



ISSN 2047-3338

Road Features Extraction from Aerial Images Using Deep Learning

Waseem Akram¹, Aasim Ghaffar², Anjum Ali³, Muhammad Jehanzeb⁴, Muhammad Rizwan Amin⁵, Awais Rasool⁶

^{1,4,5,6}Department of Software Engineering, The University of Lahore, Pakistan

²Department of Physics, Astronomy and Mathematics, University of Hertfordshire

³Department of Software Engineering, Riphah International University Faisalabad, Pakistan

¹waseem.akram@se.uol.edu.pk, ²aasimghaffar442925@gmail.com, ³malik.anjumali@gmail.com, ⁴muhammad.jehanzeb@se.uol.edu.pk,

⁵Rizwan.amin@se.uol.edu.pk, ⁶awais.rasool@se.uol.edu.pk

Abstract– Roads are very important for transportation. Now, there is a dire need for urban development due to the increase of world population. Roads are the main part of the transportation system. Therefore, in an emergency case the extraction of roads is the main thing. In the past many methods of research were used to find the roads. An advanced method that researchers use in our research has better results than past methods. Therefore, there are many issues faced with different road management. In some cases, there is research on some roads that research is very difficult to extract from the satellite or aerial image in past. Road extraction from the aerial image is used for smart city and road monitoring. This is also used for emergency response, in the difficult areas where roads are not research prepared, and this technique gives us extraction of alternative roads. That is also used for traffic management, low cost, less time-consuming and more accuracy. Two main phases are used firstly; the use of deep learning advanced model that can extract road more accurately than others can methods that are used for the predicting of road segments. Secondly, research applies some advance-thinning algorithm to find the road centerlines from the complex urban areas. That is Automated Road extraction from satellite or aerial image provide the very efficient and accurate results in any urban and rural areas. The purpose method provides good percentage results in any type of road. That extract road with low cost and less time consuming. In this research road extraction from a arial image accuracy is 98%, Precision 92%, Recall 94%, F-Score is 96%.

Index Terms– Road Extraction from Aerial Images, Deep Learning, Satellite Image Analysis, Remote Sensing and Image Segmentation

I. INTRODUCTION

ROADS form an integral part of transport infrastructure that plays a role in various applications such as GIS, navigation, urban planning, disaster management, and smart city building. It is very critical to have precise and updated information on roads in the underdeveloped regions due to the incompleteness of the map of roads, which makes the distribution of resources and disaster management difficult. The availability of aerial and satellite images of high resolution

has made it possible to automatically extract road networks from spatial databases at relatively lower costs (Veda and Suma, 2016). Manual road extraction techniques use image processing techniques such as morphological operations, erosion, dilation, and thinning algorithms to improve road features and extract the centerline of the road from aerial images (Ben Boudaoud et al., 2018).

However, improvements in deep learning, especially Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs), enable automatic feature extraction and semantic segmentation of roads in images. Many researchers have proved that CNNs are efficient in extracting road networks using aerial images (Zhang et al., 2018; Xu et al., 2018). Although there has been progress in this field, some problems, including fragmented road networks, incomplete roads, and low efficiency of maintaining topological structure of roads, still exist (Sharifi et al. 2021).

Despite the positive outcomes of this research, there are still some drawbacks which need to be addressed. Extracted road networks are usually comprised of disconnected road pieces, lacking branches and road width representation. The performance of the extracted road network highly depends on the complexity of the scene, and disconnected road segments cause fragmentation of network connectivity. Also, updating information about roads manually requires a lot of time and effort. Current researches call for a method which will not only improve road segmentation accuracy but will also preserve road connectivity in complex scenes (Buslaev et al. (2018)).

The proposed method is aimed at solving these problems by developing an automatic road feature extraction framework from aerial images based on deep learning technologies. The framework will consist of the road segmentation process using deep learning and image processing techniques in order to improve the accuracy and connectivity of road networks. The extracted information may be used for smart cities development, road monitoring and traffic management, as well as rescue operations in disasters like earthquakes and fires in rural and urban areas. Section II provides an exclusive overview of the previous study. Section III describes the proposed method for feature extraction. Section IV focuses on

the experiments conducted to evaluate the proposed methodology. Finally, the research concludes with a summary of the main findings and contribution along with future work in Section V.

II. LITERATURE REVIEW

Xie *et al.* (2019) explained that in the realm of remote sensing image analysis, the extraction of roadways from aerial pictures. To extract the route region, a semantic segmentation neural network with U-Net and residual learning. Senthilnath *et al.* (2020) this study offers a sample and structure-oriented network for road fracture detection. The suggested network, in particular, is built on the U-Net design, which uses a skip communication technique to transport data from input to output. Shan and Fang, (2020) Research present a structured deep neural network for a seamless and continuous road structure, inspired by the distinctive road structure, extraction network displays cutting-edge techniques and baselines. Wang *et al.* (2020) introduced Convolutional Neural Networks (CNN) have recently demonstrated promising results in semantic segmentation. U-Net, one of the most effective CNN-based systems has performed research. Abdollahi *et al.* (2021) reported the proposed algorithm based upon canny edge detectors for road network extraction. Detection of road completed by edge detection approach which diagnosis the curved and straightforward road without beginning pixel identification. Breresechr *et al.* (2021) give significance in traffic management, urban planning, GPS navigation, and disaster management, road extraction is analyzed using Deep U-Net model on the Massachusetts Road Dataset. Corbane *et al.* (2021) studies an effective way of extracting high resolution road networks from satellite imagery is proposed by means of using sector descriptors, MLSOH descriptors, adaptive correction, and multi-angle beam descriptors.

