DHANALAKSHMI SRINIVASAN
ENGINEERING COLLEGE
(AUTONOMOUS)
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai)
Re-Accredited with 'A' Grade by NAAC, Accredited by TCS.
Accredited by NBA (AERO, BME, CSE, ECE, EEE, IT & MECH)
PERAMBALUR - 621 212. TAMIL NADU

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

HYBRID WEB AUGMENTERS FOR FACILITATING ENVIRONMENTAL OPEN DATA CONSUMPTION WILD ROAD CONDITIONS

PRESENTED BY

810422104085 KIRUBASREE M
810422104094 MANJUPARKAVI S
810422104098 MITHUNAPRIYA M
810422104115 PADMAPRIYA R

UNDER THE GUIDANCE OF

Mrs.Francis Shamili
Assistant Professor,
Department of CSE,
DSEC(A),Perambalur

HYBRID WEB AUGMENTERS FOR FACILITATING ENVIRONMENTAL OPEN DATA CONSUMPTION WILD ROAD CONDITIONS

ABSTRACT

- Citizens increasingly rely on environmental open data to monitor and respond to local ecological conditions and Sustainable Development Goals (SDGs). However, heterogeneous datasets, inconsistent metadata, and dispersed data sources make discovery and meaningful consumption difficult for non-experts.
- This project proposes a model-driven development (MDD) approach to automatically generate web augmenters that enhance the accessibility and usability of environmental open data. A prototype implementation demonstrates automatic generation of a web augmenter for road traffic datasets. .
- The proposed model achieved 95.19% accuracy on TuSimple dataset. To evaluate the system with respect to unstructured road scenarios, we created an Indian lane dataset with 6149 labeled images from India driving dataset (IDD). We reported intersection over union of 0.31 on IDD, which was higher than other state-of-the art lane detection models.
- Finally, we undertook an ablation study to understand effectiveness of two parallel branches, i.e., encoder–decoder and dilated convolution branches and the proposed loss function. Evaluation with real-world datasets shows that the approach reduces the time required to discover and interpret environmental data while improving comprehension and engagement for non-technical users

Introduction

- Environmental open data has emerged as a critical resource for monitoring ecological conditions, supporting policy-making, and promoting citizen engagement toward Sustainable Development Goals (SDGs). Governments and public institutions publish large volumes of data on air quality, water quality, green spaces, and other environmental indicators.
- While the availability of such datasets has increased, heterogeneity in data formats, metadata standards, and access mechanisms significantly limits their effective consumption by non-technical users. Citizens, researchers, and policymakers often face difficulties in discovering relevant datasets, integrating information from multiple sources, and visualizing it in a meaningful context.
- To address these challenges, web augmenters—interactive overlays or embedded widgets—can provide contextualized visualizations directly within users' browsing experience, enabling seamless exploration of environmental data. However, manually creating these augmenters for diverse datasets is time-consuming and requires programming expertise, making them inaccessible for many potential users

Literature Survey

S.NO	TITLE	TECHNIQUE	DRAWBACK
1	Web Augmentation: A systematic mapping study Inigo Aldalur – 2023	Systematic Mapping Study	• Limited to Controlled Environments • Complexity in Designing Universal Tools • Requires Continuous Updates Due to Web Changes.
2	Supporting Natural Language Interaction with the Web Marcos Baez	Integrated Web Architec	• Depends on Complex NLP Accuracy. • May Struggle With Diverse Web Structures. • Requires Significant System Integration
3	Species awareness days: Do people care or are we preaching to the choir? 2 Marcus A, H. Chua – 2021	Google Trends and Wikipedia Page View Analys	• Limited Impact on Many Species. • 2Resources May Be Inefficiently Used. • Success Depends on Message Quality

