Audio-based Quran Verses Recognition Using Deep Learning Models

Adham Ayman(*Digital Media Engineering*) and **Prof. Mohammed Salem** (*Digital Media Engineering*)

The German university in Cairo

adham.almadany@student.guc.edu.eg mohammed.salem@guc.edu.eg

Introduction

The Holy Quran is the central religious text of Islam. It is a collection of the Holy Prophet Muhammad's spoken revelations during a twenty-three-year period. The Holy Quran is divided into 114 Chapters, each of which has its own set of verses. The Chapters (called Surah in Arabic) which consists of 6236 verses. Since Quranic verses are frequently used in a lot of speeches and sometimes cited in court rooms for reference, so people should be able to use a system that would listen to the verse and recognize the original text from the Holy Book
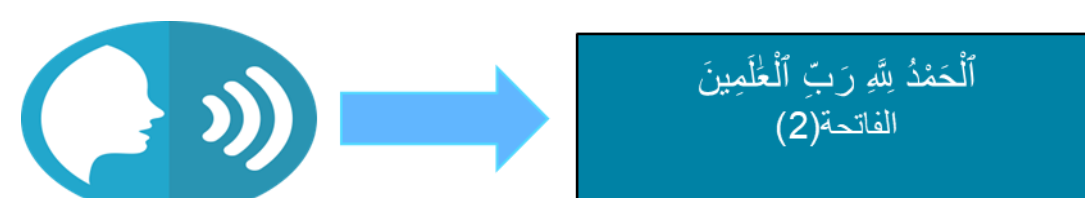
Literature Review

Zaidi Razak et al. [2] discussed the different methods and approaches to find the most acceptable method to be used in the Quran verse recognition system. They discussed the system that has been established by Hassan Tabbal et al. [3] that was built using a pre-emphasis filter for preprocessing, MFCC for feature extraction, and Hidden Markov Model (HMM) [4] for classification and recognition techniques. Using these methods, they reached a performance of 85%–92% . Another system that has been discussed, which was introduced by A. Youssef et al[5], was established using MFCC for feature extraction and Hidden Markov Model (HMM) for classification and recognition techniques. They reached a performance of 90.2% using these techniques. The system introduced by A. M. Ahmad et al.[6] also been discussed. It uses digital filtering for preprocessing, MFCC and LPCC for feature extraction, and a recurrent neural network for classification and recognition techniques. They achieved a performance of 95.9% using MFCC and 94.5%–99.3% using LPCC. S.K. Podder et al. [7] also proposed a system that was built using LPC for feature extraction and Vector Quantization (VQ) and Hidden Markov Model (HMM) were used for classification and recognition techniques. They reached a performance of 62%–96% Using these techniques. After those mentioned researches and others had been discussed the methods was decided to be used in this topic was MFCC for feature extraction and using HMM for training and testing phase.

Mohammad Mosleh et al [19] did a study on speaker recognition algorithms based on feature selection and classification methods. VQ (Vector Quantization), GMM (GaussianMixture Model), SVM (Support Vector Machines), and mixture methods of GMM and SVM were among the classification approaches. The evolutionary algorithm, Ant colony, and Relief algorithms were among the algorithms used for feature selection and speaker modeling. Using mixed approaches to classify issues has raised multi-class complexity and had little effect on recognizing a large number of speakers, according to studies. Then, by combining feature selection approaches, we can employ different classifiers to improve speaker recognition efficiency.

Methodology

Using a pre-trained model offered by Tark-eldeeb [20] along with DeepSpeech [1] we built a system that takes an input an audio file of recitation of any verse of the Holy Quran and using the DeepSpeech model we convert the speech in this audio file to Arabic words with "Tashkeel" and using this output words we are trying to build a matching algorithm to find the best match verse that include these words .

