

DRIVABLE AREA SEGMENTATION AND LANE DETECTION WITH TWINLITENET+

Hoàng Công Danh
240101039

Tóm tắt

- Lớp: CS2205.FEB2025
- Link Github của nhóm: [GitHub](#)
- Link YouTube video: [Youtube](#)



Hoàng Công Danh - 240101039

Giới thiệu

- Việc hiểu và nhận diện được môi trường xung quanh là yếu tố cực kỳ quan trọng và thiết yếu của hệ thống vận hành xe tự động. Quan trọng nhất là **Drivable Area Segmentation** và **Lane Detection**.
- Yêu cầu được đặt ra không chỉ có độ chính xác cao trong khả năng tính toán mà còn là khả năng xử lý trong thời gian thực, điều mà tiêu tốn khá nhiều tài nguyên để xử lý.
- Chính vì thế phiên bản TwinLiteNet+ được ra đời dựa trên phiên bản trước đó là TwinLiteNet với những thay đổi sau:
 - DESP (Depthwise Dilated Convolution) trong Encoder
 - PCAA (Partial Class Activation Attention) sau khi Encoder
 - Xử lý độc lập 2 Decoder để cho ra 2 bài toán yêu cầu.
 - Sử dụng 2 hàm Loss: Focal Loss và Tversky Loss

Mục tiêu

- Phân tích và tái hiện kiến trúc TwinLiteNet+ từ bài báo gốc.
- Huấn luyện mô hình trên tập dữ liệu BDD100K.
- Triển khai và đánh giá suy luận thời gian thực trên dữ liệu paper.
- Mở rộng đề tài với các hướng sau:
 - Tích hợp thêm cảm biến (depth, optical flow) để tăng độ chính xác.
 - Dùng nhãn giả từ mô hình lớn (SegFormer, DeepLabV3+) để hỗ trợ học đặc trưng.
 - Áp dụng attention nhẹ hơn hoặc MobileViT để nâng hiệu quả phân đoạn làn.
 - Chuyển mô hình sang TensorRT/ONNX để tăng tốc suy luận thực tế.

Nội dung và Phương pháp

Nội dung:

1. Nghiên cứu 2 bài toán trong mô hình xe tự hành: Drivable Area Segmentation và Lane Detection.
2. Phân tích về kiến trúc của TwinLiteNet+: Encoder, Decoder, PCAA.
3. Tiền xử lý dữ liệu BDD100K với 100k frames và gán nhãn cho 10 task
4. Huấn luyện mô hình và áp dụng 2 hàm Focal Loss, Tversky Loss.
5. Đánh giá mô hình bằng các chỉ số: mIoU cho Drivable Area Segmentation, IoU cho Lane detection và Latency của mô hình.

