Volume 2, 2026, Pages 033-042 *Journal of Smart and Sustainable Farming*
| Regular Article | Open Access

Deep Residual UNet with Transfer Learning for Fine Grained Agricultural Land Segmentation

Agnar Alfons Ramel
Faculty of Physical Sciences Mathematics, University of Iceland, Reykjavík, Iceland.
aalfonsramel32@gmail.com

Article Info
Journal of Smart and Sustainable Farming
https://www.ansispublications.com/journals/jssf/jssf.html

Received 16 January 2026
Revised from 26 February 2026
Accepted 03 March 2026
Available online 18 March 2026
Published by Ansis Publications.

https://doi.org/10.64026/JSSF/2026004

Corresponding author(s):
Agnar Alfons Ramel, Faculty of Physical Sciences Mathematics, University of Iceland, Reykjavík, Iceland.
Email: aalfonsramel32@gmail.com

Abstract – The accurate mapping of agricultural landscapes based on the high-resolution remote sensing images is a basic requirement of successful application of the policies of the precision farming and land management. This paper presents TL- ResUNet, which is an adapted U-Net network, that integrates a ResNet-50 encoder, supplemented with transfer learning methods, to be used in semantic segmentation in heterogeneous agricultural scenes. We used land-use data that were annotated according to the polygonal boundaries at a 0.5m spatial resolution thus accounting to the fragmented smallholder farming systems. The architecture combines the use of vegetation indices, attention, and dilated convolutional layers with the objective of enhancing the fine-scale crop boundary and sensitivity to the linearities. Empirical comparisons show that domain-specific pre-training is an effective approach that produces a significant increase in the Intersection over Union (IoU) measure, reaching as high as 81. The proposed model, therefore, is better than the state-of-the-art models in heterogeneous and mixed land-use environment.

Keywords – Agricultural Segmentation, Remote Sensing, Remotely Sensed Imagery, UNet, Transfer Learning, Land-Use Mapping.

I. INTRODUCTION

Proper mapping of agricultural lands is necessary to achieve effective management of the resource as well as a variety of significant tasks including crop mapping, estimate of yields and sustainable agricultural planning. The basis of the geographic units of agrarian activities is the agricultural land. Accurate and quality parcel delineations are essential to most geospatial agricultural investigations, which has a direct effect on the further operation in precision agriculture, land-use observation, and policy enforcement.

Traditionally land-cover mapping with the use of remotely sensed imagery has been based mainly on per-pixel classification where each pixel is categorized in one or more land-cover categories. The major challenge posed by the implementation of a fine spatial-resolution imagery in the per- pixel land-cover classification is over-sampling; the fine detail will not easily differentiate between areas of interest, but rather it will represent within-feature variation [1].

Agriculture is faced with an inferno of issues, such as land degradation, the disappearance of biodiversity, and the detrimental effects of climate change. These issues are significantly increased by the fact that the monoculture of one species of crop is long-term. At the same time, the high rate of population growth all around the world requires the diversified range of agricultural products to meet human needs of sustenance, energy and clothing. Crop rotation, in this regard, has been widely implemented as a management technique. Crop rotation refers to the process of using two or more types of crops sequentially in the same field of operation on a pre-defined succession program. The practice could significantly boost the agronomic performance through interference with the cycle of insects and diseases, inhibition of weeds, increased grain production, and soil nutrient replenishment.

Automated mapping of agricultural field boundaries is one of the most challenging problems in remote sensing that has not been yet resolved successfully since crop varieties and fragmented land cover have similar spectral characteristics. Traditional edge-detection methods (e.g., Sobel, Prewitt) are much more prone to noise and have a lower ability to see gradients in homogeneous areas. Superpixel methods, e.g. SLIC, enhance segmentation through a combination of pixels to coherent units; that being said, they lack the same scalability as repetitive computational requirements. SNIC (Simple Non-Iterative Clustering) in contrast has linear computational complexity, produces compact super-pixels through priority queues, preserving edges, but eliminating spectral noise [2]. However, a single use of SNIC might not be able to reflect subtle boundaries covered by vegetation or soil texture. Additional algorithms, like the Canny edges detector, that uses Gaussian smoothing and hysteresis thresholding to detect edges can be used to complement edge detection, but cannot deal with discontinuities in agricultural applications.

Remote sensing (RS) technology has led to a significant change in agricultural reform. RS enables high frequency capture of vast amount surface information on the earth hence providing unmatched efficiency in production and management of agriculture. With a rich set of sensors, RS has the ability to provide primary and secondary data regarding virtually all vital aspects of agronomic practice, such as crop development indicators, soil moisture, detection of insects, warning of diseases, and yield forecasts. The large-scale geographical coverage and different resolution of RS technologies provide plenty of data that can be used in production and management of agricultural activities.

As shown in **Fig. 1**, satellites with varying spatial resolutions play different and vital roles in the diverse precision agriculture (PA) strategies, and use their individual attributes to approach the specified needs of PA in a holistic manner. With the ever-developing technology on remote-sensing sensors, the future agricultural managers and agricultural practitioners are likely to enjoy the reimbursement of advanced RS applications; a case in point is that RS data has been proven to be incredibly useful and viable in evaluation and monitoring of agricultural practices.

