1. Company Overview

Progeny Systems has developed a number of innovative machine learning and computer vision technologies and in turn used these to create state-of-the-art image/video analytics solutions. The primary "tools" or product lines emerging from these efforts to date include *TurboVideo* for multimedia (images and video, including metadata) decoding and ingest; *Vision* Annotation Framework for efficient labeling of image and video data; *Surveillance, Persistent Observation, and Target Recognition (SPOTR)* framework for image/video analytics; *Solution0* for non-cooperative face recognition and soft biometrics; the *ClusterUI* for organizing and querying visually similar objects; and *Blue Hotel*, a container-based microservices framework for flexible composition and deployment of customized video analytics solutions. These tools have been combined, adapted, modified, and extended to create solutions that are deployed in collaboration with government partners in the Department of Defense (DoD) and Intelligence Community (IC) organizations, as well as commercial customers, for applications ranging from intelligence, surveillance, and reconnaissance (ISR) to non-cooperative face recognition to manufacturing quality control. We have worked with imagery from satellite and aerial platforms to ground-based and maritime sensors, moving platforms, portable sensors, social media, and more.

2. Technology Overview

2.1. Vision Annotation Framework

Vision is an image/video annotation tool developed by Progeny Systems that has been used in numerous government programs, including Project Maven, to label large volumes of data. Progeny has been creating state-of-the-art computer vision solutions for the DoD and IC partners for over a decade, and part of this success can be attributed to and correlated with development of *Vision*. The system consists of three main components:

- Database back-end for storing and organizing both the raw data and annotations
- Web-based interface for managing and monitoring annotation tasks
- Web and Java application client for accessing and performing annotation tasks from a REST-based service

Each of these has evolved significantly since inception (around 10 years ago) and is still continuously updated and improved as new lessons are learned and tasks become known. From the start, *Vision* was designed *for* data scientists/computer vision researchers in order to quickly pivot and create testing, training, and validation data sets with programmatic ways to push and pull data as models continue to evolve. New task types can be developed quickly and currently there are bounding boxes, four points, and others for still images and video annotation is also available. *Vision* leverages our image analytics and machine learning algorithms to pre-seed annotations with bounding boxes so that annotators need only to confirm or group visually similar clusters together. An example can be seen in Figure 1. The user has several options for navigating this video, such as scrubbing, moving forward or backward by 1 or 10 frames, playing and pausing, or jumping to the start of a specific track. Additional options are available including "flag for review" (discussed earlier), "skip" (if there are no objects of interest visible), and "submit task" (when finished).

