PRECISION LANE DETECTION WITH SEGNET: *MULTI-SCALE FEATURES* AND *ATTENTION*

22BAI1017 ROHITH ANAND
22BAI1095 SHIREEN ALIZA ALI
22BAI1192 SANYAM JINDAL
22BAI1217 SHREYA GOYAL

TABLE OF CONTENTS

01 OBJECTIVE

02 BASE PAPER

03 ABSTRACT

04 PROPOSED ARCHITECTURE

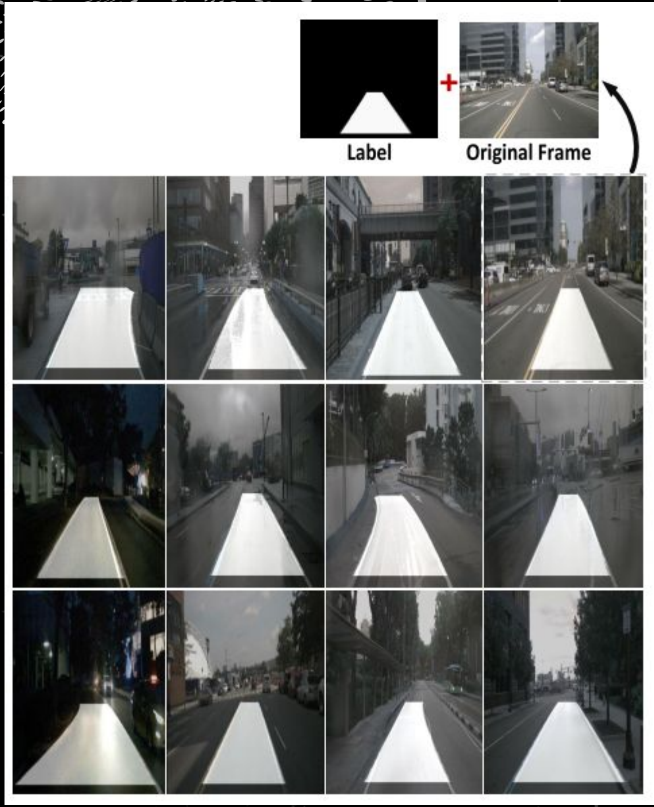

Fig. 01 Generated labels after Lane Segmentation

OBJECTIVE

The objective of this research is to develop a *highly accurate and computationally efficient lane detection* model tailored for autonomous vehicles, leveraging the strengths of both *U-Net and SegNet architectures*. This domain has the potential to revolutionize transportation, improve road safety, and reduce human error in driving. Autonomous vehicles rely heavily on advanced perception systems, where accurate lane detection is a critical component for safe and reliable navigation.

Our goal is to create a novel architecture that balances precision and real-time performance, addressing the challenges posed by varying lighting conditions, diverse road surfaces, and limited labeled data.

BASE PAPER

"A Deep Learning-Based Benchmarking Framework for Lane Segmentation in the Complex and Dynamic Road Scenes"

R. Yousri, M. A. Elattar and M. S. Darweesh, "A Deep Learning-Based Benchmarking Framework for Lane Segmentation in the Complex and Dynamic Road Scenes," in IEEE Access, vol. 9, pp. 117565-117580, 2021, doi: 10.1109/ACCESS.2021.3106377.

This research addresses the challenge of reliable lane detection in complex and dynamic road scenes, crucial for autonomous vehicles. Traditional computer vision techniques and deep learning methods are integrated to create a benchmarking framework for lane detection. An automatic segmentation algorithm using traditional methods generates weak labels from the nuScenes dataset. These labels are then used to train and benchmark five FCN-based architectures, with ResUNet++ showing the best performance in various challenging road conditions.

https://ieeexplore.ieee.org/document/9519709

GAPS IN THE RESEARCH:

The study uses a standard SegNet architecture without significant modifications. It lacks *exploration of custom modifications* for lane detection, such as incorporating spatial and channel attention mechanisms, multi-scale feature extraction , and boundary-aware loss functions. Additionally, it does not fully address *extreme environmental conditions*. These gaps highlight the need for developing tailored, robust models that integrate these custom modifications to improve accuracy and efficiency in diverse and dynamic road scenarios.

OUR FOCUS:

1. Custom Modifications: Implement spatial and channel **attention mechanisms** to enhance focus on lane regions.

2. **Multi-Scale Feature Extraction**: Integrate pyramid pooling and atrous convolution to capture context at various scales.

3. Using boundary-aware, dice, and focal **loss functions** for accurate lane edge detection and handling class imbalance.

