



Statistical Study for Enhancing the MFCC Algorithm for Real-World and Highly Noisy Environments in Voiceprint Extraction

Asmaa Barakat¹

¹ Faculty of Electrical and Electronic Engineering, University of Idlib, Idlib, Syria.

Nada Abu Nokra¹

¹ Faculty of Electrical and Electronic Engineering, University of Idlib, Idlib, Syria.

Abdul Rahman Hussian²

² Faculty of Electrical and Electronic Engineering, Istanbul University, Istanbul, Turkey.

²corresponding author Email: abdul.hussian@istanbul.edu.tr

ABSTRACT

The aim of this study is to present a novel exploration for obtaining an enhanced version of the MFCC algorithm to capitalize on its superior quality in sound analysis and voiceprint extraction. The objective is to overcome challenges in using the MFCC algorithm in real-world and highly noisy environments, where speech is susceptible to noise and interference in natural settings. This, in turn, diminishes the performance of the MFCC-based system in real-world applications.

The proposed study successfully identifies speakers in noisy conditions by leveraging the leading features of MFCC in sound analysis, addressing the challenge of noise variation. The study involves the application of various mathematical equations in the process of voiceprint extraction, comparison between them, and obtaining the optimal equation for extracting voiceprints in high-noise environments. The study applies to directly recorded sounds from the environment, without relying on pre-existing databases.

Test results indicate the speaker discrimination rates among users (10, 20, 30, 40, 50, 60, 70, 80, 90, 100). Using the traditional MFCC algorithm with the logarithmic function, the discrimination rates were (73.5%, 72.63%, 71.33%, 70.625%, 68.3%, 68.66%, 68.78%, 68.75%, 68.22%, 67.65%), respectively. With the improvement of MFCC and the application of square root, discrimination rates improved, yielding the following enhanced results (94.5%, 95%, 95.3%, 94%, 92.4%, 91.16%, 91.14%, 90.93%, 90.55%, 89.85%). Applying the cubic root equation for the same users resulted in the following order (88.5%, 88.94%, 89.66%, 85%, 84.2%, 84.41%, 84.57%, 84%, 83.94%, 83.17%). The discrimination rate with the square root from the fourth position was (83%, 83.15%, 83.83%, 81.87%, 80.30%, 80.58%, 80.35%, 79.68%, 79.33%, 78.48%). Consequently, voice discrimination rates were improved.

Keywords:

Noise Reduction - Voiceprint - MFCC Algorithm

يعد استخلاص البصمة الصوتية من المجالات الحيوية الهامة، ويعود ذلك لأهمية التطبيقات التي تستخدم فيها البصمة الصوتية، وكذلك تلقي البصمة الصوتية قبولاً من المستخدم، فأجريت العديد من البحوث والدراسات في مجال الصوت، ومن الخوارزميات الرائدة في هذا المجال خوارزمية MFCC ومن الأبحاث التي استخدمتها:

1. تم تقديم دراسة باستبدال التحويل اللوغارتمي في تحليل MFCC القياسي بمعادلة مدمجة لتحسين حساسية MFCC للضجيج، نظراً لأن خوارزمية MFCC حساسة جداً للضجيج في البيئات الصاخبة العملية، وبالتالي تؤدي إلى تدهور مستوى أداء أنظمة التعرف على الهوية بشكل كبير، بسبب عدم التطابق بين التدريب والاختبار، وبينت الدراسة أن الطريقة المستخدمة هذه تقلل بشكل كبير من معدل خطأ التعرف على الهوية في العالم الحقيقي. [1]
2. تم تصميم نظام التعرف على المتحدث، يتم تحليل الصوت باستخدام MFCC لاستخراج الميزات، ويتم التصنيف باستخدام الخوارزمية العصبية المعتمدة على "Support Vector Machinen" SVM وتم تطبيق النظام المقترح على أصوات ثمانية أشخاص من قاعدة بيانات دولية وتجاوزت الدقة 90% [2]
3. أجريت دراسة لتأثير الضجيج في ميزات GFCC و MFCC وتحليل جميع الاختلافات فيما بينها، بشكل دقيق ومقارنة النتائج، وبين البحث أن بنك الترشيح في خوارزمية GFCC أقل تأثراً بالضجيج من بنك الترشيح لخوارزمية MFCC، وأكد البحث على تفوق GFCC على MFCC من خلال استخدام الجذر التكعيبي حيث يوفر نتائج أفضل من استخدام اللوغاريتم في تحليل [3]. MFCC Computational auditory scene analysis تم تصميم نموذج للتعرف على المتحدث، يتم حذف الضجيج باستخدام "Computational auditory scene analysis" (CASA) ويتم تحليل الصوت باستخدام ثلاث خوارزميات GF و GFCC و MFCC وتم التمييز باستخدام شبكة DNN وطبق النظام على قاعدة البيانات [4]. NIST
5. صمم نظام التعرف على المتحدث مستقل عن النص، حيث يتم تحليل الصوت باستخدام تحليل المويجات المنفصلة وخوارزمية MFCC، ويتم التعرف باستخدام شبكة عصبية ذات أساس تراجعي واحتمالي وشعاعي، وقد أدى النظام المقترح إلى تحسين معدل تحديد الهوية بنسبة 15% مقارنة بنظام MFCC الكلاسيكي. بالإضافة إلى ذلك، فقد خفضت زمن التعرف بنسبة 40% مقارنة باستخدام الشبكة العصبية الخلفية ونموذج الخليط الغوسي الخلفي العصبية، تم إجراء اختبارات الأداء باستخدام قاعدة بيانات GRID ولقد أظهرت المجموعات أن هذا النهج يتميز بوقت تعريف أسرع ودقة أكبر مقارنة بالطريقة التقليدية. [5]
6. نظام التعرف على المتحدث باستخدام (Mel Frequency Cepstral Coefficient (MFCCs) وتكميم المتجهات كمصنف Vector Quantization (VQ) وتم تحليل النظام بتطبيقه على 120 متحدث من قاعدة بيانات TIDIGIT وتم الحصول على دقة 91% [6]
7. صمم نظام التعرف على المتحدث يعتمد عن النص، حيث يسمح للمتحدث التحدث بنص ثابت فقط، يتم تحليل الصوت باستخدام خوارزمية MFCC وتم استخدام مجموعة أصوات للتدريب ومجموعة للاختبار، واستخدمت الشبكة العصبية الاصطناعية كمصنف وتم الحصول على دقة 92% [7]
8. تم تصميم نظام التعرف على المتحدث بالاعتماد على MFCC الثابتة والدلتا ودلتا المزدوجة في تحليل الصوت واستخراج ميزات، ويتم التصنيف باستخدام طريقة الشبكة العصبية العميقة "Deep Neural Network" DNN، وتم تجريب النظام المقترح على قاعدة البيانات PRISM وأعطى نتائج جيدة. [8]
9. تم بناء نظام لتحديد هوية المتحدث، يتم تحليل الصوت باستخدام MFCC، واستخدام الشبكة العصبية التلافيفية CNN لتمييز المتحدثين، وتم تطبيق النظام على قاعدة البيانات TIMIT، وأثبتت النتائج التجريبية كفاءة الطريقة المستخدمة. [9]
10. صمم نظام التعرف على المتحدث باستخدام (Mel Frequency Cepstral Coefficients (MFCC); Shifted-Delta-Cepstral (SDC) حيث يتم التصنيف باستخدام Support vector machines (SVM) وجربت الدراسة على 10 مستخدمين (5 ذكور و5 إناث) من قاعدة بيانات IEMOCAP العاطفية. [10]
11. اقترح نظام التعرف على المتحدثين باستخدام مزيج من STRF و MFCC يتعلق نموذج spectral temporal receptive fields STRF بالتغيرات الطيفية والزمنية لتحليل الطيف السمي ويتم التصنيف باستخدام support vector machines (SVM) وتم تجريب النظام على 36 شخص واعطت معدل دقة للتعرف اعلى بـ 18.49% من استخدام MFCC بمفردها. [11]
12. تم تصميم نظام التعرف على المتحدث والتحقق منه باستخدام الشبكة العصبية الاصطناعية كمصنف، وتم اختبار النتيجة على 30 متحدثاً، في البداية: تم اختبار النظام باستخدام ميزات MFCC فقط وتم العثور على دقة تعرف منخفضة قدرها 50%، ولتحسين كفاءة النظام تم استخدام ميزات MFCC جنباً إلى جنب مع النقطة الوسطى وتحقيق دقة تعرف 65.3%، ويتم تحقيق دقة تعرف 65.7% عندما يكون MFCC كذلك جنباً إلى جنب مع الطرح الطيفي، وتم الحصول على دقة تعرف 75.3% عند دمج MFCC مع النقطة الوسطى. [12]
13. تم بناء نظام التعرف على المتحدث، حيث يتم الدمج بين ميزات MFCC و (Deep Belief Network) DBN واستخدام (Support Vector Machine (SVM) كمصنف، وتم تجريب النظام على قاعدة بيانات Urdu التي تضم 10 مستخدمين بلغة باكستانية. [13]