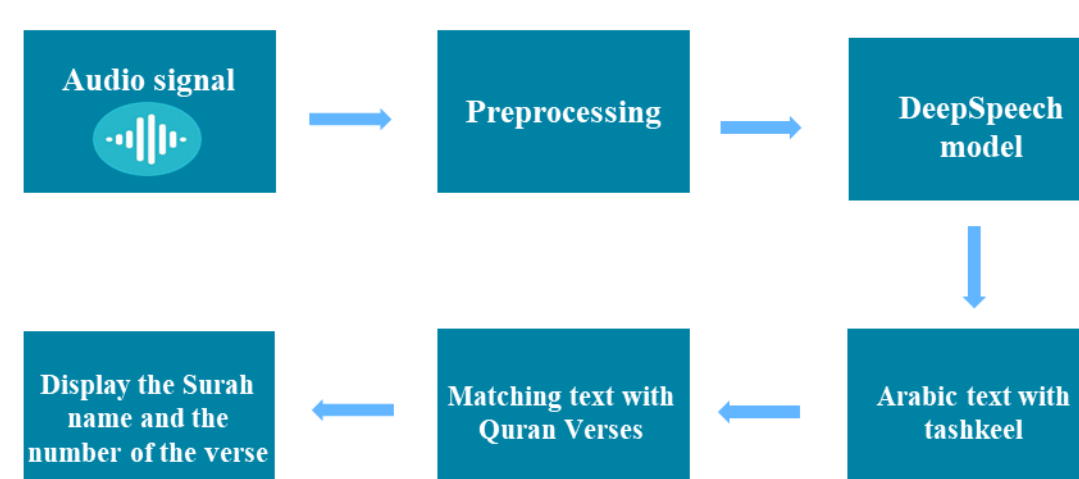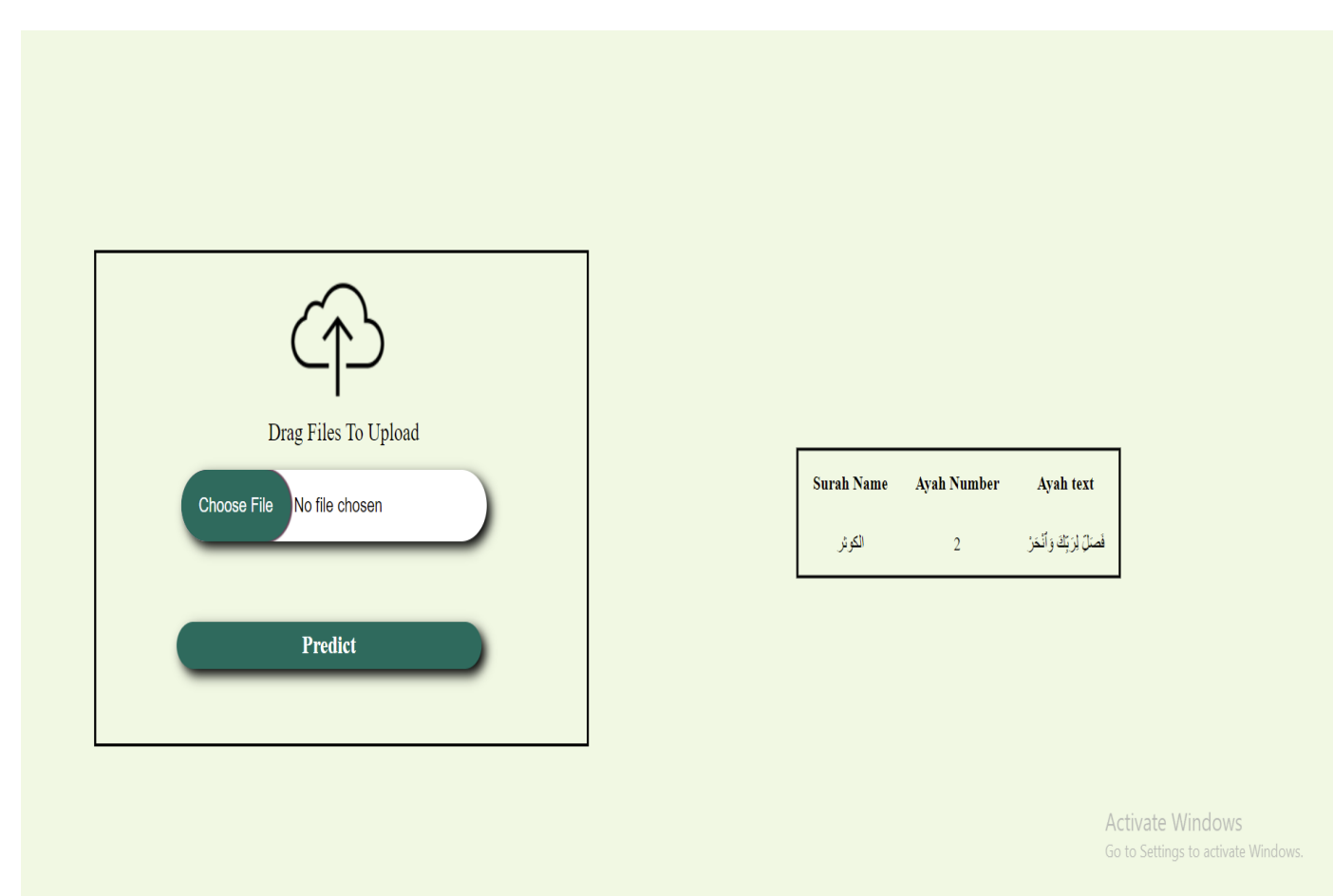A web interface was built to make the user interaction easier with the system .