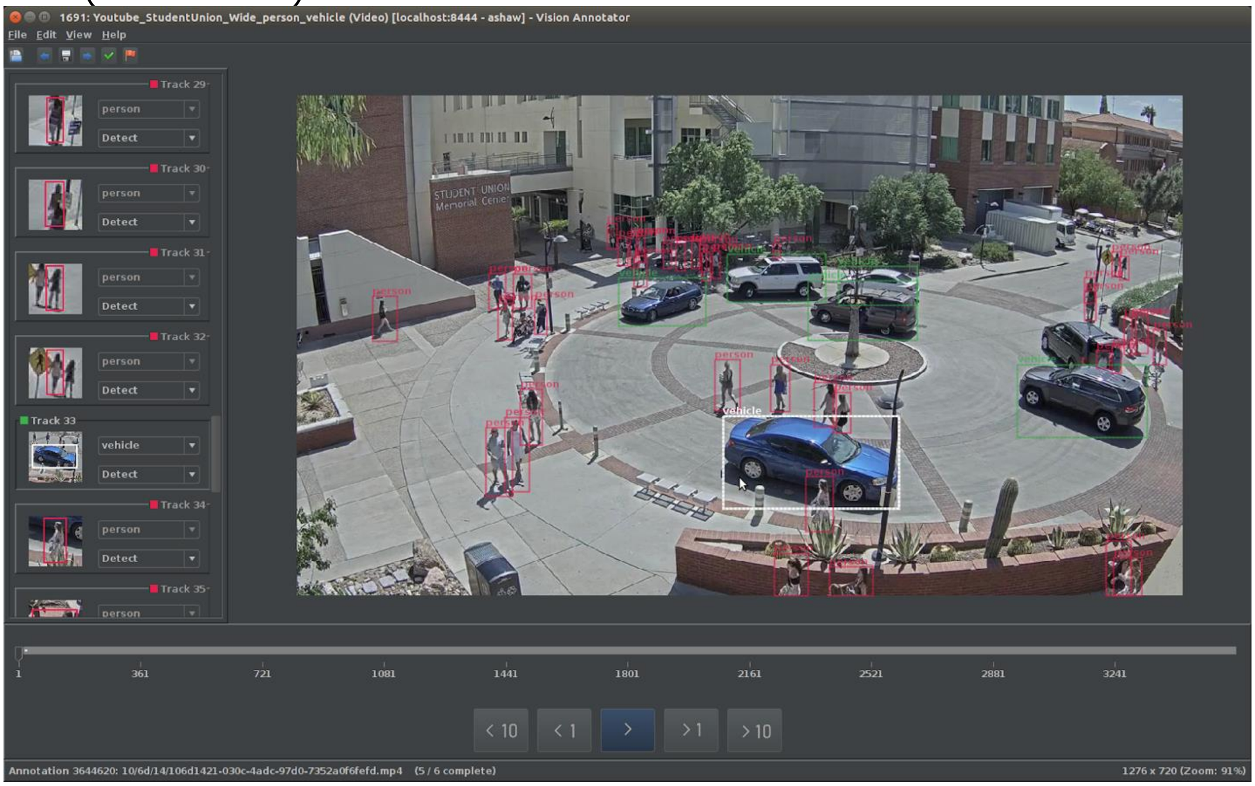

Figure 1: Vision screenshot showing person and vehicle video annotation

2.2. Face Recognition

Successful long-range identification of persons using facial recognition (>25m) has historically been a difficult problem due to challenges starting with distance to the subject through optics and sensor packages and ending with the underlying algorithms. There have been attempts to use face recognition in the past to screen people in vehicles without their knowledge, but these efforts have failed in part due to the collection methods and the face recognition algorithms used. Reflections in windows, lighting, and speed can all contribute to lower quality imagery or video, concealment (e.g., sunglasses) can prevent full face capture, and the quality and number of sensors can affect the fidelity and amount of data points that can be collected.

Solution0 is a software-based solution for face/person detection, matching / recognition, tracking, and soft biometric extraction. It can be applied in traditional verification or identification scenarios, but it truly excels in non-cooperative settings and works in conjunction with other our other software in our portfolio. *Solution0* can be deployed in various environments (desktop, server, cloud, etc.) and works with *TurboVideo* to access any data source (local storage, remote storage, live stream, etc.). It runs on commodity hardware but can leverage specialized components (e.g., GPUs) to improve processing speed if available.

2.3. Re-Identification and Keypoints

Throughout our work with DoD and IC customers, often there is not enough resolution (i.e., pixels) on the face to perform face recognition. However, there may be enough to detect a face and determine a person is present or faces may not be visible like those seen in Figure 2. In those scenarios it may only be possible to perform re-identification based on features available from a whole body until there is an opportunistic scenario wherein enough fidelity is available on the face to perform recognition.

Figure 2: Face and body detection and keypoints

Based on this experience we have invested heavily in researching and developing methods to perform person (as well as vehicle) re-identification based on visual similarity as well as leveraging keypoints. An example of person re-id is shown in Figure 3 within Progeny's *ClusterUI*.

