Int. J. Nonlinear Anal. Appl. In Press, 109–118
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.21208.4363

Large scale objects thermography and thermal imaging survey: Datasets and applications

Ahmed Abdulsatar Kareem, Adil Al-Azzawi*

Computer Science Department, College of Science, University of Diyala, Diyala, Baqubah, Iraq

(Communicated by Seyed Hossein Siadati)

Abstract

Due to machine learning-based infrared image and video collections, which enable computers to detect and categorize images with increasing accuracy, the field of image identification has experienced a revolution. This comprehensive research presents an overview of the most current advancements in the infrared image and video collections for computer vision and artificial intelligence. It has largely focused on the infrared picture and video collections that have been collected and categorized for computer vision applications such as object identification, object segmentation and classification, and motion detection. This article covers some of the most well-known machine learning methods, including deep learning, convolutional-neural-networks, support-vector machines, and decision trees. The basic problems with image identification are examined, and only a few of them include data augmentation, feature extraction, and picture segmentation. We also discuss some recent developments in the area of image identification, including ground-breaking deep learning methods like adversarial training and transfer learning. The discussion ends with a discussion of possible uses and the promise of machine learning for picture identification. Because it analyzes state of the art in machine learning for picture identification in-depth, this survey study is a vital resource for academics and entrepreneurs. We make a distinction between publicly accessible collections and those that are maintained in private, based on the various sensor types, image resolution, size, and research effort within that range. Include a glossary of words, including those for infrared radiation, infrared detectors, and infrared optics, that are crucial to comprehending infrared imaging, along with a description of their applications. This article explores the group's overall statistical relevance from a number of different perspectives. Researchers working in computer vision and artificial intelligence who are interested in managing spectra outside of the optical field might use this survey as a reference.

Keywords: artificial intelligence, deep-learning, convolutional-neural-networks, learning-based infrared, machine learning
2020 MSC: 68T05, 68T07

1 Introduction

Due to unclear symptoms or even oligosymptomatic presentation, the Covid-19 pandemic made diagnosing flu-like disorders more difficult. However, an effective screening tool is crucial for the control of highly infectious illnesses as it enables more effective medical-epidemiological operations and the wise allocation of financing for global health [18]. Face thermography, which may be used in a variety of medical, surveillance, and environmental monitoring

*Corresponding author
Email addresses: scicompms2205@uodiyala.edu.iq (Ahmed Abdulsatar Kareem), adil_alazzawi@uodiyala.edu.iq (Adil Al-Azzawi)

Received: February 2023 *Accepted:* April 2023

applications, is one of the most thoroughly investigated issues in infrared-thermal imaging. It is challenging to further research in this area due to the dearth of publicly available datasets of faces in the visible spectrum compared to faces in thermal images [4]. The Corona Virus 2019 was discovered in Wuhan, China, in December 2019, and it swiftly swept the globe. This virus killed more than 1.6 million grandfathers and infected more than 70 million individuals in a single year [17]. There have been 6,853,702 deaths and 672,789,992 confirmed cases globally as of this writing. Therefore, Covid-19 must be accurately and swiftly recognized in order to curtail its spread and lower the death toll [28]. The study found that although while Covid 19 pneumonia shares many clinical characteristics with other types of pneumonia, persons with Covid 19 have liver function loss more frequently than those who don't. This suggests that Covid 19 disease is a dangerous condition with a high likelihood of migrating to people. The disease's signature symptom is an elevated body temperature. Temperature testing is thus one of the finest ways to detect patients. Body temperature may be determined using two main techniques:

The inspector must be close to pedestrians even though the infrared short-distance body temperature meter is easy to use and reasonably priced; a thermal camera is used to remotely measure an object's temperature while an infrared thermometer is used to measure temperature over a short distance.

It is risky when the virus has infected the pedestrian. In order to provide a secure evaluation of body temperature, thermal cameras that can track body temperature from a distance have been developed [17].

2 Public database

There are several free visual-thermal face datasets available, as indicated in Table 1. There are currently no notable databases that contain the 3 data streams, which consist of synchronized-visible spectrum photos, thermal pictures, and audio recordings. A lack of participants, a dearth of noteworthy cases (which frustrates data hungry-machine-learning-algorithms), poor therm picture quality, a lack of variation in head orientations, and a lack of alignment are problems that plague the majority of the current visual-thermal face datasets [1].

