

Article

SIFT-CNN Pipeline in Livestock Management: A Drone Image Stitching Algorithm

Housseem R. E. H. Bouchekara ^{1,*}, Bashir O Sadiq ^{2,3} , Sikiru O Zakariyya ⁴ , Yusuf A. Sha'aban ^{1,3} ,
Mohammad S. Shahriar ¹ and Musab M. Isah ⁵ 

- ¹ Department of Electrical Engineering, University of Hafr Al Batin, Hafr Al Batin 31991, Saudi Arabia
² Department Electrical and Computer Engineering, Kampala International University, Kampala 20000, Uganda
³ Department of Computer Engineering, Ahmadu Bello University, Zaria 610001, Nigeria
⁴ Department of Electrical and Electronics Engineering, University of Ilorin, Ilorin 240003, Nigeria
⁵ Department of Computer Science and Engineering, University of Hafr Al Batin, Hafr Al Batin 31991, Saudi Arabia
* Correspondence: rbouchekara@uhb.edu.sa

Abstract: Images taken by drones often must be preprocessed and stitched together due to the inherent noise, narrow imaging breadth, flying height, and angle of view. Conventional UAV feature-based image stitching techniques significantly rely on the quality of feature identification, made possible by image pixels, which frequently fail to stitch together images with few features or low resolution. Furthermore, later approaches were developed to eliminate the issues with conventional methods by using the deep learning-based stitching technique to collect the general attributes of remote sensing images before they were stitched. However, since the images have empty backgrounds classified as stitched points, it is challenging to distinguish livestock in a grazing area. Consequently, less information can be inferred from the surveillance data. This study provides a four-stage object-based image stitching technique that, before stitching, removes the background's space and classifies images in the grazing field. In the first stage, the drone-based image sequence of the livestock on the grazing field is preprocessed. In the second stage, the images of the cattle on the grazing field are classified to eliminate the empty spaces or backgrounds. The third stage uses the improved SIFT to detect the feature points of the classified images to obtain the feature point descriptor. Lastly, the stitching area is computed using the image projection transformation.

Keywords: livestock management; drones; image stitching algorithms; cattle monitoring; drone image processing



Citation: Bouchekara, H.R.E.H.; Sadiq, B.O.; O Zakariyya, S.; Sha'aban, Y.A.; Shahriar, M.S.; Isah, M.M. SIFT-CNN Pipeline in Livestock Management: A Drone Image Stitching Algorithm. *Drones* **2023**, *7*, 17. <https://doi.org/10.3390/drones7010017>

Academic Editor: Fei Liu

Received: 7 November 2022

Revised: 9 December 2022

Accepted: 9 December 2022

Published: 26 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Unmanned aerial vehicles (UAVs) are routinely applied in diverse fields. One application area is the agricultural field, where UAVs can monitor grazing fields, crops, and pests; act as actuators; and observe regional geological conditions and farm produce. One of the many features that must be developed for dependable technology to aid people in using and maximizing the function of drones is the capacity for photogrammetry. However, cloud cover is a problem as long as photogrammetry is still only acquired from satellites, particularly during the rainy season. Additionally, reliance on satellite data is expensive and results in information delays due to the slow data collection rate. One of the alternative technologies to get a more thorough, fast, quick, and economic assessment is imagery taken by an unmanned aerial vehicle. Another method for locating a place for various purposes is to use a UAV equipped with a camera [1].

Both single and multiple drone technological systems have been recently employed to achieve effective and efficient surveillance of agricultural farms and cattle grazing fields to improve farming and herding operations whilst eradicating hunger [1,2]. However, the surveillance images obtained from this agricultural farmland and cattle grazing areas using

drones are limited by the drone camera's angle of view [2]. This effect often leads to overlapping images, causing challenges and obstacles in image processing and data recognition. One common technique used to solve the issue of overlapping images in a drone-based cattle surveillance system is the image stitching technique [3]. The process of integrating several overlapping images from various viewpoints to produce a high-resolution image representing the information of an entire coverage region is known as image or panorama stitching. Direct pixel-to-pixel and feature-based methods can be used to complete the procedure. Stitching images are essentially broken down into multiple processes, including interest points extraction and matching, target image registration, mixing, and special border processing. The target image's registration and integration are the most crucial [4]. The precision and speed of the finished spliced image will be significantly impacted by the image splicing algorithm used. The diagram in Figure 1 conceptualizes the processes involved in drone-based image stitching for cattle grazing fields [5,6].

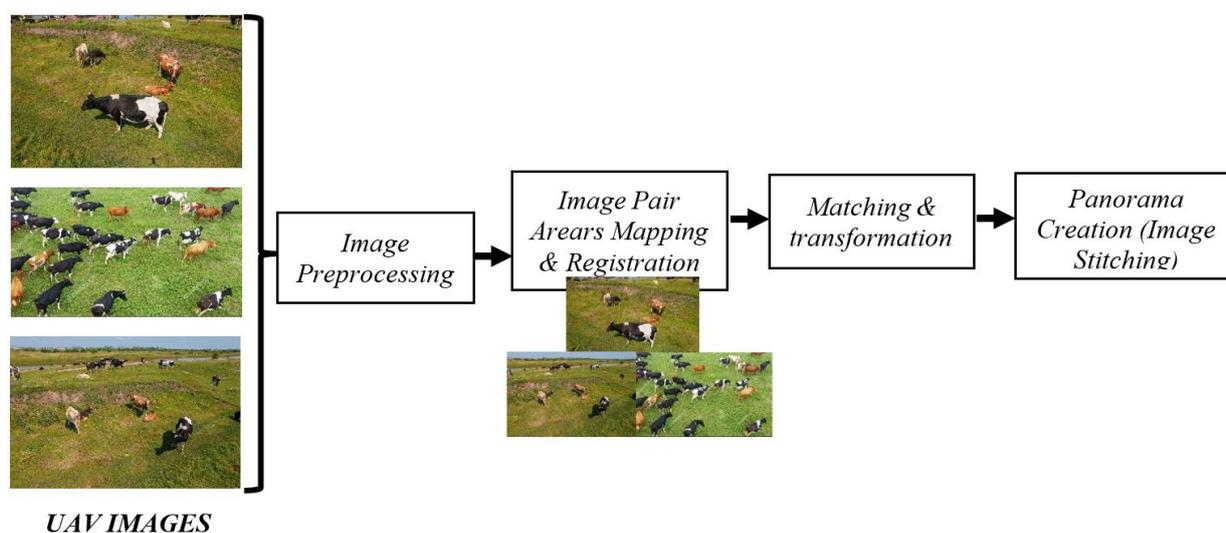


Figure 1. Generic Drone-Based Image Stitching Process for Cattle Surveillance.

Figure 1 serves as an illustration of a general drone-based image stitching method for cattle observation. The four development stages of the system under discussion are picture preprocessing, image matching and registration, image modification, and image stitching. The drones have cameras with sensors. This makes it possible to get photographic data on the cattle and the grazing area [7]. The designed system receives and preprocesses the images collected by each drone. For the vast number of images captured by these drones on grazing fields to be useful for inferential purposes, they must first be analyzed quickly. Second, wind-induced UAV vibrations and unstable aircraft control likely lead to the capture of low-quality surveillance images. The constant position changes of UAVs make it challenging to estimate camera pose, which necessitates preprocessing.

Because conventional methods, such as human watching over the grazing field, are not sustainable, drone-based cattle management and surveillance in contemporary cities are crucial [8–11]. This is because the drone/UAV can have a broader view range of the cattle on the grazing field than the human eye. The constant surveillance of cattle on grazing fields can be linked to the continuous increase in demand for and strain on the available resources to meet the goals of the international community's sustainable development agenda, which aims to eradicate world hunger by 2030. Recent research papers have made an effort to support the global community's objective for sustainable development by utilizing various cutting-edge strategies for cultivating agricultural produce and managing cattle. To enhance farming and herding operations, the smart livestock farm deployed drone technology as a surveillance approach [12–16]. The drones used for the job featured sensors for taking pictures and computer vision capabilities that oversaw the entire process.

Nevertheless, during a drone flight, a single surveillance image only contains a small amount of information due to the limited flight height and the wide angle of the aerial camera, among others. This causes some challenges and obstacles in the processing and recognition of the data presented in it. Multiple aerial livestock images must be picture-stitched together to acquire information on the cattle in a generally full surveillance area [17–21]. As a result, creating a full surveillance area will effectively inform stakeholders in the livestock farming industry of the insights obtained from the processed data and suggestions for management strategies.

The remainder of the text is organized as follows: The usage of common approaches in drone-based image stitching, usage platforms, and state-of-the-art drone-based cattle management and surveillance in contemporary cities are all reviewed in Section 2. Section 3 describes the platform's technological features, which enable the real-time use of image identification and description algorithm for cattle management and surveillance. The results obtained compared to other works of literature are presented in Section 4. Finally, Section 5 summarizes the research findings and discusses potential further research directions.

