International Journal of Advanced Research in Education and Society
e-ISSN: 2682-8138 | Vol. 7, No. 6, 255-264, 2025
https://asianscholarsnetwork.com/asnet-journals

QuranVision: Intelligent Tajweed Detection and Recognition using Convolutional Neural Network

Mazura Mat Din[1*], Hariz Zamzuri[1], Noor Rasidah Ali[2], Siti Rafidah Muhamat Dawam[1], Shaifizat Mansor[1]

[1] Center of Computing Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Kedah Branch, Merbok 08400, Kedah, Malaysia
[2] Center of Mathematical Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Kedah Branch, Merbok 08400, Kedah, Malaysia

*Corresponding Author: mazuramd@uitm.edu.my

Received: 25 September 2025 | Accepted: 30 October 2025 | Published: 1 December 2025

DOI: https://doi.org/10.55057/ijares.2025.7.6.22

Abstract: *Tajweed is a set of essential rules for Quranic recitation. Ensuring correct pronunciation and efforts in preventing the errors that could alter its meaning is an obligation that needs to be fulfilled. While traditional face-to-face learning is highly effective, it is time-consuming and may not be accessible to all learners. To address this, QuranVision, a real-time Tajweed detection mobile application, is proposed, leveraging Convolutional Neural Networks (CNNs) for image-based Tajweed rule recognition. The study evaluates YOLOv11 and SSD MobileNet V2 FPNLite, focusing on detecting Meem Sakinah-based Tajweed rules (Idgham and Izhar Shafawee). Findings show that YOLOv11 outperforms SSD MobileNet V2 FPNLite, achieving a mean Average Precision (mAP) of 0.723 at IoU 0.5:0.95 with an input size of 416. Prototype testing confirms its effectiveness in detecting Tajweed rules in both black-and-white and color-coded Quranic texts, making it the preferred model for mobile deployment. The application allows users of all backgrounds and ages to self-learn Tajweed through real-time visual and audio feedback, eliminating reliance on continuous instructor supervision. Future enhancements include expanding the dataset, improving model robustness under varied lighting conditions, and integrating speech-based Tajweed correction to further enhance learning experiences.*

Keywords: Tajweed, Mobile Application, Artificial Intelligence, CNN

---

1. Introduction

The purpose of this study is to address the challenges associated with learning and applying Tajweed rules in Quranic recitation, particularly for non-native Arabic speakers and individuals with limited access to traditional learning methods. Tajweed rules are critical for ensuring the correct pronunciation and delivery of the Quranic text, as errors in recitation can alter the meaning and sanctity of the verses. According to Alagrami and Eljazzar (2020), mistakes such as mispronouncing words during Quranic recitation can significantly change the meaning of the text, emphasizing the importance of adhering to Tajweed rules. Traditional methods of teaching Tajweed, such as *Talaqqi Musyafahah*, which involve face-to-face interaction and lip-reading between teachers and students, are effective but time-consuming and dependent on the availability of qualified instructors (Omran et al., 2023). Furthermore, Omran et al. (2023) also highlighted that non-native speakers often face difficulties in distinguishing similar-sounding words and adhering to Tajweed rules, making the learning process more challenging.

This study is motivated by the growing need for an accessible, efficient, and automated solution to facilitate the understanding and application of Tajweed rules. With advancements in technology and the increasing use of mobile applications in education, integrating these tools into Quranic learning offers a promising alternative, According to Ridwan and Majid (2018) and Kassim et al. (2021), the ubiquity of gadgets in daily life suggests that mobile applications could serve as effective tools for Quranic education, especially for individuals with limited time. Additionally, various researchers in other areas noted that Convolutional Neural Networks (CNNs) are highly effective for object detection and image classification tasks, making them a suitable choice for developing a Tajweed detection application as supported in other area of research by Melinda et al. (2024); Altim et al. (2025); Okyere-Ghamfi et al. (2024); Ma et al. (2020) and Castanyer et al. (2021). The study aims to leverage CNNs to create a mobile application that detects and classifies Tajweed rules, providing learners with instant feedback and guidance.

