Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM

ISSN: 2788-7421

The Use of DCNN for Road Path Detection and Segmentation

Nada Mohammed Murad[1,*], Lilia Rejeb[1], Lamjed Ben Said[1]

[1]SMART Lab. Institut Supérieur de Gestion de Tunis (ISG), Université de Tunis, 41 Avenue de la Liberté Bouchoucha, 2000, Tunisia

*Corresponding Author: Nada Mohammed Murad

DOI: https://doi.org/10.52866/ijcsm.2022.02.01.013
Received March2022; Accepted May 2022; Available online June 2022

ABSTRACT: In this study, various organizations that have participated in several road path-detecting experiments are analyzed. However, the majority of techniques rely on attributes or form models built by humans to identify sections of the path. In this paper, a suggestion was made regarding a road path recognition structure that is dependent on a deep convolutional neural network. A tiny neural network has been developed to perform feature extraction to a massive collection of photographs to extract the suitable path feature. The parameters obtained from the model of the route classification network are utilized in the process of establishing the parameters of the layers that constitute the path detection network. The deep convolutional path discovery network's production is pixel-based and focuses on the identification of path types and positions. To train it, a detection failure job is provided. Failure in path classification and regression are the two components that make up a planned detection failure function. Instead of laborious postprocessing, a straightforward solution to the problem of route marking can be found using observed path pixels in conjunction with a consensus of random examples. According to the findings of the experiments, the classification precision of the network for classifying every kind is higher than 98.3%. The simulation that was trained using the suggested detection failure function is capable of achieving an accuracy of detection that is 85.5% over a total of 30 distinct scenarios on the road.

Keywords: Machine Learning; Deep Learning; Road Paths; Segmentation; Detection; Neural Network; DCNN

1. INTRODUCTION

The Adaptive Driver Assistance Technology (ADAT) is dependent on a car driver's understanding of road situations. Automobiles share a road with other cars, path lines, pedestrians, and so on. Thus, recognizing road paths are crucial in the field of ADAT. Path detection may be used to create a path-departure warning structure and to steer self-driving cars in the right direction. Traditional path detection systems used complicated image processing to acquire path features that were previously created by humans [1]. Deep learning, often known as neural networks, has shown great promise in image processing in recent years. Different established image processing methodologies, e.g., deep neural networks, extract proper features after being trained on many images [2]. The use of the Deep Learning (DL) method in path detection is a topic of this study.

Most of the earlier path recognition as well as tracking studies relied on developed hypothetical path shape patterns or path color attributes. The limitations of the selected model have a large effect sitting on the detection performance of a hypothetical shape pattern. The calculation of path curvature in hyperbola [3] decides whether the path pixels may be investigated. The retrieved controller pixels are utilized to match the parametric pattern in the B-snake model [4]. Piecewise route segments [5] and the adoptive low path model [6] are two further path shape models.

The path shape is also straight lines within a bird's-eye position. The camera's characteristics are required for generating an inverse perspective transformation matrix to generate a bird's-eye view [7, 8]. In [9], the path is depicted as two parallel lines, with the camera in the middle of the automobiles. Unfortunately, shape model-based path detection is typically inflexible. Most path shape models often function well on paths with the exact geometry. The most visible feature at path

*Corresponding author: aven.jaff00@yahoo.com

boundaries in paint feature-based path detection is the area under gradient alter [10], which is magnified in gray pictures obtained via linear recognition [11].

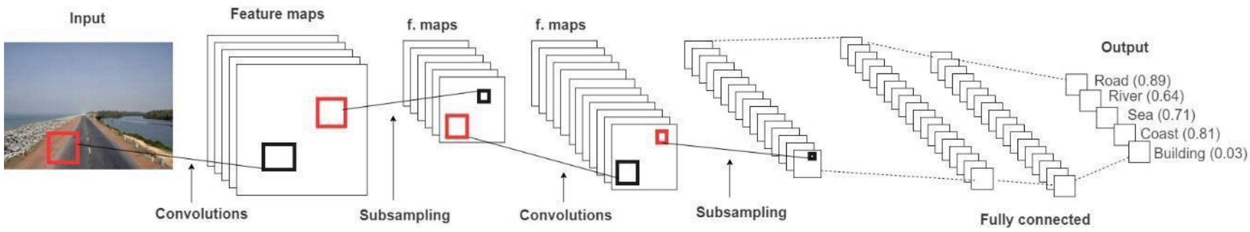


FIGURE 1. General view of the suggested approach.