Several approaches to drone-based image stitching have been widely used in literature. These are the Speeded Up Robust Feature (SURF), Oriented FAST and rotated BRIEF (ORB), and Scale-Invariant Feature Transform (SIFT) [22–25]. The development of some of the recent drone-based image stitching techniques can be traced to these three frequent approaches. The operating principle and governing equations are presented subsequently.

SURF: The SURF method is a quick and reliable approach for local, similarity-invariant representation and comparison of images. The SURF approach's essential appeal is its ability to compute operators quickly using Laplacian of Gaussian box filters, enabling real-time applications like tracking and object detection. The core of the SURF method depends on the Hessian matrix [26–28]. The SURF algorithm primarily entails creating the Hessian matrix, creating the scaling space, determining the feature points first, then using non-maximal suppression, precisely locating the feature points, and choosing the primary direction [29,30]. The governing equations to perform the SURF operations are presented in Equations (1)–(3).

The entry of an integral image $I_{\Sigma}(x)$, which is at a drone-based image location, $z = (x, y)^T$, indicates the total number of pixels in the input image I contained inside the rectangle area bounded by the origin and z .

$$I_{\Sigma}(x) = \sum_{a=0}^{\sigma \leq x} \sum_{b=0}^{b \leq y} I(a, b) \quad (1)$$

The Hessian of a given pixel is computed as

$$Hessian(f(a, b)) = \begin{bmatrix} \frac{\partial^2 f}{\partial a^2} & \frac{\partial^2 f}{\partial a \partial b} \\ \frac{\partial^2 f}{\partial a \partial b} & \frac{\partial^2 f}{\partial b^2} \end{bmatrix} \quad (2)$$

Filtering the image using a Gaussian kernel at the point $z = (x, y)^T$, the Hessian matrix $Hessian(z, \sigma)$ in z at scale σ is defined as

$$Hessian(z, \sigma) = \begin{bmatrix} L_{aa}(z, \sigma) & L_{ab}(z, \sigma) \\ L_{ab}(z, \sigma) & L_{bb}(z, \sigma) \end{bmatrix} \quad (3)$$

ORB: ORB is an alternative to SURF and SIFT, combining FAST and BRIEF descriptors. Several researchers have used ORB because it is an open-source alternative to the SURF and SIFT algorithms [31,32].

SIFT: A method for identifying conspicuous, stable feature points in an image is called SIFT. It also offers a group of features that characterize and describe a tiny area of the surrounding image for each such point. These characteristics are scale- and rotation-invariant. Four stages are used to perform the SIFT operations, each

with a governing equation. The four stages are scale-space extrema detection, key point localization, orientation assignment, and key point descriptor. Equations (4)–(6) present the mathematical expression of the stages [33–35]. To perform the scale-space extrema detection on a drone-based image $I_D(x, y)$, a transformation function is defined as a product of the convolution of a Gaussian kernel with the drone-based image.

$$L(x, y, \sigma) = GK(x, y, \sigma) \times I_D(x, y) \quad (4)$$

where $GK(x, y, \sigma)$ is the Gaussian kernel and σ is the width of the Gaussian filter.

Once the interesting point has been detected, orientation is assigned to each of the selected key points, and then the selection of the Gaussian smoothed image from the scale of the key point occurs. For the Gaussian smoothed image, $G(x, y)$, the orientation and magnitude are defined as

$$\theta(x, y) = \tan^{-1} \frac{G(x, y + 1) - G(x, y - 1)}{G(x + 1, y) - G(x - 1, y)} \quad (5)$$

$$m(x, y) = \sqrt{(G(x + 1, y) - G(x - 1, y))^2 + (G(x, y + 1) - G(x, y - 1))^2} \quad (6)$$

2. Review of Related Works

This subsection reviews the works conducted on drone-based image stitching to obtain the requisite information about the livestock on the grazing field. In the article [36], the authors presented a Speeded Up Robust Feature (SURF) algorithm designed to address the issue of a single drone not covering an image because of the drone's narrow field of view. Using the Hough transform and the existing image stitching algorithm, the article tweaked and improved the conventional SURF algorithm. The decision tree forest in the generalized Hough transform is used to filter the set of feature points after the feature points are extracted by the SURF algorithm for the images to be spliced, and the fading image fusion technique is then used for image splicing.

The authors updated the SURF method because of its reputation for quick feature point extraction. Nevertheless, when the degree of rotation and angle of view is too wide, the SURF algorithm typically has low matching accuracy. The presented drone image stitching algorithms in the literature have been classified as indoor and outdoor environmental-based algorithms. This can be seen in the work of [37], which presented a frame stitching technique for an indoor-based surveillance drone. With the goal of tackling low illumination situations and non-planar scenes, which are common in indoor environments, a selective approach was devised for merging frames of drone-captured video. The authors' indoor example of the scenario was a warehouse. As a result, it was claimed that the developed technique could handle sparse feature points and a lack of feature point correspondence in a particular scene. Both qualitatively and statistically, the proposed method was contrasted with the SURF and Scale-Invariant Feature Transform (SIFT) techniques. According to the study's presentation, the presented technique was designed to address many challenges far better than the other algorithms.

An unmanned aerial vehicle (UAV) image stitching technique was reported in [38] and was based on the optimal seam algorithm and half-projective warp. The method was created to preserve the image's original information successfully and achieve the optimum stitching effect in the UAV images. The ghosting and blurring issues on the stitched images are supposedly eliminated mostly by current seam stitching algorithms. However, the deformation and angle distortion brought on by image registration are still present in the stitched images. The author attempted to correct the deformation and angle distortion due to image registration in the stitching results. First, the correction was said to be done by creating a new difference matrix that contains the structural, color, and line difference information in the overlapped region of the aligned images.

The authors proposed a seam search algorithm by limiting the search range of the seam using a minimum and global energy value. The stitched image's shape was corrected due to the seam position and half-projective warp, preserving the original shapes of more regions. Some UAV photos were used in the experiment to demonstrate the stitching effect. Reference [18] presented a local mesh-based bundle adjustment to successfully stitch numerous overlapping drone photos into a natural panoramic image due to the limitation of parallax caused by perspective changes and the inability of the existing traditional method to handle the effect. A framework with an energy minimum as a performance metric was created to accomplish this, and the parallax faults were incorporated into it. This energy can be reduced effectively based on local bundle modification, eliminating parallax effects, and achieving precise alignment. The studies revealed that the stitching process performs better than the compared techniques in eliminating parallax effects and producing more realistic outcomes. However, ghost problems in captured surveillance scenes still occur.

To prevent the ghost problem in conventional approaches when the scene is subject to significant shifts in viewpoint, the authors of [39] integrated a robust elastic warping-based parallax-tolerant picture stitching method into the stitching of drone images. Nonetheless, some outliers were identified with this method. To properly remove outliers from a set of putative feature correspondences, a set of high-precision matched feature points were first provided for a pair of drone images. The input image was then warped according to the estimated deformation on the non-grid plane. A robust elastic warping function, which eliminates the parallax error, was also used. The global projectivity preservation approach was then used to produce resultant panoramas with high precision. Results of tests showed that the approach could produce superior panoramas to its cutting-edge rivals. The authors in [40] then provided a panoramic image that guarantees that drone surveillance images of dairy cattle are not cropped, duplicated, or missed and that the final image is practically inferential. It suggested that authors should ensure the use of a technique separating the dairy cattle's distinct regions and adding them to a panoramic image. To make a panoramic image without the difficulty of 3D reconstruction, a map of a cattle barn floor was utilized as the compositing surface. The findings of individual cattle region extraction were used to create the image.

The work of [41] guided animals on the grazing field using a multiple quadrotor UAV. The animals were guided from the grazing field to their pen with the help of the drone surveillance images incorporating noises of predators modelled with an exponential function with a view of providing a solution to the cattle roundup problem. An initial aerial picture of a big herd of cattle was obtained, and image stitching was implemented by combining the image frames. A little deviant from the work of [36–41], the authors in [42] used one of the most common drone image-based stitching techniques to monitor crop pests quickly and effectively. The main topic of discussion was using an enhanced scale-invariant feature transform (SIFT) algorithm to scan drone images and extract disease crop information quickly. Conversely, the least-squares match is employed to locate exact matches given all the rough feature matches based on SIFT features. The image stitching and mosaicking are based on an improved SIFT algorithm for the UAV images in various deformation, lighting changes, and resolutions.

