LTIMindtree

WHITEPAPER

Applying Computer Vision for **Storytelling in the Construction Industry**

Introduction

A picture is worth a thousand words and is often able to answer the questions of who, what, where and why. Construction sites are often bustling with workers and have lots of activities happening at the same time. A picture of a construction site tells a thousand tales. Analyzing such pictures using computer vision techniques can help tell a story of what's happening at site, which can help improve efficiency and safety, minimize risk, and monitor project-related progress. Construction companies are always looking for better project monitoring options and computer vision techniques provide more efficient, automated, round-the-clock monitoring.

In this white paper, we will discuss the computer vision techniques implemented to tell a story based on images from a construction site.

Problem Definition

To understand what is happening in an image, the human mind must identify the objects present and their roles. Similarly, given any image from a construction site, computer vision techniques must be able to identify objects and activities to put these pieces of information together and tell a story. The objects and activities can be translated into nouns and verbs that can help describe what is happening in the image.

Storytelling at a construction site can give the following types of information. For example:

- The number of workers present at the site
- The number of workers taking part in a welding activity at the site
- The number of workers taking part in a welding activity without the proper PPE (like hand gloves and face shields)
- The number of workers present at the site

One of our clients, a leader in the construction industry, wants LTIMindtree to apply and integrate computer vision techniques with their pre-existing CCTV infrastructure to describe and tell stories about the images captured.

In turn, the detected objects and activities in the images will be visualized on a web application dedicated to show the computer vision model's predictions on the images being analyzed and displayed on dashboards summarizing the predictions with respect to time, camera, and zone. The end user team from the construction site can utilize the dashboards to better understand what is happening on-site. For example, what times, cameras, and zones have captured what types of objects and activities more than others and require more focus towards ensuring efficiency, safety polices being followed, and project progress. Additionally, custom alerts can be set up by the user to trigger specific notifications for certain detections.

The following objects/activities are to be detected initially and more are to be added as the module scales:

- Workers
- Gas Cylinders (LPG)
- Gas Tanks (Welding)
- Welding Activity
- Grinding Activity

Scope:

The scope of work is as follows:

- Ingest images from the client's cloud cameras
- Analyze images to arrive at a feasible computer vision approach (Training)
- Apply computer vision approach on images and predict (Live)
- Creating a web-based application and give access to the customer through a secured URL
- Develop and host the necessary dashboards for visualization
- Provide the functionality of custom alerts

System Architecture

Figure 1 presents the system architecture while Table 1 summarizes each component present in the diagram.

Story Telling Architecture

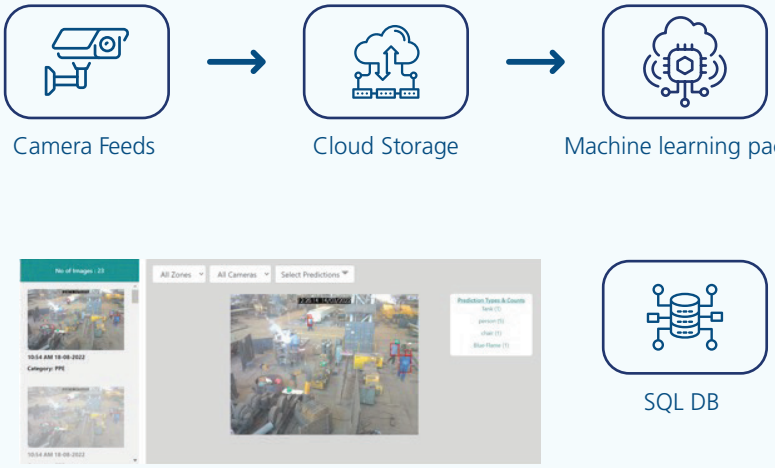

Figure 1: System Architecture

Components	Description
Camera Feeds	The images are retrieved from the cloud cameras.
Cloud Storage	The images are stored in the cloud.
Machine Learning Space	The Python computer vision model is hosted in the machine learning space and predicts on images from the cloud.
SQL Database	The database holds all the data generated by the application.
User Interface	The end user interacts with the underlying application with the user interface.

Table 1: Brief description of components

Technology Stack:

Technology	Version	Usage
Python	3.6.2	Used to perform computer vision model prediction on images.
Azure-ML	1.39.0	Used to deploy the computer vision models to the machine learning space.
MS SQL Server	2019	Stores data from the application
Angular	8.3.26	Used to develop the user interface
NodeJS	8.5.5	Used to develop the user interface.
.Net Core	3.1	Used to develop the user and API interface

Table 2: Technology Stack

Solution Approach:

By integrating computer vision techniques and pre-existing CCTV infrastructure, the events at a site can be described as a story. As shown below in the flowchart, multiple computer vision object detection models work in tandem to tell a story about the images captured. The following depicts this and the training pipeline.

