



HAL
open science

LVT Face Database: A benchmark database for visible and hidden face biometrics

Nelida Mirabet-Herranz, Jean-Luc Dugelay

► **To cite this version:**

Nelida Mirabet-Herranz, Jean-Luc Dugelay. LVT Face Database: A benchmark database for visible and hidden face biometrics. BIOSIG 2023, 22nd International Conference of the Biometrics Special Interest Group, IEEE, Sep 2023, Darmstadt, Germany. hal-04198573

HAL Id: hal-04198573

<https://hal.science/hal-04198573v1>

Submitted on 7 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LVT Face Database: A benchmark database for visible and hidden face biometrics

Nelida Mirabet-Herranz
Dept. of Digital Security
EURECOM
Biot, France
mirabet@eurecom.fr

Jean-Luc Dugelay
Dept. of Digital Security
EURECOM
Biot, France
dugelay@eurecom.fr

Abstract—Although the estimation of eHealth parameters from face visuals (images and videos) has grown as a major area of research in the past years, deep-learning-based models are still challenged by RGB lack of robustness, for instance with changing illumination conditions. As a means to overcome these limitations and to unlock new opportunities, thermal imagery has arisen as a favorable alternative to solidify different technologies such as heart rate estimation from faces. However, the reduced number of databases containing thermal imagery and the lack of health annotation of the subjects in them limits the exploration of this spectrum. Motivated by this, in this paper, we present our Label-EURECOM Visible and Thermal (LVT) Face Database for face biometrics. This database is the first that contains paired visible and thermal images and videos from 52 subjects with metadata of 22 soft biometrics and health parameters. Moreover, we establish the first study introducing the potential of thermal images for weight estimation from faces on our database.

Index Terms—Face database, Visible spectrum, Thermal spectrum, eHealth, Weight estimation.

I. INTRODUCTION

Facial processing from visual content has gained a lot of attention in the past years since it allows for non-invasive contactless monitoring of a subject’s health status, useful in numerous potential applications. Nowadays, there is a global trend to monitor eHealth parameters without the use of physical devices enabling their estimation in at-risk situations such as medical emergencies and road accidents besides at-home daily monitoring and telehealth. Automatic face recognition has consistently been one of the most active research areas of computer vision [1]. Beyond people identification and soft biometric prediction such as gender, age and ethnicity, a vast amount of health information belonging to a subject has been proved to be embedded in face visuals [2]. The estimation of health indicators such as height, weight and Body Mass Index (BMI) from a single facial shot, has been explored in the literature by training a regression method based on the 50-layer ResNet-architecture [3]. Past them, researches have extracted the called micro-signals from faces, information that has played important roles in media security and forensics [4]. An established concept in the past fifteen years is derived from the fact that blood draws more light than the ambient tissues therefore subtle changes in blood volume can be captured by cameras based on the above-mentioned light absorption. This has allowed for remote photoplethysmography (rPPG).

Researches have shown how a mobile phone camera has enough resolution to capture rPPG signal from faces leading to a successful Heart Rate (HR) estimation [5]. Following the same principle, recent works have successfully approximated the Blood Pressure (BP) of a subject thanks to the difference between the times a pulse wave reaches two different parts of the face [6]. In up-to-date investigations, the ratio of oxygenated hemoglobin with respect to total hemoglobin (SpO₂) has been computed from facial videos employing Convolutional Neural Networks (CNN) that consider the direct current and alternating current components extracted from the RGB signals of facial videos [7].

Facial eHealth models traditionally based their estimations on images acquired in the visible spectrum. Despite those networks have reached a significant level of maturity with practical success, deep learning approaches based on data images in the visible spectrum are affected by compromising factors such as occlusion and illumination changes. Thermal imagery has proved itself as a powerful caption tool [1]. Computer vision researchers have affirmed it as superior to visible imaging in hard conditions such as the presence of smoke, dust and absence of light sources [8]. Thermal imagery operates by detecting electromagnetic radiation in the medium MWIR ($3 - 8\mu m$) and long LWIR ($8 - 15\mu m$) wave infrared spectrum [9] where skin heat lays within. This capability enables thermal images to overcome the lack of illumination or some types of occlusions. However, works have highlighted how the thermal heat captured by thermal cameras can be affected by various factors such as ambient temperature or intense physical activity [1].

