

Video based Traffic Forecasting using Complication Transmit Learning Techniques

Shubham Kumar¹, Pardeep², Ravinder³ and Sunil⁴

¹M. Tech. Scholar, Department of Civil Engineering, Satpriya group of institutions, Rohtak, Haryana
sp4496738@gmail.com

²Assistant Professor, Department of Civil Engineering, Satpriya group of institutions, Rohtak, Haryana
moharpardeep@gmail.com

³PhD Scholar, Department of Civil Engineering, MMU Mullana, Haryana
ravinderpanwar7043@gmail.com

⁴M. Tech. Scholar, Department of Civil Engineering, Satpriya group of institutions, Rohtak, Haryana
sunilmalik8607@gmail.com

Abstract

Many state and local agencies are currently facing challenges concerning the collection and estimation of traffic volumes, particularly regarding “the collection of annual average daily traffic (AADT) on low-volume roads. The connection between satiate moral forecasting of traffic and video prediction is identified in this paper. With the developments in technology, traffic signals are replaced with smart systems and video streaming for analysis and maintenance of the traffic all over the city. Processing of these video streams requires lot of effort due to the amount of data that is generated. This paper proposed a simplified technique for processing such voluminous data. The large data set of real-world traffic is used for prediction and forecasting the urban traffic. A combination of predefined kernels are used for spatial filtering

and several such transferred techniques in combination will convolution artificial neural networks that use spectral graphs and time series models. Spatially regularized vector auto regression models and non-spatial time series models are the baseline traffic forecasting models that are compared for forecasting the performance. In terms of training efforts, development as well as forecasting accuracy, the efficiency of urban traffic forecasting is high on implementation of video prediction algorithms and models.

Keywords: *network-wide forecasts; spatial filtering; convolution neural networks; spatiotemporal models; urban trafficflow;*

1. Introduction

The models and principles that are observed and studied in one environment can be applied for refining the results of problems in a different environment. This concept is termed as domain adaptation or transfer learning. One example of transfer learning technique is the implementation of video prediction model for urban traffic flow monitoring and forecasting. The input feature space of a trained model can be transformed to match the inputs of the domain or outputs specific to the domain can be produced by replacing the final layer in machine learning techniques for transferring the models that are implemented commonly. Translator functions are developed for creating a link between the image and text classification issues for conversion of image and text specific features into a common feature set. Classification of animals is performed by convolution Neural system model to execute image processing which can be transferred and fine-tuned for object classification.

Along with implementation of pre-trained models in other domains, transfer learning can also be

extended to inter domain problem solving, application of the architecture of general model and so on in a more general context. Implementation of carefully tested model architectures and pre-trained models is widely acknowledges in transfer learning due to its practical advantages and resource saving features in model training. The collection of training dataset and computational power are the major resources required for model training and can be applied for lower order domains only. More focused and faster research processes are obtained by transfer learning technique that merges principles and models of various domains and advance the scientific development process sufficiently.

2. Literature Review

Detection of traffic flow in transportation is a critical area of research among researchers to assist the decision-making processes [5]. Research is done by various researchers in the field of transportation to predict traffic flow. Image and digital signal processing domains are used widely in video prediction techniques. Filtering, which is a major signal processing concept implements spatial or temporal domain convolution with kernels that are designed specifically? Spatial convolution of kernels that are coded into bi-dimensional signals from a single frame of a video stream or an image. Demising with local adaptive kernels, feature and edge detection with first- and second-order kernels, restoration and smoothing of Gaussian kernels are some of the most popular kernels in use. Based on the temporal information, data-driven kernels are developed and trained in various previous studies. Based on the theoretical background, these kernels are developed and applied on a wide scale in the domain of image processing.

Despite the improved results obtained on implementation of convolution in image processing by means of spatial kernels, the elimination of temporal information is a major limitation that prevents application of this process to video streams. Spatial information is used for prediction of images and their segments from every frame of a video. Restoration of video segment using image in painting is a major issue however, this cannot be termed as pure video prediction. Motion and such spatiotemporal features can be detected using rich data structures like 2 spatial and 1 temporal dimensions in a three-dimensional (3D) video stream. In case of multivariate signals, the prediction and processing requires advanced techniques and significant computational power. Until the early 2000s, the video prediction problem was not addressed closely. The video prediction advances are due to the improvements in hardware and the progress in artificial neural networks (ANN). Pixel-wise video prediction using feed-forward ANN architecture and video demising using temporal restricted Boltzmann machine are some of the applications of ANN in processing videos. The spatiotemporal convolution kernels that are data driven is used for horizontal and vertical neighbors with single or dual temporal lags in the input space for every pixel. The complex spatial patterns and long-term temporal dependencies are not allowed in the basic feed-forward ANN architecture.

