Optimizing the generation of orthomosaic through the fusion of UAV-based thermal and visible images. Case study: PV inspections

Fatima-zahra Bamessaoud[1*], *Nouhaila* Bahtit[1], and *Imane* Sebari[1]

[1]School of Geomatic Sciences and Surveying Engineering, Agronomic and Veterinary Institute Hassan II, BP 6202 Madinat Al Irfane, Rabat-Instituts, Rabat 10112, Morocco.

Abstract. Thermal and visible images acquired from Unmanned Aerial Vehicles (UAV) have proven to be indispensable for detecting faults in photovoltaic panels. Thermal orthomosaics enable rapid identification of defective modules, while visible imagery assists in interpreting the nature of these defects. However, accurate orthomosaic generation remains challenging due to the low spatial resolution of thermal images and the repetitive patterns of photovoltaic installations. This paper aims to optimize orthomosaic generation through the fusion of UAV thermal and visible images. An extensive literature review was conducted, followed by an analytical comparison of nine image fusion methods: ADF, CBF, FPDE, GLPF, GTF, LatLRR, MGFF, MSVD, and VGG. The methods were evaluated using qualitative visual assessment and quantitative metrics. Results indicate that MGFF outperforms the other approaches, achieving an entropy value of 1.0356 and significantly enhancing image detail, which improves tie-point detection and image alignment. Accordingly, a dedicated processing pipeline combining MGFF with a customized registration workflow was developed and validated on a challenging dataset of 500 thermal and visible image pairs, demonstrating substantial improvements in orthomosaic quality.

1 Introduction

The rapid advancement of photovoltaic (PV) technology, coupled with the increasing number and scale of PV panels, necessitates more efficient and intelligent health monitoring strategies to ensure reliable operation and high energy availability [1]. Due to the outdoor operating conditions and potential damages during manufacturing, transportation, or installation, various PV faults can occur, leading to degradation, power loss, or even fire hazards [2]. Consequently, it is crucial to implement effective health monitoring systems for PV power plants to ensure the reliability and longevity of energy production [3].
Classical techniques of visual inspections were usually performed by qualified inspectors for a long time [4]. Traditional fault detection and diagnostic techniques can be categorized into two types, namely, visual inspections and automatic fault analyses [1]. However, these

* Corresponding author: bamessaoudfatima-zahra@iav.ac.ma

classical approaches had some disadvantages, such as high manpower required, trained professionals, high rate of mistakes, risks involved in performing the job by humans, delays, and others [5]. In order to overcome these problems, various new approaches have been suggested to monitor photovoltaic systems. Consequently, many researchers and investors suggested using UAVs as a highly efficient approach owing to the following merits of these devices: minimized risks and hazards for inspectors, high-quality data gathering, rapid launching, wide application, rapid access to information, ability to perform visual inspections on the operating machines with minimum interruptions, cost-efficiency, and time-efficiency. PV inspections using UAV-based aerial surveys make use of two kinds of cameras, namely, Red-Green-Blue (RGB) cameras and thermal infrared cameras. RGB cameras are employed to obtain high-resolution images of PV modules, which are important for identifying visual defects like cracks, discoloration, dirtiness, de-lamination, and corrosion. RGB cameras require special consideration in terms of factors like resolution and flight altitude for defect identification. The second type of camera is a thermal infrared camera, which is used for obtaining temperature distribution on the surface of PV modules to help identify any hot spots that may occur due to problems like open circuits, short circuits, and shading effects [6]. When thermal cameras are used together with RGB cameras as part of the drone sensor system, proper identification becomes possible.

UAVs have proved to be a valuable asset in many industrial applications including UAV mapping, aerial thermal surveys, and photovoltaic inspection among others. The use of near-thermal imaging technology made it possible to identify issues which could not be identified before [7]. In most cases, these applications are based on the use of thermal information in creating orthophotos using structure from motion algorithms [8]. It is worth noting that the processing of thermal information using the structure from motion algorithm does not always produce accurate results.

These problems with thermal datasets are due to the lack of information provided by thermal images in comparison to RGB images, making it difficult to detect common features necessary for bundle adjustment. Actually, unlike RGB or multispectral images, almost all thermal cameras used by drones provide only one band (thermal value of each pixel covering the whole 8 to 12 micrometers range). Furthermore, the spatial resolution of thermal images is extremely low [8]. While even low-end RGB cameras have spatial resolutions of 12 MP or more, the spatial resolution of thermal cameras in UAVs ranges from 0.077 (320 x 240 pixels) to 0.33 MP (640 x 512 pixels) [8].

Image fusion has emerged as an essential topic in the field of image processing in recent times [10]. The aim of image fusion is the combination of corresponding information of each pixel or a few pixels of several image sources into one image that has higher information compared to individual images. Since image fusion came into existence in computer vision and image processing, it found numerous applications across various sectors. Among these sectors include medical imaging, night vision imaging, detection and tracking sector, and automated cars. Yet, the area of fusing visual and thermal images from drones still remains promising but underexplored.

Given these difficulties, this research provides a detailed comparative analysis of nine fusion techniques (ADF, CBF, FPDE, GLPF, GTF, LatLRR, MGFF, MSVD, VGG) on visible and thermal images captured by drones to improve the orthomosaic creation process for the inspection of photovoltaic panels. It is already known that using the fused image rather than the original visible or thermal images in the process of solving different computer vision tasks, including automated object detection and classification in terrestrial images [10], leads to better results. The goal of this paper is to develop this approach for UAV images using image fusion.