2. BACKGROUND

Observing the different path color characteristics in [12] allows for the detection of light invariant path marker candidates [13, 14]. Some researchers combine geographical information with masks to recognize a path. Some researchers employ this combination. This path recognition approach is more complex to build than others because it requires considerable image preprocessing to function properly. Other techniques do not have this prerequisite. When attempting to recognize pathways, making use of characteristics is generally successful [15]. To display the path, the random sample consensus (RANSAC) [16] is frequently used with other algorithms such as Kalman filters [5], the filter of the G.S.P.F [17], and geometry for path identification via crossroads [18].

Deep learning has spawned the development of several cutting-edge detection strategies that are truly extraordinary. It makes use of a selective search algorithm known as SS [19] in region-based convolutional neural networks (R-CNN) [20]. Moreover, R-CNN is improved using regional proposal networks by Faster R-CNN [21]. The (SSD & YOLO) [22, 23] partition a feature map of every convolution level travel, which eliminates the requirement for proposed region extraction. Huge path segments can be divided into more manageable pieces using [24]. The coordinates and depth of each little frame that makes up a path section are the outputs of the detection.

In [24], any path classifications will not be found. It also generates a deep convolutional network that does not have any deep linked layers and score regions that are evaluated based on whether they are potential candidates for item detection via a regional fully connected neural network (FCN). These deep learning detection algorithms are quite impressive, but they require extremely large and complex networks to work. Moreover, huge networks necessitate the addition of new servers, storage space, and processing capacity to keep up with the demand.

In this study, the deep convolutional neural network (DCNN) approach is given for the recognition of road paths. The suggested method involves two main elements: path classification and path detection. Path features are retrieved for path classification using neural networks rather than manually constructed characteristics. A detection loss function must be created to account for the training of the neural network results in the destruction of some or all the output feature maps. The detection loss is a combination of two different kinds of losses, i.e., the path classification failure and the path regression failure. No differentiation was made by the detection network between the suggested regions and the feature maps. Consequently, a deep convolutional network is applied to identify the existence of a path classification network. Thus, the detection network is more condensed and uses fewer available computational resources.

Other deep learning detection methods create their output as rectangular frames that cover path segments, whereas the suggested detection network generates detection on a pixel-by-pixel basis. The detection network is trained from the ground up, beginning with the unprocessed input pictures and culminating with detection maps that include the positions and classifications of pixels in the path region. The training process begins with the raw input images and ends with detection maps. Path marking with RANSAC can also be explored when utilizing a detection network model as the basis [25–29].

3. METHODOLOGY

A fundamental neural network was constructed to extract features of the dataset used for the data of the path class. When developing the path detection network, the deep-linked layers of classification nets are replaced together with the convolution levels. A detection network model can be obtained once the layers of the detection network have been initialized using the parameters of the class of the data network model, and the detection network has been trained to use the detection loss function that has been provided. Consequently, the route is denoted by the pixels that have been identified. A high-level summary of the recommended technique can be found in Fig 1.

The results of the efforts of the current study are twofold. The appropriate path features are extracted, beginning with a neural network as the foundation. The subsequent plan is to come up with a detection loss function that takes into consideration both classification and regression failures during the process of calculating the error function. The detection network can identify paths and their associated classifications using pixel-by-pixel classifications and locations. Pixel-by-pixel detection makes it possible to go beyond path form limitations.

3.1 MACHINE LEARNING

Observations, direct experience, or instruction are used to start machine learning. It searches for patterns in data to conclude from examples. Machine learning (ML) allows systems to learn and adapt without human interaction. Moreover, it analyzes and builds data-learning algorithms. ML can solve problems faster and more efficiently than the human mind. Machines may be trained to discover patterns and relationships in input data and automate regular procedures with huge computational power. ML training includes supervised learning, unsupervised learning, and reinforcement learning. Supervised machine learning algorithms use labeled samples to predict future events. The learning algorithm analyzes a known training dataset to predict output values. The system can target a new input after training. Its output may be compared with the anticipated output to discover flaws and adjust the model. Unsupervised machine learning methods use unclassified, unlabeled training data. Unsupervised learning examines how computers can infer a function from unlabeled data. The correct output is nknown becauseiinfers output from datasets. Reinforcement machine learning algorithms interact with their surroundings to find errors or rewards and feature trial-and-error search and delayed rewardsSreward feedback is needed for the agent to learn the best action.