Literature Survey

S.NO	TITLE	TECHNIQUE	DRAWBACK
4	Model-based Generation of Web Application Programming Interfaces to Access Open Data Cesar Gonz ´ alez-Mora ´ ∗ , Irene Garrigos, Jose Zubcoff and ´ Jose-Norberto Mazon -2020	Model-Based Automatic API Generation	• Depends on Quality of Source Datasets. • May Require Technical Knowledge to Use APIs. • APIfication Process Can Be Complex for Large Catalogs
5	AnATLyzer: An Advanced IDE for ATL Model Transformations Jesús Sánchez Cuadrado Universidad de Murcia jesusc@um.es Esther Guerra Universidad Autónoma de Madrid esther.guerra@uam.es Juan de Lara Universidad Autónoma de Madrid juan.delara@uam.es - 2018	Constraint Solving	• High Computational Cost. • Limited to ATL Transformations. • Requires Eclipse Environment
6	Visualization tools for Open Government Data Alvaro Graves Tetherless World Constellation Department of Cognitive Science Rensselaer Polytechnic Institute gravea3@rpi.edu James Hendler Tetherless World Constellation Department of Cognitive Science Rensselaer Polytechnic Institute hendler@cs.rpi.edu – 2013	Data Visualization	• Users Still Require Some Technical Skills. • Data Collection and Processing Difficulties Remain. • Visualization Tools May Not Cover All Data Types

Literature Survey

S.NO	TITLE	TECHNIQUE	DRAWBACK
7	A survey of biodiversity informatics: Concepts, practices, and challenges L. M. Gadelha, "A survey of biodiversity informatics: Concepts, practices, and challenges," 2021,	Data Integration using Ontologies	• Comprehensive Understanding of Biodiversity • Support for Evidence-Based Conservation Decisions • Promotion of Open Data and Collaboration
8	T. Yamane and S. Kaneko, "Impact of raising awareness of sustainable development goals: A survey experiment eliciting stakeholder preferences for corporate behavior," Feb. 2021,	Conjoint Survey Experiment	• Demonstrates Positive Impact of Awareness on Sustainable Behavior • Provides Empirical Evidence Using a Large Representative Dataset • Highlights Role of Personal Characteristics in Shaping Preferences
9	R. Saner, L. Yiu, and M. Nguyen, "Monitoring the SDGs: Digital and social technologies to ensure citizen participation, inclusiveness and transparency," Develop. Policy Rev., vol. 38, no. 4, pp. 483–500, Jul. 2020	Micro- and Macro-Monitoring Approaches	• Promotes Inclusive and Participation-Based Monitoring • Encourages Institutional Learning and Accountability • Introduces Micro- and Macro-Monitoring Frameworks

EXISTING SYSTEM

- Challenging conditions and relatively limited processing capabilities of embedded processing devices used in vehicles are considered. The structures of lane markings may be corrupted for many reasons such as reflections on the roads, erosion of markings by time, shadow and weather condition.

- The existing system for lane detection in autonomous driving mainly relies on traditional vision-based methods, which use hand-crafted low-level features and are often limited by challenging road and weather conditions, such as reflections, eroded markings, shadows, or poor lighting.
- These systems struggle with accurate detection due to the corruption of lane markings and relatively low processing capabilities of embedded devices in vehicles. As a result, traditional approaches frequently underperform in complex scenarios and require substantial computational resources, making real-time deployment difficult on low-power platforms.

Disadvantages

- Limited Accuracy in Real-World Conditions
- Traditional vision-based lane-detection methods rely on hand-crafted low-level features
- High computational requirements