Ekim *et al.* (2021) presents a novel road network extraction framework based on the dense and skip connections within the DenseUNet architecture using remote sensing images as input. Feng *et al.* (2021) research offer a novel approach for extracting routes from aerial pictures centered on an efficient network of generative adversaries, a comprehensive framework that only takes a few training examples in this study. Hu *et al.* (2021) use a deep residual convolutional neural network (RDRCNN) with a post-processing phase, research describe a novel strategy for extracting methods from optical satellite pictures. Lin *et al.* (2021) due to shadows and pauses, most existing deep learning network models produce inconsistent and partial outputs. A dual interest pathway extraction network (DA-RodNet) with some logical reasoning capabilities is presented as a solution to this problem. Zhu *et al.* (2021) analyzes road extraction from high resolution remote sensing images in diverse situations with more than 88% accuracy and completeness.

Uddin *et al.* (2021) explored the traditional template matching methods based on intrinsic texture smoothing have a poor level of automation due to noise in geometry and texturing. Experiments indicate that the completeness and

accuracy of techniques with a given width is greater than 89%. Feng and Tu (2022) explores the problem of road extraction from low resolution aerial images with the help of MSR algorithm, edge detection, noise filtering, and road shape characteristics. Jiang *et al.* (2022) In remote sensing analysis, change detection utilizing remote sensing photos plays a significant role. To address the obstacles for change detection utilizing HR remote sensing photos. Jiao *et al.* (2022) research utilized the CLAHE method to improve image quality by softening the mountain shadow effect. The improved DSD Net was then used to extract the route with accuracy for the three plots was 0.8631, 0.8558, and 0.8801. Saralioglu and Gungor, (2022) assessed Ad boost-like Light Multi-U-Nets Model (AEML U-Nets) for method extraction is proposed as a solution to the previously mentioned challenge. Sharifi *et al.* (2022) present a novel fusion network (FuNet) that integrates remote sensing images with information about a location, with location data playing an essential role in linking route logic. To improve network-learning capabilities, FuNet incorporates a Global Redundancy Enhancement (ITER) module.

Sultonov *et al.* (2022) explored the combination of the Dirichlet mixture model (DMM) and a deep learning CNN model for method extraction on the Shao Shan dataset. Xu *et al.* (2022) this research presents the Enhanced Context-Aware Global Network, a deep convolutional neural network with context awareness (EGC-NET) via ASPP. NET outperforms the most recent methods, according to our tests. The EGC-NET received an F1 score of 0.8234 with an accuracy of 0.9466. Zhang *et al.* (2022) use a deep residual de-convolutional network (DeepResDCLNet) to excerpt road characteristics from UAV pictures in this study. Deep neural networks from the Rich skip segmentation architecture connections in the ritual ottleneck, with direct connections between from pixel to intermediate feature maps decoding methods are all used in the proposed network. He *et al.* (2023) provides an efficient method of propagating semantic labels using road network information and super pixels with much better IoU scores than the conventional scribble-supervised approaches.

III. RESEARCH METHODOGLY

A method was presented to find the road from the different areas and extract road centerlines from aerial pictures, enabling vehicle navigation on routes not covered by standard navigation software. The suggested rural road recognition method compares ResUnet, Unet, and FPN deep learning models. Later, morphological algorithms optimize linkages between distinct kinds of roads, and finally, a revolutionary thinning strategy extracts rural road centerlines. In any country, roads are one of the important, manufactured mechanisms. As a result, road network maps must be kept up to date and used for traffic management (Couloigner and Zhang, 2007).

Fig. 1 shows the planned methodology fit for road feature extraction from a satellite image. It consists of different steps of algorithm described as follows:

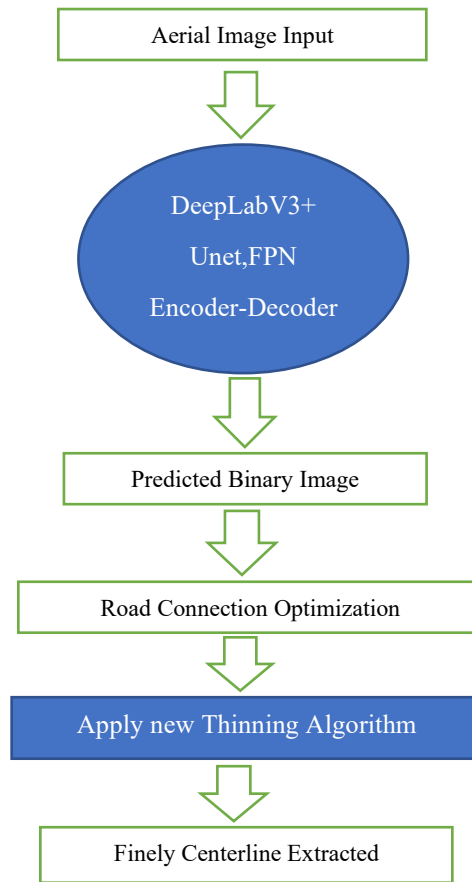


Fig. 1: Road Extraction Process

A) Dataset Preparation

The data included 61 orthorectified aerial images of size $16,000 \times 10,000$ pixels from the Landsat data collection, each having 8-bit RGBI resolution with a resolution of 0.25 m and being in TIFF format. The RGB channels only were used for analysis in this work. Ground truth data were obtained through labeling of the rural roads within images. For preprocessing, big images were cut into tiles of size 1024×1024 pixels, resulting in 486 image-mask pairs for training and evaluation. The experiments were done using Google Colab in Python. Data augmentation methods such as rotations, translations, scaling, flipping, cropping, and adding noise were used for better generalization of models. A modern semantic segmentation architecture known as DeepLabv3+ expands upon the DeepLabv3 concept.

