; pp. 336–341.
12. Asano, H.; Sakamoto, N.; Nozawa, A.; Ide, H. Presumption of transient awakening of driver by facial skin temperature. *IEEJ Trans. Electron. Inf. Syst.* **2010**, *130*, 428–432. (In Japanese)
13. Bando, S.; Oiwa, K.; Nozawa, A. Evaluation of dynamics of forehead skin temperature under induced drowsiness. *IEEJ Trans. Electr. Electron. Eng.* **2017**, *12*, S104–S109. [[CrossRef](#)]
14. Adachi, H.; Oiwa, K.; Nozawa, A. Drowsiness level modeling based on facial skin temperature distribution using a convolutional neural network. *IEEJ Trans. Electr. Electron. Eng.* **2019**, *14*, 870–876.
15. Nakanishi, R.; Imai-Matsumura, K. Facial skin temperature decreases in infants with joyful expression. *Infant Behav. Dev.* **2008**, *31*, 137–144. [[CrossRef](#)] [[PubMed](#)]
16. Ebisch, S.J.; Aureli, T.; Bafunno, D.; Cardone, D.; Romani, G.L.; Merla, A. Mother and child in synchrony: Thermal facial imprints of autonomic contagion. *Biol. Psychol.* **2012**, *89*, 123–129. [[CrossRef](#)]
17. Kopaczka, M.; Kolk, R.; Schock, J.; Burkhard, F.; Merhof, D. A thermal infrared face database with facial landmarks and emotion labels. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 1389–1401. [[CrossRef](#)]
18. Engert, V.; Merla, A.; Grant, J.A.; Cardone, D.; Tusche, A.; Singer, T. Exploring the use of thermal infrared imaging in human stress research. *PLoS ONE* **2014**, *9*, e90782. [[CrossRef](#)] [[PubMed](#)]
19. Sonkusare, S.; Ahmedt-Aristizabal, D.; Aburn, M.J.; Nguyen, V.T.; Pang, T.; Frydman, S.; Denman, S.; Fookes, C.; Breakspear, M.; Guo, C.C. Detecting changes in facial temperature induced by a sudden auditory stimulus based on deep learning-assisted face tracking. *Sci. Rep.* **2019**, *9*, 1–11. [[CrossRef](#)] [[PubMed](#)]
20. Masaki, A.; Nagumo, K.; Bikash, L.; Oiwa, K.; Nozawa, A. Anomaly detection in facial skin temperature using variational autoencoder. *Artif. Life Robot.* **2020**, *26*, 122–128. [[CrossRef](#)]
21. Cootes, T.F.; Edwards, G.J.; Taylor, C.J. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 681–685. [[CrossRef](#)]
22. Kopaczka, M.; Acar, K.; Merhof, D. Robust Facial Landmark Detection and Face Tracking in Thermal Infrared Images using Active Appearance Models. In Proceedings of the 11th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016), Rome, Italy, 27–29 February 2016; pp. 150–158.
23. Gross, R.; Matthews, I.; Baker, S. Generic vs. person specific active appearance models. *Image Vis. Comput.* **2005**, *23*, 1080–1093. [[CrossRef](#)]
24. Dollár, P.; Welinder, P.; Perona, P. Cascaded pose regression. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010), San Francisco, CA, USA, 13–18 June 2010; pp. 1078–1085.
25. Cao, X.; Wei, Y.; Wen, F.; Sun, J. Face alignment by explicit shape regression. *Int. J. Comput. Vis.* **2014**, *107*, 177–190. [[CrossRef](#)]
26. Ren, S.; Cao, X.; Wei, Y.; Sun, J. Face alignment at 3000 fps via regressing local binary features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), Columbus, OH, USA, 24–27 June 2014; pp. 1685–1692.
27. Kazemi, V.; Sullivan, J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), Columbus, OH, USA, 24–27 June 2014; pp. 1867–1874.
28. Burgos-Artizzu, X.P.; Perona, P.; Dollár, P. Robust face landmark estimation under occlusion. In Proceedings of the IEEE International Conference on Computer Vision (ICCV 2013), Sydney, Australia, 1–8 December 2013; pp. 1513–1520.
29. Steketee, J. Spectral emissivity of skin and pericardium. *Phys. Med. Biol.* **1973**, *18*, 686. [[CrossRef](#)] [[PubMed](#)]
30. Sagonas, C.; Tzimiropoulos, G.; Zafeiriou, S.; Pantic, M. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCVW), Sydney, Australia, 2–8 December 2013.
31. Zhu, X.; Ramanan, D. Face detection, pose estimation, and landmark localization in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, Rhode Island, 16–21 June 2012; pp. 2879–2886.
32. Marciniak, T.; Chmielewska, A.; Weychan, R.; Parzych, M.; Dabrowski, A. Influence of low resolution of images on reliability of face detection and recognition. *Multimed. Tools Appl.* **2015**, *74*, 4329–4349. [[CrossRef](#)]
33. Zafeiriou, S.; Trigeorgis, G.; Chrysos, G.; Deng, J.; Shen, J. The menpo facial landmark localisation challenge: A step towards the solution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–27 July 2017; pp. 170–179.
34. Ivanitsky, G.R.; Khizhnyak, E.P.; Deev, A.A.; Khizhnyak, L.N. Thermal imaging in medicine: A comparative study of infrared systems operating in wavelength ranges of 3–5 and 8–12 μm as applied to diagnosis. *Dokl. Biochem. Biophys.* **2006**, *407*, 59–63. [[CrossRef](#)]