Traffic forecasting in urban areas using spatiotemporal video prediction methodology along with the transfer learning concept is implemented in this paper and the results are validated. Popular techniques for video and image processing are selected arbitrarily and their performance is tested against the state of the art models for traffic flow data. Considering n spatial locations and X_t as $n \times 1$ vector of traffic values $(x_{1,t}, x_{2,t}, \dots, x_{n,t})$ at $i=1, \dots, n$ spatial locations at time period t . Weighted directed graphs are used for representing the one-period spatial structure in which the edges are used for coding the relationships and vertices for coding the spatial locations. Spatial structure approaches have a wide range of applications. The uncongested traffic based travel time is considered as a primary measure of relationship in this study. The VAR model in its regularized form is defined using the cross-correlation of the spatial structure. The spatial location and local neighborhood of the spatial structures are limited based on the complexity using the following formula”:

$$NB(i, r) = \{j: d_{ij} < r \text{ and } i \neq j\}$$

Here, the neighborhood's radius selected radius is represented by r and the spatial location travel time is d_{ij} . Within a specific road network, the direct flow maximum travel time is defined as the radius for forward and backward movement without loops. Two neighborhoods are also considered simultaneously for similar model specifications

with different radii. Figure 1 provides the general architectural diagram of the Graph Convolution Neural system(GCNN).

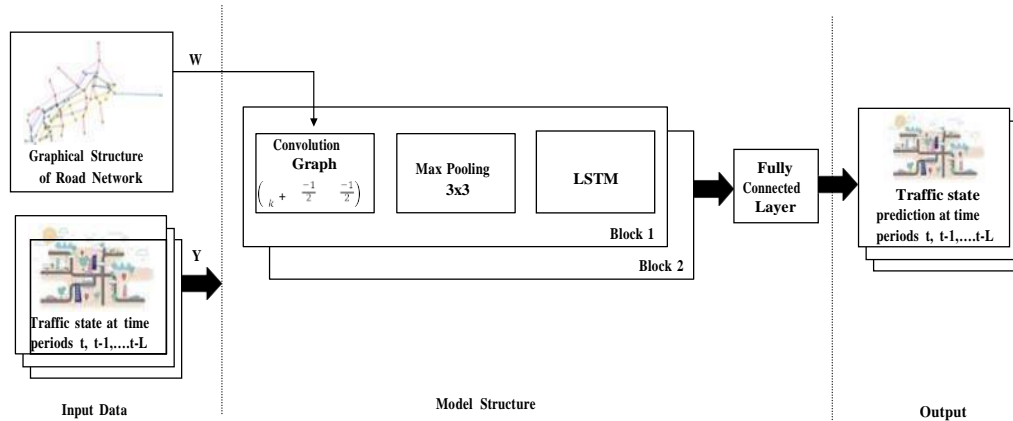


Figure 1: Graph Convolution Neural system(GCNN) general architecture

“The graph structure can be defined by matrix of weights W and $D = d(W_1^T)$ is the diagonal matrix in which the all-ones vector is represented by 1. The spectral-based graph convolution can be represented by the following expression under such conditions”:

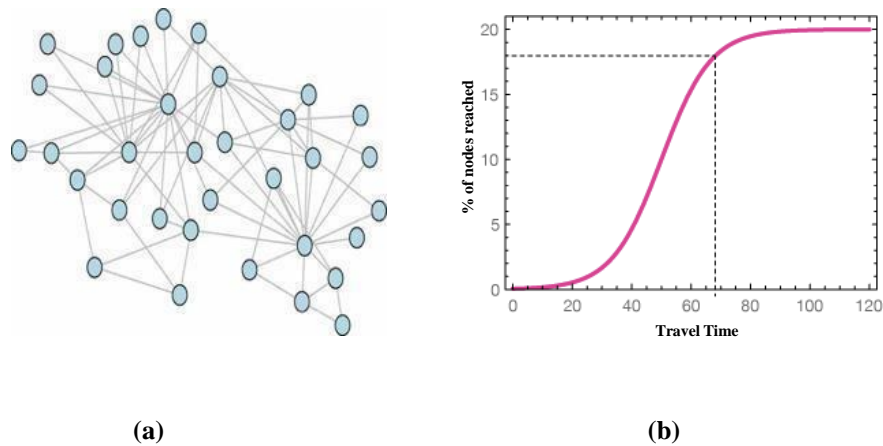
$$g_{\theta} \times Y = \theta \left(I_k + D^2 W D^2 \right) Y$$

