Cross-Spectral Image Registration: a Comparative Study and a New Benchmark Dataset

Rafael E. Rivadeneira, Henry O. Velesaca, and Angel Sappa

Abstract The field of cross-spectral imaging has significantly advanced, driven by its diverse applications, including environmental monitoring and medical imaging enhancements. The integration of images from different parts of the electromagnetic spectrum, particularly the fusion of thermal and visible images, is a crucial task for different applications. It provides a comprehensive picture of a scene, combining the clarity of visible light imaging with the contrast of thermal imaging. This research investigates the efficiency of various techniques and architectures for local feature matching between visible and thermal images, essential for accurate image registration. Through evaluating a wide array of methods against a novel acquired cross-spectral dataset encompassing varied real-world scenarios, the study provides detailed insights into their effectiveness and limitations under different conditions. It also presents quantitative benchmarks on computational speed, offering a clearer perspective on each method's performance and applicability in practical, especially resource-constrained, settings. The results indicate that these architectures exhibit remarkable capabilities in accurately and efficiently registering images from the thermal and visible domains. Their inherent flexibility in handling complex problems, along with their computational speed, suggests that these approaches hold significant promise for addressing cross-spectral imaging challenges. The dataset is available at:https://github.com/vision-cidis/CIDIS-dataset.

R. E. Rivadeneira (✉) · H. O. Velesaca · A. Sappa
ESPOL Polytechnic University, Escuela Superior Politécnica Del Litoral, ESPOL, Guayaquil, Ecuador
e-mail: rrivaden@espol.edu.ec

H. O. Velesaca
e-mail: hvelesac@espol.edu.ec; hvelesaca@correo.ugr.es

A. Sappa
e-mail: asappa@espol.edu.ec; sappa@ieee.org

H. O. Velesaca
Software Engineering Department, University of Granada, 18014 Granada, Spain

A. Sappa
Computer Vision Center, 08193 Bellaterra, Barcelona, Spain

S. Roy et al. (eds.), *Innovations in Computational Intelligence and Computer Vision*, Lecture Notes in Networks and Systems 1117,
https://doi.org/10.1007/978-981-97-6992-6_1

Keywords Thermal imaging · Deep learning · Cross-spectral registration · Cross-spectral dataset · Computer vision · Image-registration

1 Introduction

Multi-spectral imaging applications, which is based on the usage of images from different parts of the electromagnetic spectrum, are becoming quite popular. The applications go from remote sensing (e.g., Earth's monitoring [1, 2]) till medial domain (e.g., examining tissues in medical fields [3, 4]). One of the main challenges in this area is accurately combining images from different spectral bands, especially when looking at the differences between images that do not overlap in the spectral band (e.g., thermal and visible images). Combining these images correctly is essential for tasks that require a comprehensive representation of a scene, utilizing information from various sources. However, prior to merging information from different sources, it is crucial to ensure proper registration. It is worth noting that the outcome of the final application will be contingent upon the accuracy of the registration process (refer to [5] for an in-depth review of state of the art approaches).

Recent advances in image registration have been marked by the development of methods like the LightGlue architecture [6], employing deep neural networks to enhance feature matching in images. An examination of these methods reveals that the effectiveness of any given technique varies with the task at hand and the specific metrics used to evaluate the performance.

In this study, a newly acquired dataset consisting of 1000 pairs of images (visible and thermal spectrum images) is proposed, precisely captured under varying lighting conditions to ensure high quality and detail from visible spectrum images, making them especially suitable for training guided image super-resolution techniques. This dataset, with its clear images and accurate registration, also serves as a new benchmark for evaluating multi-spectral image fusion techniques. A detailed evaluation of several techniques is performed using this dataset, seeking to understand its strengths and limitations fully. Insights into the capabilities of current multi-spectral imaging methods are provided and offer direction for future enhancements, to achieve more accurate and efficient image fusion in diverse applications.