To enable the next step towards more accurate eHealth models and because we believe in the potential of thermal imagery, we introduce a new database with visuals collected using a paired thermal-visible camera and annotated with health traits from each subject. The main contributions of this work are the following: 1) We present our Label-EURECOM Visible and Thermal Face Database for face biometrics composed of 612 images and 416 videos from 52 different subjects and a compendium of 22 health metrics and soft biometrics annotated per person. 2) We propose the first study, up to the authors’ knowledge, on weight estimation from facial thermal imagery.

The rest of the paper is organized as follows, Section II lists existing databases containing thermal visuals and some descriptors of them as well as motivates the use of thermal images for health-related applications. In Section III, our LVT Face Database for face biometrics newly collected is presented in detail. Section IV includes a brief state-of-art on weight estimation from facial images and the results of an up-to-date weight estimator when re-trained with our new thermal images. Finally, Section V summarizes and concludes with the future directions of our work. The LVT Face Database for face biometrics is publicly available upon request.

II. POTENTIAL OF VISIBLE AND THERMAL PAIRED DATA

Existing biometric systems and facial eHealth applications, are based on databases acquired in the visible or, lately popular, Near InfraRed (NIR) spectrum. Particular studies have however focused on the thermal spectrum for applications such as cross-spectrum face recognition algorithms or HR estimation.

Relevant thermal databases: Interest in employing thermal face images has grown in the past years, nevertheless, this regard has been restricted mostly to tasks such as landmarks and face detection and Face Recognition (FR) [1], [10]. A relevant subset of FR is Cross-FR (CFR) discipline that aims to identify a person's image in the thermal spectrum from a gallery containing face images acquired in the visible spectrum [11]. Only a few databases have been provided involving visuals acquired in thermal spectra and among them, the ones covering health-related metadata are few. In Tab. I, we present an exhaustive selection of relevant databases that include visuals in the thermal spectrum and some key descriptors of them including their year of release, the number of subjects, images and videos present in the database and their initial intended purpose. One of the first datasets containing thermal visual data was presented in 2003 [12]. The data was acquired at the University of Notre Dame and contains images from 240 distinct subjects with four views with different lighting and facial expressions with the purpose of recognizing individuals. Beyond people recognition, Wang *et al.* establish a similar database for expression recognition containing both spontaneous and intended expressions of more than 100 subjects [13] while Gault *et al.* recorded thermal videos from 32 subjects under three imaging scenarios and their paired rPPG signals for HR estimation [14]. In 2018 two new databases were acquired for FR with multiple illuminations, pose and occlusion variations [1] and including imagery from different modalities namely visible, thermal, near-infrared and a computerized facial sketch and 3D images of each volunteer's face [15]. In the same year, Barbosa *et al.* collected thermal videos from 20 healthy subjects in two phases: phase A (frontal view acquisitions) and phase B (side view acquisitions) and the corresponding PPG and thoracic effort simultaneously recorded for HR and Respiratory Rate (RR) estimation [16]. More recently, two large-scale visible and thermal datasets have been assembled. Abdrakhmanova *et al.* gathered a combination of thermal, visual, and audio data streams to support machine

learning-based biometric applications [17] and Poster *et al.* presented the largest collection of paired visible and thermal face images up to date. Variability in expression, pose, and eyewear were recorded [18]. Following, a thermal face dataset with annotated face bounding boxes and facial landmarks composed of 2556 images was introduced [10].

Thermal data for eHealth: Although the use of facial thermal imagery has traditionally focused on face recognition tasks, some researchers have intended for eHealth parameter estimation in the thermal spectrum showing the potential of this type of data. In 2017, Rai *et al.* suggested that thermal imaging systems have the prospective of providing details regarding physiological processes using skin temperature distributions due to processes such as blood perfusion. Indeed, cameras are often used to observe minute variations in temperature in the medical field in applications including the detection of malignant tumors [9]. The assessment of eHealth parameters such as heart rate from face videos has been studied in depth in recent years. Up to the authors' knowledge, all methods need proper illumination difficult to achieve in uncontrolled environments. In 2018, Barbosa *et al.* presented a new method for remote HR monitoring based on periodic head movements caused by the cyclical ejection of blood flow from the heart to the head. This new algorithm was based on the use of thermal images as input data [16]. Moreover, they proved possible the evaluation and measurement of a subject's RR by using temperature fluctuations under the nose during the respiratory cycle. Thermal imagery proved itself of high value to overcome illumination constraints since thermal images are light invariant. In the same line, other works continue investigating the future of heart rate and blood pressure extraction from thermal images through deep-learning approaches [19].

