Comprehensive Analysis of Feature Detectors and Descriptors for Stitching UAV Images with Perspective Distortion

Mark P. B. Pacot [1,2,*] and Nelson Marcos [2]

[1] Department of Computer Science, College of Computing and Information Sciences, Caraga State University, Butuan City, Philippines

[2] Department of Software Technology, College of Computer Studies, De La Salle University, Manila, Philippines
Email: mbpacot@carsu.edu.ph (M.P.B.P.); nelson.marcos@dlsu.edu.ph (N.M.)
*Corresponding author

***Abstract*—This study analyzes traditional and deep-learning-based feature detectors and descriptors for Unmanned Aerial Vehicle (UAV) image stitching under perspective distortion. Traditional methods, including Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), and Accelerated-KAZE (AKAZE), achieve high keypoint detection and matching accuracy but struggle with geometric distortions, leading to artifacts such as ghosting and misalignment in complex UAV scenes. Deep-learning-based approaches, such as SuperPoint, Adaptive and Lightweight Key Point Detector (ALIKED), and Deep Image Structure Keypoints (DISK), offer superior alignment accuracy and visual coherence by detecting fewer but more robust keypoints. DISK and ALIKED demonstrate high spatial consistency, reducing perceptual artifacts as validated by Perceptual Image Quality Evaluator (PIQE), Naturalness Image Quality Evaluator (NIQE), and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) metrics. The results indicate that deep-learning-based detectors outperform traditional methods in UAV image stitching under perspective distortions. Future work will explore hybrid models that combine the strengths of both approaches to enhance stitching accuracy and computational efficiency.**

***Keywords*—stitching, feature detector and descriptor, perspective distortion, Unmanned Aerial Vehicle (UAV) images, no-reference metrics**

I. Introduction

In the realm of modern aerial imaging and remote sensing, Unmanned Aerial Vehicles (UAVs) have become indispensable tools for capturing high-resolution imagery of landscapes and complex scenes. The rapid expansion of UAV technology has led to an unprecedented influx of image data, which necessitates efficient and accurate image stitching methods to create seamless, panoramic views that are crucial for a wide range of applications. Image stitching, however, poses significant challenges when working with UAV imagery due to perspective distortions, which arise from the varying altitudes, angles, and flight trajectories at which these images are captured [1]. These distortions, often exacerbated by the dynamic and unpredictable movement of UAVs, demand robust feature detection and matching techniques capable of accommodating diverse image transformations.

This study focuses on the critical task of feature-based stitching for Unmanned Aerial Vehicle (UAV) images, emphasizing a comparative analysis of various feature detectors and descriptors, both low-level and deep learning-based, see Table I. As UAV technology continues to evolve and its applications expand, the findings from this study will be instrumental in optimizing the quality and accuracy of stitched images. This has broad implications for various real-world applications, including environmental monitoring, disaster response, agriculture, and urban planning. The insights gained will support the development of new algorithms and techniques tailored to the unique challenges of UAV-acquired imagery, especially in scenarios with significant perspective scene distortions.

TABLE I. Feature Detectors and Descriptors

Algorithm	Detector	Descriptor	Type
SIFT	Yes	Yes	Traditional
SURF	Yes	Yes	
MinEigen	Yes	Yes	
BRISK	Yes	Yes	
AKAZE	Yes	Yes	
DISK	Yes	Yes	Deep Learning
ALIKED	Yes	Yes	
SUPERPOINT	Yes	Yes	

Feature descriptors play a crucial role in image matching and alignment, forming the backbone of numerous stitching algorithms [2]. An in-depth understanding of the performance characteristics, computational demands, and robustness of different feature descriptors—such as Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB),

Manuscript received March 22, 2025; revised April 24, 2025; accepted June 3, 2025; published October 17, 2025.

SUPERPOINT, and Adaptive and Lightweight Key Point Detector (ALIKED)—is essential for optimizing the accuracy, speed, and reliability of stitched panoramas, particularly in the presence of challenging scene distortions.

Feature-based stitching algorithms, particularly those employing RANSAC for homography estimation and bundle adjustment for error minimization, have been widely validated in the literature [3, 4] and through extensive experimental results in various datasets, demonstrating their effectiveness in reconstructing seamless mosaics from multiple overlapping images. These algorithms play a pivotal role in compensating for distortions and enhancing overall image quality by aligning images with minimal ghosting and misalignment artifacts. Through a detailed performance assessment of feature descriptors within these stitching algorithms, this research aims to advance the precision and computational efficiency of image mosaicking for UAV applications. Specifically, the study evaluates the effectiveness of descriptors such as SIFT, SURF, ORB, and newer, deep learning-based descriptors like SuperPoint, DISK, and ALIKED. These descriptors are tested on their robustness to various distortions, including scale, rotation, affine transformations, and illumination changes, thereby providing a comprehensive overview of each method's suitability for UAV imagery.

The primary objectives of this research are to explore and evaluate feature descriptors utilized in stitching UAV images, specifically in scenarios with perspective scene distortions. However, the variations in altitude, pitch, roll, and yaw introduce complex spatial transformations in the captured images. These transformations not only complicate image alignment but also challenge traditional stitching methods, which are generally designed for planar scenes with minimal perspective changes [5–7].

By conducting a rigorous performance analysis, this study contributes to advancing image stitching techniques applicable in the fields of aerial photography and remote sensing. UAV-acquired images are often subject to geometric and radiometric distortions due to factors such as the interplay between the camera sensor, lens characteristics, and three-dimensional scene structure [8, 9]. UAVs, commonly referred to as drones, are widely used in environmental monitoring, disaster assessment, and precision agriculture due to their ability to capture high-resolution aerial images with flexible deployment [10, 11].

This paper is structured to offer an in-depth understanding of the topic. Section II provides a detailed literature review, examining the state-of-the-art techniques in feature descriptors for feature-based stitching of UAV images with scene distortions. Section III outlines the feature detectors and descriptors applied to pairwise UAV image stitching, followed by an evaluation of image quality using both subjective and objective metrics, including metrics for geometric alignment accuracy and perceptual quality. The findings underscore the impact of each feature descriptor's strengths and limitations in dealing with the challenges unique to UAV-based image stitching, paving the way for more robust and efficient solutions in this rapidly evolving field.

II. Literature Review

In computer vision, feature descriptors are critical components that enable the identification, characterization, and comparison of key features within images, facilitating robust image matching and alignment. This section provides a detailed overview of various feature descriptors, with a particular emphasis on their applicability to UAV image stitching. Additionally, the discussion focuses on the effectiveness of these descriptors in addressing perspective distortions often encountered in UAV-acquired imagery due to variations in altitude, orientation, and camera angle.

A. Low-Level Feature Detectors and Descriptors

Traditional feature descriptors form the foundation of many computer vision and image processing applications, including image stitching, object recognition, and scene reconstruction. Their role is critical in detecting and matching keypoints across images to ensure coherent mosaicking. In UAV image stitching, where scene distortion, illumination variation, and changes in scale and viewpoint are common, selecting appropriate descriptors is essential for accurate alignment and high-quality results.

Among traditional descriptors, the Scale-Invariant Feature Transform (SIFT) stands out for its robustness to scale, rotation, and illumination changes. SIFT remains a go-to solution for UAV image stitching in scenarios involving varying altitudes and camera angles, consistently delivering stable keypoint matches under complex conditions [12], as illustrated in Fig. 1.