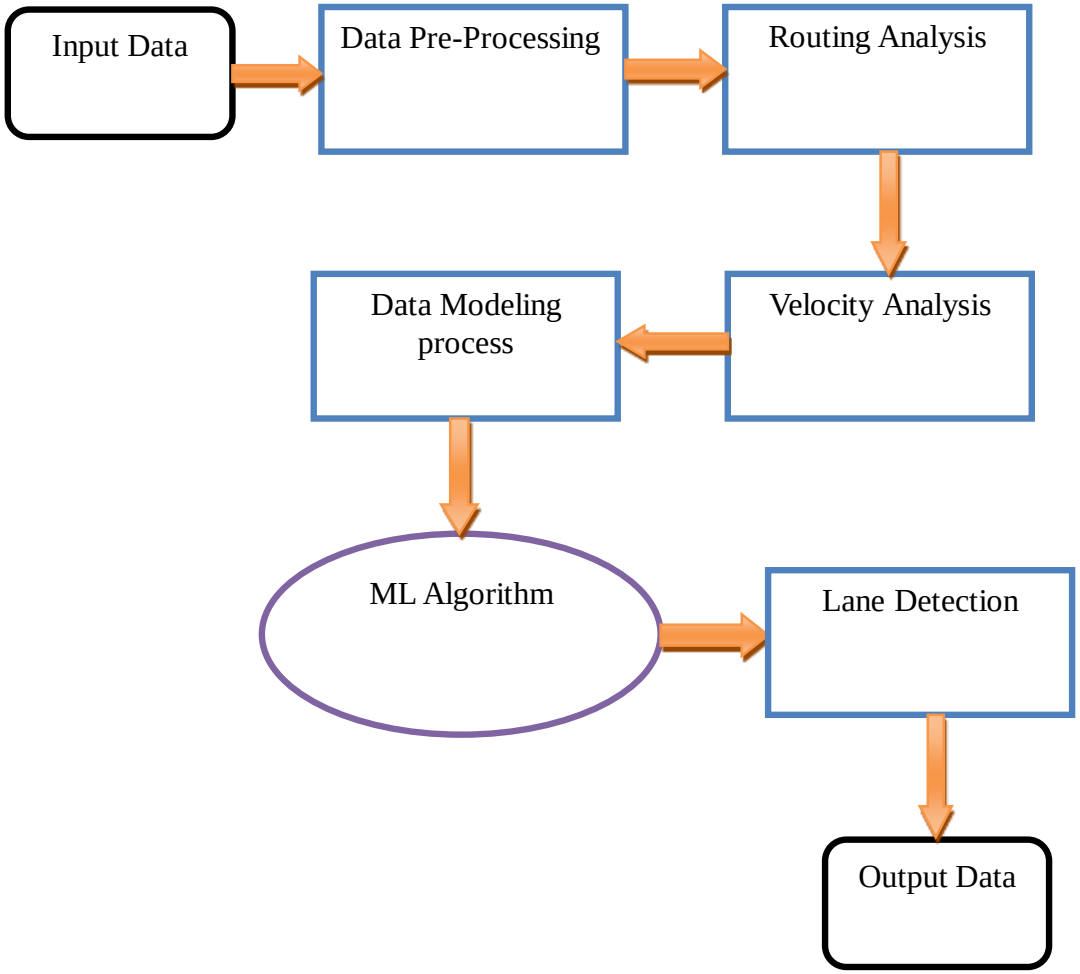
PROPOSED SYSTEM

- The main idea of the proposed method is to use one dimensional pixel intensity distributions to detect lane markings. Since the lane markings has a special pattern, this one dimensional distribution may provide sufficient discrimination for lane marking detection.
- However, in challenging road and weather conditions it becomes hard to differentiate the distribution for lane marking and non-lane marking areas. Architecture integrating the advantages of the encoder–decoder structure, SCNN-embedded single image feature extraction module, and ST-RNN module, is proposed.
- The proposed model architecture is the first attempt that tries to strengthen both spatial relation feature extraction in every single image frame and ST correlation together with dependencies among continuous image frames for lane detection. The proposed model can tackle lane detection in challenging scenes such as curves, dirty roads, serious vehicle occlusions, and so forth, and outperforms all the available state-of-the-art baseline models in most cases with a large margin.