2 Related work

B. S. Krishnan et al. (2023) present a methodology that closely aligns with our research on optimizing the processing of orthophoto generation through the fusion of UAV-based thermal and visible images, specifically for PV inspections. In their study, [12] demonstrated how the integration of visible and thermal imagery can significantly enhance the automated detection and classification of animals during drone surveys. They evaluated eight image fusion methods using thermal and visible images in conjunction with two supervised deep learning models to detect and classify white-tailed deer, domestic cows, and domestic horses. The results indicated that image fusion significantly improved the detection and classification of white-tailed deer, which are typically hard to spot against their backgrounds, by 15 to 85%. This improvement is attributed to the additional information provided by thermal images. In conclusion, amongst the eight used image fusion methods, the Sparse fusion method, in particular, stands out for its superior performance in both image quality and practical application, making it the top choice among the evaluated methods.

3 Methodology

In our current research, we have meticulously established a workflow aiming to enhance orthomosaic generation through the fusion of thermal and visible UAV-based images. The process begins with the registration of these images to produce paired thermal and visible images with the same dimensions and depicting the same scene. Subsequently, nine fusion methods are applied to generate fused images that incorporate information from both the visible and thermal images. To evaluate the quality of the fusion, seven evaluation metrics are applied to each fused image. Finally, the methods demonstrating the best overall quality are selected to generate orthomosaics. The main objective is to apply this process to a dataset that typically produces low-quality orthomosaics and, through our methodology, generate higher-quality ones. This improved orthomosaic can then be utilized in drone surveys, particularly for photovoltaic panel inspections.

Figure 1 illustrates the proposed methodology encompassing data preparation and preprocessing, image fusion methods, evaluation metrics, and validation process.

Table 1. Key technical specifications of the RGB and IR cameras used for image acquisition of PV installations.

Sensors	**Camera model**	**UAV model**	**Focal length (mm)**	**Pixel pitch (µm)**	**Resolution (pixel)**	**Angular field of view (h*v)**
RGB	DJI Zenmuse XT2	DJI Matrice M200	8	1.85	4000*300	57° * 42°
IR			19	17	640*512	32° * 26°

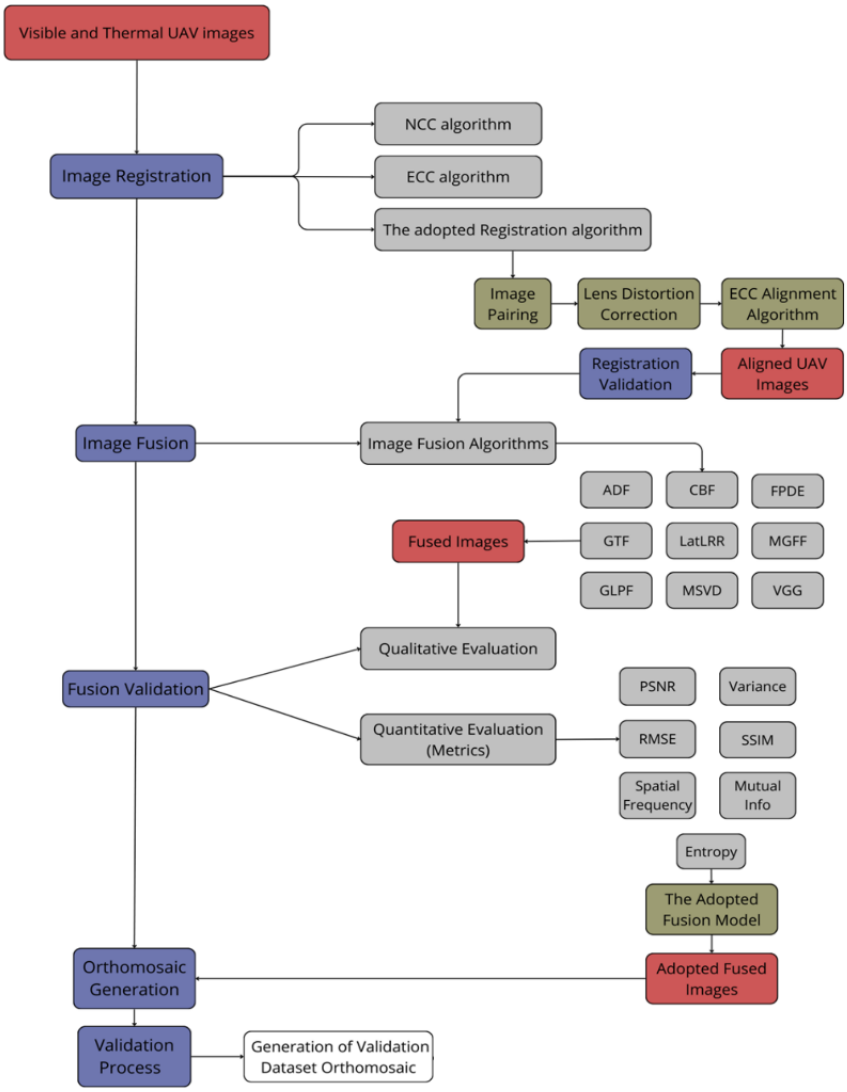


Fig. 1. A flowchart visualizing the Comprehensive Methodology employed in the current study

4 Methods

4.1 Data

The research data comprises thermal and visible drone images of photovoltaic installations, captured using rotary-wing UAVs, specifically the (Da-Jiang Innovations) DJI Matrice M200. RGB imagery were captured with the DJI Zenmuse XT2, which facilitated the simultaneous acquisition of thermal imagery. All flights employed automatic acquisition mode. No Ground Control Points (GCPs) were measured in the field; instead, the UAVs' code-based Global Positionning System (GPS) and Global Navigation Satelite System (GNSS) receivers were used to georeference the datasets during processing [11].

