Image Analysis Web Application with Leverage Azure Computer Vision Services

Kayfa Liyamiza[1], Fayza Marta Devia[2], and Apbenk Fathur Rahman[3]
[1]kayfaliyamiza07@gmail.com
[2]fayzamartadevia@gmail.com
[3]apbenkfathur@gmail.com
Informatics Engineering, The Padang State University

***Abstract*—One Azure cloud computing service called Computer Vision can identify and detect information from photos. The author of this development uses the Computer Vision Azure service to automatically detect handwriting on photographs and to detect images in general. To make it simpler for users to identify handwriting in photos, the Computer Vision Azure service's handwriting detection is implemented using a web service. Because it is so effective at detecting the information contained in images, computer vision plays a significant role in data collection, particularly data in the form of images. Users can quickly and efficiently extract text from photos using Computer Vision capabilities, and then utilize that text for a variety of applications, like looking up information in papers with handwritten text or photographs. The process of automated handwriting detection demonstrates the immense potential of Computer Vision in terms of time-saving document analysis, record digitization, and information collection.**

Keywords: Computer Vision, Azure, Cloud Computing, Image Detection, Handwriting.

I. INTRODUCTION

In this digital age, image and visual content analysis is expanding quickly. One technological advancement that significantly aids in effectively identifying and detecting the information contained in images is computer vision. One cloud computing provider that provides computer vision services using artificial intelligence for object detection and identification in images is Microsoft Azure.

In this presentation, we will talk about creating web-based applications that automatically recognize handwriting on photographs and analyze images in general using Azure Computer Vision services. The website will be simple to use for online users to navigate; all they need to do is input the image's URL address to have the image detected and the handwriting recognized. Compared to manual detection, this web application will simplify the process for users to evaluate handwriting on photos.

By using Computer Vision services to identify photos and any writing in the image, we will be able to demonstrate the steps involved in creating a web application. This web application is anticipated to facilitate and expedite the process of identifying writing on photos and aid in the effective recognition of information in images by utilizing Microsoft Azure's Computer Vision service capabilities.

The system will be designed to make it simple for users to enter the URL of the picture they wish to examine. Subsequently, the system will present data derived from the analysis regarding the identified objects or text. But, the development being done by the researchers will just make use of the already-available Computer Vision Azure service to obtain detection findings on photos and add a user-friendly website interface to make it accessible online.

II. THEORETICAL BASIC

A. *Cloud Computing*

In the field of information technology, cloud computing is a very significant and influential technology. One important aspect of cloud computing is accessibility. Users do not need to install any local software or hardware in order to access the computer resources they require via the internet. As long as they are linked to the internet, this offers freedom and convenience in accessing data and applications from any location at any time.

Additionally, cloud computing is very scalable. Users can simply adjust the capacity of the resources they utilize to suit their needs thanks to this capability. This enables them to raise carrying capacity when heavier loads are required or modify the resources used to meet shifting demand.

The fact that cloud computing expenses are customized to your needs is an additional benefit. Users only pay for the resources they really utilize under this paradigm.By doing this, they are able to pay for the resources used over a predetermined period of time and without having to make a huge upfront investment in physical infrastructure.

One key benefit of cloud computing is the flexible and effective way in which data may be managed, stored, and accessed. Customers can manage their data with an easy-to-use interface while storing it in a dependable and safe data center. Everyone has easy access to and sharing of data, facilitating improved teamwork and more mobility.

Cloud computing service models that include features like on-demand self-service, wide network access, resource pooling, quick elasticity, and measured services are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The four cloud implementation models are community, private, hybrid, and public clouds. Every model offers benefits and drawbacks, including those related to scalability, cost effectiveness, data security, and organizational reliability.

B. *Computer Vision*

- Image Recognition: This refers to the identification and categorization of faces, objects, text, and other visual patterns in pictures. Convolutional Neural Networks (CNN), one of the several image recognition models and algorithms that Azure provides, are particularly useful for this kind of task.

- Object Detection: This process locates and positions particular objects inside the picture. Users can locate significant things in photos by using Azure object detection methods including Fast R-CNN, Faster R-CNN, and Single Shot MultiBox Detector (SSD) models.

