Parnal Sinha
ITAI 2376
L02 Skillable MSLE

Reflective Journal for Lab 02

Azure AI Vision Services – Image Analysis and Object Detection

Lab Overview

When I first dove into this Azure AI Vision lab, I couldn't help but think back to our midterm projects with MNIST and diffusion models. It's pretty wild to see how far image recognition has come! What started as basic digit classification has now evolved into super sophisticated image analysis tools that can basically "see" and describe images like a human would.

Exploring Azure AI Foundry

Setting up the project was an interesting process. Creating the hub, configuring the resource group, and navigating the Azure AI Foundry portal felt like putting together a complex puzzle. The platform walks you through each step, but there's definitely a learning curve. I found myself appreciating how cloud services have become more user-friendly over time, making complex AI technologies more accessible.

Key Capabilities Explored

The lab showcased four main image analysis capabilities:

1. **Image Captioning**:
    1. The system generates a single, human-readable description of an image.
    2. In our store camera example, it could describe the scene in a concise sentence.
2. **Dense Captioning**:
    1. This was the coolest feature for me.
    2. Unlike simple captioning, it provides multiple descriptions, highlighting different objects with precise bounding boxes.
    3. It's like having an AI that can point out and describe specific elements in an image.
3. **Image Tagging**:
    1. The system identifies and tags objects, actions, and scene elements.
    2. What's neat is the confidence score that comes with each tag.
    3. It goes beyond just naming objects – it can recognize actions like "shopping" or "standing".
4. **Object Detection**:
    1. This feature can pinpoint and outline specific objects in an image.
    2. The threshold slider was particularly interesting, allowing us to adjust the confidence level for object recognition.

Parnal Sinha
ITAI 2376
L02 Skillable MSLE
Connecting to Previous Work

Comparing this to our MNIST project was eye-opening. Where MNIST was about classifying handwritten digits with basic neural networks, these Azure Vision tools represent a massive leap forward. We've gone from "is this a 7 or a 9?" to "describe everything happening in this complex scene".

Potential Future Applications

The potential downstream applications are exciting. Beyond retail scenarios like the "smart store" example, I can see this technology being crucial in:

1. Medical imaging diagnostics
2. Autonomous vehicle perception
3. Security and surveillance
4. Accessibility tools for visually impaired individuals
5. Quality control in manufacturing

Challenges and Insights

Honestly, the challenges were minimal. The biggest hurdle was getting comfortable with the Azure interface. A few key learnings:

1. Resource location matters (some regions support more features)
2. Cloud AI services are becoming increasingly user-friendly
3. Confidence scores are critical in understanding AI predictions

Closing Thoughts

This lab was a great reminder of how rapidly AI vision technologies are advancing. From our early experiments with MNIST to these sophisticated cloud-based vision services, the progress is remarkable. It's not just about recognizing patterns anymore – it's about understanding context, describing scenes, and providing actionable insights.

The future of computer vision looks incredibly promising, and labs like these give us a glimpse into that potential.

Parnal Sinha
ITAI 2376
L02 Skillable MSLE

Lab Completed:

Access to your labs will expire on Sunday, December 21, 2025 5:00 PM (UTC)

20%

1 of 5 required activities complete

☑ 1   ⚗ **Analyze images in Azure AI Foundry portal** (Expected Duration 1 hours)   Details ∨   ⓘ
AI-900T00-A Microsoft Azure AI Fundamentals [Cloud Slice Provided], Learning Path 02 (CSS)

Required: Yes
Status: Running
Started: Thursday, May 8, 2025 10:06 PM (Central Standard Time)
Ended: Thursday, May 8, 2025 10:26 PM (Central Standard Time)

| Launch | Cancel |

8 of 10 launch attempts remaining

↓

Technologies Explored: Azure AI Vision, Cloud Computing, Image Analysis

store-camera-3[1].jpg

Detected attributes   JSON

Threshold value ⓘ ———————○——— 70

person
(84.80%)

store-camera-3[1].jpg

Detected attributes   JSON

Threshold value ⓘ ——○———————— 15

person
(84.80%)

supermarket
(51.70%)

Parnal Sinha
ITAI 2376
L02 Skillable MSLE

← **Extract common tags from images** ⌄

store-camera-2[1].jpg

clothing (99.04%)
person (98.26%)
convenience store (98.00%)
retail (97.88%)
supermarket (95.77%)
grocery store (94.18%)
shopping (91.96%)
customer (91.77%)
text (90.62%)
trade (90.15%)
human face (90.05%)
indoor (89.03%)
woman (88.37%)
selling (88.15%)
shop (88.11%)
market (86.82%)
store (81.26%)
shelf (69.05%)
standing (65.64%)
marketplace (51.18%)