مجلة
بحوث الشرق الأوسط

مجلة علمية مُحَكَّمَة
(مُعتمدة) شهريـًا

العدد مائة وخمسة عشر

(سبتمبر 2025)

يصدرها
مركز بحوث
الشرق الأوسط

السنة الخمسون
تأسست عام 1974

الترقيم الدولي: (2536-9504)
الترقيم على الإنترنت: (2735-5233)

الأراء الواردة داخل المجلة تعبر عن وجهة نظر أصحابها وليست مسئولية مركز بحوث الشرق الأوسط والدراسات المستقبلية

رقم الإيداع بدار الكتب والوثائق القومية : ٢٤٣٣٠ / ٢٠١٦

الترقيم الدولي: (Issn :2536 - 9504)

الترقيم على الإنترنت: (Online Issn :2735 - 5233)

ISSN 2536-9504

مجلة بحوث الشرق الأوسط

مجلة علمية مُحكَّمة
متخصصة
في شؤون الشرق الأوسط

مجلة مُعتمَدة من بنك المعرفة المصري

موقع المجلة على بنك المعرفة المصري
www.mercj.journals.ekb.eg

- معتمدة من الكشاف العربي للاستشهادات المرجعية (ARCI) . المتوافقة مع قاعدة بيانات كلاريفيت Clarivate الفرنسية.
- معتمدة من مؤسسة أرسيف (ARCIf) للاستشهادات المرجعية للمجلات العلمية العربية ومعامل التأثير المتوافقة مع المعايير العالمية.
- تنشر الأعداد تباعًا على موقع دار المنظومة.

دار المنظومة
DAR ALMANDUMAH
البوابة الأولى للمعلومات العربية

العدد مائة وخمسة عشر (سبتمبر 2025)

تصدر شهريًا

السنة الخمسون ـ تأسست عام 1974

المطبعة | مطبعة جامعة عين شمس
Ain Shams University Press

مجلة بحوث الشرق الأوسط
(مجلة مُعتمدة) دوريَّة علميَّة مُحكَّمَة
(اثنا عشر عددًا سنويًّا)
يصدرها مركز بحوث الشرق الأوسط والدراسات المستقبلية - جامعة عين شمس

MERC

رئيس مجلس الإدارة

أ.د. غادة فاروق

نائب رئيس الجامعة لشؤون خدمة المجتمع وتنمية البيئة ورئيس مجلس إدارة المركز

رئيس التحرير **د. حاتم العبد**

مدير مركز بحوث الشرق الأوسط والدراسات المستقبلية

هيئة التحرير

أ.د. السيد عبدالخالق، وزير التعليم العالي الأسبق، مصر

أ.د. أحمد بهاء الدين خيري، نائب وزير التعليم العالي الأسبق، مصر ؛

أ.د. محمد حسام لطفي، جامعة بني سويف، مصر ؛

أ.د. سعيد المصري، جامعة القاهرة، مصر ؛

أ.د. سوزان القليني، جامعة عين شمس، مصر ؛

أ.د. ماهر جميل أبوخوات، عميد كلية الحقوق، جامعة كفرالشيخ، مصر ؛

أ.د. أشرف مؤنس، جامعة عين شمس، مصر ؛

أ.د. حسام طنطاوي، عميد كلية الآثار، جامعة عين شمس، مصر ؛

أ.د. محمد إبراهيم الشافعي، وكيل كلية الحقوق، جامعة عين شمس، مصر

أ.د. تامر عبدالمنعم راضي، جامعة عين شمس، مصر ؛

أ.د. هاجر قلديش، جامعة قرطاج، تونس ؛

Prof. Petr MUZNY، جامعة جنيف، سويسرا ؛

Prof. Gabrielle KAUFMANN-KOHLER، جامعة جنيف، سويسرا ؛

Prof. Farah SAFI، جامعة كليرمون أوفيرني، فرنسا؛

إشراف إداري

أ/ أماني جرجس

أمين المركز

إشراف فني

د/ امل حسن

رئيس وحدة التخطيط و المتابعة

سكرتارية التحرير

أ/ رانـدا نـوار — قسـم النشـر

أ/ شيماء بكر — قسـم النشـر

المحرر الفني

أ/ مرفت حافظ

رئيس وحدة الدعم الفني

تدقيق ومراجعة لغوية

وحدة التدقيق اللغوي - كلية الآداب - جامعة عين شمس

تصميم الغلاف أ/ أحمد محسن - مطبعة الجامعة

ترجى المراسلات الخاصة بالمجلة إلى: د. حاتم العبد، رئيس التحرير merc.director@asu.edu.eg