To the best of our knowledge, current literature focuses on HR, RR and BP from thermal face data. The estimation of other health traits such as SpO2 or weight from thermal images remains untouched by the community. The collection of a new database of visible face visual data and their thermal counterpart is motivated by the potential that thermal images and videos as input data have shown and by the limited number of publicly available databases containing this type of data and their associated health parameters annotation. Moreover, existing databases are limited to visual face information content and one or two parameters. We believe in the value that a database composed of more than 20 different soft biometric and health measures can add to the biometric and health research community.

III. DATABASE DESCRIPTION

In this section, we first introduce the recording setup of the database and the characteristics of the acquisition devices. We detail the data collection methodology as well as the database design and associated subjects' metadata.

Acquisition material: The visible and thermal face visual data was acquired with the dual sensor from the camera FLIR Duo R developed by FLIR Systems. The camera was designed

TABLE I
RELEVANT FACE DATABASES CONTAINING VISUALS IN THERMAL SPECTRA.

Year	Dataset	# of subjects	# of images	# of videos	Objective
2003	UND-X1 [12]	241	4584	-	FR
2010	NVIE [13]	215	Not provided	Not provided	Expression recognition
2013	TH-HR [14]	32	-	96	HR
2018	VIS-TH [1]	50	2100	-	FR
2018	TUFTS [15]	113	Over 10000	113	FR
2018	TH-HR-RR [16]	20	-	40	HR, RR
2021	Speaking faces [17]	142	-	45 hours	Biometric Authentication
2021	ARL-VTF [18]	395	549712	-	Cross-FR
2022	SF-TL54 [10]	142	2556	-	Landmarks detection
2023	Ours: LVT	52	612	416	FR, Soft biometrics, e-health

for capturing simultaneously visible and thermal visuals by unmanned aerial vehicles. FLIR Duo R dual camera has been used in recent researches due to its suitability in data collection for different tasks such as face recognition and cross-spectrum applications [1], [8]. The visible and thermal sensors of this camera are a CCD sensor with a pixel resolution of 1920×1080 and an uncooled VOx microbolometer with a pixel resolution of 640×512 respectively. Various devices were used for a health status assessment of the subjects. A contactless infrared thermometer with a precision of $\pm 0.2^\circ$ Celsius (C) between 34°C and 42.0°C and a precision of $\pm 0.3^\circ\text{C}$ in the range of 42.1°C and 43.0°C was used for computing the user's body temperature. For calculating the BP, an OMRON HEM-7155-E tensiometer was employed together with a LED finger oximeter for SpO2 measurement with a precision of $\pm 2\%$. For HR tracking, the subjects were asked to wear a Garmin Vivoactive@4 smartwatch that embeds an optical PPG sensor able to detect the heart rate by shining a green light through the subject's skin thus reflecting the red cells in the skin's blood vessels. For quantifying bodyweight related measures, we rely on the RENPHO@Body Fat Smart scale. When a subject steps on the device and after entering in the system their gender, age and height, the scale returns 13 metrics including weight and BMI.

Visuals collection protocol: Image and video acquisition were performed in an indoor environment where the ambient temperature was set to 25°C . In Fig. 1 we present the arrangements. The acquisition setup included a white wall acting as background, a chair at a fixed distance of 0.25 m from the camera which is placed at a height of 1 meter from the ground, and a two-point lighting kit placed to limit shadows allowing and easing segmentation of the subject from the background. Each volunteer participated in two separate acquisition sessions, with an average time interval of 6 weeks. Before the acquisition process, volunteers were asked to fill out and sign consent forms. The visual data includes 6 images per person (3 visible and their associated thermal pair) in each session with 3 different conditions, Neutral (N), Ambient light(A) and an occlusion in the form of eyeglasses (O) resulting in a total of 612 images. Fig. 2 illustrates example images of an individual from the database. In addition, four 60-

second videos are recorded per subject in each session with N conditions. The first pair of videos (one in visible spectrum and its paired thermal) are taken after the subject has been resting for at least 5 minutes and the second pair follows moderate exercise in the form of climbing up stairs to increase their HR values making a total of 408 60s videos.