We dispose of two datasets with the same characteristics as describe in table 1; the first one, having a well generated orthomosaic, is used in the experimental process as a mean of exploring the different paths and developing a coherent process. The second dataset, named validation dataset, present problems in orthomosaic generation. It is used for the validation of the developed process.

4.2 Image registration

Before applying image fusion algorithms, it is crucial to align the input images accurately. Consequently, we investigated various image registration algorithms and approaches, ultimately developing a customized approach tailored to our dataset represented in Figure 2.

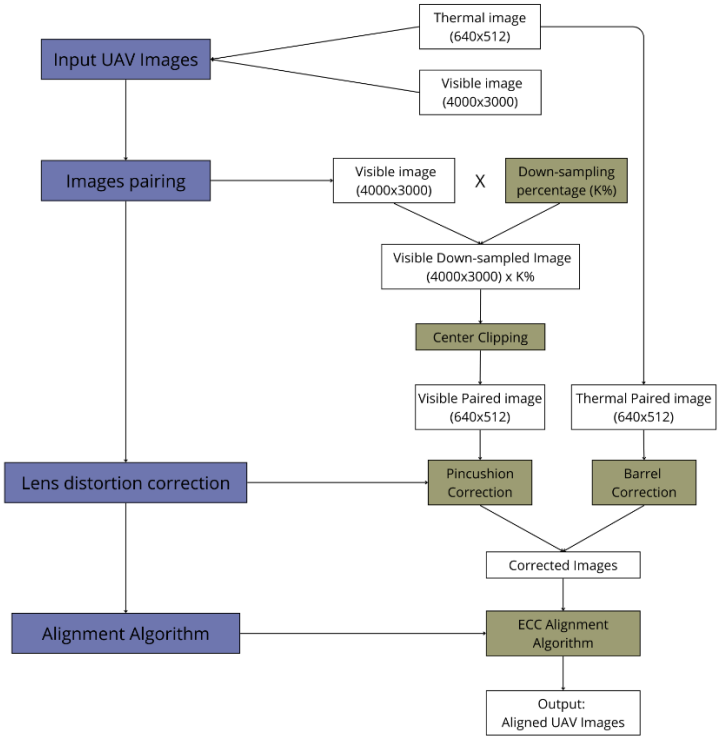


Fig. 2. A flowchart of the Registration Workflow

4.2.1 Image pairing

Due to the differing specifications of their sensors, thermal and visible images have varying extents and dimensions. However, the fusion algorithms applied to UAV data require the images to be of the same dimension. Consequently, pre-processing is necessary to produce paired thermal and visible images with the same dimensions and depicting the same scene.

The dimensions of thermal and visible images are (640x512) and (4000x3000), respectively. Thus, visible images need to be down-sampled. Additionally, the focal lengths of thermal

and visible images are 19mm and 8mm, respectively. They relate to the object distance (d) and the object size (I) as follows:

$$\frac{f_1}{I_1} = \frac{f_2}{I_2} = \frac{d}{I} \quad (1)$$

$$I_1 = P_1 \times S_1 \quad (2)$$

$$I_2 = P_2 \times S_2 \quad (3)$$

where P1 and P2 represent the pixel numbers, and S1 and S2 denote the sizes of the detector of the thermal and visible light sensors, respectively. Given S1 and S2, a specific percentage of down-sampling (k%) can be calculated between P1 and P2 to ensure that the visible image, when down-sampled, appears to have the same focal length as the thermal image. The resulting down-sampled visible image must be clipped to match the scene captured by the thermal one. Consequently, a (640x512) frame is created to clip the down-sampled image appropriately as illustrated in the figure 3 bellow.

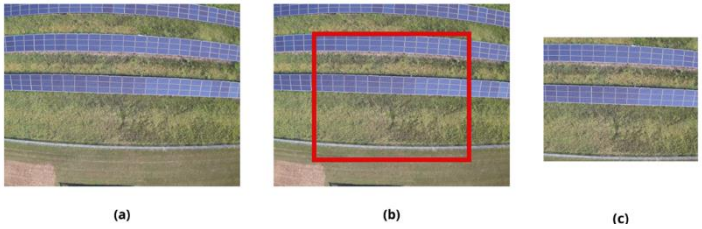


Fig. 3. Illustration of image pairing processus: (a) Down-sampled visible image. Spatial resolution: 4000 x 3000 x k%. (b) Template image. Spatial resolution: 640x512. (c) Visible pair. Spatial resolution: 640x512

At the end of this step, the visible pair image is obtained, presenting the same scene as the thermal image and having the same dimensions.

4.2.2 Lens distortion correction

The panels in the captured drone images appear to be curved, indicating the presence of radial distortion, as shown in figure 4. This type of distortion can manifest as barrel distortion, where straight lines bulge outward, or pincushion distortion, where lines curve inward, and it significantly impacts the accuracy and quality of the image registration process. To address this issue, a lens distortion correction algorithm is applied, featuring two adjustable parameters: strength, which controls the degree of distortion correction, and zoom, which manages the scaling of the image after correction. Properly tuning these parameters is crucial to ensuring that the straight lines and geometric features in the scene are accurately represented, thereby improving the overall effectiveness of the image registration process.

