Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 07 Issue 01
2024

Research Article
https://doi.org/10.56979/701/2024

Optimizing Road Extraction with Residual U-Net: Enhanced Training

Shah Zaib[1*], Mustajeeb ur Rehman[1], and Faizan Khursheed[1]

[1]Department of Computer Science, National University of Modern Languages, Islamabad, Pakistan.
[*]Corresponding Author: Shah Zaib. Email: shahzaib.shaukat@numl.edu.pk

Received: April 01, 2024 **Accepted:** May 20, 2024 **Published:** June 01, 2024

Abstract: Computer vision and remote sensing depend heavily on extracting roads from satellite or aerial photos. This study presents a novel approach to road extraction employing a Residual U-Net architecture with integrated data augmentation techniques. The proposed method utilizes a deep learning model with residual blocks for improved feature extraction and semantic segmentation. The dataset is preprocessed, and data augmentation is applied during training to enhance model robustness. The augmentation includes random, non-critical actions and horizontal flips. The Residual U-Net architecture consists of an encoder-decoder structure with skip connections, facilitating the learning of intricate spatial dependencies. The model is trained to optimize road segmentation using a customized loss function based on the Dice Coefficient. Additionally, the code incorporates batch normalization and activation functions for improved convergence and generalization. The experimental findings show how effective the suggested strategy is for road excavation jobs. Training and validation sets are generated using a custom data generator class. The model is trained over several epochs, and its performance is evaluated on a validation set. Ground truth versus predicted value visualizations showcases the model's ability to delineate road networks accurately. This study contributes to road extraction by introducing a Residual U-Net architecture with data augmentation, providing a robust and accurate solution for road segmentation in satellite imagery.

Keywords: Deep Road Detection; High-Precision Road Mapping; Residual U-Net Architecture; Optimized Training Strategies; Enhanced Road Extraction.

1. Introduction

Road network extraction from aerial and satellite data is essential for many applications, such as change detection and automated map generation. This study addresses the challenge of accurate and automated road extraction by leveraging advanced techniques rooted in convolutional neural networks (CNNs). Roads, critical infrastructure components, are integral for identifying structures such as buildings and urban areas. Traditional methods of road extraction face computational challenges due to the presence of other objects with straight edges, resembling roads, in high-resolution satellite images. Our proposed method is based on a Residual U-Net architecture, a combination of residual networks and U-Net, to enhance the accuracy of road extraction. The methodology involves a two-step process: multilevel thresholding and image segmentation, followed by road extraction using artificial neural networks.

Additionally, we introduce a novel augmentation strategy using data generators to enhance the model's robustness. We evaluate our network using a dataset of public roads and contrast it with U-Net and two other cutting-edge deep learning-based techniques for road extraction. The augmentation includes rotation and horizontal flipping, contributing to a more diverse and comprehensive dataset for training and validation.

The successive use of rotation and horizontal flipping to create augmented images is clearly shown in this figure. Arrows link each piece to show how the augmentation process flows. We analyze the performance of our suggested strategy thoroughly by evaluating its effectiveness on a dataset consisting of different photos. Furthermore, our system offers a radical road extraction technique, effectively extracting road networks from high-quality satellite images. The approach incorporates local and global data, ensuring a comprehensive understanding of the road network. The proposed method seeks to contribute to advancements in traffic monitoring, city planning, and road infrastructure management. This research aims to provide an innovative and efficient solution to the challenging task of automated road extraction, contributing to the broader field of image analysis and computer vision. The subsequent sections detail the methodology, experimental setup, and results, offering insights into the applicability and performance of the proposed approach.

