Image Segmentation in Aerial Imagery: A Review

Ade Purwanto*, Dewi Habsari Budiarti, Mukti Wibowo, Irfansyah Yudi Tanasa, Yomi Guno, Aris Surya Yunata, Asyaraf Hidayat, Fithri Nur Purnamastuti, Dede Dirgahayu
National Research and Innovation Agency (BRIN), Indonesia

Abstract
The problem of distinguishing objects has plagued researchers for many years because of low accuracy compared to human eyes' capability. In the last decade, the use of Machine Learning in aerial imagery data processing has multiplied, with the technology behind it has also developed exponentially. One of those technologies is image-based object identification, which relies heavily upon data computation. To reduce the computational load, various data segmentation algorithm was developed. This study is focused on reviewing the various image segmentation technology in aerial imagery for image recognition. Literature from as far as 1981 from various journals and conferences worldwide was reviewed. This review examines specific research questions to analyze image segmentation research over time and the challenges researchers face with each method. Machine Learning has gained popularity among segmentation methods. However, Deep Learning has been aggressively put an essential role in it by overcoming many of its weaknesses. The advanced algorithm used in Deep Learning to process the segmentation may drive more efficient and accurate data processing.

INTRODUCTION

Image segmentation divides a digital image into several image segments, called image areas or image objects (pixel sets). The objective is to change the image's representation to be simpler, more intelligible, and easier to analyze [1]. Image segmentation is customarily applied to find objects and boundaries (lines, curves, etc.) from image data. More specifically, it is the action of attaching a label to each pixel in an image so that pixels with the same label share specific properties. Segmentation extracts objects of interest for more advanced analysis, such as classification and object recognition. The technique was utilized to separate the appropriate object of interest from the image to implement a more accurate analysis, faster processing, and lighter computing resources.

There are many techniques available for segmenting images. Regarding the nature of the object to be classified, those methods could be classified into two major fields: the semantic and the instances method. The semantic process refers to actions of pixels identification, grouping, and labeling to a visual that forms an object. At the same time, the instances classify each object individually, even if the objects are from the same class. An example is given in Figure 1. To list some of the methods in image segmentation are presented in Figure 2.
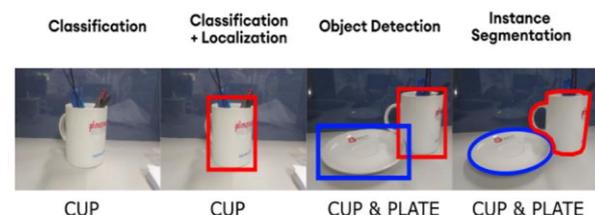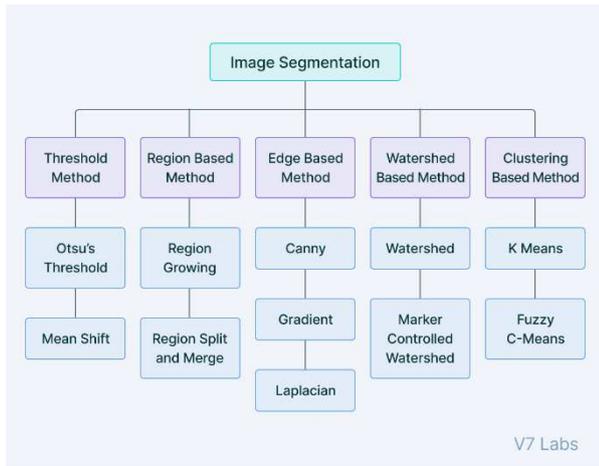

Figure 1. Example of segmentation

Figure. 2. Segmentation Methods

Table 1. Research Questions

No.	Research Questions
1.	What are the different Machine Learning-based approaches used by researchers for aerial image segmentation? What are their limitations?
2.	What are the different techniques used for image pre-processing and feature extraction?
3.	Which Machine Learning techniques are used primarily for aerial image segmentation?
4.	What are the different metrics used for evaluating the performance of the proposed Machine Learning models?
5.	How has Machine Learning-based image segmentation developed over time?
6.	What are the main challenges in implementing Machine Learning techniques in aerial imagery?

While there are already several review articles on image segmentation, very few of them discuss the specific technique used in aerial imagery data. Some discussed segmentation for medical imagery, some focused on vegetation science, and others focused on sensor and vision technology [2]–[12]. This study attempted to fill the hole in the review by focusing on image segmentation in aerial imagery technology. An extensive review of numerous sources with inclusion and exclusion criteria was done to select a suitable reference.

METHODS

This systematic review focused on various technology for aerial image segmentation, leading up to machine learning. A system has been developed to conduct this review to sort previous papers accordingly. This system is devised to maintain the review's objective to image segmentation technology for standard color (RGB) images. Also, this system aims to reduce the risk of publication bias.

The first step in the system is to frame the research questions. The relevant research publication was searched using various online databases and digital libraries. The accumulated publications were then sorted according to the inclusion and exclusion criteria.

Research Questions

The chosen methodology to be used in this review paper is by framing a few research questions. These research questions can be seen in Table 1, following the approach of Rani et al. in [13].

Search Process and Sources of Information

For this review, the authors compiled research publications mainly in segmenting images obtained from air vehicles (drone, UAV, satellite). The earliest reference was as far as 1981. This step is established to give a clear view of the development of image segmentation prior to machine learning until currently, where machine learning is heavily explored and utilized.

An online search for these publications used the keywords "image segmentation", "photogrammetry", "machine learning", "deep learning", "aerial images", "drone", "UAV", and "satellite" with an array of different combinations. Some publications that explored "classification" and "identification" were also included if they contained mentions regarding their method in image segmentation.

Inclusion and Exclusion Criteria for Article Selection

Despite the limitations set for searching the research publications related to this topic, there are still resulted in a lot of publications. To filter out unrelated publications and reduce the bias risk, the authors defined the inclusion and exclusion criteria.

The inclusion criteria are defined as:
1) Publications are written in the English language. The authors had come across a few publications written in Bahasa, Chinese, and Korean, so these are excluded purely due to the importance of internationally recognized references.
2) Publications related to image segmentation techniques for standard color (RGB and monochrome) images. Publications using images from multispectral cameras are excluded to limit the scope of the review.
3) Publications that are viewed to be able to answer or represent an answer to the research questions.

Some exclusion criteria have been mentioned briefly in the inclusion criteria. However, more exclusion criteria are:
1) Unrelated studies
2) Duplicate studies
3) Abstract-only publications
4) Publications that are unable to answer the research questions.

The flowchart of the study was as in Figure 3. Research objectives were determined, followed by the creation of exclusion and inclusion criteria for the reviewed articles. After that, a database for the selected article was established, and then an analysis of the data could be carried on. Lastly, some important points on each development step could be mentioned before concluding the study.

RESULTS AND DISCUSSION

Segmentation is a process that segments an image into parts and classifies the elements with similar traits. This process aims to separate the object in the image from the background. The classical segmentation approach consists of 4 primary methods [14]; thresholding, clustering, region-based and edge-based, which can be described as follows. The thresholding method divides the image pixels into two groups based on a threshold value; the pixels with values above the threshold value (first binary level) and pixels with values below the threshold value (second binary level). In the clustering method, the pixels of the image are organized into more than two groups; each cluster contains pixels that are considered similar.

```
┌─────────────────────────┐
│    Research Objective    │
└─────────────────────────┘
             ↓
┌─────────────────────────┐
│ Create Criteria for Article │
│        Database          │
└─────────────────────────┘
             ↓
┌─────────────────────────┐
│    Survey on Articles    │
│ Available Based on Criteria │
└─────────────────────────┘
             ↓
┌─────────────────────────┐
│    Analysis on Selected  │
│        Material          │
└─────────────────────────┘
             ↓
┌─────────────────────────┐
│ Presenting Important Point │
│  of Each Selected Article │
└─────────────────────────┘
             ↓
┌─────────────────────────┐
│     Drawing Conclusion   │
└─────────────────────────┘
```
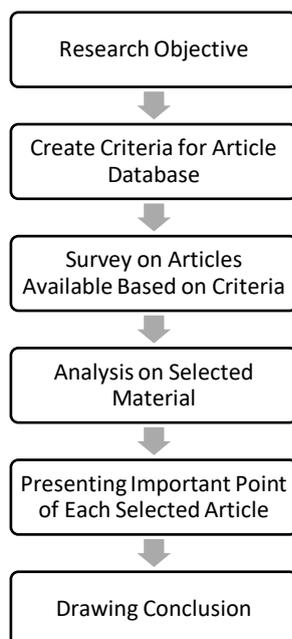
Figure 3. Course of the study

The edge-based method detects boundaries, such as points, lines, and edges, by searching for discontinuity in intensity (grey level). In Region-based segmentation, areas in the image are extracted based on specific pre-determined criteria (such as grey or color intensity).

