

Deep Learning Based Automatic Video Annotation Tool for Self-Driving Car

Referencias: N.S.Manikandan, K.Ganesan (2019). Deep Learning Based Automatic Video Annotation Tool for Self-Driving Car. [link paper](#)

El artículo propone cuatro áreas clave para los coches autónomos: detección de objetos, clasificación de objetos, detección de carriles y seguimiento de objetos. Se han evaluado y comparado diferentes algoritmos de aprendizaje profundo en cada uno de estos módulos para determinar cuál obtiene un mejor rendimiento en tiempo real. Se probaron YOLO y Retinanet-50 para la detección de objetos, VGG-19 y Resnet-50 para la clasificación de objetos, 'Finding Lane Line' de Udacity y LaneNet para la detección de carriles, y los algoritmos 'Object Detection and Tracking' de Udacity y Deep Sort para el seguimiento de objetos.

La motivación detrás de los coches autónomos es que según la WHO 1.25 millones de personas mueren cada año debido a accidentes de coches. Además, según recientes encuestas 69% afirman que los coches sin conductor son más seguros que los coches conducidos por humanos. Debido a que la anotación (*labelling*) de los datos para la computación visual es muy costosa se proponen métodos de anotación semi-automática y automática en el artículo.

Los modelos son alimentados con imágenes (*frames*) del vídeo con sus correspondientes etiquetas. Para la detección de carriles, se consideran hasta 6 carriles etiquetados: desconocido, -2, -1, 0, 1 y 2. Según la posición del vehículo se le asigna un número de carril a cada objeto. Para el seguimiento de objetos, cada objeto detectado tiene asignado un ID y se determina si el objeto está en movimiento, estacionado o detenido.

A la hora de los resultados, para la detección de objetos, Retinanet-50 obtuvo un mAP (*mean Average Precision*) de 48.3, mientras que YOLO 34.35. Además, la precisión media sobre vehículos (bus, coche, moto...) fue de 82.13% para Retinanet-50 y 60.12% para YOLO. Para la clasificación de objetos, se obtuvo que Resnet-50 tiene una precisión más alta (97%) en comparación con VGG-19 (86%). Para la detección de carriles, se encontró que el algoritmo de Udacity no era capaz de detectar carriles de carreteras asiáticas, por lo que se decidió por utilizar LaneNet (84% precisión). Para el seguimiento de objetos, en general Deep Sort tuvo un mejor rendimiento (74% precisión) que el algoritmo de Udacity.

En resumen, se han comparado algoritmos de detección de objetos, clasificación, identificación de carriles y seguimiento de objetos; obteniendo los mejores modelos: Retinanet-50, Resnet-50, LaneNet y Deep Sort, respectivamente. Además, se ha realizado un proceso de anotación semi-automática, ya que el proceso manual es demasiado costoso.