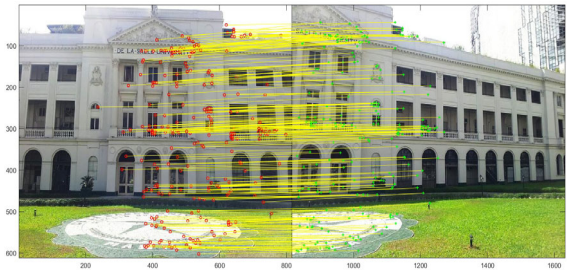


Fig. 1. Feature point matching using SIFT.

The Speeded-Up Robust Features (SURF) descriptor offers a good balance between accuracy and computational efficiency, making it well-suited for UAV applications requiring real-time performance. Its resilience to scale and transformation helps maintain accurate matches across dynamic scenes [13], as illustrated in Fig. 2.

The Minimum Eigenvalue (MinEigen) corner detector provides a fast and efficient method for identifying prominent image features, making it valuable for time-sensitive UAV applications like aerial surveillance and mapping. While simple, it enables scalable processing of

large image sequences without significant computational overhead, as illustrated in Fig. 3.

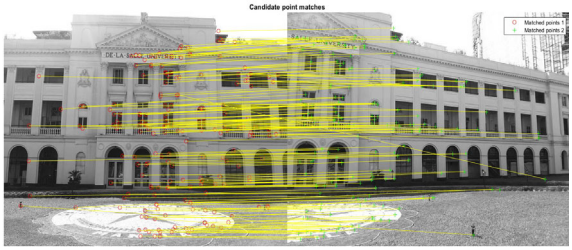


Fig. 2. Feature point matching using SURF.

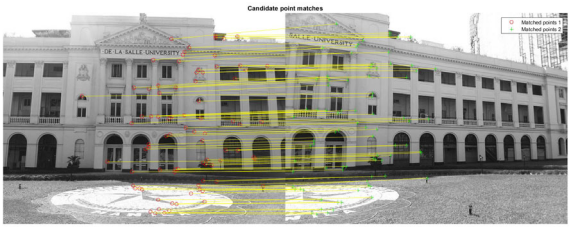


Fig. 3. Feature point matching using MinEigen.

Binary Robust Invariant Scalable Keypoints (BRISK) offers strong performance in handling distortions due to non-planar terrain in UAV-acquired imagery. Its binary descriptors and circular sampling pattern ensure resilience to geometric transformations while maintaining efficiency, making it suitable for irregular landscapes and topographic analysis [14], as illustrated in Fig. 4.

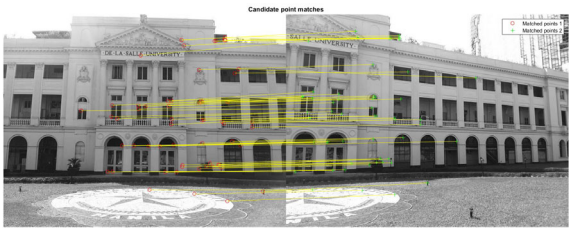


Fig. 4. Feature point matching using BRISK.

Accelerated-KAZE (AKAZE) leverages nonlinear scale spaces to extract robust features across varying resolutions and distortion levels. Its design balances computational speed and descriptor richness, enabling effective performance in applications like environmental monitoring and agricultural assessment [15], as illustrated in Fig. 5.

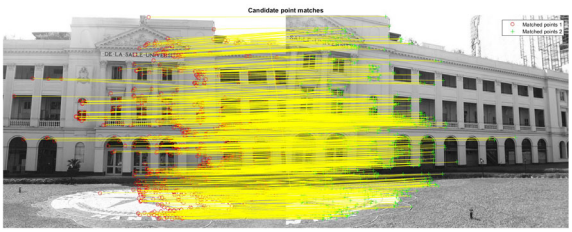


Fig. 5. Feature point matching using AKAZE.

Comparative evaluations of these traditional descriptors reveal their respective strengths: SIFT and SURF excel in robustness, while BRISK and AKAZE offer competitive accuracy with improved runtime. However, limitations persist, particularly in managing severe perspective distortions and rapidly changing viewpoints. These challenges continue to motivate improvements to traditional descriptors and the exploration of alternative approaches, including hybrid and learning-based strategies.

B. Deep Learning-Based Feature Detectors and Descriptors

Deep learning-based descriptors have redefined the landscape of feature detection and matching by enabling automatic learning of robust, highly discriminative features from image data. In UAV image stitching—where traditional descriptors often struggle with significant distortions, deep learning approaches have demonstrated considerable advantages.

Convolutional Neural Networks (CNNs) power modern detectors like DISK, ALIKED, and SuperPoint, which jointly detect keypoints and extract descriptors using a learned feature representation. These methods have proven effective in capturing complex image structures while maintaining robustness to scale, rotation, and lighting variations.

Deep Image Structure Keypoints (DISK), employs CNN-based architecture to extract dense feature maps $F = F(I)$ from an input image I, from which keypoints and descriptors are generated. Its primary strength lies in its ability to robustly identify meaningful structures, even in the presence of significant geometric distortion, as illustrated in Fig. 6. In this work, we adopt the original architecture and pre-trained model as proposed by Tyszkiewicz *et al*. [16], without applying additional fine-tuning or hyperparameter adjustments. Our objective is to evaluate the baseline performance of these models on UAV image stitching under perspective distortion, using their default configurations.

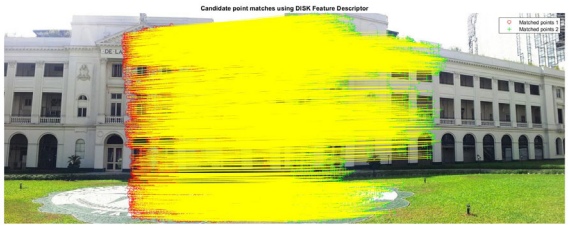


Fig. 6. Feature point matching using DISK.

Adaptive and Lightweight Key Point Detector (ALIKED) introduces a computationally efficient method that adapts its feature extraction complexity to image content. It reduces redundancy in uniform regions and focuses on textured areas, offering a balance between speed and precision, as illustrated in Fig. 7.

In this work, we adopt the original architecture and pre-trained model as proposed by Zhao *et al*. [17], without applying additional fine-tuning or hyperparameter adjustments. Our objective is to evaluate the baseline performance of these models on UAV image stitching under perspective distortion, using their default configurations.

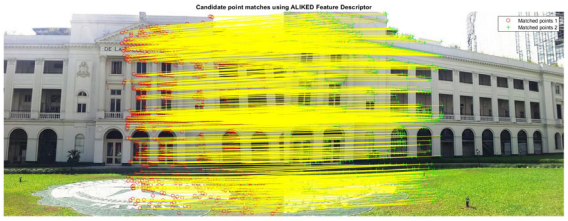


Fig. 7. Feature point matching using ALIKED.

SuperPoint leverages self-supervised learning with synthetic data to train a unified model that detects keypoints and descriptors. Its dual-decoder architecture handles keypoint localization and descriptor extraction, making it effective for dense feature matching in aerial scenes. In this work, we adopt the original architecture and pre-trained model as proposed by DeTone *et al*. [18], without applying additional fine-tuning or hyperparameter adjustments. Our objective is to evaluate the baseline performance of these models on UAV image stitching under perspective distortion, using their default configurations, as illustrated in Fig. 8.

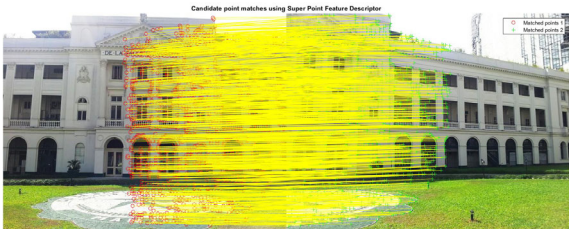


Fig. 8. Feature point matching using SuperPoint.