• وسائل التواصل:

البريد الإلكتروني لوحدة النشر: merc.pub@asu.edu.eg

جامعة عين شمس- شارع الخليفة المأمون- العباسية- القاهرة، جمهورية مصر العربية، ص.ب: 11566

(وحدة النشر - وحدة الدعم الفني) موبايل / واتساب: 01555343797 (2+)

ترسل الأبحاث من خلال موقع المجلة على بنك المعرفة المصري: www.mercj.journals.ekb.eg

ولن يلتفت إلى الأبحاث المرسلة عن طريق آخر

مجلة بحوث الشرق الأوسط

- رئيس التحرير **د. حاتم العبد**

- الهيئة الاستشارية المصرية وفقًا للترتيب الهجائي:

• أ.د. إبراهيم عبد المنعم سلامة أبو العلا	رئيس قسم التاريخ - كلية الآداب - جامعة الأسكندرية - مصر
• أ.د. أحمد الشربيني	عميد كلية الآداب السابق - جامعة القاهرة - مصر
• أ.د. أحمد رجب محمد علي رزق	عميد كلية الآثار - جامعة القاهرة - مصر
• أ.د. السيد فليفل	عميد كلية الدراسات الأفريقية العليا الأسبق- جامعة القاهرة - مصر
• أ.د. إيمان محمد عبد المنعم عامر	أستاذ التاريخ الحديث والمعاصر- كلية الآداب - جامعة القاهرة - مصر
أ.د. أيمن فؤاد سيد	رئيس الجمعية المصرية للدراسات التاريخية - مصر
• أ.د. جمال شفيق أحمد عامر	كلية الدراسات العليا للطفولة - جامعة عين شمس - مصر
• أ.د. حمدي عبد الرحمن	عميد كلية الحقوق الأسبق - جامعة عين شمس - مصر
• أ.د. حنان كامل متولي	(قائم بعمل) عميد كلية الآداب - جامعة عين شمس - مصر
• أ.د. صالح حسن المسلوت	أستاذ التاريخ والحضارة - كلية اللغة العربية - فرع الزقازيق جامعة الأزهر – مصر وعضو اللجنة العلمية الدائمة لترقية الأساتذة
• أ.د. عادل عبد الحافظ عثمان حمزة	كلية الآداب - جامعة المنيا، ومقرر لجنة الترقيات بالمجلس الأعلى للجامعات - مصر
• أ.د. عاصم الدسوقي	عميد كلية الآداب الأسبق - جامعة حلوان - مصر
• أ.د. عبد الحميد شلبي	كلية اللغة العربية بالمنصورة - جامعة الأزهر - مصر
• أ.د. عفاف سيد صبره	كلية الدراسات الإنسانية بنات بالقاهرة - جامعة الأزهر - مصر
• أ.د.عفيفي محمود إبراهيم	كلية الآداب - جامعة بنها - مصر
• أ.د. فتحي الشرقاوي	نائب رئيس جامعة عين شمس الأسبق - مصر
• أ.د. محمد الخزامي محمد عزيز	عميد كلية العلوم الاجتماعية والإنسانية - جامعة الجلالة - مصر
أ.د. محمد السعيد أحمد	كلية التربية - جامعة عين شمس - مصر
• لواء/ محمد عبد المقصود	رئيس مركز المعلومات ودعم اتخاذ القرار بمجلس الوزراء - مصر
• أ.د. محمد مؤنس عوض	كلية الآداب - جامعة عين شمس - مصر
• أ.د. مدحت محمد محمود أبو النصر	كلية الخدمة الاجتماعية - جامعة حلوان قطاع الخدمة الاجتماعية بالمجلس الأعلى للجامعات ورئيس لجنة ترقية الأساتذة
• أ.د. مصطفى محمد البغدادي	كلية التربية - جامعة عين شمس - مصر
• أ.د. نبيل السيد الطوخي	رئيس قسم التاريخ - كلية الآداب - جامعة المنيا - مصر
أ.د. نهى عثمان عبد اللطيف عزمي	كلية السياحة والفنادق - جامعة مدينة السادات - مصر

- الهيئة الاستشارية العربية والدولية وفقًا للترتيب الهجائي:

