

Preface

The field of computer vision has witnessed a rapid advancement in recent years, owing to the emergence of deep learning techniques. Deep learning has revolutionized the way we process and analyse visual data, allowing us to extract valuable information from images and videos with unprecedented accuracy and speed. One of the most exciting applications of deep learning in computer vision is text detection in natural scene images. Text detection is a critical task in many applications, such as document retrieval, license plate recognition, and social media monitoring.

In this project, I focus on scene text detection using the EAST (Efficient and Accurate Scene Text Detector) algorithm, which is a deep learning-based approach for detecting text regions in natural scene images. The EAST algorithm has been shown to achieve state-of-the-art performance on various benchmark datasets, making it a promising approach for practical applications.

The EAST algorithm is based on a deep convolutional neural network that uses a multi-level feature fusion strategy to extract relevant features from the input image. The features are then used to generate a score map and a geometry map, which are combined to produce the final text detection results. The EAST algorithm is computationally efficient, which makes it suitable for real-time applications.

In this project, I present a detailed methodology for implementing the EAST algorithm for scene text detection, using the VGG (Visual Geometry Group) model as the backbone network.

This project aims to contribute to the advancement of text detection in natural scene images, and to demonstrate the potential of deep learning-based approaches for practical applications. The results of this project can be used as a reference for researchers and practitioners working in the field of computer vision and machine learning and can also serve as a basis for further research on text detection in complex scenes.