The paper structure is organized as follows: Sect. 2 provides a comprehensive review of the existing literature relevant to image registration and cross-spectral datasets. Section 3 details the proposed methodology, outlining the processes of image registration, dataset acquisition, and metrics evaluations. The setup for the experiments, including the acquired dataset, and evaluation metrics is presented in Sect. 4. The paper concludes with Sect. 5, summarizing the findings and implications of the research.

2 Related Work

In the field of image registration, particularly with a focus on multi-spectral imaging, several techniques have emerged as significant contributors to the state-of-the-art. This section reviews the current status of notable methods such as Elastix [7], Imregister [8], LightGlue [6] (including its variants LightGlue+SP and LightGlue+D), and Nemar model [9], highlighting their unique contributions and applications in image alignment and processing.

2.1 Image Registration Techniques

The first technique reviewed is **Elastix** [7], which is a modular collection of algorithms for image registration, highly regarded for its flexibility and robustness across different imaging scenarios. It incorporates a variety of transformation models and similarity metrics, adaptable for diverse registration tasks. Particularly in medical imaging, Elastix has demonstrated its utility in ensuring precise alignment of images, contributing to improved diagnostic accuracy and treatment planning.

The next technique evaluated in the current work is **Imregister** [8]; it is a method known for its MATLAB implementation, offering a user-friendly interface for image registration. It supports multiple transformation types, including affine and non-rigid transformations. Its adaptability and ease of use has made it popular in academic research and practical applications, especially in fields where rapid prototyping of image registration solutions is required. For example, Shrestha et al. [10] tackle the registration problem with the Imregister MATLAB implementation. This technique, like Elastix, is based on the intensity image registration approach.

On the contrary to previous solutions, which are based on classical image processing approaches, in recent years different learning-based solutions have been proposed. The **Nemar model** [9] for instance, is a recent approach known for its novel use of neural networks in image registration, particularly in challenging scenarios with complex image content or noise. This unsupervised method effectively tackles multi-modal image registration by employing an innovative geometry-preserving image translation technique. It has demonstrated promising results in scenarios where traditional methods face difficulties. The approach proves its efficacy in achieving accurate and robust registration across imaging modalities without the need for annotated data, addressing the challenge of unsupervised registration in situations where obtaining reference data is challenging or expensive. For instance, Xu et al. [11] use the Nemar model among other techniques for the multi-modal image registration and fusion study.

Finally, the last technique reviewed in the current work is **LightGlue** [6], which represents a significant advancement in feature matching and image registration, utilizing deep learning techniques for enhanced accuracy and efficiency. Its variants, LightGlue+SP (LightGlue+SuperPoint) and LightGlue+D (LightGlue+DISK), introduce additional layers of sophistication. The SuperPoint technique [12] operates as a method for detecting and describing points of interest in images, using a self-learning approach (self-supervised). It is based on a convolutional neural network that is trained in an unsupervised manner to automatically learn the locations and descriptions of points of interest in the input images. The self-learning process is performed by generating attention maps and matching key points on paired images. On the other hand, the DISK (Learning Local Features with Policy Gradient) technique [13] works by learning local features in images through the optimization of a policy using the reward gradient. It uses a neural network to learn robust representations of local features, such as corners or edges, using reward signals generated during the training process.

Each of the mentioned techniques makes a significant contribution to the field of image registration, providing unique strengths and capabilities. Their ongoing development and application consistently push the boundaries in multi-spectral image analysis, showcasing the potential for even more sophisticated and accurate image alignment techniques in the future. The goal of the current work is to establish a common benchmark and evaluate the performance of these techniques using different metrics.