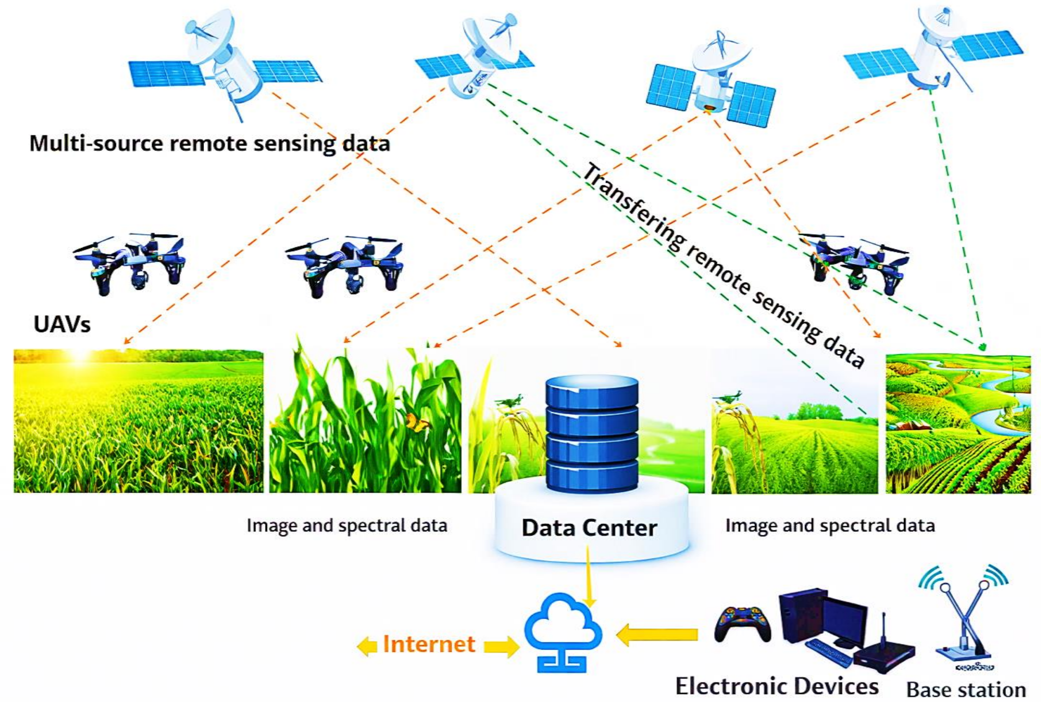

Fig 1. The Widespread Use Platform of Numerous Remote Sensing Satellites in Accuracy Cultivating.

The entire breadth and potential of the satellite imaging in achieving accurate agriculture can be achieved when hyperspectral satellite data become easily available in order to monitor agriculture. The application of hyperspectral measurements is done on the UAV (unmanned aerial vehicle) and aircraft level to monitor the biotic stresses in the olive orchards and the measurement of the nitrogen status in the wheat [3]. However, the presence of these demonstrations of remote-sensing ability does not necessarily mean that precision agriculture is a viable product. Poor data accessibility and the expensiveness of some of the tools makes their feasible implementation not possible with most farmers. Nonetheless, these restrictions do not annul the importance of such approaches as an essential contribution to the field of science and as a potential usage.

Scene labeling, also known as semantic segmentation, assigns semantic labels to every pixel in an image, which include automobile, people, and road. In comparison, instance segmentation detects and outlines the pixels related to each object instance. Panoptic segmentation incorporates semantic and instance segmentation and thus offers pixel-to- pixel annotation of a scene. Over the last few years, the intensive development of information-technological tools used in the contemporary agriculture sector has triggered the extensive usage of image-segmentation methods in monitoring crops and soils health, predicting optimal sowing, fertilization, and harvest dates, estimating crop yield, and identifying plant diseases [4].

The latest technologies, especially machine learning (ML) are reinventing the way agriculture is practiced. ML models can be used to identify agronomic problems fast and accurately thus enabling farmers to apply specific interventions. ResNet architecture addresses the vanishing-gradient issue that generally compromises the education of profound neural systems.

This mitigation is explained by the fact that residual blocks add direct shortcut connections to create an easy channel of transmitting information as well as gradients in the back propagation process. ResNet, therefore, permits building more complex topologies of networks without compromising performance.

The effectiveness of the ResNet in image-recognition tasks was proven with the help of empirical data, most notably, its high performance in the ImageNet competition. The research is aimed at formulating and testing a transfer-learning-realized residual U-Net model that has the capability to correctly segment high-resolution agricultural landscapes, especially fragmented smallholder plots, mixed-farming regions, and fine-scaled objects, such as irrigation lines and farm boundaries.

The rest of this paper will be organized in the following way: Section II will discuss the work related to land-use and land-cover segmentation and transfer learning methods. Section III contains a description of the materials and the methods, the nature of the dataset and its agricultural relevance, the architectural adjustments that are made in the proposed TL-ResUNet model, and the training procedures and the assessment outcomes. Section IV addresses the results of the experiment, including the training system and implementation strategy, the results of the performance evaluation and segmentation, the impact of transfer learning and increased optimization of the hyperparameters, and a comparison of the robustness with other methods. Lastly, in Section V, the study comes to an end with the summary of the key findings and the description of possible directions of the research in the future.

II. RELATED WORK

As Domingues, Brandão, and Ferreira [5] illustrates, techniques of segmentation models have been largely used in processing agricultural images to accurately identify the target objects, such as crops, pests, diseases and weeds, in natural environments that are typically in complex situations. Traditional methods of segmentation, which include thresholding, edge detection, clustering and region-based methods, are simple and can be understood but have significant failure points in their precision and reliability in changing field conditions.

Qadri et al. [6] described the rapid evolution of deep-learning methods that have had an immense implication on the field of image analysis, especially on crop-disease image segmentation. Deep-learning-based segmentation algorithms, including semantic, instance and panoptic algorithms, are consistently better in terms of predictive quality and computational efficiency, compared to traditional machine-learning based approaches. These new methods are more robust to perturbations, and have a more direct end-to-end training paradigm. As a result, many scholars have been considering deep-learning-based segmentation models to solve the issue of complex real-life situations.