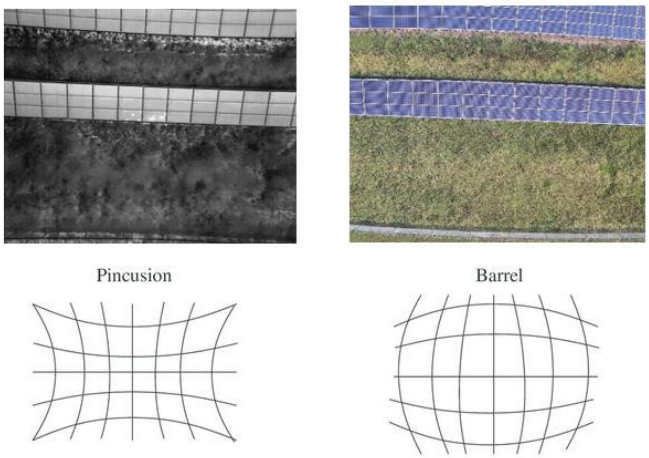


Fig. 4. Illustration of the radial distortions. Pincushion distortion remarked in the thermal image, Barrel distortion in the visible image

4.2.2.1 ECC registration algorithm

After pairing the thermal and visible images and correcting their radial distortions, the images must be aligned using an area-based registration algorithm. The adopted one, after various experiments, is the the Enhanced Cross Correlation (ECC) algorithm with its translation motion model.

4.3 Image fusion

After applying the necessary pre-processing treatments to the initial drone images, the data is now ready for the fusion process. Table 2 represents briefly the nine fusion methods applied to the dataset.

Table 2. Fusion Methods description

Method	Description
ADF	**Anisotropic Diffusion Fusion** uses anisotropic diffusion to reduce image noise without removing significant parts of the image content, such as edges, effectively preserving important features.
CBF	**Curvelet-Based Filter** employs the curvelet transform to capture edges and curves more effectively than traditional wavelet-based methods, enhancing the representation of image details.
FPDE	**Fractional Partial Differential Equation Fusion** applies fractional calculus in the partial differential equation framework, which can enhance edges and fine details in the fused image.
GLPF	**Gaussian-Laplacian Pyramid Fusion** constructs Gaussian and Laplacian pyramids of the images and fuses them at each level to combine the best features from both images, enhancing detail representation.
GTF	**Gradient Transfer Fusion** focuses on transferring gradient information from source images to the fused image, effectively preserving edge details.

LatLRR	**Latent Low-Rank Representation** uses low-rank representation to capture the essential structure of the images while ignoring noise and redundant information, achieving a balanced fusion.
MGFF	**Multi-Scale Guided Filter Fusion** utilizes guided filtering at multiple scales to enhance and fuse images, preserving edges while smoothing out textures for detailed representation.
MSVD	**Multi-Scale Singular Value Decomposition** applies singular value decomposition at multiple scales to effectively separate and combine image details in the fusion process
VGG	**VGG Neural Network-Based Fusion** employs a pre-trained VGG neural network for feature extraction and fusion, leveraging deep learning to capture complex patterns and textures in the images.

The main objective of fusion is to combine information from both UAV-based images to integrate the texture details of visible images with the salient target information in thermal images. Our focus is on photovoltaic panels, where textures are clear in visible images, while thermal images reveal critical thermal information, such as hotspots.

The fused image should contain both sets of information. It is immediately noticeable in figure 5 that all methods resulted in more textured PV panels. However, the hotspots (white spots) are not as clear as in the thermal images.

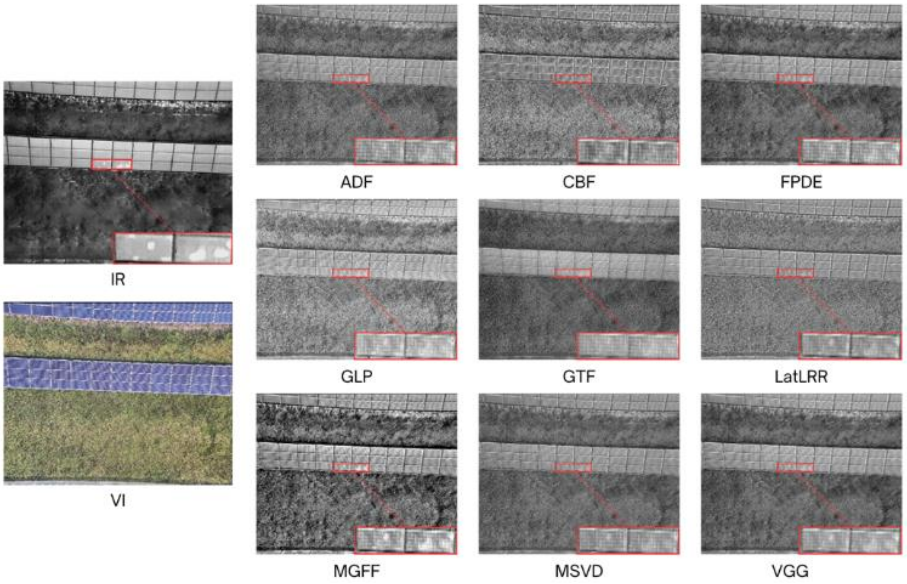


Fig. 5. Fused images

After fusing the images, we evaluate different fusion methods using quantitative evaluation metrics (figure 6). We convert all the images into grayscale with one value channel to start the computation, except for Mutual Information (MI) and Structural Similarity Index measure (SSIM), the images are normalized before computation. The evaluation was processed on a sample of fused images. For better analysis we take the mean of each evaluation metric per fusion method and we compare the results:

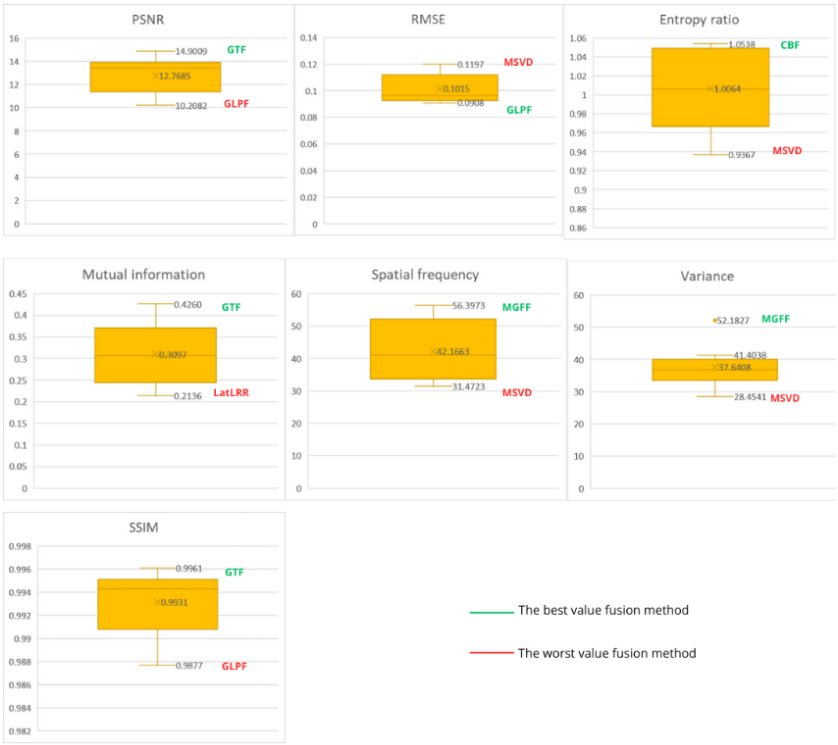


Fig. 6. Box plot representing the variation of metrics values per fusion method applied to the fused images

The results shown in figure 6 are related to the entire fused image to evaluate overall performance. To achieve more precised and relevant results, a mask was subsequently applied to target only the photovoltaic panels (figure 7). Focusing on the panel regions is particularly valuable because these areas contain the features of interest for fault detection, while background regions contribute little to the analysis and may introduce noise or artifacts. The outcomes of this procedure are presented in figure 8:

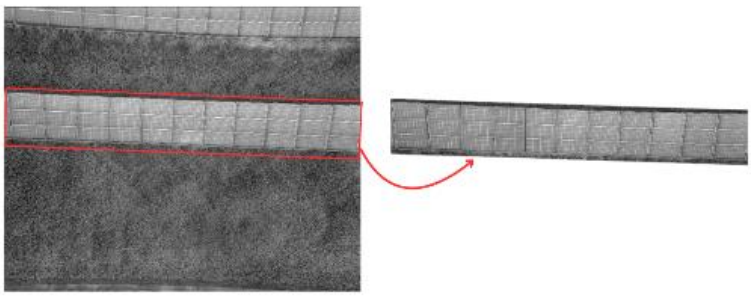


Fig. 7. Target region (PV panels)

Fig.8. Box plot representing the variation of metrics values per fusion method applied to the target region (PV panel)

- **Quality:** The PSNR (Peak Signal-to-Noise Ratio) is consistently higher and RMSE (Root Mean Squared Error) is consistently lower in the masked results compared to the full image results. This indicates that the regions containing photovoltaic panels are reconstructed with better quality compared to the entire image.
- **Detail and Contrast:** The spatial frequency (SF) and variance (Var) are generally higher in the full image results. This suggests that the overall image has more detail and contrast compared to the masked regions.
- **Information Content:** The entropy ratio (EN-ratio) and mutual information (MI) are generally higher in the full image results. This implies that the entire image contains more information content and shared information compared to the masked regions.
- **Structural Similarity:** SSIM values are slightly higher in the masked regions, indicating better structural similarity in these regions compared to the overall image.

Based on the quantitative and qualitative evaluation:

- **Best Overall Quality**: **GTF** [13] (Guided Total Variation Fusion) stands out as the best method for maintaining high image quality both in the masked regions and full images. Despite having lower spatial frequency and variance, its strong performance in image quality metrics makes it the best choice.

- **Detailed and Contrast-rich Images**: If the goal is to maximize detail and contrast, **MGFF** [14] (Multi-Scale Guided Filter Fusion) performs best in spatial frequency and variance metrics, but at the cost of some quality.

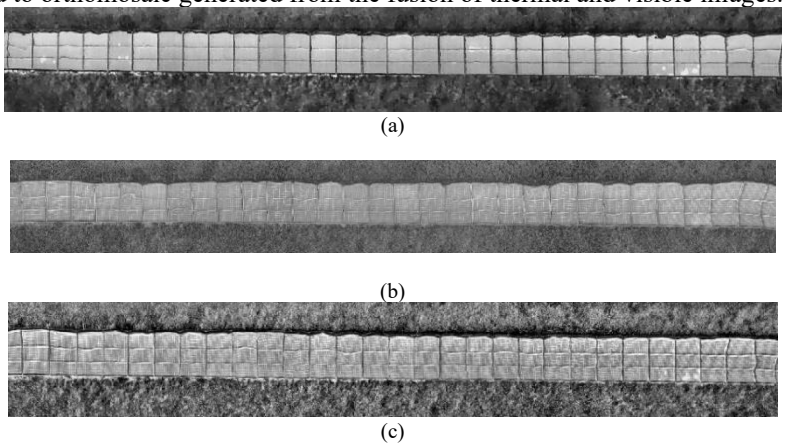
For better understanding, the table outlines the key resemblances between GTF and MGFF:

Table 3. Key resemblances between GTF and MGFF methods

Aspect	Guided total variation fusion	Multi-Scale Guided Filter Fusion
Feature preservation	Uses total variation minimization to preserve edges and textures	Uses multi-scale guided filtering to preserve important details
Decomposition approach	Decomposes images into base and difference images	Decomposes images into base and detail layers across multiple scales
Weighting mechanism	Iterative reweighting to refine difference image	Weight maps from saliency maps to combine detail layers
Combining components	Combines refined difference image with visible image	Combines final detail layers with base layers

4.4 Orthomosaic generation

The generation of the orthomosaic is intended to assess the effectiveness of our methodology. Typically, the experimental dataset produces a high-quality orthomosaic, which is then compared to orthomosaic generated from the fusion of thermal and visible images.