Table 1: lists the datasets that were concurrently captured using thermal and visual imaging and are now accessible [1].

datasets	Subjects	Image Pairs	IR Resolution	Poses	Trials	Aligned
Carl [10]	41	2460	160×120	1	1	no
VIS-TH [19]	50	2100	160×120	4	2	yes
IRIS [14]	30	4228	320×240	11	1	no
USTC-NVIE [27]	215	N/A	320×240	1	1	no
Tufts [20]	100	3600	336×256	9	1	no
UL-FMTV [13]	238	N/A	640×512	1	> 1	N/A
ARL-VTF [21]	395	549,712	640×512	3	1	no
SpeakingFaces [1]	142	4,581,595	464×348	9	2	yes

3 Face temperature measurement system

Compared to current face thermal-based technologies, a biometric system using a face thermal temperature sensors provides a number of advantages. The first explanation for this is the face thermal on the finger, an internal characteristic that is challenging to duplicate. Furthermore, environmental conditions largely have little impact on the recorded face thermal's quality. A more precise connection between temperature data and face characteristics is now possible thanks to newly developed high-resolution thermal cameras. It has been shown that combining The limitations of each stream alone can be addressed using thermal and optical data [8]. It has been discovered that adding visual information to voice signals can help speech recognition and person verification models [2, 11, 25]. The whole automated COVID-19 screening method, including the temperature check and mask check, is shown in Figure 1 [16].

3.1 Thermal face sample

The VIS-TH database had the fewest aligned image pairs and lowest thermal camera resolution while having two trials for each subject with four head rotations. Despite the smaller sample size, every participant's face was photographed from 11 different perspectives in the IRIS collection. Data from a sizable number of individuals are

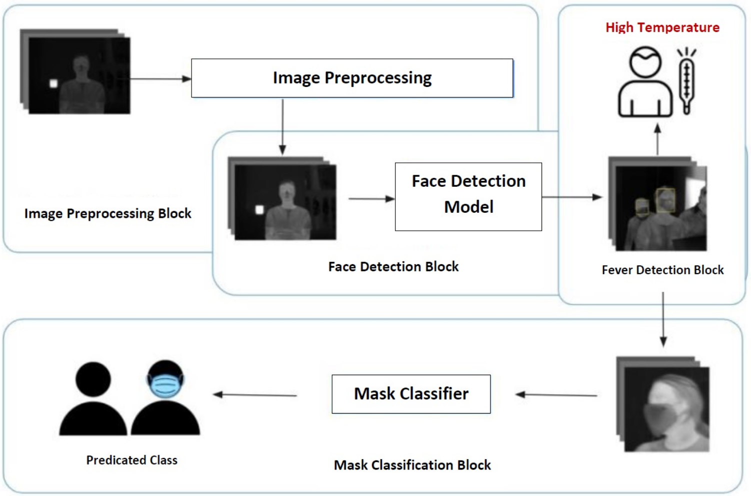

Figure 1: Block diagram showing the whole COVID-19 screening pipeline, including a mask and temperature check [16].

included in the USTC-NVIE dataset, even though there was only one trial that was caught on camera from a single location with a low-resolution camera. There are many head postures in the Tufts collection, despite the fact that each person only appears in a few images. Multiple trials are only included in UL-FMTV when seen from the front. Even though ARL-VTF is the most complete in terms of participants, pictures, and thermal resolution, in terms of trial length and head orientations, it is inferior. Figure 2 [10] is a set of images of the same person wearing glasses which were recorded in the thermal, infrared, and visible spectrums. Figure 3 [19] displays the results of integrating these images with a discrete wavelet transform (DWT).

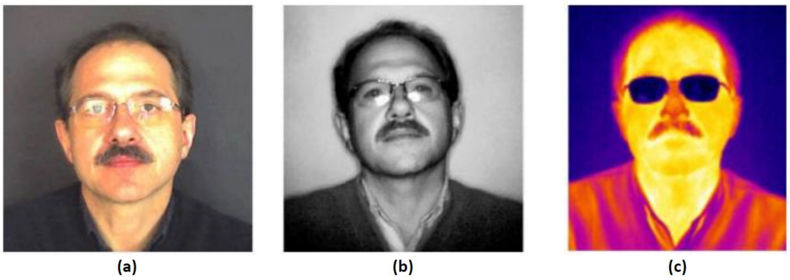

Figure 2: The same individual wearing glasses is depicted in a series of photos in (a) VIS, (b) NIR, and (c) THIR [10].

Figure 3: Three types of images were combined using the discrete wavelet transform (DWT): (a) visual, (b) thermal, and (c) sensor-level image [19].