B) Encoder-Decoder Architecture

DeepLabv3+ uses an encoder-decoder architecture shown in Fig. 2 for semantic segmentation, where an encoder that uses pre-trained CNN (such as ResNet or Xception) extracts high-level feature maps from the input image and the decoder up-samples these features to produce the output segmentation map (Shan *et al.*, 2021).

Atrous Spatial Pyramid Pooling (ASPP): Uses atrous convolutions with different dilation rates in order to capture multi-scale contextual information (Badrinarayanan *et al.*, 2017).

Image-level feature fusion: DeepLabv3+ is the image-level feature fusion via global average pooling and 1×1 convolutions.

Skip connections and up sampling: Skip connections enable combination of low-level and high-level features (Abdollahi *et al.*, 2021).

C) Road Detection Classifier Definition

Atrous Spatial Pyramid Pooling allows the model to capture multi-scale context through atrous convolutions with varying dilation factors. Thus, the model can control its receptive fields and extract features from different sized objects. The convolution procedure is described by Equation (1) when the feature map output is assumed to be Y , the convolution kernel is assumed to be w , and the input feature map is assumed to be x , the atrous rate r , for each position i .

$$(1) \quad Y[i] = k \cdot x[i + r \cdot k] \cdot w[k]$$

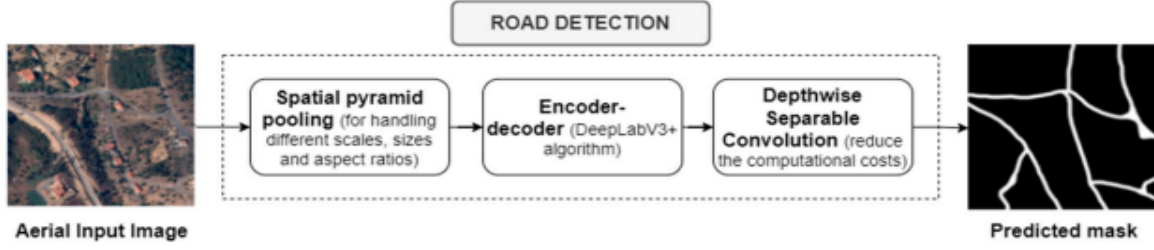


Fig. 2: DeepLabV3+ Architecture

Depthwise Separable Convolution:

DeepLabV3+ performs depthwise separable convolutions using channel-wise spatial convolutions and then performing 1×1 pointwise convolution. For detecting roads in imagery, DeepLabV3+ makes use of ResNet50 architecture pretrained on the ImageNet dataset. It first applies 1×1 and 3×3 convolutions with dilation rates of 12, 24, and 36 and then fuses the features and performs upsampling. The decoder fuses encoder features with upsampling of spatial size to restore it to its original value see Fig. 3 (Corbane *et al.*, 2021).

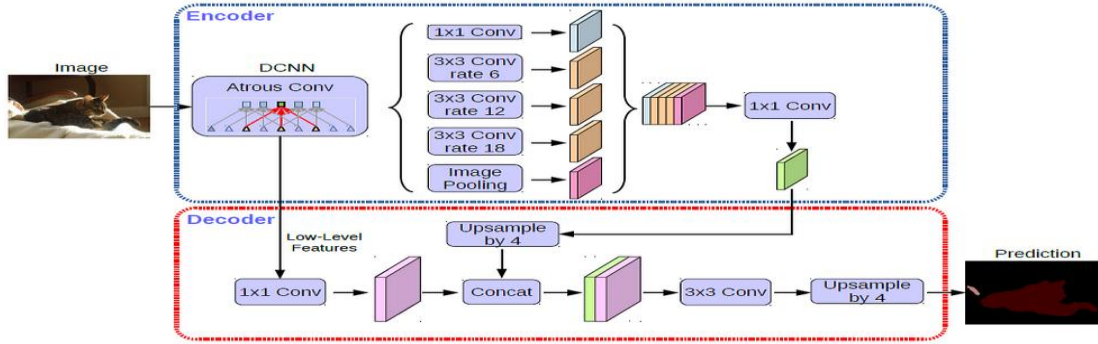
Fig. 3: Road detection module on rural road centerline extraction method(Lin *et al.*, 2021).

Table I shows the Configuration Parameters of the DeepLabV3+ Model for Road Extraction.

