AvidAnnotation

1 Introduction

Autonomous vehicle or self-deriving vehicle market is expected to boom over the next few years [1]. The future self-driving car is expected to be occupied with many sensors as shown in Figure 1.

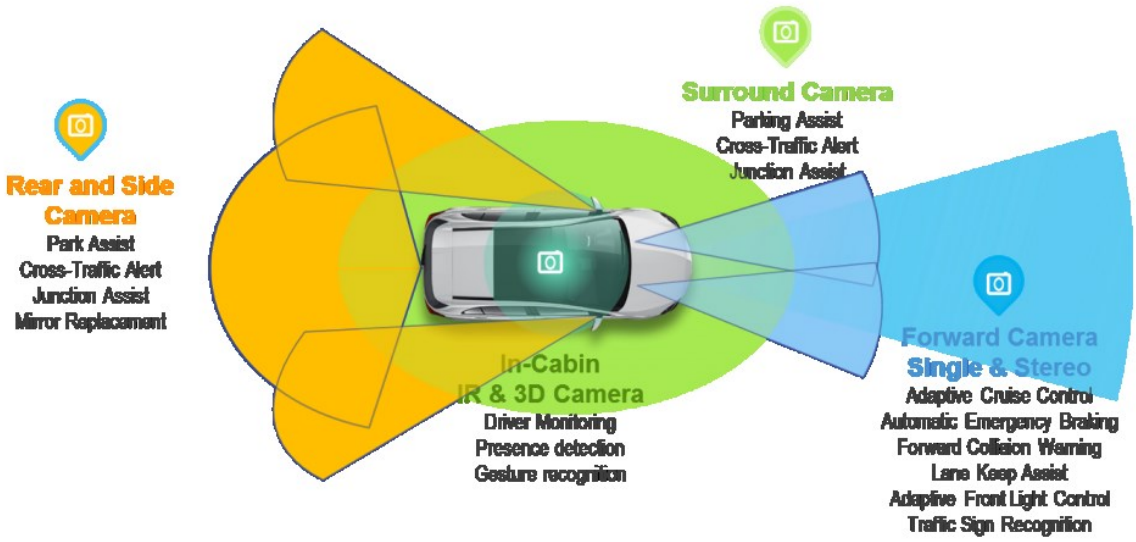

Figure 1 Autonomous Vehicle Sensors

[src: https://blog.nxp.com/automotive/radar-camera-and-lidar-for-autonomous-cars]

These sensors will be used jointly to help car driver make optimal driving decision. There are currently many solution in the market that aim to server that purpose. At present the focus in not on a fully autonomous driving vehicle, due to the complexity of the problem of detection and recognition at different environmental conditions, but the research and development is focused on advanced driver assistant solution (ADAS) while the computer and technology will assist the driver in making sound decision and in some cases protect the driver from making serious mistakes.

One of the required application in this market is an annotation tool that can extract useful information required by ADAS developer. Annotation tools are needed to generate potential challenges for ADAS solutions and validate the success/failure of existing ADAS algorithms. Currently, it is impossible to have such annotation tool as fully automated but in general, the tool is partially, semi, automated. Then manual labor are used to validate and complete the annotation for data before delivering the annotated results to ADAS developer or test laboratories.

AvidBeam® Technologies is specialized in scalable video processing, computer vision, and video analytics products. AvidBeam has joined the ADAS market and developed a tool for semi-automatic annotation (AvidAnnotation)[2]. AvidAnnotation is a scalable annotation product that

can support simultaneous users for faster processing. Vehicle data are first automatically annotated through the use of different algorithms for computer vision and image processing. Then, the annotation results are manually annotated and validated to produce 100% accurate annotation. The final results are then exported to clients. These results can later be searched, filtered, and exported for different usages such as testing new ADAS algorithms, training new detectors and recognizers for different object types, etc. AvidAnnotation is currently processes single camera (video) and LiDar data streams. AvidAnnotation is scalable ,able to process simultaneous data streams on more than one physical machine, using AvidBeam ATUN platform [3] which provides the necessary mechanism to create and launch several instances of AvidAnnotation.

