

Resolution–Accuracy Trade-offs in UAV-Based Semantic Segmentation for Precision Agricultural Imagery

Mohamed Tawhid Amin¹, Aziza I. Hussein², Mohamed Mourad Mabrook³

^{1,3}Department of Space Communication, Faculty of Navigation Science & Space Technology, Beni-Suef University, Egypt

¹Institute of Geodesy and Geoinformation (IGG), University of Bonn, Germany

²College of Engineering, Energy & Technology Research Centre, Effat University, Jeddah, Saudia Arabia

³Space & Navigation Applied Science Program, Beni-Suef National University, Egypt

Article Info

Article history:

Received Mar 23, 2025

Revised MMar 2, 2026

Accepted Jun 23, 2026

Keyword:

Deep learning

Semantic segmentation

Spatial resolution

Precision agriculture

Anomaly detection

Ground Sampling Distance

ABSTRACT

The paper explores resolution-accuracy trade-offs in UAV-based semantic segmentation for precision agricultural imagery, which is an important challenge in achieving computational efficiency while maintaining satisfactory results for semantic segmentation. It is widely believed that higher resolution imagery leads to more accurate results, but it also raises processing expenses and restricts the ability to deploy in real-time systems with unmanned aerial vehicles (UAV). This study comprehensively analyzes the effect of the spatial resolution (also known as Ground Sampling Distance (GSD)) on segmentation performance on a range of agricultural anomaly types. The agriculture-vision dataset is used in experimental runs of three different GSDs (10 cm, 20 cm and 40 cm per pixel) with UAV-acquired images. The standard semantic segmentation metrics (mean Intersection over Union (mIoU), Dice coefficient, and computational time analysis) are adopted to evaluate several deep learning models such as U-Net, R2U-Net, U-Net3+, Attention U-Net and DeepLabV3+. The results show that there is no clear correlation between achieving higher spatial resolution and achieving better segmentation accuracy. Medium resolution imagery (20 cm/pixel) can produce similar or better results for large scale anomalies, including dryness and nutrient deficiency, and at significantly lower computational costs. On the other hand, fine-grained anomalies require more fine-grained resolution to better display the anomalies, which shows that the best resolution depends on the task. Further, multi-scale feature aggregation models have higher robustness to resolution degradation. The findings offer a practical understanding of designing the resolution-aware model, which could lead to more efficient use of UAV for resolution-aware deployment, flight planning, and the implementation of edge-AI in precision agriculture. The study provides a data-driven tool for optimizing the spatial resolution versus accuracy versus efficiency of real-world monitoring systems in agriculture.

Copyright © 2026 Institute of Advanced Engineering and Science.

All rights reserved.

Corresponding Author:

M. Mourad Mabrook

Department of Space Communication, Faculty of Navigation Science & Space Technology

Beni-Suef University, Egypt

Email: mohamed.mourad@nsst.bsu.edu.eg

1. INTRODUCTION

Over the last few years, the volume of data collected every day has increased at an exponential rate, making it difficult to analyse. This explosion requires a significant amount of computing resources and hardware infrastructure, making it sometimes difficult and unsustainable. One alternative is to reconsider the data itself, that is to say, whether or not there is need to collect such vast amounts of data to meet the data analysis objectives or if less and less sophisticated data could be adequate. Precision Agriculture presents one

of the key areas, where the monitoring and management of crop conditions are vital to food security and the use of resources. This research combined the two critical issues: efficient management of large amounts of data, and the increasing relevance of precision agriculture. It investigates the need for high-resolution imagery, which is increasingly being captured by Unmanned Aerial Vehicles (UAVs), to achieve effective crop monitoring or whether lower spatial resolutions are acceptable, but require greater computational demands. The increasing interest in UAVs within precision agriculture has led to their use for capturing high-resolution images that allow the identification of minute variations within agricultural areas. Changes like early signs of disease, nutrient deficiency or water stress can be detected with deep learning, through object classification within the image. Semantic segmentation takes this to the next level by accurately segmenting anomalies at the pixel level. At the same time, the impact of spatial resolution on the accuracy of semantic segmentation models is still under-researched and it is not clear if high spatial resolution is always better for anomaly detection or if lower resolution is acceptable for practical purposes without compromising accuracy. To overcome this, the present study employed UAV images of different resolution and semantic segmentation models such as R2U-Net, DeepLabV3+ and SegFormer.

Semantic Segmentation has been proven to benefit agricultural applications, including crop mapping, disease detection and more, frequently yielding detailed results with high-resolution data. In addition, some studies have investigated lightweight versions of the model that involve reducing the number of parameters, thus boosting the model's computational efficiency. But, few have ever studied how the spatial resolution varies the accuracy especially for particular anomaly types.

This research is filling that gap by experimenting with different resolution levels of the models and comparing the performance with each other, with the datasets collected by the UAV marked with labels that indicate key anomalies for agriculture. The findings seek to offer actionable information for precision agriculture, where data resolution, computational resources and analytical accuracy are optimized. The immediate relevance lies in its application to improve the efficiency of AI systems, but it also has a broader relevance to the field of data efficiency in AI, and could be used to develop more scalable and accessible advanced tools for agricultural practitioners around the world.

Over the past few years, studies on Transformer-based segmentation models, like SegFormer, are conducted, and lightweight architectures for real-time applications are designed, but they are not suitable for deployment on UAVs due to their high computational complexity. In this paper, therefore, CNN-based models are tested for different spatial resolutions with an emphasis on deployment aspects of the models as well as the accuracy of the segmentation.

2. RESEARCH METHOD

2.1. Dataset

The Agriculture Vision dataset is a large dataset created by the Agriculture Vision Organization, which is suitable for researchers and practitioners in the field of computer vision in agriculture. This dataset has special features and difficulties, making it different from other image datasets, in that it is focused on aerial images of farms and their agricultural patterns, and is one of the first of its kind in the field of semantic segmentation.

The Agriculture Vision dataset is special in many ways, but one is the unprecedented resolution of images captured from the air, up to 10 cm/pixel. This level of detail enables researchers to analyze farmland images more and extract insights related to crop health, field conditions, and agricultural practices. These findings promise valuable computer vision applications in agriculture, with the potential for improved crop monitoring, disease detection, and yield forecasting. The dataset contains high resolution RGB channels as well as Near-infrared (NIR) channels [5] which gives additional data for image analysis. This multi-channel information enables us to better assess various agricultural phenomena including plant status, moisture level and nutrient content etc. [1],[3], [5].

2.2. Evaluation Matrix

It is crucial to assess the performance of semantic segmentation models, as these models' individual pixels with class labels, revealing significant patterns within images. Evaluation metrics are key to assessing the state of the art and proving the effectiveness of models. In this research, we have used a few evaluation metrics to get in-depth understanding of the performance of the models [6, 7, 8] as follows:

2.2.1. Pixel Accuracy (PA)

A basic metric for measuring semantic segmentation models is pixel accuracy. It gives the ratio of correctly classified pixels to the total number of pixels in the image [6].

$$PA = \frac{\sum_{i=0}^K p_{ii}}{\sum_{i=0}^K \sum_{j=0}^K p_{ij}} \quad (1)$$

2.2.2. Intersection over Union (IoU)

IoU is also called the Jaccard Index and is a commonly used measurement for semantic segmentation problems. It is the ratio of the area of the set of pixels predicted by the segmentation to the area of the set of pixels true to the segmentation, which is a measure of the overlap between the two. In mathematical terms, the intersection of A and B is divided by the union of A and B as per equation (2), IoU is calculated as described in [6],[7] with the ground truth segmentation A and predicted segmentation B.

$$IoU = J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

2.2.3. Precision or positive predictive value (PPV)

It is the relation between true positives and all positive predictions [6],[7],[8].

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

2.2.4. Recall

It indicates the completeness of the positive prediction to that of the ground truth. It decides all ground truth annotations, how many positive predictions are required.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

2.2.5. F1 score

The harmonic mean of precision and recall. It brings in a balance between precision and recall. Less false positive and less false negative means good F1 score [6] [7].

$$\text{F1 Score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

2.2.6. Dice coefficient

Two-thirds of the overlap area between the prediction map and the ground truth map divided by the total number of pixels within the prediction map and ground truth map. When the ground truth is A and the segmentation truth is B, suppose that, the above two are respectively given as [6], and [7].

$$\text{Dice} = 2 \frac{|A \cap B|}{|A| + |B|} \quad (6)$$

2.3. Semantic Segmentation Networks

Semantic segmentation is a computer vision task that involves dividing an image into multiple segments or regions and assigning a class label to each segment. This process is based on the understanding that each segment of the image represents a specific object or part of an object and is labeled accordingly [8],[12],[13],[14],[15]. Deep convolutional neural networks (DCNNs) have been a significant advancement in semantic segmentation, capable of extracting high-level features from images [8],[9],[10],[11],[12],[13],[14],[15]. In this work, we utilize many pre-built models that demonstrate high performance in semantic segmentation, and these models are optimized to fit the data feed. All the models are fully convolutional neural networks (FCNN), where FCNNs replace fully connected layers with convolutional layers [16],[17],[18], allowing them to handle input images of arbitrary sizes and produce corresponding outputs with the exact spatial dimensions [9],[10],[11],[21]. The workflow of FCNNs is such that each layer of data in a Convnet is a three-dimensional array of size $h \times w \times d$ [9],[10],[11]. Convolutional networks are built on translation invariance and operate on local input regions with relative spatial coordinates [10],[11]. FCNNs can efficiently compute feedforward computation and backpropagation over an entire image, making them popular for semantic segmentation tasks [21].