As described by Choure and Prajapat [7], convolutional neural networks, especially U-Net, DeepLabv3+, and SegNet, have significantly improved the performance of segmentation. Agricultural image tasks such as plot segmentation, crop classification and the identification of leaf-disease have been widely applied using these models. They achieve high pixel-level precision by deriving the complex image characteristics of vast and labeled data sets. Scholars often make use of transfer learning, data augmentation, and attention in situations that lack sufficient samples and thus will eventually enhance image-segmentation methods in precision-agriculture studies.

According to Lim, Loh, and Wong [8], U-Net has become a center of attention architecture especially when it comes to low-contrast-contrast (LCC) tasks. It has previously been investigated in the manner of altering its constituent parts to better utilize the information given by multispectral images (MSI) to segmentation in multifaceted situations, particularly by altering its encoder. The other option is to substitute the U-Net encoder using a Residual Network (ResNet). A CNN with residual connections (ResNet) can prevent gradient vanishing and lessen information loss in the consecutive layers, as well as, allowing deeper network training and more advanced feature learning. Based on this, U-net is largely used with ResNet backbone when operating LCC.

Zocco et al. [9] used U-Net with a lightweight ResNet-18 encoder to classify RGB-NIR images into six categories; Tao et al. [10] used the ResNet-50 as a feature extractor of RGB-NIR image to outline metropolitan areas in China in eight categories. Bu et al. [11] improved the encoder by replacing the U -Net backbone with ResNet -101 in RGB NIR multicast to six categories.

According to Gholizade et al. [12], transfer learning is the most common methodological approach in situations where the scarcity of data is an issue. Modern vision tasks tasks commonly use a convolutional network pre-trained on a large compound dataset, usually ImageNet (a collection of 1.2 million images with 1,000 categories) and then also use the pretrained network as an initializer or a fixed feature extractor on the task being run.

Singh et al. [13] show that those trained on AgriNet models are better than those trained on ImageNet models. An example is that VGG16 with AgriNet weights had an accuracy of 90, versus 83 with ImageNet weights, and VGG19 with AgriNet weights had an accuracy of 88 versus 83 with ImageNet weights. VGG19 was the best on AgriNet dataset, but VGG16 was better on rice-pest and plant-disease databases, probably because different datasets had different distributions of features.

Sahili and Awad [14] therefore suggest that agricultural data should be evaluated on various AgriNet architectures to achieve an optimal performance. The above accuracies have been obtained using the technique of freezing the base models and training only the upper layers, which can also be further optimized in further experiments.

Therefore, the appropriateness of any particular architecture depends on what context of usage it is supposed to serve. **Table 1** represents a broad comparative analysis of the models. On the one hand, Xception architecture has the smallest footprint (88MB), on the other hand, InceptionResNet-v2 has the largest top-1 accuracy (0.803) and top-5 accuracy (0.953).

Xception model has the lowest number of parameters, but the depth of VGG16 is the shallowest, with 23 layers. The use of deep learning methods has significantly improved performance in agricultural disease and pest segmentation tasks and become the methodological paradigm.

Table 1. Comparison of the ImageNet Structures in AgriNet

Model	Size	Top-1 accuracy	Top-5 accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG19	549 MB	0.713	0.900	143,667,240	26
InceptionResNe-v2	215 MB	0.803	0.953	55,873,736	572
Inception-v3	92 MB	0.779	0.937	23,851,784	159
VGG16	528 MB	0.713	0.901	138,357,544	**23**

Liu et al. [15] used deep convolutional neural network (DCNN) to construct an automatic disease diagnostic system of cucumber, hence supplying a scientific rationale of accuracy in applying insecticides. Scholars in [16] suggested a multi-task neural-network architecture to categorize and predict the extent of coffee leaf illnesses; the system provides the practitioners and farmers with methods to detect and measure the level of biotic stressor in coffee plantations.

As seen in [17], traditional deep-learning methods demand large annotated datasets to train them, which, again, have not been very substantial in agricultural image segmentation. Therefore, the research on the extensions of zero-shot segmentation to this field is a significant and demanding research problem. It should be emphasized that most approaches to classical segmentation require manual choice of threshold which makes the extraction process more difficult and prone to noise and outliers, ultimately worsening the quality of segmentation and generalization. In addition, modern image-segmentation methods are in general unable to provide the accuracy and classification capabilities required to perform effectively on complex remote-sensing images, and therefore cannot satisfy the high precision requirements of practice.

In [18], convolutional neural network models that use the power of deep learning have been trained to adjust their parameters through a loss function back-propagation process on training data, which have shown strong learning performance and high performance. The model DeepLabv3+ (classical semantic-segmentation model) has an encoder-decoder architecture that combines both shallow and deep semantic representation in a holistic form and uses depth-wise separable convolutions in a spatial pyramid pooling module which reduces the total number of parameters significantly and at the same time improve the accuracy of the segmentation outputs. There have been consistent improvements of DeepLabv3+ to meet a wide range of segmentation needs.

A hierarchical attention model proposed by Jiang and Zhou [19] was intended to improve the cross-level semantic integration process; still, there are difficulties in partitioning large and fragmented objects and splitting them into smaller parts. The combination of Atrous convolution and varying dilation rates with the receptive field enhances the Atrous space pyramid pooling element in the DeepLabv3+ system, and thus it explores the probability of exploiting global contextual data through the pyramid paradigm presented in [20].

