مؤسسة النهضة العلمية

التعرف الآلي على تلاوة القرآن الكريم: المدود نموذجاً

بلال يوسفي[1]، أكرم محمد زكي [2]، أمينة حاجي[3]
[1.2.3] الجامعة الإسلامية العالمية بماليزيا
كوالالمبور، ماليزيا
[3] haji.aminah@hotmail.com ، [2] akramzeki@iium.edu.my ، [1] yousfi.bilal@hotmail.fr

الخلاصة: قراءة القرآن الكريم وتعليمه من أهم العلوم وأرفعها، ومن هنا كان على قارئ القرآن ومعلمه مراعاة أحكام التجويد وتلقينها بشكل صحيح. يأتي هذا البحث مدعماً لهذا العلم الرفيع (علم التجويد) حيث يبين استعمال تقنية التعرف على الكلام Automatic Speech recognition (ASR) للتعرف على القراءة الصحيحة لأحد أحكام المدود من رواية حفص وهو حكم مد البدل ومد اللين لتمكين المستخدم من التعرف وتصحيح هذا النوع من المدود عن طريق مطابقة قراءته مع مجموعة من البيانات مسجلة في قاعدة البيانات من طرف مجموعة من القراء المجازين. ويقترح هذا البحث تصميم برنامج تعليمي تفاعلي لقواعد التجويد للقرآن الكريم المسموعة لقراءة حفص عن عاصم حيث تم استخدام إحدى التقنيات ألا وهي ((Mel-Frequency cepstrum coefficient (MFCC لاستخراج ميزات الأصوات وHMM من أجل هذا النظام باستخدام نماذج ماركوف الخفيةHidden Markov Models وهي عبارة عن سلسلة نهائية من الحالات والملاحظات والتي يتم توليدها عشوائياً ويتم الانتقال من حالة إلى حالة أخرى عن طريق مصفوفة الاحتمالات.

الكلمات الجوهرية: التعرف على الكلام، التجويد، MFCC، باستخدام نماذج ماركوف الخفية.

1. المقدمة

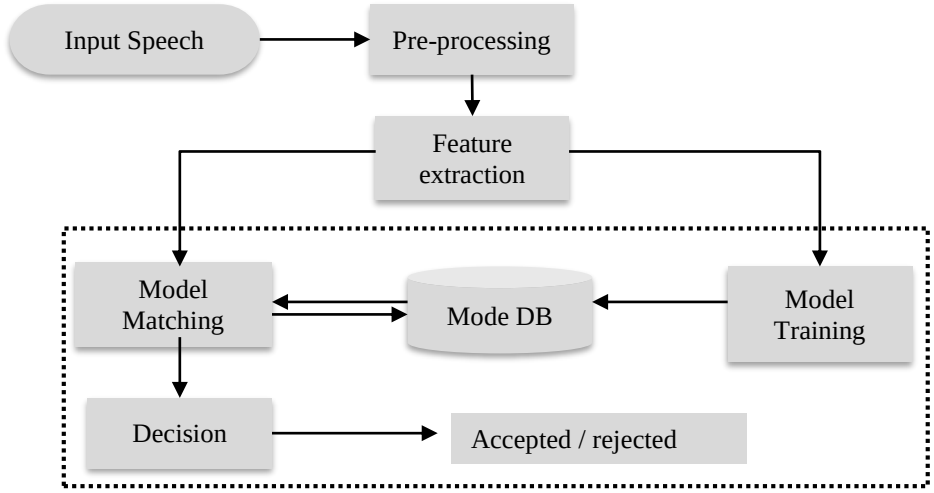
قال الله تعالى في محكم ءاياته:﴿ وَرَتِّلِ الْقُرْآنَ تَرْتِيلًا ﴾. تختلف قراءة آيات القرآن الكريم عن أية قراءة أخرى، فيتطلب ذلك ترتيل القرآن الكريم وتجويده شرط المعرفة الكاملة بأحكام التجويد، التجويد لغة هو التحسين أو الإتيان بالجيد. أما اصطلاحا فهو إعطاء الحرف حقه ومستحقه مخرجاً ومدًا، أو تلاوة القرآن على حسب ما أنزل الله سبحانه وتعالى على نبيه محمد صلى الله عليه وسلم. والهدف منه صون اللسان عن الخطأ في كتاب الله. قراءة القرآن الكريم قراءة سليمة يؤجر عليها القارئ لأن القرآن يوم القيامة يأتي شفيعًا مشفعًا استناداً لحديث أبي أمامه الباهلي رضي الله عنه قال سمعت رسول الله صلى الله عليه وسلم يقول: (اقرؤوا القرآن فإنه يأتي يوم القيامة شفيعاً لأصحابه) صحيح مسلم.