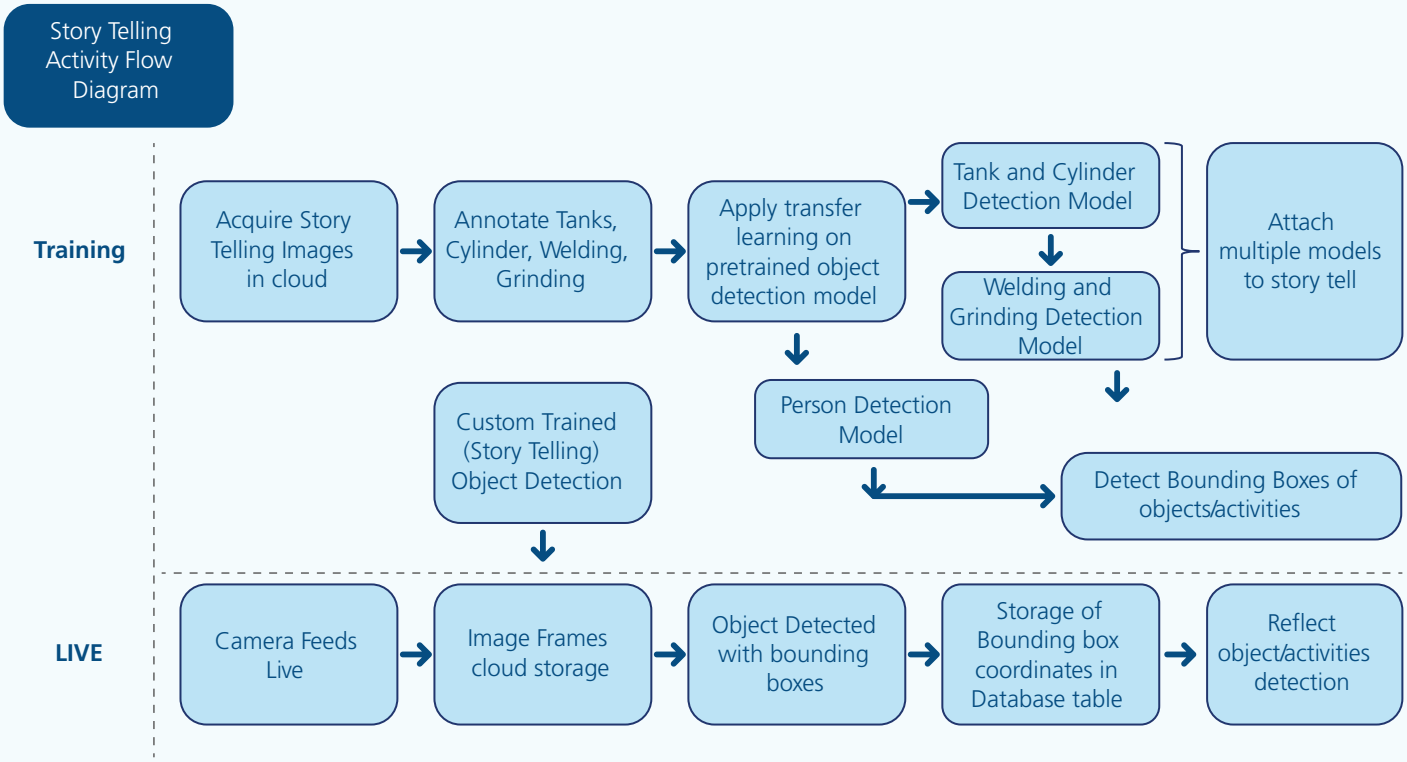

Figure 2: Outline of Story Telling Computer Vision Training Pipeline and Workflow

A person detection model is used to locate the workers in the image. For detecting the tanks, the object detection models for the cylinders, welding, and grinding activities are custom trained and used.

Storytelling is integrated into the computer vision web application as an additional description feature (an example is shown below in Figure 3). It is displayed on the web application with the other use case model predictions. Further in Figure 4 and 5, the custom alert setup page can be seen. The user can create custom queries for the system to alert the presence of specific situations. The solution was developed using LTIMindtree platform.