Fig.1 : Block Diagram of the Implementation process

Results

Build 2 web pages to interact with the system :

(A) web page to upload the wav file of the Quranic verse and it outputs the number of the verse and the name of the SURAH along with the original text of the verse

(B) Web page to take user recordings by clicking on the green button and then click again to stop recording and the outputs the number of the verse and the name of the SURAH along with the original text of the verse
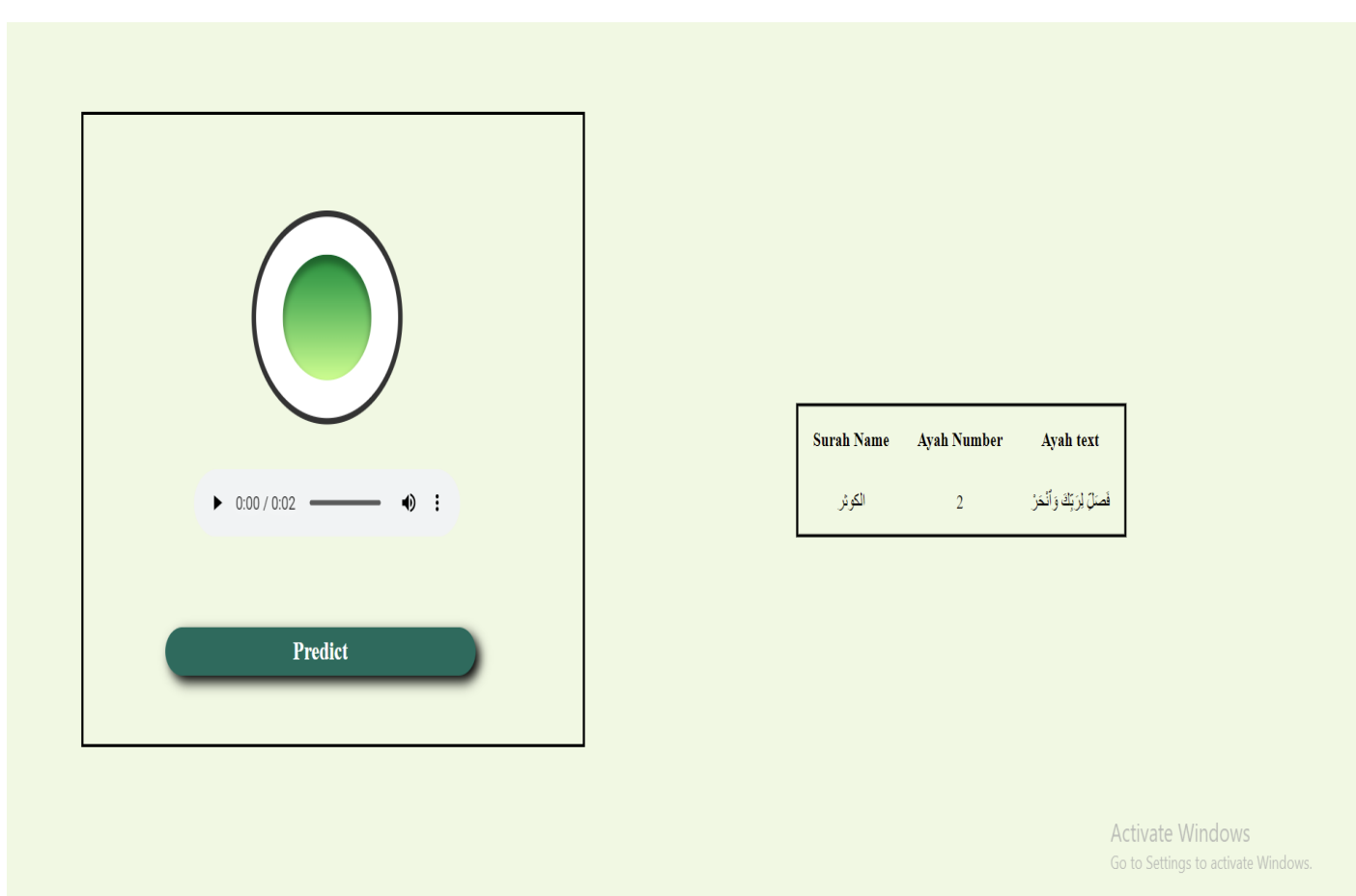
Conclusion

The aim of this project is to build a Quran verses recognition model with the highest accuracy possible. Using DeepSpeech and a pre-trained model we achieved a very good results in recognizing Arabic words with the right "Tashkeel" and the match those output words to the best matching verse of the Holy Quran .