2.2 Datasets

This section reviews the most commonly employed datasets in the non-medical cross-spectral image processing domain, intended for the evaluation of registration approaches. Having a common benchmark to be used for evaluating the different approaches is mandatory, and in the particular case of visible and thermal image spectrum there is still space for contributions. One of the first cross-spectral dataset has been presented almost 10 years ago in CVC-13 [14]; this dataset consists of three subset of visible-thermal image pairs, including scenarios with: roads, facades and smooth surfaces. These pairs of images have been acquired in different days and weather conditions. This dataset has been intended for evaluating multi-modal stereo approaches, hence a synthetic depth map is also provided for each pair of cross-spectral images. Visible spectrum images have a resolution of 640×480 pixels, while thermal images have a resolution of 534×426 pixels. Since this dataset is intended for generating 3D information, calibration information is provided together with a sequence of images with a calibration pattern to extract intrinsic and extrinsic parameters of the cross-spectral stereo system. The main limitation of this dataset is the reduced number of pairs of images. An extension of this dataset, reaching up to 100 pairs of images, is later on provided in CVC-15 ([15, 16]), but it is still reduced for nowadays learning-based approaches.

On the contrary to the limited number of pairs of images contained in the previous datasets, KAIST [17] introduces a dataset that contains about 95000 color-thermal image pairs with a resolution of 640×480 pixels, captured at 20 Hz from a vehicle. This dataset is specifically designed for pedestrian detection research, comprising color and thermal images captured in various urban environments. It includes day, night, and challenging weather conditions scenes, making it an ideal dataset for developing and evaluating multi-spectral pedestrian detection algorithms.

More recently, FLIR ADAS [18] presents a comprehensive resource for developing Advanced Driver Assistance Systems, featuring a total of 26442 annotated frames, including 9711 thermal and 9233 RGB images. These images are captured at a resolution of 640×512 pixels using a FLIR Tau2 camera, and the dataset also includes 5142 unaligned image pairs in the same resolution, making it an invaluable asset for training and validating object detection algorithms in various driving scenarios.

Another large cross-spectral dataset has been presented in M^3FD [19] (Multi-scenario Multi-modality Fusion Dataset). It is a specialized dataset designed for the fusion of infrared and visible spectrum images for object detection. It consists of 8400 images across various urban scenes, including 4200 image pairs for fusion and detection tasks. The dataset features images in both 8-bit grayscale (infrared) and 24-bit color (visible), primarily in 1024×768 pixels. It is meticulously labeled for objects like people, cars, buses, and more, although some labeling inaccuracies exist. M^3FD is invaluable for research in fusion-based object detection, particularly in urban environments. The main advantage is that it is a dataset that contains all well-registered image pairs.

Finally, another noteworthy dataset is introduced and presented in [20]. It utilizes three different sets of cameras, each capturing 1021 thermal images (LR 160×120, MR 320×240, and HR 640×480). The images depict outdoor scenarios with various objects (e.g., buildings, cars, people, vegetation) and are taken at different times of the day, including morning, afternoon, and night. In contrast to previous datasets, this dataset is specifically designed for super-resolution image evaluation. It serves as a benchmark for the PBVS-CVPR challenge from the 2020 to the 2023 edition ([21–24]).

Furthermore, the latest challenge edition [21] introduces a newly acquired cross-spectral dataset. This dataset comprises a total of 200 pairs of images captured using Basler and TAU2 cameras. The provided images have a resolution of 640×480 pixels, offering a valuable resource for enhancing and evaluating super-resolution techniques in multi-spectral imaging. Captured under daylight conditions, these images are characterized by clear edges and high quality, making them well-suited for training purposes in guided super-resolution tasks.

3 Proposed Approach

This section presents the methodology for evaluating visible and thermal image registration techniques. It employs advanced registration methods and evaluates their effectiveness using established metrics on a dedicated dataset specifically acquired for this study.

This work makes use of different image registration techniques used in state-of-the-art, such as Elastix, Imregister, LightGlue, and Nemar models, and is envisioned to enhance the precision and efficiency of image registration processes. Elastix, a versatile image registration platform, is proposed to align and compare diverse image datasets, especially in medical imaging contexts. Imregister, a MATLAB function, is employed to optimize image alignment and orientation, ensuring accurate spatial correspondence. The two aforementioned techniques are based on classical approaches and are used for the registration of multi-modal images and in this work, the results obtained have been executed and evaluated.

