Journal of Technology Research (JTR)
Volume 3, Issue 2, (2025), pp 126-132, ISSN 3005-639X

Hybrid MFCC-LSTM Models for Automatic Evaluation of Three Primary Tajweed Rules in Holy Quranic Recitation

Abdalla A. Elmasallati[1], Faraj S. Alahjal [*1], Abdelrahman M. Essa[1]

[1]Department of Electromechanical Engineering, College of Industrial Technology, Misrata, Libya.

*Corresponding author email: faraj.saleh@cit.edu.ly

Received: 30-10-2025 | Accepted: 12-12-2025 | Available online: 15-12-2025 | DOI:10.26629/jtr.2025.15

ABSTRACT

This research aims to develop an Artificial Intelligence (AI) system for the automatic evaluation and correction of Holy Quran recitation, with the goal of facilitating the learning process of Tajweed rules and ensuring the accuracy of vocal performance. The proposed system relies on processing audio signals by extracting precise features using the Mel-Frequency Cepstral Coefficients (MFCC) [1] technique, and analyzing them via a deep learning model based on Long Short-Term Memory (LSTM) networks [2]. The public QDAT dataset [3] was utilized, containing over 1,500 audio clips focused on three primary Tajweed rules: Al-Madd al-Munfasil (Separate Prolongation), Al-Ghunnah al-Mushaddadah (Heavy Nasalization), and Al-Ikhfaa (Concealment). The data was split into 70% for training, 15% for validation, and 15% for testing.The independent LSTM models achieved varied performance across the rules: the Al-Madd al-Munfasil model recorded the highest overall accuracy at 77.0%; the Al-Ikhfaa model recorded an accuracy of 66.1%; and the Al-Ghunnah al-Mushaddadah model recorded the lowest overall accuracy at 59.4%, but distinguished itself with the highest Precision level, reaching 98.7%. These results confirm the effectiveness of LSTM networks in complex Tajweed-specific speech recognition tasks, highlighting the importance of analyzing detailed performance metrics such as Precision and Recall to reveal the strengths and weaknesses in the models. As an additional contribution to support future research in this field, the researchers collected and developed a new dataset containing 1,500 recordings of the first three verses of Surat Al-Mutaffifin, which is publicly available via a GitHub repository.

Keywords: Quran Recitation, Tajweed, LSTM Networks, Mel-Frequency Cepstral Coefficients (MFCC), Artificial Intelligence**.**

نماذج MFCC–LSTM الهجين للتقييم الآلي لثلاثة أحكام تجويد أساسية في تلاوة القرآن الكريم

عبدالله علي المسلاتي[1]، فرج صالح الأحجل[1]، عبدالرحمن مفتاح عيسى[1]

[1]قسم الهندسة الإلكترونية (شعبة تقنية معلومات)، كلية التقنية الصناعية، مصراتة، ليبيا.

ملخـــــص البحـــــث

يهدف هذا البحث إلى تطوير نظام ذكاء اصطناعي لتقييم وتصحيح تلاوة القرآن الكريم آلياً، بهدف تيسير عملية تعلم أحكام التجويد وضمان دقة الأداء الصوتي. يعتمد النظام المقترح على معالجة الإشارات الصوتية عبر استخلاص سمات دقيقة باستخدام تقنية معاملات التردد الميلية (MFCC) [1] ، وتحليلها بواسطة نموذج التعلم العميق لشبكات الذاكرة طويلة

"Articles published in *J Technol Res* are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License."

