Advances in Artificial Intelligence and Machine Learning; Research 5 (1) 3409-3424 Received 11-12-2024; Accepted 16-03-2025; Published 23-03-2025

AI-Based Aerial Camera Calibration and 3D Reconstruction Accuracy Evaluation

Jittiphan Changkaew jittiphan_ch@yahoo.com
The Graduate School,
Navaminda Kasatriyadhiraj Royal Air Force Academy, Bangkok,
Thailand.

Prasartporn Wongkamchang wongkamchang@hotmail.com
The Graduate School,
Navaminda Kasatriyadhiraj Royal Air Force Academy, Bangkok,
Thailand.

Chamnan Pedchote c_pedchote@hotmail.com
The Graduate School,
Navaminda Kasatriyadhiraj Royal Air Force Academy, Bangkok,
Thailand.

Khongdet Phasinam phasinam@gmail.com
Shinawatra University,
Pathum Thani,
Thailand.

Corresponding Author: Khongdet Phasinam

Abstract

Accurate camera calibration is a cornerstone of aerial imaging, essential for precise 3D reconstruction, mapping, and motion estimation. Traditional calibration methods often depend on predefined objects and periodic recalibration, which are impractical in dynamic aerial environments. This study investigates the potential of AI-based calibration methods, specifically GeoCalib, CTRL-C, and DeepCalib, to address these challenges. Using the ISPRS Vaihingen dataset, evaluate these methods against the conventional approach. The research focuses on intrinsic parameter estimation and its impact on 3D reconstruction accuracy. Our findings reveal that CTRL-C achieved the highest precision, with a mean reconstruction error of $1.59e10^{-5}$, significantly outperforming GeoCalib (1.0549) and DeepCalib (0.2110). Additionally, DeepCalib demonstrated strong performance in minimizing Chamfer Distance (0.4220) and Hausdorff Distance (0.2502), while GeoCalib exhibited broader error distributions. These results underscore the superior capability of AI-based techniques in delivering accurate and reliable calibration for aerial imaging systems.

Keywords: Camera calibration, Deep learning, Aerial photography, Artificial intelligence

Citation: Jittiphan Changkaew, et al. AI-Based Aerial Camera Calibration and 3D Reconstruction Accuracy Evaluation. Advances in Artificial Intelligence and Machine Learning. 2025;5(1):195.

1. INTRODUCTION

In computer vision and image processing, camera calibration is an essential procedure. It plays a vital role in establishing the relationship between 2D image points and their corresponding 3D world coordinates [1]. However, camera calibration faces several significant challenges, particularly in aerial photography. One primary issue is that traditional calibration methods often rely on specific calibration objects, such as checkerboards, and require periodic recalibration [2]. This approach is ill-suited for systems operating continuously or in dynamically changing environments, such as aerial photography. Furthermore, radial and tangential lens distortions [3, 4] pose significant challenges in obtaining accurate and precise images, especially for aerial cameras operating in diverse environmental conditions. Moreover, achieving a balance between accuracy and processing speed remains a critical problem, particularly for real-time systems. Time-consuming calibration procedures may be impractical for real-world applications such as autonomous vehicles or aerial navigation systems, which demand rapid and accurate responses. Artificial Intelligence (AI), especially Deep Learning, has shown promise in recent years as a remedy for these problems [5, 6]. The application of AI in calibration not only addresses the limitations of traditional methods but also enables systems to continuously adapt to changing environmental conditions.

This paper presents the concepts and techniques in camera calibration using AI. We cover the fundamentals of camera calibration and explore the application of Deep Learning techniques in developing automated calibration systems. Additionally, we analyze methods for addressing the problems as mentioned earlier, present case studies, and discuss future development trends.

2. FUNDAMENTALS OF CAMERA CALIBRATION

2.1 Camera Model and Parameters

The central projection of points in 3D space onto a 2D plane is described by the pinhole camera model [7], as seen in FIGURE 1. The origin of a Euclidean coordinate system serves as the center of projection in this approach. The image plane, the focal plane, is defined as z = f, where f is the focal length. A location on the image plane where a line linking X to the center of projection intersects the image plane is mapped to a point X in 3D space with coordinates $(X, Y, Z)^T$. Using similar triangles, we can derive that the 3D point $(X, Y, Z)^T$ is mapped to the 2D point $(fX/Z, fY/Z)^T$ on the image plane. This mapping can be expressed mathematically as:

$$(X, Y, Z)^{T} \mapsto (fX/Z, fY/Z)^{T}$$

This equation describes the central projection mapping from world coordinates to image coordinates, transforming points from Euclidean 3-space R^3 to Euclidean 2-space R^2.

The camera center, sometimes referred to as the optical center, is the center of projection in the pinhole camera paradigm. The primary axis or principal ray of the camera is a line that extends from this camera center and intersects the image plane at a straight angle. The principal point is the intersection of this principal axis and the picture plane. Furthermore, the major plane of the camera is a plane that runs parallel to the image plane and goes through the center of the camera.