ATUN™ is an open extensible platform that efficiently extracts business intelligence from video big-data sources. ATUN™ enables developers to easily scale their computer vision algorithms without having to worry about big-data tools. ATUN™ encapsulates multiple underlying engines and provides a means to manage and utilize available hardware while minimizing any administration overhead that could otherwise be needed to carry out the features offered by such a complex system.

ATUN™ is intended for processing video big-data (camera streams and/or video files). Each media input can be associated with auxiliary data such as LiDAR sensor data. ATUN™ has an integrated cloud processing engine for computer vision, image and video processing, and a UI engine with the main purpose to provide an optimized and easy-to-use video analytics solution.

Figure 2 illustrates a high level view of ATUN's architecture. A plugin is basically the core algorithm that is to be scaled and run on the incoming media and data. Plugins could be either user-defined or prebuilt by AvidBeam®. ATUN™ then scales processing by running the plugin(s) on video frame-level in parallel on available nodes. ATUN™ allows both end-users and developers to simply use the system via a rich web interface.

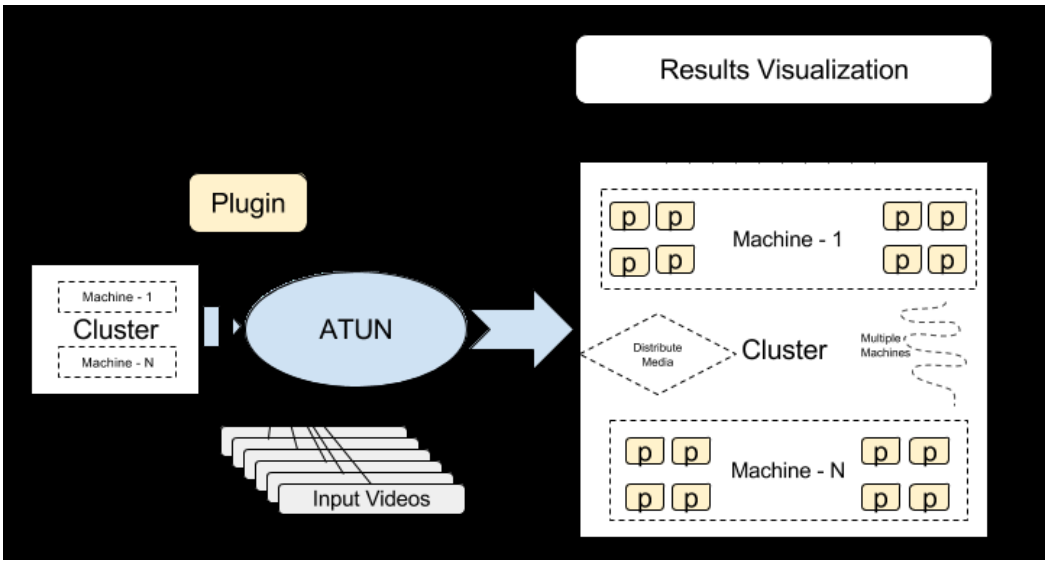

Figure 2 AvidAnnotation Solution

The following sections provide more details on how to use AvidAnnotation, its main features, and other important details.

1.1 AvidAnnotation™ Features

AvidAnnotation features are summarized as follows

1. Cloud based Solution: AvidAnnotation is a cloud scalable solution. It can be installed on any private or public cloud. It has been successfully tested on Microsoft Azure as well as Amazon AWS.
2. Scalable: AvidAnnotation is powered by AvidBeam ATUN™ platform which handles the scalability of AvidAnnotation. AvidAnnotation can process several inputs simultaneously based on the available hardware resources and is easily scalable.
3. Automatic Annotation: Tagging or labeling of objects in a video file/stream. Labeling includes object 3D position, object type.
4. Data synchronization: synchronize video, Lidar, or other sensor data based on their timestamp information.
5. Scene Management: display processed data or scene with video search capabilities using defined object tags.
6. Manual Annotation:
    - Display of video frames
    - Diplsay LiDar frames. Allows Zoom in/out, change viewing angle, or move using mouse and keyboard.
    - Tag objects in both video frame and Lidar frames using existing tags
    - Tag and adding new objects
    - Select any bounding box for any object, move, stretch, rotate the bounding box to fit the object dimensions.