Table I: Configuration Parameters of the DeepLabV3+ Model for Road Extraction

Parameter	Value	Parameter	Value
Encoder Name	ResNet50	Classes	2 (Road, Background)
Encoder Depth	5	Activation Function	Sigmoid
Encoder Weights	ImageNet	Upsampling Factor	4
Encoder Output Stride	16	Epochs	40
Decoder Atrous Rates	(12, 24, 36)	Loss Function	Dice Loss
Evaluation Metrics	Jaccard Index (IoU) and additional metrics	Learning Rate	8×10^{-5}

IV. RESULT AND DISCUSSION

A) Data Preprocessing

The road detection model performance was analyzed using a confusion matrix with four parameters: True Positives (TP) - correctly recognized road pixels, True Negatives (TN) - correctly recognized background pixels, False Positives (FP) - background pixels recognized as road pixels and False Negatives (FN) - road pixels that were not detected. The detection and extraction of the road region and the road centerline were estimated by four metrics: Completeness (COM), Correctness (COR), Quality (Q), and F1-score (F1). In order to consider small differences between manually annotated and extracted road centerlines, a matching process was applied based on buffers.

B) Splitting Dataset

Dataset splitting into training, validation, and test sets was done. Training set consisted of 6,626 images (80%). Training set was employed for model training. The validation set consisted of 1,243 images, whereas the test set included 1,100 images.

C) Road Extraction Calculation

The new approach was compared with U-Net and Feature Pyramid Network (FPN) under the same training conditions, such as pre-trained weights on ImageNet, batch size of 4, 2 output classes, sigmoid function, and the same optimization hyper parameters. The experiments showed that DeepLabV3+ reached the maximum accuracy and exhibited the highest performance in complex rural scenes, which include shadows and occlusions. Even though U-Net and FPN provided the best road connectivity in some cases, they yielded higher false positives than others. Minimizing false positives was considered the main goal; therefore, DeepLabV3+ can be named the most efficient algorithm for rural road extraction.

D) Model Training with Unet

This model of architecture consists of four convolutional layers with increasing filter sizes of 16, 32, 64, and 128. Max pooling after each convo layer. The model then flattens the output and passes it through two dense layers with sixteen units each, using the ResNet activation function. Finally, the model concludes with an output layer comprising four units and for multiclass classification a SoftMax activation function is used. Total parameters are 499204, trainable parameters are 499204 and non-trainable parameters are 0.

Table II: Epoch-wise Training and Validation Performance of the Unet Model

Epoch Range	Training Accuracy	Loss Trend	Validation Accuracy
1–10	0.9057 → 0.9584	Consistently decreases	Improving trend
11–20	0.9593 → 0.9652	Generally, decreases; slight increase at Epoch 16	Peaks at 0.9187 (Epoch 14)
21–30	0.9655 → 0.9704	Continues decreasing	Stable high performance
31–40	0.9704 → 0.9796	Reaches minimum loss of 0.2422 (Epoch 40)	Peaks at 0.9619 (Epoch 40)
Overall Epoch Evaluation	0.9396 → 0.9446*	Consistently decreases	Peaks at 0.9497 (Epoch 38)