14. تم تصميم نظام التعرف على المتحدث المعتمد على النص، وغير المعتمد على النص، ويتم استخراج ميزات الصوت بتطبيق طريقة جديدة وذلك بالدمج بين MFCC وUMRT ويتم التصنيف بتطبيق الخوارزمية العصبية الاصطناعية، وتم تسجيل الأصوات لعدد من المستخدمين، وتفوقت طريقة الدمج سابقة الذكر على استخدام MFCC فقط وبلغت الدقة 97.91%. [14]
15. تم اقتراح نظاما للتعرف على المتحدث المستقل عن النص، باستخدام الشبكة العصبية التلافيفية CNN مع بعض التعديلات لاستيعاب الطول المتغير لميزات الصوت، تم تحليل الصوت باستخدام MFCC، وتم تطبيق النموذج على قاعدة بيانات VoxCeleb والحصول على دقة تحديد المستخدم وصلت لـ 89.5%. [15]
16. أجريت دراسة لتحديد هوية المتحدث وجنسه باستخدام نموذج الخليط الغوسي GMM، وتم تحليل الصوت باستخدام MFCC و"BFCC Barfrequenc Cepstral Coefficient" وطبق النظام على 20 مستخدم 10 ذكور و10 اناث وتفوقت MFCC على BFCC، وتم التسجيل في جلسة تسجيل واحدة. [16]
17. تم اقتراح نظام التعرف على المتحدث بشكل مستقل عن النص، ويتم الحصول على ميزات الصوت بتطبيق MFCC، ويتم تحديد المتحدث بالاعتماد على الشبكة العصبية التلافيفية CNN بنسختين، شبكة مجموعة الهندسة المرئية VGG visual geometry group والشبكة العصبية المتبقية ResNets residual neural networks، وتم تطبيق الدراسة على قاعدة بيانات VoxCeleb
- وأثبتت النتائج التجريبية تفوق الطريقة المقترحة على الطرق التقليدية التي تعتمد على أساليب التعلم العميق وبهامش كبير. [17]
18. دراسة اقترحت طريقتين (smoothed envelope compensation (SEC و full envelope compensation (FEC) لمعالجة عدم التطابق في حال كان الصوت عبارة عن صراخ، من أجل حل مشكلة عدم التطابق بين التدريب والاختبار، لأن التدريب يتم باستخدام الكلام العادي، وبالتالي يقل معدل التعرف، وتطبق هاتين الطريقتين في مرحلة تحليل الصوت واستخراج الميزات لتعديل الطيف، ليكون الصراخ أقرب الى الكلام العادي، وبالتالي تعديل طيف القدرة لـ MFCC باستخدام خرائط إحصائية تعتمد على GMM، وتم تجريب النظام على قاعدة بيانات مؤلفة من 22 شخص (11 ذكر و 11 انثى) وعطت التقنيات المقترحة تحسينات كبيرة في معدلات التعرف على المتحدث. [18]
19. دراسة للحصول على افضل دقة لنظام التعرف على المتحدث، وتم تحليل الصوت باستخدام MFCC و Linear Predictive Coding (LPC) (Coding) و Zero Crossing Rate (ZCR)، واستخدم للتمييز كمصنف Feedforward Artificial Neural Network (FFANN) و Support Vector Machine (SVM)، وتم تجريب على قاعدة بيانات لـ 20 شخص تم تسجيل أصواتهم باستخدام برامج تسجيل عبر الانترنت. [19]
20. تم دراسة شبكتين عصبيتين لتحديد الجنس والكشف عن المتحدث، وهما (multi-layers perceptron (MLP و convolutional neural networks (CNNs)، وتم تحليل الصوت باستخدام MFCC، وطبق النظام على قاعدة بيانات جاهزة، وأظهرت النتائج التجريبية تفوق (MLP) على (CNNs). [20]
21. تم بناء نظام التعرف على المتحدث المعتمد على النص، المعتمد على كلمات التنبية وتنفيذه، وينقسم لمرحلتين: مرحلة التدريب ومرحلة الاختبار، في مرحلة التدريب يتم تسجيل كلمة تنبيه، ومن ثم يتم استخدام خوارزمية MFCC، كمعالجة مسبقة لاستخراج ميزات الكلام، وبعد الحصول على خصائص الكلام، يتم استخدام نموذج الخليط الغوسي (GMM) ونموذج ماركوف المخفي (HMM) في وقت واحد (معاً) للتدريب، في مرحلة الاختبار يتم بناء (GMM) و (HMM) بشكل مستمر، ويتم حساب المسافة بين مجموع البيانات وبين الكلام المدخل المجهول، وأظهرت النتائج التجريبية أن الدقة وصلت لـ 99.84% في البيئات النظيفة من الضجيج. [21]
22. تم بناء نظام للتعرف على المتحدث باقتراح نموذج يعتمد على الشبكة العصبية التلافيفية CNN، والشبكة العصبية المتكررة RNN ويتم تحليل الصوت باستخدام خوارزمية MFCC، وتم تجريب النظام المقترح على قاعدة بيانات جاهزة (Aishel)، وحقق النظام المقترح أداء جيد في دقة التعرف وصل 98.96%، ولكن العيب في النظام المقترح يحتاج الى تحسين متانته في البيئات العملية العامة. [22]
23. تم بناء نظام التعرف على المتحدث بشكل مستقل عن النص باستخدام معاملات خوارزمية MFCC لاستخراج خصائص الإشارات الصوتية، ومن ثم حقن هذه المعاملات بالشبكة العصبية التلافيفية CNN، وتم المقارنة بين الطريقة المقترحة والطرق التقليدية الموجودة في تحديد دقة المتحدث، وتظهر النتائج التجريبية أن النظام المقترح أكثر فعالية من الطرق التقليدية، ووصل معدل دقته في تحديد المتحدث الى 97.89%. [23]
24. تم اقتراح نموذج للتعرف على المتحدث مستقل عن النص في الظروف الحقيقية الصاخبة يتحقق ذلك من خلال استخدام طريقة (CASA) "Computational auditory scene analysis" لتخلص من الضجيج، ومن ثم يتم تحليل الصوت باستخدام MFCC، ويتم دمج طريقة الخليط الغوسي المتسلسل والشبكة العصبية التلافيفية (GMMCNN)، لتحديد هوية المتحدث والتعرف على المشاعر. يقترح هذا البحث خوارزمية جديدة لتحسين دقة التعرف على المتحدث في المواقف العاطفية والظروف الحساسة للضجيج المرتفع. وتبين التجارب أن النموذج المقترح يحقق نتائج واعدة بالمقارنة مع المصنفات الأخرى وذلك بتطبيقه على قاعدة بيانات (SUSAS)، وقاعدة بيانات الكلام الإماراتي (ESD)، وقاعدة بيانات (RAVDESS). [24]