Studies such as [19, 20] demonstrate the superior performance of these CNN-based methods in extracting stable and discriminative features, particularly in dynamic environments. Beyond CNNs, Siamese and Triplet networks have shown promise in learning fine-grained feature embeddings for more precise matching under challenging UAV conditions [21, 22].

Recent advancements have explored integrating deep learning with traditional descriptors (e.g., SIFT, ORB), producing hybrid models that combine the interpretability and efficiency of hand-crafted features with the learning capabilities of CNNs. This synergy improves matching precision while maintaining practical runtime efficiency, as shown in works like [23, 24].

Nevertheless, despite these advances, deep learning-based methods are not yet universally adopted in UAV stitching. Traditional descriptors continue to dominate due to their computational simplicity, established performance, and ease of deployment—particularly in real-time, resource-constrained scenarios. Moreover, techniques such as adaptive match filtering further enhance the reliability of traditional methods by reducing false matches and improving global alignment [25].

In summary, while deep learning-based feature descriptors significantly improve matching robustness and perceptual quality, practical deployment in UAV stitching still demands careful consideration of their computational cost and generalization ability under diverse real-world conditions.

III. Materials and Methods

The focus of this research is to identify the optimal feature detectors and descriptors for stitching UAV images, considering perspective scene distortions and their impact on the visual quality of the resulting mosaic. The resulting mosaics will then be evaluated using quantitative measurements to assess their effectiveness.

A. Dataset Description

The aerial photogrammetric survey was conducted using a consumer-grade camera (SONY DSC-RX100M3, 20.1-megapixel sensor with approximately 2 cm spatial resolution) with a 9 mm focal length (equivalent to 24 mm in 35 mm format) mounted on a UAV. The image in Fig. 9 was captured at an altitude of approximately 260.64 Meters above ground level, as recorded in the EXIF metadata. The GPS coordinates of the capture location are 15°5′9.54″ N and 120°50′31.69″ E, situating the survey over a predominantly agricultural area. The camera settings, including auto white balance and manual exposure, ensured consistent imaging conditions suitable for feature-based stitching and alignment. Additional datasets were provided by [25, 26] for ground-based images. Table II provides detailed characteristics of the test images used in this study.

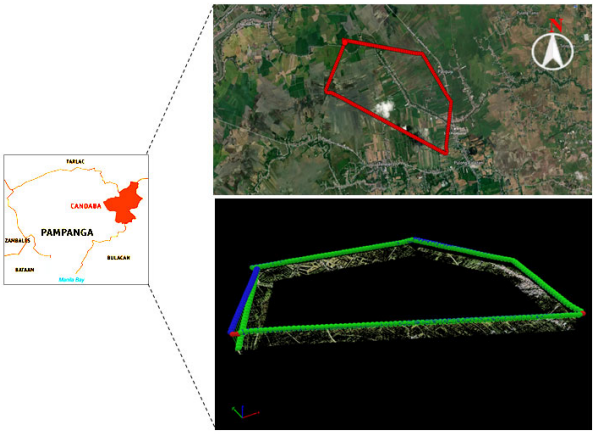


Fig. 9. Flight plan of data acquisition [4].

B. Feature-Refinement Technique

A major part of our proposed solution is based on the work, titled “Parallax tolerant image stitching based on Robust Elastic Warping (REW)” [25]. This method is particularly valuable due to its refinement technique applied to distinct detected image features. The Bayesian model introduced by Li *et al*. [25] serves to iteratively refine feature matches by probabilistically removing incorrect local correspondences. Although this Bayesian model was not initially tailored for UAV imagery, we extend its application to UAV-specific image sequences by exploring both low-level and deep learning-based feature detectors and descriptors. This approach allows us to evaluate a variety of feature detection techniques and identify those that achieve the highest accuracy and stability in the complex, perspective distorted imagery typically acquired by UAVs. We examine the

effectiveness of traditional feature descriptors and detectors for capturing fine-grained, texture-based features, as well as the capabilities of deep learning-based algorithms like SuperPoint, DISK, and ALIKED. By exploring the performance of these methods, we aim to determine the most suitable approach for accommodating the complex variations in perspective distortion encountered in UAV imagery, ultimately enhancing the overall quality of the stitched mosaics affected from such type of distortion.

TABLE II. OVERVIEW OF THE DATASETS USED IN THIS STUDY, INCLUDING GENERAL CHARACTERISTICS RELEVANT TO IMAGE ACQUISITION CONDITIONS AND SCENE DIVERSITY

Dataset	Image Source	Lightning Conditions	Weather Conditions	Camera View
General Dataset 1	General dataset (Indoor)	Uniform lighting	Indoor environment	Uniform horizontal view
General Dataset 2	General dataset (Outdoor)	Natural daylight	Clear	Oblique view
Our UAV dataset (A)	UAV-based	Natural daylight	Clear	Vertical (nadir) view, slightly tilt
Our UAV dataset (B)	UAV-based	Natural daylight	Clear	Vertical (nadir) view, slightly tilt
Fig. 19-Our UAV dataset	UAV-based	Sunny, shadowed	Cloud presence	Vertical (nadir) view, extreme geometric transformation
Fig. 20-General Dataset (A)	General dataset (Outdoor)	Natural daylight	Clear	Non-uniform horizontal view, extreme geometric transformation
Fig. 20-General Dataset (B)	General dataset (Outdoor)	Natural daylight	Clear	Non-uniform horizontal view, extreme geometric transformation

To mathematically represent the Bayesian refinement process inspired from [27, 28], we consider a set of matched points $\{(x_i, y_i), (x_i', y_i')\}$ and model the posterior probability of a match being correct as:

$$P(\text{match}\,|\,\text{correct}) \propto P(\text{match}\,|\,\text{correct})P(\text{correct}) \quad (1)$$

Assuming prior probabilities and a likelihood function based on feature descriptor distance, we can define this likelihood function as:

$$p(\text{match}\,|\,\text{correct}) = \exp\left(\frac{-\left\|d(x_i, y_i) - d(x_i', y_i')\right\|^2}{2\sigma^2}\right) \quad (2)$$

where $d(\cdot)$ represents the feature descriptor, and σ is a parameter that controls the standard deviation, effectively influencing the sensitivity of the refinement process. The Bayesian framework iteratively refines the match probability, updating the set of matches to improve the reliability of correspondences.

This refinement process is applied independently to the results from each feature detection and description method. By doing so, we can systematically evaluate the effectiveness of various feature detectors and descriptors under the unique conditions present in UAV imagery, including challenges such as perspective distortion and varying scales. Through this exploration, we seek to identify the optimal feature refinement techniques that yield high-quality, coherent image mosaics in UAV-based applications.

C. Feature Detectors, Feature Descriptors, and Pair-Wise Stitching

Feature detectors are fundamental algorithms in computer vision, playing a pivotal role in identifying significant points or regions within an image, often referred to as keypoints or interest points [29, 30]. These keypoints serve as essential reference points for subsequent image analysis tasks, including feature matching and alignment. One notable feature detector is the Harris corner detector, which evaluates the "cornerness" of each pixel in an image using the following equation:

$$R = \det(M) - k \cdot (\text{trace}(M))^2 \quad (3)$$

where R represents the corner response function, M is the structure tensor matrix, and:

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (4)$$

where, I_x and I_y are the partial derivatives of the image intensity in the x and y directions, respectively. The term $\det(M)$ is the determinant of M, calculated as $I_x^2 I_y^2 - (I_x I_y)^2$, and $\text{trace}(M)$ is the trace of M, given by $I_x^2 + I_y^2$.