- أ.د. إبراهيم خليل العلاّف — جامعة الموصل - العراق
- أ.د. إبراهيم محمد بن حمد المزيني — كلية العلوم الاجتماعية - جامعة الإمام محمد بن سعود الإسلامية - السعودية
- أ.د. أحمد الحسو — جامعة مؤتة - الأردن
  مركز الحسو للدراسات الكمية والتراثية - إنجلترا
- أ.د. أحمد عمر الزيلعي — جامعة الملك سعود - السعودية
  الأمين العام لجمعية التاريخ والآثار التاريخية
- أ.د. عبد الله حميد العتابي — كلية التربية للبنات - جامعة بغداد - العراق
- أ.د. عبد الله سعيد الغامدي — جامعة أم القرى - السعودية
  عضو مجلس كلية التاريخ، ومركز تحقيق التراث بمعهد المخطوطات
- أ.د. فيصل عبد الله الكندري — جامعة الكويت - الكويت
- أ.د. مجدي فارح — رئيس قسم الماجستير والدراسات العليا - جامعة تونس١ - تونس
- أ.د. محمد بهجت قبيسي — جامعة حلب - سوريا
- أ.د. محمود صالح الكروي — كلية العلوم السياسية - جامعة بغداد - العراق

- *Prof. Dr. Albrecht Fuess* — Center for near and Middle Eastern Studies, University of Marburg, Germany
- *Prof. Dr. Andrew J. Smyth* — Southern Connecticut State University, USA
- *Prof. Dr. Graham Loud* — University Of Leeds, UK
- *Prof. Dr. Jeanne Dubino* — Appalachian State University, North Carolina, USA
- *Prof. Dr. Thomas Ashbridge* — Queen Mary University of London, UK
- *Prof. Ulrike Freitag* — Institute of Islamic Studies, Belil Fric University, Germany

شروط النشر بالمجلة

- تُعنى المجلة بنشر البحوث المهتمة بمجالات العلوم الإنسانية والأدبية ؛
- يعتمد النشر على رأي أثنين من المحكمين المتخصصين ويتم التحكيم إلكترونيًا ؛
- تُقبل البحوث باللغة العربية أو بإحدى اللغات الأجنبية، وتُرسل إلى موقع المجلة على بنك المعرفة المصري ويرفق مع البحث ملف بيانات الباحث يحتوي على عنوان البحث باللغتين العربية والإنجليزية واسم الباحث والتايتل والانتماء المؤسسي باللغتين العربية والإنجليزية، ورقم واتساب، وإيميل الباحث الذي تم التسجيل به على موقع المجلة ؛
- يُشار إلى أن الهوامش والمراجع في نهاية البحث وليست أسفل الصفحة ؛
- يكتب الباحث ملخص باللغة العربية واللغة الإنجليزية للبحث صفحة واحدة فقط لكل ملخص، ومقدمة للبحث؛
- بالنسبة للبحث باللغة العربية يكتب على برنامج "word" ونمط الخط باللغة العربية "Simplified Arabic" وحجم الخط 14 ولا يزيد عدد الأسطر في الصفحة الواحدة عن 25 سطر والهوامش والمراجع خط Simplified Arabic حجم الخط 12 ؛
- بالنسبة للبحث باللغة الإنجليزية يكتب على برنامج word ونمط الخط Times New Roman وحجم الخط 13 ولا يزيد عدد الأسطر عن 25 سطر في الصفحة الواحدة والهوامش والمراجع خط Times New Roman حجم الخط 1؛
- مواصفات التنسيق على الترويسة (Paper) مقاس الورق (B5) 17.6 × 25 سم، (Margins) الهوامش 2.3 سم يمينًا ويسارًا، 2 سم أعلى وأسفل الصفحة، ليصبح مقاس البحث فعلي (الكلام) 13×21 سم. (Layout) والنسق: (Header) الرأس1.25 سم، (Footer) تذييل 2.5 سم ؛
- مواصفات الفقرة للبحث : بداية الفقرة First Line = 1.27 سم، قبل النص= 0.00، بعد النص = 0.00)، تباعد قبل الفقرة = (6pt) تباعد بعد الفقرة = (0pt)، تباعد الفقرات (مفرد single) ؛
- مواصفات الفقرة للهوامش والمراجع : يوضع الرقم بين قوسين هلالي مثل : (1)، بداية الفقرة Hanging = 0.6 سم، قبل النص=0.00، بعد النص = 0.00)، تباعد قبل الفقرة = 0.00 تباعد بعد الفقرة = 0.00، تباعد الفقرات (مفرد single) ؛
- الجداول والأشكال: يتم وضع الجداول والأشكال إما في صفحات منفصلة أو وسط النص وفقًا لرؤية الباحث، على أن يكون عرض الجدول أو الشكل لا يزيد عن 13.5 سم بأي حال من الأحوال ؛
- مدة التحكيم 15 يوم على الأكثر من قبول المحكمين على الموقع، مدة تعديل البحث بعد التحكيم 15 يوم على الأكثر ؛
- يخضع تسلسل نشر البحوث في أعداد المجلة حسب ما تراه هيئة التحرير من ضرورات علمية وفنية ؛
- المجلة غير ملزمة بإعادة البحوث إلى أصحابها سواء نُشرت أم لم تُنشر ؛
- تُعبر البحوث عن آراء أصحابها وليس عن رأي رئيس التحرير وهيئة التحرير ؛
- رسوم التحكيم للمصريين 650 جنيه، ولغير المصريين 155 دولار ؛
- رسوم النشر عن الصفحة الواحدة للمصريين ٣٣ جنيه، وغير المصريين ١٥ دولار ؛
- رسوم التعديل عن الصفحة الواحدة 2 جنيه ؛
- الباحث المصري يسدد الرسوم بالجنيه المصري (بالفيزا) بمقر المركز(المقيم بمحافظة القاهرة)، أو على حساب حكومي رقم : (9/450/80772/8) بنك مصر (المقيم خارج محافظة القاهرة) ؛
- الباحث غير المصري يسدد الرسوم بالدولار على حساب حكومي رقم : (EG710001000100000004082175917) (البنك العربي الأفريقي) ؛ استلام إفادة قبول نشر البحث في خلال 15 يوم على الأكثر من تاريخ سداد رسوم النشر مع ضرورة رفع إيصالات السداد على موقع المجلة؛
- المراسلات : توجه المراسلات الخاصة بالمجلة إلى : merc.director@asu.edu.eg

