Enhancing Arabic Alphabet Sign Language Recognition with VGG16 Deep Learning Investigation

A. M. Elshaer
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
ahshaer1@aast.edu

Yousef Ambioh
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
youssefbozayed@gmail.com

Ziad Soliman
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
ziadsoliman44@gmail.com

Omar Ahmed
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
omarahmedd15702@gmail.com

Miral Elnakib
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
miralelnakib7@gmail.com

Mohamed Safwat
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
mohammed.safwat99@hotmail.com

Salma M. Elsayed
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
salma54@aast.edu

Mahmoud Khalid
College of Artificial Intelligence
Arab Academy for Science, Technology, and Maritime Transport
Al Alamein, Egypt
mahmoudkahled@aast.edu

2024 14th International Conference on Electrical Engineering (ICEENG) | 979-8-3503-4342-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICEENG58856.2024.10566400

***Abstract*—This paper presents ArASL (Arabic Alphabet Sign Language) recognition, a system aimed at fostering communication between deaf and hearing individuals by converting Arabic sign language gestures into text or speech. The system utilizes visual recognition of hand gestures from image inputs, employing a novel algorithm that leverages hand geometry and distinct hand shapes for each sign. The Visual Graphics Group (VGG16) model is implemented for letter classification. Through extensive experiments conducted on real-world datasets, our algorithm demonstrates superior performance, outperforming other competitive algorithms. The system achieves an impressive accuracy rate of 96.05% in recognizing Arabic hand sign-based letters, establishing its credibility as a highly dependable solution for facilitating effective communication between the deaf and hearing communities. Further analysis of the confusion matrix and ROC curves reveals particularly strong performance with labels like "ain," "al," and "laam," indicating the model's ability to accurately classify these challenging categories with exceptional frequency.**

Keywords—Arabic Alphabet Sign Language, VGG16, Deep Learning, Deaf Culture

I. Introduction

Sign language is a profound mode of communication for the deaf, showcasing human resilience. Through intricate hand movements, thoughts and emotions are conveyed, forming unique linguistic expressions [1]. Deaf individuals historically used sign language for internal communication, resulting in distinct languages within their global communities. These languages serve as a bridge for deaf individuals to connect with each other and the broader world. However, communication challenges arise when sign language encounters understanding barriers between deaf and non-deaf individuals, leading to discrepancies and misunderstandings. Recognizing human hand movements has become crucial, and computer vision has played a pivotal role in advancing sign language recognition [2]. Deep learning, a subset of machine learning, has proven effective in deciphering the intricate nuances of sign language gestures [3]. Despite strides in unifying sign languages globally, Arabic Sign Language has received limited attention, especially in countries like Jordan, Saudi Arabia, and Egypt. Efforts to standardize and unify sign language have been made, but Arabic alphabet sign language recognition remains largely unexplored [4]. The research aims to develop a robust system for recognizing Arabic alphabet signs, adapting to linguistic nuances while integrating the VGG16 architecture. By addressing these challenges, the study aims to empower Arabic-speaking deaf communities with effective communication tools. The VGG16 model demonstrates high accuracy in classifying ArASL letters, validated through confusion matrices and ROC curves.

II. Related Work

Advancements in sign language recognition have predominantly focused on ASL and BSL, using diverse machine learning techniques to boost accuracy. However, there's a notable gap in literature concerning Arabic alphabet sign language, with limited research on its linguistic intricacies. Deep learning, especially CNNs like VGG16, has potential for image-based recognition tasks, but its application in Arabic sign language remains underexplored. Recent studies show growing interest in ArSL but often overlook the unique aspects of Arabic alphabet sign language, emphasizing general sign language gestures. The paper seeks to advance understanding of sign language recognition systems, specifically in the context of Arabic alphabet sign language. The findings are expected to inform the development of inclusive communication tools for the Arabic-speaking deaf community [5-6].

III. Methodology

The ArASL recognition pipeline involves meticulous dataset gathering and stringent preprocessing techniques to enhance data quality. It partitions the dataset into training and test sets for robust model evaluation, leveraging the potent VGG16 model as a feature extractor. Following further refinement, extracted features undergo Support Vector Classification (SVC) for effective analysis, aiming to facilitate accurate Arabic sign language recognition and communication between deaf and hearing communities [7].

A. Collecting and processing of data set

Our research utilizes the "ArASL_50K" dataset from Kaggle, comprising 50,000 instances capturing Arabic alphabet sign language nuances as represented in Figure 1. Variations in lighting, backgrounds, and spatial orientations were incorporated for real-world adaptability, serving as our research cornerstone. Preprocessing involves techniques such as image normalization, resizing, and grayscale conversion to ensure uniformity and data augmentation techniques like rotation and flipping to enhance model generalization. Resizing images to (224, 224) aligns with VGG16 requirements, contributing to efficiency and effectiveness in model training.