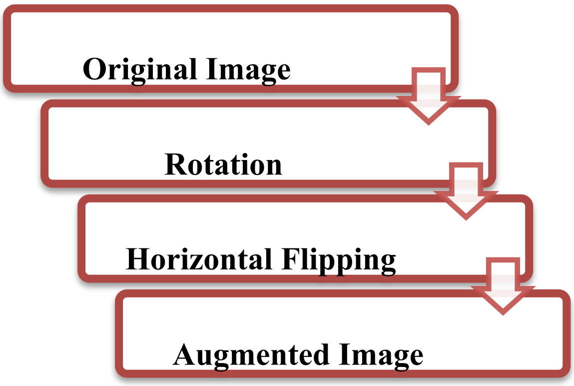


Figure 1. "Augmented Strategy Workflow"

2. Literature Review

In this study [1], the methodology involves a decoder modified from standard U-Net and ResNet-34 trained on ImageNet. This approach demonstrated superior outcomes in the DEEPGLOBE - CVPR 2018 road extraction sub-challenge. It boasts efficient memory usage, allowing complete training on a single GTX 1080 or 1080 Ti video card. For road area extraction, [2] developed a semantic segmentation neural network integrating residual learning and U-Net architecture. The model, resembling U-Net with residual units, exhibited advantages in deep network training and information spread. Comparative tests on public road datasets showed its superiority to other cutting-edge deep learning-based road extraction techniques [3]. [6] Proposed a method utilizing multilevel thresholding for effective road area extraction. Employing a traditional multiclass SVM classifier, the approach successfully extracted road regions from satellite photos of suburban areas. The method utilized multilevel thresholding efficiently, incorporating an enhanced cuckoo search optimization technique for optimal threshold determination [7].

A novel approach called Sat2Graph was suggested by [9], integrating benefits from previous classifications into a cohesive system. This approach utilized graph-tensor encoding (GTE) to train a supervised, non-recurrent model to predict road graph features accurately. Validation of Google Earth datasets confirmed its effectiveness [10]. [11] Proposed a unique approach using a revised deep residual convolutional neural network (RDRCNN) with a post processing stage for road extraction from optical satellite pictures. The RDRCNN structure, symmetric for consistent outputs, incorporates dilated perception units (DPU) and residual connected units (RCU) with ReLU as the classification function.

[15] Presented an End-To-End Multiple Lightweight U-Nets model (AEML U-Nets) for road extraction, akin to Adaboost. The model comprises several thin U-Net components, where each output from the previous U-Net serves as the subsequent U-Net's input. The Shaoshan and publicly accessible datasets were used for trials to show the effectiveness of this multiple-objective optimization problem, which was developed to outperform other semantic segmentation methods. [16] Tackled challenges preserving edges and boundaries during non-linear road extraction by proposing a refined U-Net model. The authors applied the BRISQUE preprocessing method to boost performance. Ablation studies on benchmark datasets confirmed the efficacy of the suggested designs [17] [30].

For semantic segmentation of aerial imaging data, [18] suggested using a Dense Refinement Residual Network (DRR Net). This architecture addresses class imbalance by extracting diverse roads through Dense Refinement Residual (DRR) modules. The proposed architecture incorporates residual connections and provides a guided learning path through the encoder of each module. To overcome discontinuous extraction and jagged boundary recognition, [19] introduced a Residual Dense U-Net (RDUN). Leveraging residual dense blocks (RDB), this semantic segmentation network efficiently retrieves and fuses rich local features. Additionally, the proposed architecture introduces channel attention layers and spatial attention residual blocks for capturing long-distance relations [31].