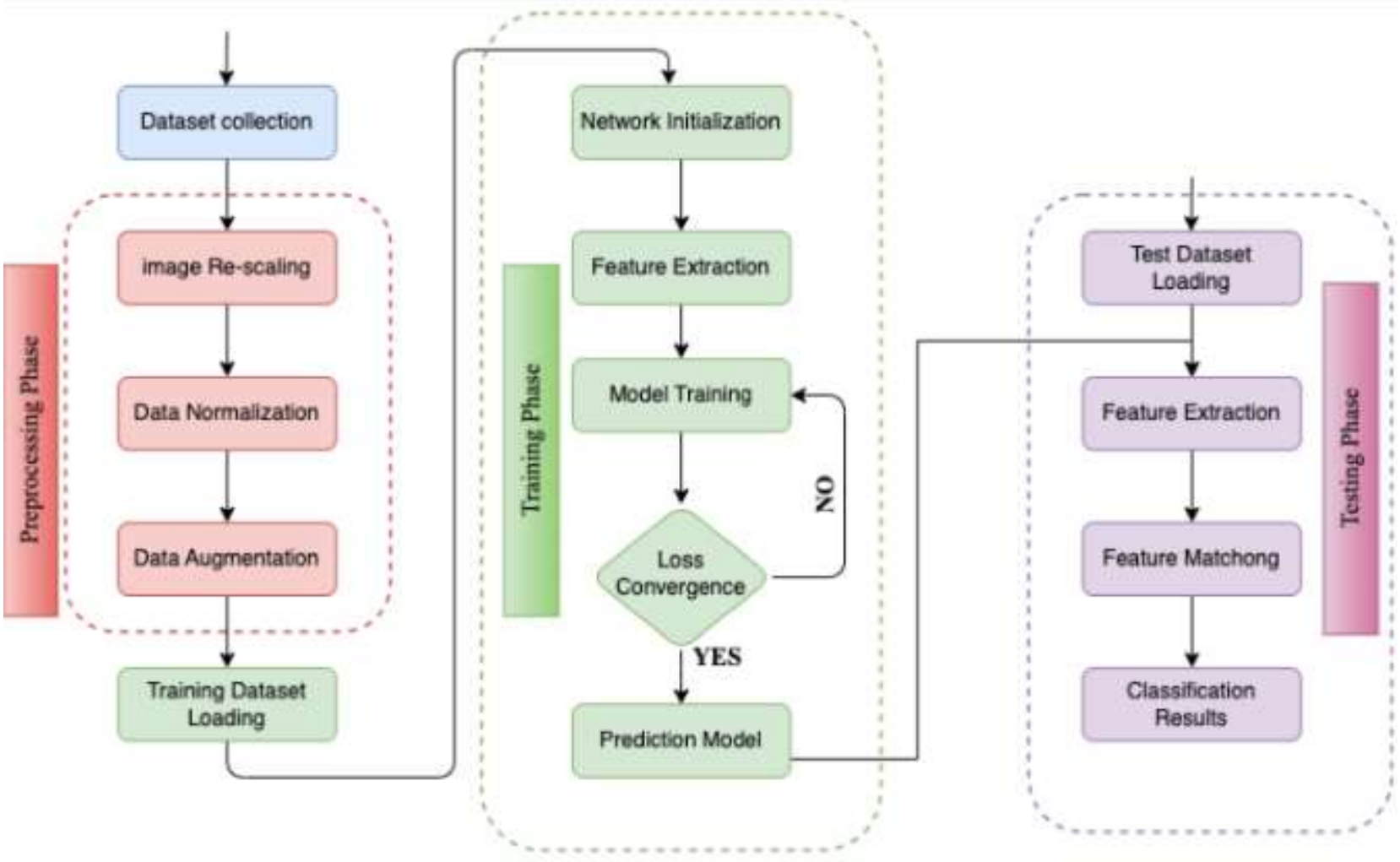
Advantages

- Information in continuous multiple frames performance
- High lane prediction accuracy
- Providing high accuracy with very low computational load compared to high-performance deep learning models.
- Suitable for low complexity embedded platforms with its very low computational load