تعد تقنيات التعرف على الكلام من أهم التقنيات الحديثة وقد تم تطوير العديد من الأنظمة المختلفة من حيث الطرق المستخدمة في استخراج السمات وطرق التصنيف. تقية التعرف على الكلام هي عملية استقبال الكلمات أو الجمل

المنطوقة من خلال ميكرفون وتحويلها بدقة وكفاءة إلى مجموعة من الكلمات [1]. ويتحقق ذلك عن طريق اتباع خطوات معينة باستخدام نظام التعرف على الكلام. يتكون النظام من ثلاثة مراحل أساسية كما هو موضح بالشكل رقم 1:

المعالجة الأولية: Pre-processing

حيث يتم تمرير الكلام المسجل (الآيات) من خلال تقنيات الترشيح (عملية معالجة الإشارة) لإزالة الضجيج والفصل بين الصوامت والمصوتات وايضا الكشف عن نقطة بداية ونهاية كل كلمة.



شكل1: المخطط التفصيلي لتقنية التعرف على الكلام.

أ. **استخلاص الصفات: Features Extraction**

تعد تقنية استخلاص الصفات من إحدى التقنيات ذوات التطبيق الواسع في مجالات عديدة من أهمها تمييز الأنماط Pattern Recognition وتعد إحدى الركائز المهمة التي يقوم عليها نظام التعرف على الكلام. الهدف من تقنية استخلاص الصفات هو استخراج معالم فريدة من نوعها لكل عينة من الكلام والتي سيتم استخدامها للتمييز بين مجموعة واسعة من الكلمات المتميزة. يتم في مرحلة استخلاص الصفات واقتباس صفات الشيء لغرض المقدرة على تمييزه فيما بعد من خلال تلك الصفات التي تصف ذلك الشيء وتميزه عن غيره [1]. تمثل مرحلة تحميل الإشارة لاستخلاص الصفات الصوتية بدقة والتي تمثل المعلومات الصوتية ((Acoustic Information لتلك الإشارة حيث يتم معالجة الإشارة من البداية إلى النهاية وذلك بتقسيم الإشارة إلى إطارات متساوية في الطول ثم يتم استخلاص الصفات لكل إطار لتكون النتيجة النهائية مجموعة من الصفات. هناك عدة طرق تستخدم لاستخلاص الصفات من بينها [2]:

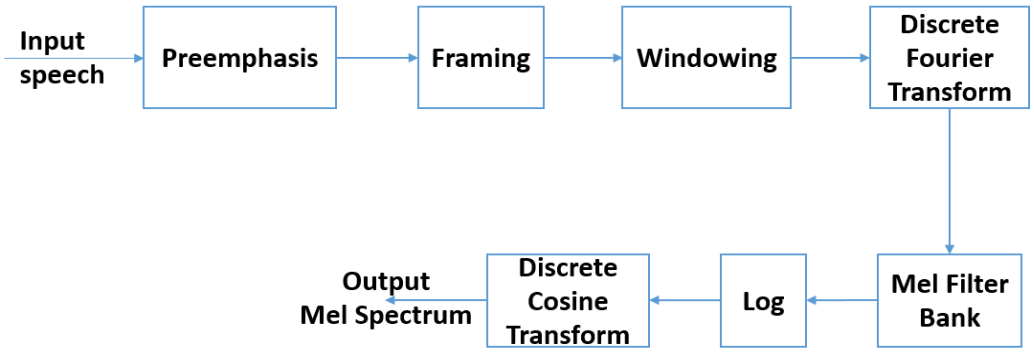
- Perceptual Linear Prediction (PLP).
- Linear Predictive Coding (LPC).
- Mel-Frequency Cepstrum Coefficients (MFCC).
- Linear Prediction Cepstral Coefficients (LPCC).
- Discrete Wavelet Transform (DWT).
- Linear Discriminant Analysis (LDA).