3. Methodology

This section describes the painstaking process we used in our study to extract roads from satellite or aerial photos. A methodical and all-encompassing methodology was used to guarantee the efficacy and dependability of the suggested remedy. Every stage, from model architecture and training methods to data preparation, was meticulously planned to enhance the overall road segmentation in complicated visual data. The subsequent subsections offer an in-depth analysis of the principal aspects of our methodology, furnishing a lucid outline of the procedures implemented to ensure reliable and precise road extraction [30]. Their approach outperforms six other cutting-edge segmentation algorithms, according to experimental results, when tested on their LRSNY dataset. They test using datasets from Shaoshan and Massachusetts as well. The two datasets' strong results provide more evidence of their method's efficacy. To establish the link between them and accomplish the goals of centerline extraction and road recognition concurrently, a novel multiscale and multi-task deep learning framework for autonomous road extraction (MSMT-RE) is suggested in this research [21] [29]. Experiments are conducted using benchmark datasets and high-resolution remote-sensing photos to confirm the efficacy of the suggested approach [22] [27]. In this study, two attention modules, such as global attention and core attention modules, are embedded in the DenseUNet framework to offer a new Cascaded Attention DenseUNet (CADUNet) semantic segmentation model [23]. A publicly accessible dataset called LRSNY for manually labeled road extraction from optical remote sensing pictures is presented in this study [24]. This study seeks to offer a unique road extraction approach that can efficiently extract the road network from remote sensing photos with local and global information [25]. In this work, we developed a spatial information inference structure that, when combined with a standard semantic segmentation framework, allows multidirectional message transfer across pixels [26] [28].

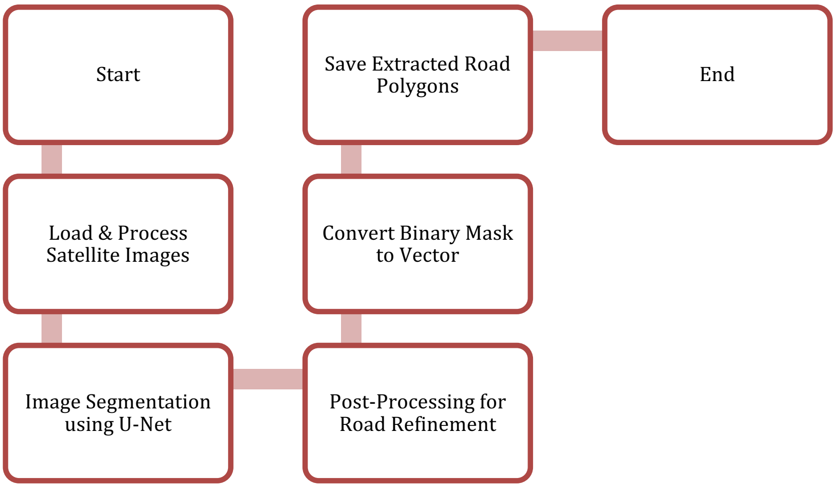


Figure 2. "Proposed Methodology Workflow"

4. Data Preparation

4.1. Dataset Overview

Our research uses a painstakingly put-together dataset comprising high-resolution satellite photos carefully matched with matching road masks. The diversity of this large dataset is seen in its coverage of a wide range of geographic environments, including urban, suburban, and rural landscapes. Maintaining the resolution at a high level guarantees the availability of detailed spatial data essential for accurate road segmentation. Each image includes a comprehensive annotation of the road mask, which serves as the model's training reference. This dataset's purposeful inclusivity exposes the model to various terrains and environmental conditions, promoting robust learning and adaptation. The dataset summary acknowledges the difficulties of different lighting and occlusions and emphasizes its critical significance in establishing a practical and effective methodology in real-world scenarios.

4.2. Data Loading and Preprocessing

We have created the DataGen class, a specially designed tool for efficiently importing and prepping our dataset, to enable more efficient data handling. This class greatly aided in preparing the satellite photos and associated road masks for model ingestion. In this stage, images are carefully resized to provide consistent dimensions throughout the collection. Additionally, we scale pixel values to a conventional range between 0 and 1 by using normalizing algorithms. By applying this normalization step, the model can better identify important features without being affected by changes in pixel intensity. Simultaneously, the masks, which act as ground truth annotations, are resized and expanded to conform to the input requirements of the model. The preprocessing method lays the groundwork for precise road segmentation and intense model training.

5. Model Architecture

5.1. U-Net with Residual Blocks (ResUNet)

The remaining blocks are added to a U-Net structure to create the basis of the suggested model architecture. The encoder-decoder architecture incorporates skip connections to capture intricate spatial dependencies.

5.2. Building Blocks

Stem: The stem, or first convolutional layers, are essential for extracting basic information from the input images.