On the other hand, LightGlue is a deep neural network-based approach that matches sparse local features across image pairs, this technique has been implemented using mainly images of the visible spectrum. Despite this restriction, for the present work, it was decided to evaluate the results of the network using the visible-thermal image pair. Finally, the Nemar model is another learning-based registration technique used in the context of visible and NIR images. For this work, all proposed techniques are evaluated using an acquired dataset with 1000 pairs of visible-thermal images.

The current work focuses on evaluating registration approaches in the cross-spectral domain. To achieve this, a large cross-spectral image dataset is needed, featuring diverse images captured in daylight conditions for a fair assessment of techniques. Addressing the limitations of existing datasets (Sect. 2), this study introduces a new dataset. It is obtained using two cameras rigidly attached to minimize the distance between their optical axes, simplifying image registration. This dataset expands on the initial PBVS-CVPR2023 challenge dataset [21], offering a more extensive image collection for a comprehensive evaluation of registration methods.

4 Experimental Results

This section presents the experimental results obtained during the evaluation of the different registration techniques on the newly acquired cross-spectral dataset.

4.1 Evaluation

The selection of metrics for assessing the registration results of thermal and visible spectrum images is crucial for determining the quality and accuracy of the alignment process between these modalities. This involves both quantitative and qualitative evaluations.

The **Quantitative Evaluation** is shown in Table 1. Each metric has been selected based on the work of different authors (e.g., [9, 11, 25]), and the metrics used in the present work are the Mutual Information (MI), the Normalized Mutual Information (NMI), the Normalized Root Mean Square Error (NRMSE), the Normalized

Table 1 Quantitative comparisons of registration accuracy (mean and standard deviation, red: best, blue: second best, green: third best)

Metric	Technique					
	Unaligned	Elastix	Imregister	LightGlue+SP	LightGlue+D	Nemar
MI ↑	0.961±0.319	1.135±0.329	1.141±0.334	1.116±0.348	0.668±0.341	0.925±0.331
NMI ↑	1.079±0.030	1.096±0.032	1.097±0.032	1.094±0.080	1.057±0.142	1.076±0.031
NRMSE ↓	0.791±0.184	0.797±0.185	0.811±0.200	0.805±0.243	0.917±0.794	0.779±0.197
NCC on edges ↑	0.236±0.067	0.310±0.087	0.293±0.090	0.284±0.092	0.140±0.080	0.193±0.064
PSNR on edges ↑	23.561±1.643	24.840±1.879	24.514±1.805	24.485±2.437	23.590±3.553	24.967±2.044
SSIM on edges ↑	0.450±0.100	0.517±0.099	0.521±0.103	0.522±0.110	0.410±0.117	0.455±0.098

Cross Correlation Coefficient (NCC), the Peak Signal-to-Noise Ratio (PSNR), and the Structural Similarity Index Metric (SSIM) [26]. The last three metrics (NCC, PSNR, and SSIM) are computed on edges detected by the Sobel-detector from both the visible and thermal registered images. Also, all six metrics are used to quantify the similarity and coherence between the corresponding pixels in both images. These metrics allow the identification of the effectiveness of the used approaches and provide an objective evaluation of the quality of the applied transformations. Furthermore, the appropriate selection of evaluation metrics contributes to the improvement of registration algorithms and facilitates decision-making in applications that require precise fusion of visible-thermal information, such as in security, surveillance, and image analysis.