3.2 GENERAL VIEW OF THE STRUCTURE OF CONVOLUTIONAL NEURAL NETWORKS

This network is known as both an FCN and a multilayer perceptron artificial neural network (MLP) because it has at least two completely connected (dense) layers of the network. The FCN moniker refers to the network's status as a fully connected neural network, while the MLP moniker refers to the network's as shown in the previous figure.

All the neurons in the feed layer are associated with each other, and an ultimate output is a completely associated layer with a score that corresponds to a specific class. Large-scale images cannot be handled by convolutional neural networks (CNN). The total number of such calculations is a given. Moreover, computations should also be done quickly. A complete network connection is not the best option.

Neurons are in three dimensions. To take advantage that images are used as inputs, CNN constricts the structure of the network rationally. Thus, Conv.Net layers contain neurons that are arranged in three dimensions instead of the usual two. In this context, depth refers to the third dimension of the activation size and not the depth of the neural network. A network's entire layer count may be referred to as layer count. Its dimensions are $32 \times 32 \times 3$, for example, and serve as an activation input for CIFAR-10. For example, neurons in a layer can only communicate with the preceding layer in a limited area, and not in a connected manner. The final CIFAR-10 output layer is $1 \times 1 \times 10$ because the image will be reduced to a single-degree vector at the end of the convolutional neural architecture (Fig. 2).

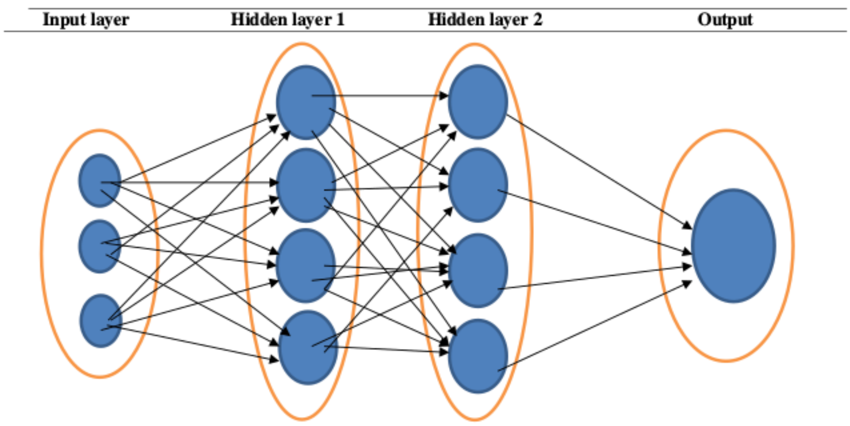


FIGURE 2. Example of hidden CNN layers.

Three-layer neural networks are the most common. CNN neurons are arranged in three dimensions (width, height, and depth) in a single layer, as seen at the bottom. An input volume is transformed into an output volume by neurons in each CN layer. If the image is stored in the red input layer, its dimensions (width and height) will be equal to the image's dimensions (depth; red, green, and blue channels; Fig. 3).