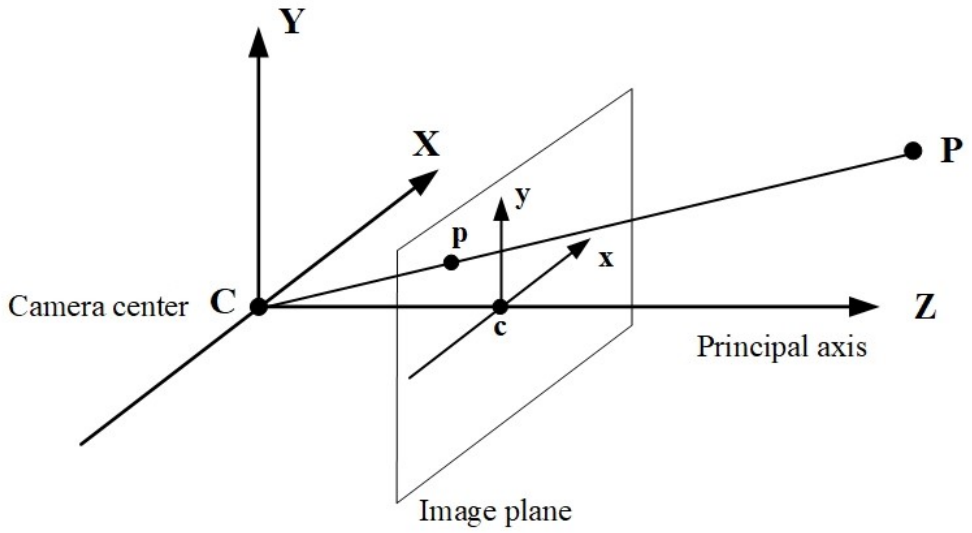


Figure 1: Geometry of a pinhole camera

The pinhole camera model is a fundamental basis for understanding camera geometry in computer vision and photogrammetry. This model explains mathematically how a perfect pinhole camera projects 3D points from the outside world onto a 2D image plane. Camera parameters define the characteristics and behavior of a camera, determining how it captures and projects 3D scenes onto a 2D image plane. These parameters are essential in computer vision and photogrammetry for accurately mapping points from the 3D world to their corresponding 2D image points. Camera parameters are typically categorized into two main groups: intrinsic and extrinsic parameters.

These two sets of characteristics are determined by camera calibration [7]. The focal length, image center, and lens distortion coefficients are examples of intrinsic parameters. Extrinsic parameters comprise rotation matrices and translation vectors. These parameters are crucial for transforming 3D world coordinates into 2D image points [8], enabling accurate spatial analysis and reconstruction in various imaging applications.

2.1.1 Intrinsic parameters

The camera's internal optical properties are described by intrinsic parameters. Among them are:

1. Focal Length (f): This represents the distance between the camera center and the image plane. In practice, we often use two focal lengths, f_x and f_y, to account for non-square pixels [1].
2. Principal Point (c_x, c_y): The intersection of the optical axis with the image plane, typically close to the image center [8].
3. Skew Coefficient: Explains the angle formed by the pixel axes x and y. In most modern cameras, this is zero [9].

These parameters form the camera intrinsic matrix K:

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Where s is the skew coefficient.

2.1.2 Extrinsic parameters

In the global coordinate system, the camera's orientation and position are determined by external factors. They include:

1. Rotation Matrix (R): A 3x3 matrix describing the camera's orientation [10].
2. Translation Vector (t): A 3x1 vector describing the camera's position [10].

The extrinsic parameters can be combined into a 3x4 matrix [R|t].

2.1.3 Projection model

The complete projection of a 3D point X = (X, Y, Z, 1)^T to a 2D image point x = (u, v, 1)^T can be described by:

$$x = K(R|t)X$$

This equation encapsulates the camera model, relating 3D world coordinates to 2D image coordinates [11].

2.1.4 Lens distortion

Real cameras deviate from the ideal pinhole model due to lens distortion. Two main types of distortion are typically considered:

1. Radial Distortion: Caused by the shape of the lens, it increases with distance from the optical center. It's usually modeled using polynomial coefficients (k_1, k_2, k_3) [3].
2. Manufacturing flaws that result in the lens not being precisely parallel to the image plane produce tangential distortion. It's modeled using parameters p_1 and p_2 [4].

The distortion model is applied after the projection and before scaling and offset by the intrinsic parameters.

2.1.5 Importance in calibration

Camera calibration aims to estimate these intrinsic and extrinsic parameters accurately. For aerial photography, precise calibration is crucial for 3D reconstruction, motion estimation, and mapping [2]. The challenge lies in accurately estimating these parameters, especially in dynamic environments or when the camera's internal characteristics may change due to temperature or vibration [12].

In Coordinate Transformation and Perspective Projection, as shown in FIGURE 2, both extrinsic and intrinsic parameters are rewritten by using homogeneous coordinates as follows:

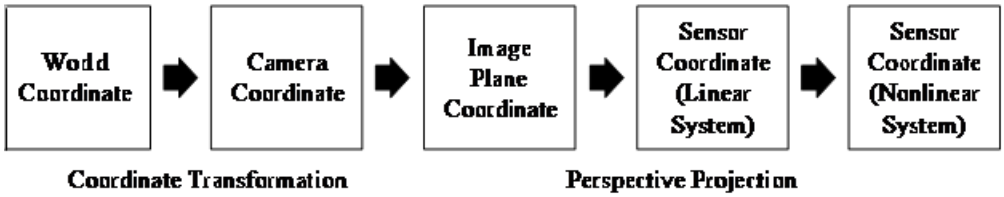


Figure 2: The Coordinated for the Camera Transform

Extrinsic Matrix: $M_{ext} = \begin{bmatrix} R_{3\times3} & t \\ 0_{1\times3} & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Intrinsic Matrix: $M_{int} = (K|0) = \begin{bmatrix} f_x & s & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Combining the above two matrixes is imperative to acquire the essential projection matrix used in camera calibration.

$$x_{image} = M_{int} M_{ext} X_{world} = P X_{world}$$

2.2 3D Reconstruction and Accuracy Metrics

Using photogrammetry and stereo-vision techniques, the multi-camera system's photos are processed to create a 3D point cloud of the scene. In order to project 3D points into the world coordinate system, each camera's intrinsic and extrinsic parameters are estimated during the camera calibration procedure [13]. We use the following metrics to evaluate the reconstructed 3D point cloud's accuracy.