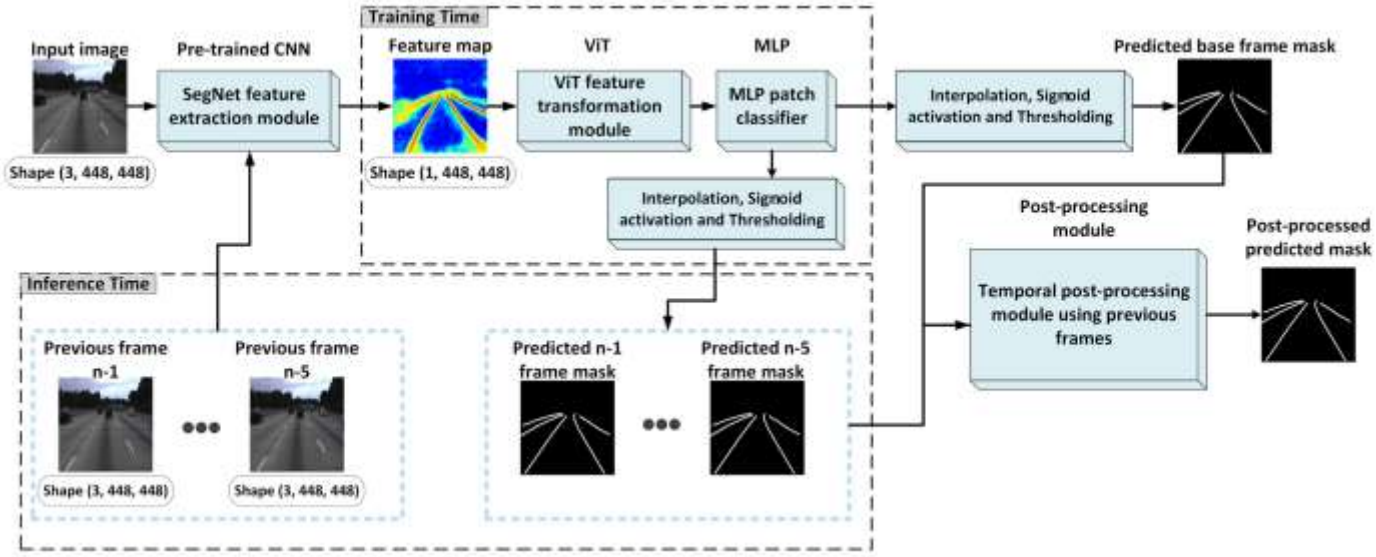
System Architecture



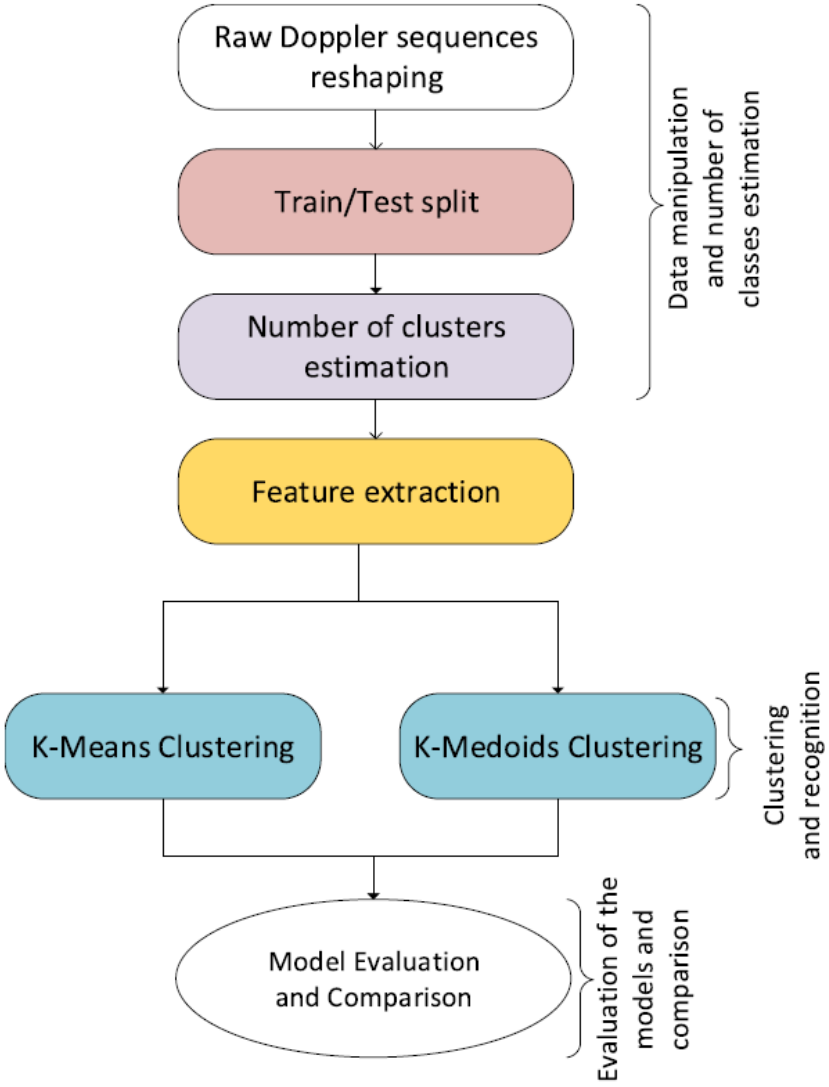
Data flow Diagram



Over all Process



Flow Diagram



SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS:

- Processor: Intel Core i5
- RAM: 8GB
- OS: Windows 10

SOFTWARE REQIUREMENTS:

- Language: Python
- Tool: Matlab 2013
- Front End: MatlabGUI (Graphical User Interface)
- Back End: CSV

MODULES

Convolutional Neural Networks Module

Image data analysis has been massively affected by the introduction of the CNN, a deep learning model architecture based on the mathematical linear operation of convolution. In order to perform image analysis with deep learning, image data are usually represented as a fixed uniform grid of parameters representing each pixel and their respective colors. The groundbreaking difference of a CNN network from an Artificial Neural Network (ANN) with fully-connected layers, is the way the connections between neurons of successive layers takes place.

