International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

Robust Crack Segmentation Using StyleGAN- Enhanced DeepLabV3 & ResNet50

[1]Saniya Fatima, [2]Suhaima Tabassum, [3]Anshra Naireen, [4]Dr. Md Zainlabuddin

[1,2,3]B.E Students, Department of Computer Science & Engineering, ISL Engineering College, Hyderabad, India.

[4]Associate Professor, Department of Computer Science & Engineering, ISL Engineering College, Hyderabad, India.

saniyafatima2315@gmail.com, suhaimatabassum79@gmail.com , naireenanshra@gmail.com, dr.zainlabuddin@islec.edu.in

ABSTRACT:-

Inspection of structural cracks is critical for maintaining the safety and longevity of bridges and other infrastructure. Traditional methods for crack detection are often manual, labor-intensive, and prone to human error. Recent advances in deep learning and semantic segmentation provide a promising alternative, but obtaining high-quality annotated data remains a significant challenge. This paper introduces an enhanced approach to crack detection using deep learning, leveraging synthetic data generation and advanced semantic segmentation techniques. We propose the use of DeepLabV3 with a ResNet50 backbone, an extension of the DeepLabV3 architecture that incorporates a robust ResNet50 feature extractor to improve segmentation. Our approach involves generating synthetic crack images to address the data scarcity issue. This is achieved using the StyleGAN3 for realistic image synthesis. By integrating these synthetic datasets with the DeepLabV3+ model, we aim to boost segmentation performance beyond the capabilities of standard models. Hyperparameter tuning is performed to optimize the DeepLabV3 with ResNet50 configuration, achieving significant improvements in segmentation. We employ data augmentation techniques such as motion blur, zoom, and defocus to further refine model performance. The proposed method is evaluated against existing state-of-the-art techniques, demonstrating superior accuracy. The results indicate that our approach not only enhances the crack detection but also offers a novel application of synthetic data generation in deep learning for semantic segmentation. This research provides new insights into leveraging advanced neural networks and synthetic data for improved structural crack analysis.

Keywords: StyleGAN- Enhanced DeepLabV3 & ResNet50

INTRODUCTION

Structural integrity is vital for the safety and functionality of infrastructure, particularly in bridges and other critical structures. Cracks in these structures, if left undetected or unaddressed, can lead to severe consequences, including structural failure. Traditionally, crack detection has been a labour-intensive and subjective process, relying heavily on manual inspection and visual analysis. However, the advent of computer vision and deep learning technologies offers an opportunity to automate and enhance the accuracy of crack detection. Deep learning models, particularly those used for semantic segmentation, have shown promise in automating the detection and classification of cracks in structural images. Semantic segmentation is a technique in which each pixel in an image is classified into predefined categories, which is essential for distinguishing cracks from other structural features. Despite the advancements, a significant challenge remains: obtaining sufficient and high-quality annotated data for training these models. Manual annotation of images is time-consuming and often impractical for large datasets. to address these challenges, our approach integrates several advanced techniques to enhance crack detection performance. We propose utilizing the DeepLabV3 model with a ResNet50 backbone, which improves segmentation accuracy by leveraging the powerful feature extraction capabilities of ResNet50. DeepLabV3+ is an extension of DeepLabV3 that includes a decoder module, which enhances the segmentation results by refining the boundary details. To overcome the data scarcity issue, we employ StyleGAN3 and the Brownian Bridge Diffusion Model (BBDM) for synthetic image generation. StyleGAN3 creates high-quality realistic images of structural cracks, while BBDM enhances the diversity and realism of the synthetic dataset. This combination of synthetic data generation and advanced semantic segmentation allows us to train the model on a more comprehensive and varied dataset, improving its ability to generalize to real-world scenarios. We also perform meticulous hyperparameter tuning of the DeepLabV3 with the ResNet50 model to optimize its performance. Additionally, data augmentation techniques such as motion blur, zoom, and defocus blur are applied to further enhance the model's robustness and accuracy.