To better understand the development of image segmentation, a thorough survey was done to track the growth of the technology. As early as 1981, Fu and Mui already categorized image segmentation into three classes: (1) characteristic highlight thresholding, (2) edge location, and (3) locale extraction [15]. Haralick and Shapiro then mentioned that segmentation techniques could be roughly classified as measurement space-guided spatial clustering, single linkage region growing schemes, hybrid linkage region growing schemes, centroid linkage region growing schemes, spatial clustering schemes, and split-and-merge schemes. Of all the methods above, the hybrid linkage schemes offer the most excellent compromise between having smooth boundaries and a few undesirable locales combined [1]. Almost a decade later, a review was done by Pal and Pal from Indian Statistical Institute in 1993. By then, it was mentioned that the literature on the segmentation of color images was not as rich as the grey tone. This research group reviewed and summarized some available techniques and attempted to cover fuzzy and non-fuzzy techniques, including segmentation for color images and neural network-based methods. The classical and fuzzy-mathematical approaches mentioned several techniques, such as histogram thresholding, edge detection, and semantic and region relaxation approaches [10].

In 1994, Skarbek and Koschan reported that Image segmentation, i.e., identification of homogeneous regions in the image, has been the subject of considerable research over the last decades. At that time, the problem of segmentation for color images, which convey much more information about objects in scenes, received much less attention from the scientific community. They report an extensive survey of algorithms for color image segmentation, categorizing them according to a well-defined list of attributes, suggestions for their improvements, and descriptions of a few novel approaches [16].

The following research on segmentation then focuses on improving segmentation methods' accuracy, precision, and computational speed and reducing the amount of manual interaction [17]. According to Pham et al., accuracy and precision can progress by consolidating earlier data from map books and combining discrete and continuous-based segmentation strategies. A parallelizable approach, such as neural networks,

could be a promising method to improve computational efficiency. A fully automatic color image segmentation was reported by Deng et al. by quantizing the data into various classes to be used to differentiate regions in the data [18]. These class labels were used to replace the pixel colors to obtain a class-map image. The step of this experiment is shown in Figure 4. A color image was quantized before the class-map was determined. After that, segmentation was carried out by marking the region to get the final result. Experiments show that JSEG provides good segmentation results on a variety of images, as shown in an example in Figure 4. The algorithm has three parameters that need to be specified by the user.

a. A limit for the color quantization process. It decides the least distance between two quantized colors.
b. The number of scales targeted for the image.
c. A limit for region merging. These parameters are essential since of the varying image characteristics in different applications.

Another research group discussed that color image segmentation was based on monochrome segmentation, which was applied to several color spaces [19]. Gray-level segmentation strategies can be directly applied to each component of a color space and then combined the result to get a final result. But, one of the issues is how to utilize the color data as an entirety for each pixel. When the color is separated into three components, it is so scattered that the color becomes essentially multispectral data, and the color data that people can understand is lost. Another issue is how to select the color representation for the segmentation. Since each color has its strengths and weaknesses, no single color representation can outperform others for segmentation purposes. Cheng et al. summarize some of the attempted research on these problems. HIS method solved it to some extent, with some issues at the hue channel, which was unstable at low saturation. Some researchers introduced physics models for color.
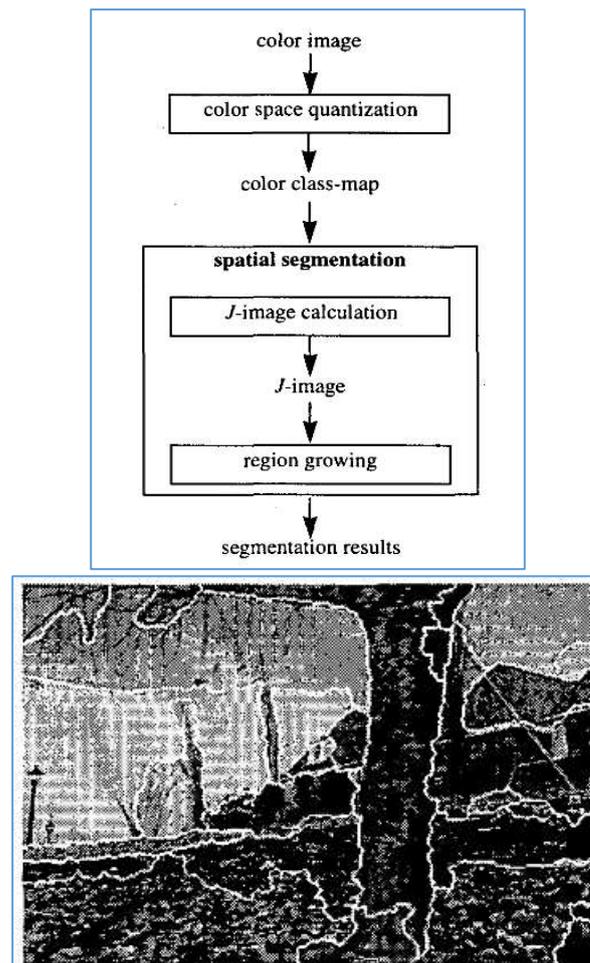

Figure 4. Schematic of JSEG by Deng et al. (top) result of JSEG (bottom)

However, by utilizing it, their many restrictions avert it from being extensively used [19]–[27].

In the case of the monochrome procedure, Cheng [14, 19, 28] made a comparison of each method in the monochrome segmentation technique and the color spaces characteristics as shown in Table 2 and Table 3, respectively. From Table 2, it is understood that the thresholding method is the simplest to execute, despite its disadvantage. Because of this, much research has been done to improve its performance, such as iterative threshold [29][30], adaptive threshold [31], and Otsu method [32]–[36]. Yanowitz & Bruckstein pioneered a technique for image segmentation by utilizing adaptive thresholding by interpolating the image grey levels at points where the gradient is high, showing the probable object edges [37]. Wang et al. tried to develop iterative thresholding for a multi-phase image. The weakness is this method depends on the Lambda constant [30]. Some methods combine the classical approach with the neural network, such as combining the region-based method with k-NN [38]. Table 3 demonstrate the characteristic of each color space with the benefit and handicap, which then opened up a greater possibility of segmentation technique by utilizing the character of the spaces. However, Cheng concludes that a universal theory applied to every color image segmentation does not exist yet. All of the segmentation methods existing for color image segmentation are ad hoc. These methods are heavily application-dependent, so there are no general algorithms and color space. On the other hand, the Fuzzy method gives an auspicious technique [19].

A research group led by Khan tried to evaluate segmentation techniques with a focus on hybrid algorithm solutions [39]. They categorize segmentation techniques into threshold-based, region-based, edge detection-based, fuzzy, and ANN-based.

Table 2. Techniques of Monochrome Segmentation [19]

Technique	Methods	Benefit	Handicap
Histogram thresholding	Inquire about several peaks of each region's data	Low computational requirement, no prior data needed	Poor result for data with no clear peaks does not count the spatial regions
Region-based approach	Pixels were grouped into regions	If the region is clear, this is the best method	Need huge computational resources
Feature clustering	Regions form separate clusters in the feature space	Easy and straightforward	No spatial information. Image dependent and challenging to determine cluster number
Edge detection	locate points with abrupt changes in the gray level	Very good for images with good contrast between each region	Noise-prone and does not perform nicely in the image with too many edges
Fuzzy	Using fuzzy operators, properties, mathematics, and inference rules to deal with various problems in segmentation	Works well in approximation approaches	Tend to have heavy computation resource
Neural Network	Apply NN in clustering or classification	Parallelable, simple	Need time to train

Table 3. Colors spaces characteristic [19]

Technique	Benefit	Handicap
RGB	Suitable for display purposes	performs poorly for segmenting color images because of a high correlation value
YIQ, YUV, I1 I2 I3	Some of the RGB correlations can be eliminated in a short amount of time. It performs well in edge detection.	It encounters the same problem as RGB, though at a lesser value.
HSI	Applicable for varying levels of illumination that create highlights, shading, and shadows in images. It can also discern different objects based on color differences.	The non-linear transformation causes low saturation that is unstable (numerically), and there exists a non-removable singularity
Nrgb	Each color component can be identified without being hindered by its brightness level to effectively counter the effect of different illumination.	At low intensities, the noise level is very high.
CIE spaces	Can process different information regarding color and intensity without any dependency. It uses geometric separation within CIE space to compare colors and can efficiently measure the difference.	There exists a non-removable singularity.

