A CNN Approach To Detect Vanishing Point By Considering Street Lane

Shruthiba A
Artificial Intelligence and Machine Learning
Bangalore Institute of Technology
Bengaluru, India
shruthibaa@bit-bangalore.edu.in

Deepu R
Computer Science and Engineering
ATME College of Engineering
Mysuru, India
rdeepusingh21@gmail.com

Amit Dhomne
CSE, NIT Rourkela
Rourkela, Odisha, India
amitdhomne97@gmail.com

2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS) | 978-1-6654-0610-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSITSS54238.2021.9683182

***Abstract*— Linear perspective is widely used in highway road images to create the impression of depth on 2D/3D photos and detect vanishing points with the help of parallel lanes lines on road images. Automated understanding of linear perspective in landscape, road, street photo collections video has a number of real-world applications. This problem of automated understanding of linear perspective in images is addressed with the help of hough transform and CNN framework architecture to detect vanishing points. However, images of the road taken from street and highway video pose a very great technical challenge, because an insufficient number of parallel edges or lines intersect and lead to false vanishing points. To solve this problem, a state of art vanishing point detection method is proposed that exploits the vanishing point and intersection of parallel lines with the help of possible parallel lines to lane lines available in the image, the center of origin, and quadrant. The proposed strategy essentially performs best in detecting vanishing points on a public road image data set.**

Keywords— Vanishing Point (VP), Convolution Neural Network (CNN), VGGNet, Hough Transform, Computer Vision, Depth Estimation.

I. INTRODUCTION

Digital devices like cameras represent the world as a two-dimensional (2D) matrix. Since, cameras represent three-dimensional scenes in two-dimensional matrices the third dimension i.e., depth is lost. Recovering scene depth has been a subject undergoing intense study in computer vision and has broad applications. Depth estimation of a scene from a single image is easier for humans, but it's very challenging for computational models to do with high precision and fewer resource requirements.

A vanishing point refers to a point on the image plane when a set of parallel lines intersect at an infinite location [Fig. 1]. In computer vision, detection of objects, calibrating cameras, steering a robot, and reconstruction of 3D images can be done by one very interesting research on estimating and detecting vanishing points in a given 2D image. In geometric properties of projection, points project to points and lines project to lines in an image plane. This geometric property provides geometric feature information in a given scene which is a strong cognitive feature for human visual perception.

Initial work of depth estimation included gathering depth information by stereo-vision, structure from motion but those results came out to be ambiguous, further work was in camera calibration where research was done capturing images with a slight displacement of the camera, in turn deriving the relationship between lens and projection plane of the camera. Nevertheless, these methods did not involve direct depth information. Subsequently, a method called depth from defocus/focus was introduced where depth could be inferred focusing only on certain objects in images. However, it failed in texture-less images. Recently, methods proposed for detection are to train deep neural networks with labeled data.

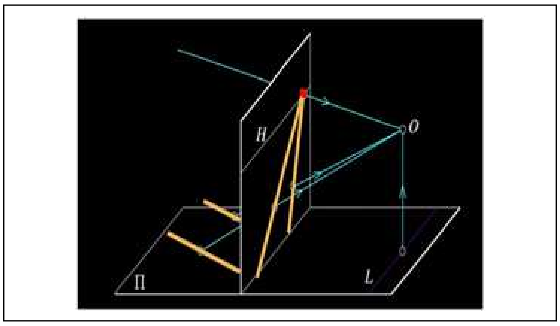


Fig 1: Parallel lines converged to a single point in the projection/image plane.

Detection of vanishing point is a crucial step in-depth estimation which can be done using various ways such as using Alex net [1] which is more effective on occluded pictures, rotated images, dark scenes, by using a single CNN for road scenes [5] and under various weather conditions [17]. Vanishing point can be detected using various networks such as conic convolution network [6], continuous CRF [20], CNN, deep learning and Markov Random Field (MRF) [10], deformable convolution, and deformable RoI pooling and in an end-to-end paradigm of multiple CNNs architectures [22]. Some of the CNN techniques used for vanishing point detection are line detectors [7], line clustering [8], the 3D orientation of omnidirectional camera [9], J-linkage algorithm [4], multiscale heatmap supervised learning [13], least-squares [11], uncalibrated camera [12], Helmholtz principle [14], segmentation of the scene into quads [3], image structure analysis [19], by measuring the performance using the metric maxima [16], by identifying similar pixels of an input image [20]. Hence, the vanishing point is the prominent method in estimating the depth with the help of a convolutional neural network.