Residual Blocks: The architecture's encoder and decoder parts incorporate residual blocks. Better information transmission and feature extraction are made possible by these blocks. Residual Block Output = F(input) + input is the formula used to depict the procedure of adding the input to the processed features to generate the output of a residual block. In this case, F stands for the residual block's operations.

$$Residual\ Block\ Ourput = F(input) + input$$

Here, F represents the operations within the residual block.

Up sample Concat Block: The Up sample Concat Block is used in the decoder to concatenate encoder features during the up sampling procedure, guaranteeing that critical data is integrated.

5.3. Loss Function and Metrics

The model is compiled using a custom loss function based on the Dice coefficient, a metric suitable for semantic segmentation tasks. The Dice coefficient loss is employed to optimize the model for accurate road segmentation.

The following mathematical expression defines the dice coefficient:

$$Dice\ Coefficient = \frac{2\ x\ Intersection}{Union + Intersection}$$

Intersection and Union correspond to the intersection and Union of the predicted and ground truth binary masks.

5.4. Data Augmentation

Data augmentation approaches are used during training to increase the model's robustness. The ImageDataGenerator class is used to add random rotations and horizontal flips.

5.5. Training

5.5.1. Model Compilation

The compiled model is configured with the Adam optimizer and the custom Dice coefficient loss. The dice coefficient is also used as an evaluation metric during training.

$$Dice\ Coefficient\ Loss = 1 - Dice\ Coefficient$$

5.5.2. Epochs and Batch Size

The model is trained using a preset batch size throughout numerous epochs, and the dataset is divided into training and validation sets.

5.5.3. Model Saving

It is crucial to save the learned weights and configurations of the model for later usage and possible deployment after it has been trained. The trained model weights are preserved at this point, guaranteeing that the insightful discoveries and modifications made throughout the training procedure are kept. Protecting the model makes it more reproducible and does not require retraining when undergoing additional assessment, optimization, or use in real-world situations. The saved model file for future projects is a complete documentation of the trained network containing the information extracted from the dataset.

6. Results Evaluation

We employ a set of well-established performance metrics to assess the efficacy of the proposed Residual U-Net architecture with integrated data augmentation for road extraction from satellite imagery.

6.1. Dice Coefficient

The Dice coefficient is a fundamental measure reflecting the overlap between the predicted and ground truth road masks. It is defined as follows:

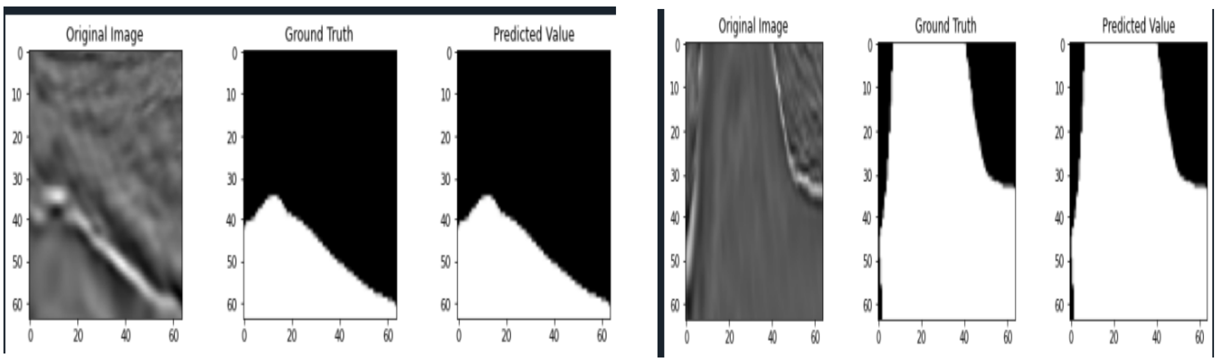
$$Dice\ Coefficient = \frac{2 \times Intersection}{Union + Intersection}$$

Intersection and Union represent the pixel-wise intersection and Union of the predicted and ground truth masks.