وقصيرة المدى .[2] (LSTM) تم استخدام مجموعة البيانات العامة QDAT [3]، التي تحتوي على أكثر من 1500 مقطع صوتي مركزة على ثلاثة أحكام تجويدية رئيسية هي :المد المنفصل، والغنة المشددة، والإخفاء .تم تقسيم البيانات بنسبة 70% للتدريب، و 15% للتحقق، و 15% للاختبار. وقد حققت نماذج LSTMالمستقلة أداءً متبايناً بين الأحكام، حيث سجل نموذج المد المنفصل أعلى دقة كلية بلغت 77.0%، بينما سجل نموذج الإخفاء دقة 66.1%، وسجل نموذج الغنة المشددة أدنى دقة إجمالية بلغت 59.4%، ولكنه تميز بأعلى مستوى دقة (Precision) وصل إلى .98.7% وتؤكد هذه النتائج على فعالية شبكات LSTM في مهام التعرف الصوتي المعقدة الخاصة بالتجويد، وتبرز أهمية تحليل مقاييس الأداء التفصيلية كالدقة والاستدعاء للكشف عن نقاط القوة والضعف في النماذج. وكمساهمة إضافية لدعم الأبحاث المستقبلية في هذا المجال، قام الباحثون بجمع وتطوير مجموعة بيانات جديدة تحتوي على 1500 تسجيلاً للآيات الثلاث الأولى من سورة المطففين، وهي متاحة للجمهور عبر مستودع GitHub .

الكلمات الدالة: تلاوة القرآن، تجويد، شبكات LSTM، معاملات التردد الميلية (MFCC)، الذكاء الاصطناعي.

1. المقدمة

تعتبر تلاوة القرآن الكريم وفق قواعد التجويد أمراً أساسياً لضمان صحة الأداء. ونظراً للحاجة الملحة إلى أدوات تقييم فورية وموثوقة، أصبح الذكاء الاصطناعي يلعب دوراً محورياً في توفير حلول تعليمية مساعدة. يتناول هذا البحث تطوير نظام آلي متقدم يعتمد على التعلم العميق للكشف عن الأخطاء في التلاوة.

تناولت الأبحاث السابقة تقنيات تقليدية للتعرف على التجويد، مثل نماذج ماركوف المخفية[4] (HMM) ، وآلات دعم المتجهات[5] (SVM) ، إلا أن هذه النماذج غالبًا ما تفتقر إلى الكفاءة في التعامل مع الاعتماديات الزمنية الطويلة [6]التي تتطلبها أحكام التجويد. يتميز هذا البحث بالتركيز على شبكات [2] LSTM، والتي تتفوق في معالجة البيانات المتسلسلة زمنياً .[7]

يهدف هذا البحث إلى تطوير نماذج LSTM لتقييم أخطاء تلاوة أحكام المد، والإخفاء، والغنة المشددة آلياً، وإظهار فعاليتها في نمذجة الاعتماديات الزمنية لهذه الأحكام باستخدام سمات.MFCC

2. الأعمال السابقة

شهد مجال التعرف الآلي على التلاوة القرآنية وتقييم أحكام التجويد تطورات مستمرة، بهدف توفير أدوات تعليمية مساعدة للمسلمين حول العالم. تقليدياً، اعتمدت معظم الأبحاث الرائدة على الأساليب الإحصائية التقليدية مثل نماذج ماركوف المخفية (Hidden Markov Models – HMMs)لنمذجة الخصائص الصوتية للفظ القرآني. على سبيل المثال، تناولت دراسات سابقة استخدام HMMs في أنظمة التعرُّف على النطق (ASR)أو نمذجة أحكام تجويدية محددة مثل المدود والغُنن، معتمدة على استخلاص المميزات الصوتية التقليدية مثل معاملات التردد الميلية .(MFCCs) في هذا السياق، ركزت بعض الأعمال على تقييم نطق الحروف المخارج، بينما تناولت أعمال أخرى تقييم دقة التلاوة الكلية.

1.2 التعرف الآلي على الكلام القرآني

انتقلت الأبحاث الحديثة نحو تطوير أنظمة تعرُّف شاملة (End-to-End)ذات مفردات كبيرة. على سبيل المثال، قدم الطنطاوي وآخرون [8] في عام 2021 مقاربة تعتمد على التعلم العميق لتحقيق معدلات خطأ منخفضة (Word Error Rate – WER) في التعرف على التلاوات القرآنية. كما اقترح الحريري وجلاد [9] (2023) نموذجاً شاملاً يعتمد على مُشفِّر CNN– Bidirectional GRU مع استخدام تصنيف التسلسل الزمني (CTC)، مما عزز من دقة التعرف على الكلام القرآني باستخدام مجموعات بيانات عامة .(Ar-DAD) كذلك، طور الفضلي وآخرون [10] في عام 2024 نموذج qArI الهجين الذي يجمع بين CTC/Attention وبنية Bidirectional LSTMP

في معمارية شـــاملة، مع تركيز على دقة التلاوة وتأثرها بقواعد التجويد.