2.2.1 Chamfer distance

The average point distance between the ground truth and the reconstructed point cloud is measured by the Chamfer Distance. It finds the nearest points in one point cloud to each point in the other

point cloud and computes the average distance. Smaller Chamfer Distance indicates more accurate 3D reconstruction.

$$D_{Chamfer}(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|^2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|^2$$

Where : S_1 and S_2 are two sets of 3D reconstruction and ground truth points. x and y are individual points in the point clouds.

2.2.2 Hausdorff distance

A point in one set and the closer point in another set are separated by the Hausdorff Distance, which is the largest of all distances. A smaller Hausdorff distance means that the two sets of reconstructed and ground truth points are closer to each other.

$$D_{Hausdorff}(S_1, S_2) = \max \left(\sup_{x \in S_1} \inf_{y \in S_2} \|x - y\|, \sup_{y \in S_2} \inf_{x \in S_1} \|y - x\| \right)$$

2.2.3 Reprojection error

The difference between the 2D projection of the reconstructed 3D points onto the picture plane and the observed 2D image points is measured by the Reprojection Error. Quantifies how precisely the 3D points that have been reconstructed are projected back onto the 2D picture planes [14, 15].

$$\text{Reprojection Error} = \frac{1}{n} \sum_{i=1}^{n} \|x_i - PX_i\|$$

Where: x_i are the observed 2D points P is the camera projection matrix. X_i are the reconstructed 3D points. $\| \cdot \|$ denotes the Euclidean distance.

2.2.4 Point-to-point errors or Euclidean distance

The Euclidean Distance compares the reconstructed 3D points directly with the ground truth 3D points. Higher accuracy is shown by a smaller average Euclidean distance between the reconstructed and ground truth points.

$$\text{Point-to-Point} = \frac{1}{n} \sum_{i=1}^{n} \left\| X^i_{reconstruction} - X^i_{ground\,truth} \right\|$$

Where : $X^i_{reconstructed}$ are the 3D points reconstructed by the algorithm. $X^i_{ground\,truth}$ are the corresponding ground truth 3D points.

3. AI TECHNIQUES FOR CAMERA CALIBRATION

3.1 Convolutional Neural Networks (CNNs)

With notable gains in speed, accuracy, and versatility over conventional techniques, Convolutional Neural Networks (CNNs) have become a potent instrument in the field of camera calibration. These networks can learn image features and use this information to accurately estimate intrinsic and extrinsic camera parameters [16]. They are particularly well-suited for this task because they can automatically learn relevant features from input images, reducing the need for manual feature engineering.

3.2 CTRL-C Algorithm

The CTRL-C algorithm [17] is a state-of-the-art calibration algorithm that combines deep learning techniques with geometric understanding to achieve robust and accurate camera calibration from a single image, outperforming previous methods in challenging real-world scenarios.

3.3 DeepCalib Algorithm

DeepCalib [19] is an algorithm for automatic intrinsic calibration of wide field-of-view cameras using deep learning. The algorithm demonstrates how deep learning can be applied to the challenging task of intrinsic calibration for wide-angle cameras. Learning from a large dataset of synthetically distorted images can generalize to real-world wide-angle cameras without requiring specific calibration patterns or multiple views. This makes it particularly useful for applications where traditional calibration methods are impractical or time-consuming.

3.4 GeoCalib

GeoCalib [20] is a new approach to single-image camera calibration that combines deep learning with traditional geometric optimization. Enhancing the ability to estimate camera parameters from a single image, such as focal length and gravity direction, is the aim. It is valuable for 3D mapping, visual localization, and augmented reality applications. Whether based purely on geometry or deep learning, they have limitations: geometry-based methods are accurate but struggle with scenes lacking clear lines or vanishing points. In contrast, deep learning approaches are more robust but often lack accuracy and generalizability.

4. EXPERIMENTS AND RESULTS

4.1 Experimental System Installation

For our experiment, the ISPRS Vaihingen dataset, as shown FIGURE 3, was selected, and evaluate the effectiveness of camera calibration using AI by comparing and analyzing its results against those obtained from Tsai's two-stage method, which is currently the most widely adopted approach for explicit camera calibration.