2.3.1. U-net

U-Net is a type of CNN architecture developed in 2015 that was thought to mark an improvement for segmentation of biomedical images [19],[22]. It is based on a symmetrical encoder-decoder structure, accompanied by skip connections and transpose convolutional upsampling, as shown in Figure 1.

The U-Net architecture's encoder extracts spatial and temporal features from the input data. It acts as the pathway for feature extraction. In contrast, the decoder is responsible for reconstructing the image resolution, and it employs transpose convolution operations to gradually reconstruct the resolution of the image. Most importantly, by integrating skip connections, the encoder and decoder feature maps are connected in a carefully designed way that is able to combine low-level and high-level features. The low

level/high level. features combination allows the network to learn both complex semantic features and basic morphological features [10],[19],[22].

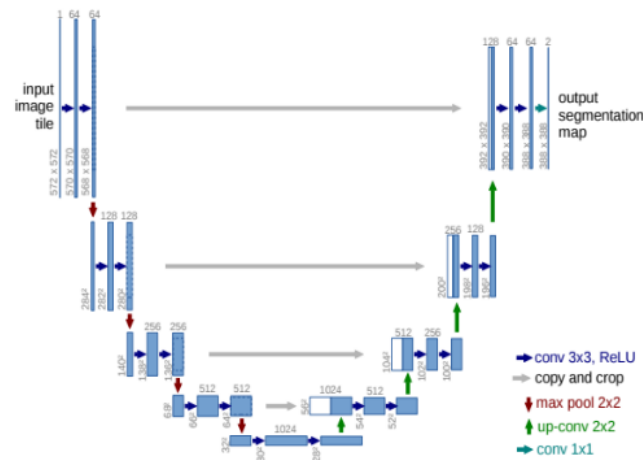


Figure 1. UNet structure [22]

2.3.2. Residual U-Net

Going deeper would improve the performance of a multi-layer neural network; however, it could hamper the training, and a degradation problem may occur [23].

To overcome these problems, He et al. proposed the residual neural network to facilitate training and address the degradation problem. the residual neural network, which consists of a series of stacked residual units. Each residual unit allows the network to learn a residual mapping of the input, rather than directly fitting the desired underlying mapping [14],[23],[24].

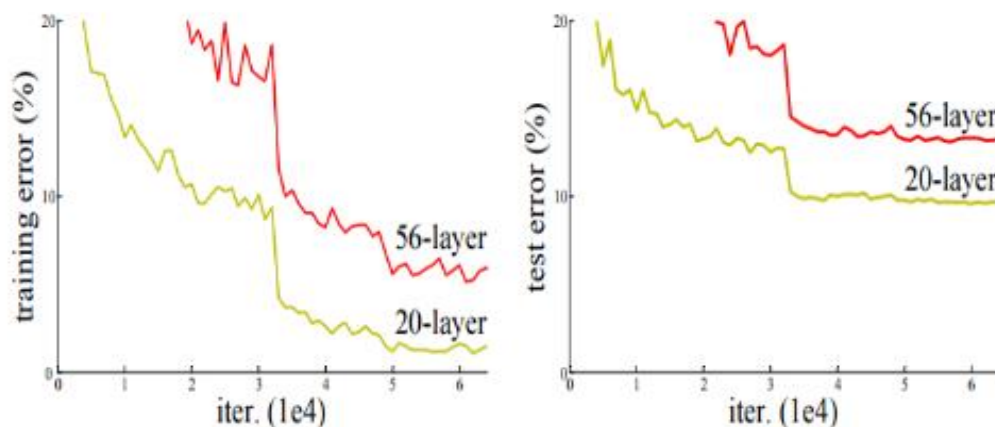


Figure 2. the deeper network has a higher training error [23]

2.3.3. UNet++ Architecture

UNet++ consists of an encoder and decoder that are connected through a series of nested dense convolutional blocks. The main idea behind UNet++ is to bridge the semantic gap between the feature maps of the encoder and decoder before fusion. Also, the using of deep supervision enables model pruning and improves or in the worst case achieves comparable performance to using only one loss layer. [20][25]

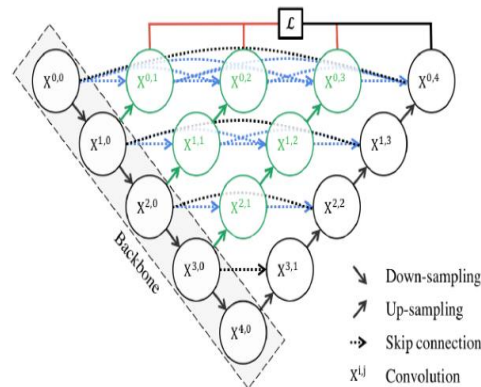


Figure 3. The Unet ++ architecture [25]

2.3.4. Unet +++ architecture

A new architecture called U-Net 3+ was proposed by Oktay et al. in 2020. UNet 3+ utilizes full-scale skip connections that incorporate low-level details with high-level semantics from feature maps in different scales, and Full-scale deep supervision learns hierarchical representations from the full-scale aggregated feature maps' deep supervisions to extract more information from all scales.[25]

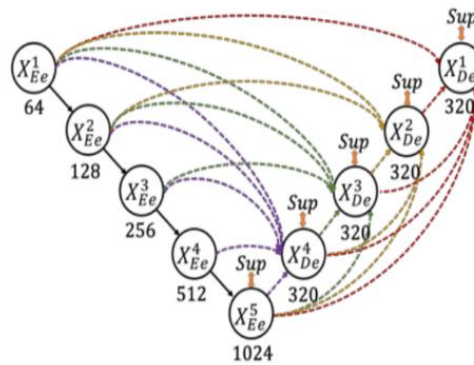


Figure 4. The Unet +++ architecture [25]

2.3.5. Attention U-Net

Attention U-Net, by Imperial College London, Nagoya University & Aichi Cancer Center, University of Luebeck, HeartFlow, and Babylon Health, is briefly reviewed. [26] With Attention Gate (AG), the model automatically focuses on learning the target structures of varying shapes and sizes

The architecture of Attention U-Net is like the U-Net One unique aspect of Attention U-Net is that there is an attention gate (AG) at each skip connection. The AG learns to selectively highlight the relevant regions of the feature maps from the contraction path, which are used to refine the corresponding feature maps from the expansion path. This mechanism helps the network to focus on important features and suppress irrelevant features, which can improve the segmentation accuracy.[26]

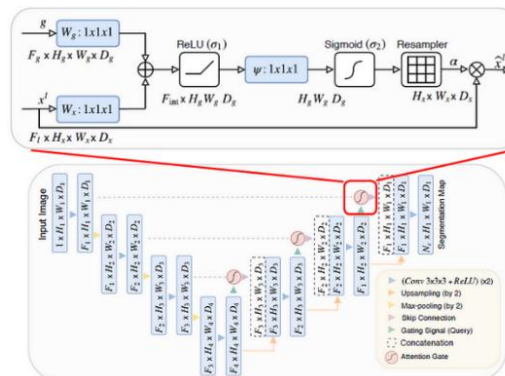


Figure 5. The Attention U-Net architecture

2.3.6. Recurrent U-Net

To capture long-term dependencies in the feature maps, a modification of the popular U-Net architecture is done, called Recurrent U-Net (RU-Net) [27,28].

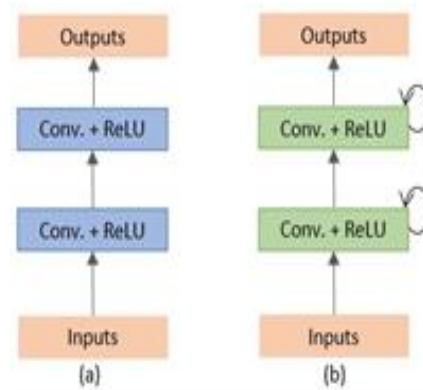


Figure 6 (a) The normal block of U-net, (b) the block of RU-Net. [27]

2.3.7. DeepLabv3+

DeepLabv3+ is an image segmentation model that produces accurate predictions at the pixel level for a given image. It is based on the atrous convolution and multi-scale processing in the DeepLabv3 architecture, which captures both local and global context in the input image.

A novel feature pyramid network (FPN) module is also utilized to fuse multi-scale feature maps, which can be used to obtain more varied spatial information from the image.

Figure 7 shows the structure of DeeplabV3+, which is composed of an encoder and decoder path with an ASPP module at the bottleneck of the structure. where Figure 8 shows The ASPP (Atrous Spatial Pyramid Pooling) technique in DeepLab V3+ uses multiple parallel filters with different rates to exploit multi-scale features and increase the effective field-of-view (FOV) of each layer.