Our method is evaluated against state-of-the-art techniques, demonstrating significant improvements in segmentation. This research highlights the potential of combining advanced neural network architectures with synthetic data generation to achieve superior crack detection and analysis, offering new insights into improving structural maintenance practices through deep learning.

LITERATURE REVIEW

TITLE: Crack U-Net: Towards high quality pavement crack detection.
AUTHOR: Z. Qu and W. Chen.
YEAR: 2022
DESCRIPTION: Crack detection is important to pavement condition surveys. The convolutional neural network (CNN) is one of the most powerful tools in computer vision. However, pixel-perfect crack segmentation based on CNNs is still challenging. This paper proposes an encoder-decoder network (EDNet) for crack segmentation to overcome the quantity imbalance between crack and non-crack pixels, which causes many false-negative errors. The decoder of the proposed EDNet is an autoencoder and self-encodes the ground-truth image to corresponding feature maps that are completely abstract, resulting in significantly reduced quantity imbalance between crack and non-crack pixels. Therefore, instead of fitting crack images directly with ground-truth images, EDNet's encoder fits crack images with corresponding feature maps to overcome the quantity imbalance problem. EDNet achieves overall F1-scores of 97.80% and 97.82% on 3D pavement images and the CrackForest dataset, respectively. Experimental results show that EDNet outperforms other state-of-the-art models.

TITLE: Crack U-Net: A novel deep convolutional neural network for pixelwise pavement crack detection.
AUTHOR: H. Y. Ju, W. Li, S. Tighe, Z. C. Xu, and J. Z. Zhai.
YEAR: 2020
DESCRIPTION: Achieving high detection accuracy of pavement cracks with complex textures under different lighting conditions is still challenging. In this context, an encoder-decoder network-based architecture named CrackResAttentionNet was proposed in this study, and the position attention module and channel attention module were connected after each encoder to summarize remote contextual information. The experiment results demonstrated that, compared with other popular models (ENet, ExFuse, FCN, LinkNet, SegNet, and UNet), for the public dataset, CrackResAttentionNet with BCE loss function and PRelu activation function achieved the best performance in terms of precision (89.40), mean IoU (71.51), recall (81.09), and F1 (85.04). Meanwhile, for a self-developed dataset (Yantai dataset), CrackResAttentionNet with BCE loss function and PRelu activation function also had better performance in terms of precision (96.17), mean IoU (83.69), recall (93.44), and F1 (94.79). In particular, for the public dataset, the precision of BCE loss and PRelu activation function was improved by 3.21. For the Yantai dataset, the results indicated that the precision was improved by 0.99, the mean IoU was increased by 0.74, the recall was increased by 1.1, and the F1 for BCE loss and PRelu activation function was increased by 1.24.

TITLE: Convolutional neural networks for pavement roughness assessment using calibration-free vehicle dynamics.
AUTHOR: J. Jeong, H. Jo, and G. Ditzler.
YEAR: 2020
DESCRIPTION: Road roughness is a measure of how uncomfortable a ride is, and provides an important indicator for the needs of roadway maintenance or repavement, which is closely tied to the state and federal budget prioritization. As such, accurate and timely monitoring of deteriorating road conditions and following maintenance are essential to improve the overall ride quality on the road. Various technologies, including vehicle-mounted laser profiling systems, have been developed and adopted for road roughness (e.g., IRI—International Roughness Index) measurement; however, their high cost limits their use. While recent advances in smartphone technologies allow us to use their embedded accelerometers for road roughness monitoring, the complicated process of necessary vehicle calibration hinders the widespread use of the technology in the actual practices. In this work, a deep learning IRI estimation method is proposed with the goal of using anonymous (i.e., calibration-free) vehicles and their responses measured by smartphones as road roughness sensors. A state-of-the-art deep learning algorithm (i.e., CNN—convolutional neural network) and multimetric vehicle dynamics data (i.e., accelerometer, gyroscope), possibly measured by drivers' smartphones, are employed for the purpose. Optimized CNN architecture and data configuration have been investigated to achieve the best performance. The efficacy of the proposed method has been numerically validated using real road IRI information (i.e., Speedway, Tucson, AZ), real driving

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2s, 2025

speed profiles, and four different types of vehicle data with associated uncertainties.