While feature detectors identify keypoints, feature descriptors complement them by providing detailed representations of the local image content surrounding these keypoints. Descriptors encode both the appearance and geometry of the region around a keypoint, making it possible to match keypoints between different images effectively. A prominent example of a feature descriptor is the Scale-Invariant Feature Transform (SIFT), which quantizes gradient orientations within a local region and creates a histogram of gradient orientations. This results in a descriptive feature vector:

$$\text{SIFT Descriptor} = \left[\text{Hist}_1, \text{Hist}_2, ..., \text{Hist}_n\right] \quad (5)$$

where each $Hist_i$ in the vector represents a histogram bin, capturing the distribution of gradient orientations within a specific angular range.

Feature descriptors like SIFT are crucial for recognizing and aligning keypoints when stitching images. Pairwise stitching is a crucial step in creating seamless panoramic images, involving a series of mathematical operations to ensure the alignment and blending of two adjacent images. This process can be broken down into several key components:

1) Keypoint matching

Pairwise stitching begins with matching keypoints between two adjacent images based on their feature descriptors. The process often employs a distance metric, such as Euclidean distance, to determine the similarity between keypoints in different images.

$$\text{Matched Keypoints}: \{(p_1, q_1), (p_2, q_2), ..., (p_n, q_n)\} \quad (6)$$

where, p_i and q_i represent placeholder key points in two images that are matched and used to calculate transformations during the stitching process.

2) Homography estimation

Once keypoints are matched, the next step is to estimate a homography matrix (H) that represents the transformation needed to align one image with the other. This transformation can be determined using a least-squares optimization approach that minimizes the reprojection error of matched keypoints:

$$H = \arg\min \sum_i d(p_i, H \times (q_i))^2 \quad (7)$$

where, H is the homography matrix, p_i and q_i are corresponding keypoints, and d is the Euclidean distance between the transformed point $H \times (q_i)$ and the corresponding point p_i.

3) Image warping

With the homography matrix computed, one of the images is warped to align with the other. The transformation involves applying the homography matrix to the pixel coordinates of one image to match the perspective of the other.

4) Blending

After alignment, the two images are blended to create a seamless transition between them. Various blending techniques can be employed, such as feathering or multi-band blending, which may involve weighted averages of pixel values.

This structured approach, combining feature detection, keypoint matching, homography estimation, warping, and blending, is essential for constructing accurate and visually appealing panoramic images from UAV-acquired data.

By carefully selecting feature detectors and descriptors, and optimizing the transformation and blending processes, we aim to address the unique challenges posed by UAV imagery, such as perspective distortion and scale variation.

D. Image Quality Metrics

In this study, the researcher employed the Perceptual Image Quality Evaluator (PIQE) to classify UAV-based image mosaics into five quality categories—Excellent ([0–20]), Good ([21–35]), Fair ([36–50]), Poor ([51–80]), and Bad ([81–100])—based on local distortions and artifacts in the image. As a no-reference quality assessment metric, PIQE is particularly suitable for evaluating mosaics where ground truth images are unavailable, offering a perceptual score that aligns well with human visual interpretation. Complementing PIQE, the study also utilizes two additional no-reference image quality metrics: the Naturalness Image Quality Evaluator (NIQE) and the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), both of which assess perceptual features to quantitatively evaluate image quality in the absence of a reference image.

IV. Experiments

This section presents experimental results evaluating the effectiveness of various feature detectors and descriptors within a feature-based stitching framework for UAV images affected by perspective distortions. The analysis highlights the strengths and limitations of each method, providing technical validation of their suitability for robust image alignment under challenging scene variations.

A. Feature Detection and Matching

In this section, we evaluate the performance of various feature detectors and descriptors by analyzing their ability to detect and match keypoints across pairs of overlapping images. Effective feature detection and matching are essential for establishing accurate correspondences, which form the foundation for robust image alignment. By comparing traditional and deep learning-based methods, we aim to determine which techniques are most resilient to perspective distortions commonly present in UAV imagery, as illustrated in Fig. 10.

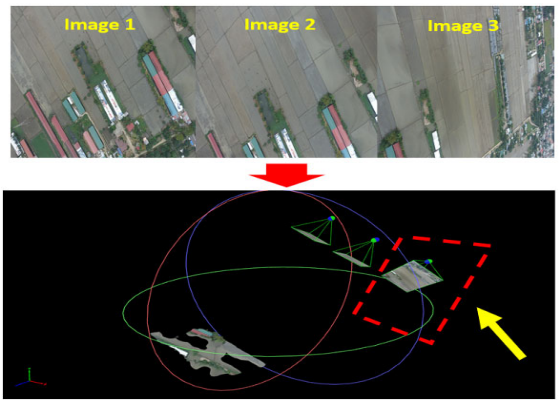


Fig. 10. Sample sequences of UAV images with perspective distortion.

Table III presents a comparative analysis of traditional and deep learning-based feature detectors and descriptors across general and UAV-specific datasets, evaluated using three key metrics: number of detected keypoints, matching accuracy (%), and average matching time (s). Traditional detectors such as AKAZE, SURF, and BRISK generally detect a larger number of keypoints and

demonstrate strong matching accuracy, with AKAZE and DISK reaching near-perfect values in specific cases.

TABLE III. SUMMARY OF FEATURE DETECTION AND MATCHING RESULTS ACROSS DATASETS

Dataset	Detector	Keypoints	Match Acc.	Time (s)
Dataset 1 (Zaragosa *et al*. [18])	SIFT	3208.2	94.94	0.84
	SURF	20632.4	91.45	1.43
	MINEIGEN	15943.0	99.70	0.77
	BRISK	20875.2	99.93	1.34
	AKAZE	24403.0	99.38	2.22
	SUPERPOINT	4459.6	98.73	2.93
	ALIKED	3834.6	99.92	1.37
	DISK	8376.8	100.0	7.51
Dataset 2 (Chen *et al.* [31])	SIFT	3337.5	92.84	0.80
	SURF	1259.2	96.31	0.01
	MINEIGEN	3269.7	99.23	0.03
	BRISK	2766.8	98.82	0.02
	AKAZE	5285.2	99.62	0.09
	SUPERPOINT	2605.5	97.62	1.04
	ALIKED	3768.2	99.47	1.28
	DISK	8476.7	99.78	7.46
UAV-Dataset (A)	SIFT	3431.9	93.73	0.89
	SURF	376.2	99.45	0.06
	MINEIGEN	1155.0	99.63	0.01
	BRISK	1024.9	100.0	0.01
	AKAZE	1608.9	100.0	0.01
	SUPERPOINT	2372.9	97.21	0.90
	ALIKED	2374.0	99.79	0.53
	DISK	6974.7	99.98	5.08
UAV-Dataset (B)	SIFT	3487.4	81.05	0.94
	SURF	113.6	99.44	0.01
	MINEIGEN	625.61	99.13	0.01
	BRISK	171.9	100.0	0.01
	AKAZE	764.8	100.0	0.01
	SUPERPOINT	2620.4	94.56	1.07
	ALIKED	1321.3	99.49	0.20
	DISK	5696.6	99.90	3.51

Deep learning-based methods—SUPERPOINT, ALIKED, and DISK—typically produce fewer but more distinctive keypoints, achieving competitive matching accuracy (often exceeding 94%) while offering favorable runtime performance in many CPU-based configurations. However, it is important to emphasize that these methods are inherently optimized for GPU execution; thus, the reported CPU runtimes may not fully reflect their computational efficiency in real-time applications.