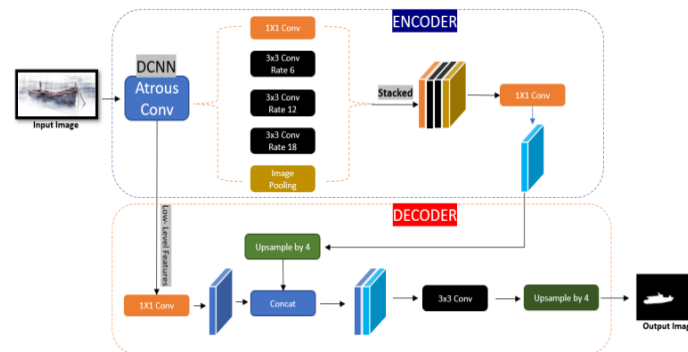


Figure 7 deeplabV3+ structure [29]

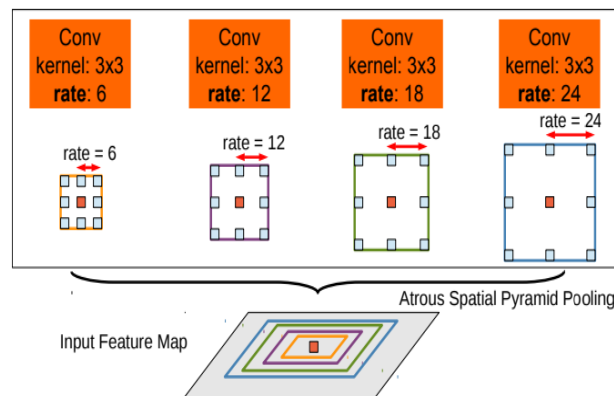


Figure 8. ASPP structure [30]

2.3.8. Training method

Each label or class is a binary semantic segmentation task to evaluate the performance of different models in anomaly detection in agricultural area station tasks. The dataset is supplied at multiple resolutions and will be downsampled to 10 cm, 20 cm, and 40 cm per pixel in order to study the effect of image resolution on the model training and performance. The data was split into training, validation, and test sets for each class for a full evaluation.

A robust pipeline for image preprocessing and enhancement is set up. This pipeline consists of normalization, resizing, and augmentation methods such as rotation and flipping to improve model generalization.

The results of each model are collected and analyzed to reveal the impact of different data structures on model performance. In addition, we will discuss the effect of different image resolutions on the training efficiency and accuracy of the models. This methodology will show how semantic segmentation can be used to identify anomalies in agricultural fields with accuracy. It also allows cost efficient treatments to be made on the basis of very accurate area measurements, and at the final stage, we will work on making the trade-off between the model accuracy and the resolution, thus emphasizing the practical advantages of the study and reassuring the audience about the accuracy of the work.

3. RESULTS AND DISCUSSION

In this section, the semantic segmentation performance is intensively evaluated on different models and resolutions. The models include U-Net, U-Net++, U-Net3+, Attention U-Net, R2U-Net and DeepLabV3+ for different resolutions of 10 cm, 20 cm, and 40 cm per pixel. The particular interest is on the relation between GSD, segmentation accuracy and computational efficiency. The analysis aims to provide both quantitative benchmarks and practical interpretation that are relevant to agricultural deployment based on UAV imagery.

3.1. Qualitative Analysis of Model Predictions

Figures 9-16 show examples of the segmentation findings for different anomaly classes, including dryness, weed clusters, nutrient deficiency, water presence, and planting irregularities. Qualitative comparison shows significant differences in representations of spatial patterns and boundary details by the models at different resolution conditions.

At high spatial resolution (10 cm/pixel), most models exhibit good ability in detecting fine-grained structures, especially in classes such as endrow and planter skip, where accurate boundary delineation is important. Architectures with multi-scale feature aggregation (e.g., U-Net3+ and DeepLabV3+) provide better continuity and less fragmentation in the predicted masks.

However, visual results also indicate that for large-area anomalies such as dryness and nutrient deficiency, increasing the spatial resolution does not necessarily improve the segmentation performance. In many cases, the lower resolution predictions look smoother and less noisy, indicating that too much spatial detail might add unnecessary complexity without helping the semantic understanding. These observations provide preliminary evidence of a strong dependence of the effectiveness of spatial resolution on the structural features of the target anomaly and motivate the quantitative analysis presented in the following sections. Figure 17 shows farmland in 2231 pixels by 5364 pixels, with a total area of 119.67 km². The total area impacted is 48.16 square kilometres. showing the degree of dryness observed.