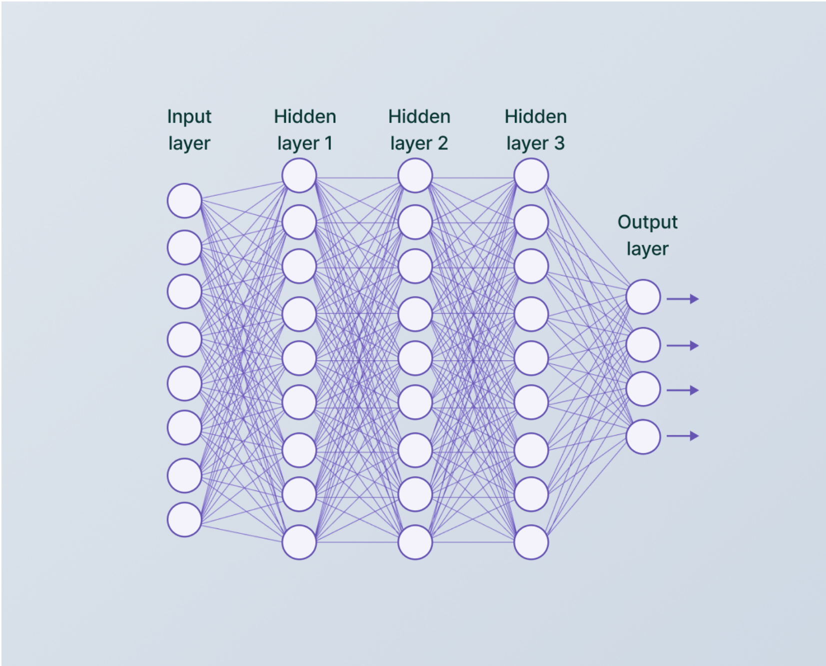


FIGURE 3. Example of three hidden CNN layers.

In this approach, these final class scores are applied one layer at a time, beginning with the values of the original image and working their way up to the values of the newly added pixels. Layers that have and do not have parameters were noted. When processing the data, the CONV/FC layers considers both the input volume activations and the parameters (the weights and bias of neurons). This allows for a more accurate data representation (the weights and biases of the neurons). Moreover, each of the RELU/POOL layers will be responsible for a particular activity. The gradient descent algorithm will be applied to train the parameters in the CONV/FC layers to guarantee that the CN class scores follow the labels included in the training set.

The equation below explained the example with an input volume with a dimension of $16 \times 16 \times 20$. Using an example-receptive field size of 3×3, every neuron in the Conv. Layer would now have a total of $3 \times 3 \times 20 = 180$ connections to the input volume. However, the connectedness is limited in space (e.g., 3×3) in the entire input depth (20).

$$\text{output } f\left(\sum_{k=0}^{n} w\, ixi + b\right) =$$
$$\text{input } (w0x0 + w1x1 \ldots.) + \text{ input depth } (b)$$

3.3 SEGMENTATION

Segmentation is a fundamental approach for image processing used in picture analysis, e.g., quantitative valuation of factors and computer-assisted analysis. In other terms, it is an imaging technology (CAD). The capacity of a lesion to be segmented serves as the foundation for determining its classification. Segmentation may need to be performed manually by specialized personnel. However, CNN is an option to consider. Processing the entire image and immediately producing a segmentation result based on the data that it processes is a typical practice for image segmentation systems. The traineeship data of the segmentation structure are photographs that contain a structure of interest as well as the segmentation outcomes.

3.4 DETECTION

The examination of strange images is a common component of many tasks that involve the search for irregularities. Anomalies are so rare and they must be identified. The DCNN efficacy for detection was investigated in this paper. The researchers identified utilizing the DCNN. A detection method was developed with the help of the dataset, and its effectiveness was evaluated.

3.5 CONVOLUTIONAL FILTERS

CNNs are a specific type of neural network model that is designed to work with 2D picture data. Even though they can also be applied to one-dimensional and 3D data, CNNs are designed to work with two-dimensional picture data. In a convolution, the input is processed similar to a typical neural network which is multiplied by a set of weights. In a two-dimensional architecture, multiplying a set of input data is accomplished using a filter or kernel, which is a 2D array of weights. This architecture is called 2D architecture.

Using a filter-sized patch of the input data and a dot product for the multiplication allows for a reduction in the overall size of the filter. A single value is produced by first multiplying each item of input and each filter item individually with a dot product, and then adding the products of those individual multiplications together. Consequently, the output of the scalar product operation is typically referred to as a single value. Using a filter that is smaller than the array of data that was being processed was decided because it is possible to multiply the same filter, which is a collection of weights, by the input array a great number of times. The filter applies a logical processing order to each overlapping patch of incoming data, beginning on the left and working its way to the right in a clockwise direction.

The outcome is a single value when a filter and input array are multiplied once. A 2D array of production values is created when a filter is utilized in numerous periods to an input array. Consequently, the *feature map* in Fig. 4 shows that the feature map is the name given to the resulting two-dimensional array. Moreover, the convolution of three-dimensional input data is demonstrated in Fig. 6 (RGB).