Our original CPU-based feature matching pipeline lacked parallel processing capabilities and relied on external libraries, limiting both efficiency and portability. In contrast, the updated GPU-based implementation leverages parallel computation to significantly accelerate feature matching—particularly for large-scale keypoint sets—while eliminating the reliance on external dependencies. Notably, this acceleration was achieved without modifying the core architecture or fine-tuning the pre-trained models of the deep learning-based detectors, thereby preserving their original design and ensuring consistency across evaluations.

These implementation choices provide a fair baseline for comparing execution time between CPU-based results (as presented in Table III) and GPU-accelerated outcomes (as visually illustrated in Table IV). Although the numerical differences in matching accuracy among the detectors are relatively small, deep learning-based methods consistently yield much better stitching results, demonstrating increased robustness to perspective distortions and producing smoother image transitions. These qualitative advantages are especially apparent in UAV imagery, both in pairwise and multiple image mosaicking scenarios.

B. Pair-Wise Stitching Technique

Pair-wise stitching builds on feature detection and matching by applying homography transformations to align and blend two adjacent images. This process is critical for creating seamless panoramic views from individual frames. In this section, we assess the quality of the stitched images using both perceptual metrics and spatial coverage evaluations, allowing us to quantify each method's effectiveness in producing visually cohesive mosaics under challenging conditions, see Table V. Additionally, Figs. 11–14 provide qualitative evaluations of the stitched images, showcasing the visual coherence and alignment accuracy achieved by each method.

TABLE IV. SUMMARY OF DATASET AND THEIR CORRESPONDING GPU-BASED AND CPU-BASED IMPLEMENTATIONS

Algorithm	Dataset 1 (Zaragosa *et al*. [18])		Dataset 2 (Chen *et al*. [31])		UAV-Dataset (A)		UAV-Dataset (B)	
	CPU-based	GPU-based	CPU-based	GPU-based	CPU-based	GPU-based	CPU-based	GPU-based
	*Measured in Seconds (s)							
SUPERPOINT	2.93	0.44	1.04	0.13	0.90	0.11	1.07	0.12
ALIKED	1.37	0.25	1.28	0.25	0.53	0.10	0.20	0.04
DISK	7.51	3.27	7.46	4.87	5.08	1.29	3.51	0.58

Table V presents PIQE, NIQE, and BRISQUE scores for stitched images across four datasets, comparing the perceptual quality of different feature detectors. In Dataset 1 [20], ALIKED achieves the lowest PIQE (36.13), SURF yields the best NIQE (2.41), and AKAZE records the lowest BRISQUE score (21.81), indicating fewer perceptual distortions. In Dataset 2 [21], ALIKED leads in all metrics PIQE (44.27), MINEIGEN performs best in NIQE (2.90), and SURF achieves the best BRISQUE score (34.74). For UAV-Dataset A, BRISK leads in all metrics PIQE (18.63), NIQE (2.96), and BRISQUE (27.66). In UAV-Dataset B, DISK leads in all metrics PIQE (21.45), BRISQUE score (33.35), and NIQE (2.62). Overall, deep-learning-based detectors show competitive performance in PIQE, BRISQUE, and NIQE across several datasets and metrics. While detectors such as DISK, SUPERPOINT, and ALIKED achieved lower scores in certain cases, the differences were not statistically significant ($p > 0.05$). These results suggest that although deep-learning-based methods are visually competitive, their perceptual advantages are not

consistently significant or generalizable across all scenarios.

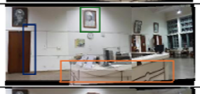
Methods	Output	Error(s)
REW with SIFT		• Ghosting • Misalignment • Discontinuous Edges
REW with SURF		• Ghosting • Misalignment • Discontinuous Edges
REW with MinEigen		• Ghosting • Misalignment
REW with BRISK		• Misalignment
REW with AKAZE		• Ghosting • Misalignment • Discontinuous Edges
REW with DISK		• Misalignment • Discontinuous Edges
REW with ALIKED		• Misalignment • Skewness
REW with SUPERPOINT		• Skewness

Fig. 11. Pair-wise image stitching results using general Dataset 1 [21].

Methods	Output	Error(s)
REW with SIFT		• Ghosting • Parallax
REW with SURF		• Ghosting • Parallax
REW with MinEigen		• Ghosting • Parallax
REW with BRISK		• Ghosting • Parallax
REW with AKAZE		• Ghosting • Parallax
REW with DISK		• Ghosting • Parallax
REW with ALIKED		• Ghosting • Parallax
REW with SUPERPOINT		• Ghosting • Parallax

Fig. 12. Pair-wise image stitching results using general Dataset 2 [18].

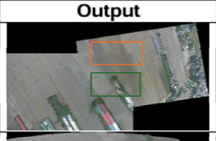
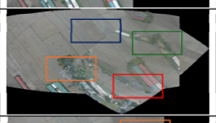
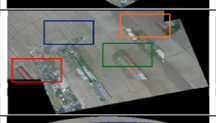
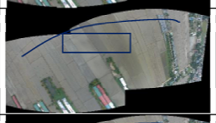
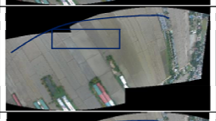
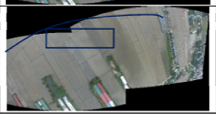
Methods	Output	Error(s)
REW with SIFT		• Ghosting • Discontinuous Edges
REW with SURF		• Ghosting • Misalignment • Discontinuous Edges • Parallax
REW with MinEigen		• Ghosting • Misalignment • Discontinuous Edges • Parallax
REW with BRISK		• Ghosting • Misalignment • Skewness • Discontinuous Edges • Parallax
REW with AKAZE		• Misalignment • Skewness
REW with DISK		• Misalignment • Skewness
REW with ALIKED		• Misalignment • Skewness
REW with SUPERPOINT		• Misalignment • Skewness

Fig. 13. Pair-wise image stitching results using our UAV Dataset (A).

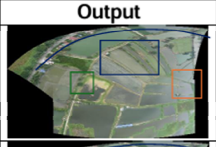
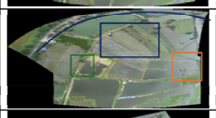
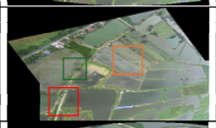
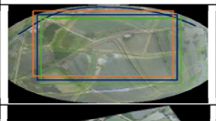
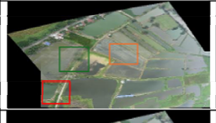
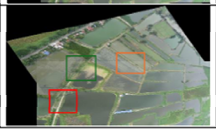
Methods	Output	Error(s)
REW with SIFT		• Ghosting • Misalignment • Skewness • Discontinuous Edges
REW with SURF		• Ghosting • Misalignment • Skewness • Discontinuous Edges
REW with MinEigen		• Ghosting • Discontinuous Edges • Parallax
REW with BRISK		• Ghosting • Misalignment • Skewness • Discontinuous Edges
REW with AKAZE	Error	–
REW with DISK		• Ghosting • Misalignment • Skewness • Discontinuous Edges
REW with ALIKED		• Ghosting • Discontinuous Edges • Parallax
REW with SUPERPOINT		• Ghosting • Discontinuous Edges • Parallax

Fig. 14. Pair-wise image stitching results using our UAV Dataset (B).

TABLE V. PIQE, NIQE, AND BRISQUE SCORES FOR PAIR WISE STITCHED IMAGES ACROSS DATASETS. THE STITCHED IMAGES ANALYZED HERE ARE DERIVED FROM FIGS [11–14].