(a)

(b)

(c)

Fig.9. Generated orthomosaics: (a) from thermal images (b) from GTF (c) from MGFF

As shown in figure 9, the GTF and MGFF-generated orthomosaics (b,c) display more details, particularly on the photovoltaic panels. However, the hotspots (white spots) are less prominent compared to those in the orthomosaic generated solely from thermal images (a).

The experimental dataset was used to identify the optimal fusion method and develop the adopted registration/fusion workflow. The established process developed using the first dataset is now applied to the validation dataset, which has been identified as problematic for successful orthomosaic generation. The goal is to evaluate the effectiveness of the established process in addressing the issues present in the validation dataset.

5 Results

Both GTF and MGFF methods are applied on the validation dataset as they both showed better results in the experimentation phase. By visually comparing the resulting fused images in figure 10, the MGFF image shows higher contrast, whereas the GTF image appears smoother. The hotspot is most prominently recognized in the MGFF image, but it is barely recognizable in the GTF image.

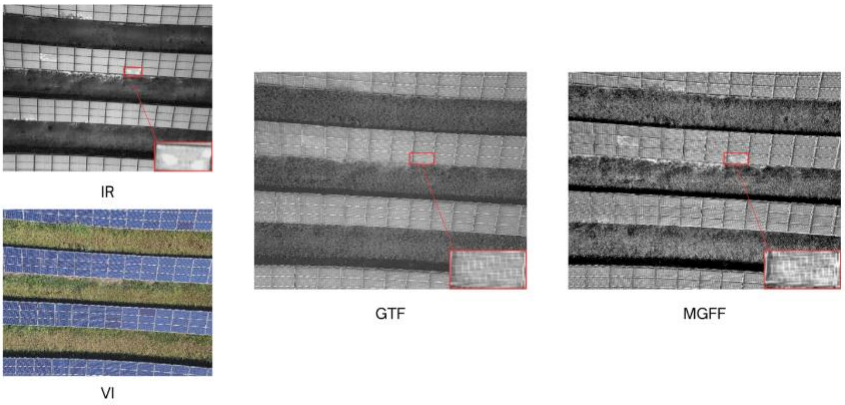


Fig.10. Fusion results

To evaluate the quality of fusion methods, performance metrics were calculated for a sample of images. For clearer analysis, we computed the mean of each metric across all images for each fusion method. The results are summarized in table 4 below.

Table 4. Comparison between GTF and MGFF methods based on evaluation metrics

Methods	PSNR	RMSE	EN-ratio	MI	SF	Var	SSIM
GTF	15.5665	0.0744	0.9830	1.0842	47.1816	42.0560	0.9971
MGFF	14.4825	0.0770	1.0356	1.1631	71.0389	59.5398	0.9971

GTF is better in terms of Peak Signal-to-Noise Ratio (PSNR) and Root Mean Square Error (RMSE), indicating it produces a fused image with less overall error and better signal quality. MGFF shows higher values in terms of Entropy (EN), Mutual Information (MI), Spatial Frequency (SF), and Variance (Var), indicating it preserves more information, texture, and contrast. This method can enhance the overall quality of orthomosaics by providing more comprehensive and detailed images, aiding in the detection of various defects, especially those that are subtle or only visible in the thermal spectrum mainly hotspots. Both methods

have identical Structural Similarity Index Measure (SSIM) values, indicating they equally preserve structural similarity.
For PV inspections, where the goal is to generate highly detailed and accurate orthomosaics, MGFF might be more suitable due to its superior ability to retain detailed information and textures, despite the slightly higher noise indicated by its lower PSNR. However, if smooth appearance and minimal noise are prioritized, GTF might be recommended.
After comparing the two methods, MGFF is adopted as the most effective image fusion technique for PV inspections. We will now generate the orthomosaic of the validation dataset using MGFF.
The initial generated orthomosaic of thermal images represented in figure 11 presents gaps due to the non-detection of tie points in certain images. Additionally, the aligned images exhibit distortion, causing lines 3 and 7 of the solar panels (counting from the upper right) to interfere with their following lines respectively. This issue is related to the mediocre resolution of the thermal images and the repetitive pattern created by the solar panels

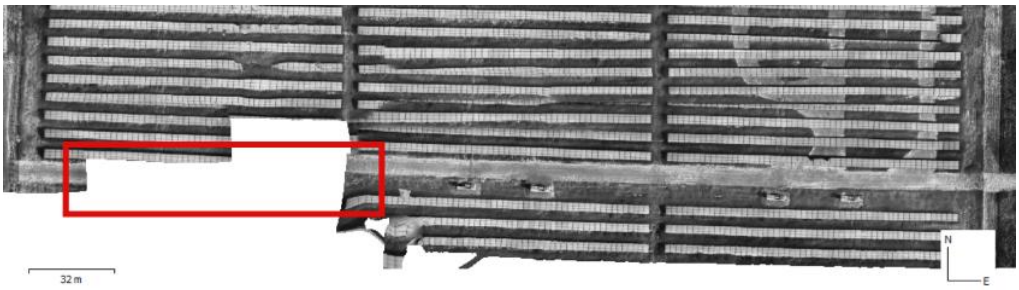


Fig.11. Initial Orthomosaic of the Validation Dataset. The red box represents the images that did not successfully complete the orthomosaic generation process.

The red frame in figure 11 represents the images that did not successfully complete the ortho generation process (27 images). For additional assurance, we process these images separately. The resulted orthomosaic is presented in figure 12.