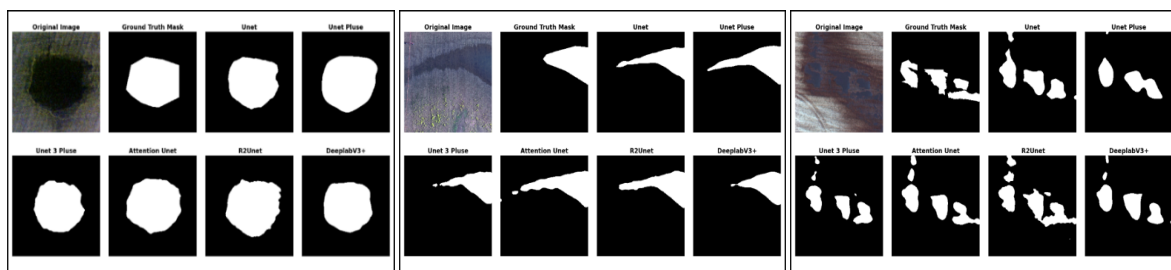


Figure 9. Water Predictions

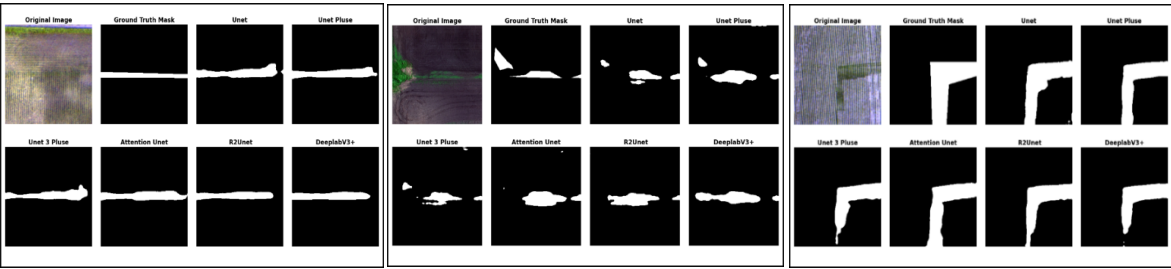


Figure 10. Weed Cluster Predictions

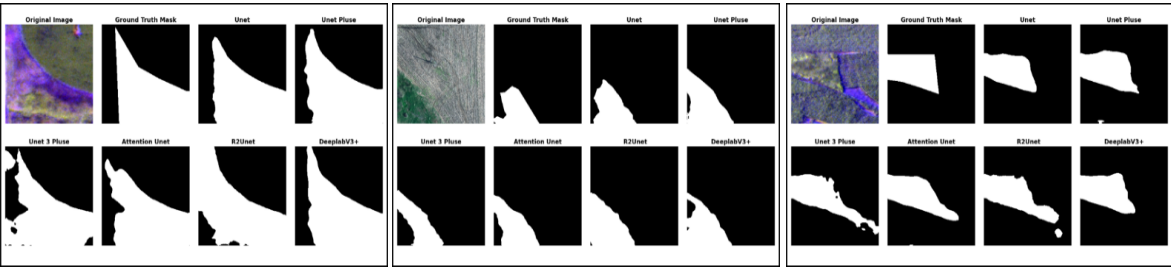


Figure 11. Water Way Predictions

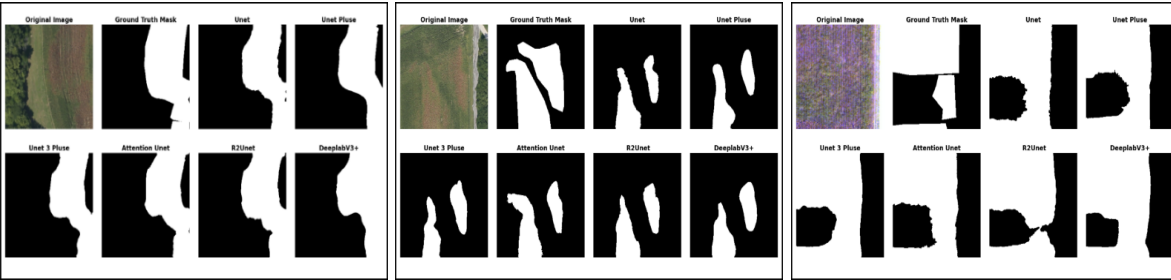


Figure 12. Dryness Predictions

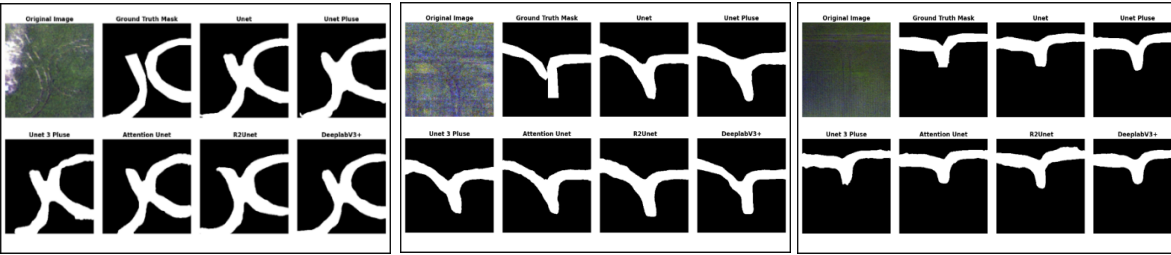


Figure 13. Endrow Predictions

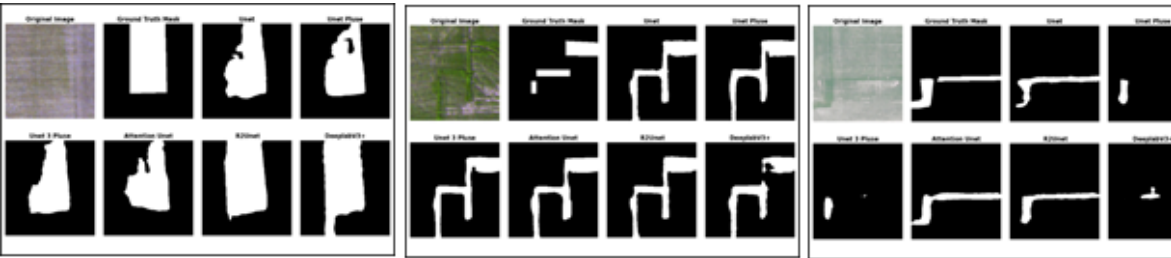


Figure 14 Double Plant Predictions

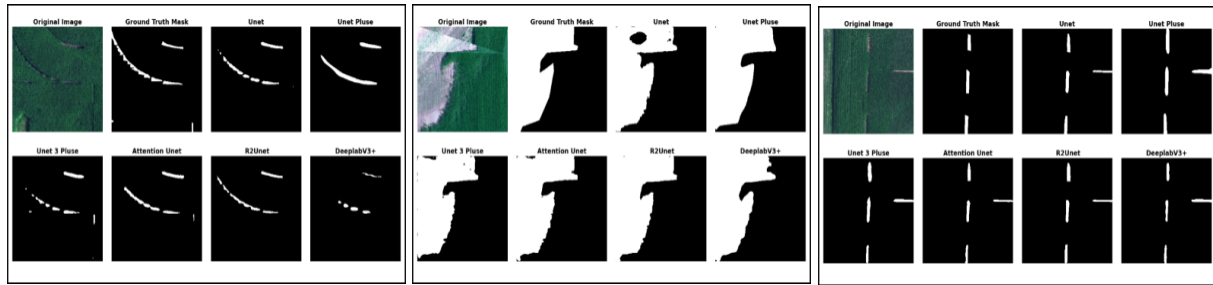


Figure 15. Planter Skip Predictions

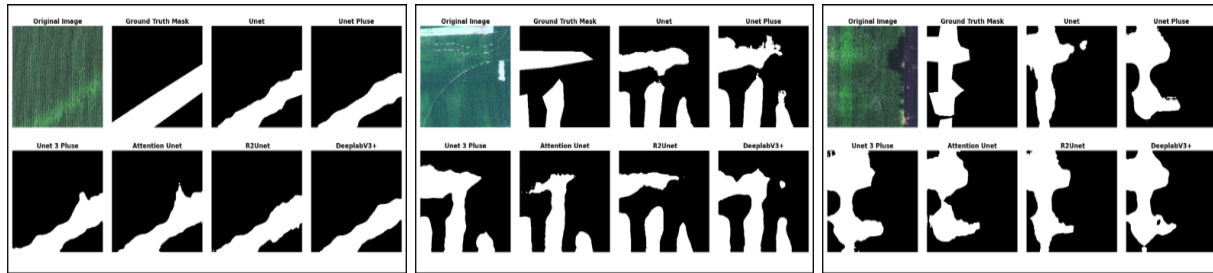


Figure 16. Nutrient Deficiency Predictions

3.2. Quantitative Performance Analysis Across Resolutions

Detailed performance at the three GSD levels (10 cm, 20 cm, and 40 cm per pixel) are provided in Tables 1, 3, and 5. The summary of the overall performance in terms of mean intersection over union (mIoU) are shown in Tables 2, 4, and 6. The U-Net model has the best average mIoU (72.89%) with high resolution (10 cm/pixel), which indicates a good performance in capturing fine spatial details. But this benefit is not consistent across all anomaly types and there is still a large variation in performance across the models. At medium resolution (20 cm/pixel) there is a remarkable change in performance. U-Net3+ achieves best mIoU (79.9%) which is better than all the other models and also better than its own performance at higher resolution. Results show that a moderate down sampling can improve the model generalization by reducing noise and redundant spatial information.

At low resolution (40 cm/pixel), U-Net3+ maintains robust performance (mIoU = 76.94%) while simpler architectures such as U-Net show performance degradation. This suggests that model design is important for achieving accuracy at lower spatial resolutions, and that features from multiple scales are an important factor.

In summary, the results show that higher resolution does not always lead to better performance, medium resolution often provides the best trade-off between efficiency and accuracy, and the resilience of models to resolution changes is very dependent on the architecture.

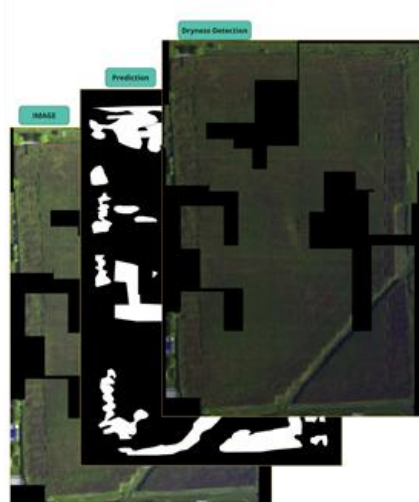


Figure 17 Farmland Dryness Detection

Table 1. Resolution at 10 cm per pixel

Label	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
Weed Cluster	U-Net	10000	555	556	59	7236	.8569	.8667	.7257	.6465	.7824	26
	U-Net++	10000	555	556	77	8288	.8562	.8864	.7275	.6609	.7929	24
	U-Net3+	10000	555	556	78	15600	.8787	.8851	.7645	.6849	.8091	20
	Attention U-Net	10000	555	556	59	7888	.8541	.8679	.7147	.6414	.7796	25
Dryness	R2U-Net	10000	555	556	40	13777	.8196	.8358	.6876	.6041	.7479	75
	DeeplabV3+	10000	555	556	46	10809	.8379	.8455	.7091	.6194	.7479	137
	U-Net	15125	840	841	73	14580	.8405	.8988	.7624	.7005	.8223	26
	U-Net++	15125	840	841	57	9955	.8417	.8772	.7774	.6986	.8211	24
	U-Net3+	15125	840	841	35	16572	.8012	.8655	.7272	.6448	.7801	20
	Attention U-Net	15125	840	841	89	16868	.8962	.9256	.8403	.7851	.8789	25
	R2U-Net	15125	840	841	33	11931	.8149	.8768	.7375	.6686	.7993	75
	DeeplabV3+	15125	840	841	23	8299	.8287	.8884	.7441	.6786	.8071	137
Nutrient deficiency	U-Net	11250	625	625	104	14826	.9334	.9173	.8877	.812	.8989	26
	U-Net++	11250	625	625	52	6572	.8488	.8155	.7408	.6311	.7719	24
	U-Net3+	11250	625	625	43	13100	.7963	.7442	.6631	.5354	.6931	20
	Attention U-Net	11250	625	625	49	7186	.8091	.7777	.6785	.5657	.7213	25
	R2U-Net	11250	625	625	53	17632	.8592	.8445	.7475	.6519	.7877	75
	DeeplabV3+	11250	625	625	22	6085	.8072	.7637	.6766	.5586	.7139	137
Double Plant	U-Net	5610	311	312	46	4714	.9585	.7395	.7429	.5714	.7237	26
	U-Net++	5610	311	312	46	3669	.9556	.7177	.7435	.5666	.7215	24
	U-Net3+	5610	311	312	53	11008	.9612	.7114	.7498	.5641	.7146	20
	Attention U-Net	5610	311	312	46	3269	.9667	.7904	.7286	.6001	.7477	25
Water	R2U-Net	5610	311	312	40	6942	.9609	.7373	.7661	.5914	.7391	75
	DeeplabV3+	5610	311	312	20	3064	.9553	.6679	.7113	.5055	.6603	137
	U-Net	1939	107	108	31	1126	.9076	.8871	.7975	.6681	.8001	26
	U-Net++	1939	107	108	33	1081	.9251	.8815	.8381	.7181	.8351	24
	U-Net3+	1939	107	108	55	3905	.9342	.8847	.8923	.7954	.8851	20
	Attention U-Net	1939	107	108	31	995	.9143	.9005	.8081	.6907	.8162	25
Water Way	R2U-Net	1939	107	108	50	3593	.9377	.9024	.6907	.8092	.8939	75
	DeeplabV3+	1939	107	108	47	3037	.9342	.9084	.8499	.7816	.8761	137
	U-Net	3509	195	196	54	3596	.9814	.9331	.8645	.9271	.9545	26
	U-Net++	3509	195	196	81	4694	.9843	.9539	.8291	.8998	.9471	24
	U-Net3+	3509	195	196	33	4741	.954	.8397	.8157	.7118	.8306	20
	Attention U-Net	3509	195	196	14	718	.9598	.9071	.8391	.765	.8661	25
Planter Skip	R2U-Net	3509	195	196	45	5867	.9501	.8611	.8111	.7131	.8316	75
	DeeplabV3+	3509	195	196	57	4635	.9721	.9018	.8834	.7921	.8831	137
	U-Net	2339	130	131	49	1723	.9901	.837	.8691	.6738	.8031	26
	U-Net++	2339	130	131	25	923	.9696	.9067	.6215	.2901	.4361	24
	U-Net3+	2339	130	131	25	3087	.9801	.7811	.7998	.6189	.7581	20
	Attention U-Net	2339	130	131	40	1473	.9871	.8701	.8691	.6946	.8181	25
Endrow	R2U-Net	2339	130	131	46	3854	.9878	.7693	.8078	.6439	.7796	75
	DeeplabV3+	2339	130	131	22	1557	.9781	.7701	.8196	.6457	.7801	137
	U-Net	4033	224	225	71	4685	.9785	.9075	.9061	.8266	.9047	26
	U-Net++	4033	224	225	67	4322	.9651	.3851	.8371	.7191	.8361	24
	U-Net3+	4033	224	225	78	13590	.9736	.8681	.8872	.779	.8569	20
	Attention U-Net	4033	224	225	113	7487	.9765	.8837	.8848	.7901	.8824	25
	R2U-Net	4033	224	225	61	8185	.9638	.8137	.8499	.7118	.8314	75
	DeeplabV3+	4033	224	225	88	10227	.9866	.9296	.9455	.8781	.9349	137

Table 2. Mean intersection over Union versus various models in 10 cm/pixel resolution

Model	mIOU
U-Net	72.89
U-Net++	68.51
U-Net3+	68.41
Attention U-Net	62.45
R2U-Net	61.49
DeeplabV3+	59.28

Interestingly, the results in Table 2 show that the U-Net model consistently achieves the highest score compared to the other models.