تقنية Mel-Frequency Cepstrum Coefficients (MFCC)

في هذا البحث تم استخدام التقنية MFCC حيث يتم في هذه المرحلة تفخيم الموجه Pre-emphasis ويتم ذلك بتطبيق مرشح Finite Impulse Filter (FIR) ويمرر هذا النوع من المرشحات الترددات العالية ويمنع الترددات الواطئة. ثم يتم تقسيم الموجة إلى عدد من الإطارات (Frames) كل إطار يتكون من N من العينات وهذه الإطارات تكون بشكل متجاور مفصولة عن بعضها بـ M من العينات فإذا كانت M N فان الإطارات المتجاورة ستكون متداخلة. من هذا المبدأ يتم تقسيم الإشارة الصوتية إلى عدة إطارات متداخلة. ومن ثم تقطيع الصوت لتحويله إلى أجزاء بواسطة تقنية النوافذ (Windowing) من بداية إلى نهاية كل مقطع لكي يتم التقليل من نسبة الخطأ التي قد تحدث نتيجة التقسيم [2]. الخطوة التالية هي تحويل كل إطار من المجال الزمني إلى المجال الترددي عن طريق استخدام DFT . ثم استخراج Mel Filter Bank ويتم ذلك عن طريق مجموعة من المرشحات (فلتر) الثلاثية Triangular Bandpass Filter التي تستخدم لكل إطار مع التردد الفعلي. يتم بواسطة هذه الطريقة تغطية جميع الترددات بدون أي فقدان للمعلومات عند إدخال الإشارة الناتجة عن FTT. الخطوة التالية هي حساب اللوغاريتم (Log) للإشارة. والهدف من هذه العملية هو التكيف مع النظام تماما مثل الأذن البشرية. للحصول على MFCC ، تتم معالجة نتيجة لوغاريتم الطاقة مع (DCT) . المعادلة التالية المبينة أدناه هي العلاقة التجريبية التقريبية لحساب الترددات Mel frequencies.

$$Mel\ (f) = 2595 * \log 10\ (1+f/700)$$

الشكل التالي يعطي المخطط التفصيلي لنموذج (MFCC)



شكل2: المخطط التفصيلي لنموذج (MFCC)

ب. تصنيف الميزات أو التعرف على الأنماط

هو عملية تحديد أوجه التشابه بين الميزات المستخرجة من عينات الكلام الجديدة ومجموعة من النماذج الصوتية المخزنة في قاعدة البيانات. استخدمت تقنيات نماذج ماركوف الخفية ((Hidden Markov Model (HMM) من قبل العديد من الباحثين للتعرف على الكلام لأنها أثبتت فعاليتها في عدد من المجالات.