II. RELATED WORKS

To detect VP in Google street-view image dataset, CNN[1] is used to recover the horizon line with 99% accuracy and achieve 92% in locating the VP in the original ALEXNet architecture by changing only the number of filters to 124 in the last two fully connected layers. Later in [2], proposed a method to perceptively project quads by rectifying the segmentation pixel-wise on designing a novel CNN architecture where segmentation is generated geometrically for each pixel orientation for a single street view image.

More recently, in [3], Zihan Zhou et al., address the problem of detecting the VP in the natural street image by proposing a method where global structures in the given image are exploited by combining contour-based edge detector with J-linkage. But fails to handle images in which the linear perspective is absent. Later In [10], Lei Geng et. al. propose a simple linear iterative clustering algorithm that segments road images into superpixels of uniform size. Using this algorithm CNN is trained to classify road and nonroad regions and based on the relationship between the superpixel neighborhood, MRF is utilized to optimize the classification results. But, the algorithm fails on images taken at night, under the rain, snow, or harsh weather conditions.

In addition, in [13], Yin-Bo Liu et. al. proposed a novel solution by combining the CNN and heatmap regression to detect unstructured road maps. Hierarchical features of the unstructured road images are extracted by adopting lightweight backbone depthwise convolution modified HRNet. But, texture detection performance strongly depends on image quality. In [15], Seokju Lee et.al., address the problem of detecting VP in images taken in rainy and low illumination conditions. VPGNet is proposed to detect and classify lanes, road markings and predict a VP with a single forward pass and use the quadrant VPP method for evaluation.

In this paper, the research problem of detecting VPs is addressed in Street Road images, parallel street images of roads connecting several cities by proposing a new method for dominant VP detection by combining a contour-based edge detector, CNN, Hough Transform, edge detection algorithm, and finding quadrants in images. CNN is composed of multiple processing layers, trained end to end to automatically learn the representation of data with multilevel abstraction.

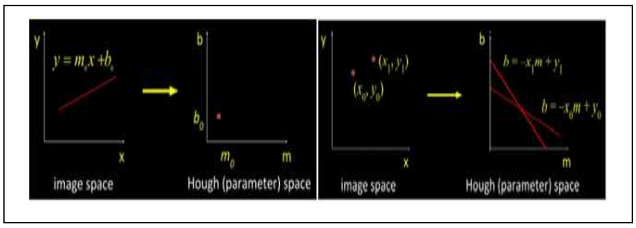
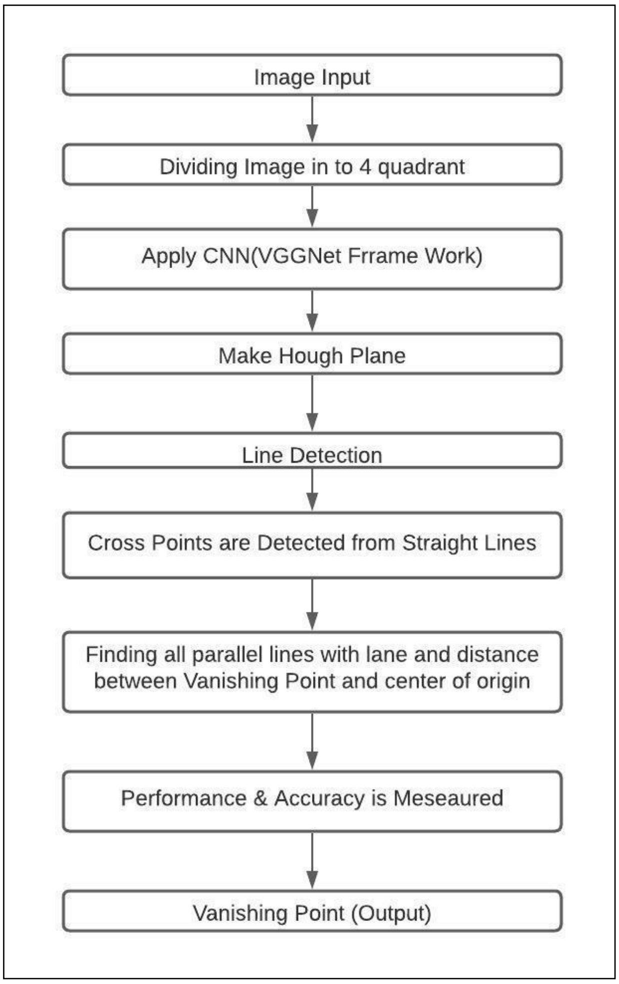


Fig 2: Line in image space corresponds to a point in Hough space and point in image space corresponds to a line in Hough space.