Using panoramic image stitching and object identification, the authors of [43] developed a drone patrol system. The authors tried to propose a technique that will stitch drone images whilst eliminating motion ghosts. The scheme integrates the SPHP algorithm with a region-growing algorithm based on various images to create a panorama image and eliminate motion ghosts. To improve the results of object classification, it also emphasizes scene classifications and employs the well-known Faster RCNN object detector. For images of straw farming, the authors of [44] recommend improving the scale-invariant feature transform (SIFT) method. The image is down-sampled before feature detection to reduce the number of feature points and improve feature extraction. Then, feature points are matched using gradient normalization-based feature descriptors to improve the matching accuracy.

The images are combined with the best stitching, fade-in, and fade-out effects to create high-quality stitching. When stitching together surveillance images from an unmanned aerial vehicle, reference [45] investigated the mismatch of the overlap areas from the UAV. A local stitching technique was suggested to ensure the evenness of the projection link between each image and the panoramic image to solve the problem as follows. The feature points were first filtered. Image rotation systems were developed to stop the error of image tilt brought on by perspective changes. A single-band grey-scale image was created for each input image by combining the bands that relate to the red, green, and blue wavelengths, according to the authors' proposal in [46], which uses deep feature matching and elastic warp. The feature points of each HSI were matched using a graph neural network, and the spectral consistency of adjacent images was then ensured using intrinsic decomposition and the determination of the transformation matrix of nearby images.

Another work [47] proposed a novel image stitching method for UAV images utilizing an adaptive natural warp for image registration. This method aims to create two aligned images in the same coordinate system. To find a perceptually optimal seam located in continuous areas with high similarity, information on the color difference, gradient difference, and texture complexity was combined using the Superpixel energy function. The Superpixel color mixing technique was utilized to remove the evident seams and produce realistic color transitions. At the same time, a graph cut algorithm was used to conceal the artefacts in the overlapping parts. The authors of [48] provide a rapid adaptive stitching method for managing numerous aerial images by analyzing the UAV image footprints' relative positions and overlapping regions using their available geotag information. All the UAV images were processed to remove the heavily correlated images using quick feature extraction and feature matching method. A local warp approach with a smooth transition for overlapping regions was introduced to eliminate blurring artefacts and achieve exact image alignment. The global alignment combined with a sharp preserving wrap to generate natural-looking panoramas was also presented by [49]. A global alignment technique was suggested to align the input images better to achieve natural-looking panoramas. The stitching results were stated to improve alignment accuracy while maintaining the shape when combined with a shape-preserving warp.

The work presented in [50] suggested a projection plane selection approach for UAV image stitching in which a weighted topological network is established. All images are divided into stable and unstable groups based on their affine model registration faults. Using this method, the reference image is selected from the stable group, maintaining a minimal total cumulative registration error. Based on the cost of the topological graph, each unstable group image has a local image feature linked to a stable group image. The authors of [51] offer a stitching technique that accurately predicts homography for images with slight parallax using a deep neural network. The fundamental components of the concept are feature maps with progressively better resolution and the hybridized matching cost volumes. The author also provides a novel loss function for stitching that considers the contents of images. To stitch a duo of images together, the authors of [52] employ SIFT and SURF as feature detectors and feature descriptors. They also use a brute force matcher with a Euclidean distance. Following the discovery of the initial match, the Random Sample Consensus (RANSAC) is used as homography to acquire the correct matching estimates. The deep neural network-based stitching approach was proposed by [51]. Three main issues that need to be fixed for aerial video stitching are the size factor in the stitched images, the frequently moving camera projection center, and severe motion blur in some frames, according to [53]. The authors further advise looking for items with apparent depth changes and selecting a zone containing them for stitching. They also suggest assessing the amount of motion blurs in each video frame and, if possible, avoiding using frames with noticeable motion blur. A new image mosaic technique for the UAV aerial image was provided by [54] using the fast explicit diffusion (FED) algorithm based on the modified KAZE algorithm. After adopting the Hamming strategy for a rough match of these feature

points, the upgraded PROSAC technique was applied for exact correct matching. The image was then stitched together using the weighted average method.

In conclusion, it is evident from the reviewed works that, through image stitching technology, high-resolution images taken from the air by drones can be processed to offer image information for the quick and precise detection of animals in grazing fields. Research works have used conventional algorithms like the SIFT, SURF, and ORB, as well as some deep learning approaches. However, using conventional algorithms presents some drawbacks, such as high dimensionality reduction of feature descriptors and low matching efficiency. Furthermore, some authors that used the deep learning approaches have an empty background that is classified as stitched points, which are challenging to distinguish livestock in a grazing area. As such, this paper proposes a pipeline for drone image stitching algorithms using SIFT-CNN (SIFT-Convolutional Neural Network). Our contributions are summarized as follows:

- We designed an improved SIFT-CNN algorithm pipeline for drone-based image stitching suitable for a livestock management system. We showed that classifying the images on the grazing field and removing empty backgrounds from the stitch points improves inference drawn from drone surveillance images.
- We simulated the design system with a set of high- and low-resolution images to learn the various impact on inference from the drone surveillance images.
- The proposed algorithm is compared with the conventional algorithms applied to grazing fields for effectiveness and efficiency.

3. Proposed System

This subsection presents the methodology adopted in this work. The grazing field represents the geographical area covered by the cattle. The problem formulation used a set of three UAVs employed to cover a grazing field in which each UAV was equipped with a camera to capture surveillance images of the livestock. The diagram in Figure 2 presents the pipeline of the proposed system.

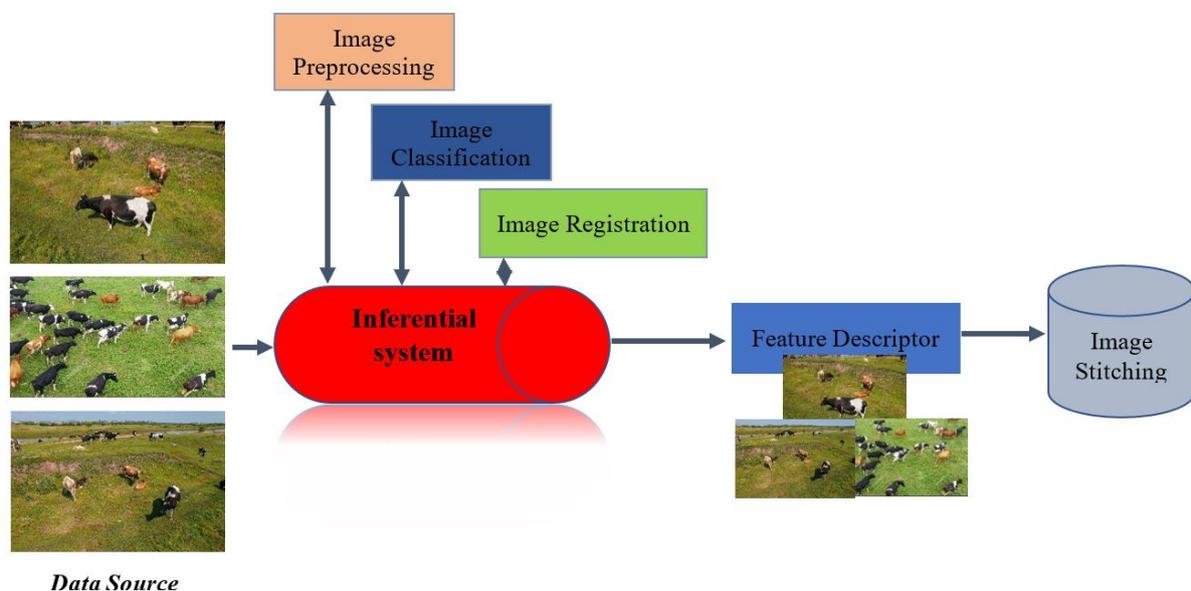


Figure 2. Conceptual Pipeline of the Proposed System of Cattle Livestock Management.

The proposed pipeline presented in Figure 2 is not limited to cattle surveillance applications. It is applicable to other UAV surveillance scenarios but with modifications.

3.1. Methodology

The step-by-step approach employed for successfully implementing the proposed method presented in this paper is discussed in the following items.

Image Preprocessing and Enhancement

Images collected on grazing fields suffer from noise inherent to these images. This type of noise is mostly speckle noise that can be attributed to the type of camera used or the weather conditions. The test images were collected and separated as those containing images of cattle and just empty backgrounds and were read in as an input into the MATLAB environment accordingly. The authors of [55] compared convolutional neural network architectures to determine bird image features efficiently in an anti-drone system. The obtained results showed that the architectures of the AlexNet, VGG-16, ResNet18, and ResNet50 obtained higher accuracy in comparison to the GoogLeNet, Inception-v3, and Squeezenet. Based on this, the ResNet architecture is used in this research work. Since we are using the already pre-trained network model, we first load the pre-trained network, perform data argumentation, replace the final layers, and train the network. Afterwards, we predicted and accessed the model's accuracy and deployed the results for the stitching process a priori. The collected data were preprocessed as the input, as shown in Figure 3. The mathematical formulae used to preprocess the input images, which are of a different dimension, are presented in Equation (7).