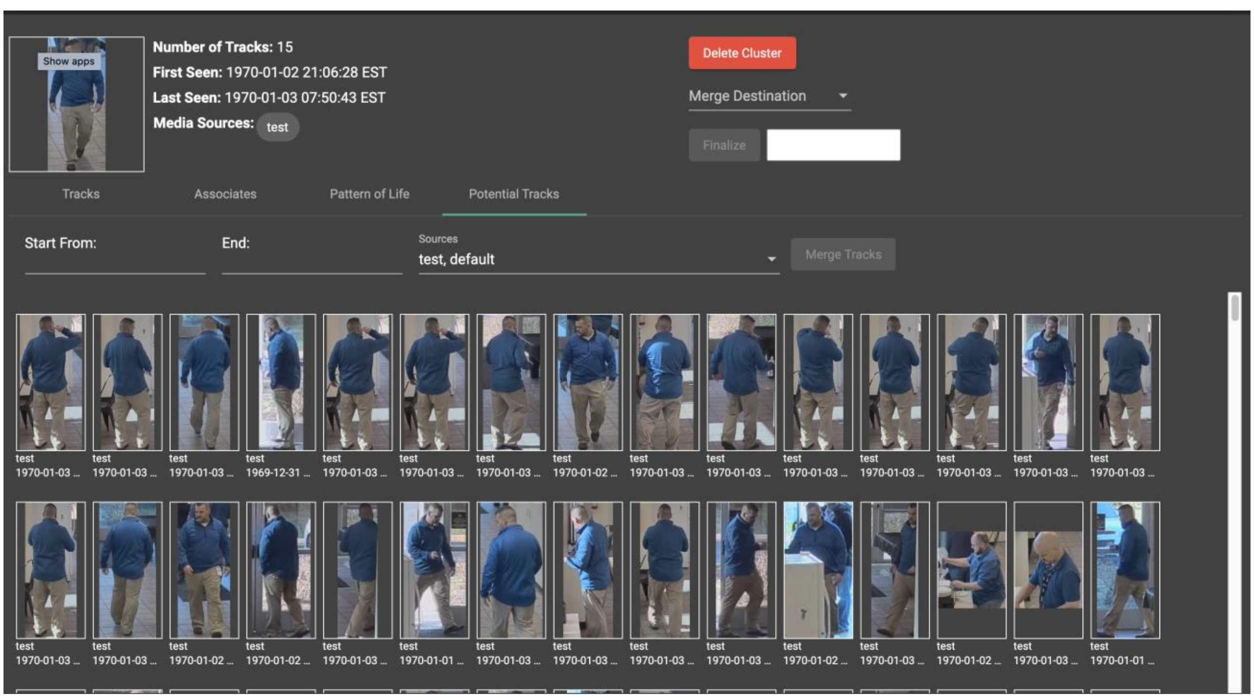

Figure 3: Person re-id in an indoor setting within the ClusterUI

2.4. Synthetic Data and Environments

Progeny Systems has created and is constantly evolving a state-of-the-art ecosystem of tools to train and deploy models using synthetic data generation. In order to generate the synthetic imagery, three-dimensional (3D) models of objects are used as a base, then covered with realistic textures and colors and inserted into a variety of environments. Through synthetically generated imagery and video, we can control all aspects of environmental conditions, resolution, viewpoint, etc. in order to fill in missing pieces of data that were not in a training corpus. The synthetic objects derived from the 3D models can be inserted into both synthetic environments or into a real imagery as well as video. While synthetic imagery has the potential to solve many problems associated with data acquisition, annotation, and model training for computer vision, there exist challenges that need to be addressed in order to advance this technology to the next level.

Our experience incorporating synthetic imagery into our workflow for model training has improved algorithm performance while at the same time providing valuable guidance for future research directions. In particular, we have identified the following challenges: realistic scene arrangements / object placement within the scene, rendering engine realism, model quality, speed of rendering, what factors do / do not improve performance, and view choice. These are all areas that Progeny believes can be improved to achieve more realistic and representative training data, which in turn should improve algorithm performance.

2.5. Edge Processing

Another focus area for our development has been on moving the processing as close to the sensor or data source as possible. This poses significant technical challenges, especially in severely size, weight, and power (SWAP) constrained environment such as small unmanned aerial system (UAS) platforms or other edge devices. To address this urgent technology need, we have developed *SPOTR-Edge*, which takes select core functionality of *SPOTR* and implements / optimizes them on an embedded processor. Both hardware and software innovation are necessary to create a system that has the required computer vision capabilities as well as the platform endurance needed for operational use.

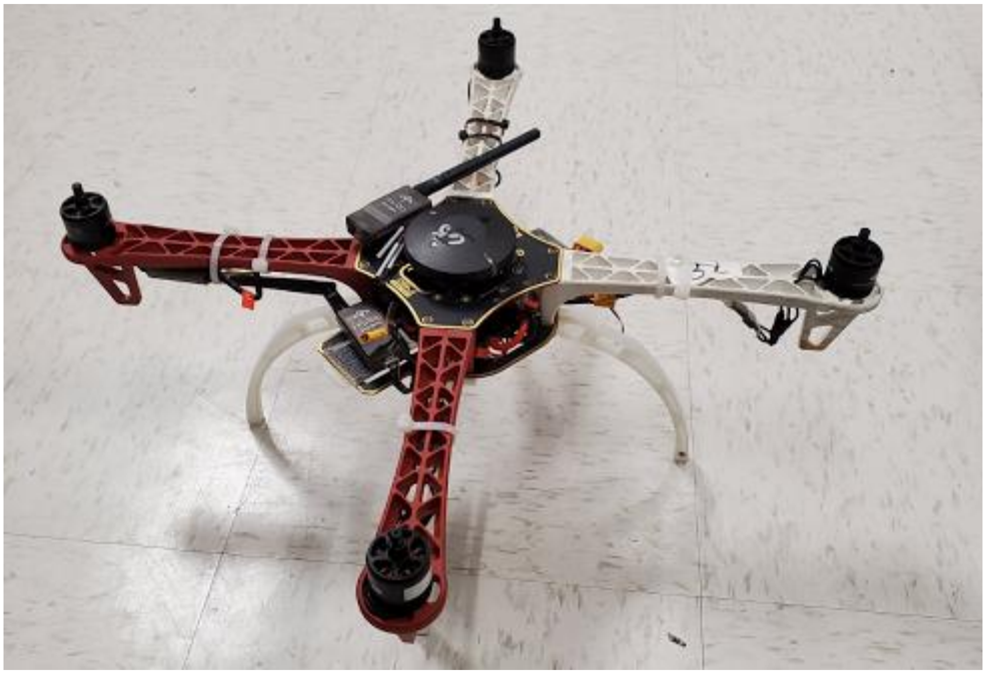

Figure 4: A test UAS platform used for collecting data

2.6. Data Collection

Throughout the many programs Progeny is a performer on, we need to collect data to train models that are tailored to specific missions of that customer's needs. Progeny owns a fleet of small UAS (sUAS) that have helped us gather large amounts of data relevant to the use cases of customers throughout the DoD and IC that can act as surrogates to operational platforms. This experience has allowed us to develop an infrastructure that is able to collect large and diverse amounts of data quickly that can then be fed into *Vision* for triage and annotation.