Flowchart :



Algorithm:

Step 1: Read the image.

Step 2: Reduce the image size and improve the quality by adding extra pixels around the image. For some images, we don’t need this step and we will set it to 0. However, other images are too big, and difficult to see where the lines converge. So, this step will help us easily visualize where the vanishing point is, by adding extra pixels around the image.

Step 3: Convert image to grayscale.

Step 4: Thresholding is the simplest method of image segmentation. From grayscale images, thresholding can be used to create binary images. We can say that it's a filtering method because in this case, we are using thresholding to remove unwanted lines. Some of the sample images have different thresholding values, while most of them have the same value.

Step 5: Applying edge detection. Edges are very important in Hough transform. Is an image processing technique for finding the boundaries of objects within images.

Step 6: Applying Hough Transform and finding edges.

Step 7: Apply CNN VGGNet architecture to detect vanishing point with 33,000 frames of image

Step 8: Obtaining distance between quadrant center (0,0) with another parallel line to check if the intersecting line is the correct line.

Step 9: Plotting the vanishing point where the lines intersect.

III. VANISHING POINT DETECTION

In this paper,

1. CNN is used to detect vanishing points.
2. With the help of the Hough transform method, we will detect possible parallel lane lines & find the distance between different edges that exist in the quadrant.

Convolutional Neural Network (CNN) VGGNet architecture:

To train deep neural networks, often a large amount of data is needed. We took images from YouTube videos including road trips, highway roads across America These videos have been captured in a variety of weather and ground conditions Eventually, we had 37,497 frames (resized to 300 × 300 pixels). To train a convolutional neural network, we need a large amount of data. We captured images from YouTube downloaded videos including highway road trips across the American region for example Sedan, buses, trucks, cars, traveling on road.

The Hough Transform is applied to the smoothed grayscale picture in the picture division venture, to identify straight lines. In the proposed technique, the greatest Hough is set to top worth to 20 and recognizes 10 lines that crossed the pinnacle point. In the proposed technique, line conditions are utilized to identify the interfacing line from the picture limit to the start and end and characterize the crossing points of these lines as applicant VP's. Fig 3, shows the consequence of applying the Hough to the input image and its corresponding info picture. The square situation in the Hough transform framework alludes to the pinnacle of greatest worth.

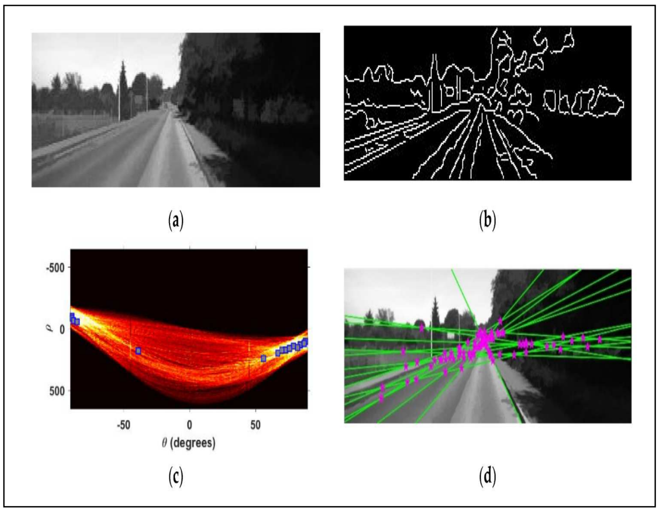


Fig. 3: Hough transforms analysis for finding parallel lines based on the quadrant.

The theta range was set to [−70: −70, 20:70] to prohibit level and vertical lines through the Hough change. Following identification of the line, convergence focuses dependent on the Hough transform, just crossing points that existed in the applicant area of the genuine VP were chosen in Fig. 2(d). Crossing point choice was performed at the convergence of the lines having a place with the space situated from the second to the fourth regions when the picture was upward partitioned into five sections, and from the second to the third regions when the picture was evenly separated into five sections.

Nonetheless, the straight-line highlights not related to vanishing points additionally exist on or close to the street, for example, stop lines, utility poles, the mainstays of street signs, also, passerby spans. Albeit electric wires stretching out corresponding to a street merge at the vanishing point, they are not straight-line highlights as a rule since they are slack.