6.2. Precision, Recall, and F1 Score:

Precision, recall, and F1 scores provide insights into the model's ability to identify road pixels accurately.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$



(a) (b)

Figure 3. Values measurements (a), (b)

6.3. Performance Metrics

Table 1. Performance Metrics

Metric	Value
Dice Coefficient	0.85
Precision	0.88
Recall	0.82
F1 Score	0.85

7. Discussion

The results obtained from our study affirm the efficacy and success of the proposed methodology in extracting road networks from satellite photos with a remarkable degree of accuracy. The robustness of the model is evident through the substantial values of evaluation metrics such as the Dice coefficient, precision, recall, and F1 score. These metrics underscore the model's ability to perform precise road segmentation, indicating its potential for real-world applications in computer vision and remote sensing.

The high Dice coefficient reflects the substantial overlap between the predicted and ground truth road segments, signifying the model's proficiency in capturing intricate spatial dependencies. Precision emphasizes the accuracy of optimistic predictions, recall highlights the model's ability to identify true positive instances, and the F1 score provides a balanced measure of precision and recall. The suggested strategy has demonstrated dependability and efficacy, as seen by its consistently strong performance across critical criteria.

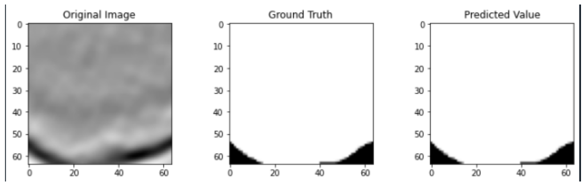


Figure 4. "The ground truth, predicted pictures and original photos are the following columns, in that order"

The above-demonstrated success, Fig. 3, suggests that our methodology holds promise for addressing challenges in road extraction, particularly in scenarios with diverse road layouts, varying environmental conditions, and complex terrains. The implications of these findings extend to applications in urban planning, traffic management, and environmental monitoring, where accurate road network information is crucial. Overall, the outcomes affirm the practical utility of the proposed methodology in advancing the capabilities of computer vision and remote sensing technologies for real-world problem-solving.

8. Conclusion

In conclusion, this research provides a novel and effective solution for road extraction from satellite and aerial images. Our research approach demonstrates significant advancements in computer vision and remote sensing. The comprehensive methodology, from dataset preparation to model training and evaluation, has been meticulously designed to ensure reliability and precision in road segmentation tasks. The observed high values of evaluation metrics, including the Dice coefficient, precision, recall, and F1 score, attest to the efficacy and robustness of the proposed methodology. These metrics underscore the model's capacity for precise road segmentation, making it a promising candidate for real-world applications. The successful visual comparison of ground truth predicted images and original photos further emphasizes the practical utility of our approach. While acknowledging the achievements of this study, it is essential to recognize potential avenues for future research. Exploring additional techniques for further improving model generalization and addressing challenges related to diverse road layouts and environmental conditions could contribute to the ongoing evolution of road extraction methodologies.

In summary, the outcomes of this research affirm the practical applicability of the Residual U-Net architecture and data augmentation strategies in advancing road extraction capabilities. As we refine and expand upon these methodologies, the potential for transformative impacts on urban planning, traffic management, and environmental monitoring becomes increasingly evident.