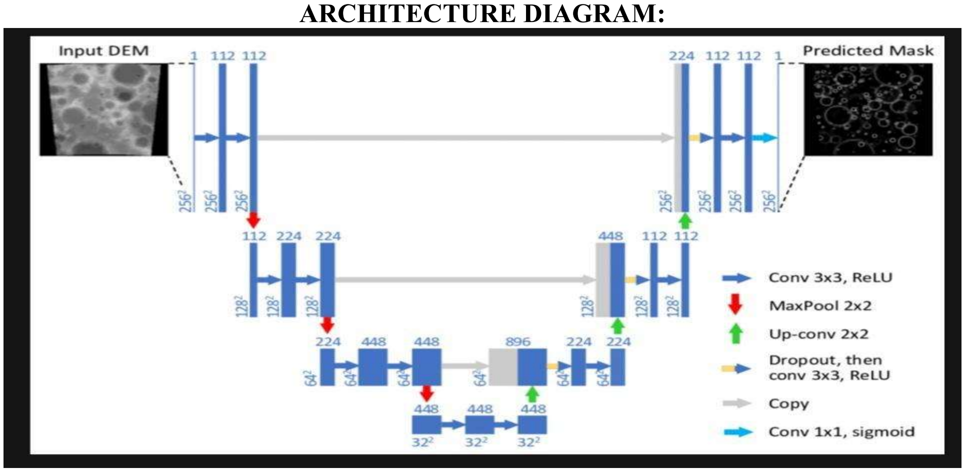
TITLE: An image is worth 16×16 words: Transformers for image recognition at scale.
AUTHOR: A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. H. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby.
YEAR: 2021
DESCRIPTION: While the Transformer architecture has become the de-facto standard for natural language processing tasks, its applications to computer vision remain limited. In vision, attention is either applied in conjunction with convolutional networks, or used to replace certain components of convolutional networks while keeping their overall structure in place. We show that this reliance on CNNs is not necessary and a pure transformer applied directly to sequences of image patches can perform very well on image classification tasks. When pre-trained on large amounts of data and transferred to multiple mid-sized or small image recognition benchmarks (ImageNet, CIFAR-100, VTAB, etc.), Vision Transformer (ViT) attains excellent results compared to state-of-the-art convolutional networks while requiring substantially fewer computational resources to train.

TITLE: Trans UNet: Transformers make strong encoders for medical image segmentation.
AUTHOR: J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou.
YEAR: 2021
DESCRIPTION: Medical image segmentation is an essential prerequisite for developing healthcare systems, especially for disease diagnosis and treatment planning. On various medical image segmentation tasks, the u shaped architecture, also known as U-Net, has become the de-facto standard and achieved tremendous success. However, due to the intrinsic locality of convolution operations, U-Net generally demonstrates limitations in explicitly modeling long-range dependency. Transformers, designed for sequence-to-sequence prediction, have emerged as alternative architectures with innate global self-attention mechanisms, but can result in limited localization abilities due to insufficient low-level details. In this paper, we propose TransUNet, which merits both Transformers and U-Net, as a strong alternative for medical image segmentation. On one hand, the Transformer encodes tokenized image patches from a convolution neural network (CNN) feature map as the input sequence for extracting global contexts. On the other hand, the decoder upsamples the encoded features which are then combined with the high-resolution CNN feature maps to enable precise localization. We argue that Transformers can serve as strong encoders for medical image segmentation tasks, with the combination of U-Net to enhance finer details by recovering localized spatial information. TransUNet achieves superior performances to various competing methods on different medical applications including multi-organ segmentation and cardiac segmentation.