The straight-line highlights related to vanishing point identification have a typical property. The principal straight-line highlights exist on the right and left rectangular areas of the lower half of the road view camera picture. As a rule, few powerful straight-line highlights exist in the space of the upper a big part of the road view camera picture. What's more, in the neighborhood of the middle in the space of the lower a large portion of the straight-line highlights related to vanishing points are less inclined to exist. All things being equal, straight-line highlights not related with the vanishing point, for example, bolts and jewels demonstrating the approach of a person on foot crossing, are bound to exist. In this way, a thin focal area in the lower half is barred, as displayed in Fig.2(a.) Also, the points of the straight-line highlights related to vanishing point identification are 100 degrees to 170 degrees in the right region, and 10 degrees to 80 degrees in the left region. Most are straight-line highlights opposite or even regardless of whether the straight-line highlights not related with vanishing point location out and about could be remembered for this area. Hence, it is thought that the vast majority of the straight-line highlights not related to vanishing point location can be eliminated by decreasing the scope of points of the straight lines utilized in the recognition as displayed in Fig. 2(b). The recognition of straight-line highlights not related to vanishing point identification is generally decreased by an impact of the space division and point constraint. Accordingly, a run roadway limit line is hard to distinguish by Hough change in examination with a strong line, yet the straight-line degree somewhat becomes higher by utilizing the proposed strategy, and it is believed that it turns out to be not difficult to identify.

Then, at that point, the Hough plane is made by utilizing $\rho = x\cos\theta + y\sin\theta$ as it were in Fig.1 region information.

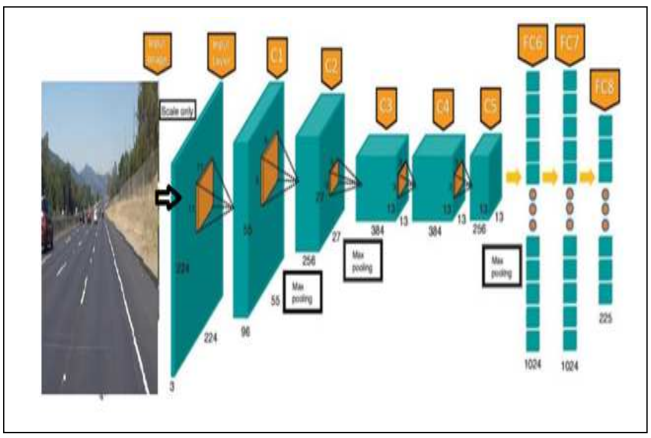


Fig 4: CNN VGGNet framework architecture.

To solve camera twisting and non-linear, a similar map is made physically from configured camera viewpoint to picture Y-axis, utilizing a polynomial norm of the third order, where parameter comes from reading road image captured from YouTube video in our dataset

$y=a\theta^3+b\theta^2+c\theta+d$

In our situation, we are working on vanishing point identification as a CNN-based framework Firstly, we discretize mathematical vanishing points into discrete labels. As it were, vanishing point identification is changed to classification problems. For a given picture, predicting the area of the vanishing point is identical to predicting the discrete label of its vanishing point. The predicted labels are remapped to mathematical coordinates in the picture, acquiring the vanishing point area.

IV. Mathematical Implementation

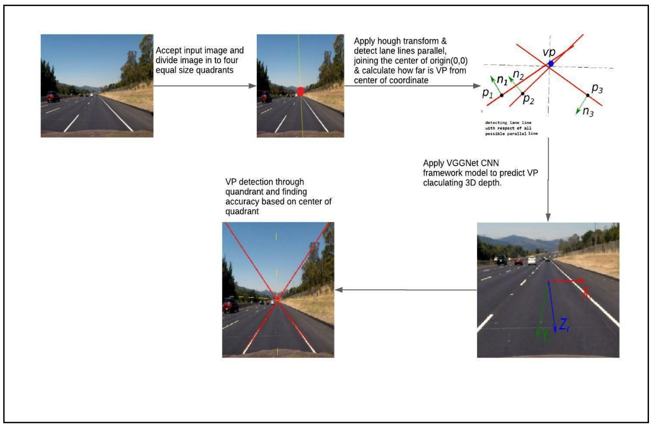


Fig 5: Mathematical overview of detecting vanishing points.