2. دراسات سابقة

في السنوات الاخيرة ظهر اهتمام كبير لدراسة تقنية التعرف على الكلام العربي من قبل العديد من الباحثين. كما قامت بعض الدراسات بتسليط الضوء على القرآن الكريم من خلال اقتراح نظريات لتعليم القرآن الكريم بشكل تفاعلي بين الإنسان والآلة. والتطلع للحصول على دقة تعرّف عالية [1] . اقترح مجموعة من الباحثين نظامًا ذكيًا لمساعدة المسلمين في تلاوة وتحفيظ القرآن E-Hafiz. وكان الهدف من (E-Hafiz) هو تسهيل التعلم والحفظ وتقليل الأخطاء في ترتيل القرآن الكريم. ويمكن للمستخدمين تعزيز مهاراتهم في التلاوة وتقصي الأخطاء. تم بناء النظام بإستخادم MFCC وVQ. وأفاد تحقيق معدل الاعتراف يتراوح بين 82% و91% [3]. مجموعة أخرى من الباحثين استخدموا أسلوب التصنيف للتعرف على القلقلة الكبرى (حكم من أحكام التجويد). ووصفت هذه الدراسة باستخدام متعدد المستقبلات MLP لتصنيف نطق القلقلة الكبرى وأيضا استخدام MFCC لاستخراج الميزات. تم تسجيل البيانات من خمس كلمات حيث حققوا نتائج تراوحت من 95٪ إلى 100٪. كما أكد الباحثون في هذه الدراسة أن استخدام التصنيف متعددة الطبقات المستقبلات MLP واستعمال MFCC لاستخراج الميزات يساعد على التفريق بين النطق الصحيح وغير الصحيح للقلقلة الكبرى [4]. في دراسة أخرى تم تصميم وتقييم عمل بحثي لتطوير المجود الآلي لاستخدامه كطريقة تفاعلية لتعلم وفحص أحكام التجويد كما يركز على قواعد تلاوة القرآن. النظام مبني على أساس MFCC ونموذج HMM. كما تم تدريب النظام على سورة الفاتحة. حيث بلغت دقة التعرف 91.95٪ (الآيات) و86،41٪ (الصوتيات) [5]. في دراسة أخرى تم تقديم نموذج جديد للتمييز بين نوع القراءة في تلاوة للقرآن الكريم. تقنية استخراج الميزة المستخدمة في هذا البحث هي (MFCC) التي استخدمت لاستخراج الخصائص من تلاوة الآيات القرآنية. وأيضا استعمال نماذج ماركوف المخفية (HMM) للتصنيف والتعرف على النمط [6] .

3. منهجية البحث

يعمل هذا النظام على التصحيح والتعرف على حكم المد لكل من المد البدل والمد اللين للقارئ حيث تم اختيار مجموعة البيانات التي هي عبارة عن مجموعة من الأصوات لعدد من الكلمات تم اختيارها من القرآن الكريم وهي خاصة بحكم المد.

كما تم جمع مجموعة من الكلمات وذلك بواسطة قراء مجازة أصواتهم. حيث تم تسجيل كل كلمة عدة مرات من كل قارئ. الجدول التالي يبين خصائص كل من المد البدل والمد اللين.

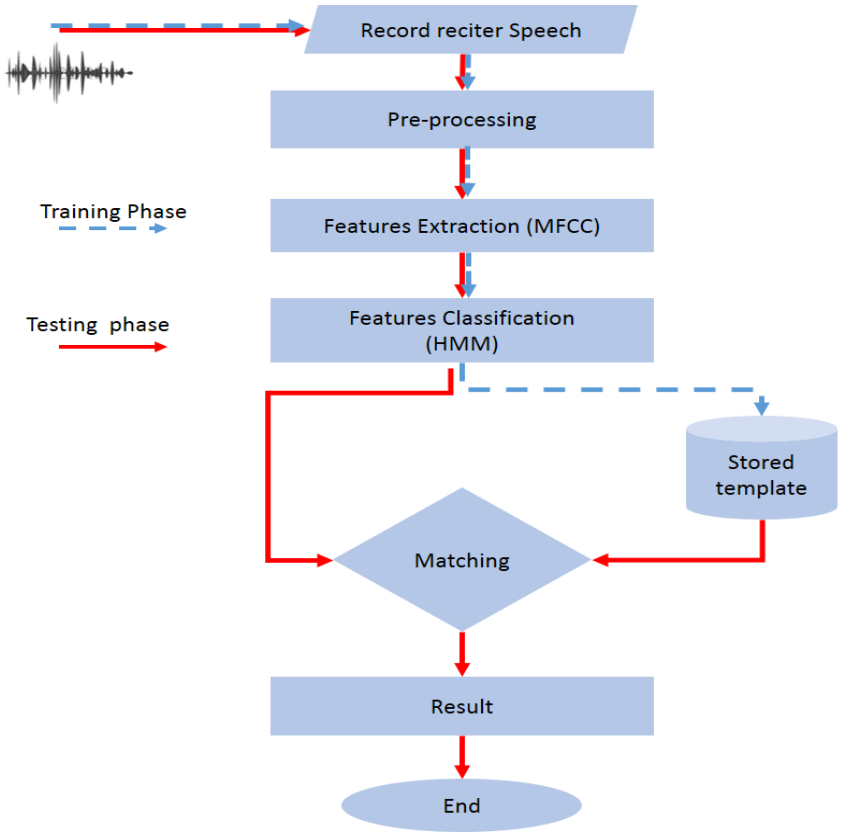
الجدول 1: خصائص كل من المد البدل والمد اللين. رواية حفص