On the other hand, Martinez et al. built a color index-based thresholding method for background and foreground segmentation images. The method is based on color index measures according to the HSV-DT and GMR-T. The color index was modified to give better data. Two fixed threshold methods were proposed to differentiate foreground (green plant) and background (soil) in plant images. The first fixed threshold was a simple fixed threshold. The second fixed threshold method applies high training data to obtain its parameters from actual environmental conditions (i.e., cloudy, water reflectance, luminance variance) [40].

Lu et al. proposed an unconstrained optimization of a linear combination of RGB component images to enhance the contrast between plant and background regions, followed by automatic thresholding of the contrast-enhanced images (CEIs) [41]. This method was demonstrated to be valid by using five plant image datasets acquired under different conditions and ground-truth data. They evaluated ten general index images using CEIs to get image contrast and the accuracy of the segmentation. A more profound threshold technique was assessed by Al-Amri, using Mean Technique, Pile, HDT, and EMT on satellite images. The result was HDT and EMT were superior to other histogram techniques. In contrast, more efficient computation resources of threshold technique were observed using the TOPSO algorithm, Particle Swarm Optimization, and the 2nd Otsu algorithm [41][42]. Dunford et al. then analyzes data using pixel-based and object-oriented methods. The data were acquired using a paraglider UAV covering 174 hectares of land to identify vegetation and count the dead wood. Analysis was done on both single images and mosaics. The data quality was varied because of the aircraft movement; thus, it decreased the accuracy of the analysis significantly [42].

Karoui introduced an unsupervised image segmentation method in the regiony utilizing level set methods and texture statistics, while Zhou brought forward a mean-shift clustering algorithm [43][44]. For edge detection, Hsiou incorporated a morphological operator with region-growing. They perform image enhancement followed by the dilation residue method to find the edges [45]. Liu et al. in 2008 utilized the fast region imaging method, with the concept that all pixels with similar features should be segmented in the same region [46]. They focus on the relationship between nearby pixels instead of the whole image feature. The algorithm, as shown in Figure 5, demonstrates the processing step, which is started by the edge detection process. After that, they over-segment the filtered image to get the

primitive partitions using watershed or region-based. Based on the results of the initial partitions obtained in the second step, a k-NN graph is constructed in the third stage. This stage also includes the computation of a new region similarity measure function proposed in the algorithm section III employing local features along region edges. The last stage is the region merging process using the k-NN graph by utilizing two keys; the merging rule and the stopping rule. The algorithm used by Liu as shown in Figure 5, was proved to be able to handle both colorful and grayscale images to get the image's partition regions. However, the method was memory expensive and had various parameters which needed to be set manually. The final result of the proposed segmentation algorithm expressed in Figure 6, with a total of 7 regions are segmented in the monochrome process and a garden was successfully segmented from the surrounding area. Xuan and Hong then improved the weaknesses of the low-contrast images by improving the noise cancellation and providing a more precise image edge [47]. The development of imaging sensors in the last two decades has provided very high image data, leading to rapid advancement in the data analysis algorithm, including edge-detection segmentation techniques. The urgent need for accurate geospatial data, including the algorithm to extract the important features from the geographical view, has led to object-based image analysis (OBIA), which is constantly growing following the world's demand [48]. This method is formulated to obtain meaningful data from remote sensing imagery instead of pixels [49]–[54]. Even though OBIA consists of a two-step procedure, image segmentation and object classification, most of the publications only discuss image segmentation.

Because of this, some researchers made review articles about the OBIA method [5, 8, 55, 56] and were extensively reviewed by Kotaridis, which presents remote sensing image segmentation advances [56].
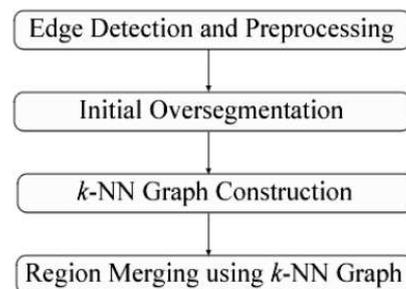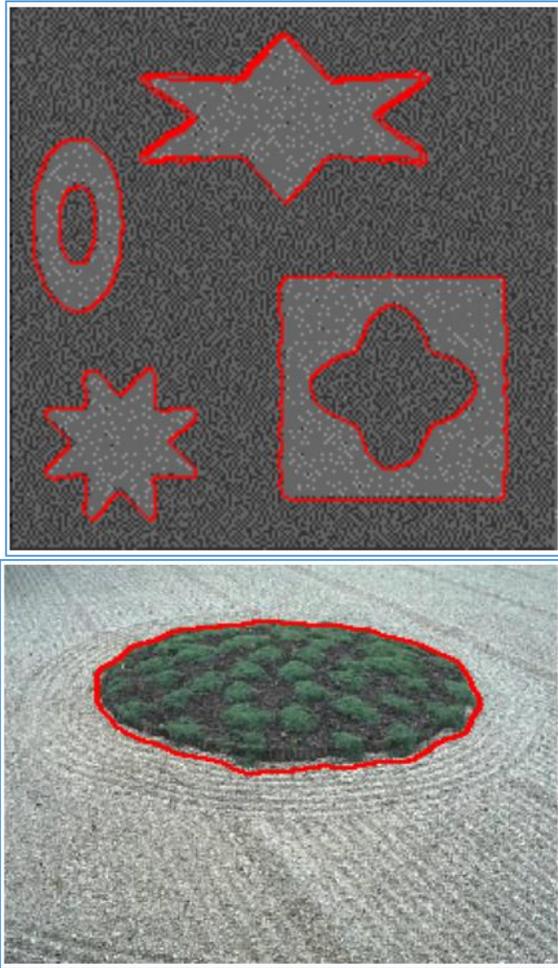

Figure 5. Flow chat of Liu et al.

Figure 6. Final result of Liu et al. in grayscale (top) and color image (bottom)

OBIA integrates knowledge from various disciplines with the target of production and utilization of geographic information. OBIA also has been applied to low spatial resolution imagery, as mentioned in [57][58].

In the last decade, a method that utilizes the growing field of Artificial Intelligence has seen a lot of potential. Deep Learning, a branch of machine learning that improves on earlier neural network systems, was explored extensively. It added more layers to the neural network, an attempt to copy the way human brains work. This method enables a system to learn from a large amount of data with improvements in optimization and accuracy. This method has been explored in many research connected to the image segmentation aerial imagery and the medical field. In this method, each pixel in an image is given a label. Researchers have explored this method for plant analysis [3, 41, 53, 59–64], monitoring the aftermath of a natural disaster[50, 65–68] and spotting areas with high water turbidity [69]. All of these researches use semantic segmentation for images obtained by a UAV or satellite. Other researchers explore different variations in the method, such as MathConvNet [70], U-Net [71], [72], TreeNet [73], and [74]. Some of the methods mentioned in semantic segmentation were Fuzzy, Bayesian network, K-nearest neighbor, Support Vector Machine, decision tree, and Random Forest. Deep Learning architecture enhanced semantic segmentation, particularly with the rise of Convolutional Neural Networks (CNN). By the robustness characteristic, Fuzzy C-Means (FCM) Clustering was the most utilized method in Fuzzy [75][92].

Later on, Puig et al. propose to simplify image analysis by reducing user interaction in supervised learning. They apply Gaussian Kernels Convolution to the RGB channel of the original image, followed by a K-means clustering algorithm. This method is often described as the soft K-means clustering method. The study case of this method was to assess the damage on sorghum plantations caused by white grub pests. Puig et al. demonstrated that their method could classify the crop field into three different areas according to the health status of the crop [76]. Popescu and Ichim built a feature selection for image segmentation using the inter-spectral co-occurrence between RGB and HSV channels by applying the LBP (local binary pattern) approach, which can describe an image's surface texture. This method has two phases: the learning phase and the segmentation phase. ROI was obtained during the learning phase, and each ROI was segmented using different features [66].

On the other hand, in the field of CNN, some researchers developed an exciting method, such as super-pixel segmentation. Yuan and Hu utilize Simple Linear Iterative Clustering [60] by utilizing a UAV to acquire forest images. Then they built the forest classifier using the input vector from a 12-dimension descriptor obtained from the super-pixels. However, the dataset was limited to only ten images, so the algorithm is not proven for a larger and condition-vary dataset. Bergado et al., on the other hand, attempt to do automatic extraction of the Spatial-Context Feature (SPF) and land cover classification using the CNN method. The data was high-resolution images with sub-decimetre resolution [77]. There are two stages in this method, a) CNN was used to learn the spatial context of the image, and b) MLP was used for the classification. This experiment shows that CNN gives better results than individual pixels classifiers or other hand-crafted methods such as GLCM and LBP. Another example of the CNN method was mentioned in [61, 62, 78–82]. Kim and his team combined U-Net with Pyramid Pooling Layers on an RGB aerial image of Seoul

City, while Rahnemoonfa et al. applied RNN dan CNN net on UAV images data of the Houston Flood. By 2019, a dual-path and lightweight DCNN was demonstrated by Zhang et al. by analyzing Vaihingen and Postdam aerial images in Germany. The dual path DCNN consists of 1) the Spatial approach, which combines a global context with multi-level features to accommodate the local and international information. 2) HED (holistically-nested edge detection) to detect boundaries for feature learning [61].