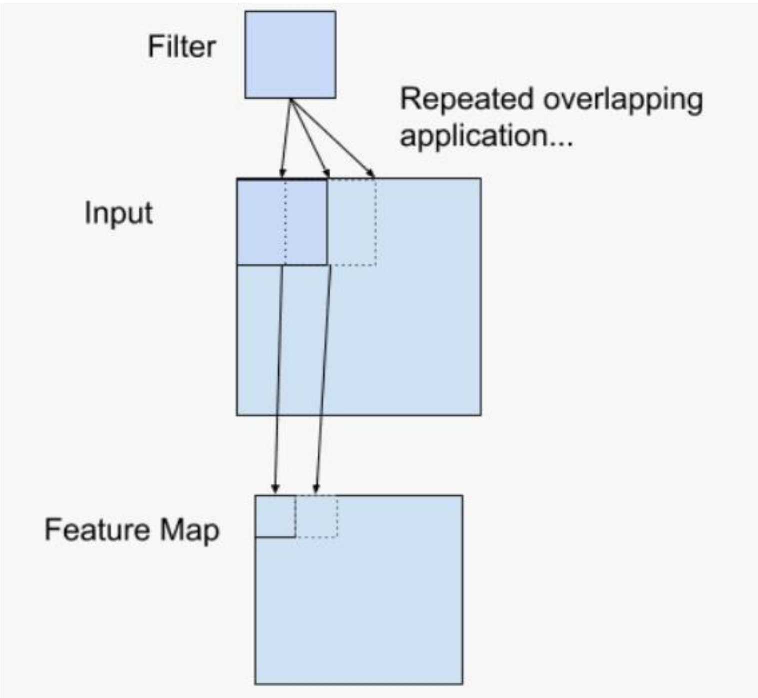


FIGURE 4. Filter used in 2D input to make the feature map.

Padding Fig. 7. No padding was noted in DCNNs. Thus, the edges of the image cannot be reached because some of the filters would be outside the image, which is not possible. Padding the input tensor with zero-valued values increases the margins.

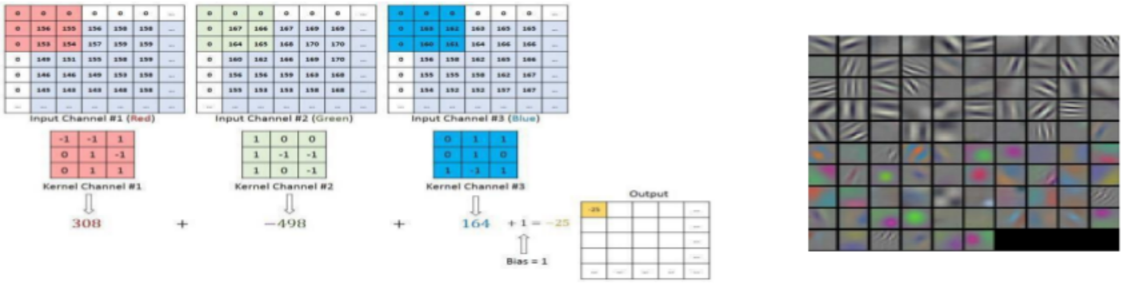


FIGURE 5. ***Left*** **multichannel input convolution [2].** ***Right*** **output of convolutional filters [8].**

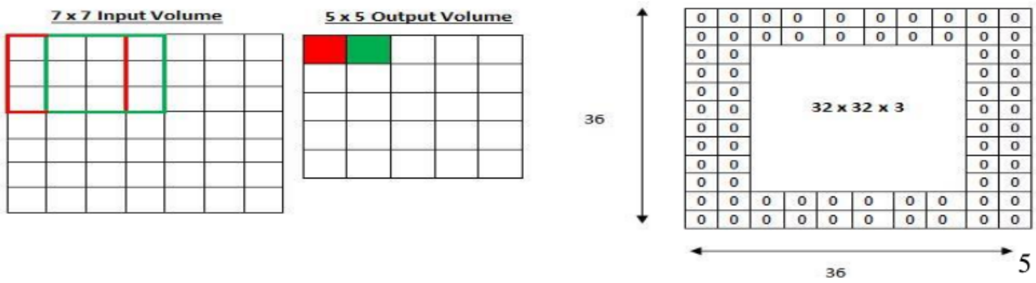


FIGURE 6. **Case of stride of volume 1 which is the default (*left*). Zero-padding of volume 2 (*right*) [8].**

4. RESULT

The test ratio is set to 2 in addition to a detection net that was trained together with a ratio to acquire the correct ratio between classification as well as regression failure. The greatest type is produced by adjusting the test ratio in the direction of ’2’, ’1’, and ’1.2’ following the experiment results. The best detection outcome is when a network can be trained to output path types and positions in various situations using the defined detection loss function. Figure 8 shows the original image in the first row and the segmentation of the original image is shown in the second and third rows. Moreover, Fig. 8 depicts the results of both the first and enhanced detection.