SpeakingFaces was created in order to get over the drawbacks of the multimodal datasets that are currently accessible. The 142 SpeakingFaces topics are varied in terms of race and gender. While being closely observed from nine various perspectives, each subject is recorded while they speak over 100 English words or provide urgent instructions. More than 13000 spoken command occurrences and more than 45 hours of footage are the outcomes (3.7 million image pairings and more). The spoken words were created using the open-source digital assistant database from Stanford University and the publicly accessible Siri command sets, both of which reflect common ways that people interact with technology. Applications involving HCI, biometrics, and recognition systems in particular are appropriate for the SpeakingFaces dataset, which may be utilized for a variety of multimodal machine-learning tasks [1].

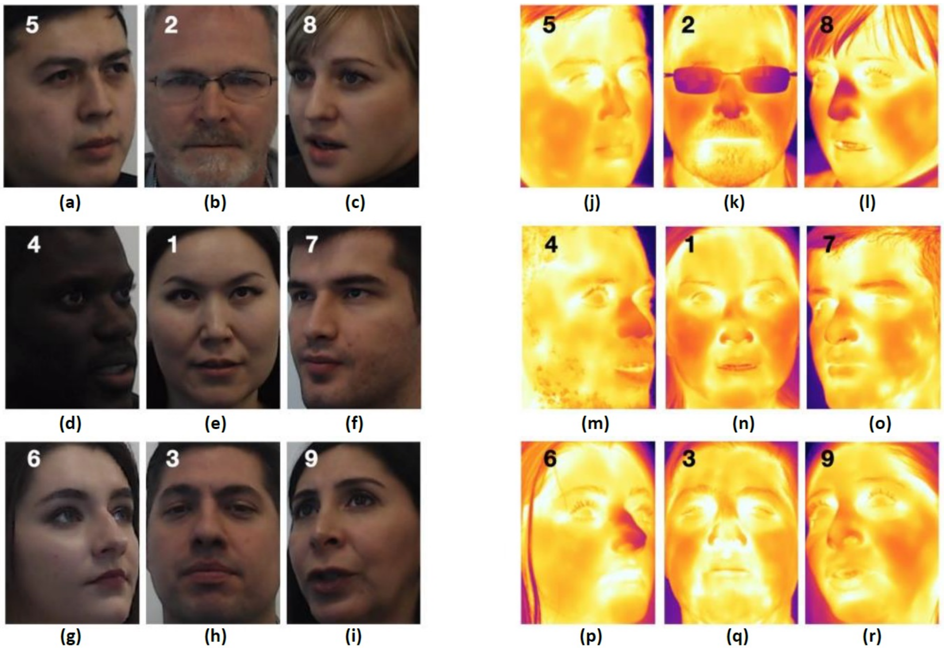

Figure 4: Shows the pairs of facial images—both thermal and visual—of nine people obtained from the predetermined positions (a), (b), (c), (d), (e), (f), (g), (h), (i) is shown visual facial images, while (j), (k), (l), (m), (n), (o), (p), (q), (r) is shown thermal facial images [1].

Figure 5: Pedestrian Infrared/visible Stereo Video Dataset (a) Infrared image, and (b) visible image [14].

Figure 6: An angry figure is depicted in (a) a visual-picture and (b) a thermal- picture [27].

3.2 Data collection

The database structure, data preparation procedures, and data gathering techniques are all covered in detail in this section. Simultaneous video and audio clips are retrieved before the data is ready for analysis in sessions

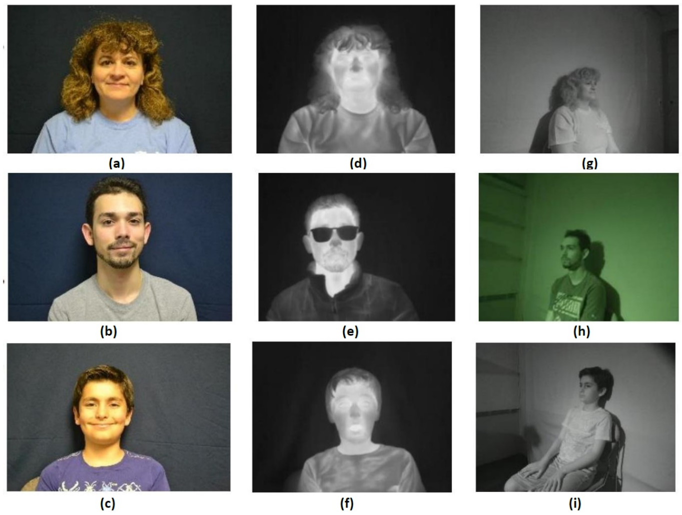

Figure 7: Illustrations images from the Tufts Face Database (a), (b), (c) 2D visible-images, (d), (e), (f) Thermal-images, and (g), (h), (i) Near-infrared (NIR) images [20].