The ideal images captured from a surveillance camera on a grazing field are prone to noise due to varying weather conditions. Hence, to improve the information quality obtained from these surveillance images, the images were enhanced using a developed adaptive histogram-based equalization technique focusing on the hue and saturation values. The underlying equations for the enhancement techniques are presented in Equations (8)–(11), whilst the algorithm is shown in Algorithm 1. The collected images were in RGB format, converted to HSV format, and filtered using a Laplacian filter. The variances of rows and columns p_v, q_v were computed and used to enhance the Laplacian filter p_a, q_a for correlation and image enhancement.

$$p_v = \sum [p(a, b) - p_a]^2 \quad (8)$$

$$q_v = \sum [q(a, b) - q_a]^2 \quad (9)$$

Equations (9)–(10) were used for correlation and enhancement of the images.

$$C(a, b) = \frac{\sum [p(x, y) - p_v], [q(x, y) - q_v]}{\sqrt{p_v, q_v}} \quad (10)$$

$$p_{enhancement} = p(a, b) + K_1 [p(a, b) - p_v] - K_2 [q(a, b) - q_v] \times C(a, b) \quad (11)$$

where K_1 and K_2 are constant values between 0 and 1.

The output of contrast after subjecting it to the Laplacian filter is enhanced using a fixed value of gamma 0.77.

$$q_{enhancement} = q^{0.77} \quad (12)$$

The enhanced luminance and contrast ($p_{enhancement}$ and $q_{enhancement}$) components of the images are added back to the hue value to get the improved HSV of the image. The HSV value is then converted back to RGB for classification.

$$(w_{224}, h_{224}) = \frac{D_M}{\max(w, h)} (w, h) \quad (7)$$

where (w, h) are the width and height of the images, respectively.

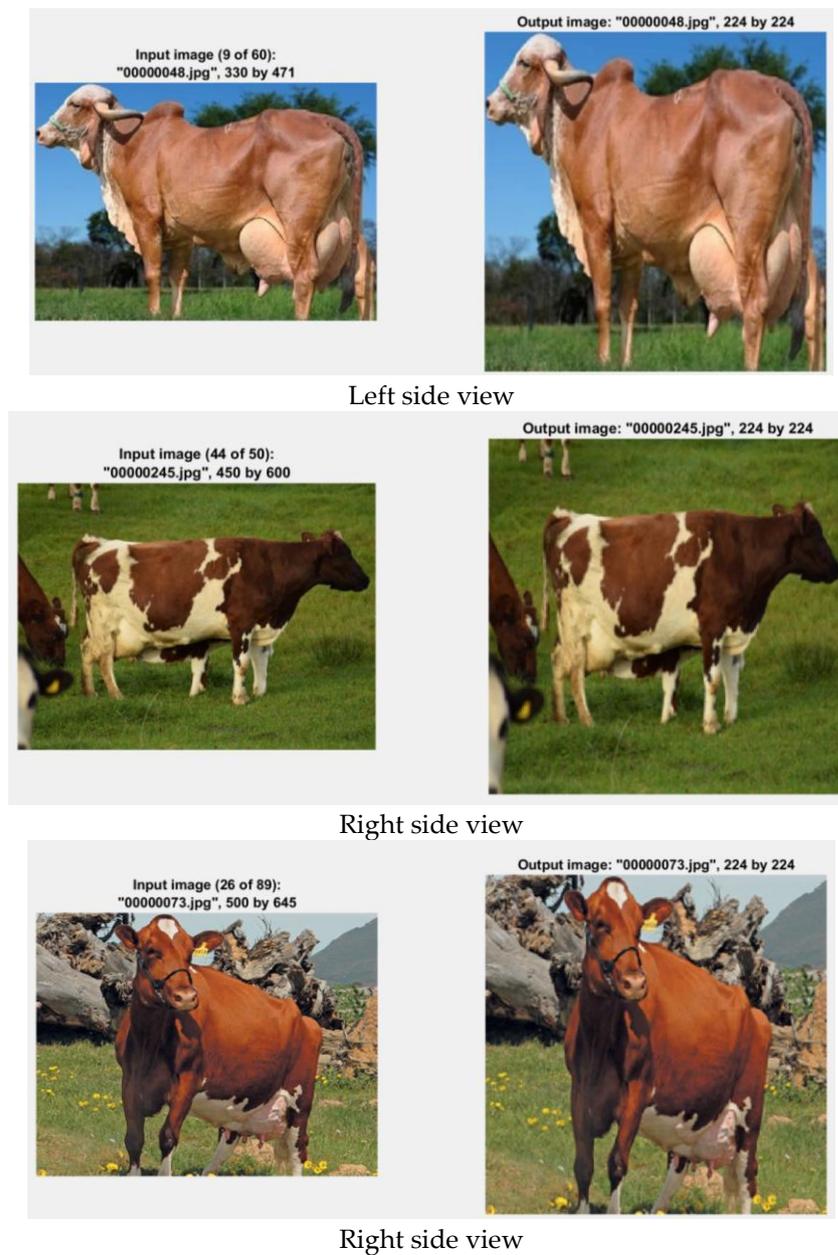


Figure 3. Processed Image.

Algorithm 1: Image Pre-processing.

Input: K_1, K_2

Output: Enhanced Images.

- 1 Survey grazing field and capture images (I)
 - 2 RGB to HSV conversion $RGB \rightarrow HSV$
 - 3 Separate colour components in images
 $H = I(:, :, 1), H = I(:, :, 2), H = I(:, :, 3),$
 - 4 Laplacian filter—Compute S and V
 - 5 Process image—compute p_v and q_v using (8) and Compute $C(a, b)$ and $P_{enhancement}$ using (9) and (10)
 - 6 Enhance colour—compute $q_{enhancement}$ using (11)
 - 7 Combine HSV and convert $RGB \leftarrow HSV$
-

3.2. Image Classification

The result of the enhanced images (R, G, B) was passed through the Resnet 50 classification algorithm, in which the architectures are depicted in Figure 4. Since the input images are a set of numeric scalars representing features, a feature layer (fc1000) is used, and the images will enter the ResNet50 layer with the pre-trained weights. The proposed model has two layers, a pre-trained ResNet layer and a Feature layer, as illustrated in Figure 4. Importing pre-trained weights for the ResNet50 model, the input data will be trained using pre-learned weights, and the feature layer will be used to detect specific global configurations of the features detected.

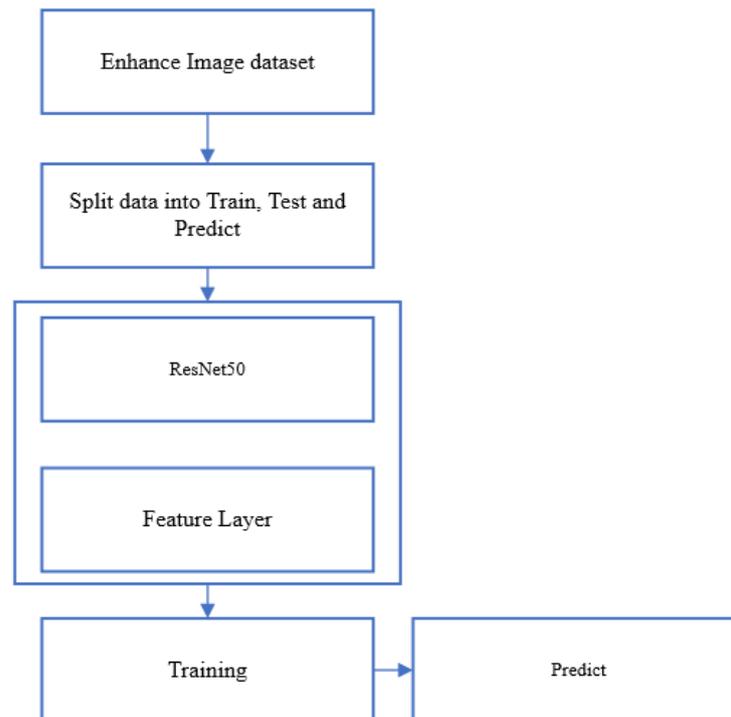


Figure 4. Image Classification Illustration.

3.3. Image Stitching

In the image stitching stage, feature detection is performed on the classified images that will be stitched. The SIFT is used for the feature detector. This procedure's first step is calculating the corner detection function to find the corners on the images since it is a grazing field image, thereby passing the coordinates into the SIFT detector. With the SIFT detector, the extent and where a key point is located in the images are identified with the aid of the Gaussian difference. Afterwards, a table of the distances between the descriptors of each image is obtained. The Euclidean distance between the image's vectors is used, and the distances smaller than a set threshold value are recorded. Using the recorded indexes, paired coordinates between images are identified. The RANSAC is then used to find the transformation matrix. Once the transform between the images is determined and the paired coordinates are known, the images can be stitched together. The diagram in Figure 5 depicts the image stitching procedure.