Table 3. Resolution of 20 cm per Pixel

Label	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
Weed Cluster	U-Net	10000	555	556	72	2720	.9186	.9224	.8467	.7778	.8751	26
	R2U-Net	10000	555	556	71	4715	.8821	.8751	.7831	.8241	.8241	20
	U-Net3+	10000	555	556	86	10860	.9402	.9359	.8941	.8411	.9135	75
Dryness	U-Net	15125	840	841	94	6273	.9502	.9532	.9298	.8868	.9411	26
	R2U-Net	15125	840	841	74	5915	.9186	.9288	.8836	.8265	.9089	20
	U-Net3+	15125	840	841	64	11126	.9412	.9536	.9115	.8725	.9319	75
Nutrient deficiency	U-Net	11977	665	666	91	3995	.9563	.9413	.9267	.8691	.9299	26
	R2U-Net	11977	665	666	62	4327	.8528	.8316	.7441	.6436	.7831	20
	U-Net3+	11977	665	666	79	10667	.9271	.8992	.8726	.7927	.8841	75
Double Plant	U-Net	5610	311	312	54	1186	.9339	.9095	.5576	.2971	.4581	26
	R2U-Net	5610	311	312	85	2955	.9731	.7727	.8235	.6663	.7997	20
	U-Net3+	5610	311	312	46	3219	.9129	.8001	.8379	.6828	.8111	75
Water	U-Net	1939	107	108	49	409	.9569	.9691	.8763	.7162	.8346	26
	R2U-Net	1939	107	108	53	542	.9547	.9086	.9004	.8125	.8966	20
	U-Net3+	1939	107	108	55	1109	.9613	.9575	.8983	.8617	.9257	75
Water Way	U-Net	3509	195	196	107	1742	.9924	.9715	.9693	.913	.9545	26
	R2U-Net	3509	195	196	52	1034	.9683	.8952	.8902	.7987	.8879	20
	U-Net3+	3509	195	196	65	2531	.9832	.9331	.9425	.8793	.9356	75
Planter Skip	U-Net	2339	130	131	49	461	.9635	.9649	.4726	.1139	.2046	26
	R2U-Net	2339	130	131	71	1135	.9888	.8294	.8743	.6999	.8217	20
	U-Net3+	2339	130	131	53	1704	.9906	.8563	.9101	.7509	.8548	75
Endrow	U-Net	4033	224	225	45	752	.8178	.9301	.3834	.2506	.4008	26
	R2U-Net	4033	224	225	78	1617	.9555	.8264	.7879	.6578	.7936	20
	U-Net3+	4033	224	225	60	2529	.9604	.8282	.382	.7115	.8313	75

In Table 3, the image resolution is reduced from 10 cm to 20 cm to evaluate how the model performs and answer some key questions. Specifically, we aim to determine whether high-resolution images are necessary in all cases and how image resolution affects the model's performance. High-resolution images contain many pixels, which provide detailed information but require significant processing time and comprehensive hardware capabilities. The goal is to see if similar results can be achieved more efficiently with lower-resolution images. While some approaches may benefit from high-resolution images, others might perform equally well with lower resolutions. The learning process and fine-tuning the models is optimized by understanding the data's characteristics, complexity, and specific requirements.

Table 4. Mean Intersection over Union of various models in 20 cm/pixel resolution

Model	mIOU
U-Net3+	79.9
R2U-Net	74.11
U-Net	60.3

In the latest analysis, table 4 showed that the U-Net3+ model consistently achieved the highest scores compared to other models, whereas in Table 2, the standard U-Net model achieved the highest scores.

Table 5. Resolution of 40 cm per pixel

Label	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
Weed Cluster	U-Net	10000	555	556	92	975	.9451	.9292	.8974	.8353	.9102	26
	R2U-Net	10000	555	556	50	869	.8409	.8406	.7185	.6188	.7645	20
	U-Net3+	10000	555	556	62	2168	.9256	.8964	.8708	.7883	.8813	75
Dryness	U-Net	15125	840	841	74	1393	.9591	.9604	.9431	.9207	.9488	26
	R2U-Net	15125	840	841	72	1582	.9041	.9102	.8614	.7933	.8847	20
	U-Net3+	15125	840	841	56	2295	.9354	.9409	.9068	.8571	.9231	75
Nutrient deficiency	U-Net	11977	665	666	108	1248	.9642	.9478	.9409	.8874	.9403	26
	R2U-Net	11977	665	666	41	900	.8283	.8044	.7036	.5974	.7831	20
	U-Net3+	11977	665	666	93	3759	.9572	.9317	.9283	.8684	.9296	75
Double plant	U-Net	5610	311	312	69	395	.9267	.8886	.5146	.2543	.4055	26
	R2U-Net	5610	311	312	49	490	.9658	.7131	.8291	.6078	.7557	20
	U-Net3+	5610	311	312	64	1153	.9751	.7557	.8641	.6715	.8034	75
Water	U-Net	1939	107	108	72	199	.9536	.9715	.8677	.7487	.8563	26
	R2U-Net	1939	107	108	46	209	.9459	.9082	.8844	.7616	.8646	20
	U-Net3+	1939	107	108	63	348	.9656	.9562	.9166	.8751	.9333	75
Water Way	U-Net	3509	195	196	76	385	.9884	.9646	.9547	.8649	.9274	26
	R2U-Net	3509	195	196	76	464	.9718	.8968	.9043	.8124	.8961	20
	U-Net3+	3509	195	196	51	562	.9786	.9076	.9386	.8515	.9197	75
Planter Skip	U-Net	2339	130	131	65	182	.9884	.9684	.8074	.3508	.5194	26
	R2U-Net	2339	130	131	52	329	.9876	.7906	.9049	.6617	.7895	20
	U-Net3+	2339	130	131	70	608	.9927	.8703	.9309	.8071	.8922	75
Endrow	U-Net	4033	224	225	38	192	.8058	.9341	.3363	.2098	.3468	26
	R2U-Net	4033	224	225	44	315	.7869	.7937	.341	.2578	.4099	20
	U-Net3+	4033	224	225	48	502	.8913	.6932	.5604	.4365	.6077	75

To understand how models perform when the image resolution is downscaled, we observed that model complexity and comprehensiveness are crucial for achieving high scores. We found that the accuracy of the U-Net model degrades compared to other models, and in some cases, it fails to learn effectively from the data. Conversely, the previous analysis shows that the U-Net3+ model consistently scores the highest. This suggests that U-Net3+ is more robust and capable of maintaining performance even with lower-resolution images, making it a better choice for these scenarios.

Table 6. Mean Intersection over Union of various models in 40 cm/pixel resolution

Model	mIOU
U-Net3+	76.94
R2U-Net	63.88
U-Net	63.39

3.3. Model Behavior and Training Efficiency

Figure 18 shows the training curves which offer additional insight into how the models learn and converge during training. U-Net and Attention U-Net converge relatively quickly, making them attractive choices when training time is an important consideration. In comparison, more sophisticated architectures, such as U-Net3+, require longer training periods but exhibit more stable convergence and deliver more consistent performance across different image resolutions.

Another important observation is that increasing model complexity does not necessarily lead to better segmentation performance. Instead, architectures designed to capture multi-scale contextual information are generally more robust to changes in image resolution. This characteristic is particularly advantageous for agricultural scenes, where significant variability in crop appearance, field structure, and background conditions can make accurate segmentation more challenging.