The **Qualitative Evaluation** of registered thermal and visible images is useful to validate the accuracy through a visual coherence analysis of the applied transformations. One effective method involves using a red-cyan visualization of the registered images—evaluation strategy used by [27]. By wearing the red-cyan method, observers can perceive depth in the fused image, allowing for a qualitative assessment of the alignment. This approach enables the detection of any disparities, misalignments, or distortions that may have occurred during the registration process. Experts carefully examine anatomical and structural details in the anaglyph image to ensure proper correspondence between features in both modalities. The red-cyan anaglyph technique provides valuable qualitative insights into visual fidelity and registration accuracy, complementing quantitative evaluations from objective metrics.

4.2 Dataset

In this study, a new dataset has been gathered, consisting of pairs of images capturing the same scene in both visible and thermal spectrums. These high-quality images, taken during the day, exhibit clear outlines, rendering them ideal for training single and guided super-resolution techniques. Figure 1 provides visual examples of this dataset (referred as Cross-spectral Image Dataset for Image Super-resolution - CIDIS).

Fig. 1 Illustrations of the acquired cross-spectral dataset, thermal and visible images of the same scenario

This dataset consists of 1000 pairs of images, captured using Balser and TAU2 visible and thermal spectrum cameras respectively. Worth mentioning that these two types of images come in different resolutions and spectrums, visible images with a native resolution of 1280×1024 (using a 13mm lens) and thermal images with a resolution of 640×480 (using an 8mm lens). The dataset is organized into three sets: training, validation, and testing. There are 700 image pairs for training, 200 for validation, and 100 for testing.

A dataset with registered pairs of visible and thermal images is required for evaluating different cross-spectral image processing approaches. Approaches that go from guided super-resolution, or guided filtering till image fusion or cross-spectral object detection.

4.3 Analysis of Registration Results

Table 1 displays the average results for each metric and technique evaluated on the visible-thermal image pair. The second column shows values for unregistered pairs. In each metric row, the top three techniques are marked in red, blue, and green. The results indicate that the Nemar and LightGlue+D models lack consistency across all metrics. In contrast, the Elastix, Imregister, and LightGlue+SP techniques consistently yield stable values. While no technique dominates across all metrics, these three techniques prove applicable in the context of multi-modal registration. It's worth noting that LightGlue+SP was evaluated without undergoing any fine-tuning process, suggesting potential for future work to enhance results.

