Analyze video content with Computer Vision and Azure Machine Learning

learn.microsoft.com/en-us/azure/architecture/ai-ml/architecture/analyze-video-computer-vision-machine-learning

Azure Machine Learning

Azure AI services

Azure Logic Apps

Azure Synapse Analytics

Azure Data Lake Storage

This article describes an architecture that you can use to replace the manual analysis of video footage with an automated, and frequently more accurate, machine learning process.

The FFmpeg and Jupyter Notebook logos are trademarks of their respective companies. No endorsement is implied by the use of these marks.

Architecture
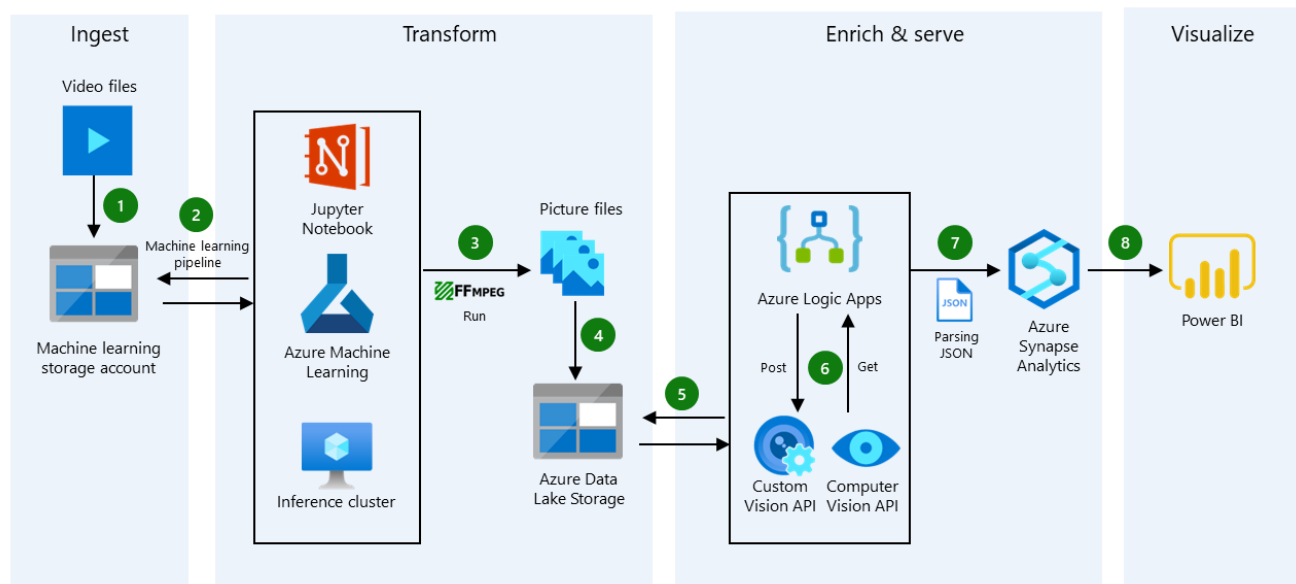
Download a PowerPoint file of this architecture.

Workflow

1. A collection of video footage, in MP4 format, is uploaded to Azure Blob Storage. Ideally, the videos go into a "raw" container.
2. A preconfigured pipeline in Azure Machine Learning recognizes that video files are uploaded to the container and initiates an inference cluster to start separating the video footage into frames.
3. FFmpeg, an open-source tool, breaks down the video and extracts frames. You can configure how many frames per second are extracted, the quality of the extraction, and the format of the image file. The format can be JPG or PNG.
4. The inference cluster sends the images to Azure Data Lake Storage.
5. A preconfigured logic app that monitors Data Lake Storage detects that new images are being uploaded. It starts a workflow.
6. The logic app calls a pretrained custom vision model to identify objects, features, or qualities in the images. Alternatively or additionally, it calls a computer vision (optical character recognition) model to identify textual information in the images.
7. Results are received in JSON format. The logic app parses the results and creates key-value pairs. You can store the results in Azure dedicated SQL pools that are provisioned by Azure Synapse Analytics.
8. Power BI provides data visualization.

Components

- Azure Blob Storage provides object storage for cloud-native workloads and machine learning stores. In this architecture, it stores the uploaded video files.
- Azure Machine Learning is an enterprise-grade machine learning service for the end-to-end machine learning lifecycle.
- Azure Data Lake Storage provides massively scalable, enhanced-security, cost-effective cloud storage for high-performance analytics workloads.
- Computer Vision is part of Azure Cognitive Services. It's used to retrieve information about each image.
- Custom Vision enables you to customize and embed state-of-the-art computer vision image analysis for your specific domains.
- Azure Logic Apps automates workflows by connecting apps and data across environments. It provides a way to access and process data in real time.
- Azure Synapse Analytics is a limitless analytics service that brings together data integration, enterprise data warehousing, and big data analytics.
- Dedicated SQL pool (formerly SQL DW) is a collection of analytics resources that are provisioned when you use Azure Synapse SQL.
- Power BI is a collection of software services, apps, and connectors that work together to provide visualizations of your data.