- تصـــنيف وتقييم أحكام التجويد المحددة :ظهرت دراسات ركزت بشكل خاص على تقييم جودة النطق لأحكام تجويدية فردية، حيث طبقت أســاليب التعلم العميق على مهام التصنيف الثنائي (صحيح/خاطئ) أو المتعدد. على سبيل المثال، ركزت أبحاث حديثة على استخدام الشبكات العصبية العميقة (DNNs) ومعماريات مثل EfficientNet-B0 لتقييم أحكام المد، والغنة، والإخفاء، وحققت نتائج دقة عالية [11].

2.2 الفجوة البحثية وإسهام العمل الحالي

على الرغم من التقدم الهائل في أنظمة ASR القرآنية والجهود المبذولة في تقييم التجويد، فإن معظم الأنظمة الشاملة الحديثة (كالمشار إليها في [8] و [9] و[10]) تركز بشـــكل أســـاســـي على مهمة النســـخ (Transcription)والكشف عن الأخطاء على مستوى الحرف أو الكلمة، وقد لا تقدم تقييماً عميقاً ومخصـصـاً لجودة الأداء الصـــوتي لأحكام التجويد التي تعتمد على صفات زمنية وشدة محددة.

يملأ هذا البحث الفجوة المنهجية المتمثلة في التقييم المُخصص والدقيق لجوهر الأحكام التجويدية الأساسية . يختلف هذا العمل عن الدراســـات الســـابقة والمقارنات الحديثة في النواحي التالية:

أ. **التركيز المخصص** : يركز البحث بشكل حصري على تقييم الأداء الصـــوتي لثلاثة أحكام تجويدية حاسمة لسلامة التلاوة، وهي :المد المنفصل (Al-Madd al-Munfasil)، والغنة المشـــددة (Al-Ghunnah)، والإخفاء (Al-Ikhfaa).

ب. **المنهجية الهجينة المُصـمَّمة** : بدلاً من اسـتخدام نموذج شامل واحد، يقترح هذا البحث منهجية هجينة تعتمد على اســـتخلاص ميزات MFCCs الدقيقة وتغذيتها لـــ ثلاثة نماذج LSTM مستقلة (نموذج مُدرب خصيصاً لكل حكم تجويدي). هذا التخصيص في بنية النموذج والتدريب يتيح تكييف البارامترات الفائقة ليناسـب الخصـائص الصـوتية الفريدة لكل حكم، مما يعزز من قوة ودقة التقييم مقارنة بالأنظمة العامة.

تُقدم هذه الدراسة مساهمة واضحة في تطوير تطبيقات التعلم التفاعلي للتجويد، من خلال تحقيق مستويات عالية من الدقة في تصـــنيف الأخطاء المتعلقة بهذه الأحكام الأساسية باستخدام نهج MFCC-LSTM الهجين.

3. المنهجية

اتبع البحث منهجية تتكون من ثلاث مراحل: جمع البيانات، اســتخلاص الســمات الصـــوتية (المعالجة الأولية)، وبناء وتدريب نموذج التعلم العميق، يوضـــح الشـــكل (1) مخطط بســيط للمراحل المختلفة للمنهجية المستخدمة.

1.3 جمع البيانات

تم استخدام مجموعة البيانات العامة QDAT [3]، التي تحتوي على أكثر من 1500 مقطع صوتي لثلاثة أحكام تجويد: المد المنفصل، النون المشددة (الغنّة)، والإخفاء. كانت الملفات الصـوتية بصـيغة WAV 11 كيلوهرتز، 16 بت، أحادي.