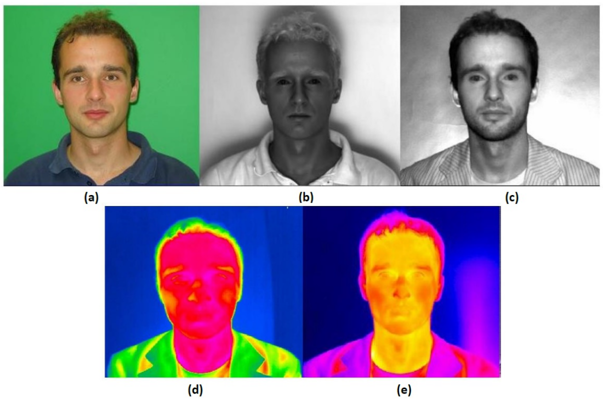

Figure 8: UL-FMTV's list of topics includes (a) visible-image, (b) SWIR-image, (c) NIR-image, (d) LWIR-image, and (e) MWIR-Image [13].

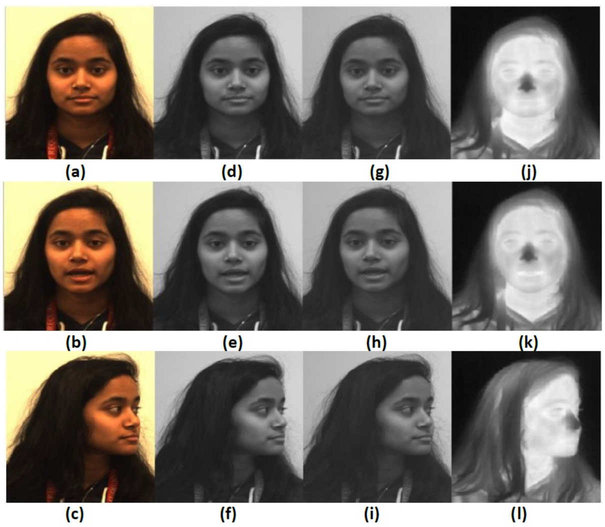

Figure 9: A set of images (a), (b), and (c) RGB-images, stereo black-and-white and LWIR cameras from the starting point (j), (k), and (l), expression, and (d), (e), (f), (g), (h), and (i), off-pose sequences [21].

when instructions are stated. From each video clip from both sessions, visual sequences were then created [1]. The visualizations were then produced in relation to their thermocouple pairs using ArUco hot tags [12].

3.3 Preprocessing

In the prior study, in trials when individuals sat without receiving any instructions, unprocessed video clips were turned into visual sequences (900 frames per position). Raw audio and video data from the speaking tests were then segmented into digestible pieces in accordance with each speech's start and end frame stamps. After that, the audio recordings were rigorously edited to leave no more than one second at the conclusion of each statement in order to account for the variations in reading speed among our participants. The files were also examined for accuracy and for text noise like hesitation or stuttering. In order to eliminate textual data noise, valid recordings were retranscribed to accurately represent voice. The length of the resulting audio files was used to determine how the movies were subsequently transformed into visual sequences. If the text noise, in addition to the routing frequencies and stumbles, was considerable, the voice was removed from the final dataset [1].

3.4 ROI extraction

The study provides models for precisely identifying faces from visual imagesFaster-RCNN [22], RFCN [23], YOLOv3 and Mobile net V2-SSD are a few examples. We considered the YOLOv3 and Mobile net V2-SSD architectures since they can run on top-tier hardware with low processing power and a high inference rate. Both of these architectures are trained using training datasets made up of thermal surveillance and enhanced surveillance datasets in order to assess the effectiveness of thermal-based face detection. Using an 8-sample batch size, the dataset samples were jumbled during training. Weight decay was set at 0.95 and 0.9 for the modified model weights, which were updated using the Adam optimizer. After 5000 trials, the learning rate, which was set at 1e-4, decreases tenfold. Performance measures such as mean Intersection Over Union (mean IOU) and mean Average Accuracy was used to compare various designs (MAP)[16]. The picture displays in figure 10 a few examples of thermal and visual photos that participants captured under varied lighting conditions.

