Allison DeCicco
September 10, 2019
Senior Design

Writing Assignment 1

Millions of cameras are installed in houses, businesses and on street corners, recording millions of images all day, every day. When a crime is committed, law enforcement searches for the person or object they are looking for by manually reviewing hundreds of hours of video surveillance footage. However, with our product, The Object Identifier™, there is a way to drastically reduce the amount of time someone spends doing this task. Using this product, the officer could teach the system what he wants it to identify and upload the footage. The system will analyze the footage and label every time the specific object that it was told to find appears in the footage. This will drastically reduce time spent doing this task and lead to solving crimes much faster. Tracking multiple objects through video is a vital issue in computer vision. It can be used in various video analysis scenarios, such as visual surveillance, sports analysis, robotic navigation, autonomous driving and medical visualization. The data that can be collected from object detection is a huge untapped potential for data analytics and could lead to advancements within many different fields.

The Object Identifier™ will automatically identify a wide variety of objects for the user. Most cameras now track an object, but they do not identify what the object is. No product exists, as of now, that can label tracked objects. Current products for object detection have trained classifiers for specific objects. It mainly includes vehicles because of the abundance of traffic cameras, but it has never been expanded to have any person, whether he is a trained computer scientist or not, to classify whatever they are interested in. The Object Identifier™ is intended to be a generalized algorithm that will train a neural network on any item they desire. The Object Identifier™ will overcome the requirement of having a professional to have to train a neural network. The objective is to simplify a complex system, so anyone can use it. Anyone using our product will be able to identify whatever object, person or image they want. It will also work on a small data set, which is not common. When training neural networks, a large data set is usually required. The impact that this product could have on so many different fields is what makes it unique. It can also progress self-driving cars by helping them identify objects in their path much faster, leading to a safer solution in self-driving cars. If researchers want to track a certain animal and figure out their daily patterns, they can identify the animal and train the network, and then the system can label the animal every time it appears on the camera. Labeling images is the next big step in computer vision because it will have major improvements in various fields.

Transitions?The user will submit a video or image and tell the program what it wants identified. The software will then run and display the results to the user. If the user isn't satisfied with how the software classified items, then the user can correct it by identifying what was marked correctly or incorrectly and what was missed. Once this is done, it will retrain the neural network and then shown the new results. If the user is satisfied with the results, they can save the classifier. If the user is not satisfied, it can continue to correct like previously stated until they are satisfied. The goal is to have every successful classifier saved in a database so that users can utilize different classifiers that have already been done previously. This will save a lot of time for researchers or any other client that is using our product.

Include transitions.The product will use the YOLO algorithm. In the paper *YOLOv3: An Incremental Improvement*, they explain the advancements in the classifier that will make this algorithm superior to other choices. Yolo uses neural networks to classify the image. It contains 54 convolutional layers. When it is trained on a set of images, it uses weights to classify the image. The YOLO algorithm solves the problem of having to use thousands of images to train a network. This algorithm is faster compared to others that are available, such as the Retina algorithm. There was a study done showing the difference and it is about four times faster than the Retina algorithm. One difficult task to tackle is when the image is not fully recognizable. The neural networks are trained on certain images, and if it is not recognizable it may be hard for the network to label the images correctly. Another challenge to overcome

is creating a user interface that will make it a user-friendly experience while also transferring the information smoothly between the model and the user. There is no mention of seeking funding?

Bibliography

Yolov3: An incremental improvement
J Redmon, A Farhadi
arXiv preprint arXiv:1804.02767