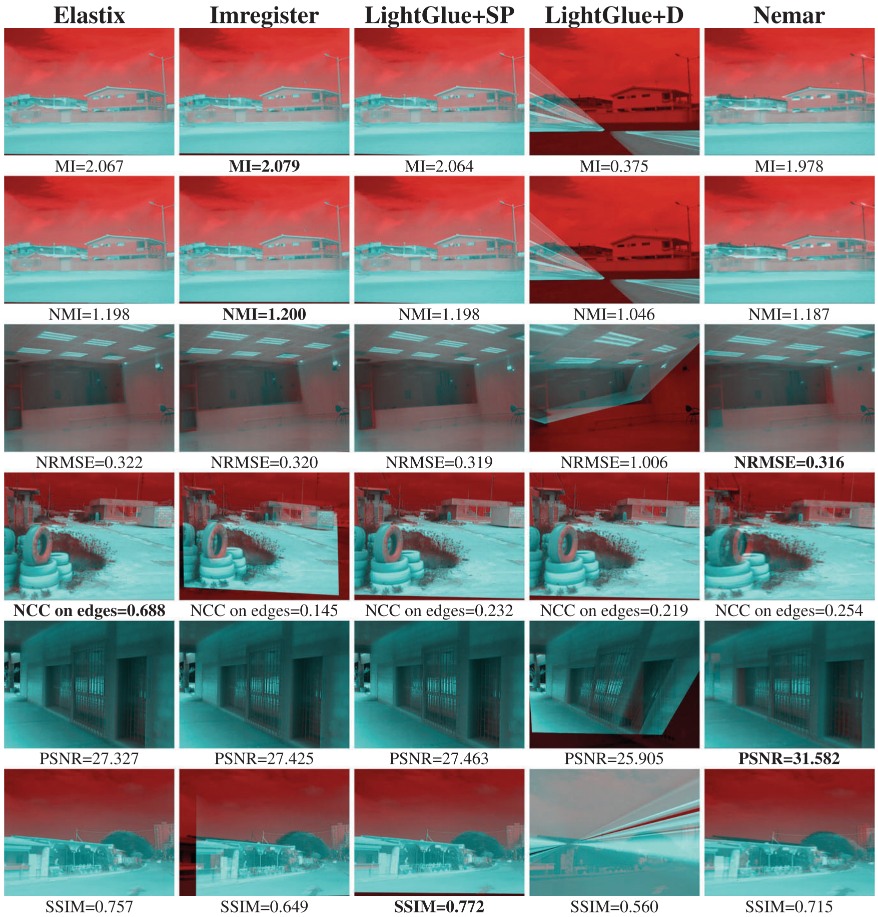

Fig. 2 Comparative results between visible (RGB) and thermal images. These example images are part of the testing set

Figure 2 illustrates results based on each metric and technique. The reference image for each metric is selected from the technique securing the first place in Table 1. For instance, row 1 displays the image with the highest mutual information value for the Imregister technique (the best for MI metric), used for evaluating results in all other techniques. This process is repeated for each metric.

On the other hand, Fig. 3 illustrates the Box Plots of each metric for each image registration technique. These graphs depict the distribution of the data, mean value, and outliers for each metric and technique. From the obtained values, it can be observed that the Elastix, Imregister, and LightGlue+SP metrics show consistent and similar results, while the LightGlue+D and Nemar techniques show less stable values and lower values compared to the average of the other metrics.

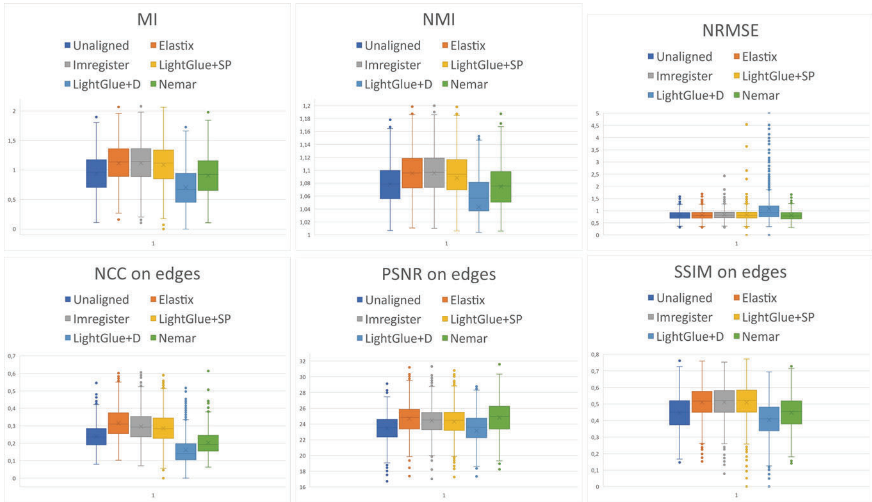

Fig. 3 Box Plots results from each of the evaluated image registration technique according for the different metrics