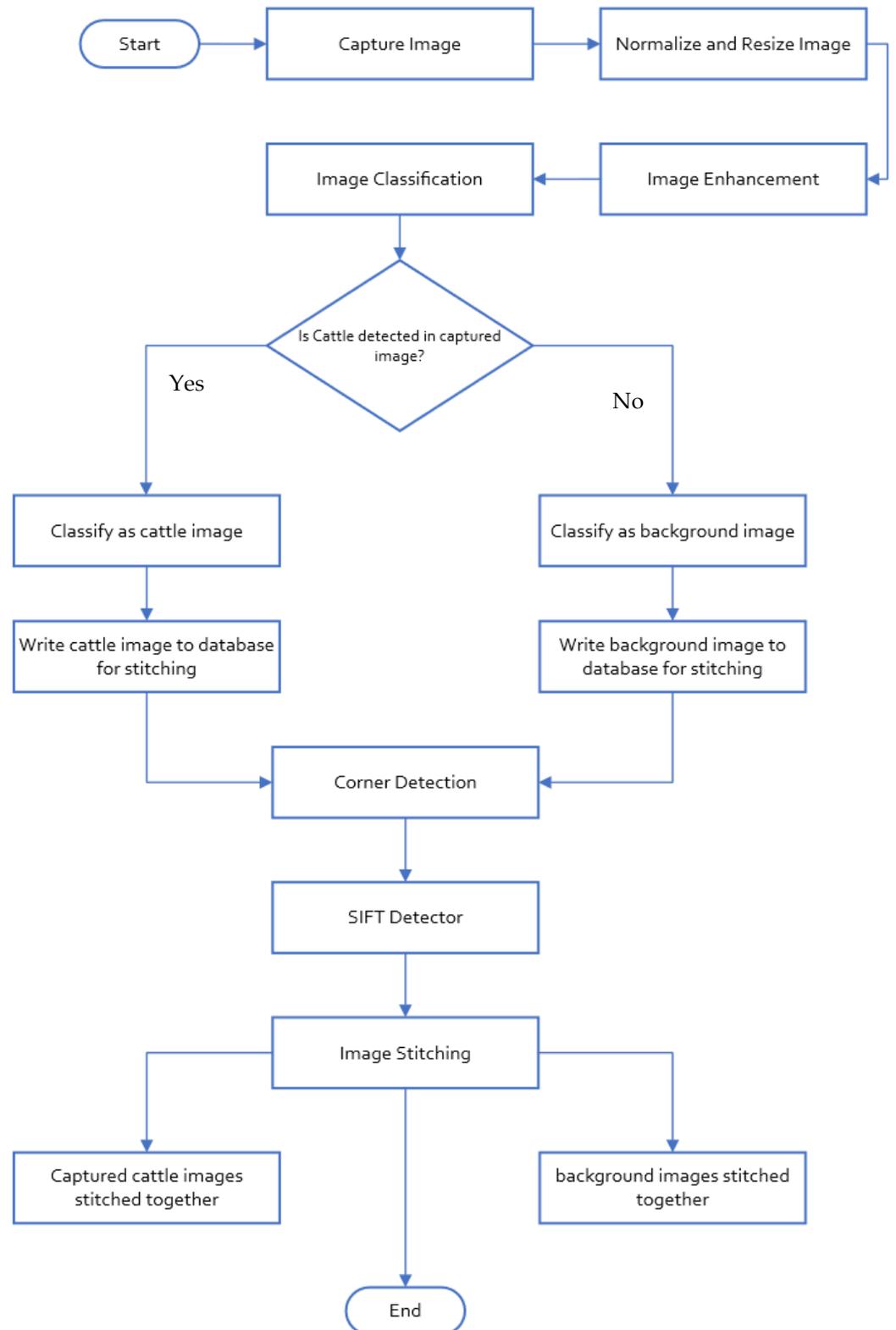


Figure 5. Stitching Procedure.

The algorithm for the stitching procedure is presented in Algorithm 2.

Algorithm 2. Stitching Procedure.**Input:** Input: Classified images, Th, dist_th, σ **Output:** Enhanced Images.

```

1   Detect corners in (I)
2   Initiate SIFT detector
3   Define transformation using Equation (4)
4   Find the orientation of the gaussian smoothed image using Equation (5)
5   Find the magnitude of the gaussian smoothed image using Equation (6)
6   Compute Euclidean distance (ED) between the images
7   If  $ED < dist\_th$ 
8     Record indexes
9   else go to line 6
10  endif
11  Determine the pair coordinates of the images using the RANSAC algorithm
12  Compute transformation matrix
13  Register commonality between the set of images
14  Stitch images

```

The Source code of the proposed methodology and pipeline are available at <https://github.com/bosadiq?tab=repositories> accessed on 2 January 2022.

4. Results and Discussion

In this section, we present the result obtained from the image preprocessing and then obtained from the pre-trained convolutional neural network model. A discussion using the SIFT-CNN algorithm is presented. After that, the performance of SIFT-CNN on cattle surveillance on grazing fields is presented. Cattle dataset images, which were separated based on the view (front view, left view and right view), were obtained from <https://www.kaggle.com/datasets/afnanamin/cow-images> (accessed on 2 January 2022). The obtained dataset consists of 89 cattle images of the front view, 60 cattle images of the side view, and 50 cattle images of the right view. The obtained dataset was split into 70% and 30% for training and testing, respectively. Cattle images as input were used to predict whether the image was that of cattle or not. The possible outcomes of the classifier are positive and negative.

4.1. Result Analysis on Image Preprocessing and Enhancement

A subjective evaluation of the images' quality using subjective visual effects was utilized as the first approach, and an objective evaluation utilizing the peak-to-signal-noise ratio (PSNR) as the second method as follows:

$$psnr = 10 \log_{10} \left[\frac{255^2}{MSE_R} \right] \quad (13)$$

The MSE_R is the mean square error and is obtained as follows:

$$MSE_R = \frac{\sum_{(a,b)} (I_{acquired}(a,b) - I_{enhanced}(a,b))^2}{ab} \quad (14)$$

where $I_{acquired}(a,b)$ is the acquired image, and $I_{enhanced}(a,b)$ is the enhanced image.

The higher the value of the PSNR, the better the reconstructed image. The subjective visual effect of the obtained image preprocessing technique on one of the sampled images is presented in Figure 6.

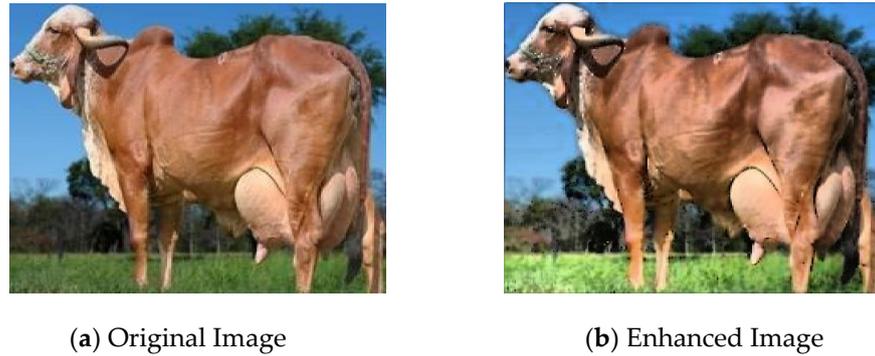


Figure 6. Subjective Visual Effect of the Obtained Image Preprocessing Technique.

Based on Equations (12) and (13), the obtained PSNR for the proposed image enhancement technique on cattle images was 65.7939. The obtained result will help improve the inference drawn from the surveyed images.

4.2. Result Analysis on the Classification of Cattle Images

The obtained images were corrupted with different noise levels because noise is inherent in cattle images obtained from cattle grazing fields using drones to test the classification accuracy before stitching. The visualization of the first section of the network architecture of the MATLAB-based image classification algorithm is depicted in Figure 7, while the diagram in Figure 8 presents the architecture of the ResNet50.