Fig. 1. Flir Duo R camera (left) and acquisition setup (right).

Subjects' metadata: Several metadata pieces of information were collected to describe the subject: gender, age and height. Other parameters were quantified to assess their health status: body temperature, HR, BP, SpO2, weight and BMI. In addition to weight and BMI, the smart scale provided other 11 variables: body fat and body water percentages, skeletal muscle, fat-free weight, muscle mass and bone mass, protein, subcutaneous and visceral fat, Basal Metabolic Rate (BMR) and metabolic age. Image and video filenames are constructed by indicating the visual data spectrum, subject id, session id



Fig. 2. Example images obtained with Flir Duo R camera. The three variations are displayed in visible (upper row) and thermal (bottom row) spectra, from left to right: N, O and A.

(1 or 2) and in the case of the images the conditions at the time of acquisition (N, O or A).

Summary: The introduced database is devised as a compendium of images, videos, soft biometrics, and health parameters recorded from 52 different subjects in two sessions. It is composed of 612 and 416 face and shoulders images and 60-second videos respectively, corresponding to a total disk space of about 285 GB. The 52 recorded participants, 38 male and 14 female are from 13 different countries from 4 continents and their ages range between 22 and 51 years. An executive summary of the dataset is presented in Tab. II.

IV. PRELIMINARY ASSESSMENT OF THE DATABASE

In this Section we present a preliminary evaluation of thermal data for eHealth parameters estimation to assess the applicability of the database. The suitability of thermal imagery for a subject's weight estimation from face images is tested.

Weight estimation from face images:

Weight is a soft biometric trait and its estimation from a single facial shot has attracted interest in the research community in the latest years [3], [20], [21]. Besides being a soft biometric trait, weight is an indicator of a person's health condition, and unlike other biometric traits such as gender and height, body weight fluctuates during a person's life and needs to be periodically re-assessed. Remote estimation of this trait has been signaled of special interest in scenarios when a subject cannot be moved onto a scale due to different disabilities or in the case of road accidents. In such cases, estimating a person's weight from facial appearance allows for an inexpensive and contactless measurement [21]. Although some researchers have intended to reduce the error presented by AI-based contactless weight models, existing methods still present several kilograms (kg) of error. Weight estimation models from face data are typically evaluated on the public dataset VIP_attribute consisting of 513 female and 513 male face images of different celebrities and their associated height, weight and BMI metadata [3]. In 2018, Dantcheva *et al.* conducted for the first time a study on the possibility of

estimating bodyweight from a subject's face by implementing a ResNet architecture with 50 layers [3] and reported a Mean Absolute Error (MAE) of 8.15 kilograms (kg) of error and a Pearson's correlation coefficient (ρ) of $\rho = 0.77$. In 2020, Han *et al.* presented an auxiliary-task learning framework for weight estimation [20] with gender and age as auxiliary traits obtaining in the same dataset a MAE of 7.20 kg. In 2023 Mirabet-Herranz *et al.* defined an optimal transfer learning protocol for a ResNet50 architecture and experimented with different influencing factors such as hair occlusions [21] achieving a MAE of 6.91 kg and a $\rho = 0.78$.

Implementation details: Weight estimation from face images has proved to be possible using deep learning structures known as Residual Neural Networks (ResNet) with 50 layers and a final regression layer [3], [21]. We selected likewise to those studies a ResNet50 structure and we carry out a two-step Transfer Learning (TL) protocol as illustrated in Fig. 3. From a largely trained model intended for age estimation from face images, we complete TL using the visible images in our LVT training set. In the second part, we continue with the pipeline by performing once more TL this time with the thermal images belonging to the LVT training set. Finally, each weight network is tested in the images of the same spectrum found in the LVT test set. Each weight model was re-trained during 10 epochs and the final regression layer during 10 more epochs. The first 20 layers in each TL step were fixed to be frozen. Adam optimizer was adopted, with a learning rate of 0.01 and Huber loss as selected in [21] with $\delta = 1$.