كمســاهمة إضـــافية لدعم الأبحاث المســتقبلية في هذا المجال، قام الباحثون بجمع وتطوير مجموعة بيانات جديدة تحتوي على 1500 تسجيل للآيات الثلاث الأولى من ســـورة المطففين، متاحة للجمهور عبر مســـتودع GitHubعلى الرابط: https://github.com/faraj-elahjal/quran-recitation-dataset

تم تجميع قاعدة البيانات شملت تلاوات قرآنية لقراء متنوعين، لتمثيل حالات الأداء الصحيح والخاطئ لأحكام المد، والإخفاء، والغنة المشددة. تم الحصول على موافقة شفهية (Consent) من جميع المشاركين والتأكد من إبقاء التسجيلات مجهولة المصدر لأغراض البحث فقط.

2.3 استخلاص السمات والمعالجة الأولية

تم تطبيق سلسلة من خطوات ما قبل المعالجة على جميع التسجيلات الصوتية لضمان جودة الإشارة وتوحيدها:

- **معدل العينة (Sampling Rate):** توحيد جميع التسجيلات عند 16 kHz.
- **المعالجة الأولية** :استخدمت تقنية MFCC لاستخلاص السمات. تم تقسيم الإشارة إلى إطارات بطول 25 ميلي ثانية (ms) مع تداخل (Hop Size) قدره 10 ميلي ثانية.
- **المعاملات :** تم استخلاص 13 معامل MFCC ، وتم إضافة معاملات الدلتا (Delta) و الدلتا-دلتا (Delta-Delta) لتمثيل التغيرات الديناميكية في الصوت. وبالتالي، أصبح إجمالي السمات لكل إطار (13 × 3) 39.
- **إثراء البيانات (Data Augmentation)** : لزيادة متانة النموذج وتقليل تأثير الضوضاء، تم تطبيق إثراء البيانات على مجموعة التدريب من خلال إضافة ضوضاء خلفية خفيفة وتغيير بسيط في سرعة النطق (Time Stretching) [12].

يتم حساب التردد على مقياس الميل (Mel-scale) عبر المعادلة (1) [13]، يليه تطبيق تحويل جيب التمام المتقطع (DCT) للحصول على معاملات MFCC عبر المعادلة (2) :

$$Mel(f) = 2595 \cdot log_{10}\left(1 + \frac{f}{700}\right) \quad (1)$$

$$MFCC_c = \sum_{k=1}^{N} log(E_k)\, cos\left[c \cdot \left(k - \frac{1}{2}\right)\frac{\pi}{N}\right] \quad (2)$$

3.3 بناء وتدريب نموذج LSTM

تم بناء ثلاثة نماذج LSTM منفصلة (لكل حكم تجويدي نموذج)، حيث تتكون الطبقة المدخلة من متجهات سمات MFCC (39 سمة) المتسلسلة زمنياً.
وهُيكل النموذج (Hyperparameters) كما يلي:

- **طبقة LSTM:** طبقة واحدة أو اثنتان من LSTM بوحدات تتراوح بين 128 إلى 256 وحدة (Neurons).
- **معدل الانسحاب (Dropout):** تم استخدام Dropout بنسبة 0.2 لطبقة LSTM لتنظيم النموذج وتجنب فرط التخصيص.(Overfitting)
- **المُحسِّن (Optimizer):** مُحسِّن Adam بمعدل تعلم أولي (Learning Rate) يبلغ 0.001 .
- **حجم الدفعة (Batch Size):** هو 32 أو 64.
- **التوقف المبكر (Early Stopping):** تم تفعيل التوقف المبكر بناءً على خسارة مجموعة التحقق (Validation Loss) مع صبر (Patience) يساوي 10 فترات (epochs) .

شكل 1. خطوات المنهجية المتبعة

- **عدد الفترات (Epochs):** تم تدريب النماذج لعدد يتراوح بين 50-100 فترة أو حتى الوصول إلى التوقف المبكر.