Dataset	Detector	PIQE	NIQE	BRISQUE
Dataset 1 (Jia *et al.* [21]), 2 imgs.	SIFT	38.23	2.73	22.06
	SURF	39.54	2.41	22.90
	MINEIGEN	36.58	2.80	22.37
	BRISK	38.71	2.71	22.54
	AKAZE	38.16	2.59	21.81
	SUPERPOINT	37.45	2.82	22.49
	ALIKED	36.13	2.77	22.31
	DISK	37.22	2.49	22.47
Dataset 2 (Zaragosa *et al.* [18]), 2 imgs.	SIFT	45.69	3.02	36.80
	SURF	46.07	3.32	34.74
	MINEIGEN	45.03	2.90	36.34
	BRISK	46.65	3.12	38.13
	AKAZE	46.32	2.96	35.13
	SUPERPOINT	47.08	3.06	35.53
	ALIKED	44.27	3.04	36.28
	DISK	46.12	3.00	35.13
UAV-Dataset (A), 2 imgs.	SIFT	29.12	5.23	36.36
	SURF	19.37	4.73	32.99
	MINEIGEN	23.47	4.71	39.06
	BRISK	18.36	2.96	27.66
	AKAZE	24.98	4.48	38.13
	SUPERPOINT	25.09	4.49	37.75
	ALIKED	24.61	4.74	35.85
	DISK	24.73	4.56	37.11
UAV-Dataset (B), 2 imgs.	SIFT	30.30	4.26	36.63
	SURF	29.77	4.07	38.07
	MINEIGEN	29.94	4.68	36.79
	BRISK	30.50	4.39	36.23
	AKAZE	32.69	8.81	44.78
	SUPERPOINT	30.09	4.50	37.13
	ALIKED	29.33	4.77	36.83
	DISK	21.45	2.62	33.35

C. *Multiple Image Stitching Technique*

Expanding beyond pair-wise stitching, this section examines the scalability of feature detectors and descriptors in the context of multiple image stitching. By sequentially aligning a greater number of overlapping images, we generate larger mosaics composed of multiple input images, effectively simulating real-world UAV imagery applications. Examples and quantitative results are provided in Table VI. Additionally, Figs. 15–18 provide qualitative evaluations of the stitched images, highlighting the visual coherence and alignment accuracy achieved by each method. This assessment offers insights into each method's ability to maintain alignment quality and spatial consistency as the complexity of the stitching task increases.

Table VI presents PIQE, NIQE, and BRISQUE scores for stitched images across four datasets, comparing the perceptual quality of different feature detectors. In Dataset 1 [20], SURF achieves the lowest PIQE (32.62), SUPERPOINT yields the best NIQE (1.56), and DISK records the lowest BRISQUE score (26.80), indicating fewer perceptual distortions. In Dataset 2 [21], DISK leads in PIQE (28.27), BRISK performs best in NIQE (2.27), and SUPERPOINT achieves the best BRISQUE score (36.12). For UAV-Dataset A, SIFT attains the lowest PIQE (22.99), MINEIGEN shows the best performance in NIQE (2.97), and SIFT also records the lowest BRISQUE (32.06). In UAV-Dataset B, AKAZE leads in PIQE (37.62), SIFT achieves the best BRISQUE score (37.03), while BRISK records the lowest NIQE (3.20). Overall, deep-learning-based detectors show competitive performance in PIQE, BRISQUE, and NIQE across several datasets and metrics. While detectors such as DISK, SUPERPOINT, and ALIKED achieved lower scores in certain cases, the differences were not statistically significant ($p > 0.05$). These results suggest that although deep-learning-based methods are visually competitive, their perceptual advantages are not consistently significant or generalizable across all scenarios.

Fig. 15. Multiple image stitching results using Dataset 1 [18].

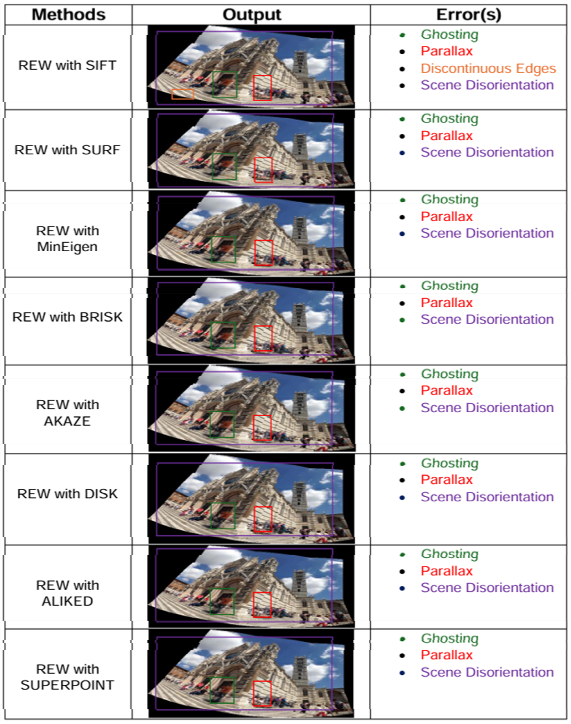


Fig. 16. Multiple image stitching results using Dataset 2 [31].

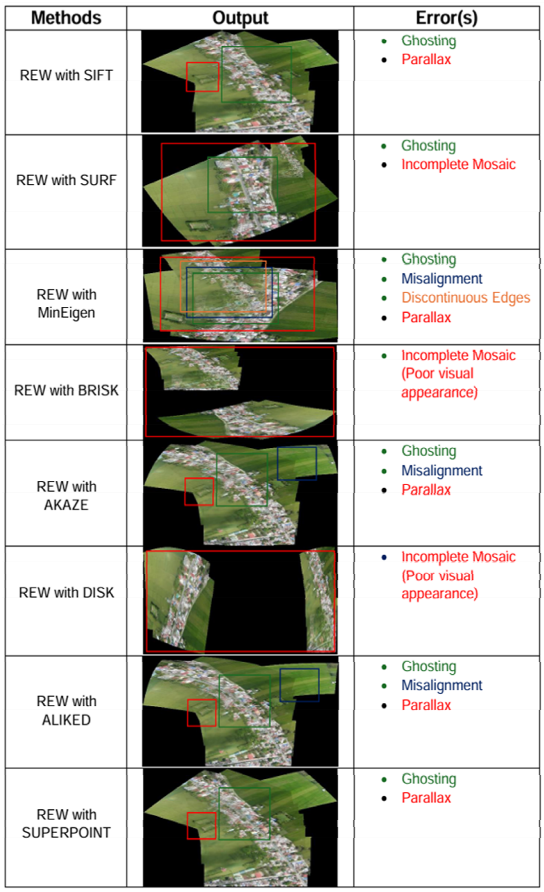


Fig. 17. Multiple image stitching results using our UAV Dataset (A).

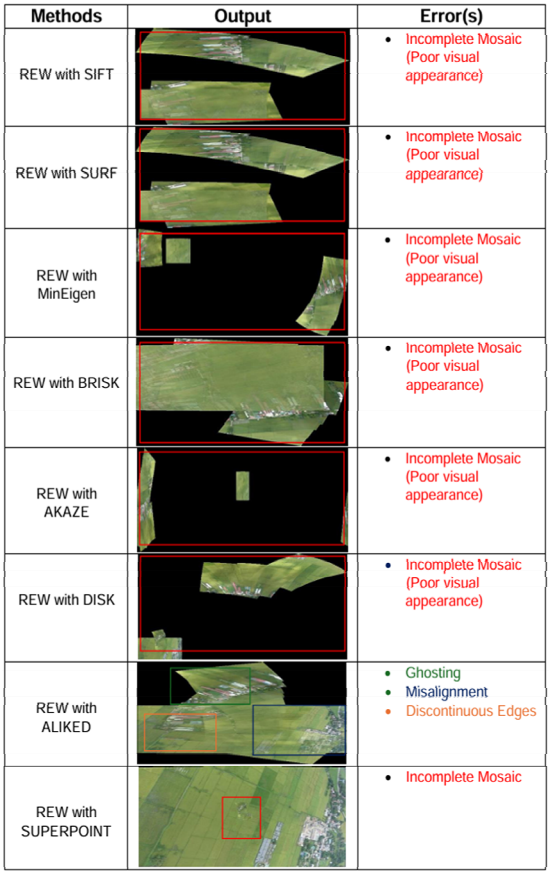


Fig. 18. Multiple image stitching results using our UAV Dataset (B).