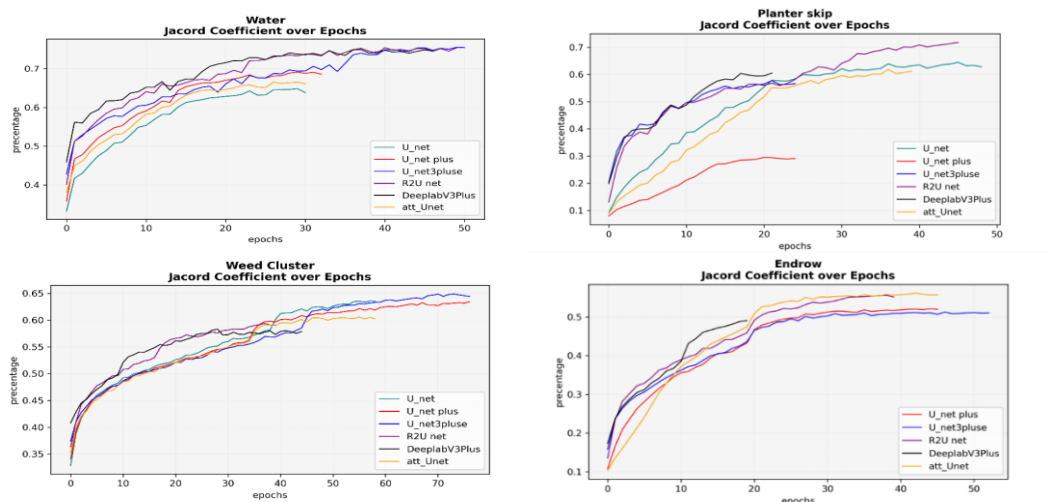


Figure 18. Epochs over accuracy for different dataset

3.4. Spatial Resolution vs Accuracy vs Computational Cost

One of the main contributions of this work is the systematic analysis of the trade-off between spatial resolution, segmentation accuracy, and computational cost, a balance that plays a key role in the practical deployment of UAV-based monitoring systems. Rather than evaluating segmentation accuracy alone, we investigate how image resolution influences both predictive performance and computational efficiency, providing insights that are directly relevant to real-world precision agriculture applications. For this purpose, we evaluated the segmentation performance of all models using the Jaccard Coefficient (JC) at three ground sampling distance (GSD) levels (10, 20, and 40 cm/pixel) and compared these results with the corresponding computational time reported in Tables 1, 3, and 5. The outcomes are summarized through the heatmap in Figure 19 and the 3D surface plots in Figure 20, which provide a comprehensive view of the interaction between image resolution, segmentation performance, and computational requirements.

An important finding is that the relationship between spatial resolution and segmentation accuracy is not strictly linear. While higher image resolution is often assumed to improve segmentation quality, our results show that this is not always the case. For several anomaly categories, e.g. dryness, nutrient deficiency, and weed clusters, the best segmentation performance is achieved at the intermediate resolution of 20 cm/pixel rather than at the finest resolution of 10 cm/pixel. In some cases, lower resolutions produce competitive or superior results. For instance, U-Net achieves a higher IoU for dryness detection at 40 cm/pixel (0.9207) than at 10 cm/pixel (0.7005), while requiring substantially less computational effort. This

suggests that extremely fine spatial detail may introduce unnecessary variability and reduce the model's ability to generalize, especially when the target anomalies exhibit relatively large spatial patterns.

Another notable observation is that reducing image resolution can substantially improve computational efficiency without compromising segmentation accuracy. Moving from 10 cm/pixel to 20 cm/pixel decreases training time by approximately three to five times, depending on the network architecture, while maintaining or improving segmentation performance. These findings indicate that medium-resolution imagery offers an attractive balance between computational cost and predictive capability, making it a practical choice for large-scale UAV-based agricultural monitoring.

The results also demonstrate that the most suitable spatial resolution depends on the type of anomaly being detected. Large and spatially continuous patterns, such as dryness and nutrient deficiency, can be identified reliably using medium or low-resolution imagery (20–40 cm/pixel). In contrast, anomalies characterized by fine structural details, as for endrow patterns and planter skips, benefit from higher-resolution imagery (10 cm/pixel), where subtle spatial features remain distinguishable. This highlights the importance of selecting image resolution according to the target application instead of relying on a single high-resolution acquisition strategy for all scenarios.

Finally, the comparative analysis reveals clear differences in how network architectures respond to resolution changes. U-Net3+ consistently maintains stable segmentation performance across all evaluated resolutions, demonstrating strong robustness to image downsampling. U-Net, on the other hand, experiences more noticeable performance degradation as the spatial resolution decreases, whereas R2U-Net provides moderate robustness but requires considerably higher computational resources. Overall, these results suggest that architectures capable of effectively aggregating multi-scale contextual information are better equipped to maintain reliable segmentation performance under varying image resolutions, making them well suited for resolution-invariant UAV applications.

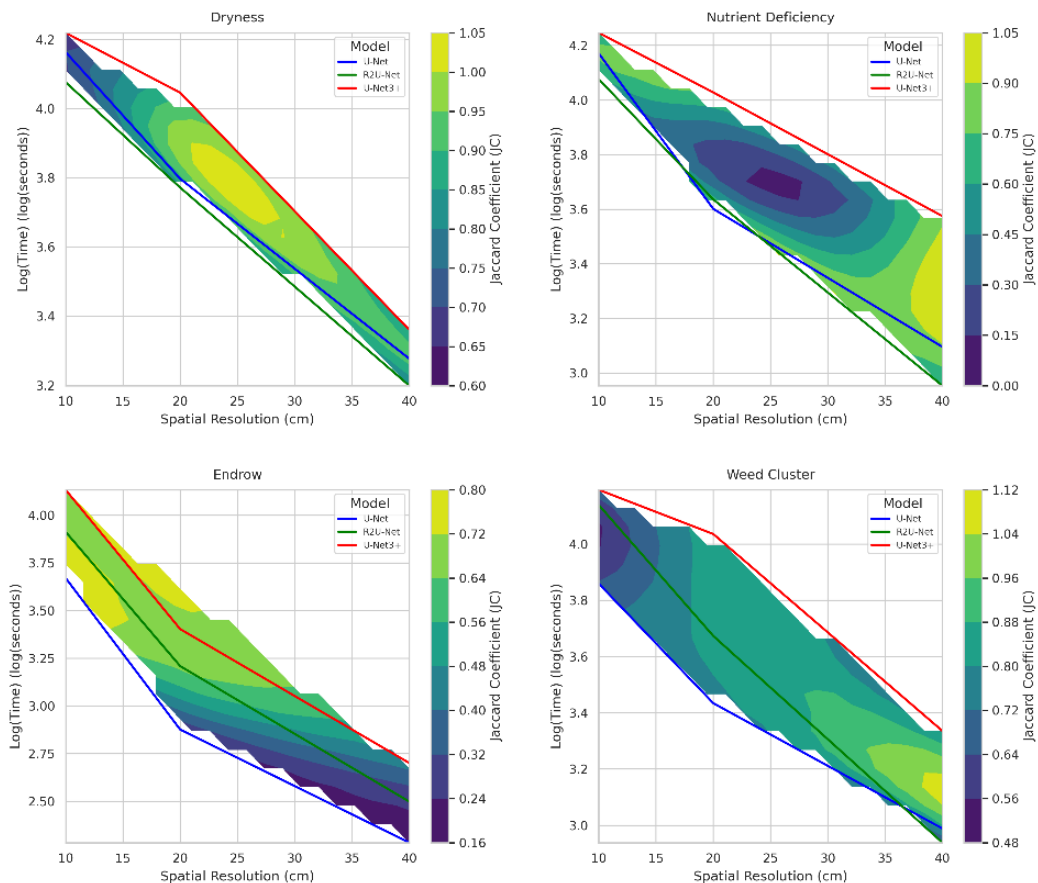


Figure 19. Heatmap line Analysis

Figure 19 shows the relationship between segmentation accuracy, spatial resolution, and computational cost across all datasets and network architectures. The heatmap visualizes the Jaccard

Coefficient (JC), with warmer colours indicating higher segmentation accuracy, while the overlaid curves illustrate how each model balances predictive performance against computational time across the three spatial resolutions. This representation provides a comprehensive view of the trade-off between accuracy and efficiency under different operating conditions.

Several consistent trends can be observed. For large-scale anomalies such as Dryness, Nutrient Deficiency, and Weed Cluster, the highest segmentation accuracy is generally achieved at medium or lower spatial resolutions. For example, U-Net reaches its best performance on the Dryness dataset at 40 cm/pixel (JC = 0.9207), whereas performance decreases considerably at 10 cm/pixel (JC = 0.7005), despite the substantially higher computational cost associated with processing higher-resolution imagery. These results suggest that increasing spatial resolution does not necessarily improve segmentation performance and may even reduce model generalization when the target anomalies are characterized by broad spatial patterns.

In contrast, anomalies containing fine spatial structures exhibit a different behaviour. The Endrow dataset benefits from higher-resolution imagery, with U-Net improving from a JC of 0.2098 at 40 cm/pixel to 0.8266 at 10 cm/pixel. Similarly, the Weed Cluster dataset achieves its best performance at the intermediate resolution of 20 cm/pixel, where U-Net3+ attains the highest JC (0.8411), while a slight reduction in accuracy is observed at 10 cm/pixel. These observations reinforce that the optimal image resolution depends on the spatial characteristics of the anomaly rather than following a single resolution strategy for all detection tasks.

The figure also highlights the behaviour difference of the evaluated architectures. U-Net consistently provides the shortest processing time across all resolutions, making it an attractive option for applications where computational efficiency is a primary concern. In comparison, U-Net3+ demonstrates greater stability as image resolution changes, maintaining consistently high segmentation accuracy across different operating conditions. This robustness is particularly evident for the Weed Cluster dataset, where U-Net3+ achieves the highest overall performance at 20 cm/pixel. Overall, the results indicate that selecting an appropriate combination of network architecture and image resolution is essential for achieving an effective balance between segmentation accuracy and computational efficiency in UAV-based agricultural monitoring.

