A Novel Fully Annotated Thermal Infrared Face Dataset: Recorded in Various Environment Conditions and Distances From The Camera

Roshanak Ashrafi\textsuperscript{a}, Mona Azarbayjani\textsuperscript{a,b}, Hamed Tabkhi\textsuperscript{c}

University of North Carolina at Charlotte

\textsuperscript{a}Department of Civil and Environmental Engineering, The William States Lee College of Engineering
\textsuperscript{b}Department of Architecture, The College of Arts and Architecture
\textsuperscript{c}Department of Electrical and Computer Engineering, The William States Lee College of Engineering

Abstract

Facial thermography is one of the most popular research areas in infrared thermal imaging, with diverse applications in medical, surveillance, and environmental monitoring. However, in contrast to facial imagery in the visual spectrum, the lack of public datasets on facial thermal images is an obstacle to research improvement in this area. Thermal face imagery is still a relatively new research area to be evaluated and studied in different domains. The current thermal face datasets are limited in regards to the subjects’ distance from the camera, the ambient temperature variation, and facial landmarks’ localization. We address these gaps by presenting a new facial thermography dataset. This article makes two main contributions to the body of knowledge. First, it presents a comprehensive review and comparison of current public datasets in facial thermography. Second, it introduces and studies a novel public dataset on facial thermography, which we call it Charlotte-ThermalFace. Charlotte-ThermalFace contains more than 10000 infrared thermal images in varying thermal conditions, several distances from the camera, and different head positions. The data is fully annotated with the facial landmarks, ambient temperature, relative humidity, the air speed of the room, distance to the camera, and subject thermal sensation at the time of capturing each image. Our dataset is the first publicly available thermal dataset annotated with the thermal sensation of each subject in different thermal conditions and one of the few datasets in raw 16-bit format. Finally, we present a preliminary analysis of the dataset to show the applicability and importance of the thermal conditions in facial thermography. The full dataset, including annotations, are freely available for research purpose at https://github.com/TeCSAR-UNCC/UNCC-ThermalFace

Keywords: Facial Thermography, Facial Dataset, Face Thermal Dataset, Thermal Comfort, Facial Landmarks, Physiological measurements

1. Introduction

The emergence of thermal imaging techniques has provided an excellent opportunity for contactless data gathering without interruption to occupants’ activities. As elevated body temperature is an important indicator of a possible underlying physiological process, thermal imaging is a great non-intrusive tool for presenting that biological state. The non-invasive nature of thermal cameras has resulted in extensive utilization of these devices for detecting and diagnostic purposes in several areas of medicine \cite{1} \cite{2} \cite{3}, and none-medical research such as fire safety \cite{4} \cite{5} \cite{6}, transportation \cite{7} \cite{8} \cite{9}, and building construction \cite{10} \cite{11} \cite{12}.

Recent accelerated innovations in thermal imaging devices make it possible to utilize lower-priced thermal cameras to collect high-quality information that can improve several aspects of our lives. Facial thermography, in particular, is one of the most widely studied areas due to its proven practical applications. Facial

Email address: rashrafi@uncc.edu (Roshanak Ashrafi)
thermal imagery has been successfully studied as an indicator of human identity [13, 14], emotions [15, 16, 17], and comfort [18, 19, 20] in several research. The current pandemic has also highlighted the facial thermography potential in the non-intrusive evaluation of human conditions more than before [21, 22].

While facial thermography has gained lots of recent attention, this is still a new research area that needs to be evaluated and studied in different domains. The lack of public datasets of facial thermal images is a significant obstacle to research improvement in this area. Currently, there are lots of RGB facial datasets available in diverse conditions [23], but the facial infrared image databases are limited and need improvement for several reasons. First, due to the lower quality of older thermal cameras, several current thermal datasets are of low resolution. They are not appropriate for use in sensitive areas, such as health-related applications [24]. Second, most of the existing public datasets are not in the original raw format and are converted frames, which results in the loss of important information [25]. The original thermal images are created utilizing 14-bit or 16-bit radiometric data that comprises information about the heat flux or temperature of each pixel. However, the majority of these datasets are transformed to lower bit resolutions, such as 8-bit RGB files, to make them smaller and more compatible with common imaging software. As a result of this process, temperature data is lost and cannot be retrieved later. Third, the current public datasets lack diversity in several aspects, such as face resolution and environmental properties. Most of the available datasets are appropriate for facial recognition purposes with no variation in environmental properties such as air temperature, relative humidity, and air speed. None of the current public datasets include data on controlled thermal variations, which is one of our top objectives in collecting this dataset. While some of the current datasets were recorded in uncontrolled thermal circumstances, resulting in temperature fluctuation, yet there is no information about the ambient temperature at the time of thermography recordings. In addition, there are no publicly available datasets with varying distances from the camera [26], while several studies have indicated the importance of distance in the infrared thermography readings [27, 28, 29]. In field thermography applications, when precision is required, the temperature variance induced by distance is substantial and needs to be considered. Fourth, while some recent research has demonstrated the great potential of using thermal cameras in smart buildings to predict occupant thermal sensations, none of them have made their datasets available for use in other projects. The improvements in thermal imaging technology enable the collection of data for use in the control systems of our smart buildings. Temperature data has been shown to be a viable variable for predicting occupant thermal preferences and controlling indoor environmental systems. While some of the current datasets were recorded in uncontrolled thermal circumstances, resulting in temperature fluctuation, there is no information about the ambient temperature at the time of thermography recordings. However, there are currently no public datasets available that include the thermal sensation of the subjects in different thermal conditions. Finally, the majority of thermal datasets are either missing any landmark annotations or include just a few facial landmarks and the bounding box. Lack of manual facial landmark annotation makes the application of this datasets limited. Facial landmarks provide an extra supervisory signal and assist in the recognition of complex cases, in addition to making it possible to align the face during the face recognition process.

To address some of the limitations mentioned above, we present the Charlotte-ThermalFace dataset with unique properties. As Figure 2 shows through some sample images, our dataset includes thermal images of the same subjects in various thermal and physical conditions, including the environmental temperature, distance from the camera, and head position. In contrast to the existing datasets that are recorded mostly at a fixed distance from the camera, our dataset includes a variation in distance, which has resulted in thermal images with several different resolutions. In addition, each subject is recorded in at least four different thermal conditions at all the specified distances. We have included the environmental temperatures data in the dataset, which we present in this sample figure for each image frame. The first column of the figure shows the lowest temperature range images, which increase as we go to the left side columns. We can also observe in this figure that varying distances from the camera have resulted in a dataset with different thermal resolutions. We recorded the data frames at 10 relative distances from the camera, of which this sample figure includes the three main ones (closest, middle, and farthest). In addition to environmental data annotations, all the images are manually annotated with extensive facial landmark localization, which include 72 points for full or semi-profile positions, and 42 points for profile faces [30]. To avoid missing temperature data, our dataset is published in the original high accuracy recording format of Flir cameras, which is 16-bit.
Figure 1: A Sample of Included Data Frames in Different Thermal Conditions

This capability stores the thermal data for each pixel in the frames, and the users can read the temperature of each pixel independently. Based on the aforementioned information, the Charlotte-ThermalFace dataset is available publicly with the following main characteristics:

- We have captured approximately 10,000 infrared thermal images in varying thermal conditions, several distances from the camera, and changing head positions. We have also controlled the air temperature to change from 20.5°C (69°F) to 26.5°C (80°F). Images are available in four different temperatures, 10 relative distances from the camera, starting at 1m (3.3 ft) to 6.6m (21.6 ft), and 25 head positions.
- The first public facial thermal dataset annotated with the environmental properties including air temperature, relative humidity, air speed, distance from the camera, and subjective thermal sensation of each person at the time.
- All the images are manually annotated with 72 facial landmarks.
- We are publishing the data in the original 16-bit radiometric TemperatureLinear format, which has the thermal value of each pixel.

This study has two main contributions. First, it presents a comprehensive comparison of the current public datasets in facial thermography. Second, it introduces the Charlotte-ThermalFace public dataset on facial thermography with a brief investigation. The rest of this paper has three main sections. In section 2. literature Review, first, we look into the applications of facial thermography in different domains. Then we study and compare the existing publicly available thermal datasets with a brief description of each. In Section 3, we introduce our developed dataset and how we address some of the existing gaps by publishing our dataset. Section 4. provides a preliminary analysis of the dataset to show its applicability for future projects.
2. Literature Review

2.0.1. Facial Thermography Applications

A significant decrease in the prices of thermal cameras and their improved quality and resolution are paving the way for the utilization of these sensors in more research and real-world applications. The applications of biometric thermal images are defined in four categories: detection, monitoring, and recognition/identification [32]. In this section, we have looked into the applications of facial thermography in these three major areas in detail.