The research adopts a structured methodology that begins with data collection, focusing on Quranic text images that feature Tajweed rules such as *Meem Sakinah* and its variations. Using these images, the study designs a CNN-based model to classify and detect Tajweed features. The developed model is then integrated into a mobile application prototype, which is the proposed QuranVision, enabling users to engage with an interactive and accessible platform for learning Tajweed rules. To ensure the reliability and effectiveness of the application, rigorous prototype testing is conducted, evaluating its accuracy, usability, and performance in real-world scenarios.

2. Literature Review

Tajweed is mandatory to ensuring accurate Quranic recitation, with errors potentially altering the text's meaning. Various studies emphasize the complexity of learning Tajweed rules, particularly for non-native Arabic speakers and those without access to traditional face-to-face teaching methods. Omran et al. (2023) implicate that the traditional *Talaqqi Musyafahah* method as effective but time-intensive, while Alagrami and Eljazzar (2020) underline the critical implications of mispronunciations. These challenges motivate the integration of technological solutions to enhance accessibility and learning efficiency (Ahmad et al., 2018).

In addressing this need, researchers have explored diverse methodologies for Quranic recitation and Tajweed rule detection. Optical Character Recognition (OCR) combined with thresholding algorithms, as employed by Meidi et al. (2021), achieved moderate success in identifying Quranic characters but faced limitations in accuracy. Similarly, Zuraiyah et al. (2020) leveraged pattern recognition and SURF algorithms for Tajweed segmentation, achieving high detection accuracy but requiring specialized computational resources. Rahman et al. (2018) explored automated Tajweed checking using audio features like Mel-Frequency Cepstral Coefficients (MFCC) and Hidden Markov Models (HMM), demonstrating effectiveness for pronunciation correction among children.

Convolutional Neural Networks (CNNs) have emerged as a transformative tool in image detection and classification. Kassim et al. (2021) and Noeman and Handayani (2020), highlights their advantages in automating feature extraction, making them ideal for complex tasks such as Tajweed rule identification. Among CNN-based object detection algorithms, the

Single Shot Multibox Detector (SSD) has proven particularly promising for balancing accuracy and speed (Liu et al., 2016). Comparative studies, such as those by Chen et al. (2018) and Xia et al. (2019), underscore SSD's efficiency in real-time applications, while retaining robust detection capabilities across diverse datasets.

Despite these advancements, several gaps persist. Existing models often struggle with generalizability across Quranic text styles and fail to address the unique challenges of detecting specific Tajweed rules, such as *Meem Sakinah*. Furthermore, prior studies have focused on isolated methods, leaving opportunities for integrating CNNs into mobile applications that offer user-friendly interfaces and real-time feedback. This study seeks to address these gaps by developing a CNN-based Tajweed detection prototype, leveraging SSD architecture to classify and detect Tajweed rules efficiently.

By critically synthesizing prior research, this review underscores the need for innovative, scalable solutions to support Quranic education. The proposed approach combines insights from multiple methodologies, integrating them into a comprehensive framework that addresses both the limitations and potential of existing systems. This ensures the study's contribution is both contextually relevant and practically impactful.

3. Research Methods

In order to complete the QuranVision application, it has to go through a few phases involving preliminary studies, knowledge acquisition which involves doing literature study in order to determine the domain, scope and suitable techniques for this application. Secondly, data acquisition and data preprocessing phase. Followed by design, development, prototype testing and documentation phase. The workflow of research phases is illustrated in Figure 1.

3.1 Preliminary Studies

In order to initiate the preliminary studies, this study has outlined three research questions and three research objectives. Below are the research questions and objectives:

Research Questions:

- How to identify the features of Tajweed rules?
- What is the suitable CNN architecture to detect and classify Tajweed rules?
- How to develop the prototype application for detecting Tajweed rules?

Research Objectives:

- To identify the features of Tajweed rules.
- To construct a model for Tajweed rules detection using CNN architecture.
- To develop a prototype application for Tajweed rules detection using CNN.