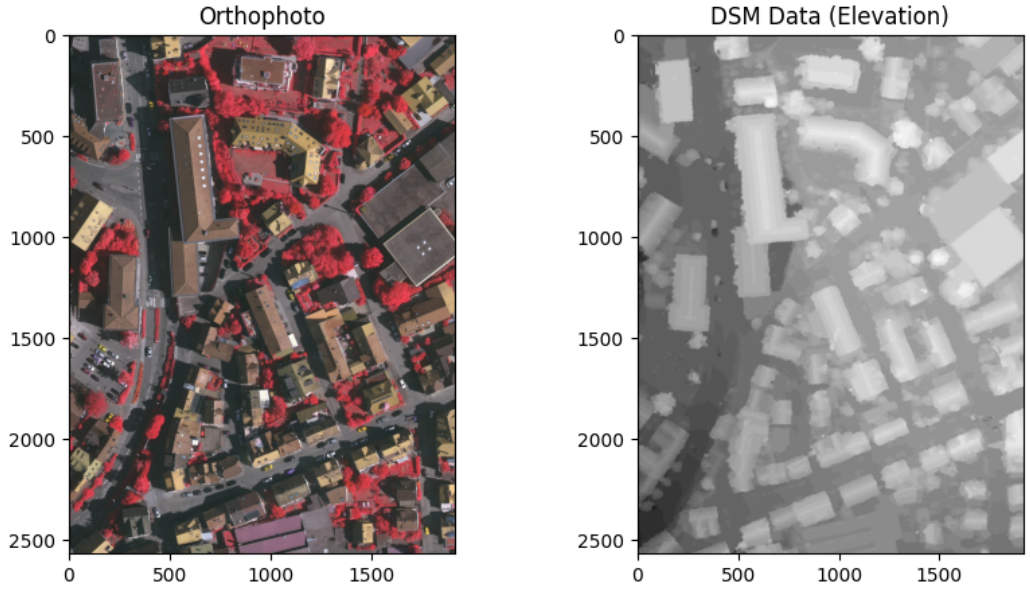


Figure 3: The ISPRS Vaihingen dataset

The performance of different camera calibration techniques is assessed in this research by comparing the average inaccuracy between the calibrated image coordinates and the corresponding real-world coordinates. The metrics we use for this comparison are the Chamfer Distance, Hausdorff Distance, Euclidean Distance, and Mean Reprojection Error.

4.2 AI-Based Camera Calibration

4.2.1 GeoCalib

GeoCalib is an AI-driven calibration approach that combines deep learning with geometric optimization to estimate intrinsic camera parameters effectively. This method leverages its ability to process single images and outputs parameters tailored for accurate 3D mapping and visual localization applications. The results of the GeoCalib calibration process are as follows in FIGURE 4, and the camera parameter in TABLE 1.

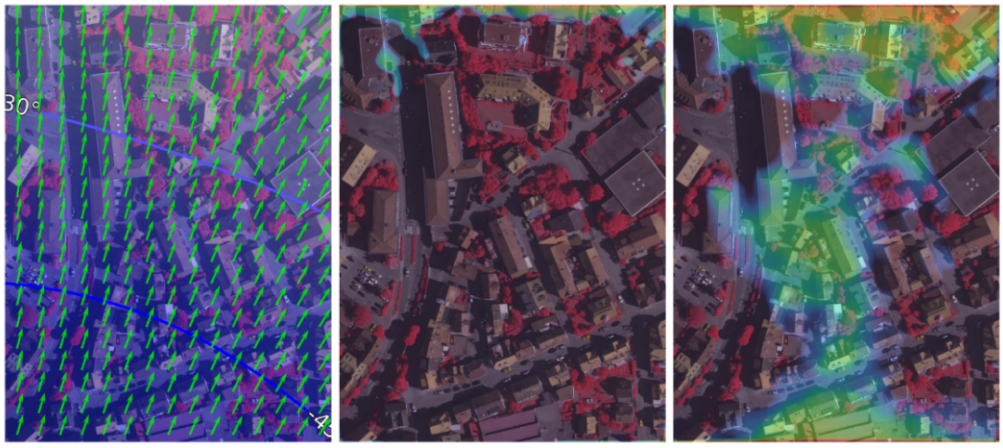


Figure 4: The Output of the GeoCalib Calibration Method

Table 1: The Camera Parameter from AI-Based Calibration Method

Models	Intrinsic				
	fx	fy	Cx	Cy	dist coef.
Geocalib	2851.097	3816.815	959.5	1284.5	[1, 0, 0, 0, 0]
DeepCalib	2438.863	3264.95	959.5	1284.5	[0.2, 0, 0, 0, 0]
CTRL-C	1358.736	1818.964	218.1047	2517.546	[0, 0, 0, 0, 0]
OpenCV	2952.650	2680.027	2377.6412	2486.843	[0.1, -0.01, 0.07, 0.07, 0]

4.2.2 CTRL-C

CTRL-C integrates deep learning with geometric constraints, achieving exceptional precision in camera parameter estimation. Unlike traditional methods, CTRL-C can perform robust calibration using a single image, making it highly suitable for dynamic aerial environments. The results of the CTRL-C calibration process are as follows in FIGURE 5 and the camera parameter in TABLE 1.

4.2.3 DeepCalib

DeepCalib is an AI-driven calibration method designed for intrinsic calibration of wide field-of-view cameras. By utilizing deep learning, it eliminates the dependency on traditional calibration patterns and adapts effectively to complex real-world scenarios. This approach makes DeepCalib particularly advantageous for aerial imaging systems operating in dynamic environments. The results of the DeepCalib calibration process are as follows in FIGURE 6 and the camera parameter in TABLE 1.