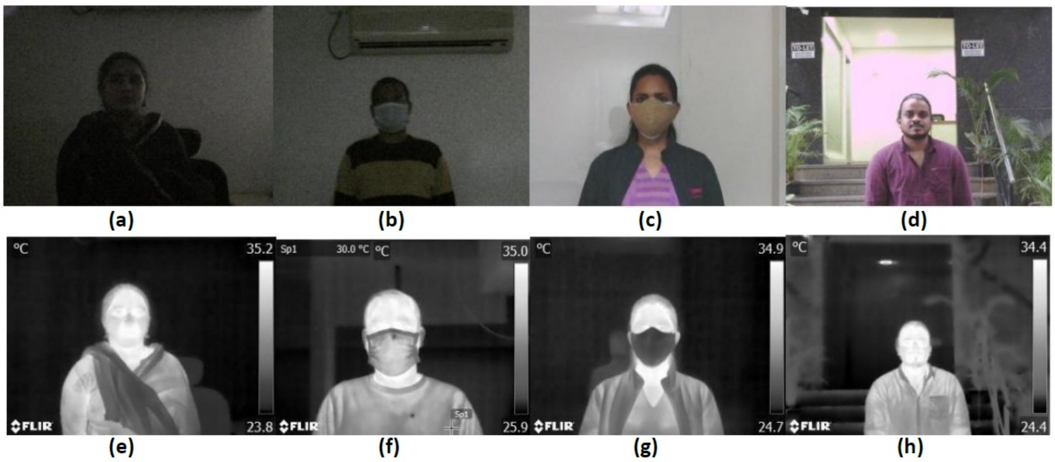

Figure 10: Examples of participant thermal and visual photos taken under various lighting conditions from the lighting dataset, (a)(b)(c), and (d) visual images, (e),(f),(g), and (h) thermal images [16].

4 Face temperature measurement system using machine learning

This section illustrates how machine learning techniques like neural networks, fuzzy logic, and evolutionary computing may aid biometrics algorithms. Machine learning's tolerance to noise and ability to systematically adjust are its two key properties. Machine learning also has a parallel structure. It has the capacity to autonomously manipulate. It works great for face extraction and conformation using heat. Machine learning has been applied in a few studies to determine the temperature of the thermal face. Since the data set for thermal films of humans are either nonexistent or extremely poor, the vast majority of earlier attempts to determine the temperature of persons by thermal face in thermal movies depended on internal databases. Table 2 lists the algorithms used in the many published machine-learning studies.

Table 2: Machine-learning-based algorithms for facial temperature measuring systems.

Reference	Database	Preprocessing	Segmentation, Feature extraction	Classifier
[16]	developed internally(NTIC)	ROI Extraction	Convolution	1. visual trained 2. Thermal trained
[5]	LFW [15]	1. ROI extraction 2. Image resize	Binary Pattern	1. MobileNet 2. DenseNet
[17]	DenseNet developed internally (using FLIR ONE Pro)	ROI Extraction	marked point matches of 10	ROC
[26]	Developed internally (AMG8833, Pi camera)	ROI Extraction	Face recognition	SVM
[24]	developed internally	ROI Extraction	–	–
[6]	developed internally	ROI Extraction	CNN	–
[18]	developed internally	ROI Extraction	Random Forest (RF)	–
[3]	developed internally	ROI Extraction	FCM clustering	AdaBoost
[1]	developed internally	ROI Extraction	ArUco markers	1. CNN 2. BRNN
[7]	–	ROI Extraction	DMF	–

A overview of each earlier work's accuracy is given in the table 3. Few studies have used machine learning techniques to quantify thermal facial temperature.

Table 3: Accuracy of thermal facial temperature detection using machine learning.

Reference	Number of subjects	Number of test images	Accuracy(%)
[16]	1,354	902	97.00
[5]	5,760	3 million	99.38
[17]	–	12,000–5,590	95.00
[26]	–	–	–
[24]	20	27,000	–
[6]	227,261	–	81.20-91.20
[18]	136	302,400	80.00
[3]	–	3,674	96.20
[1]	142	4,581,595	90.00
[7]	–	402	–

5 Literature review

The thermal imaging industry has grown significantly owing to the extraordinary rate of fresh research and technology development in recent years. As a result, it has become harder for academics to stay up to date with the most recent advancements in the profession. To address this issue, literature reviews have emerged as a critical component of thermal imaging research.

A complete study of the volume of material that has already been written about a certain subject or research issue is known as a literature review. The main objective of a literature review is to offer a comprehensive assessment of the level of knowledge currently available on a subject, highlighting key results and flagging any gaps in the literature that would call for more research.