25. تم بناء نظام التعرف على المتحدث المستقل عن النص ، يتم تحليل الصوت باستخدام خوارزمية MFCC واستخدام الخوارزمية العصبية التلافيفية كمصنف، وتطبيق النظام على ثلاث قواعد بيانات جاهزة، قاعدة بيانات عربية بلكنة امارتية، وقاعدة بيانات إنكليزية (Speech Under Simulated and Actual Stress) "SUSAS" ، وقاعدة بيانات RAVDESS "Ryerson"، وقاعدة بيانات Audio-Visual Database of Emotional Speech and Song، وتم مقارنة النتائج الجريبة بتطبيق CNN كمصنف مع عدد من النتائج بتطبيق عدة مصنفات أخرى ، فكانت أفضل دقة باستخدام CNN حيث بلغت الدقة 81.6% [25].

26. تم مقارنة بين MFCC وLPC لتحليل الصوت واستخراج الميزات، في نظام تحديد المتحدث المستقل عن النص، وتم تجريب الدراسة على قاعدة البيانات Malaysian الإنكليزية، باستخدام discriminant analysis (DA) كمصنف، وتفوقت MFCC حيث وصلت دقة التعرف على المتحدث الى 99.09% مقابل 94.75% لـ LPC. [26]

طرائق البحث

تنقية الإشارة من الضجيج: يتم تنقية الإشارة الصوتية من الضجيج، بتطبيق مرشح تمرير منخفض، تردد القطع (4kHz)، وذلك لان إشارة الصوت البشرية لا يمكن أن تتجاوز الـ (4kHz)، ويتم حذف جميع الترددات التي تكون أكبر من (4kHz)، كونها ترددات لإشارة الضجيج، ويتم تطبيق مرشح وينر لحذف الضجيج المضاف بشكل جمعي الى الإشارة الصوتية.

تحليل الإشارة الصوتية: بعد تنقية الإشارة من الضجيج نقوم بتحليلها وذلك بتطبيق الدراسة الرياضية التالية:

تطبيق خوارزمية MFCC: تعد هذه الخوارزمية الأشهر في هذا المجال، وذلك بتطبيق الخطوات التالية:

- ❖ تقسيم الإشارة الصوتية الى إطارات، زمن كل إطار (10 m.s) كون الإشارة متغيرة مع الزمن فلا نستطيع معالجتها الا بتقسيمها الى فترات زمنية متساوية وتعتبر مستقرة نوعاً ما ضمن هذه الإطارات من أجل أن نتمكن من تطبيق المعادلات عليها.
- ❖ لتركيز قيمة البيانات في مركز الإطار والتقليل منها عند الحواف، يتم ضرب الإشارة بنافاذة هامينغ. [26]
- ❖ للانتقال من المجال الزمني الى المجال الترددي نطبق FFT، وذلك للحصول على تردد التوافقيات للإشارة الصوتية.
- ❖ لتحويل كل تردد تم الحصول عليه الى تردد ميل، نطبق العلاقة التقريبية التالية [26]:

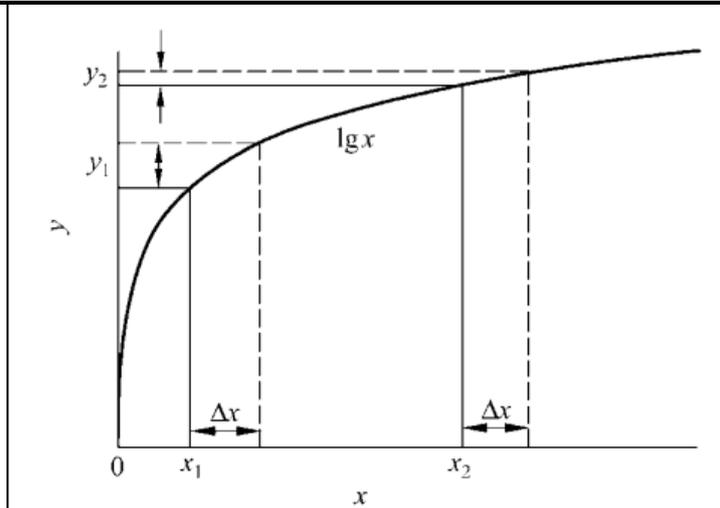
$$Mel(f) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \quad (3)$$

حيث:

- f : تردد الإشارة الصوتية.
- $Mel(f)$: قيمة تردد ميل الموافق.
- ❖ تم تطبيق عدة معادلات رياضية ودراستها احصائياً من أجل معرفة أي من هذه المعادلات هي الأفضل، ويتم باستخدامها من أجل استخلاص أفضل بصمة صوتية حقيقية في أوساط الضجيج الصاخبة، ومن هذه المعادلات الرياضية [26]:

$$Xk = \log_{10}(x)$$

وهي المعادلة الأساسية المعتمدة في خوارزمية MFCC ولكن هذه المعادلة تضم اللوغاريتم العشري، وهو حساس جداً للضجيج، فهو شديد الانحدار عندما تكون طاقة الإشارة الصوتية منخفضة، لذلك يتم الحصول على نتائج غير صحيحة عندما تكون طاقة الإشارة الصوتية منخفضة وطاقة الضجيج مرتفع كما هو موضح بالشكل (1):



الشكل (1) حساسية تابع اللوغاريتم للضجيج

لذلك نلجأ لتطبيق عدة معادلات أخرى وهي:

1. بتطبيق الجذر التربيعي على خرج مرشحات ميل:

$$X_k = \sqrt{x}$$

2. بتطبيق الجذر التكعيبي على خرج مرشحات ميل [3]:

$$X_k = \sqrt[3]{x}$$

3. بتطبيق الجذر من المرتبة الرابعة على خرج مرشحات ميل:

$$X_k = \sqrt[4]{x}$$

❖ نستخدم تحويل جيب التمام المتقطع DCT لإعادة تحويل طيف مجال ميل اللوغاريتمي إلى مجال الزمن ونتيجة هذا التحويل هو الحصول على شعاع MFCC باستخدام العلاقة [27]:

$$MFCC = \sum_{k=1}^{26} k \cdot \cos \pi \cdot i(k - 0.5)/26 \quad i = 1, 2, \dots, p \quad (5)$$

حيث:

- 26: عدد المرشحات التي قمنا باختيارها.

- X_k : قيمة الاطار.

- P: عدد معاملات MFCC الثابتة للإشارة المحللة، وفي البحث تم اختيار (P=26).

فحصل على مصفوفة التوافقيات عدد الاسطر بعدد الإطارات وعدد الاعمدة بعدد المعاملات المحللة بناء قاعدة البيانات

تم تسجيل أصوات الأشخاص الذين سيتم استخلاص بصمتهم الصوتية في بيئة صاخبة من أجل اختبار النظام المقترح، فسجلت أصوات (100) شخص (50) من الذكور و(50) من الاناث تتراوح أعمارهم من (20 الى 60) عام، قام كل واحد منهم بتسجيل عبارة باللغة العربية وهي: " السلام عليكم ورحمة الله وبركاته" مدة التسجيل (5 ثواني) حيث ان كل صوت شخص تم تسجيله (20) مرة، وعلى مجموعتين مجموعة للتدريب ومجموعة للاختبار، وبذلك اصبح عدد الملفات (2000) صوت في مجموعة التدريب و (2000) في مجموعة الاختبار، ويتم تسجيل الصوت بالتحدث أمام الميكروفون، ومن ثم يتم تسجيل إشارة دون التحدث أمام الميكروفون فيتم تسجيل إشارة الضجيج، ويتم حساب قيمة الضجيج وحذفها، وبذلك تكون الإشارة الصوتية جاهزة للتحليل، طبقت خوارزمية MFCC للحصول على التوافقيات في كل إطار. النتائج والمناقشة

بتطبيق خوارزمية MFCC التقليدية- التي تضم تابع اللوغاريتم العشري- وخوارزمية MFCC المحسنة المقترحة في هذا البحث، وذلك بتطبيق معادلة الجذر التربيعي، والجذر التكعيبي والجذر من المرتبة الرابعة وحساب نسبة التعرف، تم الحصول على النتائج التالية:

وباعتبار أن:

A: هي مجموعة الاصوات، التي يكون الفرق بين أي صوت من مجموعة الاختبار وكافة أصوات مجموعة التدريب أقل من العتبة، وذلك لنفس الشخص، أي عدد الأصوات المقبولة، وذلك بتطبيق خوارزمية MFCC التقليدية.

AA: عدد الأصوات المرفوضة رفض خاطئ، باستخدام خوارزمية MFCC التقليدية.

B: عدد الأصوات المقبولة، وذلك بتطبيق خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي.

BB: عدد الأصوات المرفوضة رفض خاطئ، باستخدام خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي.

C: عدد الأصوات المقبولة، وذلك بتطبيق خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي.

CC: عدد الأصوات المرفوضة رفض خاطئ، باستخدام خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي.

D: عدد الأصوات المقبولة، وذلك بتطبيق خوارزمية MFCC المحسنة بتطبيق معادلة الجذر من المرتبة الرابعة.

DD: عدد الأصوات المرفوضة رفض خاطئ، باستخدام خوارزمية MFCC المحسنة بتطبيق معادلة الجذر من المرتبة الرابعة.