From these results, we can include large-scale monitoring (dryness, Nutrient Deficiency, weed cluster) with no need for high resolution and can lower resolution to get better performance and fast results with less computational need.

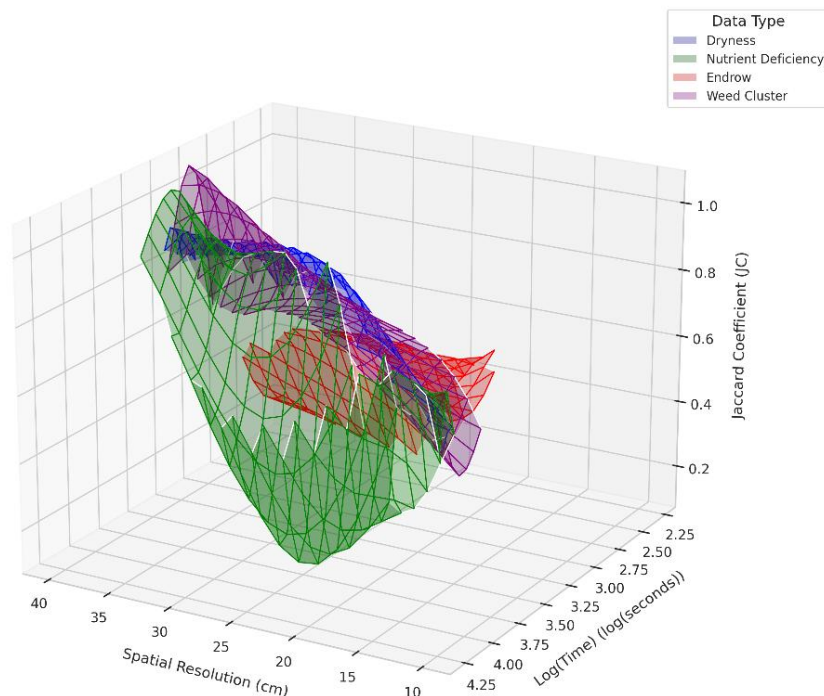


Figure 20. 3D Surface and Contour Plots

Figure 20 provides a detailed visualization of the relationship between spatial resolution, JC, and Log (Time) for four selected data types. This contrasts with Figure 20, which shows that the high spatial resolution image does not give higher Accuracy. We should consider the data type, the target from it, and the computational hardware that will process the data.

We still use relatively high-resolution images in this paper, with respect to UAVs with 1-meter resolution or bigger-scale satellites like Sentinel with 10-meter resolution. However, this research shows a comprehensive study of UAV images.

A key novelty of this study is the formulation of a hybrid combined loss function that integrates three complementary components: Binary Cross-Entropy (BCE) loss, Dice loss, and Focal loss. This combination addresses class imbalance and hard-sample mining simultaneously, which is particularly critical for agricultural anomaly detection where anomalies (e.g., dryness, nutrient deficiency) occupy a small fraction of pixels.

The proposed combined loss is defined as:

$$L_{\text{total}} = w_{\text{BCE}} \cdot L_{\text{BCE}} + w_{\text{Dice}} \cdot L_{\text{Dice}} + w_{\text{Focal}} \cdot L_{\text{Focal}} \quad (7)$$

where the weights are empirically set to $w_{\text{BCE}} = 0.3$, $w_{\text{Dice}} = 0.3$, and $w_{\text{Focal}} = 0.4$

Binary Cross-Entropy (BCE) Loss

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (8)$$

where y_i is the ground truth label, $\hat{y}_i$ is the predicted probability after sigmoid activation, and N is the number of pixels.

Dice Loss

The Dice loss is derived from the Dice coefficient and is particularly effective for handling class imbalance:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N \hat{y}_i y_i + \varepsilon}{\sum_{i=1}^N \hat{y}_i + \sum_{i=1}^N y_i + \varepsilon} \quad (9)$$

where $\varepsilon = 10^{-6}$ is a smoothing term to prevent division by zero.

Focal loss down-weights easy examples and focuses training on hard, misclassified examples:

$$\mathcal{L}_{\text{Focal}} = -\alpha(1 - p_t)^\gamma \log(p_t + \varepsilon) \quad (10)$$

$$p_t = \begin{cases} \hat{y} & \text{if } y = 1 \\ 1 - \hat{y} & \text{otherwise} \end{cases} \quad (11)$$

The parameters are set to $\alpha = 0.25$ and $\gamma = 2.0$, following standard practice in object detection literature.

Recent developments in semantic segmentation have introduced several advanced architectures, including Transformer-based models such as SegFormer and lightweight real-time networks such as BiSeNet and PIDNet. These approaches have achieved impressive performance across a wide range of computer vision tasks by effectively modeling long-range contextual information and learning richer feature representations. However, these benefits often come at the expense of increased computational complexity, higher memory consumption, and greater inference cost compared with conventional convolutional neural network (CNN) architectures.

In UAV-based agricultural monitoring, practical deployment is constrained by limited onboard computing resources, energy consumption, and the need for timely inference. Under these conditions, CNN-based models remain an attractive solution because they offer a favorable balance between segmentation accuracy, computational efficiency, and implementation complexity. For this reason, the present study focuses exclusively on CNN-based architectures to provide a realistic evaluation of models suitable for operational UAV applications. Nevertheless, extending the comparative analysis to include Transformer-based and lightweight segmentation networks represents an important direction for future research as hardware capabilities continue to evolve.

Tables 6,7,8 show a Performance comparison of semantic segmentation models for the Water Way class at 10 cm/px, 20cm/px and 40 cm/px spatial resolution respectively.

Table 6. Performance comparison of semantic segmentation models for the Water Way class at 10 cm/px spatial resolution.

Model	Accuracy (%)	Precision (%)	Recall (%)	IoU (JC)	Dice/F1 (%)
U-Net	98.14	93.31	86.45	92.71	95.45
U-Net++	98.43	95.39	82.91	89.98	94.71
U-Net3+	95.40	83.97	81.57	71.18	83.06
Attention U-Net	95.98	90.71	83.91	76.50	86.61
R2U-Net	95.01	86.11	81.11	71.31	83.16
DeepLabV3+	97.21	90.18	88.34	79.21	88.31
SegFormer	96.60	89.57	94.20	84.88	91.82
BiSeNet	91.86	77.02	85.29	67.99	80.95
PIDNet	96.28	89.91	92.00	83.39	90.94

Table 7. Performance comparison of semantic segmentation models for the Water Way class at 20 cm/px spatial resolution.

Model	Accuracy (%)	Precision (%)	Recall (%)	IoU (JC)	Dice/F1 (%)
U-Net	99.24	97.15	96.93	91.30	95.45
U-Net3+	98.32	93.31	94.25	87.93	93.57
R2U-Net	96.83	89.52	89.02	79.87	88.79
SegFormer	90.34	82.17	66.90	58.42	73.76
BiSeNet	94.69	86.18	87.94	77.07	87.05
PIDNet	93.58	81.57	88.31	73.62	84.81

Table 8. Performance comparison of semantic segmentation models for the Water Way class at 40 cm/px spatial resolution.

Model	Accuracy (%)	Precision (%)	Recall (%)	IoU (JC)	Dice/F1 (%)
U-Net	98.85	96.95	96.47	92.74	95.49
U-Net3+	97.87	90.77	93.86	85.91	91.97
R2U-Net	97.18	89.96	90.43	81.24	89.61
SegFormer	90.68	71.59	89.65	66.12	79.61
BiSeNet	90.73	73.58	84.78	64.99	78.78
PIDNet	89.72	71.86	81.05	61.52	76.18

4. CONCLUSION

This study examined the effect of spatial resolution, expressed in terms of the Ground Sampling Distance (GSD), on the performance of semantic segmentation models for UAV-based agricultural anomaly detection. We perform extensive experiments on multiple models and multiple anomaly types to show that the widely assumed “higher resolution leads to better performance” does not always hold in practice.

The results show that medium-resolution (20 cm/pixel) to low-resolution images (40 cm/pixel) are often the best compromise between segmentation accuracy and computational efficiency, especially for large-scale anomalies such as dryness, nutrient deficiency, and weed clusters. Conversely, fine-scale anomalies like endrow patterns and irregularities in planting are more responsive to high-resolution imagery (10 cm/pixel), which implies that the optimal spatial resolution is inherently task-dependent. Moreover, the results indicate that higher model complexity does not imply better performance, while some architectures (e.g., U-Net3+) have excellent robustness under different resolution levels.

One of the main contributions of this work is the quantification of the trade-off between spatial resolution, accuracy and computational cost. We demonstrated that a significant reduction in processing time can be achieved with a slight reduction in performance by using lower resolution imagery. We showed that a significant reduction in processing time can with a slight performance degradation with imagery of lower resolution. The constraints of flight time, onboard processing, data transmission and energy consumption should be considered for the UAV-based agricultural systems. This implies that UAV-based agricultural systems must consider constraints on flight time, onboard processing, data transmission and energy consumption carefully.

In practice, the work provides guidelines on how to choose appropriate image resolutions and model architectures depending on the type of anomaly targeted and operational constraints. This work offers more

efficient, scalable, and cost-effective deployment of deep learning solutions for precision agriculture by shifting the focus from model-centric optimization to resolution-aware system design.

