Image Recognition using AWS

Group Members

Aditya Goverdhana (agoverdh@kent.edu)

Balakrishna Phani Kommanaboina (bkommana@kent.edu)

Naveena Kanderi (nkanderi@kent.edu)

Jagadeesh Karri (jkarri1@kent.edu)

March 18, 2023

Instructor: Prof. Debobroto Das Robin

CAPSTONE PROJECT (CS-69099-001)

KENT STATE UNIVERSITY

Department of Computer Science

Introduction

Problem Statement

The project aims at building an elastic web application that can automatically scale out and scale in on-demand and cost-effectively by using cloud resources. The resources used were from Amazon Web Services. It is an image recognition application exposed as a Rest Service to the clients to access.

Objectives

The application takes the images and returns the predicted output by the deep learning model by using the AWS resources, an IaaS provider. AWS as an IaaS provider offers a variety of compute, storage and message services. So the tasks involved designing the architecture, implementing RESTful Web Services, a load balancer that scales in and scales out EC2 instances at App Tier according to the demand of the user.

Milestones

The initial milestones involve

- To design the System architecture
- Collect image data through web-scraping.
- Design and build an interactive system for User.
- Enable AWS Services and run some tests to see storage and computation performance.
- Finally, host the application using open-source platforms.

Tools

- AWS services (EC2, SQS, S3).
- AWS Image Recognition library (AMI).
- Web Services (HTML, CSS, JS, Flask).
- Testing for resources.
- Python for backend of the architecture.
- Heroku for hosting the application.

Problems to Address

As we know, taking in such a large amount of data for training or modelling means AWS can potentially charge for extra storage to address this issue we are planning to do progressive loading or utilize a big data platform.

Need to run some tests at the end to make sure the application is running properly on the platform.

Literature Review

Image Recognition is an essential component of various applications, such as facial recognition, object detection, and product recommendation. With the growing demand for image recognition, many researchers have focused on developing image recognition systems using cloud and edge computing. This literature review summarizes and analyzes five research papers that discuss image recognition as a service.

In the paper "Cloud Strategies for Image Recognition," da Costa and Pisa [2] discuss various cloud strategies for image recognition. The authors analyze the features of different cloud providers, such as Amazon Web Services, Google Cloud Platform, and Microsoft Azure, and evaluate their suitability for image recognition tasks. The paper emphasizes the importance of cloud computing in image recognition and provides insights into cloud-based image recognition services.

In "Real-Time Object Detection with TensorFlow Model Using Edge Computing Architecture," N et al. [3] propose a real-time object detection system using a TensorFlow model and edge computing architecture. The paper presents a system architecture that utilizes edge computing to reduce latency and improve response time. The authors evaluate the system's performance using a dataset of real-world images and demonstrate the effectiveness of the system in detecting objects in real-time.

In "The Utilization of Cloud Computing for Facial Expression Recognition using Amazon Web Services," Rafael et al. [4] discuss the utilization of cloud computing for facial expression recognition. The authors use Amazon Web Services to develop a facial expression recognition system and evaluate the system's performance using a dataset of facial expressions. The paper demonstrates the effectiveness of cloud-based facial expression recognition systems and highlights the potential of such systems in various applications, such as security and entertainment.

In "Object Detection and Recognition using Amazon Rekognition with Boto3," Sharma

[5] discusses object detection and recognition using Amazon Rekognition with Boto3, a software development kit for Python developers. The paper presents an overview of the Amazon Rekognition service and explains how to use the service with Boto3. The author demonstrates the effectiveness of the Amazon Rekognition service in detecting and recognizing objects in real-time, providing a practical guide for developers.

Finally, Suguna et al. [6] present "An Efficient Real-time Product Recommendation using Facial Sentiment Analysis." The paper proposes an efficient product recommendation system that utilizes facial sentiment analysis to determine a user's emotional state. The system uses the Microsoft Azure Face API to extract facial features and sentiment scores, which are then used to recommend products. The authors evaluate the system's performance using a dataset of user reviews and demonstrate the system's effectiveness in generating relevant product recommendations.

In conclusion, these five research papers provide insights into image recognition as a service using cloud and edge computing. The papers demonstrate the effectiveness of cloud-based and edge-based image recognition systems and provide practical guidance for developers. The papers also highlight the potential of image recognition in various applications, such as facial recognition, object detection, and product recommendation.