Here, “ θ is the trained matrix and the filter is represented with g_{θ} and identity matrix is I_k . The classical CNN architecture is modified in the GCNN along with the graph convolution filter inclusive of the pooling layers and convolution layers that are connected sequentially into a fully connected final layer.

3. Proposed Material and Methodology

After applying the available test data into the proposed traffic volume prediction model traffic volume will be predicted. For testing the accuracy of the model prediction, two error formulas are used, the first is used to test the accuracy of single sample and second is used to the accuracy of the whole sample. The difference between predicted traffic volume corresponding to the observed traffic volume of week days is shown in Fig 5 . The total error in the proposed model is calculated by the Mean

Absolute Percentage Error [35](MAPE).With the proposed method on average the MAPE error for week days is 15% The error shows that the predicted traffic volume have relatively good correlation with the actual traffic volume. This reduces the MAPE error by modifying the membership functions and rules. Indian Driving Dataset (IDD), which has about 34 classes, a large real-world dataset is used for estimating the performance of the video signal processing models. The information contains about 10,000 pixel- level annotated images and 50,000 object level annotated images, contained in 5,000 frames is used. A total of 34 labels are used for annotation from 182 drive sequences on Indian roads. The images are mostly of 1080p resolution, but there is also some images with 720p and other resolutions. The analysis is performed for a mean distance of 500 meters. Random sampling of 100 detectors is used for further analysis of the complete data set”.



“Figure 2: Settings of spatial graph: (a) 10-min neighborhood graph; (b) percentage of nodes achievable within specific travel time (min);

For a 30 second temporary aggregation, the traffic flow volume and information is analyzed for the original dataset. The road based aggregated values of lane detectors is the data preprocessing routine that is executed. In case of traffic forecasting over a short period of time, a time interval of 5 minutes is used for aggregation of the values. In the traffic flow periodical patterns, the nodes and the time interval of five minutes for a period of 30 weeks are used for calculation of median traffic values. The dataset is estimated for a period of 10 weeks along with the flow values and the periodical patterns obtained are eliminated. The training and testing of the model is performed with the detrued time series that is obtained. The static

background scene removal operation corresponds to video prediction in regular traffic conditions for forecasting the deviations using these models. Inter quartile ranges that are time period and detector specific are termed as outliers. 0.01 is the selected threshold value which is minute and filters out only the wrong observations. The congested conditions and real traffic values are analyzed. The missed values are marked by the identified outliers. The missed value imputation is performed by utilization of linear interpolation. If the missed value by the detector is greater than four hours, it is excluded from the final data set. For computational reasons, random sampling of 100 detectors is performed from the complete data set”.