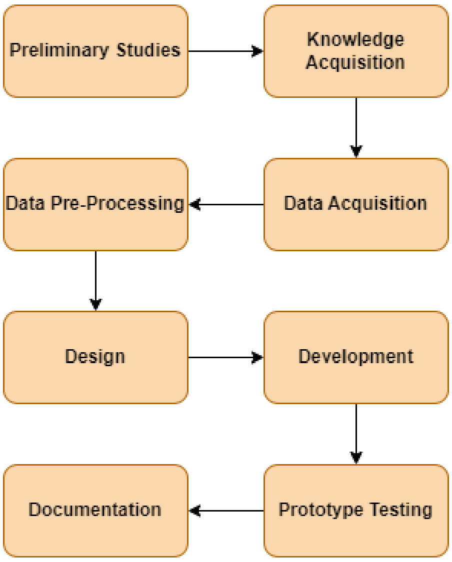


Figure 1: Research Phases of QuranVision Application

3.2 Data Acquisition Phase

As for the data acquisition phases of this study, the data collection focused on Quranic text images containing Tajweed rules, particularly *Meem Sakinah* and its associated rules (*Idgham Shafawee*, and *Izhar Shafawee*). Images were captured manually from physical and online Quranic texts using a Realme GT2 smartphone camera and snipping tools. A dataset of over 2500 images was compiled, ensuring accuracy and consistency through cross-referencing with Tajweed books and Quranic texts. Sample images are shown in Figure 2 which sources from digital Quran (left) and physical al-Quran (right).

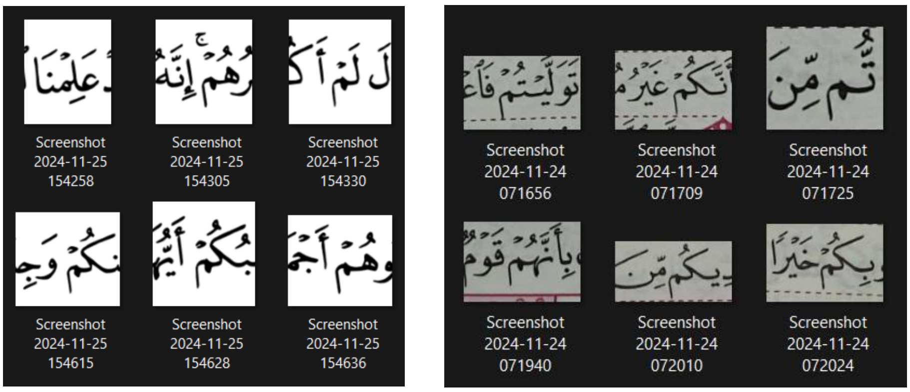


Figure 2: Example of Images from Online and Physical Quran

3.3 Pre-processing Phase

After data has been collected, the images are uploaded to Roboflow for data pre-processing phase. Annotation was performed using Roboflow, a tool designed to facilitate object detection tasks. Bounding boxes were manually created around specific Tajweed features, such as *Meem Sakinah* and its related rules (*Idgham Shafawee*, and *Izhar Shafawee*). These annotations provided precise location information for the targeted features, enabling the CNN model to learn their spatial characteristics effectively during training.

After annotations had been made to the images, data normalization was carried out to standardize the dataset and minimize variability caused by differing image properties. All images were resized to a uniform dimension of 640x640 pixels. This resizing ensured consistency across the dataset while reducing the computational load during model training.

Additionally, the images were converted to grayscales to simplify their representation, as color information was not essential for detecting Tajweed rules. This step significantly reduced the dimensionality of the data, making the model training process more efficient without compromising detection accuracy.

Data augmentation also has been made to the original annotated data, in order to increase the variability in the dataset images so that the later-trained model can detect the rules even in different conditions and scenarios. Augmentation that has been made are as said before, gray scaling the images, adjusting the hue of the images, between -10% and 10% and lastly, adjusting the brightness from -12% to 12%. Figure 3 below shows the annotation phase in Roboflow and the normalization and augmentation made to the dataset. Lastly, in data pre-processing phase, the dataset is balanced into 8:1:1 ratio which stands for 80% for training, 10% for testing and 10% for validation.