III. MATERIALS AND METHODS

Characteristics of Dataset and Agriculture Significance

The dataset is of high-resolution land-use and land-cover data, which is supposedly tailored to agricultural inferences and maps of rural landscape. It contains satellite images marking the main agricultural types like cropland, pasture, mixed farming, forest fragments, water bodies and rural settlements. Field geometries, irrigation canals, hedgers, and fragmented planting patterns as are common in smallholder agricultural systems of many developing regions were carefully marked in polygonal masks on each image. The images were up-sampled to a common spatial resolution of 0.5 m px^{-1} to have consistency among the samples used in the training.

To boost the agricultural relevance, a detailed preprocessing pipeline was used to boost vegetation related features and delineations of boundaries. First, the histogram equalization was used in order to enhance the field-texture variability. Then, the normalized vegetation indices based on the red and near-infrared spectral bands were added as auxiliary channels to facilitate the effective discrimination among the healthy crops, stressed vegetation, fallow land, and forested areas. Lastly, a stride-over grid was used to divide images into 512 x 512-patches maintaining long-range spatial features like terrace farming, irrigation networks and field boundaries. **Table 2** provides a brief synopsis of the agricultural land -cover class annotations.

Table 2. Overview of Land Agricultural Covers in The Dataset

Land-Cover Class	Number of Labeled Samples	Dominant Characteristics
Cropland	4,120	Row composition, different plant density
Mixed Farming Areas	3,870	Crops Interplanted, Small Plot Fragmentation
Forest Patches	2,940	Lush green growth, chaotic natural borders
Pastureland	2,110	Smooth texture, predominates with grass
Water Bodies	1,560	Canals, reservoirs, small streams
Rural Settlements	1,980	Constructed roofs, well-established borders

The integration of narrow rivers, small forest patches and very fragmented small-holder plots make this data agriculturally realistic and representative of those regions that are typified by mixed land-use patterns.

Adaptations in Architecture and Agriculture

The offered architecture is the altered version of the UNet-based semantic segmentation model, which is known as the TL-ResUNet, which is aimed at serving the complexity of the agricultural landscapes. The ResNet-50 backbone was incorporated at the encoding phase and it was effectively used to obtain hierarchical features of the crop textures, tree canopy layers, as well as boundary lines of the field. The remaining connections in the ResNet-50 can be also attributed to the maintainability of the spatial coherence that allows the model to sustain the fine transitions (soil-crop and narrow irrigation paths). The decoder has the transposed convolution layers that upsample the feature maps and then connects them to the encoder layers themselves and skip connections allow the network to recover fine-grained feature in agriculture.

The model had special adaptations added to it to make it sensitive to the agricultural characteristics. Attention weights followed the deepest layers of the encoder and directed the network to the significant locations in space, particularly the ones undergoing minute differences in the vegetation patterns, e.g. the margins of a farmland where the vegetation cover is light. Also, the dilated convolutions were used in the bottleneck to expand the receptive field allowing the model to incorporate contextual information needed to detect long continuous features like terraces or linear hedges that surrounded farm plots. **Fig. 2** illustrates a conceptual structural overview of the proposed architecture.

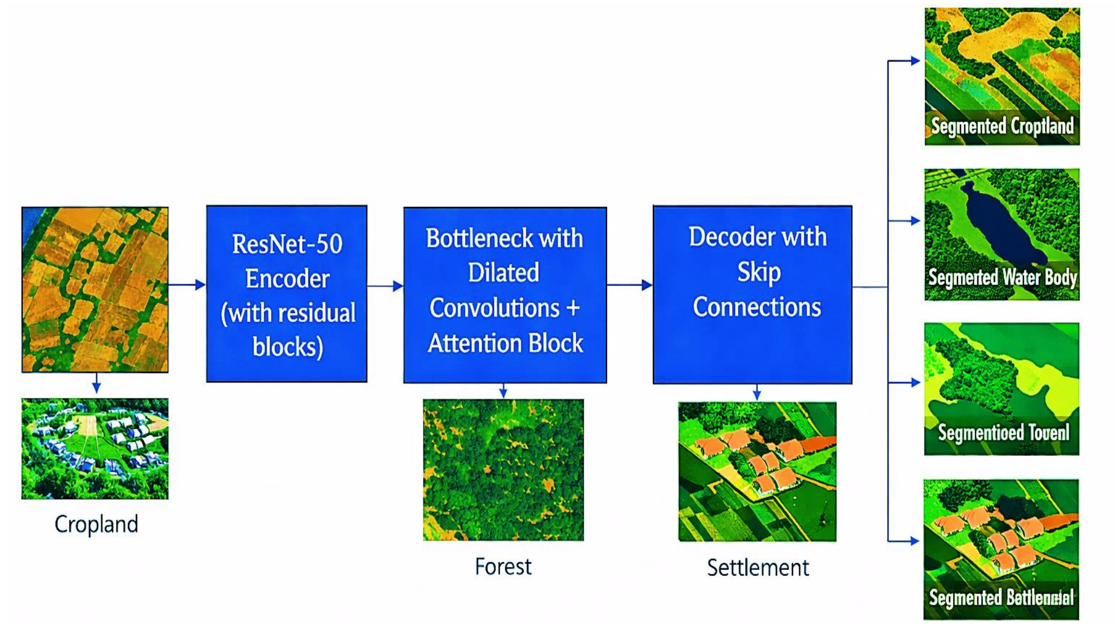

Fig 2. Architecture Structural Overview of the TL-ResUNet Architecture.

This structure ensures that macro-scale and micro-scale agricultural components are well represented to enhance the accuracy of segmentation in the real farming scenario.