يتم تحديث حالة الخلية C_t باستخدام المعادلة (5) ، والتي تدمج مخرجات بوابتي النسيان f_t والإدخال i_t ، كما يوضحان في المعادلتين (3) و (4):

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (5)$$

4. النتائج والمناقشة

تم تقييم أداء النماذج المدربة على مجموعة بيانات الاختبار باستخدام مقاييس الدقة (Accuracy)، والاستدعاء (Recall)، ودرجة (F1)، حيث تُظهر النتائج كفاءة النماذج القائمة على شبكات الذاكرة طويلة وقصيرة المدى (LSTM) في مهمة التقييم الآلي لأحكام التجويد. يوضح الشكل (2) مصفوفات الارتباك التفصيلية للأحكام الثلاثة، ويوضح الجدول (1) الدقة الإجمالية التي حققها كل نموذج:

- سجل نموذج المد المنفصل (Al-Madd) أعلى دقة كلية بلغت 77.0%.
- يليه نموذج الإخفاء (Al-Ikhfaa) بدقة 66.1%.
- بينما سجل نموذج الغنة المشددة (Al-Ghunnah) أدنى دقة إجمالية بلغت 59.4%.

هذا الأداء، على الرغم من أنه أقل من التوقعات الأولية، يظل تنافسياً مقارنةً بالمنهجيات التقليدية المستخدمة سابقاً في تقييم التجويد، مثل نماذج ماركوف المخفية (HMM) [4] وآلات دعم المتجهات (SVM) [5]. يبرز تفوق شبكات LSTM في هذا البحث قدرتها الفائقة على التعامل مع طبيعة التلاوة كبيانات متسلسلة طويلة، مما يمكنها من التقاط الاعتماديات الزمنية المعقدة الضرورية لتقييم أحكام المد والغنة بدقة أعلى [7].

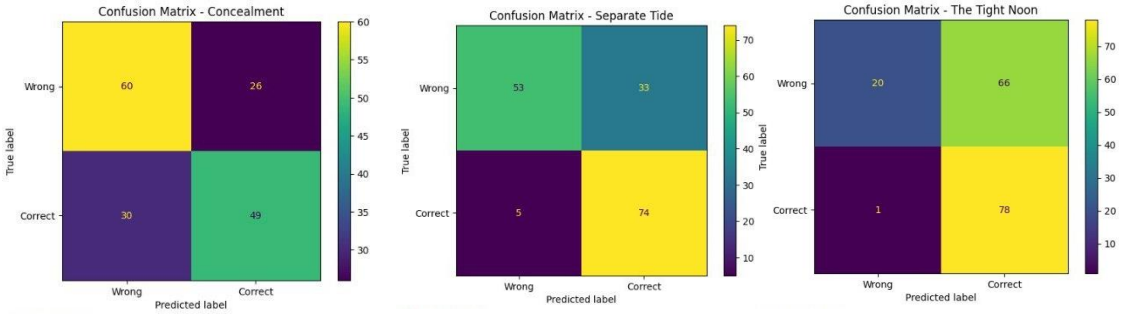
على الرغم من أن نموذج المد المنفصل حقق أعلى دقة كلية، فإن التباين في الأداء بين الأحكام يستدعي تحليلاً أعمق لمصفوفات الارتباك:

1. **نموذج الغنة المشددة:** سجل أعلى دقة (Precision) بلغت 98.7%، مما يعني أن النموذج كان موثوقاً به للغاية في توقعاته الإيجابية. لكنه عانى من انخفاض حاد في الاستدعاء (Recall) بنسبة 54.2%، حيث سجل 66 سلبية كاذبة (FN). هذا يشير إلى أن النموذج كان حذراً للغاية (Over-Cautious) في تصنيف التلاوات على أنها صحيحة، مفضلاً تصنيف التلاوة الصحيحة على أنها خاطئة بدلاً من العكس.