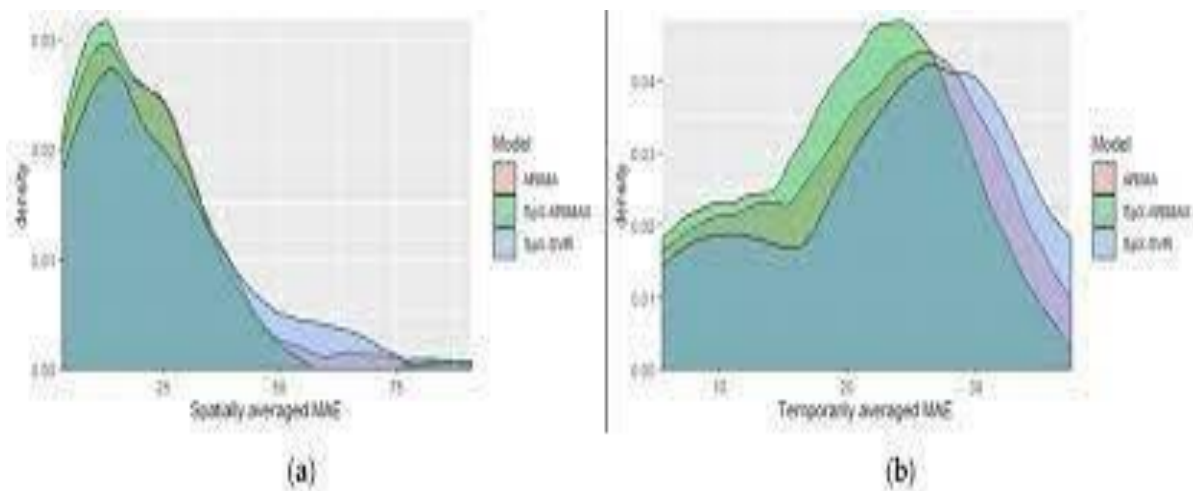


Figure 3: Spatially gathered Mean absolute error (MAE) cost concentration

“Language markdown routines obtained from online sources are executed for source codes that are publicly available in order to test the obtained experimental results for ensuring reproducibility. These repositories assist in experimentation specific for models with sampling, data preprocessing and so on for the downloaded data. Figure 2 represents the settings of spatial graph for 10-min neighborhood graph and the percentage of nodes achievable within specific travel time. This helps in demonstration of the performance comparison of the forecasting of baseline and transferred models. In comparison of

the modern models for spatiotemporal traffic forecasting, the efficiency of transferred models is high in short term forecasting. For further improvement of performance of forecasting, a good baseline can be set up by means of the non-spatial ARIMA model. However, from the nearby road segments, the spatial information cannot be utilized for forecasting horizons for a longer duration and the performance of the system is degraded. The performance stability can be improved with all the spatiotemporal model specifications other than the GCNN”.

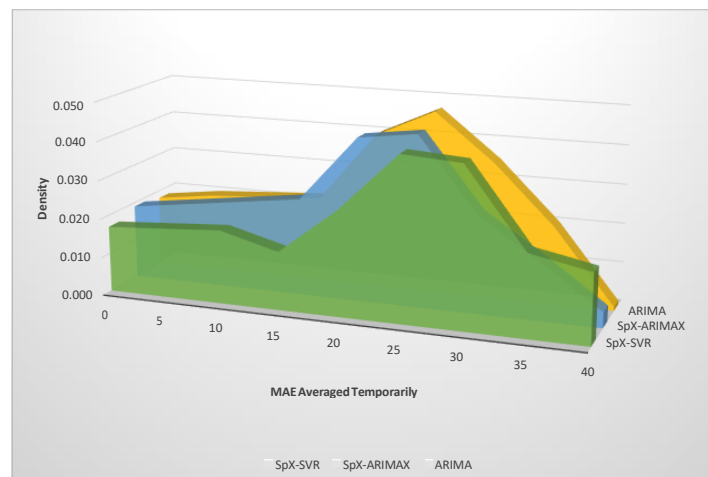


Figure 4: Temporal aggregated Mean absolute error (MAE) value density

“Multiple radii are used for using two spatial neighborhoods simultaneously based on the non-linearity’s. For all specifications of the model, a travel time of ten to thirty minutes is set by tuning the radii with the cross- validation. For longer forecasting horizons or gradual increase in the step size to a value of around five can be used for improving the time duration. Figure 3 and 4 represent the Mean absolute error (MAE) value density representation in terms of spatial as well as temporal distribution respectively. Smaller values of right tails are observed on implementation of the SpX-ARIMAX model for extreme forecasting

temporal distribution. In case of spatiotemporal traffic forecasting, video and image processing, this system is widely and efficiently implemented based on predefined spatial kernels.

distribution of errors which is the major benefit of the SpX-model. Implementation of the ARIMA model offered certain drawbacks in comparison with the SpX-ARIMAX model in terms of spatial distribution. For various time periods like peak hours and congested traffic conditions, day time and regular traffic as well as night time and free traffic flow, the SpX-ARIMAX model offers stable results in