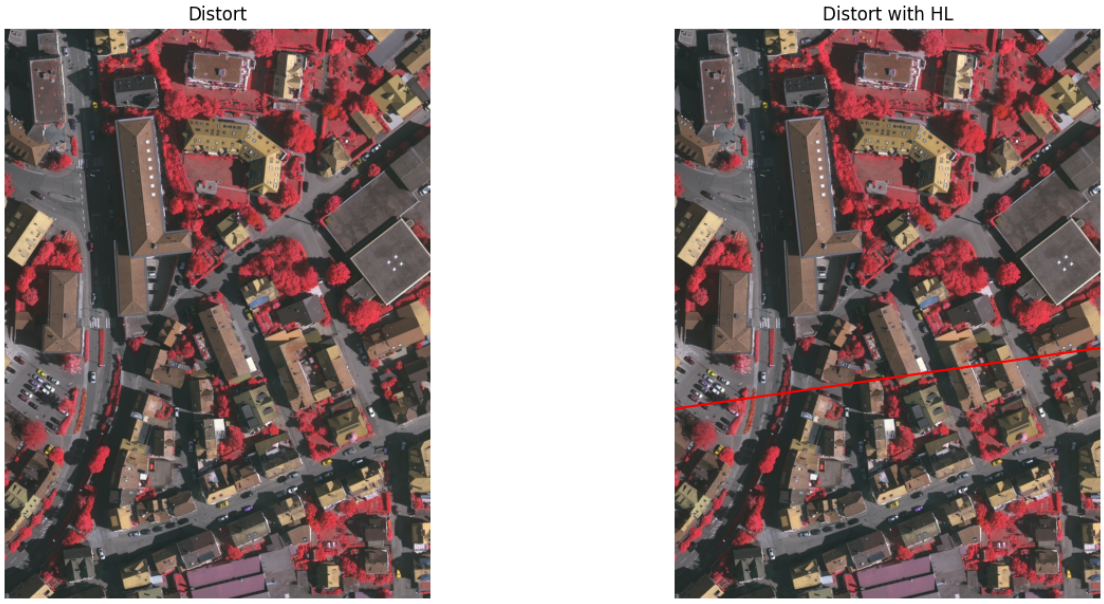


Figure 5: The Output of the GeoCalib Calibration Method

4.3 3D Reconstruction Accuracy Evaluation

The flowchart as shown in FIGURE 7, illustrates a process for evaluating 3D reconstruction accuracy using orthophoto imagery and Digital Surface Model (DSM) data. It begins with loading orthophoto images and DSM files, which serve as the input data for subsequent steps. The DSM data is then normalized for visualization, preparing it for further processing. Pixels from the orthophoto are converted into corresponding 3D world coordinates using the DSM data and transformation parameters. These 3D world coordinates are subsequently projected onto a 2D image plane using camera calibration parameters to simulate how the 3D points would appear in the image shown in FIGURE 8-10. Using the projected 2D image points, 3D world points are reconstructed based on depth values and the given camera parameters. Reconstruction errors, including the mean error, are calculated by comparing the original and rebuilt 3D points in order to measure the reconstruction's accuracy. Finally, the process visualizes the original and reconstructed 3D world points using a 3D scatter plot for a clear comparison shown in FIGURE 11. This visual representation aids in assessing how closely the reconstructed points match the ground truth, providing a comprehensive evaluation of the 3D reconstruction process.

5. CONCLUSION

This study demonstrates the transformative potential of AI-based camera calibration methods in addressing the limitations of traditional techniques for aerial imaging applications. By leveraging advanced algorithms such as GeoCalib, CTRL-C, and DeepCalib. The experimental results, using the ISPRS Vaihingen dataset, revealed that CTRL-C delivered the most precise calibration. However, while AI-based calibration methods demonstrated superior performance in our experiments,