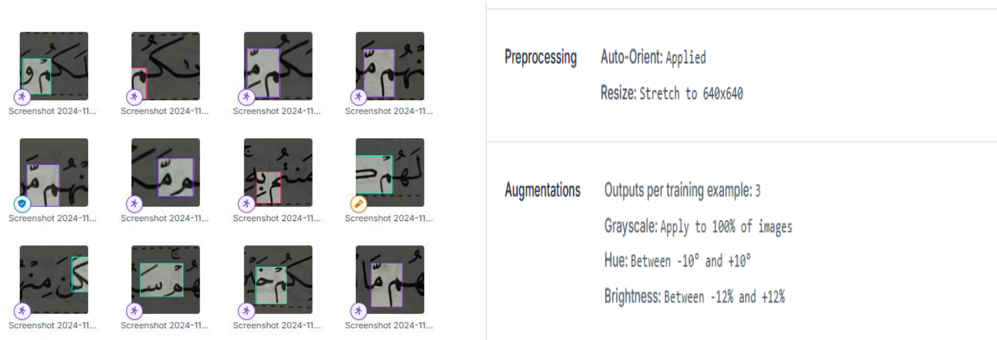


Figure 3: Annotation and Augmentation in Roboflow Tools.

The study evaluated two object detection architectures: YOLOv11 and SSD MobileNet V2 FPNLite, both of which are well-suited for mobile applications due to their efficiency and accuracy. YOLOv11 was selected for its advanced CSP2SA (Convolutional block with Parallel Spatial Attention) and SPPF (Spatial Pyramid Pooling – Fast) modules, which enhance feature extraction. Hyperparameters were fine-tuned across different experiments, including variations in input size (128, 320, 416 for YOLOv11 and 224, 320, 416, 512 for SSD MobileNet V2), batch sizes (16, 32, 64 for YOLOv11 and 8, 16, 32 for SSD MobileNet V2), and epochs (50 and 100). Training was conducted in Visual Studio Code using PyTorch and the Ultralytics YOLOv11 framework, leveraging a GPU-based environment for accelerated learning.

3.4 Design Phase

Following model selection, the best-performing YOLOv11 model was converted to TensorFlow Lite (TFLite) for mobile deployment using Android Studio with Flutter/Dart. Several optimizations were implemented to ensure efficient real-time inference, including model quantization techniques to reduce the model size while maintaining accuracy and on-device inference optimizations to enhance processing speed. Additionally, the mobile application's user interface was designed using Figma to create an intuitive user experience as shown in Figure 4.

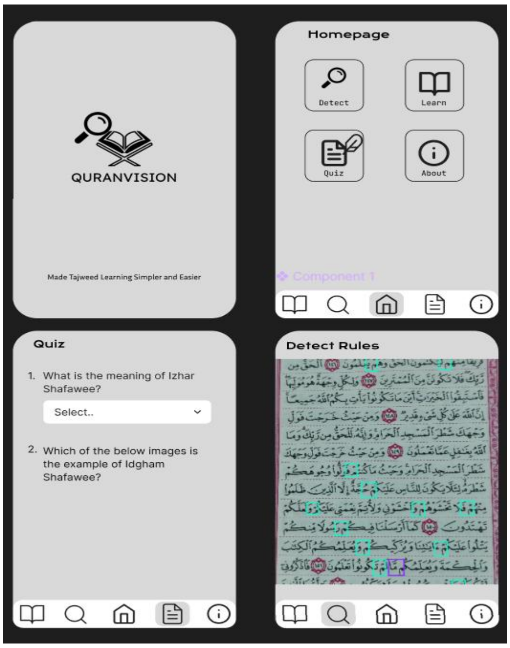
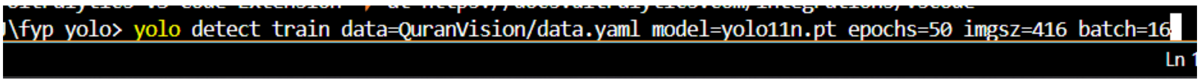