- Semantic Segmentation: The process of breaking up a picture into a few sections and classifying them into relevant categories enables users to view a specific area inside a picture. Semantic segmentation models such as Fully Convolutional Networks (FCN) and U-Net are available in Azure to be used in this situation.

- Face Processing: Emotion analysis, face detection, identity verification, and facial image emotion recognition are all included in face processing. Azure offers the Azure facial API for facial processing.

- OCR (Optical Character Recognition): Thanks to OCR technology, computers can now read and identify text included within photographs. You can use Azure's OCR service to extract text from pictures and documents.

- Video Processing: Azure is also capable of real-time video processing and analysis for Computer Vision applications, such as object tracking, object detection, face recognition, and action recognition. Azure offers the Video Indexer service, which integrates these functionalities.

C. *Cloud Azure Shell*

The Azure Cloud Shell Service gives users instant access to an Azure-managed cloud-based development environment together with the tools and software needed to develop applications. Users can use a shell client or a web browser to access this environment.

A ready-to-use shell environment with a range of tools and frameworks frequently used in application development is offered by Cloud Shell Azure. This environment is integrated with Azure and lets you use other services, manage cloud resources, and perform Azure CLI commands without requiring you to install or set up extra software locally.

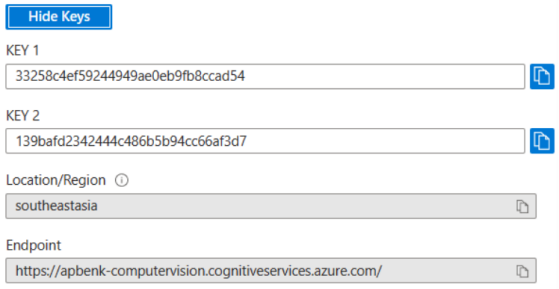
Python is a well-liked programming language that's frequently utilized to create applications. You can develop, test, and execute Python code straight from the shell with Azure Cloud Shell since it comes with an environment set up with an interpreter.

You can use the Azure Cloud Shell service to manage cloud resources, including virtual machine creation and configuration, web service management, database management, and other Azure services. The Azure Cloud Shell service connects directly with Azure. These resources can also be used with PowerShell or Azure CLI commands. The deployment of Python applications can also be automated.

III. DISCUSSION

A. *Computer Vision Resources*

Making an Azure account, configuring it, and choosing the appropriate subscription plan should be your initial steps. Subsequently, generate computer vision resources; if necessary, Azure will furnish an API key and endpoint, which will subsequently serve as a platform for web apps through the Python program.



B. *Making Web Apps with Python Programs*

After getting the API key and Endpoint from Azure, the API key and Endpoint will be used in this program:

```
import azure.ai.vision as sdk

# Replace the placeholders with your actual vision endpoint and key
VISION_ENDPOINT = "https://apbenk-computervision.cognitiveservices.azure.com/"
VISION_KEY = "33258c4ef59244949ae0eb9fb8ccad54"

service_options = sdk.VisionServiceOptions(VISION_ENDPOINT, VISION_KEY)

vision_source = sdk.VisionSource(
    url="https://i.redd.it/pf440nc480921.jpg"
)

analysis_options = sdk.ImageAnalysisOptions()
analysis_options.features = (
    sdk.ImageAnalysisFeature.CAPTION |
    sdk.ImageAnalysisFeature.TEXT
)
analysis_options.language = "en"
analysis_options.gender_neutral_caption = True

image_analyzer = sdk.ImageAnalyzer(service_options, vision_source, analysis_options)

result = image_analyzer.analyze()

if result.reason == sdk.ImageAnalysisResultReason.ANALYZED:
    captions = []
    texts = []
    if result.caption is not None:
        captions.append(result.caption.content)
    print("Captions : ")
    print(" ".join(captions))
    print(" ")
    if result.text is not None:
        for line in result.text.lines:
            words = []
            for word in line.words:
                words.append(word.content)
            texts.append(" ".join(words))
    print("Text : ")
    print(" \n".join(texts))
```

Fig. 1: Program screenshot

For AI-enabled processing of images, the program above makes use of the azure-AI-Vision library from the Azure Cognition Service service. The program's output will show the caption and any text that is identified from the image that has been submitted using the URL based on the source code.