In the field of thermal imaging, literature evaluations are crucial for establishing research questions, pinpointing knowledge gaps, and directing the development of novel theories and hypotheses. A thorough examination of the literature may offer insightful insights into how research findings are used and suggest directions for future research.

This evaluation of the literature's objectives aims to (insert specific objective of the literature review). This study attempts to identify the major results, emphasize any gaps in the literature that may call for more research, and offer a complete account of the present state of knowledge on this issue by reviewing the pertinent literature (insert topic or research question). This study's objective is to offer pertinent implications and potential areas for future research based on the findings of the reviewed literature. Overall, this literature review provides a strong foundation for future research in this area and advances our knowledge of thermal imaging.

- **Detection of Objects in Thermal images**

  The method used in their study's search for traits that stand out or attract attention in photographs and videos. Tracking, object segmentation and recognition, picture and video content reduction, and video detection are just a few of the numerous applications that SOD is utilized in. Although RGB properties are sensitive to light intensity, the difficulties of distinguishing important items in dimly lit and congested areas limit the practical applications of the SOD approach. In recent times, RGB-based algorithms have performed significantly better. Due to the fact that infrared thermography is unaffected by light levels and can significantly increase SOD's efficiency in poorly lit and congested environments, RGB-Thermal SOD has gained a lot of interest and has been successful [29]. Figure 11 displays the salinity maps for 10 typical SOTA SOD units and a proposed LSNet.

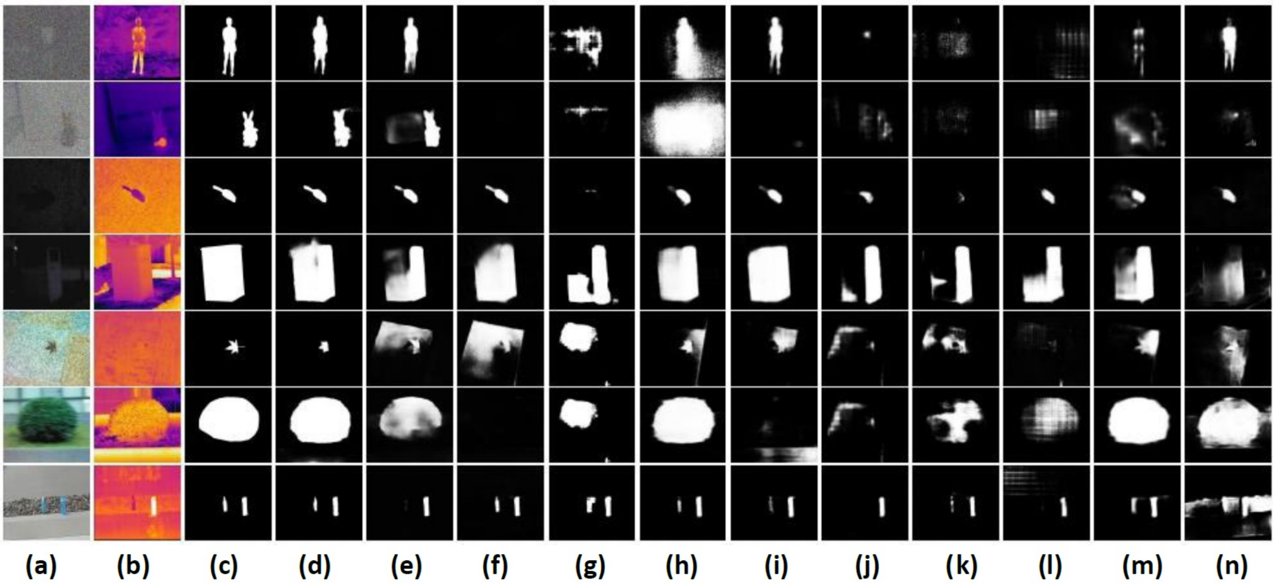

Figure 11: Ten examples of SOTA SOD saliency maps and the suggested LSNet, (a) RGB, (b) Thermal, (c) GT, (d) Ours, (e) MIDD, (f) ADF, (g) FMCF, (h) JLDCF, (i) S2MA, (j) TANet, (k) AFNet, (l) MMCI, (m) CPD, and (n) R3Net images [29].