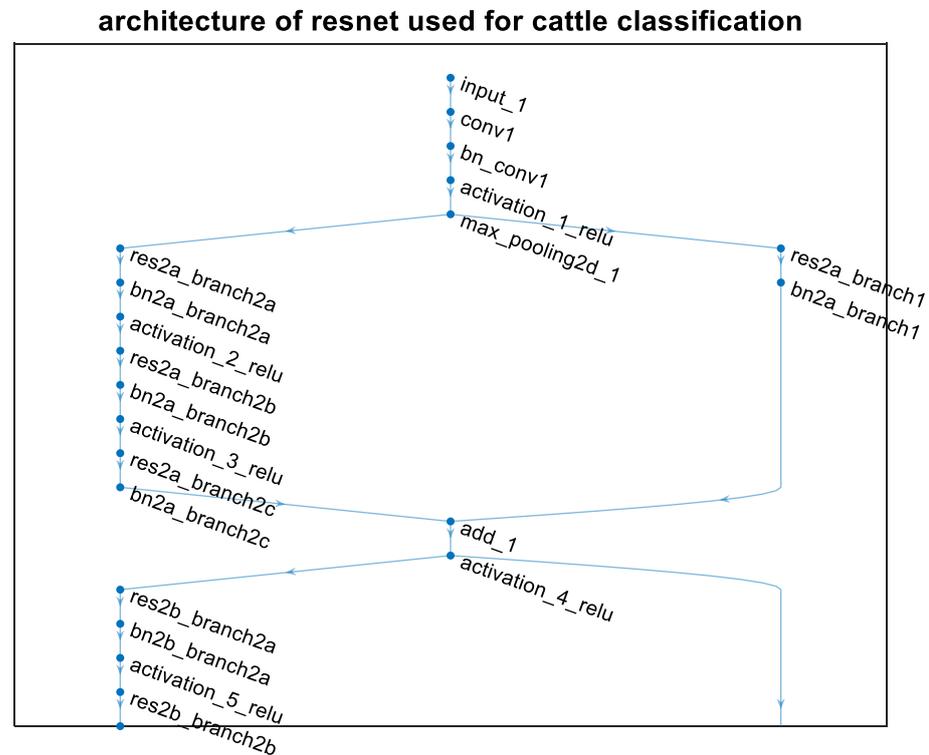


Figure 7. The First Section of the ResNet Network Architecture Used.

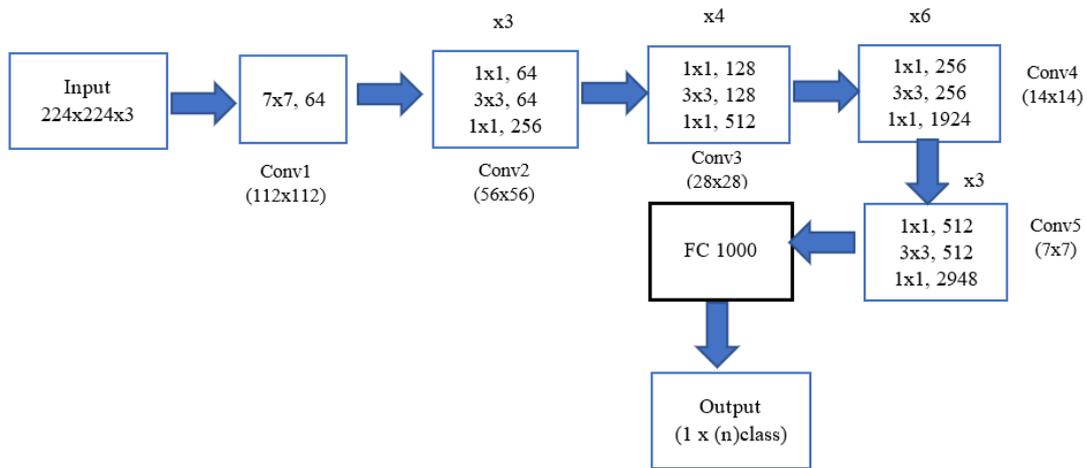


Figure 8. ResNet50 Architecture.

Table 1 presents the obtained results for the image preprocessing and classification algorithm.

Table 1. Result of Image Preprocessing and Classification.

Noise Level (%)	Classification Accuracy (%)	PSNR
0	95.56	65.7939
5	62.20	65.7986
10	60.00	65.8245
15	60.00	65.8243
20	60.00	65.8465
25	55.60	65.8465
30	55.56	65.8465

From the results presented in Table 1, it can be observed that the enhancement algorithm helped to improve the classification of images of cattle on the grazing field captured by drones. As the level of noise captured with the images increases, the classification accuracy decreases when the captured images are not passed through the enhancement algorithm. The PSNR increases as the level of captured noise increases, implying that the corrupted images are reconstructed with improved quality. The noise level graph against the classification accuracy and PSNR is shown in Figure 9.

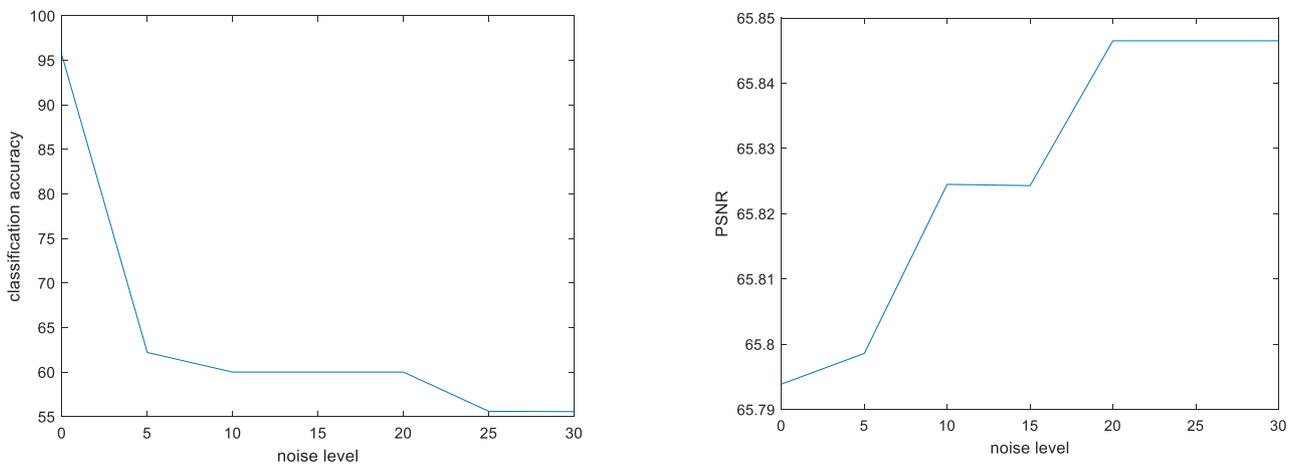


Figure 9. Noise Level against Classification Accuracy and PSNR.

4.3. Result Analysis of the Stitching Procedure

The SIFT-CNN algorithm uses the generalized Harris corner detection to extract corners and infer feature information of the cattle on the grazing field to achieve the screening and match objectives. It first harvests feature points using the SIFT technique. In order to analyze the performance of various feature-matching methods and weigh the advantages and disadvantages of each methodology, this study assesses the effectiveness of feature extraction and matching for the SIFT, SURF, SURF-GHT, and SIFT-CNN algorithms, respectively. In the test set of this paper, MATLAB's underlying library function is called by the SURF and SIFT algorithms. However, using the MATLAB-based SURF and SIFT algorithm, the procedure has to be done in grayscale and not RGB as used with the SIFT-CNN. If two random images of the cattle on the grazing fields are to be matched, then the result obtained for the SIFT-CNN matching is depicted in Figure 10.

As shown in Figure 10, the red circles in the images are sample matching points between the two images used as input in step a. In step b, the distance between matching points was further explored to determine the correlation/corresponding points between the images. Step c presented the stitched images together after the corresponding/correlation points were detected. The comparison of the matching effect is presented in Table 2.

Table 2. Matching Effect Comparison.