وتطبيق النسب المئوية للبارامترات A,AA,B,BB,C,CC,D,DD

جدول (1) نتائج تفصيلية لدقة القبول والرفض الخاطئ من أجل عشرة مستخدمين

رقم المسخدم	الج نس	A	A	B	B	C	C	D	D	A%	AA%	B%	BB%	C%	CC%	D%	D%
1	ذكر	1	8	2	0	2	0	2	0	90%	10%	100%	0%	100%	0%	100%	0%
2	ذكر	1	8	2	7	3	7	1	3	90%	10%	85%	15%	85%	15%	85%	15%
3	ذكر	1	9	1	0	2	0	1	1	95%	5%	100%	0%	100%	0%	95%	5%
4	ذكر	4	6	1	8	2	4	1	9	20%	80%	90%	10%	70%	30%	45%	55%
5	ذكر	1	3	7	6	1	0	1	0	65%	35%	80%	20%	50%	50%	50%	50%
6	انثى	2	8	0	9	1	9	1	1	60%	40%	100%	0%	95%	5%	95%	5%
7	انثى	1	9	1	0	2	0	0	2	95%	5%	100%	0%	100%	0%	100%	0%
18	انثى	1	8	2	0	2	0	0	2	90%	10%	100%	0%	100%	0%	100%	0%
9	انثى	1	6	4	0	1	9	1	8	80%	20%	100%	0%	95%	5%	90%	10%
10	انثى	1	0	1	8	2	8	7	4	50%	50%	90%	10%	90%	0%	70%	30%

جدول (2) نتائج تفصيلية لدقة القبول والرفض الخاطئ من أجل عشرين مستخدمين

رقم المسخدم	الج نس	A	A	B	B	C	C	D	D	A%	AA%	B%	BB%	C%	CC%	D%	D%
1	ذكر	1	8	2	0	2	0	2	0	90%	10%	100%	0%	100%	0%	100%	0%
2	ذكر	1	8	2	7	3	7	1	3	90%	10%	85%	15%	85%	15%	85%	15%

5 %	95 %	0%	100 %	0 %	100 %	5 %	95 %	1 1	1 9	0 0	2 0	0 0	2 1	1 9	ذكر	3
55 %	45 %	30 %	70 %	10 %	90 %	80 %	20 %	1 1	9 9	6 4	1 2	2 8	1 6	1 4	ذكر	4
50 %	50 %	50 %	50 %	20 %	80 %	35 %	65 %	1 0	1 0	1 0	1 4	4 0	1 6	7 3	ذكر	5
55 %	45 %	35 %	65 %	20 %	80 %	70 %	30 %	1 1	9 9	7 3	1 4	4 6	1 4	1 6	ذكر	6
5 %	95 %	0%	100 %	0 %	100 %	15 %	85 %	1 1	1 9	0 0	2 0	0 0	2 3	3 7	ذكر	7
0 %	100 %	0%	100 %	0 %	100 %	10 %	90 %	0 0	2 0	0 0	2 0	0 0	2 2	2 1	ذكر	8
5 %	95 %	5%	95 %	5 %	95 %	10 %	90 %	1 9	1 9	1 9	1 9	1 9	1 9	2 8	ذكر	9
25 %	75 %	15 %	85 %	0 %	100 %	25 %	75 %	5 5	1 5	3 7	1 0	0 2	2 5	5 1	ذكر	10
5 %	95 %	5%	95 %	0 %	100 %	40 %	60 %	1 9	1 9	1 9	1 0	0 2	8 8	1 2	انثى	11
0 %	100 %	0%	100 %	0 %	100 %	5 %	95 %	0 0	2 0	0 0	2 0	0 0	2 1	1 9	انثى	12
0 %	100 %	0%	100 %	0 %	100 %	10 %	90 %	0 0	2 0	0 0	2 0	0 0	2 2	2 8	انثى	13
10 %	90 %	5%	95 %	0 %	100 %	20 %	80 %	2 8	1 8	1 9	1 0	0 2	4 6	1 6	انثى	14
0 %	70 %	0% 1	90 %	10 %	90 %	50 %	50 %	7 4	1 4	2 8	2 8	1 8	1 0	1 0	انثى	15
5 %	95 %	0%	100 %	0 %	100 %	10 %	90 %	1 9	1 9	0 0	2 0	0 0	2 2	2 8	انثى	16
45 %	55 %	35 %	65 %	10 %	90 %	65 %	35 %	9 1	1 1	7 3	1 2	2 8	1 3	1 7	انثى	17
0 %	90 %	5%	95 %	5 %	95 %	50 %	50 %	2 8	1 8	1 9	1 9	1 9	1 0	1 0	انثى	18
0 %	100 %	0%	100 %	0 %	100 %	0 %	100 %	0 0	2 0	0 0	2 0	0 0	2 0	0 0	انثى	19
0 %	100 %	0%	100 %	0 %	100 %	10 %	90 %	0 0	2 0	0 0	2 0	0 0	2 2	1 8	انثى	20

جدول (3) المتوسطات الحسابية لدقة قبول والرفض الخاطئ

عدد المستخ دمين في كل مجموع ة	المتوسط الحسابي لمدل الدقة باستخدام اللوغارتم	المتوسط الحسابي لمدل الدقة باستخدام معادلة الجذر التربيعي	المتوسط الحسابي لمدل الرفض الخاطئ باستخدام معادلة الجذر التربيعي	المتوسط الحسابي لمدل الرفض الخاطئ باستخدام معادلة اللوغارتم	المتوسط الحسابي لمدل الدقة باستخدام معادلة الجذر التكعيبي	المتوسط الحسابي لمدل الدقة باستخدام معادلة الجذر التكعيبي	المتوسط الحسابي لمدل الدقة باستخدام معادلة الجذر التكعيبي	المتوسط الحسابي لمدل الدقة باستخدام معادلة الجذر التكعيبي
10	14.7	18.9	17.7	16.6	5.3	1.1	2.3	3.4

3.36842	2.2105 3	1	5.4736 8	16.63158	17.7894 7	19	14.526 32	20
3.2333	2.0666 7	0.9333 3	5.7333 3	16.76667	17.9333 3	19.06 667	14.266 67	30
3.625	3	1.2	5.875	16.375	17	18.8	14.125	40
3.88333	3.16	1.52	6.34	16.06	16.84	18.48	13.66	50
3.88333	3.1166 7	1.7666 7	6.2666 7	16.11667	16.8833 3	18.23 333	13.733 33	60
3.92857	3.0857 1	1.7714 3	6.2428 6	16.07143	16.9142 9	18.22 857	13.757 14	70
4.0625	3.2	1.8125	6.25	15.9375	16.8	18.18 75	13.75	80
4.13333	3.2111 1	1.8889	6.3555 6	15.86667	16.7888 9	18.11 111	13.644 44	90
4.302080	3.3645 8	2.0208 3	6.4687 5	15.69792	16.6354 2	17.97 917	13.531 25	100