- **Detection of animals in thermal images**

  Using infrared thermography, the short-beaked echidna's point count was calculated remotely. Sea urchins might be seen in the Wheatbelt region of Western Australia, some 170 kilometers south of Perth, at the Boyagin Nature Reserve and the Dryandra Woodland. The scientists shot thermal films from a distance of 1 to 20 meters Using an 83.4mm, 12° telephoto lens on a FLIR T1050sc infrared camera. The camera was calibrated using a global standard (FLIR, Oregon, the USA). For a total of 34 days in the years 2020 and 2021, thermal measurements were gathered on 124 echidnas once per 12 months (from January to December). Although it was impossible to completely rule out the possibility that some echidnas were measured more than once on different days because they couldn't usually be identified individually, in order to reduce the risk of people being measured twice, recordings were done across a wide region in both reserves. The microclimate data Wet bulb globe temperature (WBGT), relative humidity (RH), and ambient temperature (Ta) were measured using a Kestrel 5400 portable weather station between 0 and 20 meters from the scene of the echidna shooting and under similar sun radiation circumstances [9].

- **Human faces detection in thermal pictures**

  The methods, techniques, and algorithms utilized to recognize people in thermal images or videos as well as the datasets that were used to create the models were all disclosed in this study. This sector has expanded

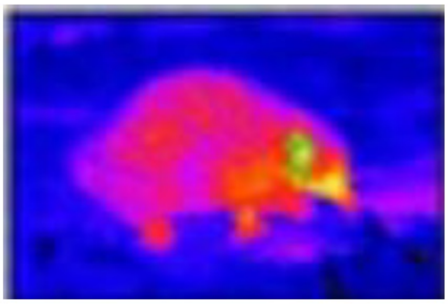

Figure 12: In the West Australian wheatbelt, a representative dataset for the active short-beaked echidna species Tachyglossus aculeatus was captured using an infrared camera [9].

significantly over the last three years as a result of the Covid 19 virus explosive expansion.

6 Conclusion

We provide a list of publicly accessible databases with infrared photos and videos for researchers involved in this project who are working in the fields of artificial intelligence and computer vision. Our main research interests are in the infrared image and video collections that are collected and processed for computer vision tasks such as object identification, object segmentation and classification, and motion detection. In accordance with the sensor kinds, picture resolution, size, and research effort within that range, we classify publically accessible or private collections. Infrared radiation, infrared detectors, infrared optics, and their potential applications are also covered in this introduction to important infrared imaging topics. From several angles, we examine the statistical significance of the entire group. Researchers in computer vision and artificial intelligence who desire to explore working with spectra outside of the visual field might use this survey as a reference.

References

[1] 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 **21** (2021), no. 10, p. 3465.

[2] T. Afouras, J.S. Chung, A. Senior, O. Vinyals and A. Zisserman, *Deep audio-visual speech recognition*, IEEE Trans. Pattern Anal. Machine Intel. **44** (2018), no. 12, 8717–8727.

[3] H. Al-Sadr, M. Popescu and J.M. Keller, *Early sepsis recognition based on infrared thermography*, Int. J. Data Min. Bioinf. **22** (2019), no. 4, 301–327.

[4] R. Ashrafi, M. Azarbayjani and H. Tabkhi, *A novel fully annotated thermal infrared face dataset: recorded in various environment conditions and distances from the camera*, Infrared Phys. Technol. **124** (2022).

[5] A. Barnawi, P. Chhikara, R. Tekchandani, N. Kumar and B. Alzahrani, *Artificial intelligence-enabled internet of things-based system for COVID-19 screening using aerial thermal imaging*, Future Generat. Comput. Syst. **124** (2021), 119–132.

[6] M.L. Brioschi, C.D. Neto, M. de Toledo, E.B. Neves, J.V.C. Vargas and M.J. Teixeira, *Infrared image method for possible COVID-19 detection through febrile and subfebrile people screening*, J. Thermal Bio. **112** (2023), 103444.

[7] Q. Cai, D. Gallup, C. Zhang and Z. Zhang, *3d deformable face tracking with a commodity depth camera*, Comput. Vision–ECCV 2010: 11th Eur. Conf. Comput. Vision, Heraklion, Crete, Greece, Proc. Part III 11, Springer Berlin Heidelberg, 2010, pp. 229–242.

[8] Z. Chen, S. Wang and Y. Qian, *Multi-modality matters: A performance leap on VoxCeleb*, INTERSPEECH, October 25–29, 2020, Shanghai, China, 2020, pp. 2252–2256.

[9] C.E. Cooper and P.C. Withers, *Postural, pilo-erective and evaporative thermal windows of the short-beaked echidna (Tachyglossus aculeatus)*, Bio. Lett. **19** (2023), no. 1, p. 20220495.