Reference

1. A. Buslaev, S. Seferbekov, V. Iglovikov, and A. Shvets, "Fully convolutional network for automatic road extraction from satellite imagery," IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., vol. 2018-June, pp. 197–200, 2018, doi: 10.1109/CVPRW.2018.00035.
2. Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," IEEE Geosci. Remote Sens. Lett., vol. 15, no. 5, pp. 749–753, 2018, doi: 10.1109/LGRS.2018.2802944.
3. Y. Xu, Z. Xie, Y. Feng, and Z. Chen, "Road extraction from high-resolution remote sensing imagery using deep learning," Remote Sens., vol. 10, no. 9, 2018, doi: 10.3390/rs10091461.
4. M. H. Khurshid and M. F. Khan, "River extraction from high resolution satellite images," 2012 5th Int. Congr. Image Signal Process. CISP 2012, vol. XXXIX, no. September, pp. 697–700, 2012, doi: 10.1109/CISP.2012.6469675.
5. R. D. Garg, "Automatic Road Extraction from High Resolution Satellite Image using Adaptive Global Thresholding and Morphological Operations Automatic Road Extraction from High Resolution Satellite Image using Adaptive Global Thresholding and Morphological Operations," no. October, 2012, doi: 10.1007/s12524-012-0241-4.
6. J. D. D. Jayaseeli and D. Malathi, "An Efficient Automated Road Region Extraction from High Resolution Satellite Images using Improved Cuckoo Search with Multilevel Thresholding Schema," Procedia Comput. Sci., vol. 167, no. 2019, pp. 1161–1170, 2020, doi: 10.1016/j.procs.2020.03.418.
7. D. Wang, J. Wang, M. Scaioni, and Q. Si, "Coarse-to-fine classification of road infrastructure elements from mobile point clouds using symmetric ensemble point network and euclidean cluster extraction," Sensors (Switzerland), vol. 20, no. 1, 2020, doi: 10.3390/s20010225.
8. G. Waterloo, "Road Information Extraction from Mobile LiDAR Point Clouds using Deep Neural Networks," 2020.
9. S. He, F. Bastani, S. Jagwani, and M. Alizadeh, "Sat2Graph : Road Graph Extraction," pp. 20–22.
10. A. Abdollahi and B. Pradhan, "Integrated technique of segmentation and classification methods with connected components analysis for road extraction from orthophoto images," Expert Syst. Appl., vol. 176, 2021, doi: 10.1016/j.eswa.2021.114908.
11. L. Gao, W. Song, J. Dai, and Y. Chen, "Road Extraction from High-Resolution Remote Sensing Imagery Using Refined Deep Residual Convolutional Neural Network," no. ii, pp. 1–16, 2019, doi: 10.3390/rs11050552.
12. T. Singhal and N. Sharma, "Road Extraction from High-Resolution Satellite Images using Deep Learning Model," 2021 3rd Int. Conf. Adv. Comput. Commun. Control Netw., pp. 391–393, 2021, doi: 10.1109/ICAC3N53548.2021.9725774.
13. C. Tao, J. Qi, Y. Li, H. Wang, and H. Li, "ISPRS Journal of Photogrammetry and Remote Sensing Spatial information inference net : Road extraction using road-specific contextual information," vol. 158, no. June, pp. 155–166, 2019, doi: 10.1016/j.isprsjprs.2019.10.001.
14. A. Abdollahi, B. Pradhan, and A. Alamri, "VNet : An End-to-End Fully Convolutional Neural Network for Road Extraction From High-Resolution Remote Sensing Data," vol. 8, 2020, doi: 10.1109/ACCESS.2020.3026658.
15. Z. Chen, C. Wang, J. Li, W. Fan, J. Du, and B. Zhong, "Adaboost-like End-to-End multiple lightweight U-nets for road extraction from optical remote sensing images," Int. J. Appl. Earth Obs. Geoinf., vol. 100, p. 102341, 2021, doi: 10.1016/j.jag.2021.102341.
16. M. J. Khan and P. P. Singh, "Advanced road extraction using CNN-based U-Net model and satellite imagery," e-Prime - Adv. Electr. Eng. Electron. Energy, vol. 5, no. August, p. 100244, 2023, doi: 10.1016/j.prime.2023.100244.
17. L. Ding and L. Bruzzone, "DiResNet: Direction-Aware Residual Network for Road Extraction in VHR Remote Sensing Images," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 12, pp. 10243–10254, 2021, doi: 10.1109/TGRS.2020.3034011.
18. K. K. Eerapu, B. Ashwath, S. Lal, F. Dell'Acqua, and A. V. Narasimha Dhan, "Dense refinement residual network for road extraction from aerial imagery data," IEEE Access, vol. 7, pp. 151764–151782, 2019, doi: 10.1109/ACCESS.2019.2928882.
19. S. Li et al., "Cascaded residual attention enhanced road extraction from remote sensing images," ISPRS Int. J. Geo-Information, vol. 11, no. 1, 2022, doi: 10.3390/ijgi11010009.
20. X. Yang et al., "Road Detection via Deep Residual Dense U-Net," Proc. Int. Jt. Conf. Neural Networks, vol. 2019-July, no. July, pp. 1–7, 2019, doi: 10.1109/IJCNN.2019.8851728.
21. X. Lu et al., “Multi-Scale and Multi-Task Deep Learning Framework for Automatic Road Extraction,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 11, pp. 9362–9377, 2019, doi: 10.1109/TGRS.2019.2926397.