The classification type is created by training a classification system with 200 periods, then evaluating it using the accuracy misunderstanding matrix in Table 1. Moreover, as shown in Table 1, a white path has the highest accuracy rate of 90.2%. In contrast, the lowest accuracy rate is 90.5%. The wrong fraction is extremely small, ensuring that the trained simulation has retrieved an appropriate quality for path recognition.

Table 1. **Accuracy Misunderstanding Matrix of Path Classification Net (in percentage).**

	Background	Yellow line	White line
Background	90.2	0.3	20..
Yellow line	0.2	90.5	2.01
White line	0.40	0.90	90.73

Table 2. **Accuracy Misunderstanding Matrix of Path Classification Net (in percentage).**

Source	Algorithm	Accuracy
[30]	Recurrent neural network	97.71%
[31]	Support vector machine	89.26%
Proposed	**Deep convolutional neural network**	**98.3%**

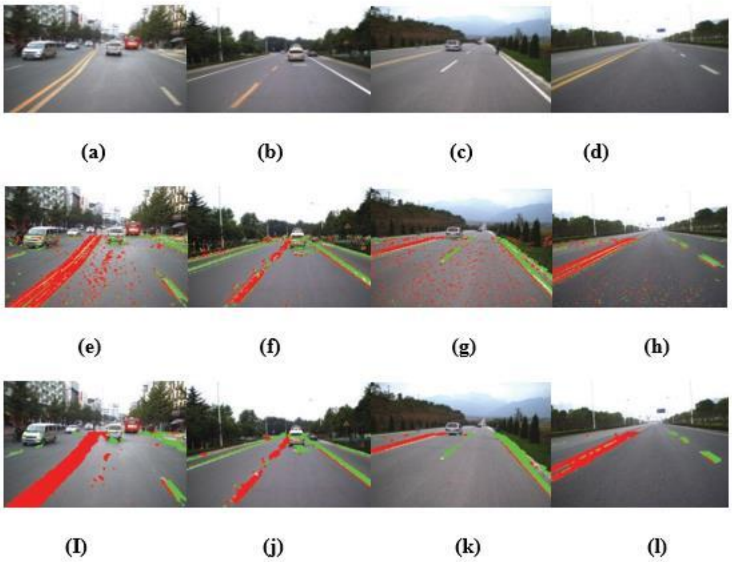


FIGURE 7. **The initial detection findings as well as the enhanced detection results. The test photos in the first row are a–d. The detection findings are shown in the second row in e–h. The enhanced detection outcomes via the neighboring pixels are shown in i–l in the row. The color observed yellow path pixels red as well as the identified white path pixels green.**