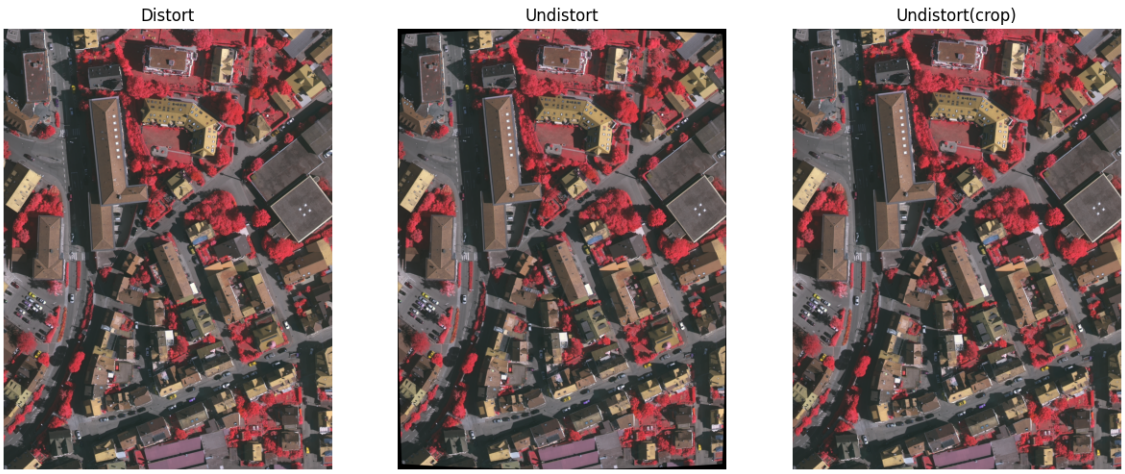


Figure 6: The Output of the DeepCalib Calibration Method

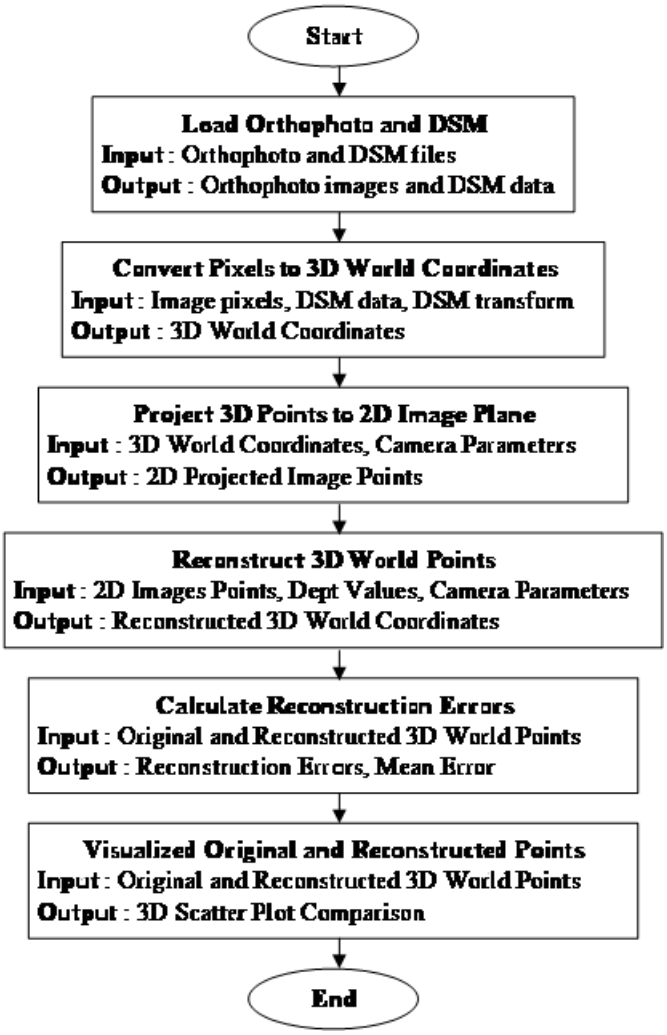


Figure 7: The Flowchart for 3D Reconstruction Accuracy Evaluation

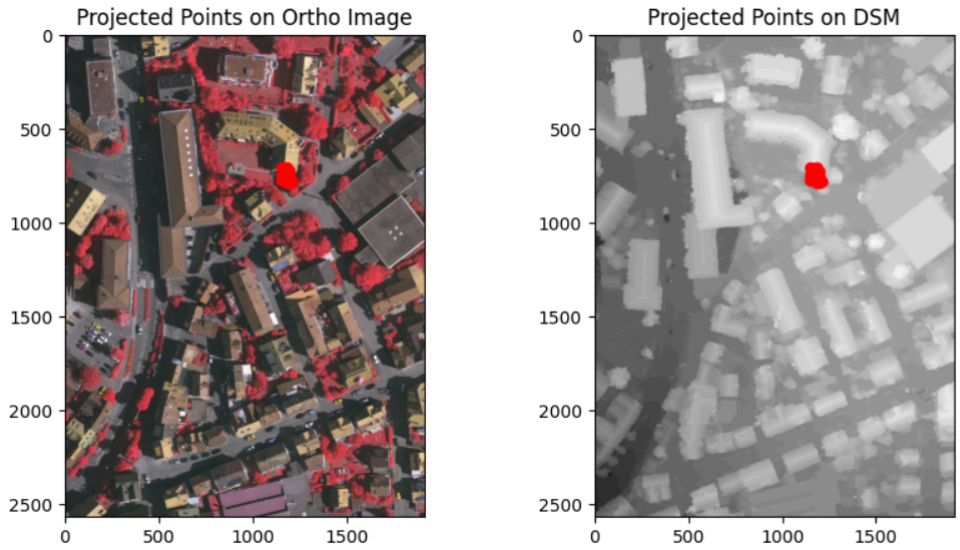


Figure 8: The Projected point from Camera Parameter Calibration using GeoCalib

Table 2: Comparison of 3D Reconstruction Accuracy Across Methods

Point No.	3D World Points			Reconstruction Error by Method			
	x	y	z	Conventional	GeoCalib	DeepCalib	CTRL-C
1	20.8931	-35.3078	275.0902	0.6893	0.9125	0.1825	7.86e-06
2	22.2318	-35.2495	274.6358	0.6881	0.9596	0.1919	9.54e-06
3	22.6658	-37.3656	279.9971	0.7557	1.0647	0.2129	2.36e-05
4	23.9325	-36.0450	278.4471	0.7117	1.0447	0.2089	1.16e-05
5	22.2917	-38.1258	275.3761	0.7984	1.1359	0.2272	2.36e-05
6	20.9197	-35.9311	275.4399	0.7125	0.9473	0.1895	1.91e-06
7	17.9687	-35.6786	272.4594	0.7199	0.8587	0.1718	4.26e-06
8	18.5804	-39.0061	281.7344	0.8349	1.0161	0.2032	4.26e-06
9	20.4926	-40.0544	269.8176	0.9044	1.2510	0.2502	2.73e-05
10	21.5907	-38.2605	266.7164	0.8292	1.1919	0.2384	2.73e-05
11	20.7966	-37.9104	273.8202	0.7972	1.0783	0.2157	2.29e-05
12	18.4388	-41.5047	279.5869	0.9573	1.1984	0.2397	2.68e-05
Mean Reconstruction Error				0.7832	1.0549	0.2110	1.59e-05
Maximum Error				0.9573	1.2510	0.2502	2.73e-05
Standard Deviation					0.1201	0.0242	8.76e-06

several limitations and considerations should be acknowledged. For instance, AI-based methods, particularly deep learning approaches like CTRL-C and DeepCalib, require significant computational resources for training and inference. Additionally, their effectiveness is highly dependent on the quality and diversity of training datasets. When applied to scenes significantly different from their training data, these methods may exhibit degraded performance.