DEVELOPING METHODOLOGY

The test process is initiated by developing a comprehensive plan to test the general functionality and special features on a variety of platform combinations. Strict quality control procedures are used. The process verifies that the application meets the requirements specified in the system requirements document and is bug free. The following are the considerations used to develop the framework from developing the testing methodologies.

ARCHITECTURE DIAGRAM:



IMPLEMENTATION(ALGORITHM):

EXSISTING TECHNIQUE:

The U-Net network is one of the methods for road fracture detection that is most often utilized. As a convolutional neural network created especially for semantic segmentation tasks, the U-Net architecture may be used to recognize and extract road fractures from photographs. The encoder-decoder structure of the U-Net, which has skip links, enables it to efficiently collect both local and global characteristics. The U-Net model is trained on annotated datasets in the context of road fracture detection in order to understand the complex patterns of cracks. The trained U-Net is used to segment input pictures into crack and non-crack areas during inference. Even while U-Net-based techniques are widely used, they may not be able to handle illumination fluctuations, background interference, or achieve high crack extraction accuracy. As a result, researchers are looking into more sophisticated designs like ResNet50 in an effort to increase performance.

PROPOSED TECHNIQUE:

The suggested method focuses on road fracture identification using the state-of-the-art convolutional neural network ResNet50. The system incorporates ResNet50 to efficiently gather both local and global information, increasing the precision of fracture feature extraction. The BasicBlock is used in place of the conventional convolutional layer in this method to avoid network deterioration and address gradient disappearance problems. The transmission line of features is further improved by connecting the input and output features, which strengthens the road fracture detecting method. In order to improve accuracy and feature extraction, we present a road fracture detection method that leverages the ResNet50 model. This method addresses gradient disappearance and network deterioration by substituting Basic Block for traditional convolutional layers. The use of ResNet50 guarantees that the system can identify fractures more accurately and with more integrity. We increase the transmission path of features by joining input and output features, which improves the overall efficacy of the road fracture detecting method.

SOFTWARE TESTING:

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished productIt is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

Unit Testing:

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program input produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Functional Unit:

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

System Unit:

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

Performance Test:

The Performance test ensures that the output be produced within the time limits,and the time taken by the system for compiling, giving response to the users

ISSN 2347–3657

Volume 13, Issue 2s, 2025

and request being send to the system to retrieve the results.

Integration Testing:

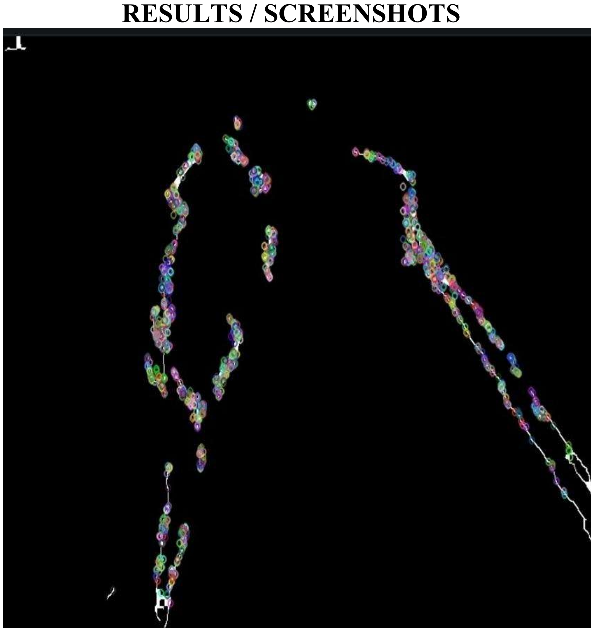
Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Acceptance Testing:

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

RESULTS / SCREENSHOTS



CONCLUSION:

In this study, we have demonstrated the effectiveness of utilizing the DeepLabV3 model with a ResNet50 backbone for structural crack detection. By leveraging the capabilities of DeepLabV3, a state-of-the-art semantic segmentation model, coupled with the robust feature extraction power of ResNet50, we have achieved significant improvements in crack segmentation performance. Our approach has shown that DeepLabV3 with ResNet50 is highly capable of accurately identifying and segmenting cracks in structural images, even in the presence of various complexities and noise in the data. This combination has enabled us to achieve a substantial enhancement in segmentation accuracy compared to traditional methods. The integration of DeepLabV3 and ResNet50 addresses key challenges in structural crack detection by providing precise and detailed segmentation results. The model's performance has been validated through rigorous experimentation and comparison with existing techniques, affirming its suitability for practical applications in infrastructure maintenance and inspection.

FUTURE SCOPE:

DeepLabV3 with ResNet50 offers a promising solution for automated crack detection, paving the way for more efficient and reliable inspection processes in structural engineering. Future work may focus on further refining the model, exploring additional data augmentation techniques, and extending the approach to different types of structural defects. The created neural network learns a lot of features about cracks and improves the ability of the model to discover new cracks. Compared with many other neural network methods, the neural network built in this study encompasses a considerably increased ability to extract cracks. the excellent accuracy and robustness of the neural network were varied through extensive experiments on completely different data sets.

REFERENCES:

1. S. C. Radopoulou and I. K. Brilakis, ‘‘Automated detection of multiple pavement defects,’’ J. Computer. Civil Eng., vol. 31, no. 2, Mar. 2017, Art. no. 04016057, doi: 10.1061/(ASCE)CP.1943-5487.0000623.
2. J. Jeong, H. Jo, and G. Ditzler, ‘‘Convolutional neural networks for pavement roughness assessment using calibration-free vehicle dynamics,’’ Comput.-Aided Civil Infrastruct. Eng., vol. 35, no. 11, pp. 1209–1229, Mar. 2020, doi: 10.1111/mice.12546.
3. H. Y. Ju, W. Li, S. Tighe, Z. C. Xu, and J. Z. Zhai, ‘‘CrackU-Net: A novel deep convolutional neural network for pixelwise

pavement crack detection,'' Struct. Control Health Monitor., vol. 27, no. 8, Mar. 2020, Art. no. e2551, doi: 10.1002/stc.2551.