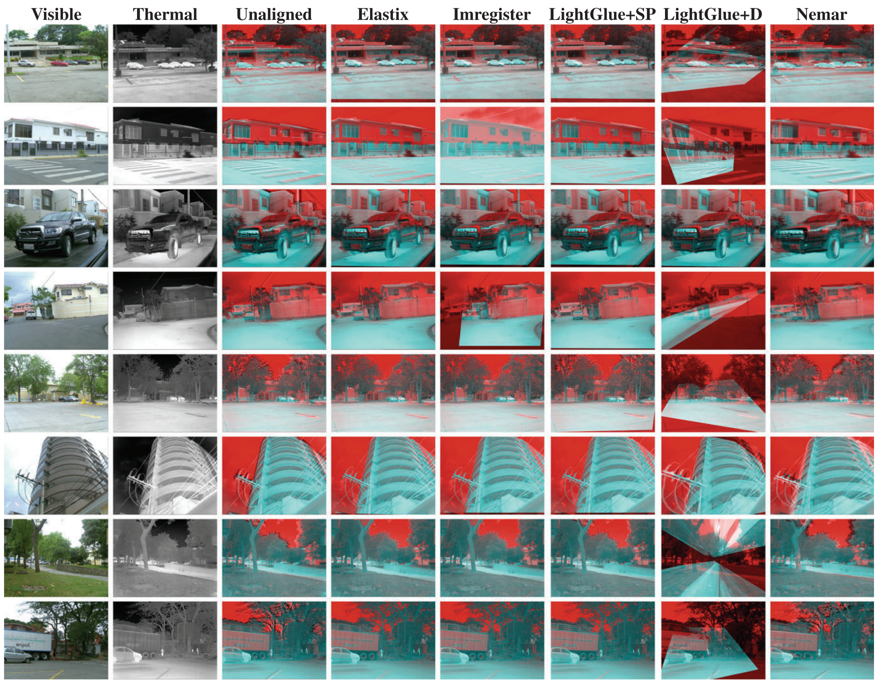

Fig. 4 Comparative results between visible (RGB) and thermal images. These example images are part of the testing set

Finally, Fig. 4 shows examples of results of each of the techniques that are part of this study (from *col. 4* to *col. 8*); these results are presented in red-cyan format for a qualitative evaluation. Also, in *col. 1* and *col. 2* the visible-thermal image pair is shown; in *col. 3* the overlap between the visible and thermal unaligned images is provided for a qualitative comparison.

5 Conclusions

In conclusion, the current study indicates that no single technique outperforms all others across every metric. However, Elastix shows consistent top-three performance, making it a robust choice for a variety of situations. For applications where edge alignment is crucial, LightGlue+D and LightGlue+SP perform exceptionally well. Nemar's low NRMSE indicates its potential for high accuracy in certain conditions. While the "best" technique depends on specific image registration task requirements, Elastix, Imregister, and LightGlue+SP emerge as solid options.

Acknowledgements This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-22-1-0261; and partially supported by the Grant PID2021-128945NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe"; and by the ESPOL project CIDIS-12-2022. The third author acknowledge the support of the Generalitat de Catalunya CERCA Program to CVC's general activities, and the Departament de Recerca i Universitats from Generalitat de Catalunya with reference 2021SGR01499.

References

1. Bouallal D, Bougrine A, Douzi H, Harba R, Canals R, Vilcahuaman L, Arbanil H (2020) Segmentation of plantar foot thermal images: application to diabetic foot diagnosis. In: 2020 International conference on systems, signals and image processing (IWSSIP). pp 116–121. IEEE
2. ElMasry G, ElGamal R, Mandour N, Gou P, Al-Rejaie S, Belin E, Rousseau D (2020) Emerging thermal imaging techniques for seed quality evaluation: principles and applications. Food Res Int 131:109025
3. Kumar P, Gaurav A, Rajnish RK, Sharma S, Kumar V, Aggarwal S, Patel S (2022) Applications of thermal imaging with infrared thermography in orthopaedics. J Clin Orthop Trauma 24:101722
4. Wilson A, Gupta K, Koduru BH, Kumar A, Jha A, Cenkeramaddi LR (2023) Recent advances in thermal imaging and its applications using machine learning: a review. IEEE Sens J
5. Velesaca HO, Bastidas G, Rouhani M, Sappa AD (2024) Multimodal image registration techniques: a comprehensive survey. Multimedia Tool Appl pp 1–29
6. Lindenberger P, Sarlin PE, Pollefeys M (2023) Lightglue: local feature matching at light speed. arXiv preprint arXiv:2306.13643
7. Marstal K, Berendsen F, Staring M, Klein S (2016) Simpleelastix: a user-friendly, multi-lingual library for medical image registration. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp 134–142