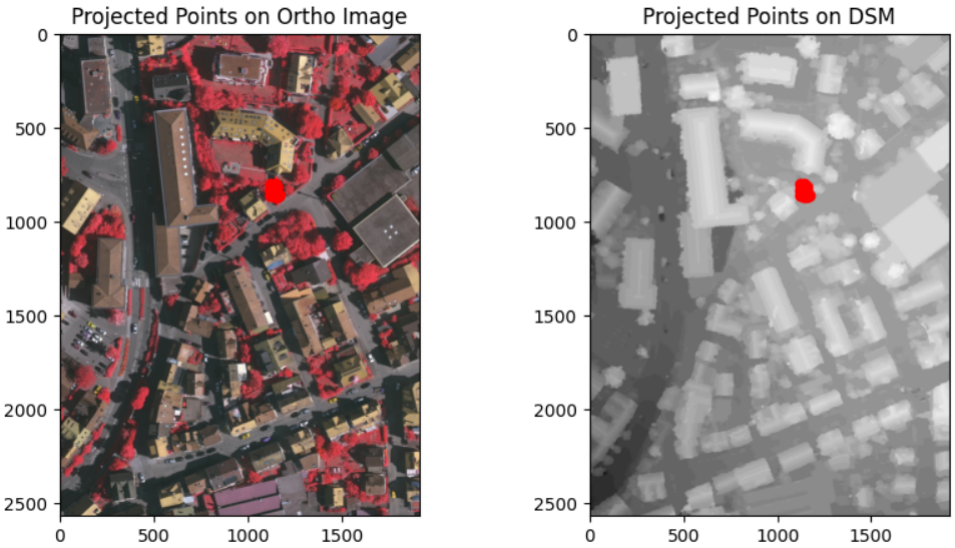


Figure 9: The Projected point from Camera Parameter Calibration using DeepCalib

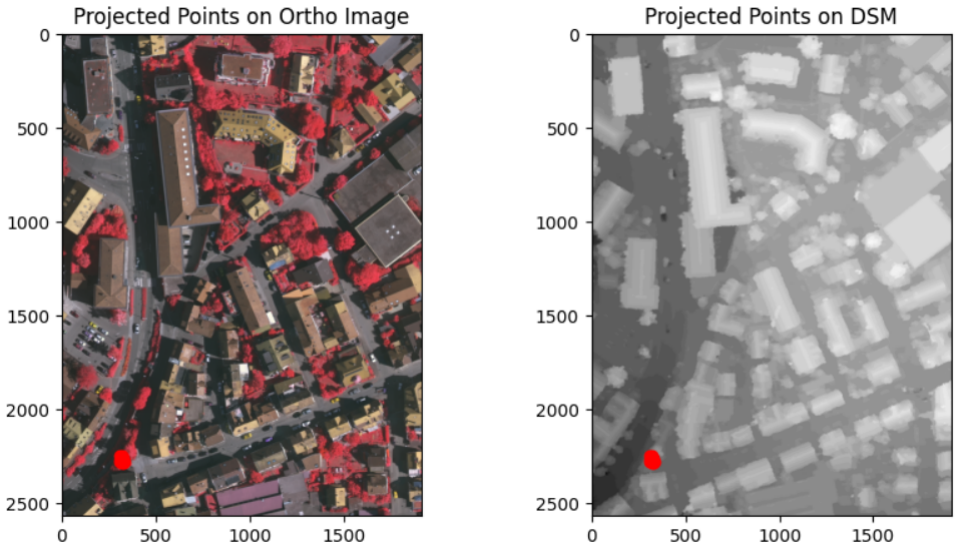
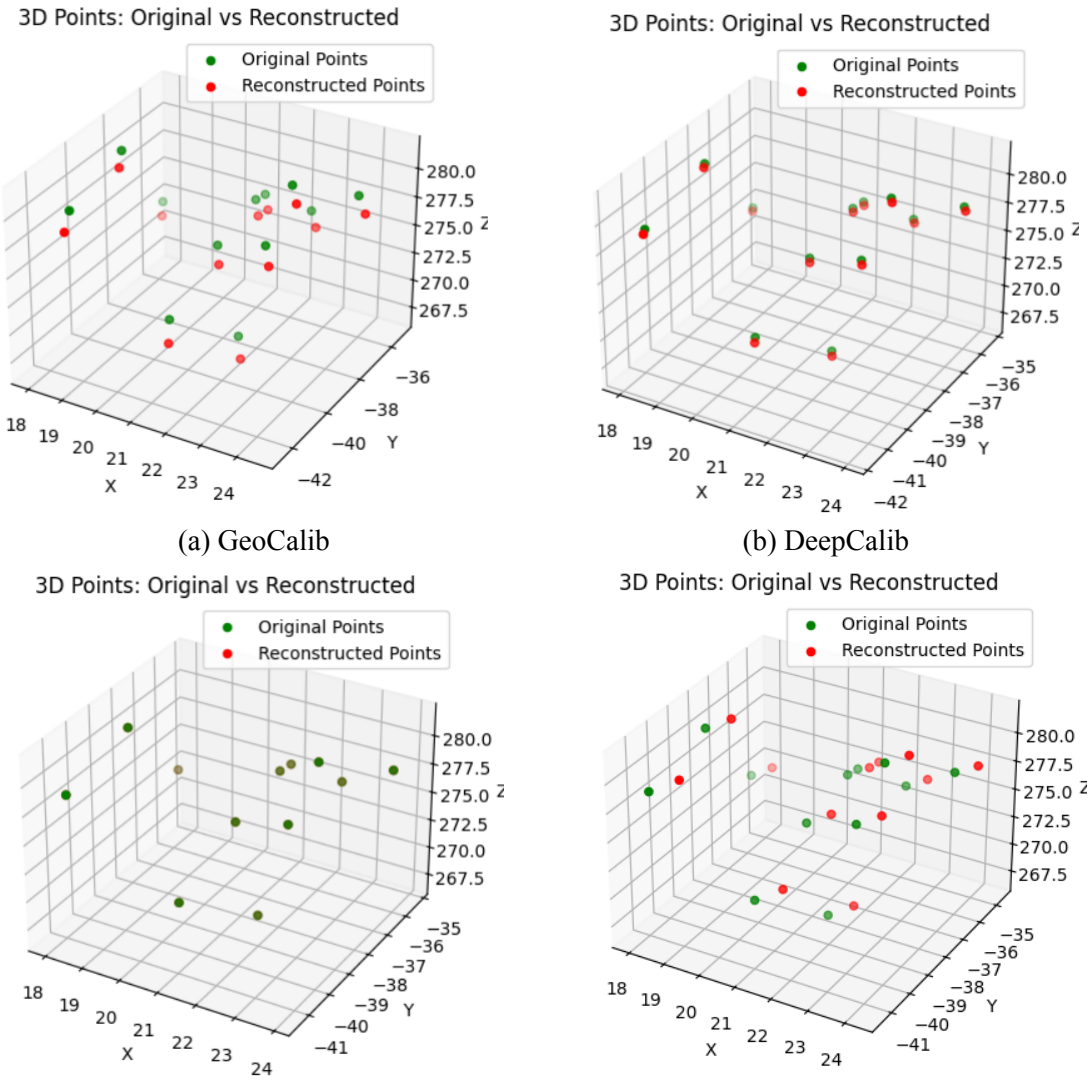