جدول (3) النسب المئوية لدقة قبول والرفض الخاطئ

عدد المستخد مين في كل مجموعة	النسبة المئوية لمعدل الدقة باستخدام اللوغارتم	النسبة المئوية لمعدل الدقة باستخدام التربيعي	النسبة المئوية لمعدل الدقة باستخدام الجزر المرتبة الرابعة	النسبة المئوية لمعدل الدقة باستخدام معادلة الجزر المرتبة الرابعة	النسبة المئوية لمعدل الدقة باستخدام معادلة الجزر المرتبة الرابعة	النسبة المئوية لمعدل الدقة باستخدام معادلة الجزر المرتبة الرابعة	النسبة المئوية لمعدل الدقة باستخدام معادلة الجزر المرتبة الرابعة	النسبة المئوية لمعدل الدقة باستخدام معادلة الجزر المرتبة الرابعة
10	73.5%	94.5%	88.50%	83%	26.50%	5.5%	11.5%	17
20	72.63%	95%	88.94%	83.15%	27.36%	5%	11.05%	16.84%
30	71.33%	95.3%	89.66%	83.83%	28.66%	4.66%	10.33%	16.16%
40	70.625%	94%	85%	81.87%	29.37%	6%	15%	18.12%
50	68.3%	92.4%	84.2%	80.30%	31.70%	7.6%	15.8%	19.41%
60	68.66%	91.16%	84.41%	80.58%	31.33%	8.83%	15.58%	19.41%
70	68.78%	91.14%	84.57%	80.35%	31.21%	8.85%	15.42%	19.64%
80	68.75%	90.93%	84%	79.68%	31.25%	9.06%	16%	20.31%
90	68.22%	90.55%	83.94%	79.33%	31.77%	9.44%	16.05%	20.66%
100	67.65%	89.85%	83.17%	78.48%	32.34%	10.10%	16.82%	21.51%

المناقشة

تم اعتبار أن العتبة المثلى، هي العتبة التي تحقق معدل قبول خاطئ يساوي الصفر، فمن خلال التجارب والتسجيلات السابقة نستنتج أن أفضل عتبة تم التوصل لها هي (0.13) من أجل تطبيق خوارزمية MFCC التقليدية، والعتبة (0.11) من أجل تطبيق خوارزمية MFCC بالأشكال المحسنة، نستنتج مما سبق أن أفضل نسبة تعرف تم الحصول عليها بتطبيق النسخة المحسنة ل MFCC وذلك بتطبيق معادلة الجذر التربيعي، ومن ثم تطبيق معادلة الجذر التكعيبي، ومن ثم تطبيق معادلة الجذر من المرتبة الرابعة، وحلت نتائج تطبيق خوارزمية MFCC التقليدية بالمرتبة الأخيرة.

فمن أجل 10 مستخدمين، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.5%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 21%.

أما من أجل 20 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.06%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.85%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.37%.

ومن أجل 30 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 5.64%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.47%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 23.47%.

ومن أجل 40 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 9%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 12.13%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 23.38%.

ومن أجل 50 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 8.2%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 12.1%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 24.1%.

ومن أجل 60 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.75%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 10.58%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.5%.

ومن أجل 70 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.57%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 10.79%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.36%.

ومن أجل 80 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.93%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.25%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.18%.

ومن أجل 90 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.61%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.22%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.33%.

أما من أجل 100 مستخدم، تفوقت خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي على الخوارزمية بتطبيق الجذر التكعيبي بفارق 6.68%، كما تفوقت على خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التكعيبي بفارق 11.37%، كما تفوقت على خوارزمية MFCC التقليدية بفارق 22.2%.

وبالتالي فإن أفضل معدل تعرف تم الحصول عليه بتطبيق خوارزمية MFCC المحسنة بتطبيق معادلة الجذر التربيعي ومن ثم تطبيق معادلة الجذر التكعيبي ومن ثم تطبيق معادلة الجذر من المرتبة الرابعة، وتوقفت جميعها على خوارزمية MFCC التقليدية.

أما معدل الرفض الخاطئ:

فمن أجل 10 مستخدمين، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 5.5%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.5%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 21%.

ومن أجل 20 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 5%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.05%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.84%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.36%.

ومن أجل 30 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 4.66%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 5.67%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.5%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 24%.

ومن أجل 40 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 6%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 9%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 12.12%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 23.37%.

ومن أجل 50 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 7.6%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 8.2%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.81%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 24.1%.

ومن أجل 60 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 8.83%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.75%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 10.58%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.5%.

ومن أجل 70 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 8.85%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.57%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 10.79%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.36%.

ومن أجل 80 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 9.06%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.94%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.25%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.19%.

ومن أجل 90 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 9.44%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.61%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.22%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.33%.

ومن أجل 100 مستخدم، أعطت خوارزمية MFCC المحسنة بتطبيق الجذر التربيعي أصغر معدل رفض خاطئ 10.1%، وهو أقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر التكعيبي بمقدار 6.72%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC المحسنة باستخدام الجذر من المرتبة الرابعة بمقدار 11.41%، وأقل من معدل الرفض الخاطئ باستخدام خوارزمية MFCC التقليدية بمقدار 22.24%.