Alternatives

- <u>Azure Video Indexer</u> is a video analytics service that uses AI to extract actionable insights from stored videos. You can use it without any expertise in machine learning.
- <u>Azure Data Factory</u> is a fully managed serverless data integration service that helps you construct ETL and ELT processes.
- <u>Azure Functions</u> is a serverless platform as a service (PaaS) that runs single-task code without requiring new infrastructure.
- <u>Azure Cosmos DB</u> is a fully managed NoSQL database for modern app development.

Scenario details

Many industries record video footage to detect the presence or absence of a particular object or entity or to classify objects or entities. Video monitoring and analyses are traditionally performed manually. These processes are often monotonous and prone to errors, particularly for tasks that are difficult for the human eye. You can automate these processes by using AI and machine learning.

A video recording can be separated into individual frames so that various technologies can analyze the images. One such technology is *computer vision*: the capability of a computer to identify objects and entities on an image.

With computer vision, monitoring video footage becomes automatized, standardized, and potentially more accurate. A computer vision model can be trained, and, depending on the use case, you can frequently get results that are at least as good as those of the person who trained the model. By using <u>Machine Learning Operations (MLOps)</u> to improve the model continuously, you can expect better results over time, and react to changes in the video data over time.

Potential use cases

This scenario is relevant for any business that analyzes videos. Here are some sample use cases:

- **Agriculture.** Monitor and analyze crops and soil conditions over time. By using drones or UAVs, farmers can record video footage for analysis.

- **Environmental sciences.** Analyze aquatic species to understand where they're located and how they evolve. By attaching underwater cameras to boats, environmental researchers can navigate the shoreline to record video footage. They can analyze the video footage to understand species migrations and how species populations change over time.

- **Traffic control.** Classify vehicles into categories (SUV, car, truck, motorcycle), and use the information to plan traffic control. Video footage can be provided by CCTV in public locations. Most CCTV cameras record date and time, which can be easily retrieved via optical character recognition (OCR).

- **Quality assurance.** Monitor and analyze quality control in a manufacturing facility. By installing cameras on the production line, you can train a computer vision model to detect anomalies.

Considerations

These considerations implement the pillars of the Azure Well-Architected Framework, a set of guiding tenets that you can use to improve the quality of a workload. For more information, see Microsoft Azure Well-Architected Framework.

Reliability

Reliability ensures your application can meet the commitments you make to your customers. For more information, see Overview of the reliability pillar.

A reliable workload is one that's both resilient and available. *Resiliency* is the ability of the system to recover from failures and continue to function. The goal of resiliency is to return the application to a fully functioning state after a failure occurs. *Availability* is a measure of whether your users can access your workload when they need to.

For the availability guarantees of the Azure services in this solution, see these resources:

- SLA for Storage Accounts
- SLA for Azure Machine Learning
- SLA for Azure Cognitive Services
- SLA for Logic Apps
- SLA for Azure Synapse Analytics
- SLA for Power BI

Security

Security provides assurances against deliberate attacks and the abuse of your valuable data and systems. For more information, see Overview of the security pillar.

Consider the following resources:

- Identity management
- Protect your infrastructure
- Application security
- Data sovereignty and encryption

- Security resources

Cost optimization

Cost optimization is about reducing unnecessary expenses and improving operational efficiencies. For more information, see Overview of the cost optimization pillar.

Here are some guidelines for optimizing costs:

- Use the pay-as-you-go strategy for your architecture, and scale out as needed rather than investing in large-scale resources at the start.
- Consider opportunity costs in your architecture, and the balance between first-mover advantage versus fast follow. Use the pricing calculator to estimate the initial cost and operational costs.
- Establish policies, budgets, and controls that set cost limits for your solution.

Operational excellence

Operational excellence covers the operations processes that deploy an application and keep it running in production. For more information, see Overview of the operational excellence pillar.

Deployments need to be reliable and predictable. Here are some guidelines:

- Automate deployments to reduce the chance of human error.
- Implement a fast, routine deployment process to avoid slowing down the release of new features and bug fixes.
- Quickly roll back or roll forward if an update causes problems.

Performance efficiency

Performance efficiency is the ability of your workload to scale to meet the demands placed on it by users in an efficient manner. For more information, see Performance efficiency pillar overview.

Appropriate use of scaling and the implementation of PaaS offerings that have built-in scaling are the main ways to achieve performance efficiency.

Contributors

This article is maintained by Microsoft. It was originally written by the following contributors.

Principal author:

Oscar Shimabukuro Kiyan | Senior Cloud Solutions Architect – Data & AI

Other contributors:

To see non-public LinkedIn profiles, sign in to LinkedIn.

Next steps

Related resources