Figure 3: Computer Vision Model Description Output for Story Telling

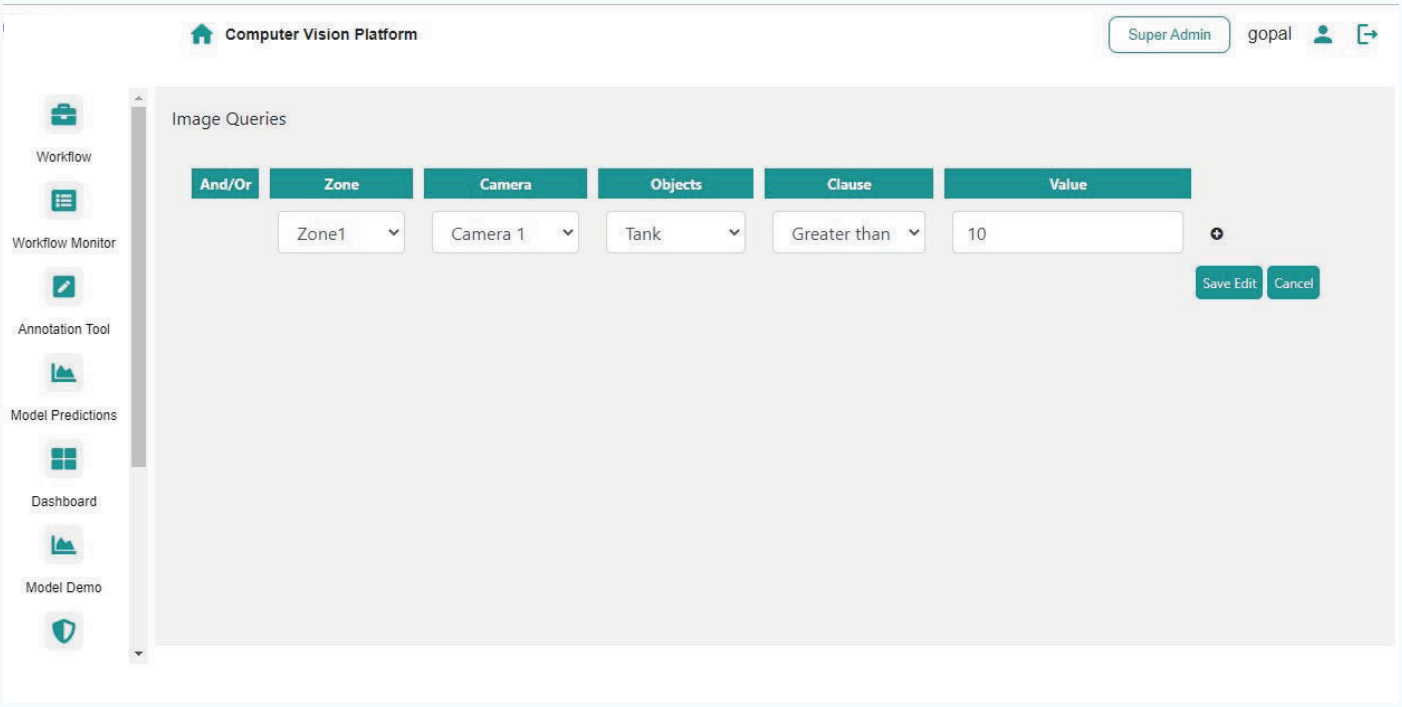

Figure 4: Setting up Custom Alerts for Story Telling Detections

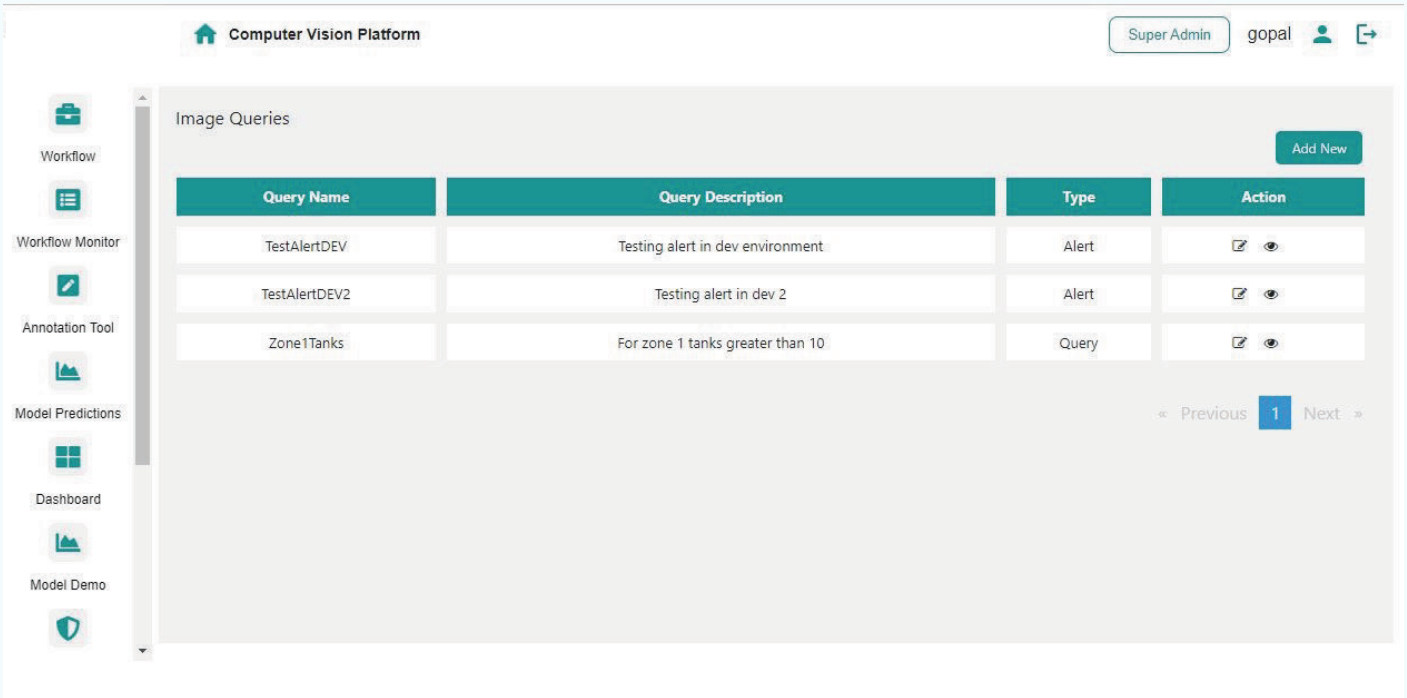

Figure 5: Displaying List of Custom Alerts for Story Telling Detections

Business Benefits

This computer vision solution aims at helping the end-user monitor their construction sites by telling a story based on the images captured using the existing CCTV infrastructure.

The beneficial aspects of the solution can be listed as below:

- User-friendly application to view non-compliance image predictions and raise tickets for initiating the right course of action.
- Efficient, automated, round-the-clock capabilities can replace monitoring manpower required by the client.
- It is deployed in the cloud with 24x7 uptime.
- Highly intuitive dashboards which publish summarized views of the predictions at the construction sites.

- Helps the client create awareness that the construction site is being monitored
- Customizable and scalable to handle more cameras and sites at any time.

Using computer vision to monitor a construction site can help in reducing risk and increasing awareness amongst construction workers to follow safety policies. This may subsequently improve efficiency on site and ensure that the project progress is in line with the timeline.

Conclusion

This white paper demonstrates the need for using computer vision to monitor construction sites by telling a story from the site's images. Subsequently, it minimizes risk and increases awareness on all fronts. Therefore, a service that can provide monitoring solutions can be beneficial for the construction industry. This same solution is applicable across industries.

About the Authors:

Aakash Agrawal

Senior Software Engineer

A result-oriented software engineer with 4+ years of experience in developing and deploying Deep learning projects and in objective optimization of construction foundations. Currently working on various image and video analytics use-cases.

Somsuvra Chatterjee

Director Product Engineering

A committed and passionate data scientist with overall 15+ years of experience, 12+ in the field of Analytics and Data Science. Worked across diversified domain starting with CPG, Retail & Manufacturing, Energy, etc. Engaged in developing solutions for various Computer Vision and Image Processing based products and projects.

He can be reached at somsuvra.chatterjee@ltimindtree.com

Abirami Vina

Senior Software Engineer

An IT professional having 4+ years of experience in R&D in the field of Artificial Intelligence, Deep Learning and Computer Vision. She completed Bachelor's in Information Technology, with an interest in Computer Vision and Image Processing algorithms. She is currently developing solutions using various computer vision and data science techniques.

Vishwanathan Raman

Prinicpal Director

Having 27+ years of experience across Geographies, expertise in the field of Business Intelligence, Data Warehousing, Artificial Intelligence, Machine Learning, Deep Learning and Data Sciences. Completed Masters from BITS Pilani on Data Analytics.

He can be reached at
vishwanathan.raman@ltimindtree.com