Nội dung và Phương pháp

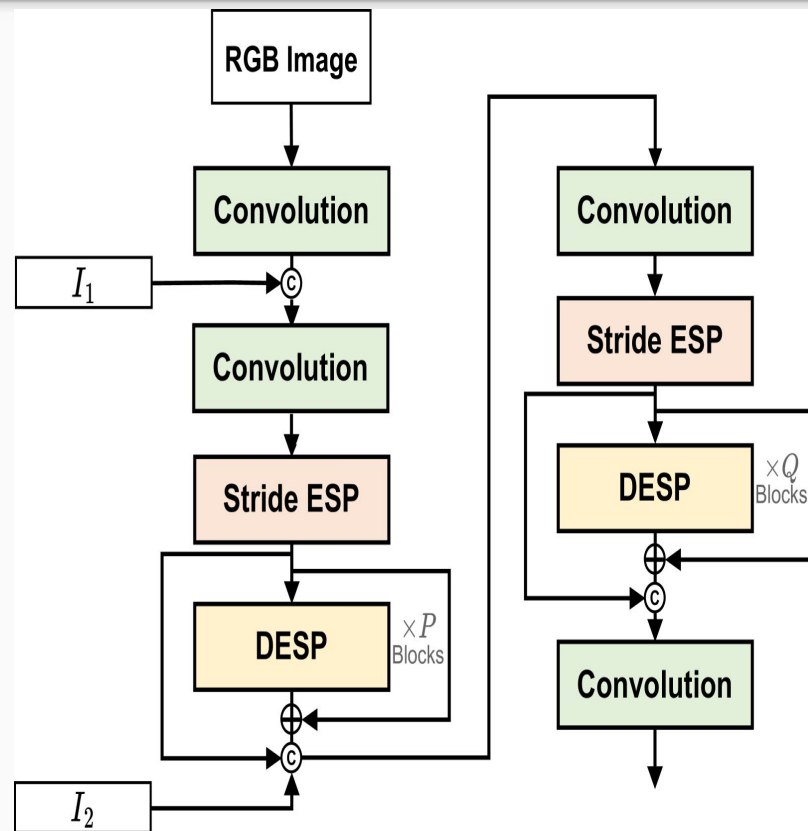
Phương pháp

1. Phân rã bài toán

- Phân đoạn vùng có thể lái (Drivable Area Segmentation).
- Phát hiện làn đường (Lane Detection).

2. Trích xuất đặc trưng ảnh đầu vào dùng:

- Encoder (DESP – Depthwise Dilated Separable Convolution)
- PCAA module (Partial Class Activation Attention)



Nội dung và Phương pháp

Phương pháp

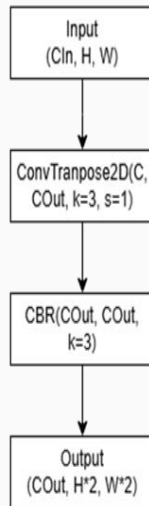
3. Decoder

- Sử dụng các **UpConvBlock** (CBR + ConvTranspose2D).
- Kết hợp **skip connection** giúp khôi phục độ phân giải đầu ra.

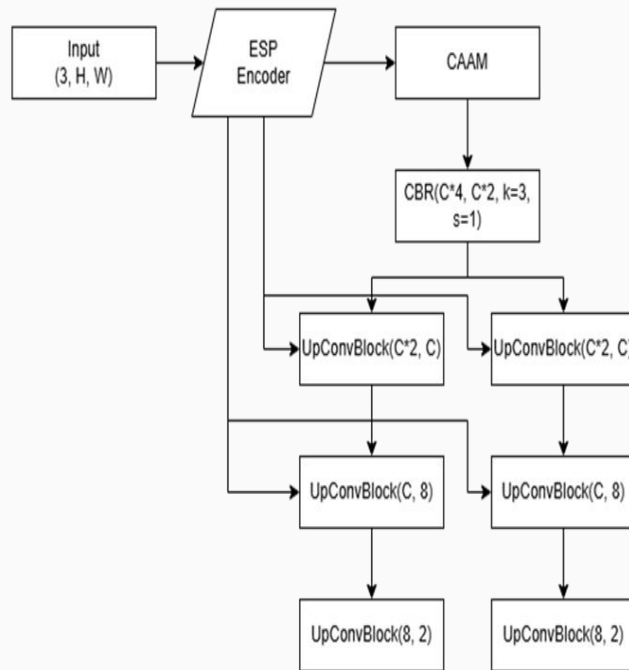
4. Loss function

- Sử dụng tổ hợp **Focal Loss** và **Tversky Loss** để tối ưu phân đoạn trong điều kiện mất cân bằng lớp