Face Detection:

While variations in lighting conditions can easily degrade the performance of face detection algorithms, infrared thermography provides accurate results even in complete darkness. This quality has resulted in better performance of thermal images in comparison to visual images in pedestrian detection for autonomous vehicles. Their results show the Missing Rate (MR) when using Histogram of Oriented Gradients (HOG) features for visual images is %73, which decreases to %50 by using thermal image [33]. In regards to face detection, it is demonstrated that when employing visual photos for face identification, non-uniform illumination and fake faces may easily make cascade classifiers inoperable. However, the thermal frames do not have this downside. Through utilizing thermal images, machine learning-based face detection algorithms use classifiers such as AdaBoost and Support Vector Machine for considering facial and non-facial patches as positive and negative regions, which have been successful [34]. Additionally, other research teams have conducted field experiments to evaluate and compare the proposed approaches to face detection using RGB photos and have demonstrated the benefits of utilizing thermal images to detect faces [35]. Researchers have shown promising results using Haar-like features combined with a cascade of boosted tree classifiers [36], and Faster R-CNN and YOLO [37] as viable strategies for face detection in the thermal domain. It is important to mention that the accuracy of these face detection algorithms was shown to be significantly reduced when pictures are flipped and rotated. [38].

Face Monitoring:

The two major applications of facial monitoring is in medicine and building science, which are described below:

Medical Facial Thermography: The feasibility of using thermal scanners for reading skin temperature instead of traditional oral or rectal thermometers has been studied in several research [39][40]. The consensus, also supported by recent data, indicates that the temperature of the eye area has the highest temperature on the head, which makes it suitable for detecting fever [41]. It has also been shown that the side temple, inner eye area, and ear have shown the best correlation of internal body temperature [42]. At the same time, because these two areas are most affected by environmental temperature [43], cheeks and nose temperature are promising measured facial areas for calibrating and offsetting the impacts of environmental factors. Thermal imaging has also been used for respiration rate detection by monitoring the temperature change frequency in the nasal area [44]. This approach can eliminate the adverse effect of the lighting condition; however, the relatively lower resolution of thermal cameras requires lower distances to the camera even in the recent research [45]. The combination of RGB and thermal imaging has led to higher accuracy of respiratory rate monitoring [46][47][21]. Due to the COVID-19 pandemic, the accuracy of infrared thermography in the medical sector has become more critical than ever. Some early studies have shown a few successful cases of using thermal cameras to detect febrile patients worldwide [48, 49, 50]. In this regard, the U.S. Food and Drug Administration (FDA) has enforced policies for Telethermograph systems during the COVID-19 pandemic based on a recent international standard (IEC 80601-2-59) [51].

Environmental Monitoring: Recent research has also proved thermal imaging to be a successful indicator of the building’s occupants’ thermal preferences. The infrared cameras can be installed at a distance from the occupant and capture the skin temperature by reading the pixel values of the desired facial regions, proving the feasibility of this technique with 94% -95 % accuracy when using FLIR A655 [52]. A real-time feedback system using FlirA35 thermal cameras was developed in 2018, analyzing both face temperature and occupants’ position [53]. A lower-cost could also replace the previously mentioned expensive cameras and smaller infrared camera, Flir Lepton, with an acceptable accuracy of 85% for predicting the skin temperature [54]. The accuracy of thermal comfort prediction was compared between different monitoring
devices, including air temperature sensors, wearable skin temperature wristbands, and thermal infrared cameras. The findings highlight a slight improvement in the prediction accuracy by adding physiological sensors to the environmental sensors. However, it questions the efficiency of using physiological sensors for this slight accuracy increase (%3-%4) [55]. In another recent study in this area, Li et al. had successful experience monitoring and recording the skin temperature of two occupants simultaneously with two thermal camera nodes, while each camera captured some parts of the faces [56].

**Facial recognition:**

Facial recognition is one of the most studied areas of facial thermography [32]. Face recognition in the thermal spectrum has gained more attention after showing successful results of Long Wave Infrared Imaging (LWIR), which was even superior to the visible spectrum face recognition results [57]. The research in this area has been ongoing ever since through utilizing several image processing methods, which have proved to be promising even in the outdoor environments, where illumination and face alignment vary substantially [58]. Multi-modal face recognition was also introduced in 2005 by utilizing both thermal and RGB images to develop a three-dimensional model of the face, and its thermal texture map [59]. Research has performed experiments involving data fusion of multi-spectral imagery and was able to confirm the improvement in face recognition rate from 0.789 to 0.870 as a result of image fusion [60]. Furthermore, dual camera setups were designed to be used in dynamic illumination conditions to switch between the visible and infrared spectrum and select the most confident result in different lighting situations [61].

Emotion and expression detection as two other application of face recognition, which have also shown promising results [62], for detecting conditions such as anxiety, fear, and alertness [63]. The main reason for effectiveness of thermal images for this application is variations in the skin temperature while expressing different emotions, which had made thermal data an essential source of additional information to improve in the evaluation of facial expressions and emotions [64]. This quality has helped with creation and robot, capable of communicating with children and recognition of their emotions including disgust, fear, happiness, sadness, and surprise. These emotions could be recognized with an 85% accuracy rate using low-cost hardware and low-cost approaches for visual and thermal image processing [65]. Another research has shown the increased expression prediction accuracy from %89.5 when using visual images to %93.7 by using thermal frames. Multi-modal fusion through utilizing thermal, visual, and depth domains have also been proved to be successful in emotion recognition through utilizing more sophisticated late fusion approaches, such as fuzzy inference systems and Bayesian inference [23]. Intoxication detection is another application of facial thermography in the facial recognition category. Alcohol consumption results in abnormally dilated blood vessels and increased blood pressure. In the facial area, this biological behavior shows itself by increasing the temperature around the nose, forehead, and eye area [66]. Despite several challenges, recent research has shown an average percentage accuracy of %99.63 in identifying drunk individuals through thermal images [67].

In conclusion, we have shown that thermal images can be of great importance in several domains. Although both visual and thermal images are influenced by the change in lighting and thermal conditions, the intensity of the change varies extensively based on the cameras’ dynamic range. As a consequence, when using an visual camera, the change in pixel values in the facial regions may be close to 100% of the dynamic range of the sensor. However, the thermal domain changes are much less than the dynamic range that a typical thermal camera can capture, which will result in the higher consistency of thermal images, even in varying thermal conditions [35].
2.1. Existing Facial Thermal Datasets

Increased application of facial thermography in the mentioned sectors has highlighted the need for a comprehensive database for further studies. This section presents the current publicly available thermal face datasets and compares them to their key characteristics and limitations. This information is also available in Table 2.1 for better comparison.

**IRIS**: The IRIS (Imaging, Robotics, and Intelligent Systems Lab) database contains 4228 pairs of images in both thermal and visible domains from 32 subjects, captured simultaneously. Similar to the Equinox database, each subject demonstrated three facial expressions: smiling, frowning, and surprised in five different illumination settings. This database also includes a 4-second video in each scenario at 10fps as the subjects pronounce the vowels. Each scenario is captured in 11 different positions by rotating the camera 36 degrees for each modality, at a fixed distance of 183 cm from the subject [68].

**IRIS-M3**: This upgrade of the IRIS lab database contains 2624 pairs of images from 82 subjects in multiband spectrum information, including one thermal and 25 visual bands per participant, in both outdoor and indoor environments. Another significant contribution of this upgrade is a diverse ethnic collection of Caucasian, Asian, Asian Indian, and African participants, with 24% female and 76% male participants. However, the database includes only one variation in facial expression and one frontal position toward the camera, while the distance from the camera is 1.2 meters for all the images [60].