Ghassemi et al. developed a deep neural network training method using the convolutional encoder-decoder network. Their research object was satellite images. Their result was that residual encoders provide a better generalization capability than an ordinary convolutional network; furthermore, this capability increases the depth of the encoder. Experimental results also show that adapting a trained network improves its performance significantly with a lower level of complexity than training a network from scratch. The architecture used by Ghassemi et al. gives better results than other references by using two datasets with many different characteristics between training data and test data [81]. Wu et al. from the Beijing University of Posts and Telecommunications research group developed an image semantic segmentation method using the Attention Dilation-LinkNet (AD LinkNet) Neural Network. It adopts an encoder-decoder structure and serial-parallel combination dilated convolution. The goal is to enlarge the receptive field and assemble features with various scales, such as long walks and small pools.

A channel-wise attention mechanism is also designed to sharpen contextual information in satellite images and a pre-trained encoder. AD LinkNet uses three satellite datasets which are augmented using horizontal, vertical, and diagonal folding methods. After that, color augmentation was also carried out to deal with different lighting due to differences in image capture time. They conclude that for surface classification and road extraction, AD-LinkNet significantly improves segmentation accuracy. In the same year, Wurm et al. also conducted research to analyze the ability of transfer learning by using a fully convolutional network (FCN) that has been trained and applied to datasets that are different from the training dataset. They took a case study of mapping slum areas in urban areas. Their method is that the FCN was trained using image data from the QuickBird satellite. Then the FCN is transferred to be used on the low-resolution dataset from Sentinel-2 and the high-resolution dataset from TerraSAR-X. The conclusion was the experimental results show that the high geometric resolution (0.5 m) on the QuickBird image gives the best experimental results, followed by Transfer learning to the Sentinel-2 image, which significantly improves the quality and accuracy of segmentation. However, Transfer learning to TerraSAR-X images did not improve the segmentation results and even decreased the results obtained [80]. By 2020, research using the CNN method was carried out by Bhatnagar et al. by comparing the results of segmentation using Machine Learning (ML) and Deep Learning (DL). The ML method uses Random Forest as a pixel classifier and Graph Cut for image segmentation, while the DL method uses CNN with a combination of ResNet50 and SegNet architectures. As a result, segmentation using ML has an accuracy value close to 85%, while DL has an accuracy value close to 90%. However, more labeled data (ground truth) is needed to train the system, more computation time, and higher hardware requirements to perform segmentation using the DL method. The conclusion that can be drawn is that by looking at the comparison of needs and the value of the difference in accuracy, it can be concluded that the ML method is more profitable to use. This is because to map the bog area, the network needs to be trained for each area, topography, season, and other natural conditions, which will be very inconvenient [62].

On the other hand, Valente et al. developed an automated method to count the number of plants using high-resolution photos from UAVs with a case study for spinach plants. The technique used is machine vision, namely the Excess Green Index and the Otsu method, and transfer learning uses CNN to identify and count the number of plants. Two shots were taken for each data collection area to produce two original images with different spatial resolutions (8 and 16 mm/pixel). This is done to optimize the computation time by reducing the image size. This method is validated using ground truth data covering an area of 1/8. They concluded that the calculation accuracy value is 95% for a spatial resolution of 8 mm/pixel in 172 m2. However, for a land area of 3.2 hectares, this method gives an error of 42.5% [82].

Ulmas & Liiv also use convolutional machine learning with a modified U-Net structure to map land cover on satellite images. The data used is from the BigEarthNet satellite. The conclusion obtained by this team is that CNN is possible to be used in satellite data segmentation. Some notes on their method are that the land cover category needs to be changed to get high-accuracy results. The improvised U-Net model proved to give better results for low-resolution maps. The presence of noise and mislabels in

BigEarthNET data needs to be considered because this greatly affects the results [72]. Lu et al. attempted to retrieve plant physical data (canopy area, width, position) by segmenting the image using Naïve Bayes with a case study using an apple orchard. Their method is that the original image is first pre-processed using SLRCM (Shadow Region Luminance Compensation Method) to fix the lighting in the image. The performance results of the proposed method give a precision rate of 95.30%, a recall rate of 84.45%, and an F1-score of 89.53%. The quality of the segmented image is also better than the results using the usual GMM or K-means algorithm [83]. There is also research to use a VGG-based fully CNN to map flooded areas using drones by Gebrehiwot et al., with their finding that a pre-trained and fine-tuned model may result in superior performance, specifically when domain-specific training data are scarce [84][85]. Yang et al. then introduced Mask R-CNN-based model by training it on Google Earth imagery, gaining superior performance compared to ResNet-100 [86]. In the same year, Guo et al. developed a way to extract water out of GaoFen-1 data in a multi-scale dimension with comparable results to U-Net and DeepLab-V3+ [87]. Water edge detection was detailly obtained by Pan et al. by incorporating a modified U-Net with a super-pixel method. Furthermore, in high-resolution imagery [88], Wieland et al. successfully separated normal water from flood water in various scenarios by using the fuse of U-Net, DeepLab-V3+ with Mobilenet-V3, ResNet-50, and EfficientNet-B4 [58]. Meanwhile, Li and Cheng, in late 2022 combined CNN and a transformer method to try to enhance the segmentation result of DeepLabV3+ and UCTransNet [89]. Tang et al made distinctive research to solve a massive need for ground truth data by utilizing mCL LC as the argument to determine various image data and use it in the encoder-decoder structure of CNN [90]. Chen et al made a review on satellite data imagery segmentation. However, unfortunately, only nine papers were reviewed [86].

With all of the mentioned papers, the trend of research in the field of segmentation was shown with the reach of 12k publications regarding segmentation in Scopus-indexed articles, as presented in Figure 7 by providing "image segmentation" as the searched word in the article title, abstract, and keywords. With the keywords "image" and "machine learning", the trend shows an aggressive increase from 2016 in the publication using machine learning as their base image processing method, as shown in Figure 8.

All in all, we have reviewed 149 references comprised of reports, articles, conference papers, and books from 1982 to 2022 related to image segmentation, citing 78 articles regarded as important advances in this field. From the reviewed articles, the data in Figure 8 shows that machine learning is gaining popularity among researchers. Specifically, deep learning has obtained an important breakthrough in solving many problems in image segmentation. In our reference, CNN is still the most popular method, while at the same time, OBIA is caching up aggressively, as shown in Figure 9. From the viewpoint of the reviewed reference type, there are five books cited in this article. In contrast, 17% of the references were conference papers from various fields, more than a hundred journal articles, and a couple of thesis and reports, as shown in Figure 10.

. In terms of application, the reviewed paper consisted of applications in Earth and Planetary Science as the majority, with 28%, followed by Vegetation Science, with 23%. Other fields were Vision, Medical, mathematics, Archaeology, and Material Science. Figure 11 displays the proportion of the reviewed article's field, while Figure 12 expresses the comparison of the aerial imagery articles' data source. Of the Aerial data from the reviewed article, there are 63% of images from satellites and 37% from UAVs.
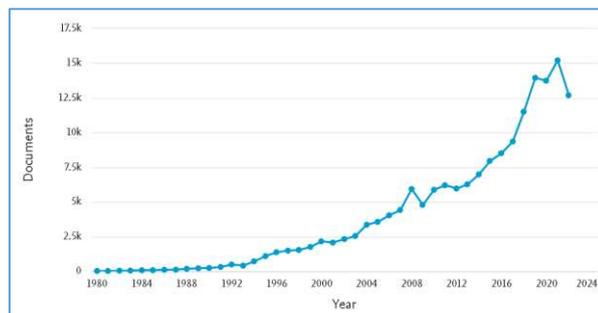