Also a web interface was built to make the interaction of the users with the system more easy .

There is two ways to the user to interact with the system the first one is by uploading a wave file of the desired verse which he want to search for it's number and SURAH name the other way is to record a new audio file with the user's voice to search for a specific verse .

References

1. "Deepspeech." https://deepspeech.readthedocs.io/en/r0.9/.
2. Z. Razak, N. J. Ibrahim, M. Idris, E. Tamil, M. Yakub, Z. Yusoff, and N. Rahman,"Quranic verse recitation recognition module for support in j-qaf learning: A review," International Journal of Computer Science and Network Security (IJCSNS), vol. 8, pp. 207–216, 08 2008.
3. H. Tabbal, W. El Falou, and B. Monla, "Analysis and implementation of a "quranic" verses delimitation system in audio files using speech recognition techniques," in 2006 2nd International Conference on Information Communication Technologies, vol. 2, pp. 2979–2984, 2006.
4. L. Rabiner and B. Juang, "An introduction to hidden markov models," IEEE ASSP Magazine, vol. 3, no. 1, pp. 4–16, 1986.
5. O. E. Amr Youssef, "An arabic tts system based on the ibm trainable speech synthesizer," Le traitement automatique de l'arabe, vol. 3, no. 1, 2004.
6. A. Ahmad, S. Ismail, and D. Samaon, "Recurrent neural network with backpropagation through time for speech recognition," in IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004., vol. 1, pp. 98–102 vol.1, 2004.
7. M. Ostendorf, V. Digalakis, and S. Kimball, "From hmm's to segment models: A unified view of stochastic modeling for speech recognition," Speech and Audio Processing, IEEE Transactions on, vol. 4, pp. 360 – 378, 10 1996.
8. A. Mohammed, M. S. Sunar, and M. S. Salam, "Quranic verses verification using speech recognition techniques," Jurnal Teknologi, vol. 73, 03 2015.
9. B. Putra, B. Tris Atmaja, and D. Prananto, "Developing speech recognition system for quranic verse recitation learning software," IJID (International Journal on Informatics for Development), vol. 1, p. 14, 12 2012.
10. M. Y. El Amrani, M. H. Rahman, M. R. Wahiddin, and A. Shah, "Building cmu sphinx language model for the holy quran using simplified arabic phonemes," Egyptian Informatics Journal, vol. 17, no. 3, pp. 305–314, 2016.
11. W.Walker, P. Lamere, P. Kwok, B. Raj, R. Singh, E. Gouvea, P.Wolf, and J. W"olfel, "Sphinx-4: A flexible open source framework for speech recognition," Sun Microsystems, 12 2004.
12. J.-C. Wang, J.-F. Wang, C.-B. Lin, K.-T. Jian, and W.-H. Kuok, "Content-based audio classification using support vector machines and independent component analysis," vol. 4, pp. 157–160, 01 2006.
13. G. Kour and N. Mehan, "Music genre classification using mfcc, svm and bpnn," International Journal of Computer Applications, vol. 112, pp. 12–14, 2015.
14. J. Lee, T. Kim, J. Park, and J. Nam, "Raw waveform-based audio classification using sample-level cnn architectures," 12 2017.
15. E. Law, K. West, M. Mandel, M. Bay, and J. Downie, "Evaluation of algorithms using games: The case of music tagging.," pp. 387–392, 01 2009.
16. P. Warden, "Speech commands: A dataset for limited-vocabulary speech recognition," 04 2018.
17. A. Mesaros, T. Heittola, A. Diment, B. Elizalde, A. Shah, E. Vincent, B. Raj, and T. Virtanen, "Dcase 2017 challenge setup: Tasks, datasets and baseline system," 11 2017.
18. J. Das, A. Ghosh, A. Pal, S. Dutta, and A. Chakrabarty, "Urban sound classification using convolutional neural network and long short term memory based on multiple features," 11 2020.
19. M. M. Hasti Baharipour, Mohammad EbrahimShiri, "Feature selection and classification techniques for speaker recognition," International journal of Computer Science Network Solutions, vol. 7, 01 2014.
20. https://github.com/tarekeldeeb/DeepSpeech-Quran