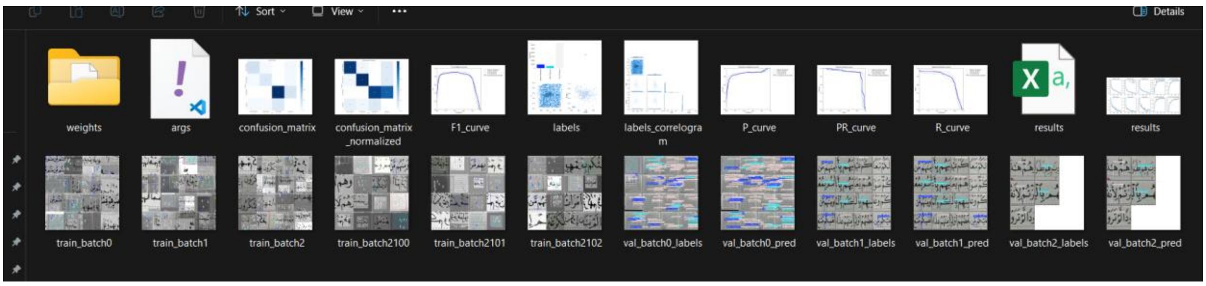
Figure 4: GUI Design using Figma.

3.5 Development Phase

The next phase of this research is the development phase which is done through 2 types of platforms namely Visual Studio Code for the model development and Android Studio for the mobile application development. For the model development, first thing that needs to be done is creating a clean python environment for installing dependencies needed to run the model. Next step is installing pytorch, torchvision, and ultralytics. After environment is ready, dataset is downloaded from Roboflow. Before start training the data, the data.yaml file needs to be edited first in order to update the train, test and validation folder location according to their locations on the local machine. After that, the training process starts by writing the command in python terminal in VS Code. Upon completion of data training, the result is saved in the file system for best model, graph curves, train and validation images, training epoch results in .csv format. Figure 5 shows the training code and results folder.



(a)



(b)

Figure 5: Training Code (a) and Results Folder (b).

The mobile application development is conducted in Android Studio using Flutter:Dart language. This enables the QuranVision mobile application to run on both Android and IOS devices. In this process, the integration between the produced image detection model and the mobile application is conducted. Figure 6 shows the environment of the QuranVision mobile application development.

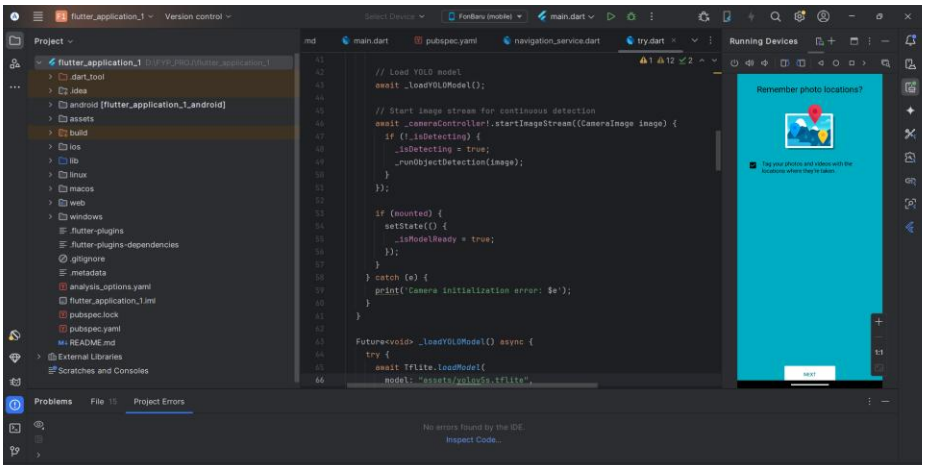


Figure 6: Android Studio environment for developing mobile application.

3.6 Prototype Testing Phase

Prototype testing was conducted to evaluate the application's reliability and accuracy in detecting Tajweed rules. The testing phase included assessing detection performance on different Quranic texts, including printed, handwritten, black-and-white, and color-coded formats. The application's speed and accuracy in real-time detection scenarios were also examined. Furthermore, failure cases, such as misclassification of overlapping Tajweed symbols, were analysed to identify areas for further improvement.

4. Results and Discussions

Findings in this research was based on hyperparameters settings, which are decided upon executing the built-up configurations. Those hyperparameters are fine-tuned as tabulated in Table 1 and 2, in order to achieve the highest level of performance and accuracy fuelled by 2500 data involved to complete the whole experiments.