Training Strategy and Measures

The training approach aimed at maximizing the detection capability of the model to different agricultural lands. Three weight initialize scenarios had been investigated. In the former, the pre-initializations of all the encoder and decoder weights were performed using LeCun uniform distribution, to obtain a pure data-driven basis. ResNet-50 and ImageNet pre-trained encoder weight and randomly initialized decoder weight were the second one. Finally, weights in scenario three were pre-trained using an existing agricultural segmentation model on a sub-sampling of the data, and therefore, provided advantages of domain-specific transfer learning.

The training was done in two phases to stabilize learning. In the Stage 1, encoder layers were frozen, and decoder and classification layers were only updated during 13 epochs to quickly enable the model to conform to the agricultural output masks. In Stage 2, the encoder layers were again unfrozen and further fine-tuned during 7 further epochs to further curtail the identification of the slightest amounts of crop textures and boundaries. Initial Learning rate was 0.0001 and cosine annealing schedule were Adam optimizers which were used to scale parameter updates progressively.

The performance of each model was evaluated by the Intersection over Union (IoU) of each agricultural type to assess the quality of the land parcels, small fields or narrow water channels that were represented by the network. The 20% validation split was applied in order to make sure that it was tested on the pictures of the agricultural landscapes that were not seen. It was implemented using the PyTorch architecture on a GeForce Tesla V100 and 43GB of RAM in a server with a lot of memory to conduct batch training at large scales. Training performance summaries are provided in **Table 3**.

Table 3. Training and Validation IoU Scores Under Various Initialization Methods

Epoch	Random Initialization (IoU)	ImageNet Pre-train (IoU)	Domain-Specific Pre-train (IoU)
10	0.48	0.55	0.59
20	0.57	0.67	0.73
30	0.66	0.76	0.82

IV. RESULTS AND DISCUSSION

Semantic segmentation of DeepGlobe dataset was done using we used improved UNet model. DeepGlobe dataset was used in the training and evaluation of optimized TL-ResUNet architecture that incorporated a fine-tuned and initialized ResNet-50 encoder. The PyTorch architecture is the training framework of choice in most cases when it comes to deep learning and machine learning model training.

Training System and Model Implementation

The modified encoder UNet architecture (see **Fig. 3**) is realized by employing the PyTorch model by incorporating ResNet-50 elements as an encoder. Servers that that were used during the training process were the GPU servers which had Tesla V100 graphics cards and 43GB of RAM. ResNet-50 is a model that is trained using DeepGlobe datasets and adjusted using ResNet-50 weights that is pretrained using ImageNet dataset. This architecture is learned into two phases as the first phase of training only integrates training the final layers but the second stage will involve the unfreezing of the entire layers and thereafter training. This architecture was trained using 20 epochs of which 13 epochs were utilized in the first stage and 7 epochs in the second stage.

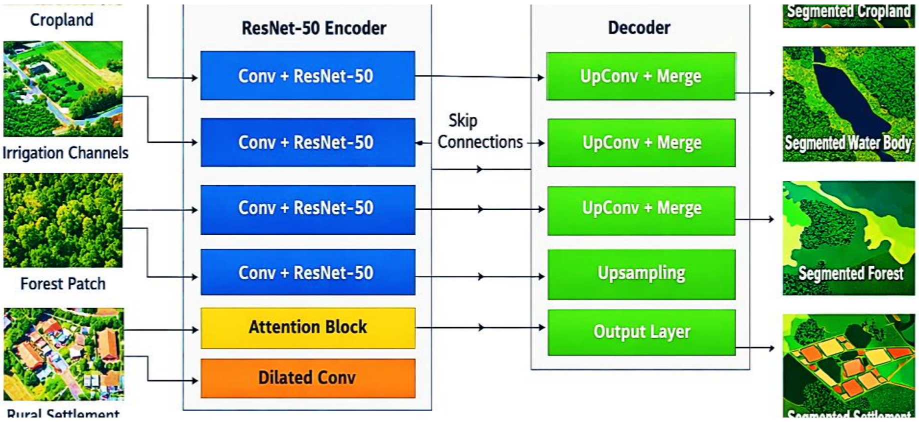

Fig 3. UNet Modified Design.

There are three initializing methodologies of weight that were discussed during the training. Initializer The initializer was the LeCun uniform initializer, which was used to set the initial weights. This initializer is uniformly distributed on the interval [-L, L] and L refers to the input unit number in the tensor and fin is the square-root of 1/fin. Second, we have used ResNet-50 encoder and our model that is already ImageNet-trained. The uniform initializer of LeCun was used to initialize all the layers of the decoder. As it is shown in **Fig. 4**, we employed the current segmentation architecture by pre-training the encoder using ResNet-50 and using it to get an initialization.

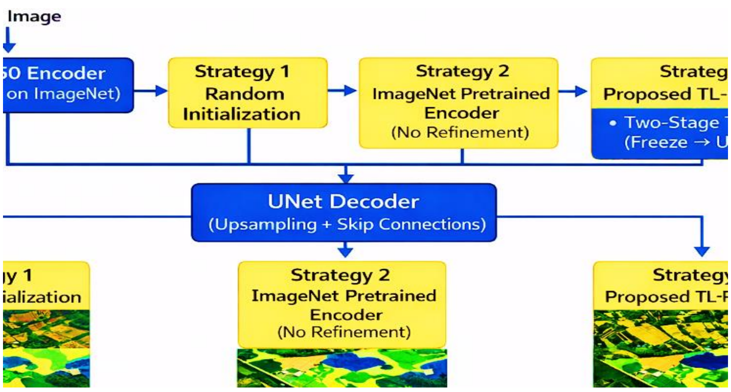

Fig 4. TL-ResUNet Architecture and Strategies.