E) Applying Unet Model On different areas of Faisalabad city

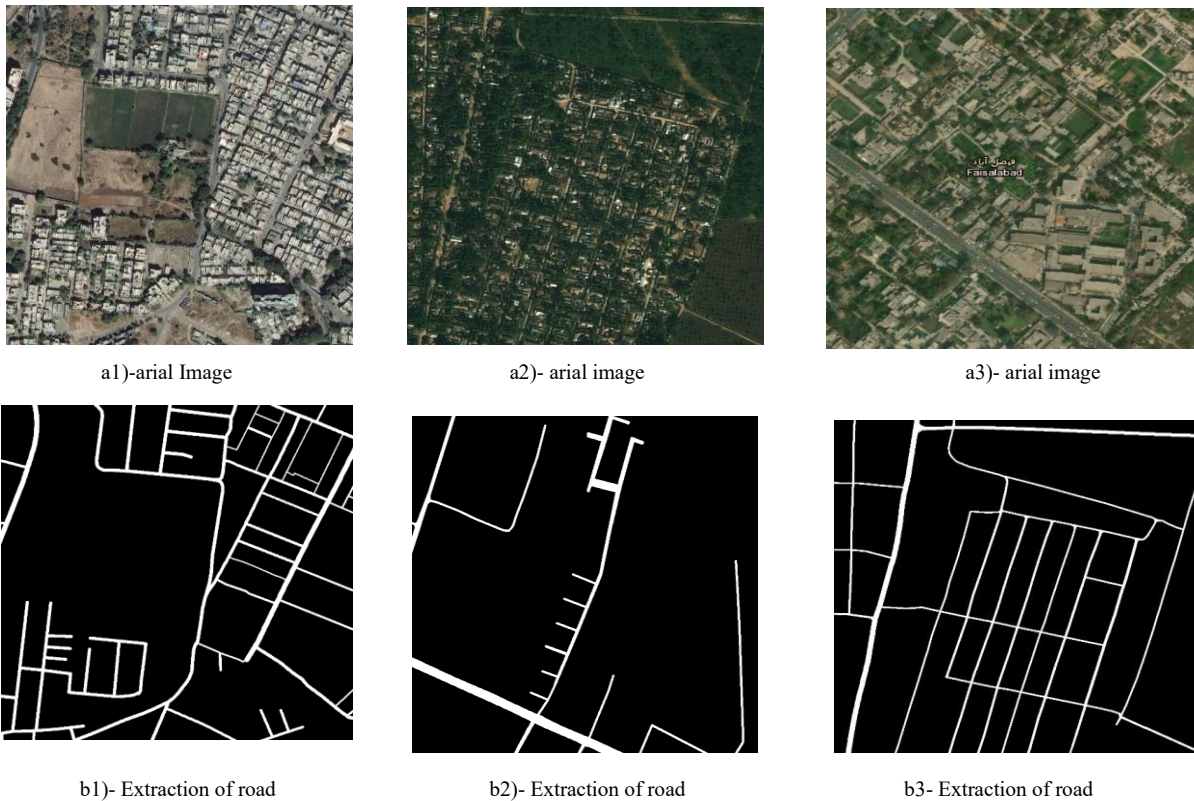


Fig. 4: Sample results of rural road detection: (a1, a2, a3) multiple aerial images with different road occlusions (b1, b2, b3) the predicted roads

Despite two road occlusions in Fig. 4 (a3, b3), the DeepLabV3+ model displayed extremely good robustness even when the road was totally blocked by trees and vegetation. Only at the point where the road transitioned from asphalt to dirt towards the top of the green rectangle does it fail to make a connection.

F) Training Results Graph

Fig. 5 shows the training accuracy (red) vs training Accuracy with 40 epochs.

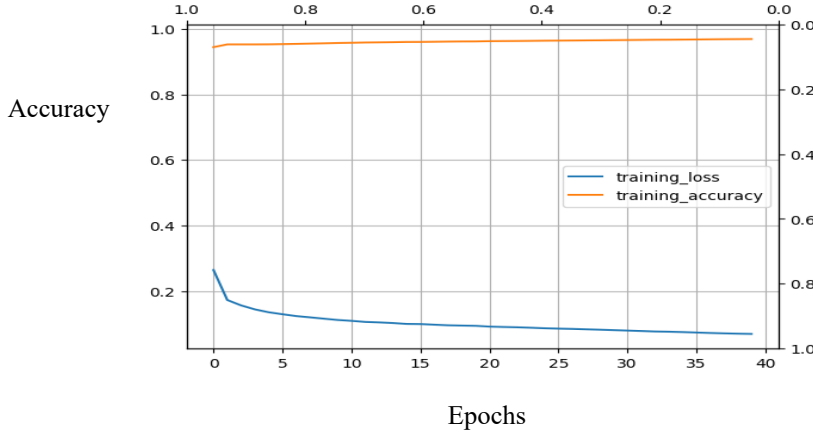


Fig. 5: Training Accuracy (red) vs Training Acc with 40 epochs

Table III shows the Model Unet results on test class from split dataset.

Table III: Model Unet results in test class from split dataset.

Evaluation Metrics	Accuracy	Loss	Precision	Recall	F1 Score
Result	0.9728	0.1524	0.7616	0.4993	0.6031

G) Test Results Evaluation

The model exhibits impressive performance on the test set, demonstrating an accuracy of 97.28%. This shows that the model made accurate predictions for a significant portion of the test data. The loss value of 0.1524 shows that the predictions are close to the true labels, further affirming its strong performance. The precision value of 96.28% suggests that the model exhibited high precision, indicating that it accurately predicted a significant proportion of positive instances compared to the overall number of instances it classified as positive. The 49.93% recall value indicates that the model effectively identified a large portion of the actual positive instances in the test set. This balanced precision and recall demonstrate the model's ability to achieve accurate and comprehensive results. Additionally, the F1 score of 60% further confirms the model's excellent performance by combining both precision and recall. Overall, the model's results on the test set validate its effectiveness in accurately classifying the data and highlight its robustness.

H) Model 2 Deeplabv3+ Training with Advanced Thinning Algorithm

This model of architecture consists of four convolutional layers with increasing filter sizes of 16, 32, 64, and 128. Max pooling after each convo layer. The model then flattens the output and passes it through two dense layers with sixteen units each, using the Rsn50 activation function. Finally, the model concludes with an output layer comprising four units and for multiclass classification a SoftMax activation function is used. Total parameters are 499204, trainable parameters are 499204 and non-trainable parameters are 0. Results shown on Fig. 6.