Figure 10: The Projected point from Camera Parameter Calibration using CTRL-C

Table 3: The 3D World Coordinate Reconstruction Performance Matrics

	Mean Reconstruction Error	Hausdorff Distance	Chamfer Distance	RMSE
GeoCalib	1.0549	1.2510	2.0521	1.0615
DeepCalib [18]	0.2110	0.2502	0.4220	0.2123
CTRL-C [16]	1.59e-05	2.7309e-05	3.1817e-05	1.865e-05
Conventional	0.7832	0.9573	1.5284	0.7876



(a) GeoCalib (b) DeepCalib

(c) CTRL-C (d) Conventional approach

Figure 11: The 3D plot for Reconstruction Errors for each point (a) GeoCalib (b) DeepCalib (c) CTRL-C and conventional approach

Traditional calibration methods offer clear geometric interpretations of their parameters, facilitating troubleshooting and manual adjustment. In contrast, AI-based methods may operate as "black boxes," making it difficult to diagnose issues or make targeted refinements to the calibration. Therefore, future research should explore hybrid approaches that combine the precision of AI-based methods with the interpretability and theoretical foundations of conventional techniques.

6. ACKNOWLEDGEMENT

The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) [Cramer, 2010]: http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html.

7. FUNDING

This research is supported by Graduate School , Navaminda Kasatriyadhiraj Royal Air Force Academy, Thailand.

8. TRANSPARENCY

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

9. COMPETING INTERESTS

The authors declare that they have no competing interests.

10. AUTHORS' CONTRIBUTIONS

All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

References

[1] Zhang Z. A Flexible New Technique for Camera Calibration. IEEE Trans Pattern Anal Mach Intell. 2000;22:1330-1334.

[2] Remondino F, Fraser C. Digital Camera Calibration Methods: Considerations and Comparisons. Int Arch Photogramm Remote Sens Spat Inf Sci. 2006;36:266-272.

[3] Brown DC. Decentering Distortion of Lenses. Photogramm Eng. 1966;32:444-462.

[4] Fryer JG, Brown DC. Lens Distortion for Close-Range Photogrammetry. Photogramm Eng Remote Sens. 1986;52:51-58.

[5] Ou Q, Xie Q, Chen F, Peng J, Xiong B. Reinforcement Learning Based Calibration Method for Cameras With Large Fov. Measurement. 2022;202:111732.

[6] Suewongsuwan K, Angsuseranee N, Wongkamchang P, Phasinam K. Comparative Analysis of UAV Detection and Tracking Performance: Evaluating YOLOV5, YOLOV8, and YOLOV8 Deepsort for Enhancing Anti-UAV Systems. Edelweiss Appl Sci Technol. 2024;8:708-726.

[7] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision. Cambridge University Press. 2003.

[8] Heikkila J, Silvén O. A Four-Step Camera Calibration Procedure With Implicit Image Correction. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition. IEEE Comput Soc. 1997:1106-1112.

[9] Tsai R. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-The-Shelf TV Cameras and Lenses. IEEE J Robot Automat. 1987;3:323-344.

[10] Ma Y, Soatto S, Kosecka J, Sastry SS. An Invitation to 3-D Vision: From Images to Geometric Models. New York: Springer. 2004.

[11] Faugeras O. Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT press. 1993.

[12] Dang T, Hoffmann C, Stiller C. Continuous Stereo Self-Calibration by Camera Parameter Tracking. IEEE Trans Image Process. 2009;18:1536-1550.

[13] Wang Q. Towards Real-Time 3D Terrain Reconstruction From Aerial Imagery. Geographies. 2024;4:66-82.

[14] Dai C, Han T, Luo Y, Wang M, Cai G, et al. NMC3D: Non-overlapping Multi-Camera Calibration Based on Sparse 3D Map. Sensors. 2024;24:5228.

[15] Wahbeh W, Müller G, Ammann M, Nebiker S. Automatic Image-Based 3D Reconstruction Strategies for High-Fidelity Urban Models—Comparison and Fusion of UAV and Mobile Mapping Imagery for Urban Design Studies. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2022;43:461-468.

[16] Hold-Geoffroy Y, Sunkavalli K, Eisenmann J, Fisher M, Gambaretto E, et al. A Perceptual Measure for Deep Single Image Camera Calibration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:2354-2363.

[17] Lee J, Go H, Lee H, Cho S, Sung M, et al. Ctrl-c: Camera Calibration Transformer With Line-Classification. In: International Conference on Computer Vision; 2021:16228-16237.

[18] Jau YY, Zhu R, Su H, Chandraker M. Deep Keypoint Based Camera Pose Estimation With Geometric Constraints. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020:4950-4957.

[19] Bogdan O, Eckstein V, Rameau F, Bazin JC. Deepcalib: A Deep Learning Approach for Automatic Intrinsic Calibration of Wide Field-Of-View Cameras. In: Proceedings of the 15th ACM SIGGRAPH European conference on visual media production. ACM; 2018:1-10.

[20] Veicht A, Sarlin PE, Lindenberger P, Pollefeys M 2024. Geocalib: Learning Singleimage Calibration With Geometric Optimization. 2024. ArXiv preprint: https://arxiv.org/pdf/2409.06704v1.