4. Expanding **training dataset** to improve performance in extreme weather and lighting conditions.

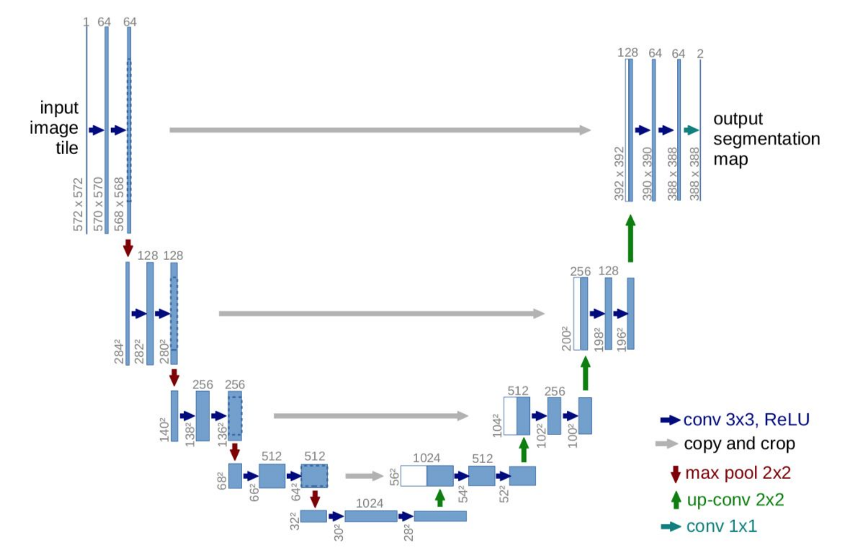

Fig. 02 Simple Convolutional Network (U-Net)

OUR ABSTRACT

This research aims to *enhance lane detection accuracy* in autonomous vehicles by addressing key limitations of existing models. We propose a *custom U-Net architecture with advanced modifications* to overcome gaps identified in prior work. Our focus includes integrating spatial and channel attention mechanisms to refine lane region detection, employing multi-scale feature extraction techniques such as pyramid pooling and atrous convolution for better contextual understanding, and implementing boundary-aware, dice, and focal loss functions to improve lane edge delineation and class imbalance handling. Additionally, we address performance in extreme environmental conditions. This approach aims to develop a robust and efficient lane detection system capable of operating effectively in diverse and challenging road scenarios.

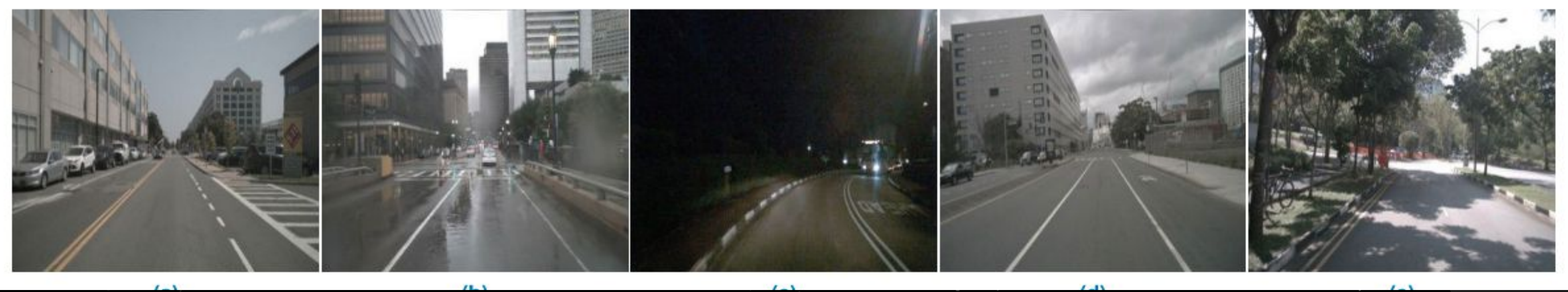

Fig. 03 Data samples (Dataset : https://nuscenes.org/)

PROPOSED ARCHITECTURE

SegNet is a deep learning architecture designed for *semantic segmentation* tasks. It features an **encoder-decoder network structure** with a focus on retaining spatial information and reducing computational overhead. Key components of SegNet include:

Encoder: Consists of convolutional layers followed by max-pooling layers that capture and compress feature maps(downsampling) while reducing spatial dimensions while retaining the most important information. SegNet stores the max-pooling indices (the positions of the maximum values) during this process.

Decoder: Utilizes the max-pooling indices from the encoder to upsample the feature maps. This process helps recover spatial resolution lost during encoding, ensuring accurate localization of features. This method helps in precisely reconstructing the segmentation map.

Pixel-wise Classification: The final layer performs pixel-wise classification to generate segmented output.

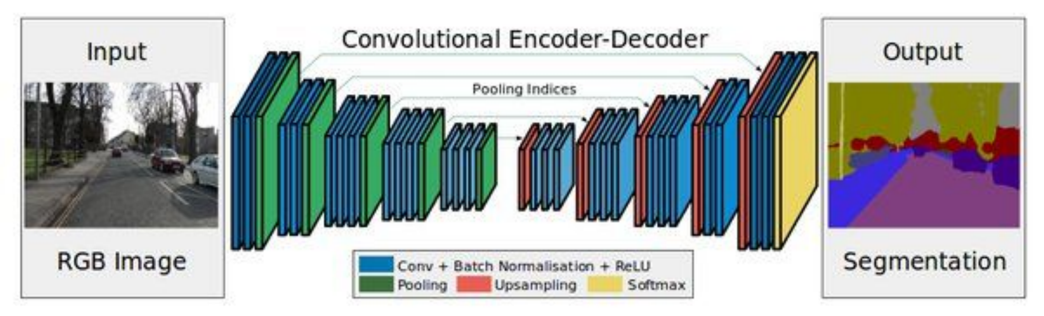

Fig. 04 SegNet

THANK YOU