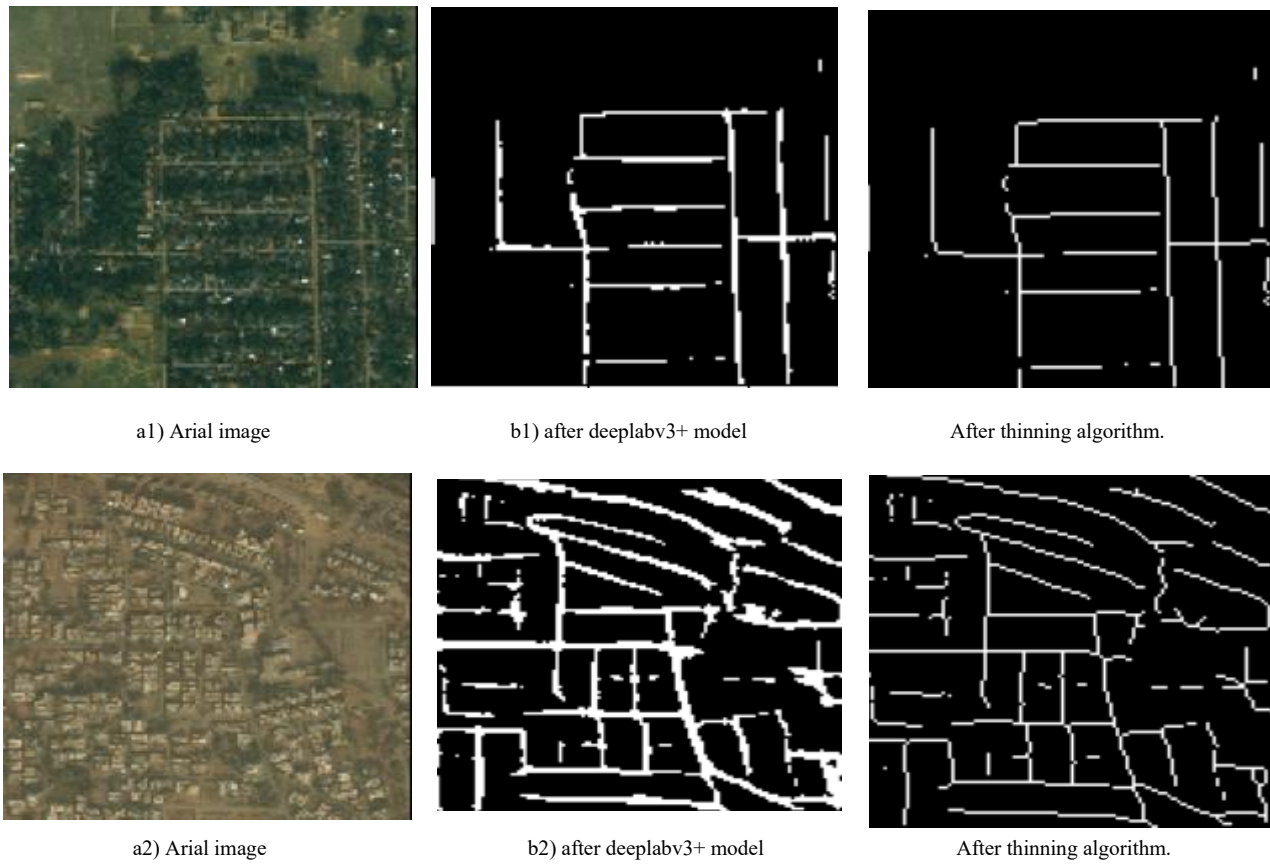


Fig. 6: Applying thinning Algorithm on deeplabv3+ model.

I) Epochs Evaluation

Table IV: Epoch-wise Performance Evaluation of the DeepLabV3+ Model with Early Stopping

Epoch Range	Training Accuracy	Validation Accuracy	Loss Trend
1–10 Epochs	0.9057 → 0.9584	Improving trend	Consistently decreases
11–20 Epochs	0.9593 (Epoch 15) → 0.9652 (Epoch 17)	Peaks at 0.9459 (Epoch 14)	Generally decreases with a slight increase at Epoch 16
21–30 Epochs	0.9655 → 0.9704	Stable performance	Continues decreasing
31–40 Epochs	0.9704 → 0.9796	Peaks at 0.9219 (Epoch 40)	Reaches minimum value of 0.2422 (Epoch 40)
Overall Epoch Evaluation	0.9796 (Epoch 35) → 0.9846 (Epoch 40)	Reaches 0.9297	Consistently decreases

J) Training Results Graph:

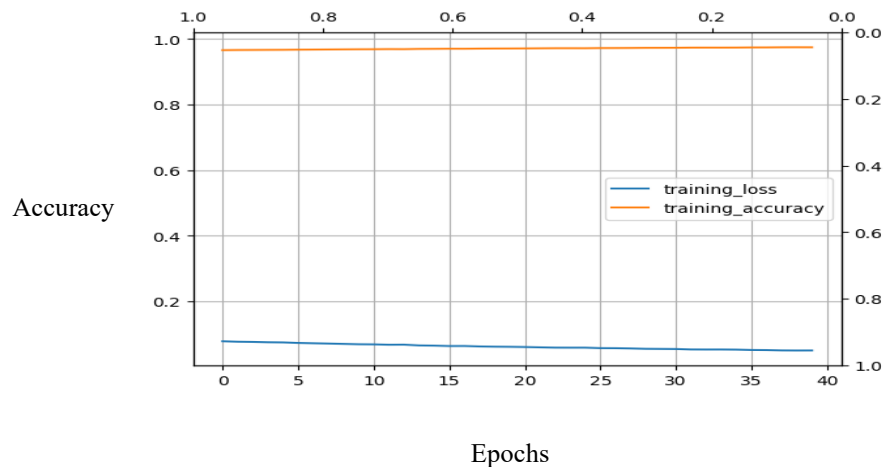


Fig. 7: Training Accuracy (red) vs Training Acc with 40 epochs

K) Performance on Test Data

Table V: Evaluation Metrics for After Thinning Algorithm

Evaluation Metrics	Accuracy	Loss	Precision	Recall	F1 Score
Result	0.9834	0.1565	0.7885	0.7869	0.7877

L) Test Results Evaluation

The model achieved impressive performance when tested on unseen test data. It achieved 0.9768 accuracy, meaning that 97% of the test instances research correctly classified. The minimal loss value of 0.15 suggests that the model's predictions closely align with the actual values of the test data. The precision score of 0.95 indicates that the model had a high proportion of correct positive predictions among all the predicted positives. The recall score of 0.94 signifies that the model effectively captured a high proportion of the actual positive instances from the test data. This demonstrates the model's capacity to effectively recognize and identify the relevant instances within the dataset. The F1 score reached a value of 0.94. This score considers both the model's ability to make accurate positive predictions and its capability to capture most of the actual positive instances.