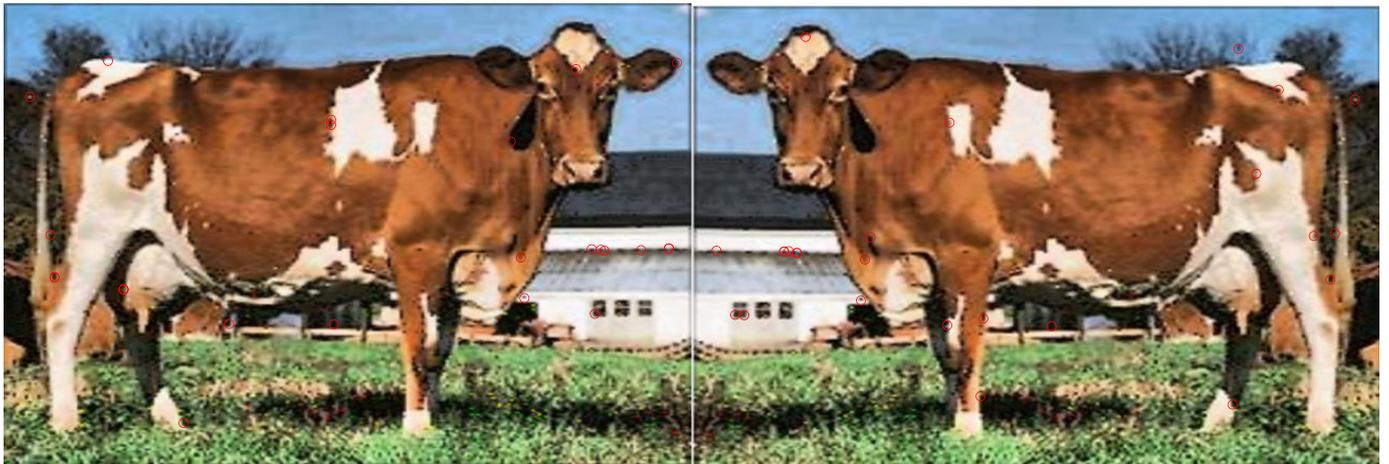
Algorithm Type	Number of Characteristic Points Right Graph	Number of Characteristic Points Left Graph	Matching Pair	Matching Time/s
SIFT	2010	1658	128	36.997
SURF	2000	1169	96	23.316
SIFT-CNN	2010	1658	128	123.004

The table above demonstrates that when all other requirements are satisfied (images captured with no noise present), the SIFT-CNN method matches well but takes the longest duration due to its preprocessing stage. The SURF matching is slightly worse even if the matching duration is shorter than that of the SIFT and SIFT-CNN. Table 3 presents the matching comparison when noise is present in the image.

Table 3. Matching Effect Comparison with Noise.

Algorithm Type	Number of Characteristic Points Right Graph	Number of Characteristic Points Left Graph	Matching Pair	Matching Time/s
SIFT	2010	1658	112	36.997
SIFT-CNN	2010	1658	128	123.004

The proposed SURF-CNN technique, when noise is present in the images inherent to drone images, as shown in Table 3, guarantees the same speed of feature extraction while increasing the number of matching point pairs. Furthermore, the algorithm was able to achieve the same matching time as when noise was not present in the images. This implies that the presented algorithm is robust to noise.



Step a



Step b



Step c

Figure 10. SIFT-CNN Matching Result.

5. Conclusions

This study proposes a SIFT-CNN pipeline for livestock management method in stitching drone images captured from the different drones on the grazing field. This is in response to the need to obtain adequate information from the grazing field that one drone cannot provide due to camera coverage angle, imaging breadth, and flight height. To do this, this study used the SIFT algorithm, which can extract features to match them. The proposed system relies on information extraction from aerial photographs to offer a stitching solution for livestock management operations. Hence, giving a direct and practical way to monitor cattle for perimeter coverage, health amongst others, and under other natural circumstances. The method's key contribution is that it practically and formally showed that image preprocessing by removing noise and then classifying the images on the grazing field with a view to eliminating empty backgrounds from the stitch points improves inference drawn from drone surveillance images. It is worth noting that the technique requires more matching time in total due to the image preprocessing stage. Further work should consider real-world implementation (prototype development) using the FPGA or Raspberry pi for full-scale deployment on the grazing field, as well as comparing the performance of the different types of CNN architectures on the proposed pipeline.

Author Contributions: Conceptualization, H.R.E.H.B. and Y.A.S.; methodology, B.O.S., S.O.Z., and Y.A.S.; software, B.O.S. and S.O.Z.; validation, Y.A.S., M.S.S. and H.R.E.H.B.; formal analysis, Y.A.S. and M.S.S.; investigation, B.O.S. and Y.A.S.; resources, H.R.E.H.B. and M.S.S.; writing—original draft preparation, H.R.E.H.B., Y.A.S., S.O.Z., and B.O.S.; writing—review and editing, H.R.E.H.B., Y.A.S., S.O.Z., M.M.I. and M.S.S.; visualization, M.S.S.; supervision, H.R.E.H.B. and Y.A.S.; project administration, H.R.E.H.B. and Y.A.S. and M.S.S.; funding acquisition, H.R.E.H.B., M.S.S., M.M.I., and Y.A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by institutional fund projects under no (IFP-A-01-2-1-2022). Therefore, authors gratefully acknowledge technical and financial support from the ministry of education and University of Hafr Al Batin, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rivas, A.; Chamoso, P.; González-Briones, A.; Corchado, J.M. Detection of Cattle Using Drones and Convolutional Neural Networks. *Sensors* **2018**, *18*, 2048. [[CrossRef](#)] [[PubMed](#)]
2. Yang, F.; Zhu, N.; Pei, S.; Cheng, I. Real-Time Open Field Cattle Monitoring by Drone: A 3D Visualization Approach (2021). In Proceedings of the 15th International Conferences Computer Graphics, Visualization, Computer Vision and Image Processing, Virtual, 21–23 July 2021.
3. Dhana Lakshmi, M.; Mirunalini, P.; Priyadharsini, R.; Mirnalinee, T.T. Review of Feature Extraction and Matching Methods for Drone Image Stitching. In *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*; Springer: Cham, Switzerland, 2019. [[CrossRef](#)]
4. Bai, Z.; Li, Y.; Chen, X.; Yi, T.; Wei, W.; Wozniak, M.; Damasevicius, R. Real-Time Video Stitching for Mine Surveillance Using a Hybrid Image Registration Method. *Electronics* **2020**, *9*, 1336. [[CrossRef](#)]
5. Bang, S.; Kim, H.; Kim, H. UAV-based automatic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching. *Autom. Constr.* **2017**, *84*, 70–80. [[CrossRef](#)]
6. Widiastuti, I.; Muna, N.; Purnomo, F.E.; Lutfi, F.; Soelaksini, L.D. Automatic Image Stitching of Agriculture Areas based on Unmanned Aerial Vehicle using SURF. In Proceedings of the 1st International Conference on Food and Agriculture, Nusa Dua, Bali, Indonesia, 20–21 October 2018.
7. Alanezi, M.A.; Mohammad, A.; Sha'Aban, Y.A.; Bouchekara, H.R.E.H.; Shahriar, M.S. Auto-Encoder Learning-Based UAV Communications for Livestock Management. *Drones* **2022**, *6*, 276. [[CrossRef](#)]
8. Alanezi, M.A.; Sadiq, B.O.; Sha'Aban, Y.A.; Bouchekara, H.R.E.H. Livestock Management on Grazing Field: A FANET Based Approach. *Appl. Sci.* **2022**, *12*, 6654. [[CrossRef](#)]