4. Conclusion

- Modeling a traffic volume prediction model has proven to be a very difficult task. The highly nonlinear nature of traffic behavior necessitates the use of fuzzy logic for the modeling and development of traffic volume prediction model. Fuzzy logic can be a potential method dealing with structural and parametric uncertainties in the traffic Volume behavior human expert knowledge expressed by linguistic variables is a powerful tool enabling fuzzy models to deal with uncertainties and inaccuracies.
- In the current study membership functions, fuzzy control rules of fuzzy logic is used to make traffic prediction model. The above model has limitation that fuzzy control rules were designed on small data set.
- The result shows this approach is satisfactory and wide applicability in the area of Intelligent Transport System is still gaps to do improve the reliability and efficiency of the fuzzy logic model. In addition, new extensive data sets will be necessary to provide the basis for constructing a high quality model.
- Spatiotemporal and video prediction based transfer learning techniques for forecasting urban traffic is analyzed and promoted in this paper. Computational resources are optimized in terms of scientific aspects as research and intellectual resources as well as practical aspects. This leads to methodological enhancement through retrained models, model structures, algorithms and ideas transfer in an efficient manner.
- In a city wide traffic data, the video streams are analyzed and the data structures are studied for similarities. Forecasting techniques for urban traffic is performed by using the video prediction techniques in historical developments that can be applied in the modern states. The traffic forecasting domain makes use of convolution network and spatial filtering video processing techniques for studying the experimental part.
- The urban traffic data is predicted using the video tools as supported by the hypothesis made in this paper and city

wide traffic data set is used for the experimentation. Future work is focused on improving the datasets”, in addition to reducing the processing interval for cinematographic doling out.

References

- [1] Cheslow, M., Hatcher, S. G. and Patel, V. M.1992. An initial evaluation of alternative intelligent vehicle highway systems architecture. Rep. No. 92w0000063, MITRE Corporation, Bedford, Mass (2002).
- [2] Topuz, V.2002.Traffic Demand Prediction Using ANN Simulator, In: KES 2007/WIRN 2004, Part I, LNAI 4692, Springer-Verlag, Berlin Heidelberg, 864-870.
- [3] Okutani, I. and Stephanedes, Y. J.2004..Dynamic prediction of traffic volume through kalman filtering theory, *Transp. Res., Part B: Methodol*, 18(1), 1–11.
- [4] Xie, Y., Zhang, Y. and Ye, Z.2005. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. *Computer-Aided Civil and Infrastructure Engineering*, 22 326–334.
- [5] Ross, P.2005 .Exponential filtering of traffic data. *Transportation Research Record* 869, 43–49.
- [6] Smith, B. L., and Demetsky, M. J.2005. Short-Term Traffic Flow Prediction: Neural Network Approach, In *Transportation Research Record* 1453, TRB, National Research Council, Washington, D.C., 98–104.
- [7] Van Lint, J. W. C., Hoogendoorn , S. P., and Van Zuylen, H. J, *Accurate Freeway Travel Time Prediction with State-Space Neural Networks Under Missing Data*.2005. *Transportation Research Part C*, Vol. 13, No. 5-6 347–369.
- [8] Park, D., and Rilett , L. R.2005. Forecasting Freeway Link Travel Times with a Multilayer Feedforward Neural Network, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 14, No. 5, 357–367.
- [9] Yin, H. B., Wong, S. C., Xu, J. M., and Wong, C. K. 2005. Urban Traffic Flow Prediction Using a Fuzzy-Neural Approach. *Transportation Research Part C*, Vol. 10, No. 2, 85–98.
- [10] Vlahogianni, E. I., Karlaftis, M. G. and Golias, J. C.2005.Optimized and Meta-Optimized Neural Networksfor Short-Term Traffic Flow Prediction: A Genetic Approach, *Transportation Research Part C*, Vol. 13, No. 3, 211–234
- [11] Li , L., Lin ,W.H. and Liu ,H.2006. Type-2 fuzzy logic approach for short-term traffic forecasting. *Intelligent Transport Systems, IEE Proceedings*, 153(1), 33 – 40.
- [12] Xie, Y., Zhang, Y. and Ye, Z.2007. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. *Computer-Aided Civil and Infrastructure Engineering*, 22