Table 1: Parameters for YOLOv11 Experiments

Parameters	Experiment 1			Experiment 2		
Dataset size	4.5k+			4.5k+		
Learning rate	0.01			0.01		
Train: Valid: Test Ratio	8:1:1			8:1:1		
Input size	128	320	416	320		
Epoch	50			50		100
Batch	16			16	32	64

Table 2: Parameters for SSD MobileNet V2 FPNLite Experiments

Parameters	Experiment 1				Experiment 2		
Dataset size	4.5k+				4.5k+		
Learning rate	0.01				0.01		
Train: Valid Ratio	8:2				8:2		
Input size	224	320	416	512	320		
Epoch	50				50		100
Batch	8				8	16	32

The study evaluated the performance of two object detection models, YOLOv11 and SSD MobileNet V2 FPNLite, for Tajweed rule detection. The experiments focused on optimizing hyperparameters such as input size, batch size, and the number of training epochs. Results showed that YOLOv11 consistently outperformed SSD MobileNet V2 FPNLite in accuracy, with the best results achieved using an input size of 416, reaching a mean average precision (mAP) of 0.723 at IoU 0.5:0.95. The model with an input size of 320 also performed well, balancing detection accuracy and training time, making it a viable option for real-time applications. SSD MobileNet V2 FPNLite, despite improvements with larger input sizes, remained less effective, with its best configuration (input size 416) achieving a lower mAP of 0.646. The study highlighted that increasing batch size reduced training time but did not significantly enhance accuracy, suggesting that a batch size of 16 for YOLOv11 and 8 for SSD was the most effective setting. Figure 7 shows the visualizations of both models in experiment 1 which shows that YOLO outperforms SSD model in both mAP at threshold 0.5:0.95 (a) and in terms of recall values (b). Although when it comes to mAP values at IoU threshold of 0.5 (c), SSD and YOLO are almost on par with each other. This finding shows that YOLO is much more suitable for integration with mobile applications in this study.

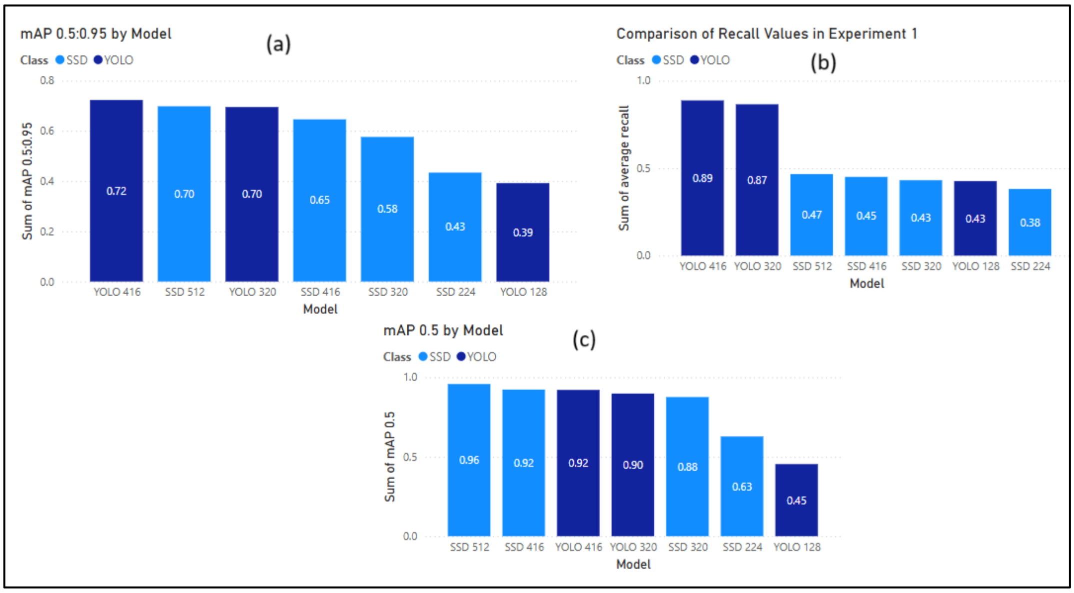


Figure 7: Comparison of Models' Performance.