Result of Performance Evaluation and Segmentation

Our results on DeepGlobe's validation subset were as shown below after 30 training cycles:

- The best result of the IoU = 0.68 is attained when randomly initialized weights are used.
- The desirable result with the prepared encoder weights of the ImageNet encoder is the Intersection over Union (IoU) of 0.81.

Fig. 4 depicts that the model works effective within the majority of conditions, whereas it could completely fail at considering other categories, including narrow bodies of water. An example of this is as shown below. The same can be said with the misclassification of small, forested areas [21] (see **Fig. 5**). Even those that are found close to the agricultural plains are considered with the correct classification of forests that are densely covered. The line of demarcation of the agricultural land and the forested areas is not an easy task. In addition, the model works extremely well in certain classes like farm and city land.

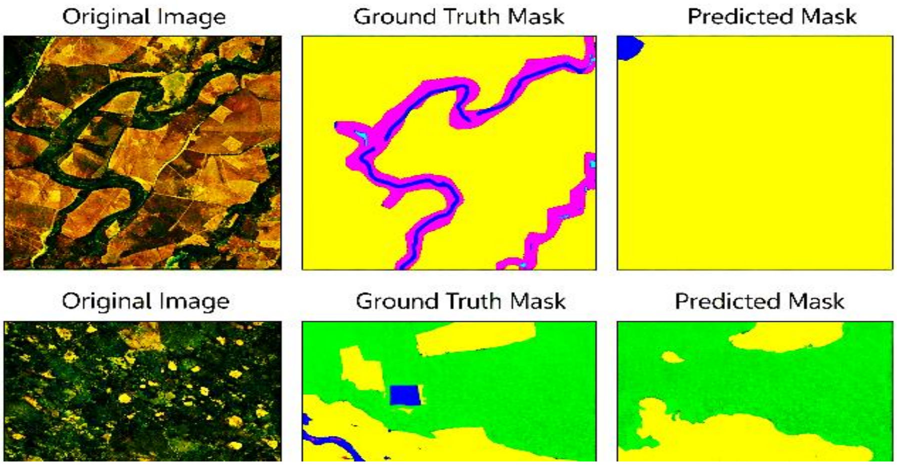

Fig 5. Samples of Incorrect Model Classifications.

The validation learning curves represented in **Fig. 6** below represent the outcomes of every strategy. The pre-trained and randomly initialized networks both stabilize at a constant value much faster and the former appear to stabilize at a larger constant value as well.

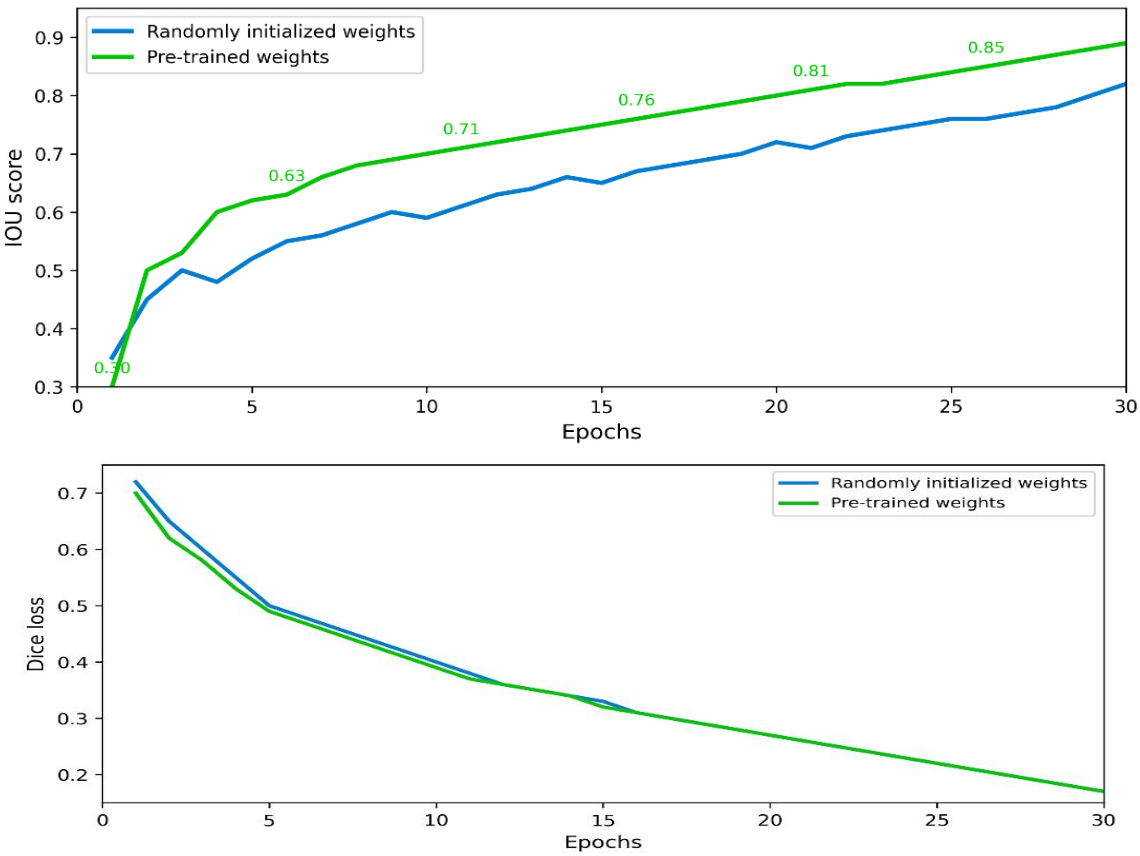

Fig 6. Evolution of Learning Curves of The Learned Model During Training and Validation Stages.