- 326–334.
- [13] Su, H., Yu, S. 2007. Hybrid GA Based Online Support Vector Machine Model for Short-Term Traffic Flow Forecasting In: APPT LNCS 4847, Springer-Verlag Berlin Heidelberg, 743-752.
- [14] Zeng, D., Xu, J., Gu, J., Liu, L. and Xu, G. 2008. Short term traffic flow prediction using hybrid ARIMA and ANN models, Workshop on Power Electronics and Intelligent Transportation System, 621-625. Zheng, W., Lee, D. H. and Shi, Q. 2006. Short-term freeway traffic prediction: Bayesian combined neural network approach, Journal of Transportation Engineering, 114–121.
- [15] Kajackas, A., Vindašius and Š., Stanaitis. 2009., Vehicle Communication: Emergency Message Delay Distribution, Electronic and Electrical Engineering. – Kaunas: Technologija, 8(96), 33–38.
- [16] CHEN, B.P. and MA, Z. 2009. Short-term Traffic Flow Prediction Based on ANFIS, International Conference on Communication Software and Networks ICCSN, ISBN 978-0-7695-3522-7, IEEE, 791-793.
- [17] SADEK, A. W. 2009. Artificial Intelligence Applications in Transportation, In: Transportation Research E- C113. Artificial Intelligence in Transportation Information for Application. Transportation Research Board, ISSN 0097-8515, 1-6.
- [18] Molchanov, P., Tyree, S., Karras, T., Aila, T., & Kautz, J. (2016). Pruning convolutional neural networks for resource efficient transfer learning. arXiv preprint arXiv:1611.06440, 3.
- [19] Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S., & Zhang, G. (2015). Transfer learning using computational intelligence: A survey. Knowledge-Based Systems, 80, 14-23.
- [20] Chaabouni, S., Benois-Pineau, J., & Amar, C. B. (2016, September). Transfer learning with deep networks for saliency prediction in natural video. In 2016 IEEE International Conference on Image Processing (ICIP) (pp. 1604-1608). IEEE.
- [21] Jayashree, S. and D. A. Janeera. "Real-Time Fire Detection, Alerting and Suppression System using Live Video Surveillance." (2016).
- [22] Lucena, O., Junior, A., Moia, V., Souza, R., Valle, E., & Lotufo, R. (2017, July). Transfer learning using convolutional neural networks for face anti-spoofing. In International Conference Image Analysis and Recognition (pp. 27-34). Springer, Cham.
- [23] Kaya, H., Gürpınar, F., & Salah, A. A. (2017). Video-based emotion recognition in the wild using deep transfer learning and score fusion. Image and Vision Computing, 65, 66-75.
- [24] Diba, A., Fayyaz, M., Sharma, V., Karami, A. H., Arzani, M. M., Yousefzadeh, R., & Van Gool, L. (2017). Temporal 3d convnets: New architecture and transfer learning for video classification. arXiv preprint arXiv:1711.08200.
- [25] Su, Y. C., Chiu, T. H., Yeh, C. Y., Huang, H. F., & Hsu, W. H. (2017). Transfer learning for video recognition with scarce training data for deep convolutional neural network. arXiv preprint arXiv:1409.4127.
- [26] Qian, Y., Dong, J., Wang, W., & Tan, T. (2017, September). Learning and transferring representations for image steganalysis using convolutional neural network. In 2016 IEEE international conference on image processing (ICIP) (pp. 2752-2756). IEEE.
- [27] Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z., & Klette, R. (2018). Deep-anomaly: Fully convolutional Neural system for fast anomaly detection in crowded scenes. Computer Vision and Image Understanding, 172, 88-97.
- [28] Kumar, T. S. (2019). A Novel Method for HDR Video Encoding, Compression and Quality Evaluation. Journal of Innovative Image Processing (JIIP), 1(02), 71-80.
- [29] Manoharan, S. (2019). A Smart Image Processing Algorithm for Text Recognition Information Extraction and Vocalization for the Visually Challenged. Journal of Innovative Image Processing (JIIP), 1(01), 31-38.
- [30] D. A. Janeera and Sasipriya.S. "A Brain Computer Interface Based Patient Observation and Indoor Locating System with Capsule Network Algorithm" In International Conference on Image Processing and Capsule Networks (ICIPCN 2020), Springer, Thailand, 2020.
- [31] Ruth Anita Shirley D, Ranjani K, Gokulalakshmi Arunachalam, Janeera D.A., "Distributed Gardening System Using Object Recognition and Visual Servoing" In International Conference on Inventive Communication and Computational Technologies [ICICCT 2020], Springer, India, 2020.