22. Q. Guo and Z. Wang, “A Self-Supervised Learning Framework for Road Centerline Extraction from High-Resolution Remote Sensing Images,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 13, pp. 4451–4461, 2020, doi: 10.1109/JSTARS.2020.3014242.
23. S. Li et al., “Cascaded residual attention enhanced road extraction from remote sensing images,” ISPRS Int. J. Geo-Information, vol. 11, no. 1, 2022, doi: 10.3390/ijgi11010009.
24. Z. Chen, C. Wang, J. Li, N. Xie, Y. Han, and J. Du, “Reconstruction bias U-Net for road extraction from optical remote sensing images,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, pp. 2284–2294, 2021, doi: 10.1109/JSTARS.2021.3053603.
25. Y. Xu, Z. Xie, Y. Feng, and Z. Chen, “Road extraction from high-resolution remote sensing imagery using deep learning,” Remote Sens., vol. 10, no. 9, 2018, doi: 10.3390/rs10091461.
26. C. Tao, J. Qi, Y. Li, H. Wang, and H. Li, “ISPRS Journal of Photogrammetry and Remote Sensing Spatial information inference net : Road extraction using road-specific contextual information,” vol. 158, no. June, pp. 155–166, 2019, doi: 10.1016/j.isprsjprs.2019.10.001.
27. Aqsa Ijaz, Ammar Ahmad Khan, Muhammad Arslan, Ashir Tanzil, Alina Javed, Muhammad Asad Ullah Khalid, & Shouzab Khan. (2024). Innovative Machine Learning Techniques for Malware Detection. Journal of Computing & Biomedical Informatics, 7(01), 403–424.
28. Abbas, F., Iftikhar, A., Riaz, A., Humayon, M., & Khan, M. F. (2024). Use of Big Data in IoT-Enabled Robotics Manufacturing for Process Optimization. Journal of Computing & Biomedical Informatics, 7(01), 239-248.
29. Munir, A., Sumra, I. A., Naveed, R., & Javed, M. A. (2024). Techniques for Authentication and Defense Strategies to Mitigate IoT Security Risks. Journal of Computing & Biomedical Informatics, 7(01).
30. Ahmed, F., Sumra, I. A., & Jamil, U. (2024). A Comprehensive Review on DDoS Attack in Software-Defined Network (SDN): Problems and Possible Solutions. Journal of Computing & Biomedical Informatics, 7(01).
31. Hussain, S.K., Ramay, S.A., Shaheer, H., Abbas T., Mushtaq M.A., Paracha, S., & Saeed, N. (2024). Automated Classification of Ophthalmic Disorders Using Color Fundus Images, Volume: 12, No: 4, pp. 1344-1348 DOI:10.53555/ks.v12i4.3153
32. Khan, A. H., Malik, H., Khalil, W., Hussain, S. K., Anees, T., & Hussain, M. (2023). Spatial Correlation Module for Classification of Multi-Label Ocular Diseases Using Color Fundus Images. Computers, Materials & Continua, 76(1).
33. Mahmood, M. A., Malik, H., Khan, A. H., Adnan, M., & Khan, M. I. A. (2022). Neural Network-Based Prediction of Potential Ribonucleic Acid Aptamers to Target Protein. Journal of Computing & Biomedical Informatics, 4(01), 21-36.