Transfer learning and HP optimization Effect

The illustration of superimposing a replica of the initial image explains the benefit of using pre-learned architectures in training. It should be noted that hyperparameter optimization methods and data pretreatment can be employed in order to improve the work of the architectures. The overall scores are outlined in **Table 4**.

Table 4. UNet Training Comparison Results in Terms of the IoU Measure Per Epoch

Epoch	Train/Validation	UNet Trained by the Noise Cancellation Weights	UNet Trained Weights on ImageNet without Residual Layers	UNet with ImageNet Trained Weights
10	train	0.51	0.52	0.54
	validation	0.49	0.50	0.51
20	train	0.59	0.62	0.69
	validation	0.58	0.61	0.64
30	train	0.68	0.75	0.84
	validation	0.68	0.74	0.81

In general, it could be viewed that UNet learned with the help of a residual layers and transfer learning accelerated feature learning and more efficiently. UNet random weighted of 30 epochs achieves the same ionu of 0.68 as UNet weighted randomly on ImageNet weight of 0.74. Lastly, the Unet using the ImageNet weights of the ResNet50 achieved a score of 0.81 in IoU.

To demonstrate how our model is effective compared to others, we present the comparison data in **Table 5**. ClassifierNet generates equitable segmentation results in large regions, but does not work in segmentation of small features, small regions, and field boundaries. DeepLabv3 + is a better process that enhances the extraction of features by using Atrous convolution, Atrous spatial pyramid pooling and upsampling. One can obtain a more accurate localization of an object by categorizing all pixels within a particular object.

Up to now, DeepLabv3+ architecture has not been applied in detection and localization tasks. Moreover, DeepLabv3+ is a single-stage detector, therefore, continuing the problem of class imbalance. DeepLabv3+ and DeepLabv3 enhance information on complex specifics; nonetheless, they are also prone to artifacts and they do not always provide consistent information on a larger scale. The multi-level characteristics are well incorporated in our model and the segmentation outcomes are more accurate in larger and detailed areas.

Table 5. TL-ResUNet Comparison and Other Models

Algorithms	IoU
TL-ResUNet	81.02
DeepLabv3+	75.61
DeepLabv3	74.52
DFCNet	71.31
ClassmateNet	69.87
Baseline	55.19

Comparison of Robustness Analysis With the Existing Methods

We associated the limitations and advantages of the previous approaches with the suggested methodology in numerous aspects, both qualitative and quantitative performance indicators, as presented in **Table 6**. The tested ratings show that the proposed approach was doing well in densely wooded areas, being able to classify the latter correctly even when they were close to areas full of farmland. In addition, the model records high performance in certain types such as the urban and agricultural environments.

Table 6. Evaluation of Segmentation Approaches Using Attributes

Criterion	Proposed Method	DeepLabv3+	DeepLabv3	DFCNet
Small Land Segmentation	robust	robust	standard	robust
Multiple Land Segmentation	powerless	powerless	powerless	standard
Processing Time	robust	robust	standard	powerless
Robust to Color	standard	powerless	standard	standard
Robust to Noise	robust	robust	robust	standard
Object Independence	standard	robust	robust	standard
Scene Independence	robust	standard	robust	standard

The results of segmentation approaches may be categorized into three (standard, ineffective, and robust) groups. Thorough measurements also reveal that the process is appropriate in dividing every type of soil and field types. The technique can be inefficient in certain circumstances, e.g., small water bodies or small wooded areas, conventional standards. There is inconclusive evidence that comparing algorithms to noise or color is inconsistent, and the process of categorizing the land usually distorts the original geometry of the dynamically changing objects.

V. CONCLUSION

This paper has shown that transfer learning and residual feature extraction implemented as a UNet can significantly improve the land-use segmentation of agricultural land. The suggested TL-ResUNet has the capacity to reproduce larger scales of land patterns and minute details, confronting the problems of heterogeneous agricultural systems and discontinuous farm plots. Findings validate that encoder pre-training, especially in a domain-specific agricultural dataset speed up the convergence rate and enhance the segmentation of agriculture data. Though there are still limitations in the detection of very narrow water bodies and small forest patches, the general strength of the model among the various classes of land-cover opportunities is the demonstration of the appropriateness of this model to be applied in the real world of agriculture. Future research will concentrate on the better optimization of hyperparameters, integration of multi-temporal data and a better management of class imbalance to further improve the accuracy of segmentation and generalization.

CRediT Author Statement
The author reviewed the results and approved the final version of the manuscript.

Data Availability
The datasets generated during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interests
The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding
No funding was received for conducting this research.

Competing Interests
The authors declare no competing interests.