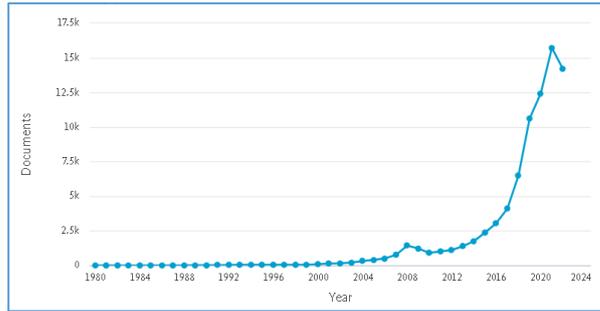Figure 7. Publication trend in image segmentation
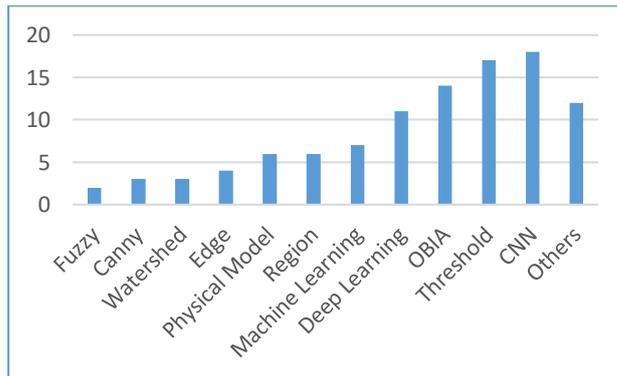
Figure 8. Publication trend in Machine Learning

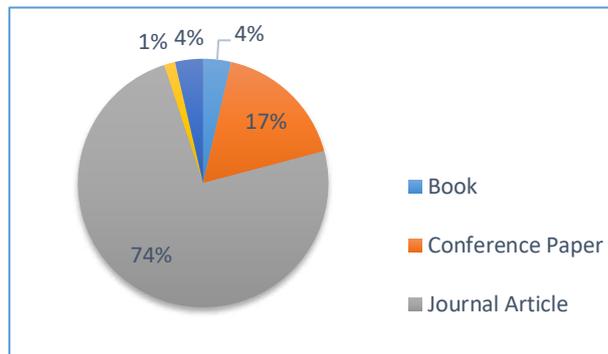Figure 9. Segmentation method reviewed in this study
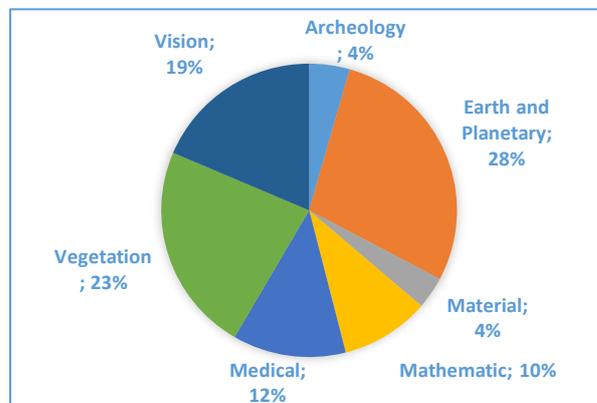

Figure 10. Type of reference's source

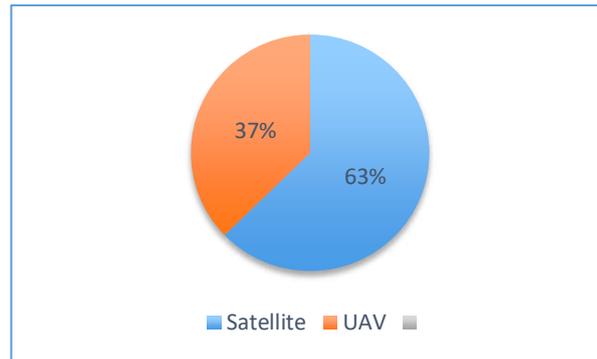Figure 11. Application field of the reviewed articles

Figure 12. Comparison between UAV and Satellite application of the reviewed articles

Some of the exciting studies in Machine Learning in aerial imagery were done, for example, by Heffels et al. in 2020. They successfully solved popular datasets of DroneDeploy with an accuracy of IOU 70%. This number is the highest publicly available validation set. They use DeepLabv3+ Xception65 architecture [91] to employ U-net neural network architecture. Figure 13 shows how ground truth labeling was done. Every object as cars, buildings, and vegetation, was labeled and then overlayed on the original picture with the mask's transparency of 50%. This ground truth labeling was then used to train the system. The conclusion of their study was augmentation, batch normalization, and the ResNet50 encoder enhanced the segmentation using the DeepLabv3+ Xception65 architecture. An example of the result is in Figure 14, which is an overlayed image predicted label by Keras CCE ResNet50 U-net baseline.

Table 4. Summary of Methods

Research Questions	Findings
1. What are the different Machine Learning-based approaches used by researchers for aerial image segmentation? What are their limitations?	Method: CNN, RCNN, DCNN, GLCM, LBP, U-Net, FCN, MathConvNet, TreeNet, Bayesian-Net, AD-LinkNet, K-nearest neighbor, SVM, Decision Tree, Random Forest, Fuzzy C-Means (FCM), Gaussian Kernels Convolution, SLIC, SPF-CNN, SLRCM, ... Limit: There is no single algorithm able to serve as an all-purpose method. Each and every method has its own benefit and weaknesses. Mostly, by combining different methods, the optimum segmentation purposes will be achieved Ground truth data is scarce in to train model
2. What are the different techniques used for image pre-processing and feature extraction?	Histogram thresholding, Region limitation, Feature clustering, Edge separation, Super-pixel
3. Which Machine Learning techniques are used primarily for aerial image segmentation?	CNN and its variation, U-Net, SVM
4. What are the different metrics used for evaluating the performance of the proposed Machine Learning models?	Performance is measured by accuracy percentage compared to ground truth data. Mostly in mIoU (mean Intersection over Union) dan mPA (mean Pixel Accuracy)
5. How has Machine Learning-based image segmentation developed over time?	From monochrome imagery to complex inter-spectral interaction data acquisition, algorithms developed to be more complex and demanding. However, the advance in hardware technology support it to accelerate the development over time.
6. What are the main challenges in implementing Machine Learning techniques in aerial imagery?	• To acquire a single algorithm that can serve as an all-purpose method. Every method has its benefit and weaknesses, and every segmentation process has its purposes. Mostly, by combining different approaches, the optimum expected result will be acquired. • To acquire enough data to train the model in various environments and requirements so that the algorithm could get the optimum result from a certain dataset

Figure 13. (Left) Original picture, (Center) labeling, (Right) overlay label on top of the original picture [91]
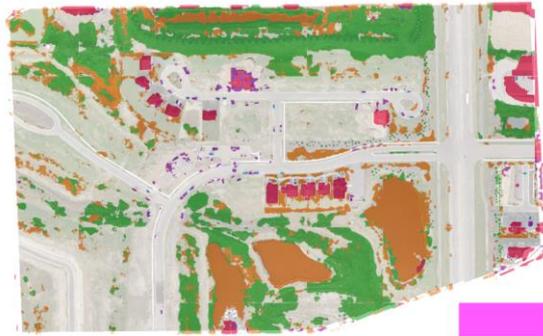

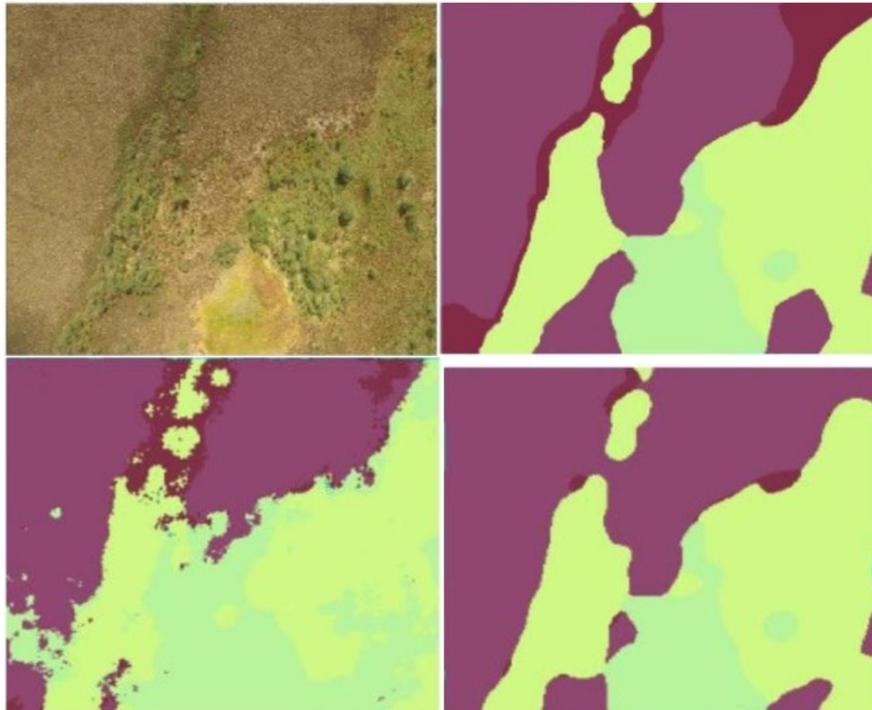Figure 13. The Drone Deploy dataset overlaid with 65% transparency [91]

Figure 14. (top-left) Drone image, (top-right) ground truth labeled image, (bottom-left) DL semantic segmentation using U-Net with ResNet50 model, (bottom-right) DL semantic segmentation using SegNet and ResNet50 model [62]

Bhatnagar et al., on the other hand, apply a deep-learning algorithm. CNN architecture was utilized with the combination of esNet50 and SegNet [63]. The study shows that in comparing six different state-of-the-art pixel-based classifiers, the best ML algorithm for the given dataset was shown to be the RF classifier with an accuracy of 92% on Clara Bog drone images. The ResNet50 base model with both U-Net reached 91.5% and SegNet 89.9%. An example of their images is shown in Figure 15.