UpConvBlock

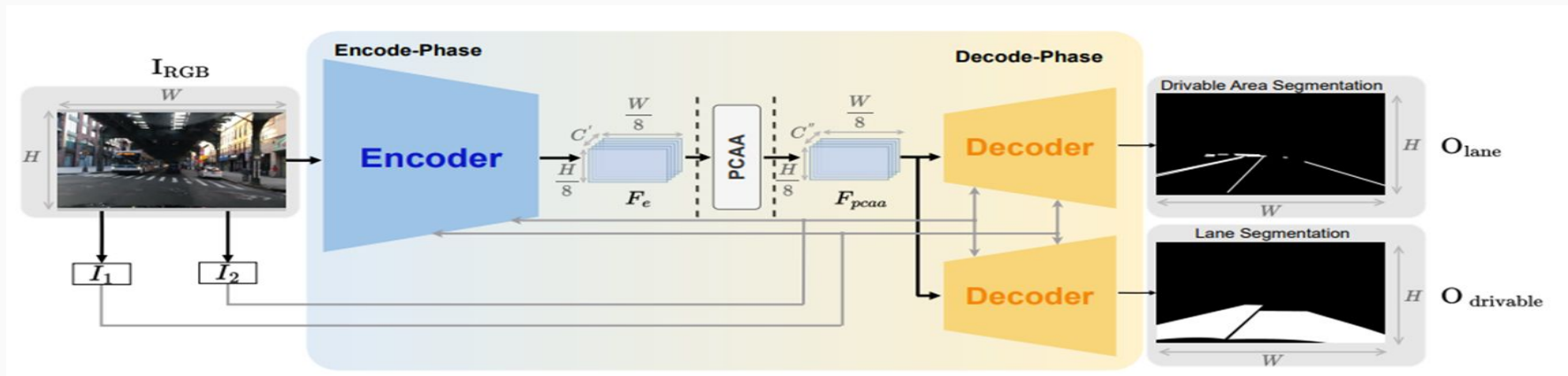


Model



Nội dung và Phương pháp

Phương pháp



Tổng quát lại, **Model TwinLiteNet+** sẽ cấu thành từ **ESP Encoder**, **PCAA** và **Decoder**. Ảnh đầu vào sẽ được đưa qua **ESP Encoder** để downsample xuống 8 lần sau đó **feature maps** này sẽ được đưa qua **khối PCAA** để tập trung vào các class quan trọng mà mô hình đang nhắm đến là làn đường và lane line và sau đó **feature maps** này sẽ được đưa qua **Decoder** kết hợp với các ảnh đầu vào được **giảm chiều để upsample** về kích thước ảnh ban đầu về label của ảnh ban đầu bao gồm làn đường và lane line

Kết quả dự kiến

- Mong muốn $\geq 92\%$ mIoU (Drivable Area), $\geq 34\%$ IoU (Lane).
- Latency của mô hình nhanh.
- Có báo cáo so sánh chi tiết hiệu suất các phiên bản mô hình (Nano \rightarrow Large).
- Mô hình sau cải tiến kỳ vọng đạt độ chính xác cao hơn hoặc tốc độ suy luận tốt hơn.
- Bổ sung khả năng mở rộng cho các hệ thống perception thực tế.

	Paper(Epoch 100)		Our(Epoch 25)	
Model	Drivable Area	Lane	Drivable Area	Lane
	mIoU (%)	IoU (%)	mIoU (%)	IoU (%)
TwinLiteNet+Nano	87.3	23.3	87.0	23.0
TwinLiteNet+Small	90.6	29.3	89.0	27.0
TwinLiteNet+Medium	92.0	32.3	90.0	30.0

Tài liệu tham khảo

- [1] Che, Quang-Huy; Le, Duc-Tri; Pham, Minh-Quan; Nguyen, Vinh-Tiep; Lam, Duc-Khai. TwinLiteNet+: A Stronger Model for Real-time Drivable Area and Lane Segmentation.
- [2] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” 2018.
- [3] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, “Segformer: Simple and efficient design for semantic segmentation with transformers,” 2021.
- [4] D. Wu, M.-W. Liao, W.-T. Zhang, X.-G. Wang, X. Bai, W.-Q. Cheng, and W.-Y. Liu, “Yolop: You only look once for panoptic driving perception,” Machine Intelligence Research, pp. 1–13, 2022.
- [5] Q.-H. Che, D.-P. Nguyen, M.-Q. Pham, and D.-K. Lam, “Twinlitenet: An efficient and lightweight model for driveable area and lane segmentation in self-driving cars,” in 2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR), pp. 1–6, 2023.
- [6] S.-A. Liu, H. Xie, H. Xu, Y. Zhang, and Q. Tian, “Partial class activation attention for semantic segmentation,” in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16815–16824, 2022.