The Hough methodology is particularly used to find global descriptions of features where it is not mandatory to know the number of solution classes in prior, which is mostly noise local measurements. The motivation behind the application of the Hough transform method is to detect that the line for each input value (coordinate point) helps in finding an accurate solution, for example, the straight line which will generate image points. Here, n1,n2 are coordinates of the line concerning different angles parallel to the possible lane line and VP is the vanishing point, and p1,p2,p3 represents the parallel line that is formed through hough transform.

Consider a normal situation, where an issue of adjusting a multiple line segment or set of line segments to a set of discrete image points which is pixel location as output generated based on edge detection Hough transform method.

The figure below shows a set of possible solutions to a given problem.

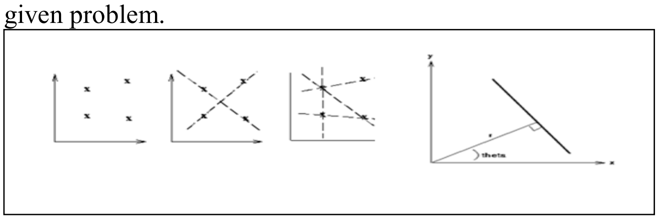


Fig 6: Line segment based on a different angle concerning the origin.

There are different ways to describe line segments. The equation for describing a set of lines uses *parametric* or *normal* notions:

$$x\cos\theta + y\sin\theta = r$$

Where, **r** = length of a normal from the origin to this line,,**θ** is the orientation of **r** taking into consideration into x-axis as shown in Fig.5 For any point (x,y) on this line, **r** and **θ** are constant.When viewed in Hough parameter space, points that are collinear in the cartesian image space become readily apparent as they yield curves that intersect at a common (r,θ) point. As the algorithm runs, each **(X_i, Y_i)**is transformed into a discretized **(r,θ)** curve and the accumulator cells which lie along this curve are incremented. For instance, in the case of *circles*, the parametric equation is

$$(x-a)^2 + (y-b)^2 = r^2$$

Where r is the radius of the circle and a, b are the coordinates of the center of the circle. This hough transform helps in detecting all possible parallel lines in 3-D or 2-D space so concerning the center of origin so we can draw parallels to lanes and can check what are different parallel line is intersecting and joining the nearby center of origin. Distance between Center of origin & r (meeting point of vanishing or parallel line) is stated as

|Coordinate of new vanishing point - center of origin | == range (0 ,1) -------final equation

V. Result

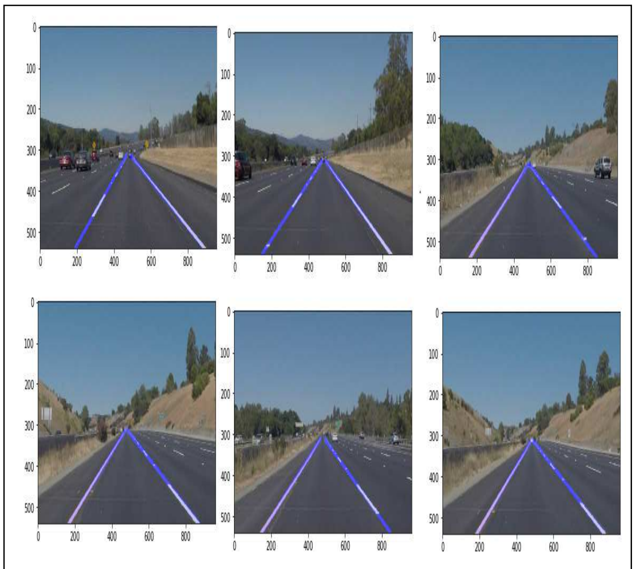


Fig 7: Snapshots of detected Vanishing Points

Using the concept of drawing possible parallel line segments based on a different angle, origin from fig 6, and projected vanishing point from the final equation as shown above we can find vanishing point considering lane line and trying to show in fig 7.

Fig 7 shows how lane lines converge to give vanishing point and accuracy graph Fig 8 shows how to change on a number of nodes in a Fully connected layer in CNN will impact the accuracy of CNN network detecting vanishing point

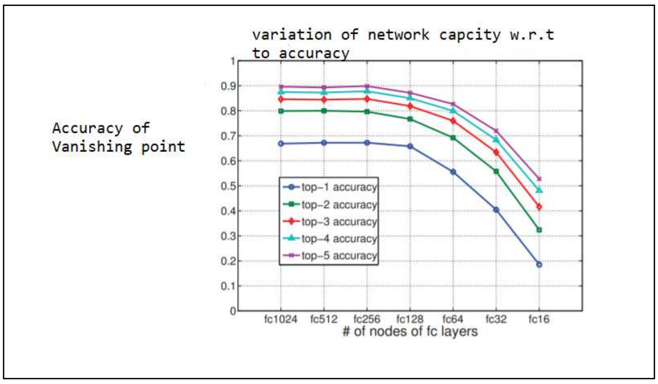


Fig 8: Accuracy graph of the Vanishing point.