The following is an example of an image the the program above will try to detect:

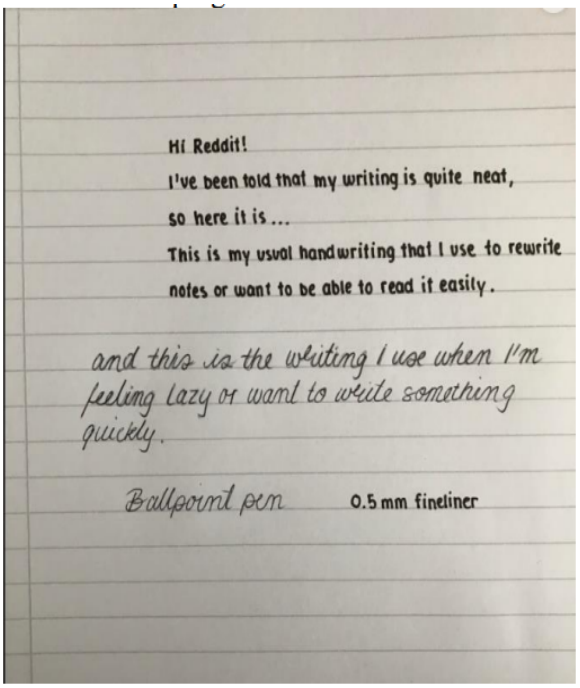


Fig. 2: Input to the Python program

The image input must be the URL of the iamge, by replacing the URL variable in the source code with the URL:

So The program will display the following output:

```
Captions :
a piece of paper with writing on it

Text :
Hi Reddit!
I've been told that my writing is quite neat,
so here it is ...
This is my usual handwriting that I use to rewrite
notes or want to be able to read it easily.
and this is the writing I use when I'm
feeling lazy or want to write something
quickly
Ballpoint pen
0.5 mm fineliner
```

Fig. 3: Program output results

Two variables, namely Text and Caption, are visible based on the program outcomes. The output from the program will look like this: ”A piece of paper with writing on it” is the caption that appears as a consequence of general image detection. The line in the image that the Text variable displays corresponds to; if a new line is found in the image, the text will also be created on a new line in the program’s output. This is due to the fact that Azure computer vision’s AI can identify photos with bounding boxes, like the one below:

This bounding box is what makes the output results line up according to the bounding box detected by AI cloud computing on handwritten images.

C. *Deploy Aplication*

To deploy this application , the Flask framework is used, along with a Python program that has been integrated with the Flask library