Prototype testing further validated YOLOv11's superiority in real-world applications. The model with an input size of 416 demonstrated the highest accuracy, effectively detecting Tajweed rules across different Quranic text formats, including full pages and isolated rule snippets. However, smaller input sizes, particularly 128, struggled with larger or more complex images. A key limitation identified was the challenge of detecting rules in close-up views, suggesting the need for additional training data for better localization. Overall, YOLOv11 was deemed the most suitable model for integration into the QuranVision mobile application due to its superior performance, making it the recommended choice for real-time Tajweed detection. Figure 8 shows the final output of the mobile application after the integration with TFLite model where (a) shows the homepage, (b) is the detection module page, (c) is the learning module and (d) is the quiz module of the application.

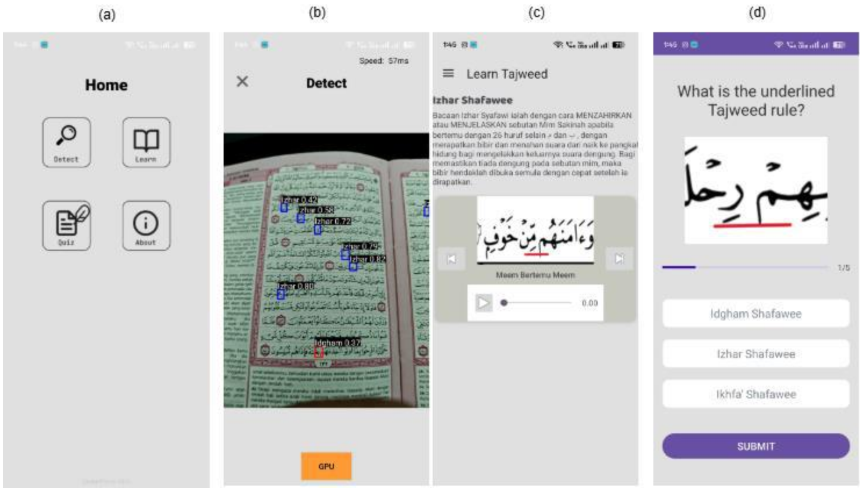


Figure 8: Overview of QuranVision Mobile Application

5. Conclusion

The findings of this study confirm that QuranVision offers a scalable and effective solution for self-learning Tajweed. By providing real-time visual and audio feedback, the application reduces reliance on instructor-led sessions, making Tajweed learning more accessible for users of all backgrounds and ages. The integration of deep learning techniques into Quranic education demonstrates the potential of AI-driven applications in enhancing religious studies.

Future enhancements for QuranVision include expanding the dataset to incorporate different Quranic scripts and handwritten texts, integrating speech-based Tajweed correction to provide pronunciation feedback, and improving model generalization through adversarial data augmentation. These improvements will further enhance the application's accuracy and usability, ensuring a more effective learning experience for users. Ultimately, this study highlights the feasibility of using deep learning in Quranic education, paving the way for innovative AI-driven learning applications.

Acknowledgement
We would like to acknowledge Universiti Teknologi MARA for their support especially Jawatankuasa Penyelidikan Negeri, UiTM Cawangan Kedah for funding under Tabung Penyelidikan Am.

Conflict of Interest Statement
The authors declare that there is no conflict of interest regarding the publication of this study.

References

Ahmad, F., Yahya, S. Z., Saad, Z., & Ahmad, A. R. (2018). Tajweed Classification Using Artificial Neural Network. *2018 International Conference on Smart Applications, Communications and Networking, SmartNets 2018*, 25–28. https://doi.org/10.1109/SMARTNETS.2018.8707394

Altim, M. Z., Basalamah, A., & Kasman. (2025). Implementation of Convolutional Neural Network (CNN) Based on Mobile Application for Rice Quality Determination. *Protokol: Journal of Engineering.*

Alagrami, A. M., & Eljazzar, M. M. (2020). SMARTAJWEED Automatic Recognition of Arabic Quranic Recitation Rules, 145–152. https://doi.org/10.5121/csit.2020.101812