**UND**: The University of Notre Dame database had published this database, which includes 2492 image pairs from 241 subjects, which is one of the most significant subject populations in the thermal/visual datasets. Each subject demonstrates two facial expressions of smiling and a natural face under three dif-
| Dataset | Thermal | RGB | Subjects | P | I | E | Resolution | Distance (cm) | Landmarks | Size(images) |
|---------|---------|-----|----------|---|---|---|------------|--------------|------------|-------------|
| IRIS [68] | ✓ | ✓ | 32 | 11 | 5 | 3 | 320 × 240 | 183 | - | 4228 |
| IRIS-M3 [69] | ✓ | ✓ | 82 | 25 | S* | 1 | 640 × 480 | 120 | - | 2624 |
| UND [69] | ✓ | ✓ | 241 | 1 | 3 | 3 | 320 × 240 | ? | - | 2624 |
| UH [70] | ✓ | | 138 | 5 | 1 | I-5 | 640 × 512 | ? | - | 7590 |
| FSU [71] | ✓ | | 10 | 0-20 | 1 | 0-20 | 320 × 240 | ? | - | 234 |
| Carl [13] | ✓ | | 41 | 1 | 3 | 1 | 160 × 120 | 135 | - | 7.380 |
| NVIE [72] | ✓ | ✓ | 215 | 1 | 3 | 7 | 320 × 240 | 75 | - | 234 |
| KTFE [74] | ✓ | ✓ | 26 | 1 | 1 | 7 | 320 × 240 | 85 | - | 126 GB |
| ARL-M3 [73] | ✓ | ✓ | 60 | 1 | 1 | video | 640 × 480 | 250 | 6 | ? |
| MMFDV1 [73] | ✓ | ✓ | 51 | 1 | 1 | video | 640 × 480 | 250 | 6 | ? |
| ARL-MF [74] | ✓ | ✓ | 236 | 1 | 1 | 7 | 640 × 512 | 100 | - | ? |
| Eurocom [75] | ✓ | ✓ | 50 | 4 | 5 | 7 | 160 × 120 | 150 | - | 4200 |
| Tuft [76] | ✓ | ✓ | 113 | 9 | 2 | 5 | 336 × 256 | 150 | - | 10,000 |
| RWTH [77] | ✓ | ✓ | 90 | 9 | 2 | 8 | 1024 × 786 | 90 | 68 | 10,000 |
| ARL-VTF [78] | ✓ | ✓ | 395 | 3 | 1 | 2 | 640 × 512 | 210 | 6 | 500,000 |
| Sejong-A [78] | ✓ | ✓ | 30 | 1 | 1 | 1-3 | 768 × 756 | 200 | - | 1,500 |
| Sejong-B [78] | ✓ | ✓ | 70 | 1 | 5-15 | 1 | 13 | 768 × 756 | 200 | - | 23,000 |
| I2BVSD [79] | ✓ | ✓ | 75 | 1 | 1 | 7 | 720 × 576 | ? | - | 681 |
| Sober-Drunk | ✓ | | 41 | 1 | 1 | 2 | 128 × 160 | 30 | - | 4,100 |
| Database [80] | ✓ | | 46 | 1 | 1 | 4 | 640 × 480 | ? | - | 11,500 |
| PUCV- [81] | ✓ | ✓ | 147 | S* | S* video | 464 × 348 | S* | 5 | 9,982 |
| TFW [82] | ✓ | ✓ | 142 | 9 | 1 | S* | 464 × 348 | 100 | - | 4,581,595 |
| SpeakingFaces [83] | ✓ | ✓ | 20 | 3 | S* | 3 | 320 × 240 | ? | - | 21,676 |

Table 1: Public Available Datasets, P: Position, I: Illumination, E/D: Expression/Disguise,S*: Several

Different indoor lighting conditions. Another unique feature of this dataset is capturing images in 4 sessions throughout the month. However, there are only four image pairs per subject [69].

**UH**: The University of Houston database contains 7590 images of 138 subjects, captured in the Mid Wave MW thermal IR domain. There are 55 images of each subject in the database, demonstrating different facial expressions and arbitrary facial positions [70].

**FSU**: The State University of Florida database contains 234 images from 9 subjects in thermal infrared in 7-14 µm spectral range information, which includes 25 visual bands and one thermal per participant. The images include varying angles and facial expressions for all the subjects. The dataset is at 320 × 240 resolution and in 8-bit BMP format [71].

**Carl**: This dataset includes a total of 7,380 images that were recorded simultaneously with visible, near-infrared, and thermal sensors. 41 subjects participated in this study under three different illuminating conditions: natural, infrared, and artificial. The snapshots are all in frontal face position with neutral facial expressions. The Thermographic camera TESTO 880-3, equipped with an uncooled detector, captured both thermal and visible images. The images are captured in four separate sessions in which were two days apart [13].

**NVIE**: The Natural Visible and Infrared Facial Expression database mainly focuses on capturing different facial expressions among several subjects. The dataset is from 215 subjects and includes two sub-databases: (1) Spontaneous and (2) Pose. The spontaneous dataset contains sequences from starting an expression to the final frame, and the Pose dataset contains only the last expression in both the visual and
thermal spectrum. The researchers have captured all the images under three different indoor illumination settings.

**KTFE:** Kotani Thermal Facial Emotion dataset includes simultaneous thermal and visible images in seven spontaneous emotions, including neutral, anger, happiness, sadness, fear, disgust, and surprise. 26 subjects participated in the study, with an age range of 11-years-old to 32-years-old. An infrared camera, the NEC R300, was used for capturing both thermal and visual videos. The dataset includes 126 gigabytes of visible and thermal facial emotion facial expression data frames. This database is one of the few datasets that mentions the recording room air temperature, kept between 24 °C and 26 °C.

**ARL-MMFD:** Army Research Laboratory Multi-Modal Face Database was recorded simultaneously by a long-wave infrared camera and a visible spectrum camera. The utilized polymetric sensor is capable of recording geometric and textural facial details. The dataset included 60 subjects in the first published version that was then extended to 111 subjects. The researchers have used LWIR polarimetric for capturing images in both datasets, which are named Volume1 and Volume2. The expression change in the subjects’ faces was created by counting out loud numerically. The first volume dataset is the only dataset collected at three different distances from the camera, which are 2 m, 5 m, and 7.5 m. However, the second volume is captured only at a single range of 2.5 m from the camera.

**ULFMT:** Université Laval Face Motion and Time-Lapse Video Database is recorded within four years from 238 subjects. This experiment includes different facial poses and expressions, ethnicities, ages, and sexes. The researchers have recorded the images in 4 different spectrum ranges with different cameras, including a Jenoptik camera for LWIR, a Phoenix Indigo IR camera produced by FLIR for MWIR, a CMOS made by Goodrich for SWIR, and a standard CCD made by Much for the NIR/Visible spectrum. The subjects had changed their facial expression arbitrarily and changed their head position from full-frontal face to complete profile for the video frames captured at 30 fps for 10 seconds.

**Eurocom:** The dataset includes a total of 4200 images from 50 subjects, captured simultaneously in the visible and thermal spectrum. The dataset contains six illumination settings: ambient light, rim light, key light, fill light, all lights on, all lights off; seven expressions: neutral, happy, angry, sad, surprised, blinking, yawning; four head positions: up, down, right at 30°, left at 30°; and occlusion: eyeglasses, sunglasses, cap, mouth occluded by hand, eye occluded by hand. The thermostat temperature was at the average temperature of 25°C for the test room. The thermal camera is a Flir Duo with an uncooled VoX microbolometer and a thermal resolution of 160x120 pixels.

**Tufts:** The Tufts University database focuses on capturing images in various modalities from subjects in 15 different countries, genders, ages (4–70 years old), and ethnicities. A total of 10,000 images from 113 subjects had participated in the dataset. Images are under different scenarios of simultaneous visible and thermal images, near-infrared (NIR) images, a recorded video, a computerized facial sketch using the FACES software, and 3D images of all the subjects. The researchers have utilized FLIR Vue Pro camera for recording the thermal images. The facial expressions for each subject are neutral, with a smile, eyes closed, an exaggerated shocked expression, and wearing sunglasses.

**RWTH:** The RWTH Aachen University database contains 2935 images in the thermal domain from 90 subjects. The subjects’ distance from the camera is 90 centimeters. Each subject demonstrates six facial expressions of contempt, disgust, anger, fear, surprise, sadness, happiness, and a neutral face. The neutral scenario is captured in 9 different positions by rotating the head in vertical and horizontal positions. The database is a manually annotated dataset with 68 facial landmark points, emotions, and positions.

**ARL-VTF:** DEVCOM Army Research Laboratory Visible-Thermal Face Dataset (ARL-VTF) presents 500,000 images from 395 subjects, captured simultaneously with spectrum information. The dataset was recorded with a long wave infrared LWIR camera and three visible spectrum cameras. The FLIR Boson uncooled VOX microbolometer camera captured the thermal images with a spectral band of 7.5 µm to 13.5 µm. The database includes a variation in facial expressions created by counting out loud the numbers and facial positions created by rotating the head from left to right. The distance from the camera was 2.1 meters for all the images. The database is annotated with face bounding boxes and 6 points of facial landmarks, including left eye center, right eye center, the base of the nose, left mouth corner, right mouth corner, and center of the mouth.

**Sejong:** Sejong is a recent multi-modal disguise face database, which contains images recorded in four
modalities including visible, infrared, thermal, and visible-plus-infrared. The database contains two subsets and subset B were captured one year after the first one (subset A). Subset-A has 30 participants (16 men and 14 females) and the total of 1,500 images, whereas Subset-B contains 70 subjects (44 males and 26 females) with the total number of 23,000 images. The highlight of this database is the add-on images that were captured in all four modalities. In the thermal images as they are not a part of the human body, disguise add-ons have a lower temperature than human skin and so seem darker than human skin or hair. In addition, subset B contains five to fifteen different poses for each subject. For the recordings the camera box were placed in the fixed distance of two meters from the subject and the room temperature was kept at 25 ± 5 °C. The thermal images were captured by a Therm-App camera with the resolution of 768x756 pixels [78].

I2BVSD: The IIITD In and Beyond Visible Spectrum Disguise (I2BVSD) face database includes disguised/obfuscated face images in both thermal and visual spectrum. The database’s disguise modifications are listed as follows. (1) No disguise: unambiguous appearance, (2) Hairstyles: wigs come in a variety of styles and hues. (3) Beard and mustache: many kinds of beards and mustaches (4) Eyeglasses: sunglasses and spectacles (5) Cap and hat: several types of caps, turbans, veils, and bandanas (6) Mask-related variation: disposable mask; and (7) Multiple variations: A mix of disguised accessories. The dataset includes 75 subjects, with one neutral and at least five disguised images for each subject. There are 681 images in each spectrum, which includes 6 to 10 images for each individual. The resolution of thermal images is 720 x 576 pixels.