8. Muthukumaran D, Sivakumar M (2017) Medical image registration: a matlab based approach. Int J Sci Res Comput Sci Eng Inf Technol 2(1):29–34
9. Arar M, Ginger Y, Danon D, Bermano AH, Cohen-Or D (2020) Unsupervised multi-modal image registration via geometry preserving image-to-image translation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp 13410–13419
10. Shrestha P, Lee C, Fahy K, Balakrishnan M, Ge N, Bazylak A (2020) Formation of liquid water pathways in pem fuel cells: a 3-d pore-scale perspective. J Electrochem Soc 167(5):054516
11. Xu H, Ma J, Yuan J, Le Z, Liu W (2022) Rfnet: unsupervised network for mutually reinforcing multi-modal image registration and fusion. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp 19679–19688
12. DeTone D, Malisiewicz T, Rabinovich A (2018) Superpoint: self-supervised interest point detection and description. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp 224–236
13. Tyszkiewicz M, Fua P, Trulls E (2020) Disk: learning local features with policy gradient. Adv Neural Inf Process Syst 33:14254–14265
14. Campo FB, Ruiz FL, Sappa AD (2012) Multimodal stereo vision system: 3d data extraction and algorithm evaluation. IEEE J Sel Top Sign Proc 6(5):437–446
15. Aguilera C, Barrera F, Lumbreras F, Sappa AD, Toledo R (2012) Multispectral image feature points. Sensors 12(9):12661–12672
16. Barrera F, Lumbreras F, Sappa AD (2013) Multispectral piecewise planar stereo using manhattan-world assumption. Pattern Recogn Lett 34(1):52–61
17. Hwang S, Park J, Kim N, Choi Y, So Kweon I (2015) Multispectral pedestrian detection: benchmark dataset and baseline. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 1037–1045
18. Flir thermal dataset. https://www.flir.com/oem/adas/adas-dataset-form/, accessed: Oct. 12, 2023
19. Liu J, Fan X, Huang Z, Wu G, Liu R, Zhong W, Luo Z (2022) Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp 5802–5811
20. Rivadeneira RE, Sappa AD, Vintimilla BX (2020) Thermal image super-resolution: a novel architecture and dataset. In: VISIGRAPP (4: VISAPP). pp 111–119
21. Rivadeneira RE, Sappa AD, Vintimilla BX, Bin D, Ruodi L, Shengye L, Zhong Z, Liu X, Jiang J, Wang C (2023) Thermal image super-resolution challenge results-pbvs 2023. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops
22. Rivadeneira RE, Sappa AD, Vintimilla BX, Kim J, Kim D, Li Z, Jian Y, Yan B, Cao L, Qi F et al (2022) Thermal image super-resolution challenge results-pbvs 2022. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp 418–426
23. Rivadeneira RE, Sappa AD, Vintimilla BX, Nathan S, Kansal P, Mehri A, Ardakani PB, Dalal A, Akula A, Sharma D et al (2021) Thermal image super-resolution challenge-pbvs 2021. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp 4359–4367
24. Rivadeneira R, Sappa A, Vintimilla B, Guo L, Hou J, Mehri A, Ardakani P, Patel H, Chudasama V, Prajapati K et al (2020) Thermal image superresolution challenge-pbvs 2020. in 2020 ieee. In: CVF conference on computer vision and pattern recognition workshops (CVPRW). pp 432–439
25. Ordun C, Raff E, Purushotham S (2023) Vista-morph: unsupervised image registration of visible-thermal facial pairs. arXiv preprint arXiv:2306.06505
26. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP et al (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
27. Mascarich F, Alexis K (2020) Visual-thermal camera dataset release and multi-modal alignment without calibration information. arXiv preprint arXiv:2012.14833