7. Exportation of annotated data: annotated and selected data can be exported externally for different usage. Data are exported as JSON files.
8. Model training for new objects: client can select new objects from video/LiDar frames and re-train AvidAnnotation to recognize and detect them in future executions.

AvidAnnotation Web Interfaces

AvidAnnotation web interfaces display the results of individual job processing and allow client to search, modify these results. Currently AvidAnnotation has three main views

1. Annotation view (Scene management):
2. Manual annotation

Annotation View

This view display the results of automatically/manual annotated media. The results include objects found in each media. Client can filter scenes by object tags or input media source as shown in Figure 3

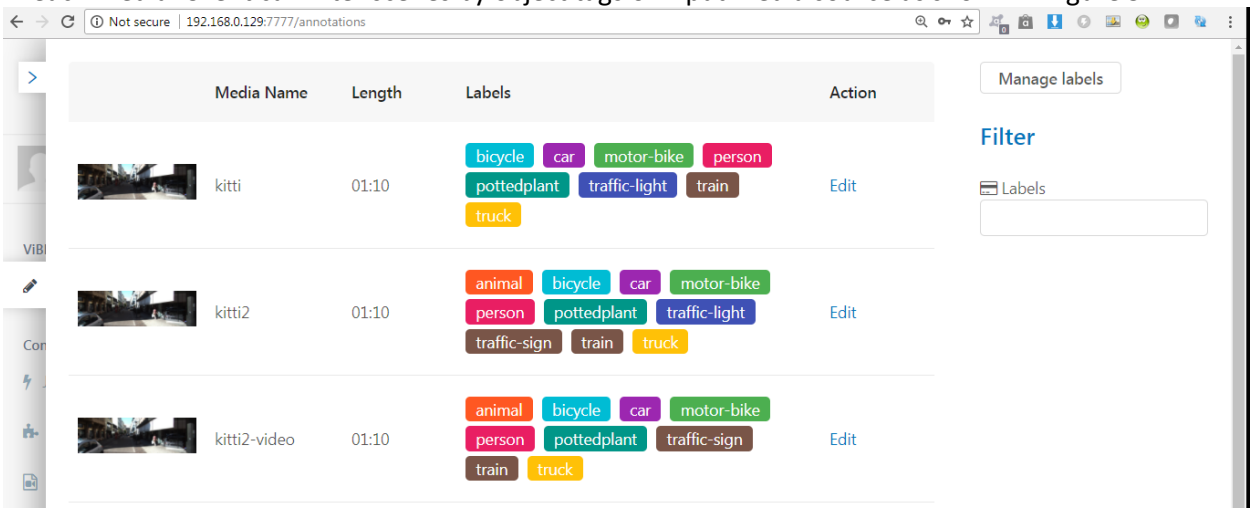

Figure 3 Annotation View

User can also select edit action to edit specific media for manual tagging and annotation.

Manual Annotation View

From manual annotation view as shown in Figure 4, user can

1. Display frame by frame
2. Select certain tagged object and edit its bounding box from either of the video or Lidar view. User can move, scale, or rotate the bounding box. Once the user is satisfied by the changes, he/she can select sync button to synchronize changes between video and Lidar views. Finally, the user can select save button to save new object position in the database.
3. Add or create new labels. User can label or tag new objects and stored these objects in the database as well (Figure 5).