Sober-Drunk Database: The dataset include images of drunk and sober individuals in thermal modality and was created at Electronics Laboratory, Physics Department, University of Patras, Greece. The images were captured with the FLIR Thermo Vision Micron A10 Model infrared camera with a resolution of 128x160 pixels. The dataset contains thermal images of both drunk and sober states for each individual. 41 individuals were included in this experimental method, 31 males and 10 females. The recording distance was 30 centimeters from the camera. The first 50 frames were acquired for each individual immediately before to initiating alcohol intake, and the second 50 frames were acquired 30 minutes after the fourth glass of wine was consumed. A total of 100 frames were captured for each participant, resulting in a database of 4100 pictures [80].

PUCV-DTF: The Pontificia Universidad Católica de Valparaíso-Drunk Thermal Face database is also a drunk classification dataset. 40 men and 6 women, with an average age of 24 years, have participated in this study. The images were captured with the FLIR TAU 2 thermal camera, which had a resolution of 640x480 pixels, a thermal sensitivity of 50 mK, and a wavelength range of 7.5–13.5 m. The dataset includes 250 images for each subject, which results in 11,500 total images. The room temperature and distance to the camera were not mentioned in the dataset publication [81].

TFW: Thermal Faces in the Wild (TFW) dataset contains thermal images of people both in outdoor and indoor environments with manually labeled bounding boxes and five-point facial landmarks (eye centers, nose tips, and mouth corners). The outdoor images were recorded in a variety of weather conditions, which included thermal photos acquired in both bright and cloudy conditions. The dataset has 9,982 frames and 16,509 labelled faces from 147 subjects. 5,112 images (5,112 faces) are recorded in a controlled setting and 4,870 image frames (11,397 faces) are collected outdoor. Occlusion, head position changes, various scales, face masks, diverse settings, and weather conditions are all part of the set. The FLIR T540 thermal camera, with a resolution of 464 x 348 pixels, a waveband of 7.5–14 m, a field of view of 24, and an iron color palette, was used to capture the datasets [82].

SpeakingFaces: SpeakingFaces dataset combines high-resolution thermal and visual spectral picture streams of fully-framed faces with audio recordings of each participant speaking about 100 sentences. The dataset includes 142 individuals, in nearly 13,000 instances of synced data from 9 different positions (3.8 TB). During a single trial, each individual participated in two different sessions. Subjects were requested to stay quiet and steady throughout the first session, while the operator captured visual and thermal video feeds from a sequence of nine collecting angles. The other session had the subject reading a sequence of words while visual, thermal, and audio data were recorded from the same nine camera angles. The images were recorded with FLIR T540 thermal camera with a resolution of 464x348 pixels. Subjects were seated in a distance of one meter from the camera, and the room temperature was set at 25 degrees Celsius [83].
TIV: Terravic Facial IR Database contains facial images in both thermal and visual spectrum. This database contains twenty individuals, each of whom has a unique set of frames with a variety of modifications, including front, left, and right orientations, indoor/outdoor setting and using glasses or a hat. It contains 21676 thermal facial photos of 20 different individuals. The image frames were recorded with Raytheon L-3 Thermal-Eye 2000AS with a resolution of 320x240 pixels and delivered in the 8-bit gray scale JPEG format.

As presented, most of the current datasets include a diverse number of facial expressions, emotions, and head positions, which makes them suitable for facial recognition purposes. However, there is a substantial lack of facial landmark annotation in these datasets. Only RWTH and ARL-VTF have included manual facial annotations, while the ARL-VTF dataset contains only 6 main landmark points. Landmark detection is one of the most important tasks required for extracting biometric data from face thermal images. One main approach to detecting facial landmarks is utilizing visual images, which requires calibrating an RGB camera and thermal camera together. Other approaches rely on the pixel value differences between the individual’s faces and their backgrounds. However, this approach does not work properly when the subject’s head position changes. As presented in RWTH dataset analysis, learning-based methods improve the facial landmark detection accuracy. However, there is currently only the RWTH dataset that includes a complete facial landmark annotation for 2935 images, which was inspired by similarly annotated datasets such as HELEN. However, the data is gathered at a fixed distance of 90 centimeters from the camera and is published in 8 bit PNG format, which makes it loose some thermal information in the conversion process. The Charlotte-ThermalFace dataset provides all the images in a 16 bit TIFF format. Furthermore, none of the currently available datasets covers controlled thermal variation, which results in temperature variation, but there is no information about the ambient temperature while thermography. Another value of our dataset is the ambient temperature and relative humidity variation, which is not been considered previously. We have recorded the thermal images for each subject in at least four different room air temperatures and included that information in the dataset annotations.

In addition, the only dataset that has included distance in the facial thermography variation is the recent ARL-VTF dataset that has been developed by the DEVCOM Army Research Laboratory in 2021 and includes images at 3 different distances from the camera. However, this dataset does not include any controlled ambient temperature variation, and only 6 facial landmarks are tagged.

Finally, as most of the presented datasets were recorded for computational purposes, and the human aspects such as the thermal sensation of the subjects were not considered. With the recent advances in thermography and its application in smart buildings, thermal imaging is going to be part of the future infrastructures. Our dataset would be a great help for researchers in this area.

3. Charlotte-ThermalFace: UNC Charlotte Thermal Face Database Overview

As discussed in the previous section, the current facial thermography datasets are still very limited in quality and quantity and need improvement in many aspects. The facial thermal data is provided based on variation in environmental temperature, distance, and head position. We have gathered 10000 infrared thermal images from 10 healthy subjects in varying thermal conditions, several distances from the camera, and different head positions. We have annotated the data with the ambient temperature, relative humidity, and air speed of the room at the exact time of capturing each image. In addition, 72 facial points are manually marked and added to the annotations.

UNCC-ThermalFace is the first publicly available thermal database annotated with the thermal sensation of each subject in different thermal conditions. It also enables radiometric enabled raw data frames. The radiometric option defines retaining the electromagnetic radiation in the data frame files. By enabling this option, our dataset provides information about the captured radiance in each pixel of the recorded images. All the data is in the original 16-bit radiometric raw format, with a thermal value for each pixel. Flir A700 has recorded the frames, which is one of the most recent Flir Systems’ cameras. The thermal sensor is an uncooled microbolometer with the temperature range of 0°C to 650°C and accuracy of ±2°C (±3.6°F) or ±2% of reading for ambient temperature 15°C to 35°C. The images are captured at 10 relative
distances to the camera for each temperature range, as is shown in Figure 3. The original resolution of the thermal sensor is 640x480 pixels, and the resolution of the cropped facial area varies for each distance range.

| ID | Sex | Age | Height(cm) | Weight(kg) | BMI(kg/m²) |
|----|-----|-----|------------|------------|------------|
| 1  | Female | 34  | 168        | 69         | 24.4       |
| 2  | Male | 42  | 170        | 87         | 30.1       |
| 3  | Female | 30  | 170        | 54         | 18.7       |
| 4  | Male | 33  | 173        | 70         | 23.4       |
| 5  | Female | 34  | 163        | 57         | 21.5       |
| 6  | Female | 35  | 171        | 89         | 29.4       |
| 7  | Male | 32  | 168        | 70         | 24.8       |
| 8  | Male | 34  | 183        | 78         | 23.3       |
| 9  | Female | 33  | 168        | 72         | 25.5       |
| 10 | Male | 27  | 175        | 72         | 23.5       |
| AVG | -    | 33.4 | 170.9      | 71.8       | 23.3       |

Table 2: Subjects’ Information

3.1. Data Collection Methodology

The Office of Research Protections and Integrity UNCC-IRB183845 has approved this study. The data collection process took place in Jun 2021. We have collected the data in a one-day long session or two shorter sessions over two days based on the subject’s preference. Participants are five males and five females, healthy adults. We made certain that the subjects did not have any thermoregulatory illnesses, heat intolerance, colds, flu, or infections. The participants did not use any makeup or facial cream and removed their glasses for the recordings. All participants wore light-colored short-sleeve shirts and pants. Two of the male participants had facial hair. Table 2 shows the information for each subject in detail. Each recording session is designed for capturing thermal recordings from one subject at ten relative distances from the camera and 25 different head positions in each interval. In addition, the temperature of each session is different from the previous session, which have resulted in several variations in the thermal condition. All the ten subjects have participated in at least four recording sessions.

After taking the informed consent, the researchers recorded the participant’s age, gender, height, and weight. Adaptation is a common part of human-related thermal condition research, which helps subjects’ bodies reach a steady metabolic rate and thermal state, as it is influenced by their prior activities and environments. Earlier research has suggested that the mean skin temperature and thermal sensation stabilize after 40 minutes of being exposed to a new thermal environment with a temperature difference of less than 10°C [91, 92, 93]. The adaptation time period is sometimes selected to be less than this period in similar previous research, such as 20 minutes[94] or 30 minutes[95]. As one purpose of our experiment has been the development of a facial thermography dataset, we have decided to choose a longer acclimatization period so we can be sure of the reliability of our dataset. The participants had entered the test room and stayed in a seated position for 60 minutes before the test, so their metabolic rate had reached a stable state, and any influence from the prior outdoor temperature was eliminated. The experiment for each participant was a combination of five sub-sessions in an approximately one-hour time frame for each sub-session. The recording was initiated at a temperature of 21°C while the thermostat held the ambient temperature constant with a 1°C fluctuation allowance in each steady state session. Each subsequent session started at a 1.5°C higher temperature in the same one-hour time frame.