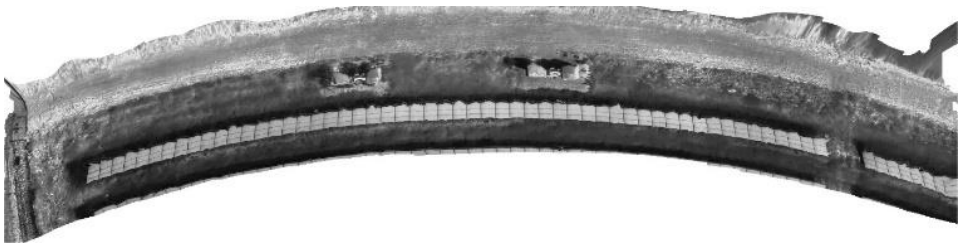


Fig.12. Thermal orthomosaic of the failed images

As illustrated in figure 12, the processing of the thermal images from the gap separately generated an erroneous orthomosaic. To improve this result and evaluate the effectiveness of our fusion process, we apply it on these 27 images.
The validation dataset is firstly pre-processed and aligned using the registration workflow developed in the experimentation phase. After pairing and aligning the images, the dataset is fused. As determined during the experimentation phase, MGFF is the fusion method that yield the best results both qualitatively and quantitatively. Using this method, the gap images are fused before proceeding to the generation of their orthomosaic. The processing is

accurate. Tie points are detected on the fused images, resulting in successful generation of the orthomosaic of gap images (figure 13).

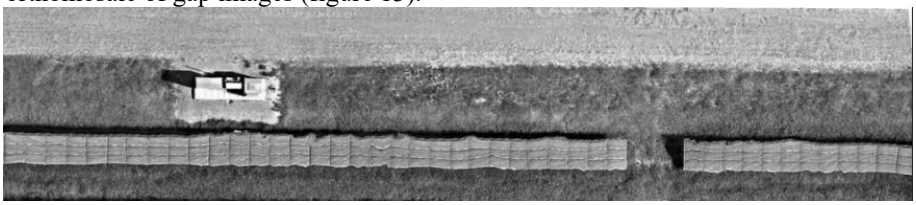


Fig.13. Orthomosaic of gap images fused using the MGFF method

Evidently, MGFF image fusion process enhanced the tie points detection on the gap images, permitting the generation of the orthomosaic (figure 13). We apply the fusion algorithm on the entire dataset comprising 499 images, and generate the orthomosaics of the fused images presented in figure 14.

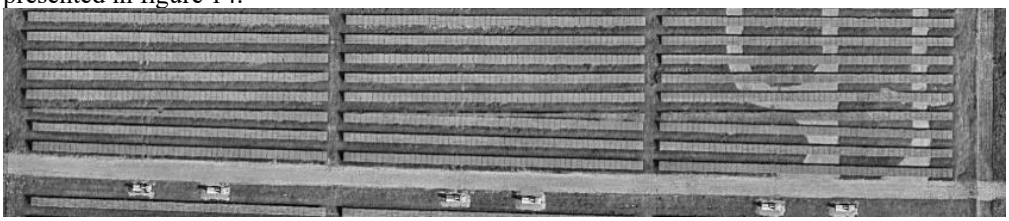


Fig.14. Orthomosaic of the MGFF-fused validation dataset

The generation of the orthomosaic of the full validation dataset reveals satisfactory insights. The MGFF fusion method significantly improved the generated orthomosaic. All fused images are aligned, eliminating the gaps present in the initial generation. The panel lines are straight, with the interference in line 3 from the initial generation corrected. However, line 7 still overlaps with its adjacent line.

Overall, the MGFF method proves effective in improving the orthomosaic generation of the validation dataset by eliminating gaps and correcting the interference between some lines.

6 Discussion

While thermal and visible image fusion is commonly used in security and inspection applications, its implementation for UAV-based images remains very limited, particularly for the inspection of photovoltaic installations. This underscores the significance of our research, as it makes a substantial contribution to this emerging field.

The fusion algorithms in our research require pairs of thermal and visible images to be precisely aligned for successful completion of the fusion process, highlighting the critical role of the registration step. However, well-known feature-based registration algorithms such as Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), and Oriented FAST and Rotated BRIEF (ORB) are unsuitable for our data due to challenges in detecting and matching keypoints caused by the repetitive pattern of the photovoltaic panels in the images. As a result, we are compelled to employ area-based registration algorithms to align the thermal and visible images.

From our comprehensive literature review, existing area-based registration algorithms typically apply a geometric transformation, regardless of the method used to estimate their transformation matrix. However, the data used in this study present significant challenges because they were captured with different focal lengths, resulting in varying extents between

thermal and visible images, as well as differing resolutions and dimensions. To address this, a customized registration process was developed. Initially, thermal and visible images were paired using a down-sampling percentage (k%) derived from the relationship between their pixel counts and the size of the detectors. This approach effectively mitigated the impact of differing focal lengths. Furthermore, lens distortion correction was employed to rectify curvature in the solar panel lines. Finally, the Enhanced Correlation Coefficient (ECC) algorithm executed the necessary geometric transformations to achieve precise alignment of paired thermal and visible images, ensuring accurate overlay.

Alignment of the data enabled the application of image fusion algorithms. Following a comparative analysis of nine methods (ADF, CBF, FPDE, GLPF, GTF, LatLRR, MGFF, MSVD, VGG) using evaluation metrics, MGFF and GTF yielded satisfactory results on the experimental dataset.