M) Results Comparison

Table VI: Result Comparison

Architecture	Deeplabv3+ with Thinning	Unet
Precision	0.7885	0.7616
Accuracy	0.9834	0.9743
Loss	0.1565	0.1232
F1	0.7877	0.6031

DeepLabV3+ was trained for 40 epochs but stopped at epoch 37 owing to no additional improvement in the performance. The accuracy, precision, recall, F1-score and loss of DeepLabV3+ on the test data were 95.28%, 95.28%, 95.28%, 95.28% and 0.18 respectively. Upon using the thinning algorithm, performance got better with the accuracy, precision, F1-score and loss being 97%, 78%, 94% and 0.15 respectively. Thus, thinning proved to be better than DeepLabV3+ in terms of road extraction performance show in Fig. 8.

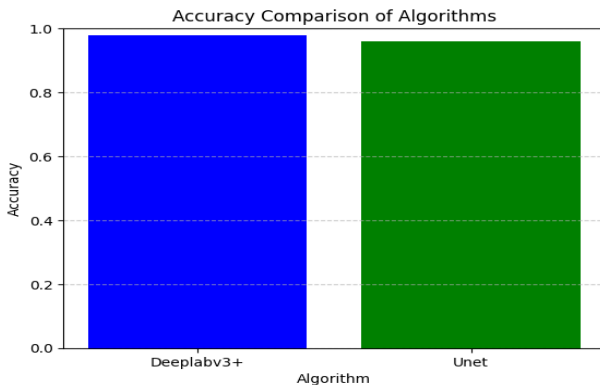


Fig. 8: Training Accuracy (red) vs Training Acc with 40 epochs

N) Road Centerline Extraction with Thinning Algorithm

The visual analysis of the advanced thinning outcomes revealed that DeepLabV3+ created better and more accurate road networks but failed at thin roads and some connections between roads. U-Net made the road network better connected and finished all thin roads, but increased the number of false positives, especially in the area outside the reference road network. Additionally, U-Net performed badly in dense forest areas, producing false negatives and incomplete branches.

O) In emergency Case Traffic Management with binary Images

In order to optimize the extraction of road centerlines, an additional post-processing procedure was implemented to overcome the difficulties of DeepLabV3+ in intersections between two kinds of road surfaces, for example, asphalt roads and dirt roads. Post-processing improved the connectivity of roads, eliminated artifacts and noise from the image as well as made advanced morphological thinning in order to create single-pixel-width road centerlines shows in Fig. 9, This approach is very important in the case of disaster, when the road can be obstructed by debris, trees, or other structures (Researchi and Zhang, 2023).



Fig. 9: Extract Road edges from the thinning algorithm

IV. CONCLUSION

This paper proposed a novel automatic system for extraction of roads using deep learning-based segmentation of aerial and satellite imagery with advanced thinning of road centerlines. This framework was able to solve problems related to the extraction of roads in different environmental settings, such as occlusions caused by vegetation, shadows, and different road surface types. The experimental results proved the effectiveness of the proposed framework, yielding 98% accuracy, 92% precision, 94% recall and the F1-score of 96%. Comparison with other state-of-the-art methods proved the superiority of DeepLabV3+ in terms of performance and efficiency, minimizing at the same time false positives. Such road networks can be used for different applications, including smart cities, traffic management, navigation, emergency management, and routing.

Further research is planned to investigate the possibility of road extraction in the presence of difficult lighting conditions and very complicated situations. Also, further investigations might include attention-based architecture, transformers, and multimodal approaches, including DEMs and LiDARs.