نوع المد		عدد الحركات	القاعدة
البدل	البدل الاصلي	حركتان	حرف المد بعد همزة قطع، وكان بدلا لهمزة قطع
	الشبيه أو الملحق بالبدل		كل حرف مد بعد همزة، ولم يكن بدلا لهمزة قطع
اللين		6.4.2 حركات وقفا	إطالة الصوت بالواو والياء الساكنتين المفتوح ما قبلهما، وما بعدهما ساكن سكونا عارضا عند الوقف ولا يمد عند الوصل

الجدول 2: الكلمات المستخدمة في تدريب واختبار النظام.

نوع المد	المد البدل	المد اللين
الكلمات المختارة	آدم	كهيئة
	إيمانا	شيء
	أوتُوا	سوءاتهما
	ءَامَنُوا	يايئس
	آزر	سوء

الخطوة الأولى هي عملية تسجيل الصوت عن طريق جمع عينات تلاوة القرآن الكريم من مجموعة من القراء لديهم إجازة في حفص، كل منهم يقرأ آيات معينة تحتوي على نوعين من أحكام المد وهما مد البدل ومد اللين لحفص مرات عديدة وبطريقة صحيحة. الخطوة التالية بعد جمع البيانات هي ما قبل المعالجة حيث يتم تمرير إشارات الكلام التي تم جمعها من خلال كتلة ما قبل المعالجة لإزالة الضوضاء الواردة في إشارة الكلام وصوت منفصل عن خطاب صامت للحد من مجموعة من الصفات التي نؤكد فقط على المعلومات التي يراد أن تنقل. ثم يتم تطبيق الخوارزمية MFCC لاستخراج الميزات التي تستخدم على نطاق واسع كمدخل لغرض التعرف. بعد ذلك يتم تصنيف الميزات والتدريب، وحساب عوامل HMM. تتم مطابقة بيانات الاختبار (القارئ الجديد) مع الميزات الصوتية المخزنة في قاعدة البيانات (التي يتم بناؤها من بيانات التدريب) لتقديم نتائج على أساس ما إذا كانت القراءة بشكل صحيح أو بشكل غير صحيح. الشكل رقم 3 يوضح الخطوات المتبعة في منهجية البحث.



شكل3: المخطط يبين منهجية البحث.

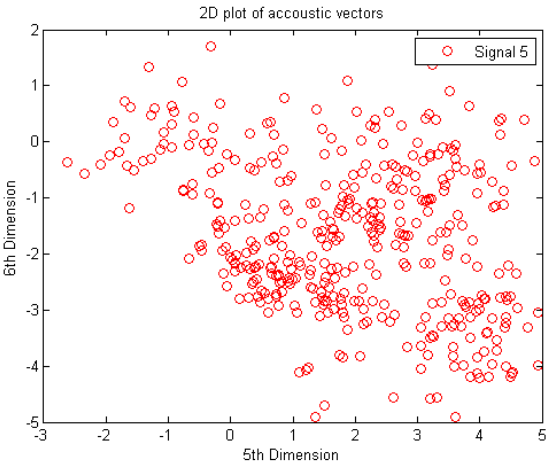
4. تحليل النتائج:

بعد تحديد العينات تم عرضها علىMATLAB لتحديد الخصائص الفيزيائية لها. قمنا بحساب متوسط الزمن اللازم لكل نوع من المدود. الجدول التالي يبين متوسط الزمن اللازم لمد البدل.