We also discussed the effect of using lightweight models like transformer-based ones, but the results still need further investigation. For future work, we will extend this analysis to include lightweight real-time segmentation models and explore adaptive resolution strategies and edge deployment scenarios to further enhance the applicability of UAV-driven agricultural monitoring systems.

ACKNOWLEDGEMENTS

Many thanks for Naptaplaya co. For providing datasets and images for our team. Also, many thanks for faculty of navigation science and space technology (NSST) in Beni-Suef University for their academic and technical supporting this research.

REFERENCES

- [1] L. Dietmann And J. Stålhamma, "Adoption Of Digital Precision Agriculture Technology And Farm Data: A Case Study On Swedish Grain Farmers And On-Combine Near-Infrared Spectroscopy For Quality Measurement," Master's Thesis, Faculty of Engineering LTH Lund University, 2020.
- [2] B. M. Kayrouz, "Precision Agriculture: Realizing Increased Profit And Reduced Risk Through Cost Map And Lightbar Adoption," Master's Thesis, University of Kentucky, Graduate School, 2008.
- [3] L. P. Hafsal, "Precision Agriculture With Unmanned Aerial Vehicles For Smc Estimations – Towards A More Sustainable Agriculture," Master's Thesis, Department Of Applied Ecology And Agriculture, Hedmark University of Applied Sciences. 2025.
- [4] G. Bucci, D. Bentivoglio, And A. Finco, "Precision Agriculture As A Driver For Sustainable Farming Systems: State Of Art In Literature And Research," *Quality - Access to Success*. Vol. 19. pp. 114-121. 2018
- [5] M. T. Chiu et al. "Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 2825-2835, doi: 10.1109/CVPR42600.2020.00290.
- [6] T.S. Rajalakshmi and R. Senthilnathan, "Dataset and Performance Metrics Towards Semantic Segmentation," *International Journal of Engineering and Management Research*, Vol. 13, No. 1, Pp. 40, February 2023. <https://doi.org/10.31033/ijemr.13.1.5>
- [7] E. Fernandez-Moral, R. Martins, D. Wolf And P. Rives, "A New Metric for Evaluating Semantic Segmentation: Leveraging Global and Contour Accuracy," 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 2018, Pp. 1051-1056. <https://doi.org/10.1109/IVS.2018.8500497>
- [8] Zhang, Sanxing, Zhenhuan Ma, Gang Zhang, Tao Lei, Rui Zhang, and Yi Cui. 2020. "Semantic Image Segmentation with Deep Convolutional Neural Networks and Quick Shift" *Symmetry* 12, no. 3: 427. <https://doi.org/10.3390/sym12030427>
- [9] K. Simonyan And A. Zisserman, "Very Deep Convolutional Networks For Large-Scale Image Recognition," *Arxiv Preprint Arxiv:1409.1556*, 2014.
- [10] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 3431-3440, <https://doi.org/10.1109/CVPR.2015.7298965>.
- [11] L. -C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834-848, 1 April 2018, <https://doi.org/10.1109/TPAMI.2017.2699184>
- [12] H. Noh, S. Hong and B. Han, "Learning Deconvolution Network for Semantic Segmentation," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015, pp. 1520-1528, <https://doi.org/10.1109/ICCV.2015.178>
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, And A. Rabinovich, "Going Deeper With Convolutions," In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 1-9, <https://doi.org/10.1109/CVPR.2015.7298594>
- [14] K. Simonyan And A. Zisserman, "Very Deep Convolutional Networks For Large-Scale Image Recognition," *arXiv:1409.1556v6*, 2014, <https://doi.org/10.48550/arXiv.1409.1556>
- [15] C. Farabet, C. Couprie, L. Najman and Y. LeCun, "Learning Hierarchical Features for Scene Labeling," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915-1929, Aug. 2013, <https://doi.org/10.1109/TPAMI.2012.231>

- [16] S. Jegou, M. Drozdal, D. Vazquez, A. Romero, And Y. Bengio, "The One Hundred Layers Tiramisu: Fully Convolutional Densenets For Semantic Segmentation," Montreal Institute for Learning Algorithms, Ecole Polytechnique De Montreal, Imagia Inc., Montreal, And Computer Vision Center, Barcelona.
- [17] A. Paszke, A. Chaurasia, S. Kim, And E. Culurciello, "Enet: A Deep Neural Network Architecture For Real-Time Semantic Segmentation," arXiv:1606.02147v1, <https://doi.org/10.48550/arXiv.1606.02147>
- [18] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, And A. L. Yuille, "Semantic Image Segmentation With Deep Convolutional Nets And Fully Connected Crfs," arXiv:1412.7062v4 <https://doi.org/10.48550/arXiv.1412.7062>
- [19] V. Iglovikov And A. Shvets, "Ternausnet: U-Net With Vgg11 Encoder Pre-Trained On Imagenet For Image Segmentation," arXiv:1801.05746v1, <https://doi.org/10.48550/arXiv.1801.05746>
- [20] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, And J. Liang, "Unet++: A Nested U-Net Architecture For Medical Image Segmentation," Lecture Notes in Computer Science, vol 11045. Springer, Cham. https://doi.org/10.1007/978-3-030-00889-5_1
- [21] J. Long, E. Shelhamer, And T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, Pp. 3431-3440.
- [22] O. Ronneberger, P. Fischer, And T. Brox, "U-Net: Convolutional Networks For Biomedical Image Segmentation," In International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI), 2015, Pp. 234-241.
- [23] K. He, X. Zhang, S. Ren, And J. Sun, "Deep Residual Learning For Image Recognition," In Proceedings of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2016, Pp. 770-778.
- [24] Z. Zhang, Q. Liu, And Y. Wang, "Road Extraction By Deep Residual U-Net," IEEE Transactions on Intelligent Transportation Systems, Vol. 20, No. 9, Pp. 3379-3390, 2019.
- [25] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, And J. Wu, "Unet 3+: A Full-Scale Connected Unet For Medical Image Segmentation," In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020, Pp. 68-74.
- [26] Q. Zuo, S. Chen, And Z. Wang, "R2au-Net: Attention Recurrent Residual Convolutional Neural Network For Multimodal Medical Image Segmentation," IEEE Transactions on Medical Imaging, Vol. 40, No. 1, Pp. 145-157, 2021.
- [27] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, And V. K. Asari, "Recurrent Residual Convolutional Neural Network Based On U-Net (R2u-Net) For Medical Image Segmentation," In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 837-844.
- [28] S. Zheng, X. Li, And H. Li, "Road Information Detection Method Based On Deep Learning," Journal Of Physics: Conference Series, Vol. 1827, No. 1, pp. 012181, 2021.
- [29] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, And A. L. Yuille, "Deeplab: Semantic Image Segmentation With Deep Convolutional Nets, Atrous Convolution, And Fully Connected Crfs," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 40, No. 4, pp. 834-848, 2018.
- [30] R. J. Hemalatha, T. R. Thamizhvani, A. J. A. Dhivya, J. E. Joseph, B. Babu, And R. Chandrasekaran, "Active Contour Based Segmentation Techniques For Medical Image Analysis," In 2020 International Conference on Computational Intelligence in Data Science (ICCIDS), 2020, pp. 112-11.

BIOGRAPHY OF AUTHORS



Mohamed Tawhid Amin

He is M.Sc. Student at Institute of Geodesy and Geoinformation, University of Bonn, Germany. Received B.Sc. at Faculty of Navigation Science & Space Technology, Beni-Suef University, Egypt. He is a Teaching Assistant, Department of Space Communication, Faculty of NSST, Beni-Suef University, Egypt from 2024 till now.

mohatawheed@gmail.com, s57mamin@uni-bonn.de



Aziza I. Hussein    received her Ph.D. degree in Electrical & Computer Engineering from Kansas State University, USA in 2001 and the M.Sc. and B.Sc. degrees from Assiut University, Egypt in 1989 and 1983, respectively.

She joined Effat University in Saudi Arabia in 2004 and established the first Electrical and Computer Engineering program for women in the country and taught related courses. She was the head of the Electrical and Computer Engineering Department at Effat University from 2007-2010. She was the head of Computer and Systems Engineering Department, Faculty of Engineering, Minia University, Egypt from 2011-2016. She was a professor and chair of the Electrical & Computer Engineering Department and director of the Master of Energy program at Effat University Saudi Arabia from 2016-2021. Currently she is a professor and researcher at the same department. Her research interests include microelectronics, analog/digital VLSI system design, RF circuit design, high-speed analog- to-digital converters design, and wireless communications systems design. azibrahim@effatuniversity.edu.sa



M. Mourad Mabrook    received his B.Sc. and M.Sc. degrees in Electrical and Communication Engineering from Assiut University, Egypt, in 2008 and 2013, respectively. He received Ph.D. degree in communication Dept., Menia University, Egypt in 2017. He is Associate Professor at Space communication Engineering Dept. in faculty of Navigation Science & Space Technology (NSST), Beni-Suef university, Beni-Suef, Egypt. Head of Applied sciences of space and navigation Program, Beni-Suef National University. He is a part time Assistant prof. at faculty of Engineering, Nahda university (NUB), Beni-Suef, Egypt. He Published more than 40 journal and conference papers in the fields of wireless communications, 5G networks, cognitive radio, Artificial intelligence, circuit design and sensors. He is a reviewer in many international journals related to Elsevier, Springer and IEEE Publishers'. mohamedmourad2008@gmail.com, mohamed.mourad@nsst.bsu.edu.eg