مقارنة الدراسة الحالية مع دراسات سابقة

يقارن هذا المقطع النتائج التي تم التوصل إليها في هذه الدراسة، والنتائج في بعض الدراسات السابقة، باستخدام خوارزميات تحليل صوت مختلفة، وخوارزميات اختبار، فكانت نتائج معدل التعرف على المستخدم، عند الأعداد المذكورة، وباستخدام الطرق الموضحة كما في الجدول التالي:

عدد المستخدمين في كل مجموعة	معدل الدقة باستخدام ام اللوغار تم	معدل الدقة باستخدام الجذر التربيعي	معدل الدقة باستخدام الجذر التكعيبي	معدل الدقة باستخدام م معادلة الجذر من المرتبة الرابعة	MFC C و ANN [7]	MFC C [12]	MFCC +SS [12]	MFCC+ Centroid [12]
10	73.5%	94.5%	88.50%	83%	92%	84.7%	85.3%	85.6%
20	72.63%	95%	88.94%	83.15%	82%	75.2%	77.3%	78.3%
30	71.33%	95.3%	89.66%	83.83%	76%	50%	60%	65.3%
40	70.625%	94%	85%	81.87%	72%	-----	-----	-----
50	68.3%	92.4%	84.2%	80.30%	70%	-----	-----	-----
60	68.66%	91.16%	84.41%	80.58%	-----	-----	-----	-----
70	68.78%	91.14%	84.57%	80.35%	-----	-----	-----	-----
80	68.75%	90.93%	84%	79.68%	-----	-----	-----	-----
90	68.22%	90.55%	83.94%	79.33%	-----	-----	-----	-----
100	67.65%	89.85%	83.17%	78.48%	-----	-----	-----	-----

1. استخدمت الدراسة المرجعية من أجل التحقق من هوية المتكلمين، خوارزمية MFCC، وذلك من أجل الحصول على ميزات الصوت، وهي نفسها الخوارزمية التقليدية المستخدمة في هذه الدراسة، واستخدمنا خوارزمية ANN وعلى مرحلتين، مرحلة للتدريب، ومرحلة

للاختبار، وتم تطبيق الدراسة على مجموعات مختلفة الأعداد من المتكلمين، تم تسجيل أصواتهم، وكانت النتائج كما هو موضح بالجدول السابق، تفوقت النتائج التي تم التوصل إليها في هذا البحث على الدراسة المرجعية المذكورة، فكان من أجل أعداد المستخدمين المذكورة (10,20,30,40,50)، معدل التعرف أفضل بـ (2.5%، 13%، 19.3%، 22%، 22.4%) باستخدام معادلة الجذر التربيعي، أما باستخدام معادلة الجذر التكعيبي والجذر من المرتبة الرابعة والخوارزمية التقليدية كانت نتائج الدراسة المرجعية أفضل [7].

2. وبمقارنة نتائج الدراسة الحالية مع البحث المرجعي الذي تم فيه تصميم نظام التعرف على المتحدث والتحقق منه باستخدام الشبكة العصبية الاصطناعية كمصنف، وتم اختبار النتيجة على (10, 20, 30) متحدثاً، في البداية: تم اختبار النظام باستخدام ميزات MFCC فقط وتم العثور على دقة تعرف منخفضة قدرها 50%، ولتحسين كفاءة النظام تم استخدام ميزات MFCC جنباً إلى جنب مع النقطة الوسطى وتحقيق دقة تعرف 65.3%، ويتم تحقيق دقة تعرف 65.7% عندما يكون MFCC كذلك جنباً إلى جنب مع الطرح الطيفي، وتم الحصول على دقة تعرف 75.3% عند دمج MFCC مع النقطة الوسطى. [12]

تفوقت النتائج التي تم التوصل إليها في هذا البحث على الدراسة المرجعية المذكورة عندما استخدم الخوارزمية MFCC، فكان من أجل أعداد المستخدمين المذكورة (10,20,30)، معدل التعرف أفضل بـ (9.8%، 19.8%، 45.3.4%) باستخدام معادلة الجذر التربيعي، أما باستخدام معادلة الجذر التكعيبي كانت النتائج أفضل بـ (3.8%، 13.74%، 39.66%) عند نفس أعداد المستخدمين وعلى الترتيب، وباستخدام معادلة الجذر من المرتبة الرابعة كانت النتائج من أجل 10 مستخدمين تفوقت نتائج الدراسة المرجعية المذكورة على هذه الدراسة بـ (1.7%)

أما من أجل (20,30) مستخدم تفوقت نتائج البحث الحالي على الدراسة المرجعية بـ (7.95%، 33.83%)، أما باستخدام خوارزمية MFCC التقليدية، فمن أجل استخدام (10,20) مستخدم تفوقت نتائج الدراسة المرجعية على البحث الحالي بالقيم التالية: (11.2%، 2.57%) على الترتيب، أما عند استخدام 30 مستخدم تفوقت الدراسة الحالية بـ (21.33%) على الدراسة المرجعية، ولتحسين النتائج استخدمت الدراسة المرجعية النقطة الوسطى جنباً إلى جنب مع خوارزمية MFCC، ولكن تفوقت نتائج هذه الدراسة الحالية باستخدام معادلة الجذر التربيعي على الدراسة المرجعية، عند استخدام (10,20,30) مستخدم بالقيم (9.2%، 17.7%، 35.3%)، كما تفوقت هذه الدراسة أيضاً باستخدام معادلة الجذر التكعيبي بالقيم (2.3%، 11.64%، 29.66%)، أما باستخدام معادلة الجذر من المرتبة الرابعة، من أجل 10 مستخدمين تفوقت الدراسة المرجعية بـ (2.3%) على الدراسة الحالية، أما من أجل (20,30) مستخدم تفوقت الدراسة الحالية على المرجعية بـ (5.85%، 23.83%) على الترتيب، وعند استخدام خوارزمية MFCC باستخدام معادلة اللوغاريتم، من أجل (10,20) مستخدم تفوقت الدراسة المرجعية على الدراسة الحالية بـ (4.67%، 11.8%) على الترتيب، أما عند استخدام (30) تفوقت الدراسة الحالية بـ (11.33%) على الدراسة المرجعية، وبالمقارنة بين الدراسة الحالية والدراسة المرجعية باستخدام MFCC والطرح الطيفي، تفوقت الدراسة الحالية باستخدام معادلة الجذر التربيعي، على الدراسة المرجعية عند مجموعات المستخدمين (10,20,30) بالقيم التالية (8.9%، 16.7%، 30%)، وكذلك الأمر تفوقت الدراسة الحالية باستخدام معادلة الجذر التكعيبي، على الدراسة المرجعية عند نفس مجموعات المستخدمين بالقيم التالية (2.9%، 10.64%، 24.36%)، أما باستخدام معادلة الجذر من المرتبة الرابعة، من أجل 10 مستخدمين تفوقت الدراسة المرجعية بـ (2.6%) على الدراسة الحالية، أما من أجل (20,30) مستخدم تفوقت الدراسة الحالية على المرجعية بـ (4.85%، 18.53%) على الترتيب، وعند استخدام خوارزمية MFCC باستخدام معادلة اللوغاريتم، من أجل (10,20) مستخدم تفوقت الدراسة المرجعية على الدراسة الحالية بـ (5.67%، 12.1%) على الترتيب، أما عند استخدام (30) تفوقت الدراسة الحالية بـ (6.03%) على الدراسة المرجعية.