Figure 3 displays the test room layout and the positioning of the following: (1) thermal camera, (2) subject’s stations, (3) data loggers, (4) temperature and relative humidity sensors, and (5) air diffuser and airflow sensor. This figure shows that the temperature and humidity sensors are placed at the number(4) and mounted on a pole at three different heights (0.1, 1.1, and 1.7 meters). The recordings are at ten different relational distances from 1 to 6.6 meters, which this figure shows by number(2). Participants were seated in front of the camera during the study and changed their head positions as instructed while recording RGB and thermal frames pairs.
The researchers performed the recordings in each station from the camera and in 25 different head positions instructed to the subjects (Figure 3). The user logged the thermal sensation in each station from the camera through a "Google Form" with three standard levels: cool, slightly cool, neutral, slightly warm, and hot. An approximate number of 1,000 frames were captured for each subject, with a total of 10,000 thermal frames for the dataset. Table 3 shows the recorded variables and recording sensors’ information in more detail. The air temperature and relative humidity are measured with HOBO Pro v2 temperature/relative humidity data logger sensors, which were calibrated with ice water before the experiments. The airspeed is recorded at the air diffuser proximity with an average distance of 3 meters from the subject’s station. The dataset is annotated with the face landmarks coordinates, distance to the camera, room air temperature, relative humidity, airflow, and subject thermal sensation.
3.2. Dataset Analysis

This section presents a preliminary study of the gathered dataset. The goal of this evaluation is to investigate the applicability of the data. The analysis of the subjects’ thermal sensation based on their skin temperature is not in the scope of this study and will be covered in our next publication as our objective in this section is to show the key statistical properties of the whole dataset. First, we analyze the main independent variables, including environmental temperature, relative humidity, and distance from the camera. Then, we look into the correlation of facial skin temperature in different facial areas with the environmental temperature and each other. The authors will also compare sample frames together to study skin temperature differences in diverse thermal conditions and among different subjects. The provided results for facial area correlation are based on our developed method for detecting facial landmarks. The complete explanation of this method is outside the scope of this paper and will be explained in detail in another relevant publication. In addition, we have not included the utilized visual images in the published
Some recent research has succeeded in detecting facial thermal images without using RGB data \(^{37}\). Since our dataset includes images with various resolutions, it would be a great fit to evaluate the prediction accuracy of such algorithms, which would be one of our future objectives. Here we provide a brief explanation of data preparation and the results’ analysis. We have first cropped the images to the facial area with some margins. We have utilized the Dlib-based model for face recognition, which is based on a 29 convolutional layer in Residual Networks ResNet. This model is a version of the ResNet-34 network that works by removing some layers and reducing the number of filters per layer to half \(^{96}\);\(^{97}\). For locating the facial areas, the position of the facial landmarks has been transferred from the RGB image to the thermal image by calculating the homography matrix between the two frames, as is presented in detail by Negishi et al. \(^{98}\). As the covered distance is 1-6.6 meters from the camera and the facial area resolution gets as low as 25*30 pixels, the HRNet algorithm has been used to identify facial landmarks in the RGB images \(^{99}\). Four facial areas (nose, cheeks, forehead, and chin) have been selected based on the literature to be studied individually in detail \(^{55},\ 56\).

The authors have utilized the estimated facial landmarks to define the selected facial areas, as shown in Figure 5. Since the recorded frames are raw radiometric data, the value of each pixel presents the temperature of the pixel’s representing area in the actual world. According to the Flir camera’s datasheet, when using the TemperatureLinear (TLinear) option, the temperature of each pixel in degrees of Kelvin can be defined with Formula \(^{1}\). This is also very important to set the emissivity of the measuring target to the correct number before the captures, which is 0.98 for the human skin. We have set the TLinear resolution is set to 100 mK(millikelvins) in our recordings \(^{31}\).

\[
\text{SkinTemperature} = \left( \frac{\text{PixelValue}}{\text{TLinearresolution}} - 273.15 \right) \text{K} \tag{1}
\]

4. Dataset Evaluation

First, the dataset frames’ environmental properties are studied to ensure the proper coverage of different thermal conditions in the dataset. Figure 6 displays the (a) temperature and (b) relative humidity coverage based on the number of images for each range. The temperature range is between 20.6 °C and 26.6 °C, divided into four groups with a 1.5 °C range for each. This figure shows an approximately uniform distribution in the room (a) air temperature, which means the number of data frames in different temperature ranges is approximately the same. Although we have not controlled the (b) relative humidity in the recording process, controlling the air temperature has resulted in changes in relative humidity, which were recorded and added to the annotations. Additional statistical information about the environmental properties, including air temperature, globe temperature, relative humidity, and airflow, is provided in Table 4. Most of the sessions had zero airflows to ensure image consistency; however, the air conditioner was at 0.5 m/s airspeed at the diffuser proximity in extremely cold conditions.
Figure 6: Covered Environmental Temperature and Relative Humidity in the Dataset

Table 4: Environmental Properties Measurement Results Overview

Furthermore, we have studied the changes in the skin temperature in each region for the whole dataset. The temperature of each pixel in the desired Region of Interest (ROI) is calculated by the equation 1 used to identify the average temperature in each facial area. Table 5 shows an overview of the dataset’s skin temperature measurement results for all the recorded frames. As the table shows, the temperature in the forehead area has the highest mean value (34.3 °C), while other facial areas’ mean values are closer together in the range of 33.17 °C to 33.82 °C. The low standard error numbers (0.02-0.03) in all facial regions demonstrate that a sample data set can be an acceptable representation of the entire dataset. In addition, the close amounts of mean and median numbers show the relatively symmetrical distribution of the data. Additionally, Figure 7 presents box plots for the temperature variation of each facial region based on the environmental temperature. As expected by the literature, the nasal area has the highest variation, and the forehead temperature shows the lowest variation.

Table 5: Skin Temperature Measurement Results Overview For All Subjects
To have another holistic view of the changing patterns of different facial parts with room temperature, figure 8 shows the relationship between skin temperature and room temperature. The changing pattern of these charts shows that the temperature in all facial areas for the whole dataset has a positive weak linear association with the room ambient temperature. The $R^2$ values for these linear patterns are chin 0.13, forehead 0.4, nose 0.4, cheeks’ average 0.6, left cheek 0.25, and right cheek 0.25, which backs up this linear relation. The linear correlation values with the Pearson method are also presented in Table 6, which will be further discussed. As it is shown the forehead area has the lowest linear slope (0.1918) and the temperature range as the air temperature increases to the highest amount. On the other hand, the nose temperature shows the most significant change as the air temperature increases with a linear slope of 0.5742. These findings are in line with the literature [43], highlighting that the nose area has the highest correlation with the room temperature. In contrast, the forehead area changes the least and can be a good indicator of internal body temperature. As shown in this figure, we have looked into both right and left cheek temperature and their average temperature. The average cheek temperature is the average temperature of the left and right sides when both sides are visible in the thermal frame. If the head position is a full profile, the invisible cheek is excluded from the calculation. Therefore, although the number of data points for the right and left cheek is less than the other facial parts, the cheek average has the same data point as the other face areas.
Furthermore, Table 6 presents the "Pearson Correlation Values" for different facial regions with the environmental temperature and the subjects’ thermal sensations across the whole dataset. The Pearson correlation coefficient indicates the strength and direction of a linear relationship between different variables, and it would be a relevant factor for our investigation. Table 6 shows that the cheek average and nose area have the highest correlation with the environmental temperature, which is 0.76 and 0.67, respectively. On the other hand, the forehead area has the lowest correlation with the environmental temperature, which proves the applicability of this facial area as an indicator of core body temperature, as environmental properties less influence this area. This result is in line with the previous findings from the literature[43] on the correlation of different facial regions with the environmental temperature. In addition, we can see that the cheeks and the nose area are the best indicators of subjective thermal sensation, while the forehead area is not that suitable for that purpose.
Table 6: Correlation of Facial Regions With Room Temperature and Thermal Sensation

|                      | Right Cheek | Left Cheek | Cheeks Average | Nose | Forehead | Chin |
|----------------------|-------------|------------|----------------|------|----------|------|
| Room Temperature     | 0.5         | 0.5        | 0.76           | 0.67 | 0.35     | 0.55 |
| Thermal Sensation    | 0.42        | 0.43       | 0.63           | 0.55 | 0.22     | 0.45 |

The relationship between different facial regions among all the subjects is also calculated based on the Pearson Correlation Coefficient and visualized in a heat map figure. As Figure 9 shows, most of the facial regions have a moderate to relatively strong correlation. The right and left cheeks show no correlation, which results from the change in the head positions that makes the two sides have different angles with the camera in most of the data frames. When the head is in a semi-profile position, one of the cheeks is right in front of the thermal camera, while the other side has an angle with the thermal camera, which makes the reading of that side lower than the actual temperature. As a result, the two facial sides can only have an approximately similar reading condition in the full frontal face position. This correlation heatmap also shows that the cheek area has the highest correlation with other facial regions, especially the nose and chin region (0.62). On the other hand, the forehead has the lowest correlation values with other facial regions, which was expected based on both literature and the correlation values of the forehead area with the environmental temperature from Table 6.
Moreover, as one of the main focuses of this study, the importance of environmental temperature in changing the facial thermal images is highlighted by studying some sample images from the dataset of three selected subjects. Figure 10 shows how the facial skin temperature changes in three different thermal conditions for these three individuals. The first row has the lowest Room Temperatures (RT), which increases gradually as we reach the last row. The RT for each row is approximately the same for all three subjects. In addition, the reported Thermal Sensation (TS) of each subject is included, which is also the same for images of subjects in the same row. Ironbow A false-color pallet is used for easier identification of changes in different facial areas.