VI. Conclusion

Thess outcomes show that VGGNet based vanishing point detection with the help of the Hough transform method performs extremely well in detecting vanishing points considering the lane line of the road. VGGNet based CNN framework with 33,000 frames for predicting vanishing point areas has a wide scope. Additionally note that we gathered the biggest vanishing point dataset to date, with more than 10K pictures, to give better insight, data preparing, testing, and advancement for future vanishing points. This to some extent clarifies that our CNN VGGNet framework performs well with our dataset with the help of hough transform.

References

[1] Shuai, Yan, et al. "Regression convolutional network for vanishing point detection." *2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE, 2017.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Shi, Yongjie, et al. "Three Orthogonal Vanishing Points Estimation in Structured Scenes Using Convolutional Neural Networks." *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019. K. Elissa, “Title of paper if known,” unpublished.

[3] Zeng, Zhiliang, et al. "Deep Recognition of Vanishing-Point-Constrained Building Planes in Urban Street Views." *IEEE Transactions on Image Processing* 29 (2020): 5912-5923.

[4] Tardif, Jean-Philippe. "Non-iterative approach for fast and accurate vanishing point detection." *2009 IEEE 12th International Conference on Computer Vision*. IEEE, 2009.

[5] Borji, Ali. "Vanishing point detection with convolutional neural networks." *arXiv preprint arXiv:1609.00967* (2016).

[6] Zhou, Yichao, et al. "NeurVPS: neural vanishing point scanning via conic convolution." *arXiv preprint arXiv:1910.06316* (2019).

[7] Zhai, Menghua, Scott Workman, and Nathan Jacobs. "Detecting vanishing points using global image context in a non-manhattan world." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

[8] Bazin, Jean-Charles, et al. "Globally optimal line clustering and vanishing point estimation in manhattan world." *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012.

[9] Shi, Yongjie, et al. "3D Orientation Estimation and Vanishing Point Extraction from Single Panoramas Using Convolutional Neural Network." *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

[10] Geng, Lei, et al. "Combining CNN and MRF for road detection." *Computers & Electrical Engineering* 70 (2018): 895-903.

[11] Antunes, Michel, and Joao P. Barreto. "A global approach for the detection of vanishing points and mutually orthogonal vanishing directions." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013.

[12] Kogecka, J., and Wei Zhang. "Efficient computation of vanishing points." *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*. Vol. 1. IEEE, 2002.

[13] Liu, Yin-Bo, Ming Zeng, and Qing-Hao Meng. "Unstructured road vanishing point detection using convolutional neural networks and heatmap regression." *IEEE Transactions on Instrumentation and Measurement* 70 (2020): 1-8.

[14] Almansa, Andrés, Agnes Desolneux, and Sébastien Vamech. "Vanishing point detection without any a priori information." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.4 (2003): 502-507.

[15] Lee, Seokju, et al. "Vpgnet: Vanishing point guided network for lane and road marking detection and recognition." *Proceedings of the IEEE international conference on computer vision*. 2017.

[16] Tsai, Yi-Min, Yu-Lin Chang, and Liang-Gee Chen. "Block-based vanishing line and vanishing point detection for 3D scene reconstruction." *2006 international symposium on intelligent signal processing and communications*. IEEE, 2006.

[17] Liu, Fayao, Chunhua Shen, and Guosheng Lin. "Deep convolutional neural fields for depth estimation from a single image." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

[18] Hu, Junjie, Yan Zhang, and Takayuki Okatani. "Visualization of convolutional neural networks for monocular depth estimation." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.

[19] Garg, Ravi, et al. "Unsupervised CNN for single view depth estimation: Geometry to the rescue." *European conference on computer vision*. Springer, Cham, 2016.

[20] Dai, Jifeng, et al. "Deformable convolutional networks." *Proceedings of the IEEE international conference on computer vision*. 2017.

[21] Deepu, R., S. Murali, and Vikram Raju. "A Mathematical model for the determination of the distance of an object in a 2D image." Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2013.