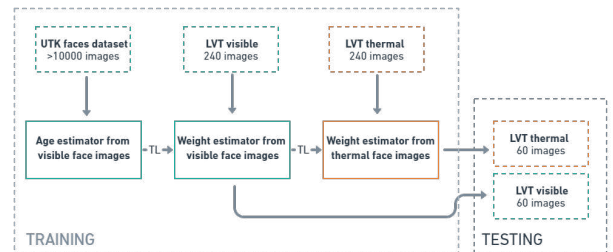


Fig. 3. Transfer learning protocol for weight estimation from visible and thermal images.

Visible-thermal experimental results:

It is known in the research community that bone, muscle and body fat do not conduct equally temperature [22]. Heat emission patterns can be used to characterize a person since they give information about the location of major blood vessels, skeleton thickness, amount of tissues, and muscle and fat amount¹. Therefore we believe thermal imagery will access crucial information for weight estimation from faces

¹<https://biometrics.mauguet.org/types/face.htm#thermogram>

TABLE II
SUMMARY OF THE INFORMATION CONTAINED IN THE LVT FACE DATABASE.

Identities	Metadata			
	Soft biometrics	Health parameters		
52 subjects	Gender	Body temperature	BMI	Body mass
2 sessions	Age	SpO2	Body fat (%)	Bone mass
Visuals (Thermal and visible)	Height	HR resting	Body water (%)	Proteins
	6 paired images in three conditions (N, O, A)	Weight	HR activity	Skeletal muscle
1 paired 60s videos subject rested	Biometric	BP maximum	Fat-free weight	Visceral fat
1 paired 60s videos after physical activity		ID	BP minimum	BMR
				Metabolic age

neglecting skin tissues-related noise and the impact of certain occlusions namely hair. The weight distribution associated with the subjects present in our database has a maximum value of 116.2 kg, a minimum value of 52.3 kg, a Mean=73.54 kg and a STD=14.03 kg. We perform a subject-exclusive split of the training set (480 images from 40 subjects) and the testing set (120 images from 12 remaining subjects). Several metrics are reported in our experiments: the above-mentioned correlation coefficient ρ and the MAE in kg, which are the most common units of measurement in weight estimation research; the root-mean-square error (RMSE) and the Percentage of Acceptable Predictions (PAP) used in [21] representing the percentage of the prediction whose error is smaller than 10% of the initial weight, i.e. a reasonable error in medical applications. In Tab. III the results of the weight network are presented. The metrics show that ResNet50 has a small advantage in the performance of weight estimation when re-trained using thermal data. Both the MAE and RMSE are lower for the thermal network at around 0.3kg. Moreover, the correlation coefficient between the predicted and original weight from the subjects is slightly higher for the thermal spectrum. This confirms the potential of thermal imagery for capturing hidden and more detailed information from human faces.

TABLE III
COMPARISON OF WEIGHT ESTIMATION FROM FACES BETWEEN THERMAL AND VISIBLE SPECTRA IMAGES.

Spectrum	MAE	RMSE	Correlation	PAP
Visible	8.31	15.03	0.43	61.6%
Thermal	7.98	14.73	0.49	61.6%

V. CONCLUSION

This paper presents the novel LVT Face Database for face biometrics. This database contains visuals from 52 subjects under different conditions, resulting in a total of 306 visible and 306 thermal images in addition to 204 visible and 204 thermal videos collected simultaneously using a paired camera (FLIR Duo R) allowing comparison or fusion of those different data types. The visuals acquired are associated with metadata belonging to the subjects both biometric- and health-related. To the best of our knowledge, this is the first database

to provide visible-thermal face images and recordings with accompanying gender, age, body temperature, SpO2, BP, HR (resting and after physical activity), height, weight, BMI and 11 additional health metrics. We believe the extensive amount of parameters annotated by every subject will help unlock the potential of thermal data for assessing a person’s health status. In addition, we provide preliminary experimental results of weight estimation from facial images using a baseline algorithm with ResNet50 architecture as a backbone, pre-trained with visible images. Results exhibit the potential of thermal data for contactless weight estimation. Based on this promising outcome, future work will focus on considering thermal imagery not only as an alternative to visible but also as a complement. The estimation of other parameters such as SpO2 or height from thermal depictions will be explored.

ACKNOWLEDGMENT

This work has been partially supported by the European CHIST-ERA program via the French National Research Agency (ANR) within the XAIface project (grant agreement CHIST-ERA-19-XAI-011).