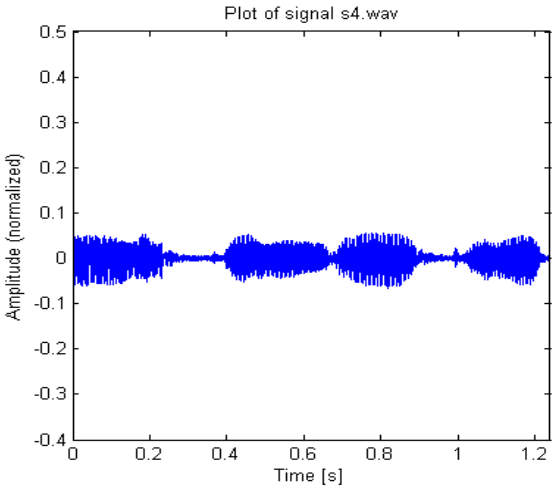
المد البدل	الفتحة القصيرة المرققة	الفتحة القصيرة المفخمة	الكسرة القصيرة	الضمة القصيرة
المتوسط الزمني	1.755	2.81	1.74	1.70

الجدول 3: جدول يبين متوسط أزمان مد البدل في رواية حفص

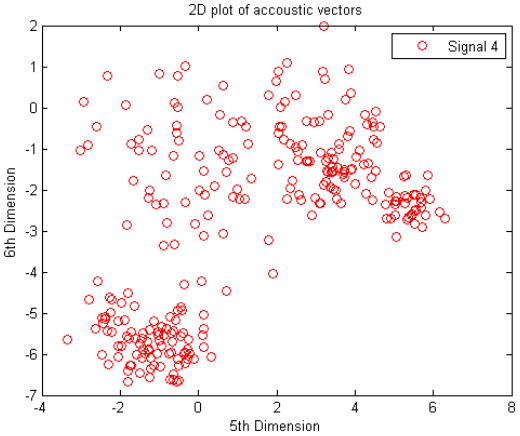
الاشكال 4-5-6 تبين بعض النماذج لبعض الخطوات التي طبقت على الخوارزمية المقترحة



شكل4: 2D Plot لكلمة ءَامَنُوا



شكل5: عرض الإشارة الأصلية لكلمة ءَامَنُوا



شكل6: 2D Plot لكلمة كهيئة.

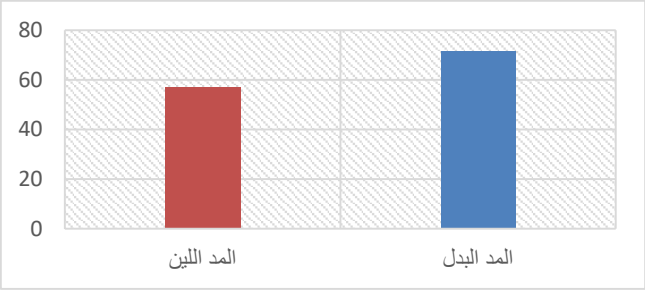
إن نسبة صحة الكلمة في نظام تمييز الكلام تعرّف كما يأتي:

$$\text{نسبة صحة الكلمة} = \frac{\text{عدد الكلمات المميزة بشكل صحيح}}{\text{العدد الكلي للكلمات}} \times 100$$

يمثل الجدول التالي نسبة دقة التعرف المتحصل عليها بعد القيام بمجموعة من التجارب. من خلال النتائج تبين أن الخوارزمية أعطت نسبة تعرّف مقبولة. حيث كانت نسبة التعرّف بالنسبة لمد اللين كانت 57.14% أما بالنسبة لمد البدل فقد وصلت النسبة إلى 71.43%.

الجدول 3: نسبة التعرّف لنفس الأصوات التي تم التدرب عليها

نوع المد	المد البدل	المد اللين
عدد العينات	7	7
عدد العينات ذات النتائج الصحيحة	5	4
عدد العينات ذات النتائج غير الصحيحة	2	3
نسبة التعرّف	%71.43	%57.14



شكل7: مدرج تكراري يمثل نسبة التعرّف على أحكام المدود.