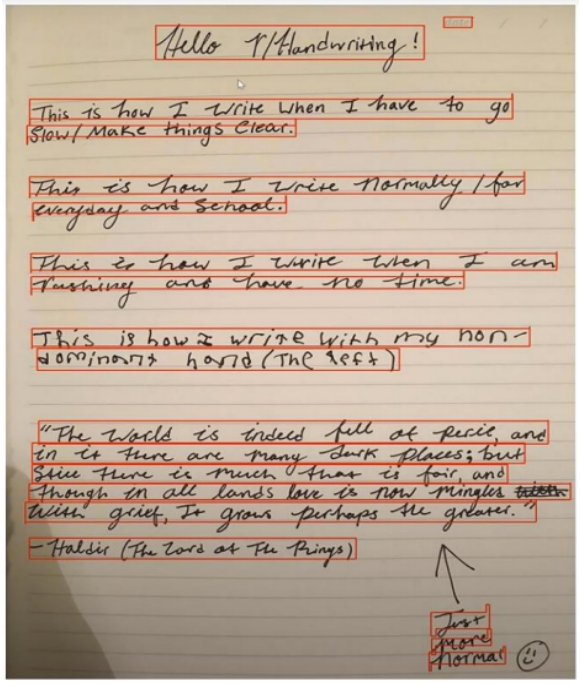


Fig. 4: Bounding box in handwriting detection

```python
@app.route('/')
def home():
    return render_template('index.html')

@app.route('/analyze', methods=['POST'])
def analyze():
    url = request.form['url']

    service_options = sdk.VisionServiceOptions(VISION_ENDPOINT, VISION_KEY)
    vision_source = sdk.VisionSource(url=url)

    analysis_options = sdk.ImageAnalysisOptions()
    analysis_options.features = (
        sdk.ImageAnalysisFeature.CAPTION |
        sdk.ImageAnalysisFeature.TEXT
    )
    analysis_options.language = "en"
    analysis_options.gender_neutral_caption = True

    image_analyzer = sdk.ImageAnalyzer(service_options, vision_source,
                                       analysis_options)
    result = image_analyzer.analyze()

    if result.reason == sdk.ImageAnalysisResultReason.ANALYZED:
        captions = []
        texts = []
        if result.caption is not None:
            captions.append(result.caption.content)

        if result.text is not None:
            for line in result.text.lines:
                words = []
                for word in line.words:
                    words.append(word.content)
                texts.append(" ".join(words))
```

Fig. 5: Python program using flask

For the results of the program (Figure 5), the website will be displayed on localhost http://127.0.0.1:5000. Here is a simple view of the website:

The program output section will be displayed as follows using the same url as Figure 2:

To deploy this application so that it can be used by the public, github and the cloud shell storage service from Azure are used, following the link from github:

Fig. 6: Main appearance of the website

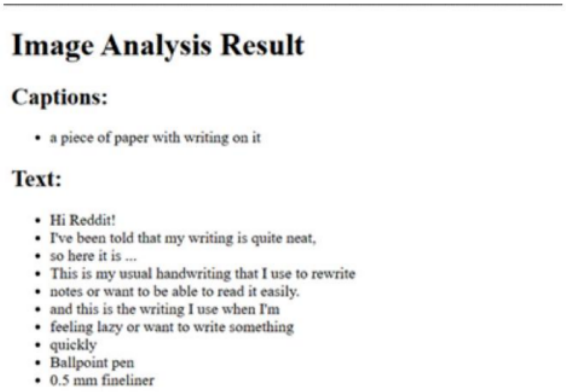


Fig. 7: Display of output on the website

https://github.com/apbenk/ apbenkcv.git and the link from the application that has been deployed using cloud shell st

IV. CONCLUSION

This paper is one example of how to develop a web application that uses Azure's Computer Vision service to detect common images and automatically identify handwriting in this paper. This application speeds up the process of digitizing archives and data collection, saves time for document analysis, and extracts data from images with artificial intelligence.

Finally, handwriting recognition and general image detection are possible through the use of Azure Computer Vision services while developing this web application. This speed up the process of data collection, archive digitization and document analysis. Azure Computer Vision is very effective in collecting data, especially images, with high efficiency.

REFERENCES

[1] Ministry, W., Siti, K., Sholikhatin, A., & Prasetyo, AB (2020). Integration of Telemedicine with Cloud Computing in. *JOURNAL OF INFORMATICS*, 7(2), 91–96. http:// ejournal.bsi.ac.id/ejurnal/index.php/ji.

[2] Chen, S., Saiki, S., & Nakamura, M. (2020). *Toward flexible and efficient home context sensing: Capability evaluation and verification of image-based cognitive APIs. Sensors (Switzerland), 20(5).* https://doi.org/10.3390/s20051442

[3] Wijaya, W., Tolle, H., & Utaminingrum, F. (2018). Grid-Double Block Method for Detecting the Left Margin of Handwriting in Graphology Applications. *Journal of Information Technology and Computer Science*, 5(1), 69. https://doi.org/10.25126/jtiik.201851575

[4] Guntara, RG (2022). Facial Image Recognition Application on KTP Using Google Cloud Vision API and Kairos API Based on Android. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 4(2), https://doi.org/ 10.28926/ilkomnika.v4i2.504

[5] Chellsya, Ananda Ayu, Suci Aulia, and Sugondo Hadiyoso. "Implementasi Computer Vision Dalam Mendeteksi Pelanggaran Tidak Menggunakan Helm Pada Pengendara Motor." *eProceedings of Applied Science* 9.1 (2023).

[6] https://learn.microsoft.com/en-us/azure/cognitive-services/computer-vision/overview/computer-vision-for-digital-asset-management, (accessed on 5 June 2023)