The authors have selected an ellipse-shaped Region of Interest (ROI) in the facial area. The histogram of the pixel value distribution in the defined ROI also shows how skin temperature changes on the whole face for each subject. As mentioned in Equation 1, the skin temperature is calculated linearly from the pixel value, so the larger amounts of this value show a higher skin temperature. We can observe that the nose and cheek area show more apparent changes than the forehead and chin for all three subjects. In addition, we can see in both the false-colored images and histograms subjects (a) and (b) have a noticeable increase in their skin temperature as the room temperature increases.

The individual differences in facial regions’ temperatures can also be detected easily in images of different subjects in the same row when the Room Temperatures (RT) are relatively close together. For instance, in the first row, the room temperature is in the range of (21.5°C-21.7°C), but the user’s skin temperature shows apparent differences in different facial regions. It is also important that all three subjects have reported
their thermal sensation as "Slightly Cool," while their facial thermal image and the ROI histogram vary to a great extent. While subjects (a) and (b) show an increase in the facial skin temperature, subject (c) shows a different pattern in skin temperature changes, which is an initial increase followed by an identifiable decrease when the temperature reaches 25.2°C (77°F). The explanation for this behavior change is that the subject was beginning to sweat in the forehead area, which was not noticeable even by himself. The higher ambient temperature was the cause of the sweating. As mentioned by the individual, he typically has a lower tolerance for increased ambient temperatures, which results in him sweating more frequently in a hot room temperature. However, the perspiration was insignificant and therefore not noticeable by the researchers or the subject at the time of the experiment. This behavior highlights the importance of including ambient temperature variation in the facial thermography dataset. We can observe that the forehead area of this subject shows an increase when the room temperature changes from 21.6 °C to 23.5 °C and the subject’s thermal sensation changes from "Slightly Cool" to "Neutral". However, as the unnoticeable sweating begins in the next two rows, the forehead temperature decreases, while the subject’s thermal sensation changes to "Slightly Warm". These individual differences in the facial thermal images have reoccurred in the three next rows, as well as the whole dataset. The personal differences in facial thermal images are one of the main reasons we need a comprehensive dataset in several different thermal conditions for our investigations into facial thermography.

Although the subjects show personal differences in their facial thermal images, we can observe a linear correlation between the temperature of different facial areas and the environmental temperature in all of them. Figure 11 presents this linear correlation in all the facial regions and for all the included subjects. The horizontal axis presents the room temperature, and the primary vertical axis shows the facial skin temperature. In addition, each person’s reported subjective thermal sensation as the temperature increases is displayed in the secondary vertical axis. As shown in Figure 11 for all the individual subjects, the highest skin temperature variations as the room temperature increases are in the nose and cheeks area, and the lowest variation is in the forehead and chin area. This is important to note that subject 2’s forehead temperature decreases as the room temperature increases, which was previously explained in Figure 10 for the same subject in column (c). Another important observation from this figure is the higher temperature differences in facial areas when the room temperature is low compared to higher room temperatures. This pattern is repeated in all the subjects, while the differences are much greater in some of them, such as subjects 3 and 10. These figures also show that all the subjects have experienced different thermal sensations; however, the pattern and frequency of each thermal vote are different and based on individual differences.

5. Discussion and Future Work

This paper presents a novel facial infrared thermal dataset with variation in environmental properties, distance from the camera, and head position in raw 16-bit data frames. The data is annotated with each person’s environmental conditions, facial landmarks, and at the time of recording each frame. The subjective thermal sensation annotations are a new addition to the face thermal image datasets. The comparative study of the temperature in different facial areas shows the importance of environmental temperature in facial thermography. Variation in environmental temperature, facial resolution, and head position makes our dataset an excellent fit for training facial recognition algorithms. Another important consideration is identifying facial landmarks for a more precise evaluation of facial skin temperature. Although we have used the RGB paired images for facial area identification in this paper, there is other state-of-the-art research for facial landmark detection. Since our dataset includes images with various resolutions, it would be a great fit to evaluate the prediction accuracy of such algorithms, which would be one of our future research objectives. The explanation for this behavior change is that the subject was beginning to sweat in the forehead area, which was not noticeable even by himself.

In the data analysis section for this paper, we have focused on the reliability of the dataset for future research. Therefore, we have looked into the temperature range of the images and the correlation of ambient temperature with the skin temperature of each subject. However, other analyses can be performed on this dataset, including investigation of the performance of different learning algorithms on facial landmark detection. In addition, studies can be performed to investigate the accuracy of detecting subjects’ thermal
comfort based on their skin temperature. Although previously included research has already proved this possibility [56], it is important to study this feature in a farther distance from the camera if we want to apply it in real buildings. Another subsequent study would be on the prediction of subjects’ thermal sensations based on the infrared thermal images by utilizing the current dataset. We will analyse the influence of distance and head position on the thermal camera readings and the prediction accuracy of thermal comfort based on the subjects’ skin temperature. The manual annotation of the facial frames will provide us with the ground truth facial landmarks in the thermal frames, which will be used for the validation of our presented approach.

More projects are in progress for the improvement of this dataset based on its current limitations. Our dataset does not include different facial expressions, which are included in some datasets and will be added to our updated versions. Currently, due to the high variety of recording conditions for each subject, we have included ten subjects in our experiment. We are planning on adding more subjects in the future version of our dataset from different ethnic groups. The current dataset was recorded in a controlled indoor environment. Including images that were taken outdoors could make this dataset more suitable applicable. In addition, our current dataset does not include RGB images, which can be included in our following versions to make the dataset helpful for a broader range of projects.

The complete dataset is available for downloading on our Github page [100].
Funding Sources
This work was supported by the National Science Foundation [grant number NSF 2104223].

References

[1] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, J. Philip, Medical applications of infrared thermography: A review. Infrared Physics and Technology 55 (4) (2012) 221–235. doi:10.1016/j.infrared.2012.03.007
URL http://dx.doi.org/10.1016/j.infrared.2012.03.007

[2] L. Antognoli, P. Marchionni, S. Nobile, V. Carnielli, L. Scalise, Assessment of cardio-respiratory rates by non-invasive measurement methods in hospitalized preterm neonates. MeMeA 2018 - 2018 IEEE International Symposium on Medical Measurements and Applications, Proceedings doi:10.1109/MeMeA.2018.8438772

[3] M. Chakraborty, S. K. Raman, S. Mukhopadhyay, S. Patsa, N. Anjum, J. G. Ray, High precision automated face localization in thermal images: an early case study, Medical Imaging 2017: Image Processing 10133 (February 2017) (2017) 1013326. doi:10.1117/12.2254236

[4] T. Giitsidis, E. G. Karakasis, A. Gasteratos, G. C. Sirakoulis, Human and fire detection from high altitude UAV images, Proceedings - 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2015 (2015) 309–315. doi:10.1109/PDP.2015.118

[5] A. Szajewska, Development of the Thermal Imaging Camera (TIC) Technology, Procedia Engineering 172 (2017) 1067–1072. doi:10.1016/J.PROENG.2017.02.164

[6] Y. Ma, X. Feng, J. Jiao, Z. Feng, S. Qian, H. Xue, H. Li, Smart Fire Alarm System with Person Detection and Thermal Camera, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12143 LNCS (2020) 353–366. doi:10.1007/978-3-030-50436-6_26
URL https://link.springer.com/chapter/10.1007/978-3-030-50436-6_26

[7] L. Chen, N. Liu, M. Hu, G. Zhai, RGB-Thermal Imaging System Collaborated with Marker Tracking for Remote Breathing Rate Measurement, 2019 IEEE International Conference on Visual Communications and Image Processing, VCIP 2019 (2019) 12–15. doi:10.1109/VICIP47243.2019.8965987

[8] B. Metthig, A. Liu, S. Habibi, M. V. Mohrenschildt, Leveraging Thermal Imaging for Autonomous Driving, IETE 2019 - 2019 IEEE Transportation Electrification Conference and Expo (2019). doi:10.1109/ITCExpo.2019.8790493