9. Gallacher, D. Drone-based vegetation assessment in arid ecosystems. In *Sabkha Ecosystems*; Springer: Cham, Switzerland, 2019; pp. 91–98.
10. Alanezi, M.A.; Shahriar, M.S.; Hasan, B.; Ahmed, S.; Sha'Aban, Y.A.; Boucekara, H.R.E.H. Livestock Management With Unmanned Aerial Vehicles: A Review. *IEEE Access* **2022**, *10*, 45001–45028. [[CrossRef](#)]
11. Aquilani, C.; Confessore, A.; Bozzi, R.; Sirtori, F.; Pugliese, C. Review: Precision Livestock Farming technologies in pasture-based livestock systems. *Animal* **2022**, *16*, 100429. [[CrossRef](#)]
12. Al-Thani, N.; Albuainain, A.; Alnaimi, F.; Zorba, N. Drones for sheep livestock monitoring. In Proceedings of the 2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON), Palermo, Italy, 16–18 June 2020; pp. 672–676.
13. Juyal, P.; Sharma, S. Crop Growth Monitoring Using Unmanned Aerial Vehicle for Farm Field Management. In Proceedings of the 2021 6th International Conference on Communication and Electronics Systems (ICCES), Coimbatre, India, 8–10 July 2021; pp. 880–884.
14. Barbedo, J.; Koenigkan, L.; Santos, P. Cattle Detection Using Oblique UAV Images. *Drones* **2020**, *4*, 75. [[CrossRef](#)]
15. Aburasain, R.Y.; Edirisinghe, E.A.; Albatay, A. Drone-based cattle detection using deep neural networks. In *Proceedings of SAI Intelligent Systems Conference*; Springer: Cham, Switzerland, 2020; pp. 598–611.
16. Zuo, A.; Wheeler, S.A.; Sun, H. Flying over the farm: Understanding drone adoption by Australian irrigators. *Precis. Agric.* **2021**, *22*, 1973–1991. [[CrossRef](#)]
17. Colorado, J.; Mondragon, I.; Rodriguez, J.; Castiblanco, C. Geo-Mapping and Visual Stitching to Support Landmine Detection Using a Low-Cost UAV. *Int. J. Adv. Robot. Syst.* **2015**, *12*, 125. [[CrossRef](#)]
18. Wan, Q.; Chen, J.; Luo, L.; Gong, W.; Wei, L. Drone Image Stitching Using Local Mesh-Based Bundle Adjustment and Shape-Preserving Transform. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 7027–7037. [[CrossRef](#)]
19. Xu, Q.; Chen, J.; Luo, L.; Gong, W.; Wang, Y. UAV image stitching based on mesh-guided deformation and ground constraint. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 4465–4475. [[CrossRef](#)]
20. Yang, J.; Liu, L.; Lu, L.; Deng, F. Image registration and selection for unmanned aerial vehicle image stitching. *J. Appl. Remote Sens.* **2020**, *14*, 046512. [[CrossRef](#)]
21. Renwick, J.D.; Klein, L.J.; Hamann, H.F. Drone-based reconstruction for 3D geospatial data processing. In Proceedings of the 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), Reston, VA, USA, 12–14 December 2016; pp. 729–734.
22. Zhu, H.; Jiang, Y.; Zhang, C.; Liu, S. Research on Mosaic Method of UAV Low-altitude Remote Sensing Image based on SIFT and SURF. *J. Phys. Conf. Ser.* **2022**, *2203*, 012027. [[CrossRef](#)]
23. Liu, J.; Wei, Q.; Bai, Y. Fast stitching of UAV images based on improved SURF algorithm. In Proceedings of the 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Changsha, China, 20–22 October 2021; pp. 1310–1313.
24. Wu, T.; Hung, I.-K.; Xu, H.; Yang, L.; Wang, Y.; Fang, L.; Lou, X. An Optimized SIFT-OCT Algorithm for Stitching Aerial Images of a Loblolly Pine Plantation. *Forests* **2022**, *13*, 1475. [[CrossRef](#)]
25. Yan, Q.; Qiu, H. An improved ORB algorithm for the Unmanned Aerial Vehicle (UAV) image stitching task. In *Second International Conference on Optics and Image Processing (ICOIP 2022)*; SPIE: Bellingham, WA, USA, 2022; Volume 12328, pp. 328–334.
26. Sheena, S.; Sheena, M. A comparison of SIFT and SURF algorithm for the recognition of an efficient iris biometric system. *Int. J. Adv. Res. Comput. Commun. Eng.* **2016**, *5*.
27. Zhu, Z.; Zhang, G.; Li, H. SURF feature extraction algorithm based on visual saliency improvement. *Int. J. Eng. Appl. Sci.* **2018**, *5*, 257267.
28. Juan, S.; Qingsong, X.; Jinghua, Z. A scene matching algorithm based on surf feature. In Proceedings of the 2010 International Conference on Image Analysis and Signal Processing, Zhejiang, China, 9–11 April 2010; pp. 434–437.
29. Yuan, M.; Lei, T.; Liu, X.; Li, S. Fast image stitching of unmanned aerial vehicle remote sensing image based on SURF algorithm. In *Eleventh International Conference on Digital Image Processing (ICDIP 2019)*; SPIE: Bellingham, WA, USA, 2019; Volume 11179, pp. 346–351.
30. Byrne, J.; Laefer, D.; O'Keefe, E. Maximizing feature detection in aerial unmanned aerial vehicle datasets. *J. Appl. Remote Sens.* **2017**, *11*, 025015. [[CrossRef](#)]
31. Luo, C.; Yang, W.; Huang, P.; Zhou, J. Overview of Image Matching Based on ORB Algorithm. *J. Phys. Conf. Ser.* **2019**, *1237*, 032020. [[CrossRef](#)]
32. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
33. Wang, X.; Fu, W. Optimized SIFT image matching algorithm. In Proceedings of the 2008 IEEE International Conference on Automation and Logistics, Qingdao, China, 1–3 September 2008; pp. 843–847.
34. Wang, Y.; Yuan, Y.; Lei, Z. Fast SIFT Feature Matching Algorithm Based on Geometric Transformation. *IEEE Access* **2020**, *8*, 88133–88140. [[CrossRef](#)]
35. Kupfer, B.; Netanyahu, N.S.; Shimshoni, I. An Efficient SIFT-Based Mode-Seeking Algorithm for Sub-Pixel Registration of Remotely Sensed Images. *IEEE Geosci. Remote Sens. Lett.* **2014**, *12*, 379–383. [[CrossRef](#)]
36. Duan, H.; Liu, Y.; Huang, H.; Wang, Z.; Zhao, H. Image stitching algorithm for drones based on SURF-GHT. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019; Volume 569, p. 052025.

37. Ramaswamy, A.; Gubbi, J.; Raj, R.; Purushothaman, B. Frame Stitching in Indoor Environment Using Drone Captured Images. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 91–95. [\[CrossRef\]](#)
38. Chen, J.; Li, Z.; Peng, C.; Wang, Y.; Gong, W. UAV Image Stitching Based on Optimal Seam and Half-Projective Warp. *Remote Sens.* **2022**, *14*, 1068. [\[CrossRef\]](#)
39. Chen, J.; Wan, Q.; Luo, L.; Wang, Y.; Luo, D. Drone Image Stitching Based on Compactly Supported Radial Basis Function. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 4634–4643. [\[CrossRef\]](#)
40. Li, R.; Furuta, R.; Taniguchi, Y. Distortion Correction and Stitching of Overlapping Cattle Barn Images. Available online: <https://scholar.google.com.ng> (accessed on 15 September 2022).
41. Jung, S.; Ariyur, K.B. Strategic cattle roundup using multiple quadrotor UAVs. *Int. J. Aeronaut. Space Sci.* **2017**, *18*, 315–326. [\[CrossRef\]](#)
42. Yue, J.; Lei, T.; Li, C.; Zhu, J. The Application of Unmanned Aerial Vehicle Remote Sensing in Quickly Monitoring Crop Pests. *Intell. Autom. Soft Comput.* **2012**, *18*, 1043–1052. [\[CrossRef\]](#)
43. Zhou, Y.; Rui, T.; Li, Y.; Zuo, X. A UAV patrol system using panoramic stitching and object detection. *Comput. Electr. Eng.* **2019**, *80*, 106473. [\[CrossRef\]](#)
44. Liu, Y.; He, M.; Wang, Y.; Sun, Y.; Gao, X. Farmland Aerial Images Fast-Stitching Method and Application Based on Improved SIFT Algorithm. *IEEE Access* **2022**, *10*, 95411–95424. [\[CrossRef\]](#)
45. Sun, Y.; Lv, Y.; Song, B.; Guo, Y.; Zhou, L. Image Stitching Method of Aerial Image Based on Feature Matching and Iterative Optimization. In Proceedings of the 2021 IEEE 40th Chinese Control Conference (CCC), Shanghai, China, 26–28 July 2021; pp. 3024–3029.
46. Mo, Y.; Kang, X.; Duan, P.; Li, S. A Robust UAV Hyperspectral Image Stitching Method Based on Deep Feature Matching. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–14. [\[CrossRef\]](#)
47. Yuan, Y.; Fang, F.; Zhang, G. Superpixel-Based Seamless Image Stitching for UAV Images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1565–1576. [\[CrossRef\]](#)
48. Pham, N.T.; Park, S.; Park, C.-S. Fast and Efficient Method for Large-Scale Aerial Image Stitching. *IEEE Access* **2021**, *9*, 127852–127865. [\[CrossRef\]](#)
49. Guo, D.; Chen, J.; Luo, L.; Gong, W.; Wei, L. UAV Image Stitching Using Shape-Preserving Warp Combined With Global Alignment. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [\[CrossRef\]](#)
50. Xie, R.; Tu, J.; Yao, J.; Xia, M.; Li, S. A robust projection plane selection strategy for UAV image stitching. *Int. J. Remote Sens.* **2019**, *40*, 3118–3138. [\[CrossRef\]](#)
51. Zhao, Q.; Ma, Y.; Zhu, C.; Yao, C.; Feng, B.; Dai, F. Image stitching via deep homography estimation. *Neurocomputing* **2021**, *450*, 219–229. [\[CrossRef\]](#)
52. Wei, Q.; Lao, S.; Bai, L. Panorama Stitching, Moving Object Detection and Tracking in UAV Videos. In Proceedings of the 2017 International Conference on Vision, Image and Signal Processing (ICVISIP), Osaka, Japan, 22–24 September 2017; pp. 46–50. [\[CrossRef\]](#)
53. Cai, J.; Lee, I. The stitching of aerial videos from UAVs. In Proceedings of the 2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013), Wellington, New Zealand, 27–29 November 2013; pp. 448–452.
54. Cui, H.; Li, Y.; Zhang, K. A Fast UAV Aerial Image Mosaic Method Based on Improved KAZE. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 2427–2432.
55. Oh, H.M.; Lee, H.; Kim, M.Y. Comparing Convolutional Neural Network (CNN) models for machine learning-based drone and bird classification of anti-drone system. In Proceedings of the 2019 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Republic of Korea, 15–18 October 2019; pp. 87–90.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.