Transformer Networks Module

In an effort to push the boundaries of the at the time state-of-the-art models for sequential data, such as Recurrent Neural Networks (RNNs) or Long-short term memory (LSTM) neural networks, the Transformer architecture was proposed (Vaswani et al., 2017). The Transformer architecture immediately offered increased parallelization capabilities, compared to the RNNs, as it relies solely on attention mechanisms to extract global dependencies between input and output (Vaswani et al., 2017).

Performance enhancement Module

In this stage, The preceding sections of the Related Work chapter describe various end-to-end complete approaches for lane detection. Nevertheless, a significant percentage of recent works focus on exploring different performance enhancement modules, striving for increased predictive accuracy or reduced model complexity. This section provides a detailed summary of these recently proposed modules. An interesting alternative to standard convolutional layers for the CNN architectures, was proposed).

Dataset and pre-processing Module

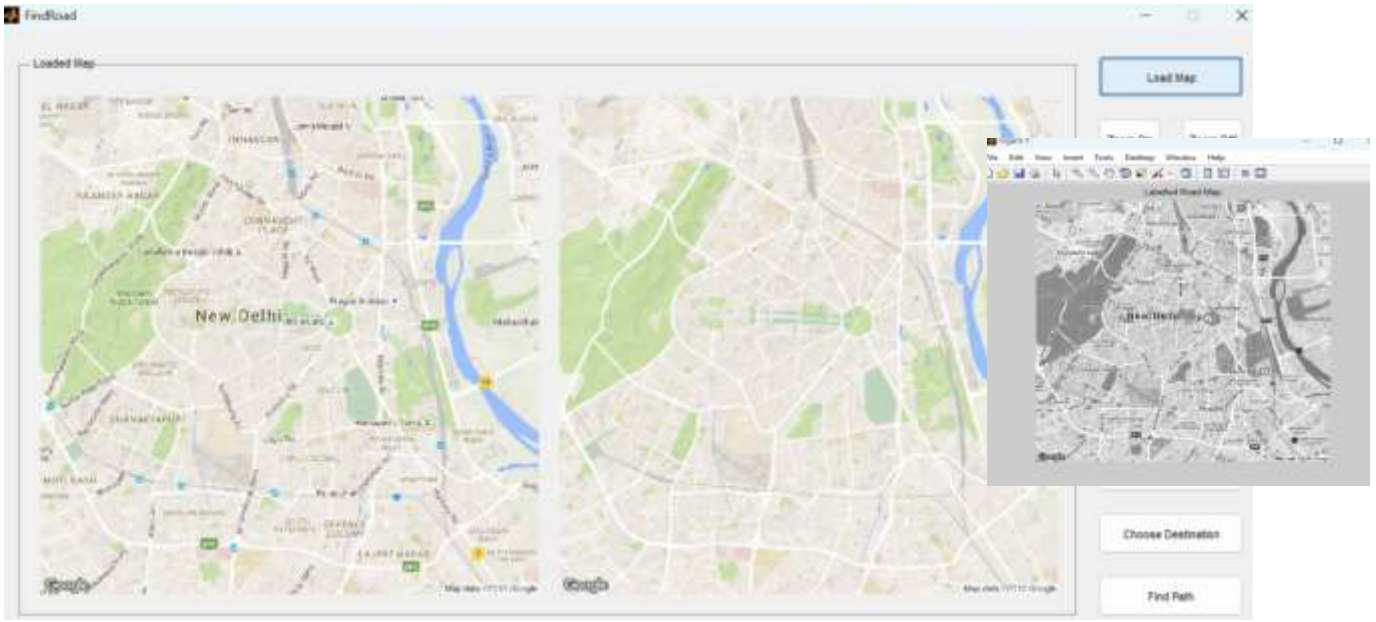
- For our experimentation we have chosen to use the image dataset.As previously mentioned in the Scientific Method section, TUSimple is one of the most widely used datasets for the lane detection task. The full dataset consists of 6408 images, divided in approximately 320 onesecond long clips of 20 frames each.

- The original resolution for each frame image of the clip is 1280x720 pixels. For every clip, only the 20th frame comes with annotated lane markings for a non-standard number of lanes (2 to 5 lanes). According to the creators of the dataset, the clips were recorded under good or mediocre weather conditions, during different daytime and traffic conditions. Moreover, the dataset is already split into test and training sets by the creators. Out of the total 6408 clips, 3626 are given as the train and validation split whereas 2782 are provided for testing purposes.