CONCLUSION

Image segmentation from aerial imagery taken with satellites, aircraft, and UAVs has been essential for various purposes. For civilian purposes, it plays a big part in disaster mitigation and data gathering for crops and forestry, amongst other things. It is inexpensive when no specific camera, such as hyperspectral, is involved. Depending on the area of observation, images from satellites are preferred when the area is immense. While images from UAVs are chosen for a smaller area. However, aerial images have some disadvantages as they are very susceptible to the lighting condition when the images were taken, cloud cover, etc. The time of day can influence whether certain areas fall under the shadows of certain mountains, clouds, or buildings. Particular objects are also possible to appear similar from a distance, such as a river and a road. There is also the problem of low contrast, such as the difference between a field of grass and a forest. For years, only the eyes of a human can discern between them as it is difficult to define in an algorithm to differentiate it.

The vision technology research to recognize objects started as early as the 1980s. However, a breakthrough was apparent after Machine Learning was introduced. It offers a more straightforward way to solve many problems in the field of image segmentation as long as an abundance of images and sufficient computational resources are available. A system can then be trained to segment images, extracting features from them and classifying them using a method similar to training a human observer. Even though its early development was plagued with limitations due to the hardware technology, resulting in the preference for using grayscale images, the advanced development of hardware and algorithm was able to counter some of the problems. Methods such as FCN and CNN, determined by the layers they used in training, provide better results than the conventional methods, though still not without some challenges. The summary of our findings is presented in Table 4. The deep learning method made a more drastic improvement over

time when its performance overcame many of its machine-learning counterparts. It uses much more layers as it exploits the advances of the current hardware technology. The challenge is to get enough ground truth data so that the algorithm can be trained sufficiently to extract the optimum information out of a dataset. Deep Learning can potentially be used as a major segmentation method in image recognition in various fields that were once only ventured by using the hyperspectral camera.

REFERENCES

[1] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 1, pp. 100–132, Jan. 1985, doi: 10.1016/S0734-189X(85)90153-7.

[2] K. K. D. Ramesh, G. Kiran Kumar, K. Swapna, D. Datta, and S. Suman Rajest, "A Review of Medical Image Segmentation Algorithms," *EAI Endorsed Transaction on Pervasive Health Technology*, vol. 7, no. 27, 2021, doi: 10.4108/eai.12-4-2021.169184.

[3] S. S. Lomte and A. P. Janwale, "Plant Leaves Image Segmentation Techniques: A Review," *International Journal of Computer Sciences and Engineering*, vol. 5, no. 5, 2017.

[4] P. Ganesan, B. S. Sathish, K. Vasanth, V. G. Sivakumar, M. Vadivel and C. N. Ravi, "A Comprehensive Review of the Impact of Color Space on Image Segmentation," *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, Coimbatore, India, 2019, pp. 962-967, doi: 10.1109/ICACCS.2019.8728392.

[5] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A Review of Supervised Object-based Land-cover Image Classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 277–293, Aug. 01, 2017. doi: 10.1016/j.isprsjprs.2017.06.001.

[6] M. R. S. Zawad, M. F. S. Zawad, M. A. Rahman, and S. N. Priyom, "A Comparative Review of Image Processing Based Crack Detection Techniques on Civil Engineering Structures," *Journal of Soft Computing in Civil Engineering*, vol. 5, no. 3, pp. 58–74, 2021, doi: 10.22115/SCCE.2021.287729.1325.

[7] A. Wang, W. Zhang, and X. Wei, "A Review on Weed Detection Using Ground-based Machine Vision and Image Processing Techniques," *Computers and Electronics in Agriculture*, vol. 158, pp. 226–240, Mar. 01, 2019. doi: 10.1016/j.compag.2019.02.005.

[8] M. D. Hossain and D. Chen, "Segmentation for Object-Based Image Analysis (OBIA): A Review of Algorithms and Challenges from

Remote Sensing Perspective," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 115–134, Apr. 01, 2019. doi: 10.1016/j.isprsjprs.2019.02.009.

[9] X. Liu, Z. Deng, and Y. Yang, "Recent Progress in Semantic Image Segmentation," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 1089–1106, Aug. 2019, doi: 10.1007/s10462-018-9641-3.

[10] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, Sep. 1993, doi: 10.1016/0031-3203(93)90135-J.

[11] S. Saifullah, "K-Means Clustering for Egg Embryo's Detection Based-On Statistical Feature Extraction Approach of Candling Eggs Image," *SINERGI*, vol. 25, no. 1, pp. 43-50, Nov. 2020, doi: 10.22441/sinergi.2021.1.006.

[12] R. Lionnie, R. C. Ramadhan, A. S. Rosyadi, M. Jusoh, and M. Alaydrus, "Performance analysis of various types of surface crack detection based on image processing," *SINERGI*, vol. 26, no. 1, pp. 1-6, Feb. 2022, doi: 10.22441/sinergi.2022.1.001.

[13] P. Rani, S. Kotwal, J. Manhas, V. Sharma, and S. Sharma, "Machine Learning and Deep Learning Based Computational Approaches in Automatic Microorganisms Image Recognition: Methodologies, Challenges, and Developments," *Archives of Computational Methods in Engineering*, vol. 29, no. 3, pp. 1801–1837, May 01, 2022. doi: 10.1007/s11831-021-09639-x.

[14] F. U. Siddiqui and A. Yahya, *Clustering Techniques for Image Segmentation*. Springer International Publishing, 2022. doi: 10.1007/978-3-030-81230-0.

[15] K. S. Fu and J. K. Mui, "A survey on image segmentation," *Pattern Recognition*, vol. 13, no. 1, pp. 3–16, Jan. 1981, doi: 10.1016/0031-3203(81)90028-5.

[16] W. Skarbek and A. Koschan, "*Colour Image Segmentation - A Survey -*," Technische Universitat Berlin, Berlin, 1994.

[17] D. L. Pham, Y. Xu, J. L. Prince, and C. Xu, "*A Survey of Current Methods in Medical Image Segmentation*," Baltimore, Jan. 1998.

[18] Y. Deng, B. S. Manjunath, and H. Shin, "*Color Image Segmentation*," Santa Barbara, 1999.

[19] H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, no. 12, pp. 2259–2281, Dec. 2001, doi: 10.1016/S0031-3203(00)00149-7.

[20] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985, doi: 10.1002/col.5080100409.

[21] G. Healey, "Using color for geometry-insensitive segmentation," *Journal of the Optical Society of America A*, vol. 6, no. 6, p. 920, 1989, doi: 10.1364/josaa.6.000920.

[22] G. Healey, "Segmenting Images Using Normalized Color," *IEEE Trans Syst Man Cybern*, vol. 22, no. 1, pp. 64–73, 1992, doi: 10.1109/21.141311.

[23] A. P. Petrov and L. L. Kontsevich, "Properties of color images of surfaces under multiple illuminants," *Journal of the Optical Society of America A*, vol. 11, no. 10, p. 2745, 1994, doi: 10.1364/josaa.11.002745.

[24] G. J. Klinker, S. A. Shafer, and T. Kanade, "Image Segmentation and Reflection Analysis Through Color," *Applications of Artificial Intelligence VI*, vol. 0937, p. 229, 1988, doi: 10.1117/12.946980.

[25] B. A. Maxwell and S. A. Shafer, "Physics-Based Segmentation: Moving Beyond Color," in *Computer Vision and Pattern Recognition*, 1996, pp. 742–749.

[26] B. A. Maxwell and S. A. Shafer, "*Physics-Based Segmentation: Looking Beyond Color*," Carnegie Mellon, 1996.

[27] S. L. Guth, "Model for color vision and light adaptation," *Journal of the Optical Society of America A*, vol. 8, no. 6, 1991, doi: 10.1364/josaa.8.000976.