Fig. 1. Arabic Alphabet Sign Language ArASL data set

B. VGG16 archirecture

The architecture of our Arabic alphabet sign language recognition model is meticulously designed to leverage the strengths of both the VGG16 convolutional neural network (CNN) for feature extraction and Support Vector Machines (SVM) for classification [8]. VGG16 features a robust 16-layer structure, including 13 convolutional layers utilizing 3x3 kernels for feature extraction as shown in Figure 2a. These layers progressively increase filter count, extracting complex features. Figure 2b, aquire an Additionally, three fully connected layers integrate features for pattern recognition, crucial for discerning Arabic alphabet signs [5-6]. To tailor the model specifically for Arabic alphabet sign language, the final output layer is configured to match the number of classes corresponding to individual Arabic alphabet letters. This ensures that the model is capable of distinguishing between the distinct gestures associated with each letter of the alphabet.

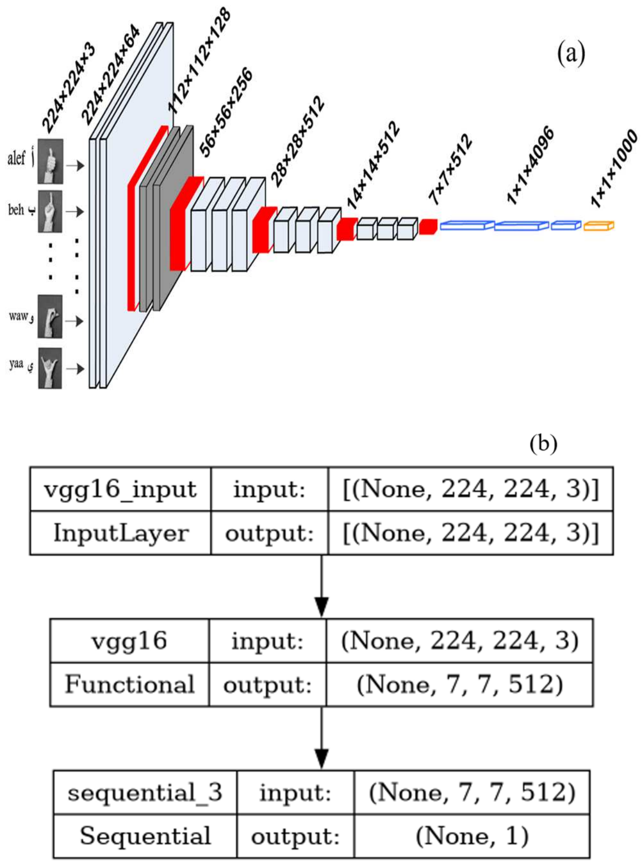

Fig. 2. (a) VGG16 architecture with ArASL as inputs and (b) Three fully connected layers of VGG16 pipeline

IV. Result and Discussion

The confusion matrix in Figure 3 illustrates the VGG16 model's ArASL classification into 35 categories. Despite an overall accuracy of 96.05%, scrutiny reveals areas for improvement, including seven False Negatives (FN) indicating missed positive cases. Certain categories like "ain," "al,"

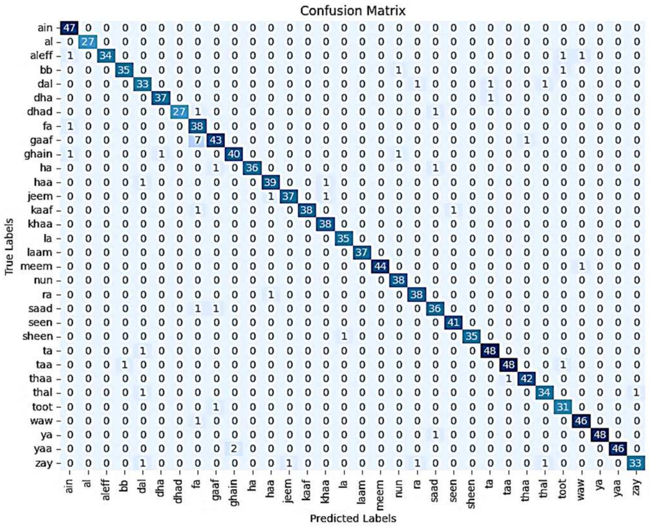

Fig. 3. Confusion matrix of

and "laam" perform well, while "dal" and "jeem" need improvement. Further analysis could address class imbalance and feature normalization for enhanced learning. The ROC curve, depicted in Figure 4, showcases the VGG16 model's discriminative ability across unique ArASL handshapes like "alif" and "jim." Analyzing the Area Under the Curve (AUC) provides quantitative insights, with lower AUC values indicating areas for refinement. Ultimately, the ROC curve serves as a roadmap to enhanced ArASL recognition, guiding targeted improvements. Table 1 summarizes VGG16 accuracy compared to other models on varied ArASL dataset sizes, achieving 96.05% accuracy on a 50,000-image dataset. Despite differences in dataset sizes and architectures, VGG16 demonstrates competitive accuracy against CNN and AlexNet, slightly surpassed by PCANET on a smaller dataset.