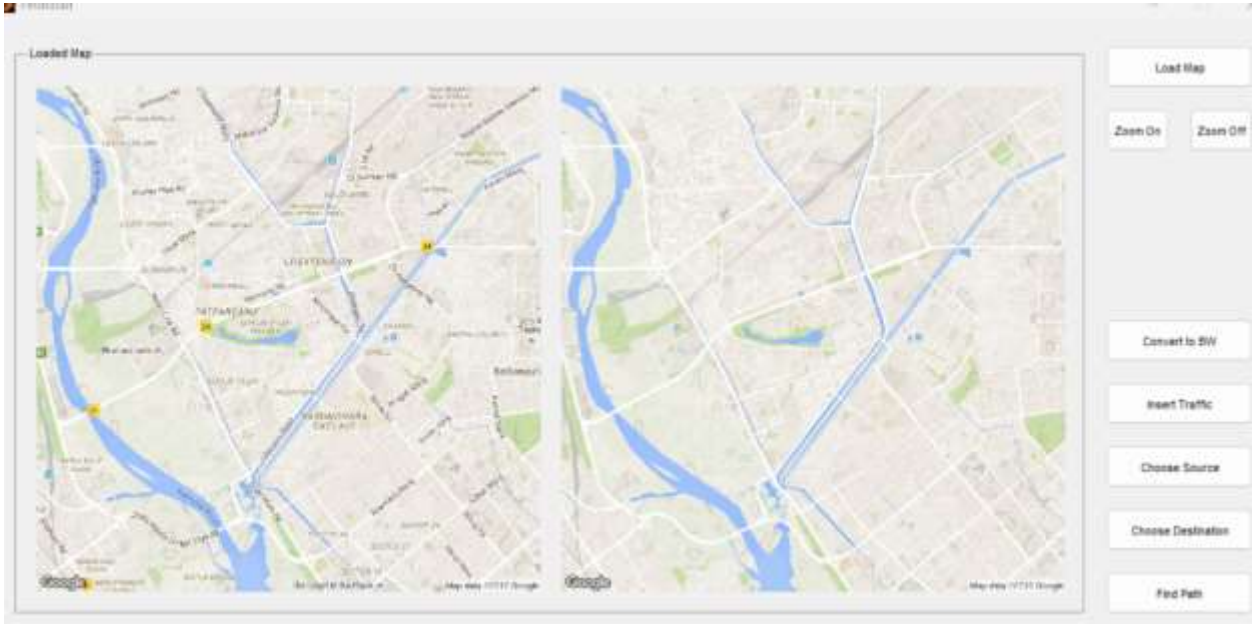
CNN backbone network Module

- After applying the pre-processing techniques on the TUSimple dataset, we incorporate the SegNet backbone network. SegNet is a CNN model designed for semantic segmentation of images, whose goal is to assign a class label to each pixel in the input image.
- As previously mentioned, it was firstly introduced. Their model consists of an encoder and a decoder, with skip connections between them. Typically, the input data is transformed into a higher-dimensional representation by the encoder block, sequentially and repeatedly from the previous layer while down-sampling in size.
- Once the input data has been processed by the encoder, the resulting highest-dimensional representation is sent to the decoder that performs the reverse process.

