CAIP

AI Project Outline

Scenario: **Face Recognition**

What is the business problem that needs solving?

As an organization that provides security services to large corporations via mainly visual surveillance, and motion tracking as predicators to alarming the client, we are seeing an increased demand for the integration of atomization of the surveillance process, in addition to increased reliability of intruder identification. Our clients are interested in enhanced technologies that can provide increased efficiency, and faster processing; things that our current process cannot provide.

Our current process looks as such:

- When an anomaly is detected, there becomes a need to determine whether the alarm is a true threat or a false positive. Security operators do this via reviewing video from the camera that triggered the event.
- Security operators then need to decide whether it is a true threat, an employee on their lunch break, or just a visitor.
- If the security officer validates the alarm, the officer must then turn their attention to assessing the situation. They may choose to follow the suspect on camera and alert ground control in order to observe their behavior and determine the risk level.
- Furthermore, security officers will now look to utilize their resources to identifying the suspect, and recording the license plate number of an unauthorized vehicle.
- Head of security will then deploy security team to execute their response, provided the various assessments have determined one is necessary; or, they will instruct their security team to stand down.

How can AI solve this problem?

Implementing AI as part of our facial recognition software suite will allow for quicker, more efficient analysis of these various assessments. The benefits that can be seen from a more efficient, quicker risk analysis include:

- AI will enable our clients easily identify what is a human face on their surveillance, and in addition to recognizing the aesthetics of a person's face, the machine can also be programmed to analyze attributes such as mood.
- The ability to detect that a person is a potential threat, and further identify that the potential threat is: irritated, angry, or even hostile, within the matter of seconds, provides an invaluable resource for our clients.
- Not only does this eliminate the time needed for human assessment of risk and further human observation to assess the threat level, it mitigates the risk for those assessments and observations to be skewed by bias, bad judgement, preconceived notions, or prejudice.
- AI has the ability to accurately scan people being seen on surveillance, in real time, against millions of training data images in the mater of seconds.
- This kind of powerful processing will provide our clients with the highest level of security, as a financial institution for instance, will be able to detect an incoming threat almost immediately and take decisive action. The ability to search through millions of records in seconds, whereas individual analysis was previously required saves invaluable time.
- There are numerous examples of how AI has successfully increased the power and potential of facial recognition in the security industry. With applications across Law Enforcement, National Security, Government, Retail, Banking, and many others. AI enabled facial recognition software is currently being used in capacities such as identifying and locating: human trafficking victims, known criminals, wanted persons and vehicles, abducted children, terrorists, and many other potential security threats at every level.

What are tools/resources needed to implement the solution?

To implement AI as a solution, we may need to make a number of adjustments to our technology and data infrastructure, including storage networks, data pipelines, and/or processing systems. While each client's needs may vary slightly, security is straight forward concept that requires forward thinking concepts and analysis of historical data.

We will require the following:

- Identify in-house experts with the skillsets that will be required at different phases throughout the project.
- Raw data that can be turned into actionable business intelligence. This data is necessary in order to train the new model on human behaviors. We will need to ensure that this data is plentiful and contains what we would consider an equal amount of "normal" examples, as it does "abnormal" examples.

- The machine will use this data to effectively “learn” facial characteristics, features, body language, signs of emotional or mental distress, and characteristics associated with emotions like: joy, happiness, anxiety, sadness, fear, adrenaline, anger, irritability, and hostility.
- Once this data has been gathered from internal sources, as well as external sources, it will need to be assessed by our data engineers to ensure it is quality data that has been labeled appropriately. If the data point indicates a threat, it needs to be labeled as such. Furthermore, images or video will need to be further drilled down with the labels that indicate other data points we want it to lear, i.e., when using a video image of a bank robber as training data, we would want the machine to note the following: black hat obstructing visual, gait is abnormal, head positioned toward floor, perspiring, appears nervous, hands in pockets.
- Note that using too much identifying information in training data, such as male/female, brown/blonde hair, Caucasian/African American can lead to the machine being trained too specifically on the data, as opposed to utilizing the data as a reference to predict. This creates bias within the algorithm and unreliable outcomes, so special care needs to be taken to prevent this during the training stages.
- Once the model has been effectively trained, we will enter a test phase to gauge accuracy of the prediction model prior to deployment. This is a necessary step that should not be avoided, as this will ensure that the algorithm is working to spec, and as it has been designed to. The data used for testing should be a controlled sample, not used for training the model.

What ethical challenges might arise?

Some of the ethical concerns that may arise with the implementation of AI enabled facial recognition include:

- **The great debate of machines being unable to replace humans:** While some may argue that people can match faces to photos better than machines, the National Institute for Standards and Technology (NIST) recently shared a study of facial recognition technologies that are at least two years behind the models being used by companies using the most advance AI capabilities, and concluded that even older technologies could outperform human facial recognition capabilities.
- **The argument that the machine may produce false positives**: The process of avoiding false positives needs to be part of the model’s design. By following measures, as those outlined above regarding preventing bias, false positives can be great minimized. One can also argue that people do not have a way of eliminating these same biases or prejudices, and therefore have far more instances of falsely identifying someone as a potential threat. If false positives do occur, unlike people, one of the advantages of this technology is that it continuously learns and improves, so false positives would be greatly reduced over time.
- **Data Privacy:** Biometric data is a subset of PII which specifically refers to an individual's unique physical or behavioral characteristics that can be used to identify them. As an organization that specializes in surveillance technology, we are already bound by the laws regarding this sensitive data classification. If we plan to enable our technology with AI, we will need to examine the laws further

regarding the storage of this data and how it is being used, as the machine will keep this data indefinitely, and continuously use it as part of its analysis process for prediction.

What are some tactics for addressing these ethical challenges?

Though requirements for security are different then those of other applications, and certain clients requiring security will be considered exempt from certain regulations, i.e., government and law enforcement; to address these potential challenges, we must ensure the following:

- It is a requirement of GDPR that organizations adhere to the rules of privacy by design. The plan for development and the development process must take privacy into account throughout every phase.
- CCPA and California's AB 1950, require that organizations provide reasonable security to protect customers data privacy. We may want to consider what levels of access to PII will enable us to maintain our legal obligations, and have the same conversation with our clients regarding their internal servers. The choice to create more rigorous restrictions surrounding who has access to the data, and ways in which the data can be accessed may be required to minimize risks.
- All clients are required to make it publicly known that they are recording pedestrians, employees, and visitors. Entrance into an establishment that has this language clearly posted is legal acknowledgement and acceptance of the fact that your image will be captured during your time on the property.
- We must ensure that we have a well documented plan, outlining our processes for managing risk, privacy, and the solution being provided. Well maintained records to prove that these measures have been taken will also be necessary.