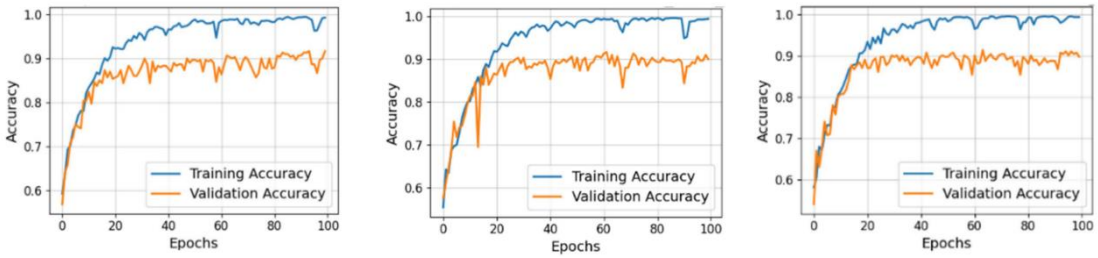
2. **نموذج الإخفاء:** سجل أدنى دقة (Precision) بلغت 62.0% بسبب ارتفاع عدد الإيجابيات الكاذبة (FP=30). هذا يدل على أن النموذج واجه صعوبة في التمييز بين الخصائص الترددية للغنة وتلك المتعلقة بغيرها من الأحكام، مما جعله يبالغ في تصنيف التلاوات الخاطئة على أنها صحيحة [13].

كما يوضح الشكل (3) (منحنيات التدريب والتحقق) مسار التعلم للنماذج الثلاثة؛ حيث يُلاحظ تقارب سريع ومستقر بين منحنيات خسارة التدريب (Training Loss) وخسارة التحقق (Validation Loss)، مما يؤكد نجاح استراتيجيات تنظيم النموذج (Regularization) مثل Dropout والتوقف المبكر (Early Stopping) في تجنب فرط التخصيص (Overfitting) وضمان وصول النموذج إلى نقطة تعميم مثلى على البيانات الجديدة. يُعزى الأداء المستقر للنماذج إلى التكامل الفعال بين عنصرين أساسيين في المنهجية: الأول هو قدرة شبكات LSTM الفائقة على النمذجة الزمنية، حيث تستخدم بواباتها (المعادلات 3-5) لالتقاط الاعتماديات الزمنية الطويلة الضرورية لتقييم مدة الأحكام التجويدية

بدقة، مما يمنحها ميزة واضـــحة على النماذج المعتمدة على HMM [4]؛ والثاني هو جودة السمات المُستخدمة، حيث أتاحت سمات MFCC [1] المدعومة بسمات الدلتا والدلتا–دلتا (39 سـمة لكل إطار) للشـبكة التعرف على التغيرات الحركية والسـريعة في الترددات، وهي ضـرورية لتمييز حالات مثل الإخفاء والغنة المشددة. بالإضافة إلى ذلـك، لـعـبـت تـقـنـيـة إثـراء الـبـيـانـات (Data Augmentation) دوراً هاماً في زيادة متانة النموذج ضـد التباين في جودة التسـجيلات أو ضـوضـاء الخلفية [11]. وبالمقارنة مع الأعمال السـابقة [5, 6]، يُظهر نظامنا أداءً تنافسـياً أو متفوقاً مع معالجة مجموعة من الأحكام التجويدية الصـعبة، مما يؤكد أهمية الاتجاه نحو اسـتخدام نماذج التعلم العميق في تطوير أدوات تعليم التجويد الآلية.



شكل 2. مصفوفات الأرتباك للأحكام الثلاثة.



شكل 3. منحنيات التدريب والتحقق للنماذج.

جدول 1. مقارنة بين أداء النماذج المختلفة.

مقياس الأداء	المد المنفصل (Al-Madd)	الإخفاء (Al-Ikhfaa)	الغنة المشددة (Al-Ghunnah)
الدقة (Accuracy)	%77.0	%66.1	%59.4
الدقة (Precision)	%93.7	%62.0	%98.7
الاستدعاء (Recall)	%69.2	%65.3	%54.2
درجة F1 (F1 Score)	%80.0	%63.6	%70.1

5. الاستنتاجات

أثبت هذا البحث بنجاح فعالية اسـتخدام شـبكات الذاكرة طويلة وقصـيرة المدى (LSTM) المقترنة بسـمات معاملات التردد الميلية (MFCC) في تطوير نظام آلي لتقييم دقة تلاوة القرآن الكريم. لقد حققت النماذج أداءً متفاوتاً، حيث أظهر نموذج المد المنفصـل الأداء الأكثر توازناً بدقة كلية بلغت .%77.0بينما سجل نموذج الغنة المشـددة أدنى دقة كلية، لكنه تميز بدقة تصـنيف (Precision)عالية جداً بلغت 98.7%، مما يشير إلى

موثوقيته العالية عند التصنيف الإيجابي، على الرغم من حذره الشديد في اكتشاف جميع الحالات الصحيحة.