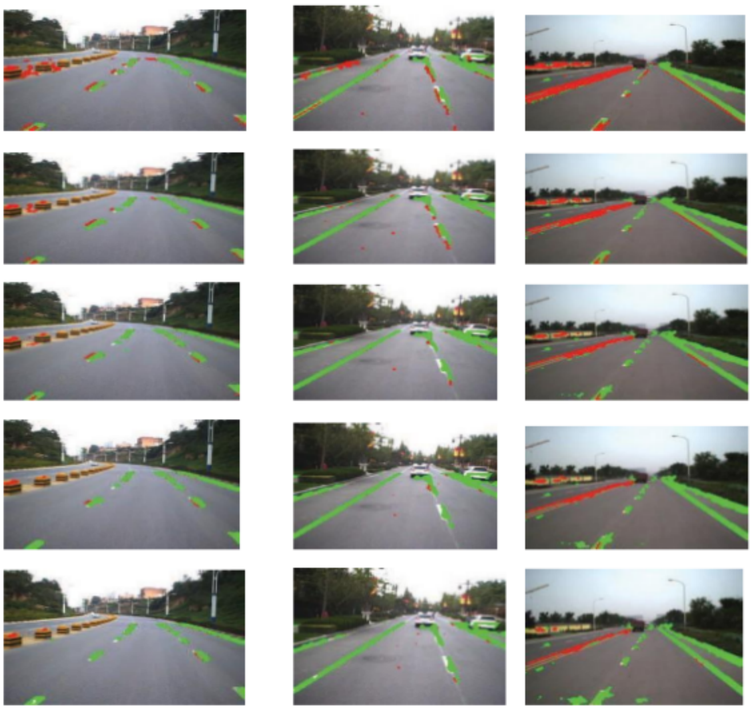
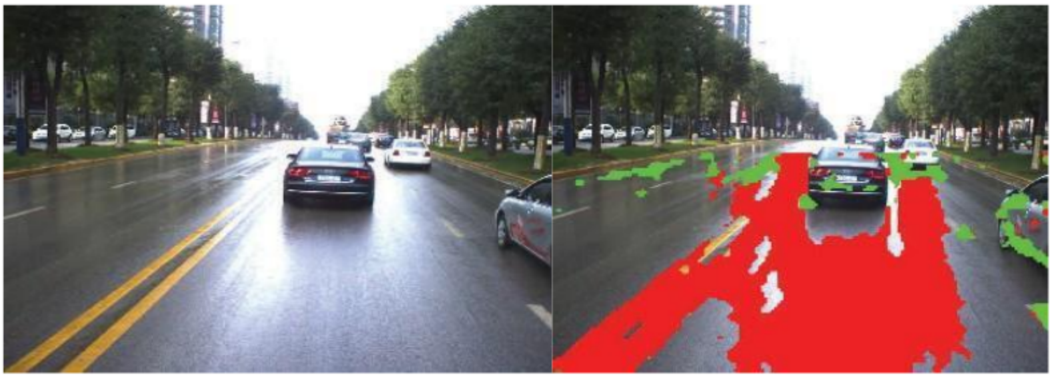


FIGURE 8. **Performing type appraisal through various ratios between (alpha (α) and beta (β)).**



(a) Original image (b) Segment image

FIGURE 9. The fragmented scene with numerous reflection.

5. CONCLUSION

This study offers a case for the installation of a road path identification system that is underpinned by DCNN. To extract features of acceptable pathways, a neural network that is both tiny and shallow is developed. The accuracy of categorization is higher than 90.2% on each path. Developing an effective detection network simulation is fairly difficult for route detection after a detection network has been trained on images of various road conditions. This adds to the intricacy of the process. Despite adverse detection outcomes in one high ranking-reflection picture, a leading model's detection accuracy in performance can nevertheless attain 85.5% across the board for the rest of the dataset under consideration. At the end of the operation, the detected path pixels are joined to represent the whole path.

5.1 FUTURE WORK

In the future, photos in various weather and lighting circumstances will be used to fine-tune the derived network model to increase the detection performance. Furthermore, utilizing more advanced deep learning architecture, the path marking process will be altered to change the variation of fitting paths.

ACKNOWLEDGEMENT

The first author would like to thank the reviewers for providing useful suggestions, allowing for the improved presentation of this paper.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] H. Tan, Y. Zhou, Y. Zhu, D. Yao, and J. Wang, "Improved River flow and random sample consensus for curve path detection," *Advances in Mechanical Engineering*, vol. 7, no. 7, pp. 1687814015593866–1687814015593866, 2015.
[2] Y. Wang, D. E. K. Teoh, and Shen, "Path detection and tracking using b-snake," *Image & Vision Computing*, vol. 22, no. 4, pp. 269–280, 2014.
[3] D. Schreiber, B. Alefs, and M. Clabian, "Single camera path detection and tracking," in Intelligent TransportationSystems, 302-307," 2005.
[4] R. Jiang, K. Reinhard, V. Tobi, and S. Wang, "Path detection and tracking using a new path model and distance transform," *Machine Vision & Applications*, vol. 22, pp. 721–737, 2017.
[5] A. Borkar, M. Hayes, and M. T. Smith, "Robust path detection and tracking with ransac and kalman filter," *IEEE International Conference on Image Processing*, pp. 3225–3228, 2014.
[6] Z. Ying and G. Li, "Robust path marking detection using boundary-based inverse perspective mapping," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1921–1925, 2016.
[7] D. Ding, J. Yoo, J. Jung, and S. Jin, "Various path marking detection and classification for vision-based navigation system?," *IEEE International Conference on Consumer Electronics*, pp. 491–492, 2015.
[8] J. Li, M. Xue, D. Prokhorov, and D. Tao, "Deep neural network for structural prediction and path detection in road scene," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 28, no. 3, pp. 690–703, 2017.
[9] H. Jung, J. Min, and J. Kim, "An efficient path detection algorithm for path departure detection."
[10] A. Parajuli, M. Celenk, and H. B. Riley, "Robust path detection in shadows and low illumination conditions using local gradient features," *Open Journal of Applied Sciences*, vol. 3, no. 1, pp. 68–74, 2016.