References
[1]. T. Hengl et al., “SoilGrids250m: Global gridded soil information based on machine learning,” PLoS ONE, vol. 12, no. 2, p. e0169748, Feb. 2017, doi: 10.1371/journal.pone.0169748.
[2]. R. Selvaraj and D. G. B. Amali, “A new Texture Aware—SEed demand enhanced Simple Non-Iterative Clustering (ESNIC) segmentation algorithm for efficient land use and land cover mapping on remote sensing images,” IEEE Access, vol. 12, pp. 121208–121222, Jan. 2024, doi: 10.1109/access.2024.3519612.
[3]. D. Argüello, M. Noguera, F. Vaz, A. Cordeiro, J. Silvestre, and A. Aquino, “Leaf Nitrogen Content Estimation in Olive Orchards based on Analysis and Modelling of UAV-Acquired Hyperspectral data,” Smart Agricultural Technology, p. 101754, Dec. 2025, doi: 10.1016/j.atech.2025.101754.
[4]. S. Padhy et al., “AgriSecure: A FOG Computing-Based security framework for agriculture 4.0 via Blockchain,” Processes, vol. 11, no. 3, p. 757, Mar. 2023, doi: 10.3390/pr11030757.
[5]. T. Domingues, T. Brandão, and J. C. Ferreira, “Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive survey,” Agriculture, vol. 12, no. 9, p. 1350, Sep. 2022, doi: 10.3390/agriculture12091350.
[6]. S. A. A. Qadri, N.-F. Huang, T. M. Wani, and S. A. Bhat, “Advances and Challenges in Computer Vision for Image-Based Plant Disease Detection: A Comprehensive survey of machine and deep learning approaches,” IEEE Transactions on Automation Science and Engineering, vol. 22, pp. 2639–2670, Apr. 2024, doi: 10.1109/tase.2024.3382731.
[7]. P. Choure and S. Prajapat, “Exploring U-Net, FCN, SegNet, PSPNet, Mask R-CNN and using DeepLabV3+ for multiclass semantic segmentation on satellite images of Western Ghats,” in Lecture notes in networks and systems, 2024, pp. 783–806. doi: 10.1007/978-3-031-74443-3_46.
[8]. C. C. Lim, Y. P. Loh, and L.-K. Wong, “LAU-Net: A low light image enhancer with attention and resizing mechanisms,” Signal Processing Image Communication, vol. 115, p. 116971, Apr. 2023, doi: 10.1016/j.image.2023.116971.
[9]. E. Zocco, C. Witharana, I. M. Ortega, and W. Ouimet, “Automated detection of Beaver-Influenced floodplain inundations in Multi-Temporal aerial imagery using deep learning algorithms,” ISPRS International Journal of Geo-Information, vol. 14, no. 10, p. 383, Sep. 2025, doi: 10.3390/ijgi14100383.
[10]. C. Tao et al., “MSNet: multispectral semantic segmentation network for remote sensing images,” GIScience & Remote Sensing, vol. 59, no. 1, pp. 1177–1198, Aug. 2022, doi: 10.1080/15481603.2022.2101728.
[11]. Y. Bu et al., “ResNet incorporating the fusion data of RGB & hyperspectral images improves classification accuracy of vegetable soybean freshness,” Scientific Reports, vol. 14, no. 1, p. 2568, Jan. 2024, doi: 10.1038/s41598-024-51668-6.
[12]. M. Gholizade, H. Soltanizadeh, M. Rahmanimanesh, and S. S. Sana, “A review of recent advances and strategies in transfer learning,” International Journal of Systems Assurance Engineering and Management, vol. 16, no. 3, pp. 1123–1162, Feb. 2025, doi: 10.1007/s13198-024-02684-2.
[13]. G. Singh, R. S. Singh, R. Mehta, and V. K. Nayak, “AgriNet: Leveraging MobileNetV2 with Explainable AI and Drone Imaging for Efficient Plant Disease Diagnosis,” in Lecture notes in networks and systems, 2025, pp. 254–262. doi: 10.1007/978-3-031-91005-0_24.
[14]. Z. A. Sahili and M. Awad, “The power of transfer learning in agricultural applications: AgriNet,” Frontiers in Plant Science, vol. 13, p. 992700, Dec. 2022, doi: 10.3389/fpls.2022.992700.

[15]. J. Liu, X. Wang, Q. Chen, P. Yan, and X. Liu, "Deep learning method for cucumber disease detection in complex environments for new agricultural productivity," BMC Plant Biology, vol. 25, no. 1, p. 888, Jul. 2025, doi: 10.1186/s12870-025-06841-y.
[16]. R. S, S. R, S. M, and K. L, "Multi-Task Distillation Learning for Coffee Corticium Salmonicolor Pink Berry Disease for Real-Time Prediction," 2024 International Conference on Sustainable Communication Networks and Application (ICSCNA), pp. 1065–1071, Dec. 2024, doi: 10.1109/icscna63714.2024.10863987.
[17]. L. Lei, Q. Yang, L. Yang, T. Shen, R. Wang, and C. Fu, "Deep learning implementation of image segmentation in agricultural applications: a comprehensive review," Artificial Intelligence Review, vol. 57, no. 6, May 2024, doi: 10.1007/s10462-024-10775-6.
[18]. N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 0588–0592, Apr. 2017, doi: 10.1109/iccsp.2017.8286426.
[19]. X. Jiang and X. Zhou, "Multiple semantic processes at different levels of syntactic hierarchy: Does the higher-level process proceed in face of a lower-level failure?," Neuropsychologia, vol. 50, no. 8, pp. 1918–1928, Apr. 2012, doi: 10.1016/j.neuropsychologia.2012.04.016.
[20]. Z. Lyu, S. Liu, F. Sun, Z. Guan, and Z. Lyu, "Enhanced DeepLabV3+ for wheat grain segmentation: Addressing tiny targets, blurred edges, and class imbalance," Journal of Agriculture and Food Research, vol. 26, p. 102680, Jan. 2026, doi: 10.1016/j.jafr.2026.102680.
[21]. F. Safarov et al., "Improved Agricultural Field Segmentation in Satellite Imagery Using TL-ResUNet Architecture," Sensors, vol. 22, no. 24, p. 9784, Dec. 2022, doi: 10.3390/s22249784.

ISSN: 3104-4654