كما تم تصميم واجهة بسيطة باستخدام مكتبة Tkinter لتجربة أداء النماذج، وأظهرت فعالية واضحة في عرض النتائج بشكل سهل وواضح للمستخدم.

يوصى في العمل المستقبلي بمعالجة التحديات المنهجية التي ظهرت، خاصة تحسين مؤشر الاستدعاء (Recall)لنموذج الغنة المشددة، وتوسيع نطاق النظام ليشمل أحكام تجويد إضافية ودمجه مع منصات التعليم الإلكتروني لتعظيم الفائدة المرجوة.

6. المراجـــــع

[1] V. Tiwari, "MFCC and its applications in speaker recognition," International Journal on Emerging Technologies, vol. 1, no. 1, pp. 19–22, 2010.

[2] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.

[3] Al Harere, A., & Al Jallad, A. (2022). Holy Quran Recitation Evaluation Using Deep Learning Algorithms. International Journal of Advanced Computer Science and Applications.

[4] A. Alghamdi, A. Al-Ajmi, and S. Al-Saeed, "Quran Recitation Error Detection using Hidden Markov Models," in Proc. of the Int. Conf. on Artificial Intelligence and Soft Computing (ICAISC), 2019, pp. 45-52.

[5] K. Alshanaq, M. Barhoush, and K. Nahar, "AN EFFICIENT HOLY QURAN RECITATION RECOGNIZER BASED ON SVM LEARNING MODEL," Jordanian Journal of Computers and Information Technology, vol. 6, pp. 392–414, Oct. 2020. doi: 10.5455/jjcit.71-1593380662.

[6] A. M. Alagrami and M. M. Eljazzar, "SMARTAJWEED Automatic Recognition of Arabic Quranic Recitation Rules," Computer Science & Information Technology (CS & IT), pp. 145–152, Dec. 2020. doi: 10.5121/csit.2020.101712.

[7] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," Neural Computation, vol. 12, no. 10, pp. 2451–2471, 2000. doi: 10.1162/089976600300014993.

[8] I. K. Tantawi, M. A. M. Abushariah, and B. H. Hammo, "A deep learning approach for automatic speech recognition of The Holy Qur'ān recitations," International Journal of Speech Technology, vol. 24, pp. 1017–1032, 2021.

[9] A. A. Harere and K. A. Jallad, "Quran recitation recognition using end-to-end deep learning," arXiv preprint, arXiv:2305.07034, 2023.

[10] S. Alfadhli et al., "qArI: A Hybrid CTC/Attention-Based Model for Quran Recitation Recognition Using Bidirectional LSTMP in an End-to-End Architecture," IEEE Access, 2024.

[11]) Placeholder: Use the full citation for the study mentioned in the search results that focuses on classifying the three Tajweed rules using DNN/EfficientNet - e.g., "Evaluation of the Pronunciation of Tajweed Rules Based on DNN as a Step Towards Interactive Recitation Learning" or a similar relevant paper.(

[12] M. Al-Hussaini and F. M. Al-Shargabi, "A Comprehensive Review on Arabic Speech Recognition Systems using Deep Learning," Applied Sciences, vol. 13, no. 7, 2023. doi: 10.3390/app13074609.

[13] Z. S. Al-Hafidh, S. Z. Al-Mutairi, and M. I. Al-Jarallah, "A Comparative Study of Machine Learning Classifiers for Tajweed Rule Detection," Int. J. of Advanced Computer Science and Applications (IJACSA), vol. 12, no. 3, 2021. doi: 10.14569/IJACSA.2021.0120387.

[14] D. D. C. K. M. Z. P. L. H. B. M. "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition," Interspeech, 2019. doi: 10.21437/Interspeech.2019-2680