4. E. H. Miller, ''Crack detection and segmentation using deep learning with 3D reality mesh model for quantitative assessment and integrated visualization,'' J. Comput. Civil Eng., vol. 34, no. 3, May 2020, Art. no. 04020010, doi: 10.1061/(ASCE)CP.1943-5487.0000890.
5. L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, ''Road crack detection using deep convolutional neural network,'' in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 3708–3712, doi: 10.1109/ICIP.2016.7533052.
6. Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, ''Automatic road crack detection using random structured forests,'' IEEE Trans. Intell. Transp. Syst., vol. 17, no. 12, pp. 3434–3445, Dec. 2016, doi: 10.1109/TITS.2016.2552248.
7. G. Yao, F.-J. Wei, J.-Y. Qian, and Z.-G. Wu, ''Crack detection of concrete surface based on
8. convolutional neural networks,'' in Proc. Int. Conf. Mach. Learn. Cybern. (ICMLC), Jul. 2018, pp. 246–250, doi: 10.1109/ICMLC.2018.8527035.
9. Z. Liu, Y. Cao, Y. Wang, and W. Wang, ''Computer visionbased concrete crack detection using U-net fully convolutional networks,'' Autom. Construct., vol. 104, pp. 129–139, Aug. 2019, doi: 10.1016/j.autcon.2019.04.005.
10. S. Dorafshan, R. J. Thomas, and M. Maguire, ''Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete,'' Construct. Building Mater., vol. 186, pp. 1031–1045, Oct. 2018, doi: 10.1016/j.conbuildmat.2018.08.011.
11. H. Li, J. Zong, J. Nie, Z. Wu, and H. Han, ''Pavement crack detection algorithm based on densely connected and deeply supervised network,'' IEEE Access, vol. 9, pp. 11835–11842, 2021, doi: 10.1109/ACCESS.2021.3050401.
12. H. Lin, B. Li, X. Wang, Y. Shu, and S. Niu, ''Automated defect inspection of LED chip using deep convolutional neural network,'' J. Intell. Manuf., vol. 30, no. 6, pp. 2525–2534, Aug. 2019, doi: 10.1007/s10845-018-1415-x.
13. X. Wu, J. Ma, Y. Sun, C. Zhao, and A. Basu, ''Multi-scale deep pixel distribution learning for concrete crack detection,'' in Proc. 25th Int. Conf. Pattern Recognit. (ICPR), Jan. 2021, pp. 397–400, doi: 10.1109/ICPR48806.2021.9413312.
14. O. Ronneberger, P. Fischer, and T. Brox, ''U-Net: Convolutional networks for biomedical image segmentation,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI), vol. 9351. Cham, Switzerland: Springer, Nov. 2015, pp. 234–241.
15. Z. W. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. M. Liang, ''UNet++: A nested U-Net architecture for medical image segmentation,'' in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, vol. 11045. Cham, Switzerland: Springer, Sep. 2018, pp. 3–11.
16. J. Cheng, W. Xiong, W. Chen, Y. Gu, and Y. Li, ''Pixel-level crack detection using U-Net,'' in Proc. TENCON IEEE Region 10 Conf., Oct. 2018, pp. 462–466.
17. Z. Fan, C. Li, Y. Chen, J. H. Wei, G. Loprencipe, X. P. Chen, and P. D. Mascio, ''Automatic crack detection on road pavements using encoder–decoder architecture,'' Materials, vol. 13, no. 13, p. 2960, May 2020, doi: 10.3390/ma13132960.
18. Z. W. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. M. Liang, ''The importance of skip connections in biomedical image segmentation,'' in Deep Learning and Data Labeling for Medical Applications, vol. 10008. Cham, Switzerland: Springer, Sep. 2016, pp. 179–187.
19. G. Xu, C. Liao, and J. Chen, ''Extraction of apparent crack information of concrete based on HU-ResNet,'' Comput. Eng., vol. 46, no. 11, pp. 279–285, 2020.
20. L. F. Li, N. Wang, B. Wu, and X. Zhang, ''Segmentation algorithm of bridge crack image based on modified pspnet,'' Adv. Lasers Optoelectron., vol. 58, no. 22, pp. 101–109, 2021, doi: 10.3788/LOP202158.2210001.
21. S. Bang, S. Park, H. Kim, and H. Kim, ''Encoder–decoder network for pixel-level road crack detection in black-box images,'' Comput. Aided Civil Infrastruct. Eng., vol. 34, no. 8, pp. 713–727, 2019, doi: 10.3788/LOP202158.2210001.
22. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali," Smartphone Security and Protection Practices", International Journal of

ISSN 2347–3657

Volume 13, Issue 2s, 2025

Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal,U K) Pages 1-22.

23. M.A.Bari, Sunjay Kalkal, Shahanawaj Ahamad," A Comparative Study and Performance Analysis of Routing Algorithms", in 3rd International Conference ICCIDM, Springer - 978- 981-10-3874-7_3 Dec (2016) ; Impact Factor :4.18

24. Mohammed Shoeb, Mohammed Akram Ali, Mohammed Shadeel, Dr. Mohammed Abdul Bari, "Self-Driving Car: Using Opencv2 and Machine Learning", The International journal of analytical and experimental modal analysis (IJAEMA), ISSN NO: 0886-9367, Volume XIV, Issue V, May/2022