These methods were further evaluated using the validation dataset, where the MGFF method effectively enhanced the orthomosaic by emphasizing details in thermal images, thereby improving tie point detection in the orthomosaic generation. This success is attributed to the use of multi-scale guided filtering to preserve critical features from both thermal and visible light images. The orthomosaic of the validation dataset, initially marred by gaps, image misalignment, and solar panel line interference -due to tie point non-detection caused by thermal image resolution and repetitive solar panel details-, was successfully processed without gaps. It preserved radiometric information from thermal images crucial for fault detection in solar panels. Figure 15 resumes the proposed workflow to enhance the generation of orthomosaic applying visible and thermal image fusion technique.

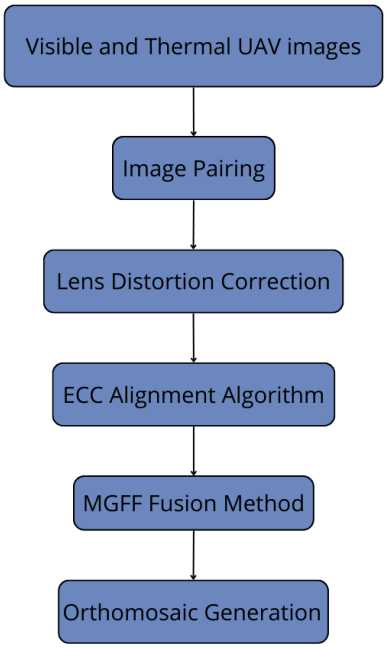


Fig.15. The proposed workflow to optimize the process of orthomosaic generation through the fusion of UAV-based thermal and visible images.

7 Conclusion

In conclusion, the integration of visible and thermal UAV-based image fusion offers significant potential for enhancing orthomosaic generation, particularly in the inspection of photovoltaic panels. With careful consideration of dataset characteristics, advanced image registration techniques, and ongoing customization of fusion algorithms, the effectiveness of these methodologies can be greatly improved. By focusing on preserving critical thermal information and exploring alternative metadata manipulation techniques, the remote sensing community is well-positioned to contribute significantly to the sustainable and efficient monitoring of photovoltaic installations, ultimately supporting the broader goal of advancing green energy solutions.

The Code developed for the registration workflow, all the fusion algorithms and the metrics computation are available in Github: https://github.com/fatima-zahra-bamessaoud/Image-Fusion-applied-to-PV-inspections

Fatima-zahra BAMESSAOUD and **Nouhaila BAHTIT**: Conceptualization, Methodology, Data curation, Visualization, Writing - Original Draft.
Imane SEBARI: Conceptualization, Methodology, Project administration, Supervision, Resources. Writing - Review & Editing.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. B. Li, C. Delpha, D. Diallo, A. Migan-Dubois, Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review. Renew. Sust. Energ. Rev. **138**, 110512 (2021). https://doi.org/10.1016/j.rser.2020.110512
2. J.A. Tsanakas, L. Ha, C. Buerhop, Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: A review of research and future challenges. Renew. Sust. Energ. Rev. **62**, 695-709 (2016). https://doi.org/10.1016/j.rser.2016.04.079
3. S.K. Firth, K.J. Lomas, S.J. Rees, A simple model of PV system performance and its use in fault detection. Sol. Energy **84**, 624-635 (2010). https://doi.org/10.1016/j.solener.2009.08.004
4. S. Naveen Venkatesh, V. Sugumaran, Fault Detection in aerial images of photovoltaic modules based on Deep learning. IOP Conf. Ser. Mater. Sci. Eng. **1012**, 012030 (2021). https://doi.org/10.1088/1757-899x/1012/1/012030
5. P. Guerriero, G. Cuozzo, S. Daliento, Health diagnostics of PV panels by means of single cell analysis of thermographic images, in Proceedings of the EEEIC conference (2016)
6. Y. Zefri, A. Elkettani, I. Sebari, S.A. Lamallam, Thermal infrared and visual inspection of photovoltaic installations by uav photogrammetry—application case: Morocco. Drones **2**, 1-24 (2018)
7. U. Pruthviraj, Y. Kashyap, E. Baxevanaki, P. Kosmopoulos, [Insert title of the article here]. Remote Sens. **15** (2023)
8. W.H. Maes, A.R. Huete, K. Steppe, Optimizing the processing of UAV-based thermal imagery. Remote Sens. **9** (2017). https://doi.org/10.3390/rs9050476
9. H. Hoffmann, H. Nieto, R. Jensen, R. Guzinski, P. Zarco-Tejada, T. Friborg, Estimating evaporation with thermal UAV data and two-source energy balance models. Hydrol. Earth Syst. Sci. **20**, 697-713 (2016). https://doi.org/10.5194/hess-20-697-2016

10. Y. Liu, S. Liu, Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation. Information Fusion **24**, 147-164 (2015). https://doi.org/10.1016/j.inffus.2014.09.004
11. Y. Zefri, I. Sebari, H. Hajji, G. Aniba, In-depth investigation of applied digital photogrammetry to imagery-based RGB and thermal infrared aerial inspection of large-scale photovoltaic installations. Remote Sens. Appl. **23** (2021) doi: 10.1016/j.rsase.2021.100576
12. B.S. Krishnan et al., Fusion of visible and thermal images improves automated detection and classification of animals for drone surveys. Sci. Rep. **13** (2023), doi: 10.1038/s41598-023-37295-7
13. J. Ma, C. Chen, C. Li, J. Huang, Infrared and visible image fusion via gradient transfer and total variation minimization. Information Fusion **31**, 100-109 (2016). https://doi.org/10.1016/j.inffus.2016.02.001
14. D.P. Bavirisetti, G. Xiao, J. Zhao, R. Dhuli, G. Liu, Multi-scale Guided Image and Video Fusion: A Fast and Efficient Approach. Circuits Syst. Signal Process. **38**, 5576-5605 (2019). https://doi.org/10.1007/s00034-019-01131-z