[9] A. Alarabzadeh, M. A. Notani, A. Kazemiyan Zadeh, A. Nahvi, A. Sassani, H. Ceylan, Electrically Conductive Asphalt Concrete: An Alternative for Automating the Winter Maintenance Operations of Transportation Infrastructure, Composites Part B: Engineering 173 (2019) 106985. doi:10.1016/J.COMPOSITESB.2019.106985

[10] I. G. Dino, A. E. Sari, O. K. Iseri, S. Akin, E. Kalfaoglu, B. Erdogan, S. Kalkan, A. A. Alatan, Image-based construction of building energy models using computer vision, Automation in Construction 116 (2020) 103231. doi:10.1016/J.AUTCON.2020.103231

[11] R. Gupta, A. Kotopoules, Magnitude and extent of building fabric thermal performance gap in UK low energy housing, Applied Energy 222 (2018) 673–686. doi:10.1016/J.APENERGY.2018.03.096

[12] A. G. Entrop, A. Vasenev, Infrared drones in the construction industry: designing a protocol for building thermal energy procedures, Energy Procedia 132 (2017) 63–68. doi:10.1016/J.ENPYRD.2017.09.636

[13] V. Espinosa-Duró, M. Faundez-Zanuy, J. Mekyska, A New Face Database Simultaneously Acquired in Visible, Near-Infrared and Thermal Spectrums, Cognitive Computation 5 (1) (2013) 119–135. doi:10.1007/s12559-012-9163-2

[14] M. Peng, C. Wang, T. Chen, G. Liu, NIRFaceNet: A convolutional neural network for near-infrared face identification, Information Sciences (Switzerland) 7 (4) (2016) 1–14. doi:10.3390/info7040061

[15] M. Kopaczka, R. Kolk, D. Merhof, A fully annotated thermal face database and its application for thermal facial expression recognition, I2MTC 2018 - 2018 IEEE International Instrumentation and Measurement Technology Conference: Discovering New Horizons in Instrumentation and Measurement, Proceedings (2018) 1. doi:10.1109/I2MTC.2018.8409768

[16] C. Goulart, C. Valadão, D. Delisle-Rodriguez, D. Funayama, A. Favarato, G. Baldo, V. Binotte, E. Caldeira, T. Bastos-Filho, Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction, Sensors (Switzerland) 19 (13). doi:10.3390/s19132844

[17] O. Cevher, E. Raff, S. Purushotham, The Use of AI for Thermal Emotion Recognition: A Review of Problems and Limitations in Standard Design and Data. URL http://arxiv.org/abs/2009.10589

[18] A. C. Cosma, R. Simha, Using the contrast within a single face heat map to assess personal thermal comfort, Building and Environment 160 (2019). doi:10.1016/j.buildenv.2019.106163

[19] D. Lai, Q. Chen, A two-dimensional model for calculating heat transfer in the human body in a transient and non-uniform thermal environment, Energy and Buildings 118 (2016) 114–122. doi:10.1016/j.enbuild.2016.02.051
URL http://dx.doi.org/10.1016/j.enbuild.2016.02.051

[20] S. Lu, W. Wang, S. Wang, E. Cochran Hameen, Thermal Comfort-Based Personalized Models with Non-Intrusive Sensing Technique in Office Buildings, Applied Sciences 9 (9) (2019) 1768. doi:10.3390/app9091768
URL https://www.mdpi.com/2076-3417/9/9/1768

[21] Z. Jiang, M. Hu, L. Fan, Y. Pan, W. Yang, G. Zhai, Y. Lu, Combining Visible Light and Infrared Imaging for Efficient Detection of Respiratory Infections such as COVID-19 on Portable Device, Sensors (Switzerland) 7 (4) (2017) 1–14. doi:10.3390/s17040061

[22] S. N. Al-Humairi, A. A. Kamal, Opportunities and challenges for the building monitoring systems in the age-of-pandemic of COVID-19: Review and prospects, Innovative Infrastructure Solutions 6 (2) (2021) 1–10. doi:10.1007
[23] C. A. Corneanu, M. O. Simón, J. F. Colin, S. E. Guerrero, Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (8) (2016) 1548–1568. doi:10.1109/TPAMI.2016.2515606

[24] D. Foster, M. Thielke, R. Nguyen, S. Rajaraman, X. Di, C. N. Fondej, V. M. Patel, N. J. Short, B. S. Riggan, N. M. Narasabadi, S. Hu, A Large-Scale, Time-Synchronized Visible and Thermal Face Dataset, in: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2021, pp. 1558–1567. doi:10.1109/WACV48630.2021.00160

URL: https://ieeexplore.ieee.org/document/9423096

[25] A. Kwasniewska, J. Ruminski, M. Szankin, M. Kaczmarek, Super-resolved thermal imagery for high-accuracy facial areas detection and analysis, Engineering Applications of Artificial Intelligence 87 (October 2019) (2020) 103263. doi:10.1016/j.engappai.2019.103263

URL: https://doi.org/10.1016/j.engappai.2019.103263

[26] H. Zhang, B. S. Riggan, S. Hu, N. J. Short, V. M. Patel, Synthesis of High-Quality Visible Faces from Polarimetric Thermal Faces using Generative Adversarial Networks, International Journal of Computer Vision 127 (6-7) (2019) 845–872. doi:10.1007/s11263-019-01176-5

URL: https://doi.org/10.1007/s11263-019-01176-5

[27] N. Playà-Montmany, G. J. Tattersall, Spot size, distance and emissivity errors in field applications of infrared thermography, Methods in Ecology and Evolution (2021) 2041–2043. doi:10.1111/2041-210X.13563

URL: https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.13563

[28] E. Faye, O. Dangles, S. Pincoeurbde, Distance makes the difference in thermography for ecological studies, Journal of Thermal Biology 56 (2016) 1–9. doi:10.1016/j.jtherbio.2015.11.011

[29] R. Vardasca, A. R. Marques, J. Díaz, A. Seixas, J. Mendes, E. F. J. Ring, The influence of angles and distance on assessing inner-canthi of the eye skin temperature, Tech. rep. (2017).

[30] M. Bodini, A review of facial landmark extraction in 2D images and videos using deep learning, Big Data and Cognitive Computing 3 (1) (2019) 1–14. doi:10.3390/bdc3010014

[31] FLIR Company, A Resource Guide for Using Infrared in the Research and Development Industry.

[32] M. Kristo, M. Iväkos, An overview of thermal face recognition methods, 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings (2018) 1098–1103. doi:10.23919/MIPRO.2018.8400200

[33] Z. Chen, X. Huang, Pedestrian detection for autonomous vehicle using multi-spectral cameras, IEEE Transactions on Intelligent Vehicles 4 (2) (2019) 211–219. doi:10.1109/TIV.2019.2904389

[34] C. Ma, N. T. Trung, H. Uchiyama, H. Nagahara, A. Shimada, R. I. Taniguchi, Mixed features for face detection in thermal image, Thirteenth International Conference on Quality Control by Artificial Vision 2017 10338 (May 2017) 103380E. doi:10.1117/12.2266836

[35] C. Ma, N. T. Trung, H. Uchiyama, H. Nagahara, A. Shimada, R. I. Taniguchi, Adapting local features for face detection in thermal image, Sensors (Switzerland) 17 (12) (2017) 1–31. doi:10.3390/s17122747

[36] S. Sumriddetchkajorn, A. Somboonkaew, Face detection in thermal imagery using an Open Source Computer Vision library, Thermosense XXXI 7290 (April 2009) (2009) 729006. doi:10.1117/12.819996

[37] M. Vukovic, R. Petrovic, M. Pavlovic, S. Stankovic, Thermal Image Degradation Influence on R-CNN Face Detection Performance, 27th Telecommunications Forum, TELFOR 2019 (2019) 10971128. doi:10.1109/TELFOR48224.2019.8971128

[38] R. Hughes, Non-Contact Thermometers for Detecting Fever: A Review of Clinical Effectiveness, Journal of Chemical Information and Modeling 55 (9) (2008) 287. doi:10.1021/ci8001534

[39] A. M. Solent, N. M. Molinari, A. Baxter, A. V. Nguyen, H. Siegelson, C. M. Brown, A. Plummer, A. Higgins, S. Podolsky, P. Spandorfer, N. J. Cohen, D. B. Fishbein, Mass screening for fever in children: a comparison of 3 infrared thermal detection systems., Pediatric emergency care 29 (3) (2013) 305–313. doi:10.1097/PEC.0b013e3182854465

[40] J. B. Mercer, E. F. J. Ring, Fever screening and infrared thermal imaging: Concerns and guidelines, Thermology International 19 (3) (2009) 67–69.

[41] E. Y. Ng, Is thermal scanner losing its bite in mass screening of fever due to SARS?, Medical Physics 32 (1) (2005) 93–97.