[28] W. Khan, "Image Segmentation Techniques: A Survey," *Journal of Image and Graphics*, vol. 1, no. 4, pp. 166–170, 2014, doi: 10.12720/joig.1.4.166-170.

[29] J. H. Syu, S. J. Wang, and L. C. Wang, "Hierarchical Image Segmentation Based on Iterative Contraction and Merging," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2246–2260, May 2017, doi: 10.1109/TIP.2017.2651395.

[30] D. Wang, H. Li, X. Wei, and X. P. Wang, "An Efficient Iterative Thresholding Method for Image Segmentation," *Journal of Computational Physics*, vol. 350, pp. 657–667, Dec. 2017, doi: 10.1016/j.jcp.2017.08.020.

[31] L. Xuan and Z. Hong, "An improved canny edge detection algorithm," *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, vol. 2017-Novem, pp. 275–278, 2018, doi: 10.1109/ICSESS.2017.8342913.

[32] B. Bayram *et al.*, "Automatic extraction of sparse trees from high-resolution ortho-images," *Arabian Journal of Geosciences*, vol. 11, no. 12, 2018, doi: 10.1007/s12517-018-3680-6.

[33] C. Huang, X. Li, and Y. Wen, "An OTSU Image Segmentation Based on Fruitfly Optimization Algorithm," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 183–188, Feb. 2021, doi: 10.1016/j.aej.2020.06.054.

[34] T. Y. Goh, S. N. Basah, H. Yazid, M. J. Aziz Safar, and F. S. Ahmad Saad, "Performance Analysis of Image Thresholding: Otsu Technique," *Measurement*, vol. 114, pp. 298–307, Jan. 2018, doi: 10.1016/j.measurement.2017.09.052.

[35] A. Akagic, E. Buza, S. Omanovic, and A. Karabegovic, "Pavement Crack Detection Using Otsu Thresholding for Image Segmentation," in *MIPRO 2018*, 2018.

[36] Y. Feng, H. Zhao, X. Li, X. Zhang, and H. Li, "A multi-scale 3D Otsu thresholding algorithm for medical image segmentation," *Digital Signal Processing*, vol. 60, pp. 186–199, Jan. 2017, doi: 10.1016/J.DSP.2016.08.003.

[37] S. D. Yanowitz and A. M. Bruckstein, "A new method for image segmentation," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 1, pp. 82–95, Apr. 1989, doi: 10.1016/S0734-189X(89)80017-9.

[38] H. Liu, Q. Guo, M. Xu, and I. F. Shen, "Fast image segmentation using region merging with a k-Nearest neighbor graph," *2008 IEEE International Conference on Cybernetics and Intelligent Systems, CIS 2008*, no. c, pp. 179–184, 2008, doi: 10.1109/ICCIS.2008.4670856.

[39] W. Khan, "Image Segmentation Techniques: A Survey," *Journal of Image and Graphics*, pp. 166–170, 2013, doi: 10.12720/joig.1.4.166-170.

[40] M. Castillo-Martínez, F. J. Gallegos-Funes, B. E. Carvajal-Gámez, G. Urriolagoitia-Sosa, and A. J. Rosales-Silva, "Color Index Based Thresholding Method for Background and Foreground Segmentation of Plant Images," *Computers and Electronics in Agriculture*, vol. 178, Nov. 2020, doi: 10.1016/j.compag.2020.105783.

[41] Y. Lu, S. Young, H. Wang, and N. Wijewardane, "Robust Plant Segmentation of Color Images Based on Image Contrast Optimization," *Computers and Electronics in Agriculture*, vol. 193, Feb. 2022, doi: 10.1016/j.compag.2022.106711.

[42] R. Dunford, K. Michel, M. Gagnage, H. Piégay, and M. L. Trémelo, "Potential and Constraints of Unmanned Aerial Vehicle Technology for the Characterization of Mediterranean Riparian Forest," *International Journal of Remote Sensing*, vol. 30, no. 19,

pp. 4915–4935, 2009, doi: 10.1080/01431160903023025.

[43] I. Karoui, R. Fablet, J. M. Boucher, and J. M. Augustin, "Region-based Image Segmentation using Texture Statistics and Level-Set Methods," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2006, pp. 693–696. doi: 10.1109/icassp.2006.1660437.

[44] Y. M. Zhou, S. Y. Jiang, and M. L. Yin, "A Region-Based Image Segmentation Method with Mean-Shift Clustering Algorithm," *Proceedings - 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008*, vol. 2, pp. 366–370, 2008, doi: 10.1109/FSKD.2008.363.

[45] Y. T. Hsiao, C. L. Chuang, J. A. Jiang, and C. C. Chien, "A Contour Based Image Segmentation Algorithm Using Morphological Edge Detection," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2005, pp. 2962–2967. doi: 10.1109/icsmc.2005.1571600.

[46] H. Liu, Q. Guo, M. Xu and I. Shen, "Fast image segmentation using region merging with a k-Nearest Neighbor graph," *2008 IEEE Conference on Cybernetics and Intelligent Systems*, Chengdu, 2008, pp. 179-184, doi: 10.1109/ICCIS.2008.4670856.

[47] W. Rong, Z. Li, W. Zhang and L. Sun, "An improved Canny edge detection algorithm," *2014 IEEE International Conference on Mechatronics and Automation*, Tianjin, China, 2014, pp. 577-582, doi: 10.1109/ICMA.2014.6885761.

[48] S. Lang, "Object-based image analysis for remote sensing applications: modeling reality-dealing with complexity," in *Object-Based Image Analysis*, Springer, 2008. [Online]. Available: http://www.unccd.int/

[49] T. Blaschke *et al.*, "Geographic Object-Based Image Analysis - Towards a New Paradigm," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 180–191, 2014, doi: 10.1016/j.isprsjprs.2013.09.014.

[50] N. Chehata, C. Orny, S. Boukir, D. Guyon, and J. P. Wigneron, "Object-based Change Detection in Wind Storm-damaged Forest Using High-resolution Multispectral Images," *International Journal of Remote Sensing*, vol. 35, no. 13, pp. 4758–4777, 2014, doi: 10.1080/01431161.2014.930199.

[51] S. Derivaux, G. Forestier, C. Wemmert, and S. Lefvre, "Supervised Image Segmentation Using Watershed Transform, Fuzzy Classification and Evolutionary Computation," *Pattern Recognition Letters*,

vol. 31, no. 15, pp. 2364–2374, Nov. 2010, doi: 10.1016/j.patrec.2010.07.007.

[52] K. Navulur, *Multispectral Image Analysis Using the Object-Oriented Paradigm*. CRC Press, 2006. doi: 10.1201/9781420043075.

[53] L. M. Moskal, D. M. Styers, and M. Halabisky, "Monitoring Urban Tree Cover Using Object-Based Image Analysis and Public Domain Remotely Sensed Data," *Remote Sensing (Basel)*, vol. 3, no. 10, pp. 2243–2262, Oct. 2011, doi: 10.3390/rs3102243.

[54] G. Hay and G. Castilla, "Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT)," in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVIII-4/C42*, Salzburg: Salzburg University, 2006. Accessed: Sep. 21, 2022. [Online]. Available: https://www.isprs.org/proceedings/xxxvi/4-c42/info.aspx

[55] G. Chen, Q. Weng, G. J. Hay, and Y. He, "Geographic Object-based Image Analysis (GEOBIA): Emerging Trends and Future Opportunities," *GIScience and Remote Sensing*, vol. 55, no. 2, pp. 159–182, Mar. 2018, doi: 10.1080/15481603.2018.1426092.

[56] I. Kotaridis and M. Lazaridou, "Remote Sensing Image Segmentation Advances: A Meta-Analysis," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173. Elsevier B.V., pp. 309–322, Mar. 01, 2021. doi: 10.1016/j.isprsjprs.2021.01.020.

[57] O. Csilik and M. Belgiu, "Cropland Mapping from Sentinel-2 Time Series Data Using Object-based Image Analysis," in *AGILE 2017*, 2017.

[58] M. Wieland, Y. Li, and S. Martinis, "Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network," *Remote Sensing of Environment*, vol. 230, Sep. 2019, doi: 10.1016/j.rse.2019.05.022.

[59] A. Safonova, E. Guirado, Y. Maglinets, D. Alcaraz-Segura, and S. Tabik, "Olive tree biovolume from uav multi-resolution image segmentation with mask r-cnn," *Sensors*, vol. 21, no. 5, pp. 1–17, 2021, doi: 10.3390/s21051617.

[60] Y. Yuan and X. Hu, "Random Forest and Object-based Classification for Forest Pest Extraction from UAV Aerial Imagery," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 2016-Janua, no. July, pp. 1093–1098, 2016, doi: 10.5194/isprsarchives-XLI-B1-1093-2016.

