شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

بسم الله الرحمن الرحيم

MONA MAGHRABY

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

MONA MAGHRABY

شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

شبكة المعلومات الجامعية
@ ASUNET

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

MONA MAGHRABY

AIN SHAMS UNIVERSITY

FACULTY OF ENGINEERING

Computer and Systems Engineering

An Improved Semantic Segmentation for Autonomous Driving

A Thesis submitted in partial fulfilment of the requirements of

M.Sc. in Department of Computer and Systems Engineering

Electrical Engineering

by

Taha Mohamed Ahmed Ibrahim Emara

B.Sc. in Electronics & Communications Department

Electrical Engineering

High Institute of Engineering, El-Shrouk Academy, 2013

Supervised By

Prof. Dr. Hazem Mahmoud Abbas

Prof. Dr. Hossam Eldin Hassan Abdelmunim

Cairo, 2021

AIN SHAMS UNIVERSITY

FACULTY OF ENGINEERING

Computer and Systems Engineering

An Improved Semantic Segmentation for Autonomous Driving

by

Taha Mohamed Ahmed Ibrahim Emara

Examiners' Committee

Name and affiliation	**Signature**
Prof. Dr. Mohsen Abdelrazek Rashwan Electronics and Communications Engineering Faculty of Engineering, Cairo university.	....
Prof. Dr. Hazem Mahmoud Abbas Computer and Systems Engineering Faculty of Engineering, Ain Shams University.	....
Prof. Dr. Mahmoud Ibrahim Khalil Computer and Systems Engineering Faculty of Engineering, Ain Shams University.	....
Prof. Dr. Hossam Eldin Hassan Abdelmunim Computer and Systems Engineering Faculty of Engineering, Ain Shams University.	....

Date: February 2021

Statement

This thesis is submitted as a partial fulfillment of Master degreee in Department of Computer and Systems Engineering, Faculty of Engineering, Ain shams University. The author carried out the work included in this thesis, and no part of it has been submitted for a degree or a qualification at any other scientific entity.

Taha Mohamed Ahmed Ibrahim Emara

Signature

....

Date: February 2021

Researcher Data

Name: Taha Moahmed Emara

Date of Birth: 01/07/1991

Place of Birth: Damietta, Egypt

Last academic degree: B.Sc.

Field of specialization: Electronics & Communications Engineering

University issued the degree : High Institute of Engineering, El-Shrouk Academy

Date of issued degree : June 2013

Current job : Deep leanring Engineer at Cisco

Summary

The First step of autonomous car is based on visual scene understanding of the surrounding environment. This visual understanding entails identification and localization of surrounding objects. Developing a semantic image segmentation architecture, for segmenting the entire view into regions and assigning a semantic label to these regions, lies at the heart of this problem. This thesis proposes an effective and efficient semantic image segmentation model for autonomous driving.

In the last several years, semantic image segmentation likes other computer vision tasks as object detection and image classification, has seen considerable advancements due to the employment of deep learning architectures, especially convolutional neural networks CNN. Training such architectures to obtain a high level of accuracy requires a very complex model. Being the autonomous driving a critical real-time application, computationally efficient models are needed. Also, edge devices as mobile phones have a low capacity of computational power. Also, this requires a specific and efficient deep neural networks. Although there are many ways to design deep neural networks and the availability of efficient training hardware, still designing a high accuracy and computationally efficient models is very challenging.

This thesis focuses on providing efficient deep neural networks for semantic image segmentation at two concurrent levels.

Computationally Efficient Model: we designed lightweight neural networks for semantic segmentation by following up the encoder-decoder structure, employing lightweight efficient backbone networks, and designing lightweight efficient decoder module.

High Accuracy Model: while considering the computational cost of the proposed models in our mind, we also consider the performance efficiency in terms of accuracy by employing long and short residual connections and designing efficient module called Deeper Atrous Spatial Pyramid Pooling (DASPP) to capture the extracted features by the encoder section at multi-level context.

We evaluated our model on the standard dataset Cityscapes. Also, in our evaluation procedure, we evaluated our model on severe weather condition on the standard dataset

Foggy Cityscapes. A three variant of semantic segmentations model are proposed to provide multiple trade-offs between accuracy and computational efficiency. Our model LiteSeg-Mobilenet can achieve 161 frame per second (FPS) with mean intersection over union (mIOU) of 67.81% while the previous state-of-the-art ESPNet on the same hardware can achieve 144 FPS with mIOU of 60.3% on the standard Cityscapes test set.

Key words: Semantic Image Segmentation, Convolutional Neural Network, Fully Convolutional Network, Encoder-Decoder Architecture

Acknowledgment

Taha Emara

Computer and Systems Engineering

Faculty of Engineering

Ain Shams University

Cairo, Egypt

February 2021

First, I am so grateful to Allah Almighty for his blessing and grace in giving me the ability to finish this work.

I want to express my appreciation and gratitude to Prof. Dr. Hossam Eldin Hassan for his guidance and continuous support over the years during the research and the writing stages. He has been supportive from the first day I began working on my research till the end. His insightful discussions and suggestions on the research helped me finish this work successfully.

I would like to thank Prof. Dr. Hazem Abbas for his guidance, support, insightful discussions, and inspirations to us. He introduced me to the field of machine learning and deep learning during his informative and inspired lectures.

Also, I would like to especially thank my parents for their unconditional love, care, and support. I also would like to thank my wife, for being my pillar of support and her understanding during this work.

Lastly, I would also like to thank the deep learning and the open-source communities for providing us tools, frameworks, and tutorials that made many things easy to understand and implement.

Contents

List of Figures