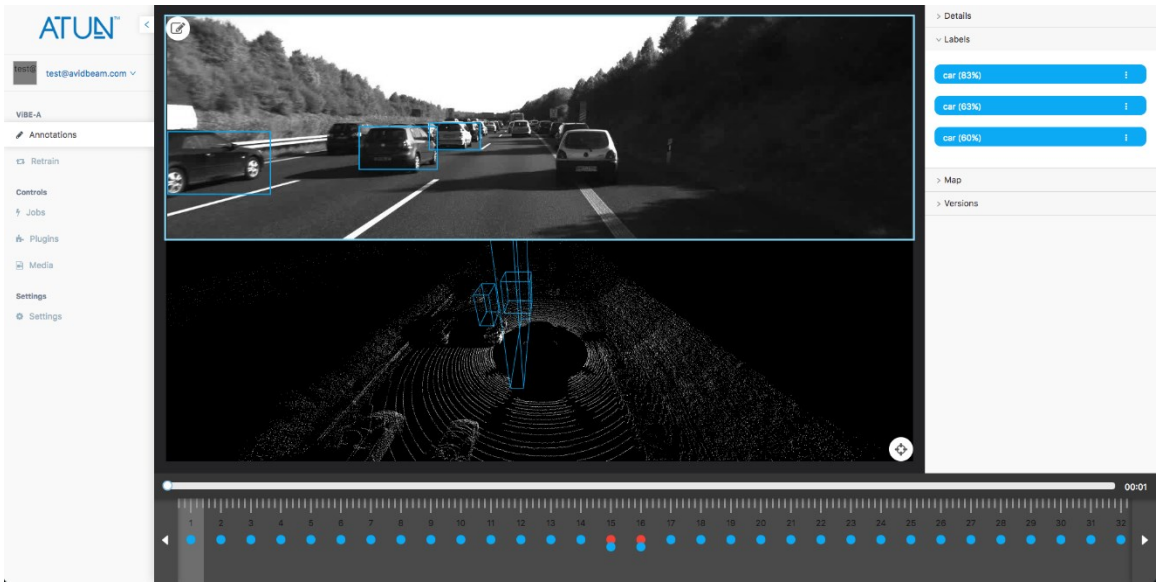

Figure 4 Annotated Objects on Video Frame and LiDAR Data.

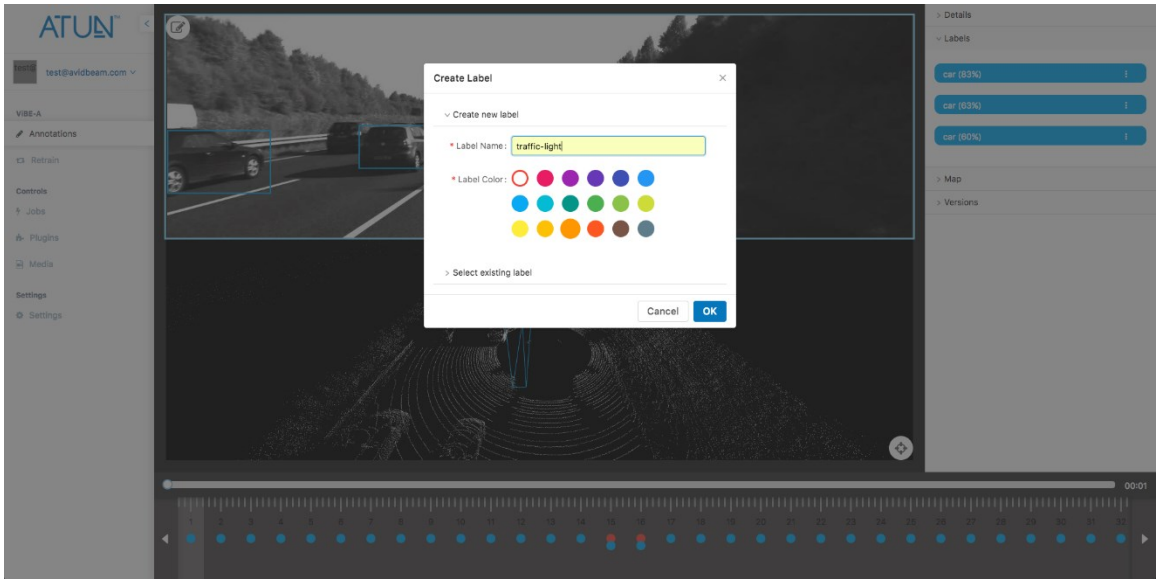

Figure 5 Adding annotation label in Frame.