[61] S. Zhang, Z. You, and X. Wu, "Plant Disease Leaf Image Segmentation Based on Superpixel Clustering and EM Algorithm," *Neural Computing and Applications*, vol. 31, pp. 1225–1232, Feb. 2017, doi: 10.1007/s00521-017-3067-8.

[62] S. Bhatnagar, L. Gill, and B. Ghosh, "Drone Image Segmentation Using Machine and Deep Learning for Mapping Raised Bog Vegetation Communities," *Remote Sensing (Basel)*, vol. 12, no. 16, Aug. 2020, doi: 10.3390/RS12162602.

[63] M. Gomez Selvaraj *et al.*, "Detection Of Banana Plants and Their Major Diseases Through Aerial Images and Machine Learning Methods: A Case Study in DR Congo And Republic of Benin," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 110–124, Nov. 2020, doi: 10.1016/j.isprsjprs.2020.08.025.

[64] W. Wang, L. Zhang, J. Wang, and Z. Long, "Locating Defects and Image Pre-processing: Deep Learning in Automated Tobacco Production," *Journal of Sensors*, vol. 2022, pp. 1–13, Aug. 2022, doi: 10.1155/2022/6797207.

[65] T. Chowdhury, M. Rahnemoonfar, R. Murphy, and O. Fernandes, "Comprehensive Semantic Segmentation on High Resolution UAV Imagery for Natural Disaster Damage Assessment," *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 3904–3913, 2020, doi: 10.1109/BigData50022.2020.9377916.

[66] D. Popescu, L. Ichim, and T. Caramihale, "Flood Areas Detection Based on UAV Surveillance System," in *19th International Conference on System Theory, Control and Computing (ICSTCC)*, 2015, pp. 14–16.

[67] M. Rahnemoonfar, R. Murphy, M. V. Miquel, D. Dobbs, and A. Adams, "Flooded Area Detection From UAV Images Based on Densely Connected Recurrent Neural Networks," in *IGARSS 2018*, IEEE, 2018.

[68] A. Gupta, S. Watson, and H. Yin, "Deep Learning-based Aerial Image Ssegmentation with Open Data for Disaster Impact Assessment," *Neurocomputing*, vol. 439, pp. 22–33, Jun. 2021, doi: 10.1016/j.neucom.2020.02.139.

[69] D. S. Bona, A. Murni, and P. Mursanto, "Semantic segmentation and segmentation refinement using machine learning case study: Water turbidity segmentation," *Proceedings of the 2019 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology, ICARES 2019*,

pp. 1–5, 2019, doi: 10.1109/ICARES.2019. 8914551.

[70] L. Ichim and D. Popescu, "Road Detection and Segmentation from Aerial Images Using a CNN Based System," *2018 41st International Conference on Telecommunications and Signal Processing, TSP 2018*, pp. 1–5, 2018, doi: 10.1109/TSP.2018.8441366.

[71] J. H. Kim *et al.*, "Objects Segmentation from High-Resolution Aerial Images Using U-Net with Pyramid Pooling Layers," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, pp. 115–119, 2019, doi: 10.1109/LGRS.2018.2868880.

[72] P. Ulmas and I. Liiv, "Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification," 2020, doi: 10.1109/ACCESS.2017.

[73] K. Yue, L. Yang, R. Li, W. Hu, F. Zhang, and W. Li, "TreeUNet: Adaptive Tree convolutional neural networks for subdecimeter aerial image segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 156, no. July, pp. 1–13, 2019, doi: 10.1016/j.isprsjprs.2019.07.007.

[74] G. Zhang, T. Lei, Y. Cui, and P. Jiang, "A dual-path and lightweight convolutional neural network for high-resolution aerial image segmentation," *ISPRS International Journal of Geo-Information*, vol. 8, no. 12, 2019, doi: 10.3390/ijgi8120582.

[75] S. Naz, H. Majeed and H. Irshad, "Image segmentation using fuzzy clustering: A survey," *2010 6th International Conference on Emerging Technologies (ICET)*, Islamabad, Pakistan, 2010, pp. 181-186, doi: 10.1109/ICET.2010.5638492.

[76] E. Puig, F. Gonzalez, G. Hamilton, and P. Grundy, "*Assessment of Crop Insect Damage Using Unmanned Aerial Systems: A Machine Learning Approach,*" 2015. [Online]. Available: www.mssanz.org.au/modsim2015

[77] J. R. Bergado, C. Persello and C. Gevaert, "A deep learning approach to the classification of sub-decimetre resolution aerial images," *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, China, 2016, pp. 1516-1519, doi: 10.1109/IGARSS.2016.7729387.

[78] J. H. Kim *et al.*, "Objects Segmentation from High-Resolution Aerial Images Using U-Net with Pyramid Pooling Layers," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, pp. 115–119, Jan. 2019, doi: 10.1109/LGRS.2018.2868880.

[79] M. Rahnemoonfa, R. Murphy, M. V. Miquel, D. Dobbs, and A. Adams, "Review paper: Flooded Area Detection from UAV Images Based on Densely Connected Recurrent Neural Networks," 2018.

[80] M. Wurm, T. Stark, X. X. Zhu, M. Weigand, and H. Taubenböck, "Semantic Segmentation of Slums in Satellite Images Using Transfer Learning on Fully Convolutional Neural Networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 59–69, Apr. 2019, doi: 10.1016/j.isprsjprs.2019.02.006.

[81] S. Ghassemi, A. Fiandrotti, G. Francini, and E. Magli, "Learning and Adapting Robust Features for Satellite Image Segmentation on Heterogeneous Data Sets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6517–6529, Sep. 2019, doi: 10.1109/TGRS.2019.2906689.

[82] J. Valente, B. Sari, L. Kooistra, H. Kramer, and S. Mücher, "Automated Crop Plant Counting from Very High-resolution Aerial Imagery," *Precision Agriculture*, vol. 21, no. 6, pp. 1366–1384, Dec. 2020, doi: 10.1007/s11119-020-09725-3.

[83] Y. Lu, S. Young, H. Wang, and N. Wijewardane, "Robust Plant Segmentation of Color Images Based on Image Contrast Optimization," *Computers and Electronics in Agriculture*, vol. 193, Feb. 2022, doi: 10.1016/j.compag.2022.106711.

[84] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, "Deep convolutional neural network for flood extent mapping using unmanned aerial vehicles data," *Sensors (Switzerland)*, vol. 19, no. 7, 2019, doi: 10.3390/s19071486.

[85] L. Hashemi-Beni and A. A. Gebrehiwot, "Flood Extent Mapping: An Integrated Method Using Deep Learning and Region Growing Using UAV Optical Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, 2021, doi: 10.1109/JSTARS.2021.3051873.

[86] Y.-N. Chen, K.-C. Fan, Y.-L. Chang, and T. Moriyama, "Special Issue Review: Artificial Intelligence and Machine Learning Applications in Remote Sensing," *Remote Sensing (Basel)*, vol. 15, no. 3, p. 569, Jan. 2023, doi: 10.3390/rs15030569.

[87] G. Guo and L. Sun, "Image Retrieval Technology of Economic Regulations Based on Semantic Segmentation," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–13, Aug. 2022, doi: 10.1155/2022/6799899.

[88] F. Pan, Z. Lu, D. Chen, and Di. Xue, "An optic disk semantic segmentation method based on weakly supervised learning," in *Proceedings of the 32nd Chinese Control and*

Decision Conference, CCDC 2020, 2020. doi: 10.1109/CCDC49329.2020.9164841.

[89] Y. Li, Z. Cheng, C. Wang, J. Zhao, and L. Huang, "RCCT-ASPPNet: Dual-Encoder Remote Image Segmentation Based on Transformer and ASPP," *Remote Sensing (Basel)*, vol. 15, no. 2, Jan. 2023, doi: 10.3390/rs15020379.

[90] K. Yuan *et al.*, "A multi-strategy contrastive learning framework for weakly supervised semantic segmentation," *Pattern Recognition*, vol. 137, p. 109298, May 2023, doi: 10.1016/J.PATCOG.2022.109298.

[91] M. R. Heffels and J. Vanschoren, "Aerial Imagery Pixel-level Segmentation," Dec. 2020, doi: 10.48550/arxiv.2012.02024.

[92] A. Adriansyah and S. H. M. Amin, "Learning of fuzzy-behaviours using Particle Swarm Optimisation in behaviour-based mobile robot," *International Journal of Intelligent Systems Technologies and Applications*, vol. 5, no. 1-2, pp. 49-67, 2008, doi: 10.1504/IJISTA.2008.018166