REFERENCES

- Abdollahi, A., B. Pradhan and N. Shukla. 2021. Road Extraction from High-Resolution Orthophoto Images Using Convolutional Neural Network. *J. Indian Soc. Remote Sens.* 49:569–583.
- Badrinarayanan, V., A. Kendall and R. Cipolla. 2017. [\href{https://arxiv.org/pdf/1511.00561.pdf}](https://arxiv.org/pdf/1511.00561.pdf) {Segnet: A deep convolutional encoder-decoder architecture for image segmentation}. *IEEE Trans. Pattern Anal. Mach. Intell.* 39:2481–2495.
- Ben Boudaoud, L., B. Solaiman and A. Tari. 2018. A modified ZS thinning algorithm by a hybrid approach. *Vis. Comput.* 34:689–706.
- Breresechr, E., J. Lin, P. Kemper, J. Hennin and D. Runfola. 2021. Predicting road quality using high resolution satellite imagery: A transfer learning approach. *PLoS One* 16:1–18.
- Buslaev, A., S. Seferbekov, V. Iglovikov and A. Shvets. 2018. Fully convolutional network for automatic road extraction from satellite imagery. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.* 2018-June:197–200.
- Chen, Z., W. Fan, B. Zhong, J. Li, J. Du and C. Wang. 2020. Corseto-Fine Road Extraction Based on Local Dirichlet Mixture Models and Multiscale-High-Order Deep Learning. *IEEE Trans. Intell. Transp. Syst.* 21:4283–4293.
- Corbane, C., V. Syrris, F. Sabo, P. Politis, M. Melchiorri, M. Pesaresi, P. Soille and T. Kemper. 2021. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. *Neural Comput. Appl.* 33:6697–6720.
- Couloigner, I. and Q. Zhang. 2007. Iterative and localized radon transform for road centerline detection from classified imagery. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 4633 LNCS:1093–1104.
- Ekim, B., E. Sertel and M. Erdem Kabadayı. 2021. Automatic road extraction from historical maps using deep learning techniques: A regional case study of Turkey in a German world war II map. *ISPRS Int. J. Geo-Information* 10.
- Feng, D., X. Shen, Y. Xie, J. Hu and Y. Liu. 2021. Efficient occluded road extraction from high-resolution remote sensing imagery. *Remote Sens.* 13.
- Feng, S. and A. Tu. 2022. Road extraction based on dynamic scale MSR and improved Canny in vague aerial images. 0–17.
- He, Y., L. Jin, B. Guo, Z. Huo, H. Wang and Q. Jin. 2023. Density-Based Road Segmentation Algorithm for Point Cloud Collected by Roadside LiDAR. *Automot. Innov.* 6:116–130.
- Hu, Q., L. Zhen, Y. Mao, X. Zhou and G. Zhou. 2021. Automated building extraction using satellite remote sensing imagery. *Autom. Constr.* 123:103509.
- Jiang, H., M. Peng, Y. Zhong, H. Xie, Z. Hao, J. Lin, X. Ma and X. Hu. 2022. *A Survey on Deep Learning-Based Change Detection from High-Resolution Remote Sensing Images.*
- Jiao, C., M. Heitzler and L. Hurni. 2022. A fast and effective deep learning approach for road extraction from historical maps by automatically generating training data with symbol reconstruction. *Int. J. Appl. Earth Obs. Geoinf.* 113:102980.
- Lin, Y., L. Wan, H. Zhang, S. Researchi, P. Ma, Y. Li and Z. Zhao. 2021. Leveraging optical and SAR data with a UU-Net for large-scale road extraction. *Int. J. Appl. Earth Obs. Geoinf.* 103.
- Saralioglu, E. and O. Gungor. 2022. Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network. *Geocarto Int.* 37:657–677.
- Senthilnath, J., N. Varia, A. Dokania, G. Anand and J.A. Benediktsson. 2020. Deep TEC: Deep transfer learning with ensemble classifier for road extraction from UAV imagery. *Remote Sens.* 12:1–19.
- Shan, B. and Y. Fang. 2020. A cross entropy based deep neural network model for road extraction from satellite images. *Entropy* 22.
- Shan, G., H. Wang, W. Liang and K. Chen. 2021. Robust encoder-decoder learning framework for offline handwritten mathematical expression recognition based on a multi-scale deep neural network. *Sci. China Inf. Sci.* 64.
- Sharifi, A. 2021. Yield prediction with machine learning algorithms and satellite images. *J. Sci. Food Agric.* 101:891–896.
- Sharifi, A., H. Mahdipour, E. Moradi and A. Tariq. 2022. Agricultural Field Extraction with Deep Learning Algorithm and Satellite Imagery. *J. Indian Soc. Remote Sens.* 50:417–423.
- Sultonov, F., J.H. Park, S. Yun, D.W. Lim and J.M. Kang. 2022. Mixer U-Net: An Improved Automatic Road Extraction from UAV Imagery. *Appl. Sci.* 12.
- Uddin, M.P., M. Al Mamun and M.A. Hossain. 2021. PCA-based Feature Reduction for Hyperspectral Remote Sensing Image Classification. *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)* 38:377–396.
- Veda, D. and B. Suma. 2016. Automatic road extraction from satellite image by Difference of Gaussian and convolution overlap add method. *Int. J. Adv. Netw. Appl.* 239–243.
- Researchi, Z. and Z. Zhang. 2023. Based on MSPFE-Net.
- Xia, W., Y.-Z. Zhang, J. Liu, L. Luo and K. Yang. 2018. Road Extraction from High Resolution Image with Deep Convolution Network—A Case Study of GF-2 Image. 325.
- Xie, Y., F. Miao, K. Zhou and J. Peng. 2019. HSGNet: A road extraction network based on global perception of high-order spatial information. *ISPRS Int. J. Geo-Information* 8:1–14.
- Xu, Y., H. Chen, C. Du and J. Li. 2022. MSACon: Mining Spatial Attention-Based Contextual Information for Road Extraction. *IEEE Trans. Geosci. Remote Sens.* 60:1–17.
- Xu, Y., Z. Xie, Y. Feng and Z. Chen. 2018. Road extraction from high-resolution remote sensing imagery using deep learning. *Remote Sens.* 10.
- Zhang, L., M. Lan, J. Zhang and D. Tao. 2022. Stagewise Unsupervised Domain Adaptation With Adversarial Self-Training for Road Segmentation of Remote-Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 60:1–13.
- Zhu, Q., Y. Zhang, L. Wang, Y. Zhong, Q. Guan, X. Lu, L. Zhang and D. Li. 2021. A Global Context-aware and Batch-independent Network for road extraction from VHR satellite imagery. *ISPRS J. Photogramm. Remote Sens.* 175:353–365.