REFERENCES

- [1] K. Mallat and J.-L. Dugelay, “A benchmark database of visible and thermal paired face images across multiple variations,” in *2018 International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 2018, pp. 1–5.
- [2] A. Ross, S. Banerjee, and A. Chowdhury, “Deducing health cues from biometric data,” *Computer Vision and Image Understanding*, vol. 221, p. 103438, 2022.
- [3] A. Dantcheva, F. Bremond, and P. Bilinski, “Show me your face and i will tell you your height, weight and body mass index,” in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018.
- [4] M. Wu, “Exploiting micro-signals for physiological forensics,” in *Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security*, 2020, pp. 1–1.
- [5] H. Rahman, M. U. Ahmed, S. Begum, and P. Funk, “Real time heart rate monitoring from facial rgb color video using webcam,” in *The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS), 2–3 June 2016, Malmö, Sweden*, no. 129. Linköping University Electronic Press, 2016.
- [6] Y. Lu, C. Wang, and M. Q.-H. Meng, “Video-based contactless blood pressure estimation: A review,” in *2020 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2020, pp. 62–67.
- [7] Y. Akamatsu, Y. Onishi, and H. Imaoka, “Blood oxygen saturation estimation from facial video via dc and ac components of spatio-temporal map,” in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

- [8] M. J. Eddine and J.-L. Dugelay, "Gait3: An event-based, visible and thermal database for gait recognition," in *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 2022, pp. 1–5.
- [9] M. Rai, T. Maity, and R. Yadav, "Thermal imaging system and its real time applications: a survey," *Journal of Engineering Technology*, vol. 6, no. 2, pp. 290–303, 2017.
- [10] A. Kuzdeuov, D. Koishigarina, D. Aubakirova, S. Abushakimova, and H. A. Varol, "Sf-tl54: A thermal facial landmark dataset with visual pairs," in *2022 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, 2022, pp. 748–753.
- [11] D. Anghelone, C. Chen, P. Faure, A. Ross, and A. Dantcheva, "Explainable thermal to visible face recognition using latent-guided generative adversarial network," in *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*. IEEE, 2021, pp. 1–8.
- [12] X. Kevin and W. Bowyer, "Visible-light and infrared face recognition," in *Workshop on Multimodal User Authentication*. Citeseer, 2003, p. 48.
- [13] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen, and X. Wang, "A natural visible and infrared facial expression database for expression recognition and emotion inference," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 682–691, 2010.
- [14] T. Gault and A. Farag, "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*, 2013, pp. 336–341.
- [15] K. Panetta, Q. Wan, S. Agaian, S. Rajeev, S. Kamath, R. Rajendran, S. P. Rao, A. Kaszowska, H. A. Taylor, A. Samani *et al.*, "A comprehensive database for benchmarking imaging systems," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 3, pp. 509–520, 2018.
- [16] C. Barbosa Pereira, M. Czaplik, V. Blazek, S. Leonhardt, and D. Teichmann, "Monitoring of cardiorespiratory signals using thermal imaging: a pilot study on healthy human subjects," *Sensors*, vol. 18, no. 5, p. 1541, 2018.
- [17] M. Abdrakhmanova, A. Kuzdeuov, S. Jarju, Y. Khassanov, M. Lewis, and H. A. Varol, "Speakingfaces: A large-scale multimodal dataset of voice commands with visual and thermal video streams," *Sensors*, vol. 21, no. 10, p. 3465, 2021.
- [18] D. Poster, M. Thielke, R. Nguyen, S. Rajaraman, X. Di, C. N. Fondje, V. M. Patel, N. J. Short, B. S. Riggan, N. M. Nasrabadi *et al.*, "A large-scale, time-synchronized visible and thermal face dataset," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 1559–1568.
- [19] K. S. Nair and S. Sarath, "Illumination invariant non-invasive heart rate and blood pressure estimation from facial thermal images using deep learning," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2021, pp. 1–7.
- [20] D. Han, J. Zhang, and S. Shan, "Leveraging auxiliary tasks for height and weight estimation by multi task learning," in *2020 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2020, pp. 1–7.
- [21] N. Mirabet-Herranz, K. Mallat, and J.-L. Dugelay, "New insights on weight estimation from face images," in *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*. IEEE, 2023, pp. 1–6.
- [22] M. Morley, "Thermal conductivities of muscles, fats and bones," *International Journal of Food Science & Technology*, vol. 1, no. 4, pp. 303–311, 1966.