CONCLUSION

The ArASL system presented in this paper serves to bridge the communication gap between deaf and hearing communities by translating Arabic sign language gestures into text or speech with remarkable accuracy. Powered by the VGG16 model, ArASL analyzes hand geometry and distinct handshapes for accurate letter classification, achieving an impressive 96.05% accuracy. This success positions ArASL as a dependable solution for inclusive communication, surpassing competitive algorithms. Further insights from the confusion matrix and ROC curves highlight areas for refinement, with some labels demonstrating consistent correct classifications while others indicate room for improvement.

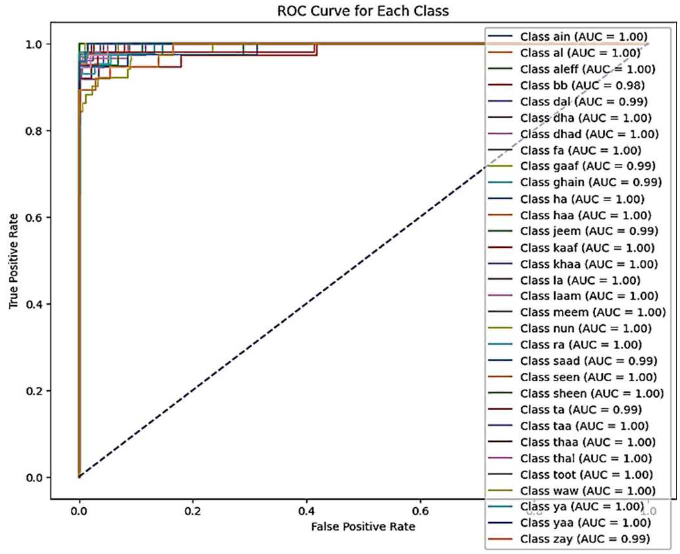

Fig. 4. ROC performance Curves of VGG16 models in ArASL Recognition

Table 1. comparison between VGG16 model and various models for ArASL recognition.

Model architecture	Accuracy (%)	Dataset size (image)	Ref.
VGG16	96.05%	50K	Present work
CNN	90%	31×125(augmented)	[3]
PCANET	99%	28×50(augmented)	[8]
AlexNET	94.81%	54 K	[9]

REFERENCES

[1] Ibrahim, Nada B., and Hala H. Zayed. "An automatic Arabic sign language recognition system (ArSLRS)," Journal of King Saud University-Computer and Information Sciences, no. 4, 470-477, 2018.

[2] Luqman, Hamzah, and Sabri A. Mahmoud. "Transform-based Arabic sign language recognition", Procedia Computer Science, no. 117, 2-9, 2017.

[3] Kamruzzaman, M. M. "Arabic sign language recognition and generating Arabic speech using convolutional neural network," Wireless Communications and Mobile Computing, no. 2020, 1-9, 2020.

[4] Maged, Heba, and Ahmed AbouEl-Farag. "Improving speaker identification system using discrete wavelet transform and AWGN," 2014 IEEE 5th International Conference on Software Engineering and Service Science, no. 1, 1171-1176, 2014.

[5] Tharwat, Gamal, and Belgacem Bouallegue. "Arabic sign language recognition system for alphabets using machine learning techniques," Journal of Electrical and Computer Engineering, no. 2021, 1-17, 2021.

[6] Nguyen, Hai Thanh, and Tran Thanh Dien. "Letter recognition in hand sign language with VGG-16," Intelligent Systems and Networks: Selected Articles from ICISN, Vietnam, no.1, 410-417, 2022.

[7] Aly, Saleh, Basma Osman, and Mahmoud Saber. "Arabic sign language fingerspelling recognition from depth and intensity images," 12th International Computer Engineering Conference (ICENCO), no. 1 99-104, 2016.

[8] Alsaadi, Zaran, and Abdelrahman Osman Elfaki. "A real time Arabic sign language alphabets (ArSLA) recognition model using deep learning architecture," Computers, no. 5, 78-84, 2022.

[9] Duwairi, Rehab Mustafa, and Zain Abdullah Halloush. "Automatic recognition of Arabic alphabets sign language using deep learning," International Journal of Electrical & Computer Engineering, no. 3, 1-12, 2022.