الاستنتاجات

- أدى تطبيق تابع الجذر التربيعي، والتكعيبي والجذر من المرتبة الرابعة الى تحسين معدل التعرف على المستخدم، وتقليل معدل الرفض الخاطئ دون التأثير على دقة النظام.
- بتطبيق معادلة الجذر التربيعي تم الحصول على أفضل معدل تعرف، وبالتالي تحسين خوارزمية MFCC.
- تم تحسين MFCC من أجل استخدامها في أوساط الضجيج الصاخبة جداً.

- Wu, Z., & Cao, Z. (2005). Improved MFCC-based feature for robust speaker identification. *Tsinghua Science & Technology*, 10(2), 158-161. .1
- Kamruzzaman, S. M., Karim, A. N. M., Islam, M. S., & Haque, M. E. (2010). Speaker identification using mfcc-domain support vector machine. *arXiv preprint arXiv:1009.4972* .2
- Zhao, X., & Wang, D. (2013, May). Analyzing noise robustness of MFCC and GFCC features in speaker identification. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 7204-7208). IEEE. .3
- Zhao, X., Wang, Y., & Wang, D. (2014). Robust speaker identification in noisy and reverberant conditions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(4), 836-845. .4
- Almaadeed, N., Aggoun, A., & Amira, A. (2015). Speaker identification using multimodal neural networks and wavelet analysis. *Iet Biometrics*, 4(1), 18-28. .5
- Bharti, R., & Bansal, P. (2015). Real time speaker recognition system using MFCC and vector quantization technique. *International Journal of Computer Applications*, 117(1). .6
- Nandyal, S., Wali, S. S., & Hatture, S. M. (2015). MFCC based text-dependent speaker identification using BPNN. *International Journal of Signal Processing Systems*, 3(1), 30-34. .7
- Matějka, P., Glembek, O., Novotný, O., Plchot, O., Grézl, F., Burget, L., & Cernocký, J. H. (2016, March). Analysis of DNN approaches to speaker identification. In *2016 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 5100-5104). IEEE .8
- Lukic, Y., Vogt, C., Dürr, O., & Stadelmann, T. (2016, September). Speaker identification and clustering using convolutional neural networks. In *2016 IEEE 26th international workshop on machine learning for signal processing (MLSP)* (pp. 1-6). IEEE. .9
- Mansour, A., & Lachiri, Z. (2017). SVM based emotional speaker recognition using MFCC-SDC features. *International Journal of Advanced Computer Science and Applications*, 8(4). .10
- Wang, J. C., Wang, C. Y., Chin, Y. H., Liu, Y. T., Chen, E. T., & Chang, P. C. (2017). Spectral-temporal receptive fields and MFCC balanced feature extraction for robust speaker recognition. *Multimedia Tools and Applications*, 76, 4055-4068. .11
- Chauhan, N., & Chandra, M. (2017, March). Speaker recognition and verification using artificial neural network. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 1147-1149). IEEE. .12
- Ali, H., Tran, S. N., Benetos, E., & d'Avila Garcez, A. S. (2018). Speaker recognition with hybrid features from a deep belief network. *Neural Computing and Applications*, 29, 13-19. .13
- Antony, A., & Gopikakumari, R. (2018). Speaker identification based on combination of MFCC and UMRT based features. *Procedia computer science*, 143, 250-257. .14
- Yadav, S., & Rai, A. (2018, September). Learning discriminative features for speaker identification and verification. In *Interspeech* (pp. 2237-2241). .15
- Kumar, C., Ur Rehman, F., Kumar, S., Mehmood, A., & Shabir, G. (2018, March). Analysis of MFCC and BFCC in a speaker identification system. In *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (pp. 1-5). IEEE. .16
- An, N. N., Thanh, N. Q., & Liu, Y. (2019). Deep CNNs with self-attention for speaker identification. *IEEE access*, 7, 85327-85337. .17
- Jokinen, E., Saeidi, R., Kinnunen, T., & Alku, P. (2019). Vocal effort compensation for MFCC feature extraction in a shouted versus normal speaker recognition task. *Computer Speech & Language*, 53, 1-11. .18
- Chauhan, N., Isshiki, T., & Li, D. (2019, February). Speaker recognition using LPC, MFCC, ZCR features with ANN and SVM classifier for large input database. In *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)* (pp. 130-133). IEEE. .19
- Mamrybayev, O., Toleu, A., Tolegen, G., & Mekebayev, N. (2020). Neural architectures for gender detection and speaker identification. *Cogent engineering*, 7(1), 1727168. .20

- Tsai, T. H., Hao, P. C., & Wang, C. L. (2021). Self-defined text-dependent wake-up-words speaker .21
recognition system. *IEEE Access*, 9, 138668-138676.
- Ye, F., & Yang, J. (2021). A deep neural network model for speaker identification. *Applied* .22
Sciences, 11(8), 3603.
- Abdiche, D., & Harrar, K. (2021, February). Text-independent speaker identification using mel-.23
frequency energy coefficients and convolutional neural networks. In *2020 2nd International
Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)* (pp. 204-
209). IEEE.
- Nassif, A. B., Shahin, I., Hamsa, S., Nemmour, N., & Hirose, K. (2021). CASA-based speaker .24
identification using cascaded GMM-CNN classifier in noisy and emotional talking
conditions. *Applied Soft Computing*, 103, 107141.
- Shahin, I., Nassif, A. B., & Hindawi, N. (2021). Speaker identification in stressful talking .25
environments based on convolutional neural network. *International Journal of Speech
Technology*, 24, 1055-1066.
- Zailan, M. K. N., Mohd Ali, Y., Noorsal, E., Abdullah, M. H., Saad, Z., & Mat Leh, A. (2023). .26
Comparative analysis of LPC and MFCC for male speaker recognition in text-independent
context. *ESTEEM Academic Journal*, 19, 101-112.
- Chen, Y. (2022). A hidden Markov optimization model for processing and recognition of English .27
speech feature signals. *Journal of Intelligent Systems*, 31(1), 716-725.