[42] N. Silawan, K. Kusukame, K. J. Kek, W. S. Kuan, A Novel Environment-Invariant Core Body Temperature Estimation for High Sensitivity and Specificity Fever Screening, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2018-2018-8512541

[43] A. H. Alkali, R. Saatchi, H. Elphick, D. Burke, Facial tracking in thermal images for real-time noncontact respiration rate monitoring - Proceedings - UKSim-AMSS 7th European Modelling Symposium on Computer Modelling and Simulation, EMS 2013 (2013) 265–271. doi:10.1109/EMBS.2013.662

[44] Y. Cho, S. J. Julier, N. Marquardt, N. Bianchi-Berthouze1, Robust tracking of respiratory rate in high- dynamic range scenes using mobile thermal imaging (2015). doi:10.1364/boe.8.004480

[45] T. Negishi, G. Sun, H. Liu, S. Sato, T. Matsui, T. Kirimoto, Stable Contactless Sensing of Vital Signs Using RGB-Thermal Image Fusion System with Facial Tracking for Infection Screening, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2018-2018-8513300

[46] L. Chen, N. Liu, M. Hu, G. Zhai, Rgb-Thermal imaging system collaborated with marker tracking for remote breathing
rate measurement, 2019 IEEE International Conference on Visual Communications and Image Processing, VCIP 2019 (2019) 12–15. doi:10.1109/VCIP47243.2019.8955987.

[48] J. K. Lee, H. W. Jeong, Wearing facemasks regardless of symptoms is crucial for preventing spread of COVID-19 in hospitals, Infection control and hospital epidemiology (2020) 1–4; doi:10.1017/ice.2020.202.

[49] L-K. Lee, C-C. Wang, M-C. Lin, C-T. Kung, K-C. Lan, C-T. Lee, Effective strategies to prevent coronavirus disease-2019 (COVID-19) outbreak in hospital. Journal of Hospital Infection 105 (1) (2020) 102–103. doi:10.1016/j.jhin.2020.02.022.

URL: https://linkinghub.elsevier.com/retrieve/pii/S0195670120300980.

[50] A. Barnawi, P. Chhikara, R. Tekchandani, N. Kumar, B. Alzahrani, Artificial intelligence-enabled Internet of Things-based system for COVID-19 screening using aerial thermal imaging. Future Generation Computer Systems 127 (2021) 119–132. doi:10.1016/j.future.2021.05.019.

URL: https://linkinghub.elsevier.com/retrieve/pii/S0378778818309629.

[51] G. Recommendations, A. C. Systems, General Recommendations Operating commercial office buildings under epidemic. http://vcipl-okstate.org/pbvs/bench/.

[52] J. Ranjan, J. Scott, ThermalSense: Determining Dynamic Thermal Comfort Preferences using Thermographic Imaging. doi:10.1145/2971648.2971659.

URL: http://dx.doi.org/10.1145/2971648.2971659.

[53] H. Metzmacher, D. Wolk, C. Schmidt, J. Frisch, C. van Treeck, Real-time human skin temperature analysis using thermal image recognition for thermal comfort assessment, Energy and Buildings 158 (2018) 1063–1078. doi:10.1016/j.enbuild.2017.09.032.

[54] D. Li, C. C. Menassa, V. R. Kamat, Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography. Energy and Buildings 176 (2018) 250–261. doi:10.1016/j.enbuild.2018.07.025.

URL: https://linkinghub.elsevier.com/retrieve/pii/S0378778818309629.

[55] A. Aryal, B. Becerik-Gerber, Skin Temperature Extraction Using Facial Landmark Detection and Thermal Imaging for Comfort Assessment (2019) 71–80. doi:10.1145/3360322.3360848.

[56] D. Li, C. C. Menassa, V. R. Kamat, Robust non-intrusive interpretation of occupant thermal comfort in built environments with low-cost networked thermal cameras. Applied Energy 251 (2019) 113336. doi:10.1016/j.apenergy.2019.113336.

URL: https://linkinghub.elsevier.com/retrieve/pii/S0306261919310104.

[57] A. Selinger, D. A. Socolinsky, Appearance-based facial recognition using visible and thermal imagery: a comparative study. Image Rochester NY 4 (2006) 28.

URL: http://www.equinooxensors.com/publications/andreas_face.pdf.

[58] H. Méndez, C. S. Martin, J. Kittler, Y. Plasencia, E. García-Reyes, Face recognition with LWIR imagery using local combination of geometry with physiological information. Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005 II (2005) 1022–1029. doi:10.1109/CVPR.2005.241.

[59] H. Chang, H. Harishwaran, M. Yi, A. Koschan, B. Abidi, A. Alzahrani. Visual and multi-illuminant database for face recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2006 (c) (2006) 0–7. doi:10.1109/CVPR.2006.28.

[60] J. Serrano-Cuerda, A. Fernández-Caballero, M. T. López, Selection of a visible-light vs. thermal infrared sensor in hospitals. Infection control and hospital epidemiology (2020) 1–4. doi:10.1109/ice.2020.3002911.

[61] I. Pavlidis, J. a. Levine, P. Bankol, Ioannis Pavlidis Honeywell Laboratories Minneapolis , MN ioannis . pavlidis @ i honeywell . com James Levine Mayo Clinic Mayo Clinic, IEEE Transactions on Biomedical Engineering (6) (2002) 315–318. doi:10.1109/TBME.2001.958491.

[62] H. Nguyen, K. Kotani, F. Chen, B. Le, A thermal facial emotion database and its analysis, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 5558 LNCS (2009) 327–336. doi:10.1007/978-3-642-01793-3_34.

[63] I. A. Kakadiaris, G. Passalis, T. Theoharis, G. Toderici, I. Konstantinidis, N. Murtuza, Multimodal face recognition: Combination of geometry with physiological information. Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005 II (2005) 1022–1029. doi:10.1109/CVPR.2005.241.

[64] C. Goulart, C. Valadão, D. Delisle-Rodriguez, D. Funayama, A. Favarato, G. Baldo, V. Binotte, E. Caldeira, T. Bastos-Filho, Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction, Sensors (Switzerland) 19 (13). doi:10.3390/s19132844.

[65] D. Robinson, J. Clay-Warner, C. D. Moore, T. Everett, A. Watts, T. N. Tucker, C. Thai, Toward an unobtrusive interaction, Sensors (Switzerland) 19 (13). doi:10.3390/s19132844.

[66] A. Sancen-Plaza, L. M. Contreras-Medina, A. I. Barranco-Gutiérrez, C. Villaseñor-Mora, J. J. Martínez-Nolasco, J. A. Padilla-Medina, Facial Recognition for Drunk People Using Thermal Imaging, Mathematical Problems in Engineering (2020) 1–12. doi:10.1155/2020/1024173.

[67] A. Barnawi, P. Chhikara, R. Tekchandani, N. Kumar, B. Alzahrani, Artificial intelligence-enabled Internet of Things-based system for COVID-19 screening using aerial thermal imaging. Future Generation Computer Systems 127 (2021) 119–132. doi:10.1016/j.future.2021.05.019.

URL: https://linkinghub.elsevier.com/retrieve/pii/S0306261919310104.

[68] J. K. Lee, H. W. Jeong, Wearing facemasks regardless of symptoms is crucial for preventing spread of COVID-19 in hospitals. Infection control and hospital epidemiology (2020) 1–4. doi:10.1109/ICIP.2019.8965987.

[69] X. Chen, P. J. Flynn, K. W. Bowyer, N. Dame, Visible-light and Infrared Face Recognition, The Proceedings of Workshop.
[93] B. Balaji, H. Teraoka, R. Gupta, Y. Agarwal, ZonePAC: Zonal power estimation and control via HVAC metering and occupant feedback, BuildSys 2013 - Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings doi:10.1145/2528282.2528304.

[94] A. C. Cosma, R. Simha, Thermal comfort modeling in transient conditions using real-time local body temperature extraction with a thermographic camera, Building and Environment 143 (2018) 36–47. doi:10.1016/j.buildenv.2018.05.052.

[95] J. H. Choi, V. Loftness, Investigation of human body skin temperatures as a bio-signal to indicate overall thermal sensations, Building and Environment 58 (2012) 258–269. doi:10.1016/j.buildenv.2012.07.003.

[96] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem (2016) 770–778. doi:10.1109/CVPR.2016.90.

[97] Davis King, dlib C++ Library: High Quality Face Recognition with Deep Metric Learning (2017). URL: http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html.

[98] T. Negishi, G. Sun, S. Sato, H. Liu, T. Matsu, S. Abe, H. Nishimura, T. Kirimoto, Infection Screening System Using Thermography and CCD Camera with Good Stability and Swiftness for Non-contact Vital-Signs Measurement by Feature Matching and MUSIC Algorithm, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBS (2019) 3183–3186. doi:10.1109/EMBC.2019.8857027.

[99] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang, W. Liu, B. Xiao, Deep High-Resolution Representation Learning for Visual Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence (March) (2020) 1–1. doi:10.1109/tpami.2020.2983686.

[100] TeCSAR-UNCC/UNCC-ThermalFace. URL: https://github.com/TeCSAR-UNCC/UNCC-ThermalFace.