Castanyer, R. C., Martínez-Fernández, S., & Franch, X. (2021, May). Integration of convolutional neural networks in mobile applications. In *2021 IEEE/ACM 1st Workshop on AI Engineering-Software Engineering for AI (WAIN)* (pp. 27-34). IEEE. https://doi.org/10.1109/WAIN52551.2021.00010

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (n.d.). *SSD: Single Shot MultiBox Detector*. https://github.com/weiliu89/caffe/tree/ssd

Kassim, I. A. A., Ab Rahman, M., Ab Rahman, T., & Mohd Muji, S. Z. (2021). Development of Automated Tajweed Checking System for Children in Learning Quran. *Evolution in Electrical and Electronic Engineering*, *2*(1), 165–176. http://publisher.uthm.edu.my/periodicals/index.php/eeee

Ma, W., Wang, X., & Yu, J. (2020). A lightweight feature fusion single shot multibox detector for garbage detection. *IEEE Access*, *8*, 188577–188586. https://doi.org/10.1109/ACCESS.2020.3031990

Meidi, D., Gerhana, Y., Atmadja, A., & Dauni, P. (2021, January). Implementation of OCR (Optical Character Recognition) Using Otsu Threshold Method for Detecting Tajweed Qur'an. In *Proceedings of the 1st International Conference on Islam, Science and Technology, ICONISTECH 2019, 11-12 July 2019, Bandung, Indonesia.*

Melinda, M., Arnia, F., Yafi, A., Andryani, N. A. C., & Enriko, I. K. A. (2024). Design and Implementation of mobile application for CNN-based EEG identification of Autism Spectrum Disorder. *International Journal on Advanced Science Engineering and Information Technology, 14*, 57.

Noeman, A., & Handayani, D. (2020). Detection of Mad Lazim Harfi Musyba Images Uses Convolutional Neural Network. *IOP Conference Series: Materials Science and Engineering*, *771*(1). https://doi.org/10.1088/1757-899X/771/1/012030

Okyere-Gyamfi, S., Akoto-Adjepong, V., Adu, K., Ayidzoe, M. A., Appiah, O., Appiahene, P., Mensah, P. K., Opoku, M., Bawah, F. U., Awarayi, N. S., Boateng, S., Nimbe, P., & Adekoya, A. F. (2024). Deployment of mobile application using a novel CNN model for the detection of COVID-19 thoracic disease. *Scientific African, 26*, e02432. https://doi.org/10.1016/j.sciaf.2024.e02432

Omran, D., Fawzi, S., & Kandil, A. (2023). Automatic Detection of Some Tajweed Rules. In *2023 20th Learning and Technology Conference (L&T)* (pp. 157-160). IEEE. https://doi.org/10.1109/LT58159.2023.10092350

Rahman, M. A., Kassim, I. A. A., Rahman, T. A., & Muji, S. Z. M. (2021). Development of automated Tajweed checking system for children in learning Quran. *Evolution in Electrical and Electronic Engineering*, *2*(1), 165-176.

Ridwan, T., & Majid, N. (2019). Development System for Recognize Tajweed in Qur'an using Automatic Speech Recognition. *Proceedings of the 1st International Conference on Science and Technology for an Internet of Things*. https://doi.org/10.4108/eai.19-10-2018.2281368

Syaifullah, M., Jannah, M., Fradila, N., Nigrum, P. P., Santi, S., & Nasution, W. (2022). The Implementation of the Science of Tajwid in Learning of the Qu'ran with the Tajwod Wheel Method to Develop the Reading of the Qu'ran in Ibtidaiyah Madrasah. *Widyagogik: Jurnal Pendidikan dan Pembelajaran Sekolah Dasar*, *9*(2), 271-280.

Xia, X., Shi, X., Lu, Q., & Fan, C. (2019). Application of Single Shot MultiBox Detector in Logistics Safety Testing. *IOP Conference Series: Materials Science and Engineering*, *569*(2). https://doi.org/10.1088/1757-899X/569/2/022002

Zuraiyah, T. A., Madenda, S., Salim, R. A., & Noviana, R. (2020). Tajweed Segmentation Using Pattern Recognition, Extraction and SURF descriptor Algorithms. *IOP Conference Series: Materials Science and Engineering*, *846*(1). https://doi.org/10.1088/1757-899X/846/1/012022