Our project aims to create an elastic web application that uses cloud resources to develop an image recognition application. The project involves designing the architecture and implementing RESTful Web Services to provide access to the application. The use of cloud resources to provide scalable and cost-effective solutions is a common theme in the above 5 research papers. The above research papers emphasize the use of cloud-based services, including AWS services, for scalable and cost-effective image recognition solutions.

Project Background

Image recognition, also known as computer vision, is a field of study in artificial intelligence that involves training computer systems to identify, interpret, and understand images or visual data. It involves developing algorithms and models that can recognize patterns, objects, and features within an image or video.

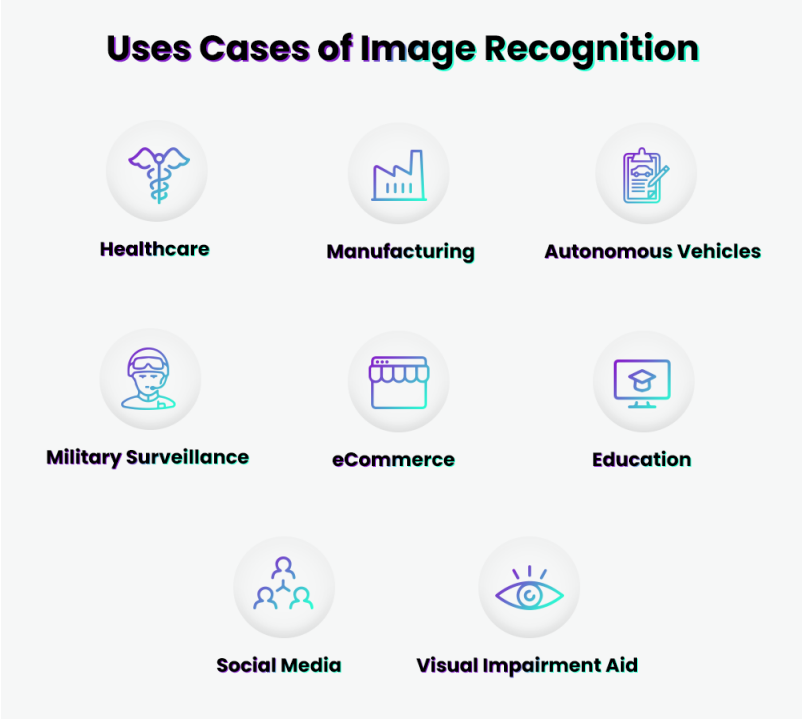

Figure 1: Some use-cases of Image Recognition.Ariwala [1]

Image recognition has a wide range of applications across various industries such as healthcare, automotive, retail, security, and entertainment. Some examples include:

- **Healthcare**: image recognition technology can be used for medical imaging analysis, disease diagnosis, and drug discovery.
- **Automotive**: image recognition can be used for autonomous driving, traffic monitoring, and vehicle safety features.

- **Retail**: image recognition can be used for product recognition, inventory management, and customer behavior analysis.
- **Security**: image recognition can be used for facial recognition, object detection, and surveillance monitoring.
- **Entertainment**: image recognition can be used for augmented reality, virtual reality, and gaming.

IaaS (Infrastructure as a Service) is a cloud computing model in which cloud service providers offer virtualized computing resources such as servers, storage, and networking to users on a pay-per-use basis. IaaS allows users to scale up or down their computing resources as needed, without having to invest in their own infrastructure. Examples of IaaS providers include Amazon Web Services, Microsoft Azure, and Google Cloud Platform.

IaaS services have a wide range of applications across various industries, such as:

- **Web hosting**: IaaS providers can host websites and web applications in the cloud, providing scalable and reliable infrastructure.
- **Big data**: IaaS providers can provide the computing resources needed to process and analyze large volumes of data.
- **Disaster recovery**: IaaS providers can provide backup and recovery solutions in the event of a disaster or outage.
- **DevOps**: IaaS providers can provide the infrastructure needed for software development and testing.
- **Machine learning**: IaaS providers can provide the computing resources needed to train and run machine learning models.

Infrastructure

The infrastructure described is a common architecture for building web applications that require image recognition capabilities. The infrastructure is divided into two tiers: the Web Tier and the Application Tier.

The Web Tier is responsible for interacting with users and displaying the user interface. It uses HTML/JS/CSS to render the interface and collects user inputs. In this architecture, a Python Flask is used for user interaction.

The Application Tier is the core of the application and performs the image recognition tasks. It handles business logic, database manipulation functions (CRUD), and communicates with other resources. In this architecture, an AWS Deep Learning model is used for image classification.