[11] H. Yoo, U. Yang, and K. Sohn, “Gradient-enhancing conversion for illumination-robust path detection,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1083–1094, 2016.
[12] J. Son, H. Yoo, S. Kim, and K. Sohn, “Real- time illumination invariant path detection for path departure warning system,” *Expert Systems with Applications*, vol. 42, no. 4, pp. 1816–1824, 2015.
[13] U. Suddamalla, S. Kundu, S. Farkade, and A. Das, “A novel algorithm of path detection addressing varied scenarios of curved and dashed pathmarks,” *International Conference on Image Processing Theory, TOOLS and Applications*, pp. 87–92, 2016.
[14] U. Ozgunalp and N. Dahnoun, “Robust path detection & tracking based on novel feature extraction and path categorization,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8129–8133, 2014.
[15] A. Borkar, M. Hayes, S. Smith, and Pankanti, “A layered approach to robust path detection at night.”
[16] A. Martin, R. C. Fischler, and Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, ACM,” 2016.
[17] Y. Wang, N. Dahnoun, and A. Achim, “A novel system for robust path detection and tracking,” *Signal Processing*, vol. 92, no. 2, pp. 319–334, 2015.
[18] S. Vitor, P. V. K. Bottazzi, B. Borges, J. Stantic, and Jo, “Adaptive regions of interest based on hsv histograms for path marks detection,” *Advances in Intelligent Systems & Computing*, vol. 274, pp. 677–687, 2014.
[19] J. R. R. Uijlings, K. E. A. V. D. Sande, T. Gevers, and A. W. M. Smeulders, “Selective search for objectrecognition,” *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2016.
[20] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmenta- tion,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, 2014.
[21] K. S. Ren, R. He, J. Girshick, and Sun, “Faster r- cnn: towards real-time object detection with region proposal networks,” *International Conference on Neural Information Processing Systems*, pp. 91–99, 2015.
[22] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, Y. Cheng, A. C. Fu, and Berg, “Ssd: Single shot multibox detector, 21-37,” 2015.
[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” *Computer Vision and Pattern Recognition*, pp. 779–788, 2016.
[24] S. Nageshrao, H. E. Tseng, and D. Filev, “Autonomous Highway Driving using Deep Reinforcement Learning,” *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019.
[25] S. R. A. Ahmed and E. Sonuç, “Deepfake detection using rationale-augmented convolutional neural network,”Applied Nanoscience,” 2021.
[26] M. T. Mahmood, S. R. A. Ahmed, and M. R. A. Ahmed, “Detection of vehicle with Infrared images in Road Traffic using YOLO computational mechanism,” *IOP Conference Series: Materials Science and Engineering*, vol. 928, pp. 22027–22027, 2020.
[27] S. Abdulateef, S. R. Ahmed, Ahmed, and M. D. Salman, “A Novel Food Image Segmentation Based on Homogeneity Test of K-Means Clustering,” *IOP Conference Series: Materials Science and Engineering*, vol. 928, pp. 32059–32059, 2020.
[28] M. R. Ahmed, S. R. Ahmed, A. D. Duru, O. N. Uçan, and O. Bayat, “An Expert System to Predict Eye Disorder Using Deep Convolutional Neural Network,” *Academic Platform Journal of Engineering and Science*, vol. 9, no. 1, pp. 47–52, 2021.
[29] M. Waleed, A. S. Abdullah, and S. R. Ahmed, “Classification of Vegetative pests for cucumber plants using artificial neural networks,” *” 2020 3rd International Conference on Engineering Technology and its Applications (IICETA)*, 2020.
[30] P. Shivakumara, T. Palaiahnakote, . D. Tang, L. Maryam, P. Umapada, and . H. Anisi, “CNN-RNN based method for license plate recognition,” *CAAI Transactions on Intelligence Technology*, 2018.
[31] S. Abdullah, Omar, Khairuddin, Sahran, and M. S. Khalid, “License plate recognition based on Support Vector Machine,” 2009.