[10] V. Espinosa-Duró, M. Faundez-Zanuy and J. Mekyska, *A new face database simultaneously acquired in visible, near-infrared and thermal spectrums*, Cognit. Comput. **5** (2013), 119–135.

[11] R. Gade and T.B. Moeslund, *Thermal cameras and applications: A survey*, Mach. Vis. Appl. **25** (2014), 245–262.

[12] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas and M.J. Marín-Jiménez, *Automatic generation and detection of highly reliable fiducial markers under occlusion*, Pattern Recognit. **47** (2014), 2280–2292.

[13] R.S. Ghiass, H. Bendada and X. Maldague, *Université laval face motion and time-lapse video database (UL-FMTV)*, Technical Report; Université Laval: Québec, QC, Canada, 2018.

[14] R.I. Hammoud, *IEEE OTCBVS WS series bench*, http://vcipl-okstate.org/pbvs/bench/.

[15] G.B. Huang, M. Mattar, T. Berg and E. Learned-Miller, *Labeled faces in the wild: A database for studying face recognition in unconstrained environments*, In Workshop on faces in'Real-Life'Images: detection, alignment, and recognition, 2008.

[16] P. Katte, S.T. Kakileti, H.J. Madhu and G. Manjunath, *Automated thermal screening for COVID-19 using machine learning*, Artific. Intell. over Infrared Images Med. Appl. Med. Image Assisted Biomarker Discovery, MIABID AIIIMA 2022, Lecture Notes in Computer Science, Springer, Cham, 2022.

[17] Z. Ma, H. Li, W. Fang, Q. Liu, B. Zhou and Z. Bu, *A cloud-edge-terminal collaborative system for temperature measurement in COVID-19 prevention*, IEEE INFOCOM 2021-IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), IEEE, 2021, pp. 1–6.

[18] A.C. Makino Antunes, A. Aldred, G.P. Tirado Moreno, J.A. de Souza Ribeiro, P.E. Brandão, G.T. Barone, J.D.A. Conselheiro, A.C. Goulart, I.C. Desuo and G. Gomes, *Potential of using facial thermal imaging in patient triage of flu-like syndrome during the COVID-19 pandemic crisis*, PloS one **18** (2023), no. 1, 0279930.

[19] K. Mallat and J.L. Dugelay, *A benchmark database of visible and thermal paired face images across multiple variations*, Int. Conf. Biomet. Spec. Interest Group (BIOSIG), IEEE, 2018, pp. 1–5.

[20] K. Panetta, Q. Wan, S. Agaian, S. Rajeev, S. Kamath, R. Rajendran, S.P. Rao, A. Kaszowska, H.A. Taylor, A. Samani and X. Yuan, *A comprehensive database for benchmarking imaging systems*, IEEE Trans. Pattern Anal. Mach. Intell. **42** (2018), 509–520.

[21] 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 and S. Hu, *A large-scale, time-synchronized visible and thermal face dataset*, Proc. IEEE/CVF Winter Conf. Appl. Comput. Vision, 2021, pp. 1559–1568.

[22] S. Ren, K. He, R. Girshick and J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, Adv. Neural Info. Process. Syst. **28** (2015), 91–99.

[23] D.J. R-FCN, *Object detection via region-based fully convolutional networks*, Adv. Neural Inf. Process. Syst., 2016, pp. 379–387.

[24] A. Scarano, F. Inchingolo and F. Lorusso, *Facial skin temperature and discomfort when wearing protective face masks: thermal infrared imaging evaluation and hands moving the mask*, Int. J. Environ. Res. Public Health **17** (2020), no. 13, 4624.

[25] S. Shon, T. Oh and J. Glass, *Noise-tolerant audio-visual online person verification using an attention-based neural network fusion*, ICASSP 2019-2019 IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP), IEEE, 2019, pp. 3995–3999.

[26] R.S.S. Singh, T.J.S. Anand, S.A. Anas and B. Acharya, *A real-time analytic face thermal recognition system integrated with email notification*, Eng. Technol. Appl. Sci. Res. **13** (2023), no. 1, 9961–9967.

[27] 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 Trans. Multimed. **12** (2010), 682–691.

[28] World Health Organization (WHO), *https://www.who.int*.

[29] W. Zhou, Y. Zhu, J. Lei, R. Yang and L. Yu, *LSNet: Lightweight spatial boosting network for detecting salient objects in RGB-thermal images*, EEE Trans. Image Process. **32** (2023), 1329–1340.