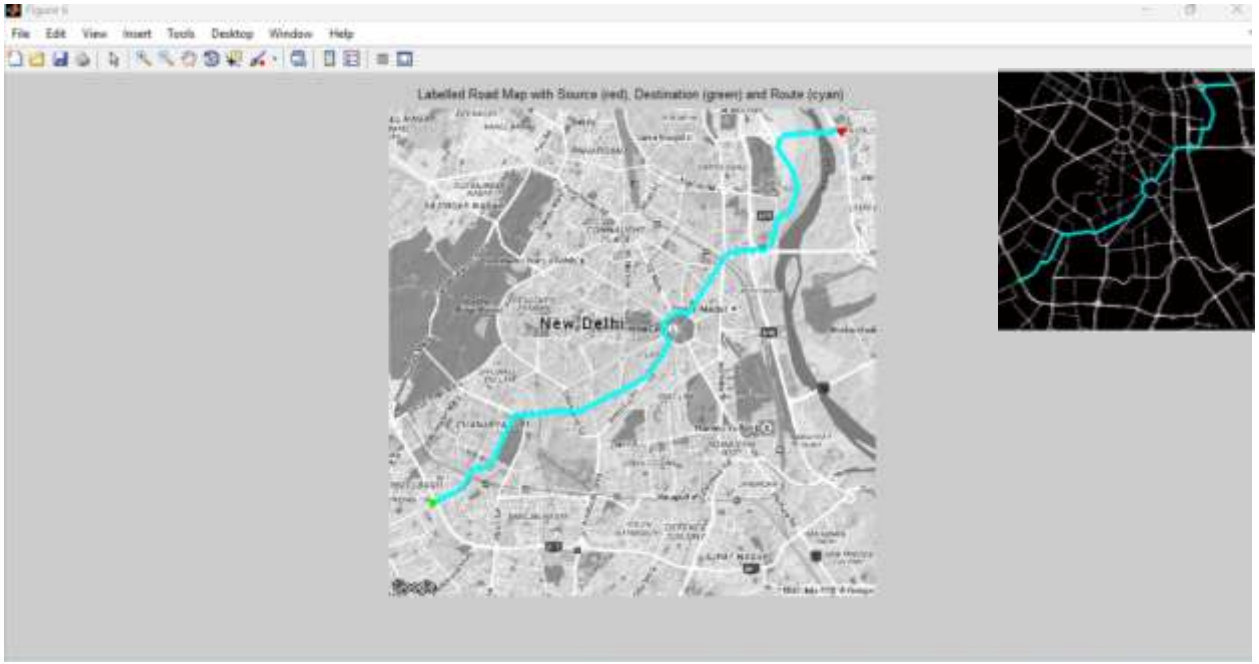
System Implementation



Input Analysis



Augmenters detection



CONCLUSION

- In conclusion, This project presented a Model-Driven Development (MDD) framework for generating web augmenters that facilitate the consumption of environmental open data by non-technical users.
- By leveraging high-level models of dataset types, spatial and temporal attributes, and interaction goals, the approach automatically produces data adapters, standardized data exports, and interactive map-based visualizations.
- The proposed system reduces manual development effort, improves consistency across heterogeneous datasets, and enhances accessibility for citizens, researchers, and policymakers.
- A prototype implementation using air and water quality datasets demonstrated the feasibility and effectiveness of the approach. Evaluation showed that generated web augmenters improved discoverability, comprehension, and engagement, validating the potential of MDD to bridge the gap between complex environmental datasets and end users.

FUTURE ENHANCEMENT

- Future enhancements for this project focus on advancing both the technical capability and user accessibility of the lane detection and web augmenter systems. Integrating real-time sensor data with advanced deep learning models can further improve accuracy on unstructured and adverse road conditions.
- Incorporating semantic enrichment and linked data standards will facilitate interoperability with various open-data platforms, making the system more robust and versatile.
- User interfaces can be enhanced to provide interactive visualization and dynamic customization options, empowering both technical and non-technical users. Crowdsourcing feedback mechanisms may be introduced to continually refine detection accuracy and dataset quality based on real-world usage.
- Expansion to support additional environmental datasets, such as air and water quality, would broaden the system’s societal impact. Optimization for low-power edge devices can enable wider deployment in cost-sensitive and resource-constrained environments.

REFERENCES

- I. Aldalur, “Web augmentation: A systematic mapping study,” Sci.Comput. Program., vol. 232, Jan. 2024, Art. no. 103045.
- F. Artigas and S. Ae Chun, “Visual analytics for open governmentdata,” in Proc. 14th Annu. Int. Conf. Digit. Government Res., Jun. 2013,pp. 298–299.
- M. Baez, C. Cappiello, C. M. Cutrupi, M. Matera, I. Possaghi, E. Pucci,G. Spadone, and A. Pasquale, “Supporting natural language interactionwith the web,” in Proc. Int. Conf. Web Eng. Cham, Switzerland: Springer,2022, pp. 383–390.
- J. P. Bigham and R. E. Ladner, “Accessmonkey: A collaborative scriptingframework for web users and developers,” in Proc. Int. Cross-DisciplinaryConf. Web Accessibility (W4A), 2007, pp. 25–34.
- M. Brambilla, J. Cabot, and M. Wimmer, Model-Driven SoftwareEngineering in Practice, vol. 1. San Rafael, CA, USA: Morgan & Claypool, Sep. 2012.