5. الخاتمة:

تم تطوير وبناء هذه الدراسة على نظريات وتدريبات تساعد على تطوير نظام التجويد للمدود كما تم استخدام MFCC وHMM نظرًا لكونهما من أهم الخوارزميات التي أعطت نتائج جيدة في الدراسات السابقة وبالتالي هذه الدراسة موصى بها في الأبحاث والأعمال المستقبلية. وعلى الرغم من ذلك، واجهت الدراسة بعض العيوب والمشاكل وذلك بسبب تداخل الموجات الصوتيه بسبب الملفات المحملة والجودة الرديئة (ضوضاء) أثناء عملية تسجيلها.

6. جدول الألفاظ

English	عربي
Automatic Speech recognition (ASR)	التعرف على الكلام التلقائي
Hidden Markov Models (HMM)	نماذج ماركورف الخفية
Mel-Frequency Cepstrum Coefficients (MFCC)	طريقة لتحليل الصوت تعتمد بالأساس على ترشيح معاملات الترددات (تقنية درجة النغم)
Filter Bank	بنك الترشيح
Triangular Bandpass Filter	ترشيح حزمة مثلثة الشكل

7. المراجع

[1]

B. Yousfi and A. M. Zeki, "Automatic Speech Recognition for the Holy Qur 'an, A Review," in The International Conference on Data Mining, Multimedia, Image Processing and their Applications (ICDMMIPA2016), 2016, p. 23.

[2]

S. K. Gaikwad, B. W. Gawali, and P. Yannawar, "A Review on Speech Recognition Technique," *Int. J. Comput. Appl.*, vol. 10, no. 3, pp. 16–24, 2010.

[3]

A. Muhammad, W. M. M. Z. ul Qayyum, S. Tanveer, A. Martinez-Enriquez, and A. Z. Syed, "E-hafiz: Intelligent system to help Muslims in recitation and memorization of Quran," *Life Sci. J.*, vol. 9, no. 1, pp. 534–541, 2012.

[4]

H. A. Hassan, N. H. Nasrudin, M. N. M. Khalid, A. Zabidi, and A. I. Yassin, "Pattern classification in recognizing Qalqalah Kubra pronuncation using multilayer perceptrons," in *Computer Applications and Industrial Electronics (ISCAIE), 2012 IEEE Symposium on*, 2012, pp. 209–212.

[5]

D. Raja-Jamilah Raja-Yusof and Fadila Grine, N. Jamaliah Ibrahim, M. Yamani Idna Idris, Z. Razak, and N. Naemah Abdul Rahman, "Automated tajweed checking rules engine for Quranic learning," *Multicult. Educ. Technol. J.*, vol. 7, no. 4, pp. 275–287, 2013.

[6]

B. Yousfi and A. M. Zeki, "Holy Qur'an speech recognition system distinguishing the type of recitation," in *Computer Science and Information Technology (CSIT), 2016 7th International Conference on*, 2016, pp. 1–6.

8. الخلاصة باللغة الانجليزية

Holy Qur'an Speech Recognition System

Mudud Tajweed Rule Checking

Bilal Yousfi[1], Akram M. Zeki[2] and Aminah haji[3]

[1,2,3] Kulliyyah of Information and Communication Technology (KICT)
International Islamic University Malaysia (IIUM)
yousfi.bilal@hotmail.fr[1], akramzeki@iium.edu.my[2], haji.aminah@hotmail.com[3]

Abstract

Learning and teaching the Quran is the most important science for Muslim. The teacher and learner in this area they should have the provisions of tajweed rules when reading the Quran. This research will present a speech recognition system to recognize, identify and correct the rules of Madd (elongate tone) or prolongation for Qira'at of HAFSS. The proposed system will be capable to recognise, identify, point out the mismatch and Discriminate between two types of madd which are: The Soft Lengthening **المد اللين** and The Exchange Prolongation **المد البدل** rules for Hafss for the verses contains the two rules which made by the expert teachers stored in a database. In addition, in this research used Mel-Frequency Cepstral Coefficient (MFCC) and Hidden Markov Models (HMM) as feature extraction and feature classification respectively.

Keywords: speech recognition, Tajweed, MFCC, HMM