TABLE VI. PIQE, NIQE, AND BRISQUE SCORES FOR PAIR WISE STITCHED IMAGES ACROSS DATASETS. THE STITCHED IMAGES ANALYZED HERE ARE DERIVED FROM FIGS. 15–18.

Dataset	Detector	PIQE	NIQE	BRISQUE
Dataset 1 (Zaragosa *et al*. [18]), 5 imgs.	SIFT	32.79	1.64	27.90
	SURF	32.62	1.61	27.85
	MINEIGEN	33.43	1.59	27.44
	BRISK	33.16	1.58	27.40
	AKAZE	33.08	1.60	27.31
	SUPERPOINT	33.06	1.56	27.53
	ALIKED	33.09	1.58	27.28
	DISK	32.88	1.58	26.80
Dataset 2 (Chen *et al*. [31]), 6 imgs.	SIFT	28.98	2.72	36.48
	SURF	28.92	2.70	37.29
	MINEIGEN	29.46	2.73	36.99
	BRISK	28.39	2.57	37.12
	AKAZE	28.57	2.70	36.98
	SUPERPOINT	28.33	2.82	36.12
	ALIKED	29.34	2.82	36.50
	DISK	28.27	2.79	36.77
UAV-Dataset (A), 9 imgs.	SIFT	22.99	2.98	32.06
	SURF	44.81	4.84	49.32
	MINEIGEN	35.51	2.97	48.28
	BRISK	31.71	5.48	47.11
	AKAZE	25.07	3.57	33.30
	SUPERPOINT	25.06	3.78	32.54
	ALIKED	28.74	3.35	33.43
	DISK	31.51	8.50	47.01
UAV-Dataset (B), 17 imgs.	SIFT	40.42	3.28	37.03
	SURF	38.85	3.36	43.52
	MINEIGEN	40.09	8.53	45.82
	BRISK	37.75	3.20	46.78
	AKAZE	37.62	9.96	44.65
	SUPERPOINT	55.86	5.91	52.21
	ALIKED	47.05	4.31	39.66
	DISK	58.85	7.51	45.36

V. CONCLUSION

This study presents a comprehensive evaluation of feature detectors and descriptors for UAV image stitching under perspective distortion, leveraging the REW method integrated with both traditional and deep learning-based approaches. Results across multiple datasets demonstrate that traditional detectors—such as SIFT, SURF, MINEIGEN, BRISK, and AKAZE—typically yield high keypoint counts and strong matching accuracy. However, they often struggle with geometric distortions, leading to artifacts such as ghosting, warping, and misalignment, especially in complex urban and aerial UAV scenes. Deep learning-based detectors—namely SUPERPOINT, ALIKED, and DISK—tend to extract fewer but more distinctive and robust keypoints, resulting in improved alignment accuracy and visual coherence. DISK shows strong spatial consistency, while ALIKED performs well in multi-image stitching scenarios with smoother transitions and fewer artifacts. Perceptual quality metrics (PIQE, NIQE, BRISQUE) further support these findings, with deep learning-based methods achieving lower scores that reflect higher perceptual quality. However, these methods are not universally robust. In cases involving radiometric distortions, low-texture environments, or extreme geometric transformations, even deep learning-based approaches fail to produce coherent mosaics—leading to structural misalignments and visible artifacts, as illustrated in Fig. 19. Table VII summarizes the primary sources of error and the corresponding global

correction strategies applied across the evaluated datasets for failed cases, providing insight into how each distortion type is addressed in practice. The images shown in Figs. 19 and 20 illustrate these error-contributing factors—such as extreme geometric transformations and cloud presence—which led to noticeable artifacts and degraded visual quality in the resulting mosaics.

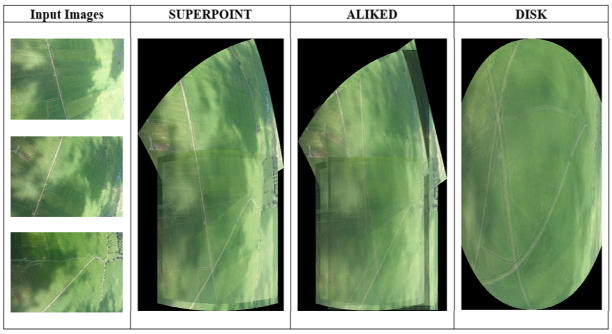


Fig. 19. Failed cases using deep learning-based feature detectors and descriptors, caused by low-texture regions, radiometric distortions, and extreme perspective transformations in the input images.

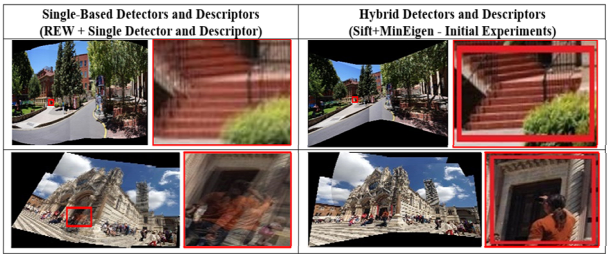


Fig. 20. Preliminary results of image stitching using hybrid feature detectors and descriptors.

TABLE VII. SUMMARY OF DATASET ERROR SOURCES AND CORRESPONDING CORRECTION STRATEGIES

Dataset	Image Source	Error Contribute	Global Correction Strategy
Fig. 19-Our UAV dataset	UAV-based	Vertical (nadir) view, extrema geometric transformation cloud presence	Camera calibration (resource-intensive), feature-based geometric correction methods, statistical and model-based radiometric correction methods
Fig. 20-General Dataset (A)-Left side	General Dataset (Outdoor)	Non-uniform horizontal view, extreme geometric transformation	Camera calibration (resource-intensive), feature-based geometric correction
Fig. 20-General Dataset (B)-Right side	General Dataset (Outdoor)	Non-uniform horizontal view, extreme geometric transformation	Camera calibration (resource-intensive), feature-based geometric correction

VI. FUTURE WORKS

In future work, exploring hybrid feature detectors and descriptors that integrate the strengths of both traditional and deep learning-based methods could further enhance stitching quality in images affected by perspective distortions. By leveraging the robustness and efficiency of traditional detectors alongside the adaptability and precision of deep learning approaches, a hybrid model may offer improved alignment and visual coherence as part of our proposed feature-based geometric correction technique—even in complex UAV imagery with challenging perspectives, as demonstrated in our preliminary experiments shown in Fig. 19. Table VIII presents the average number of matched features across different combinations of detectors and descriptors, computed using Eq. (8), where $|M_{\text{total}}|$ represents the total number of keypoint matches for the i^{th} image pair, and n denotes the total number of image pairs. The summation aggregates match across all image pairs, and the result is normalized by n to obtain the average number of matches per pair. Results indicate that hybrid pairings of traditional methods (e.g., SIFT+MinEigen, SURF+KAZE) yield the highest match counts, suggesting strong compatibility in terms of keypoint detection and descriptor robustness. Combinations involving deep learning-based detectors (e.g., ALIKED, SUPERPOINT) result in lower average matches, but may still offer advantages in terms of precision or alignment accuracy.

$$\text{Average Matcher} = \frac{\sum_{i=1}^{n} |M_{total}|_i}{n} \tag{8}$$

TABLE VIII. SUMMARY OF COMPATIBLE FEATURE DETECTOR AND DESCRIPTOR WITH PROMISING RESULTS

No.	Combined Features	Average Matcher
1	SIFT+MinEigen	312.89
2	SURF+KAZE	305.33
3	KAZE+MinEigen	292.22
4	ALIKED+SIFT	288.56
5	BRISK+KAZE	277.11
6	SUPERPOINT+SIFT	237.11
7	ALIKED+MinEigen	197.02

ETHICAL STATEMENT

UAV (drone) images used in this study were obtained from the Department of Agriculture, Philippines, via a formal request under Executive Order No. 2, s. 2016, or the Freedom of Information (FOI) program. Data collection complied with national regulations on aerial surveillance and was authorized for research use only, as granted by the department.