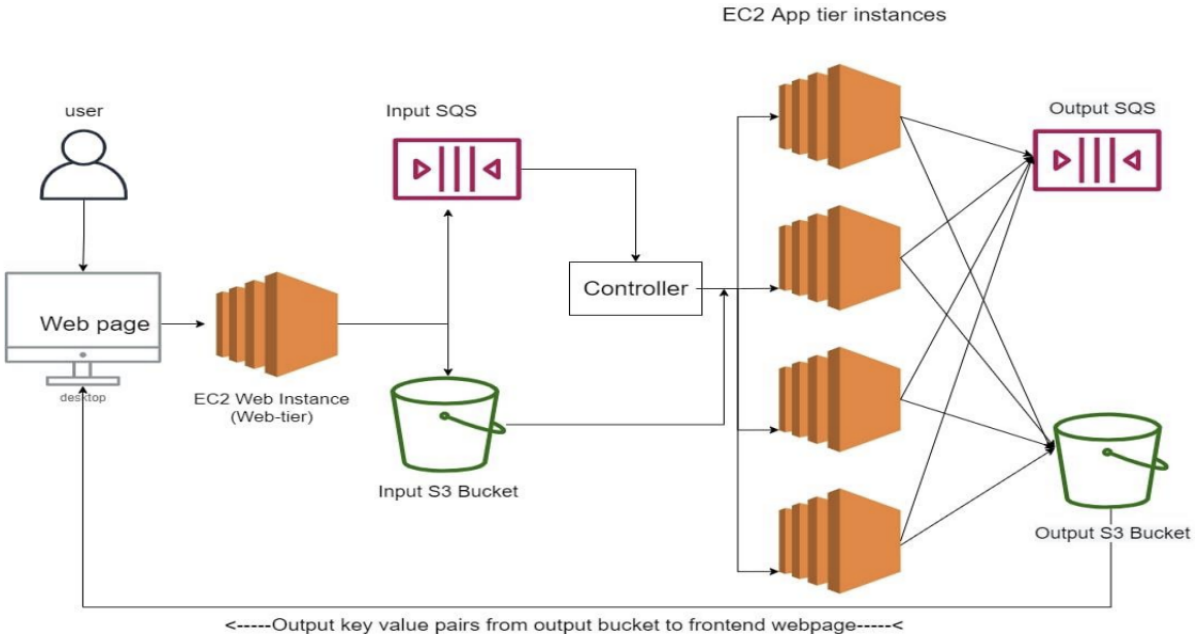

Figure 2: System Architecture.

The resources used in this infrastructure include input and output S3 buckets for loading images and displaying classified images on the user interface. The App Tier uses SQS for

input images as messages and stores them for processing. The classified image results are then passed onto the output SQS and then onto the output S3 bucket, which is then used by the Web Tier to display the classified images on the user interface.

The number of EC2 instances to be used depends on the amount data uploaded. In case if we upload images less than 15 then the number of instances running will be equal to the number of images uploaded.

Overall, this infrastructure provides a scalable and cost-effective way to build web applications with image recognition capabilities. By separating the Web and Application Tiers and using AWS services, the infrastructure can automatically scale resources up or down based on demand, improving performance and reducing costs.

References

[1] Pinakin Ariwala. What is the working of image recognition and how is it used?, 2022. URL `https://marutitech.com/working-image-recognition/`.

[2] Bernardo Botelho Antunes da Costa and Pedro Silveira Pisa. Cloud strategies for image recognition. In *2020 4th Conference on Cloud and Internet of Things (CIoT)*, pages 57–58, 2020. doi: 10.1109/CIoT50422.2020.9244200.

[3] Mahiban Lindsay N, Alla Eswara Rao, and Madaka Pavan Kalyan. Real-time object detection with tensorflow model using edge computing architecture. In *2022 8th International Conference on Smart Structures and Systems (ICSSS)*, pages 01–04, 2022. doi: 10.1109/ICSSS54381.2022.9782169.

[4] Gregorius Rafael, Hendra Kusuma, and Tasripan. The utilization of cloud computing for facial expression recognition using amazon web services. In *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pages 366–370, 2020. doi: 10.1109/CENIM51130.2020.9297974.

[5] Vivek Sharma. Object detection and recognition using amazon rekognition with boto3. In *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 727–732, 2022. doi: 10.1109/ICOEI53556.2022.9776884.

[6] R. Suguna, M. Shyamala Devi, Akash Kushwaha, and Puja Gupta. An efficient real time product recommendation using facial sentiment analysis. In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pages 1–6, 2019. doi: 10.1109/ICECCT.2019.8869300.