DATA AVAILABILITY

The UAV imagery data used in this study were obtained through a formal FOI request from the Department of Agriculture, Philippines. Access to the data is restricted and may be provided upon reasonable request, subject to approval by the data-owning agency.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

M. B. P. conducted the research under the supervision of N. M. M. B. P. analyzed the experimental results and wrote the manuscript, which was finalized by N. M. All authors approved the final version of the paper.

ACKNOWLEDGMENT

The authors would like to extend their gratitude to the CCIS Network Team at Caraga State University for providing access to their high-performance computing resources, which were essential for the development and testing of the software. Special thanks are also due to the Department of Agriculture, Manila, Philippines, for generously lending their UAV-acquired images, which significantly contributed to the success of this research.

REFERENCES

[1] H. Mokayed, P. Shivakumara, H. H. Woon, M. Kankanhalli, T. Lu, and U. Pal, "A new DCT-PCM method for license plate number detection in drone images," *Pattern Recognition Letters*, vol. 148, pp. 45–53, 2021.

[2] R. de Lima, A. A. Cabrera-Ponce, and J. Martinez-Carranza, "Parallel hashing-based matching for real-time aerial image mosaicing," *Journal of Real-Time Image Processing*, vol. 18, pp. 143–156, 2021.

[3] J. Shen, "Image stitching quality evaluation and improvement based on SIFT features and RANSAC algorithm," in *Proc. 2024 2nd Int. Conf. on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, 2024, pp. 755–766.

[4] M. P. B. Pacot, and N. Marcos, "PBIM: A patch-based image mosaicking for UAV images with natural scene distortions," in *Proc. 2023 IEEE Asia-Pacific Conf. on Geoscience, Electronics and Remote Sensing Technology (AGERS)*, 2023, pp. 91–97.

[5] A. Li, J. Guo, and Y. Guo, "Image stitching based on semantic planar region consensus," *IEEE Transactions on Image Processing*, vol. 30, pp. 5545–5558, 2021.

[6] X. Yang, Z. Y. Liu, H. Qiao, J. H. Su, D. X. Ji, A. Y. Zang, and H. Huang, "Graph-based registration and blending for undersea image stitching," *Robotica*, vol. 38, no. 3, pp. 396–409, 2020.

[7] N. Li, T. Liao, and C. Wang, "Perception-based seam cutting for image stitching," *Signal, Image and Video Processing*, vol. 12, pp. 967–974, 2018.

[8] Y. Wang, Q. Cong, S. Yao, X. Jia, J. Chen, and S. Li, "Research on geometric error correction of pushbroom hyperspectral camera carried by UAV," *Seventh Symp. on Novel Photoelectronic Detection Technology and Applications*, vol. 11763, pp. 1214–1220, 2021.

[9] P. V. Arun, I. Herrmann, K. M. Budhiraju, and A. Karnieli, "Convolutional network architectures for super-resolution/sub-pixel mapping of drone-derived images," *Pattern Recognition*, vol. 88, pp. 431–446, 2019.

[10] S. I. Deliry and U. Avdan, "Accuracy of unmanned aerial systems photogrammetry and structure from motion in surveying and mapping: A review," *Journal of the Indian Society of Remote Sensing*, vol. 49, no. 8, pp. 1997–2017, 2021.

[11] D. Popescu, L. Ichim, and F. Stoican, "Unmanned aerial vehicle systems for remote estimation of flooded areas based on complex image processing," *Sensors*, vol. 17, no. 3, p. 446, 2017.

[12] D. G. Lowe, "SIFT—The scale-invariant feature transform," *International Journal of Computer Vision*, vol. 2, no. 2, pp. 91–110, 2004.

[13] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in *Proc. 9th European Conf. on Computer Vision (ECCV)*, 2006, pp. 404–417.

[14] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. 2011 International Conference on Computer Vision (ICCV)*, 2011, pp. 2548–2555.

[15] P. F. Alcantarilla and T. Solutions, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1281–1298, 2011.

[16] M. Tyszkiewicz, P. Fua, and E. Trulls, "DISK: Learning local features with policy gradient," *Advances in Neural Information Processing Systems*, vol. 33, pp. 14254–14265, 2020.

[17] X. Zhao, X. Wu, W. Chen, P. C. Chen, Q. Xu, and Z. Li, "ALIKEd: A lighter keypoint and descriptor extraction network via deformable transformation," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–16, 2023.

[18] J. Zaragoza, T. J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2339–2346, 2013.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv Preprint, arXiv:1409.1556, 2014.

[21] Q. Jia, Z. Li, X. Fan, H. Zhao, S. Teng, X. Ye, and L. J. Latecki, "Leveraging line-point consistence to preserve structures for wide parallax image stitching," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 12186–12195.

[22] M. Tyszkiewicz, P. Fua, and E. Trulls, "DISK: Learning local features with policy gradient," *Advances in Neural Information Processing Systems*, vol. 33, pp. 14254–14265, 2020.

[23] Y. Ono, E. Trulls, P. Fua, and K. M. Yi, "LF-Net: Learning local features from images," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[24] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-supervised interest point detection and description," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 224–236.

[25] J. Li, Z. Wang, S. Lai, Y. Zhai, and M. Zhang, "Parallax-tolerant image stitching based on robust elastic warping," *IEEE Transactions on Multimedia*, vol. 20, no. 7, pp. 1672–1687, 2017.

[26] C. H. Chang, Y. Sato, and Y. Y. Chuang, "Shape-preserving half-projective warps for image stitching," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 3254–3261.

[27] M. Guo, M. Guo, J. Su, J. Chen, J. Yu, J. Wang, and F. Jin, "Bayesian iterative prediction and lexical-based interpretation for disturbed Chinese sentence pair matching," in *Proc. ACM Web Conf.*, 2024, pp. 4618–4629.

[28] J. Yu and F. Da, "Projection model-driven image stitching: A novel warping method using epipolar displacement field," *Machine Vision and Applications*, vol. 36, no. 3, p. 66, 2025.

[29] Q. Huang, X. Guo, Y. Wang, H. Sun, and L. Yang, "A survey of feature matching methods," *IET Image Processing*, vol. 18, no. 6, pp. 1385–1410, 2024.

[30] S. M. Javidan, A. Banakar, K. Rahnama, K. A. Vakilian, and Y. Ampatzidis, "Feature engineering to identify plant diseases using image processing and artificial intelligence: A comprehensive review," *Smart Agricultural Technology*, 100480, 2024.

[31] Y. S. Chen and Y. Y. Chuang, "Natural image stitching with the global similarity prior," in *Proc. European Conf. on Computer Vision (ECCV)*, 2016, pp. 186–201.