السيد الدكتور/ مدير مركز بحوث الشرق الأوسط والدراسات المستقبلية، ورئيس تحرير المجلة جامعة عين شمس– العباسية– القاهرة – ج. م.ع (ص. ب 11566) للتواصل والاستفسار عن كل ما يخص الموقع : محمول / واتساب: 01555343797 (+2)

(قسم النشر merc.pub@asu.edu.eg) رُسل الأبحاث من خلال موقع المجلة على بنك المعرفة المصري: www.mercj.journals.ekb.eg ولن يلتفت إلى الأبحاث المرسلة عن طريق آخر.

الرؤيـــة

السعي لتحقيق الريادة في النشر العلمي المتميز في المحتوى والمضمون والتأثير والمرجعية في مجالات منطقة الشرق الأوسط وأقطاره .

الرسالـــة

نشر البحوث العلمية الأصيلة والرصينة والمبتكرة في مجالات الشرق الأوسط وأقطاره في مجالات اختصاص المجلة وفق المعايير والقواعد المهنية العالمية المعمول بها في المجلات المُحَكَّمَة دوليًا.

الأهـــداف

- نشر البحوث العلمية الأصيلة والرصينة والمبتكرة .
- إتاحة المجال أمام العلماء والباحثين في مجالات اختصاص المجلة في التاريخ والجغرافيا والسياسة والاقتصاد والاجتماع والقانون وعلم النفس واللغة العربية وآدابها واللغة الانجليزية وآدابها ، على المستوى المحلى والإقليمي والعالمي لنشر بحوثهم وإنتاجهم العلمي .
- نشر أبحاث كبار الأساتذة وأبحاث الترقية للسادة الأساتذة المساعدين والسادة المدرسين بمختلف الجامعات المصرية والعربية والأجنبية ،
- تشجيع ونشر مختلف البحوث المتعلقة بالدراسات المستقبلية والشرق الأوسط وأقطاره ،
- الإسهام في تنمية مجتمع المعرفة في مجالات اختصاص المجلة من خلال نشر البحوث العلمية الرصينة والمتميزة .

محتويات **العدد** (115)

	عنوان البحث		الصفحة
	الدراسات القانونية		
1	تاريخ التنظيم القضائي في مصر	محمود أمين عبد الحافظ	28-1
2	نظرة حول قوانين الاستثمار المصرية المتعاقبة ماذا أضاف قانون الاستثمار رقم ٧٢ لسنة ٢٠١٧	إسلام مراد جابر مبارك	78-29
3	مفاهيم ومعايير الحكومة في المشروعات الاقتصادية	مؤمن صلاح صالح محمد نصير	122-79
4	الذكاء الاصطناعي ومشكلة البطالة في مصر الواقع والتحديات	محمد حمزة منصور	162-123
5	مثال عملي لتطبيق منازعة تنفيذ أمام القضاء الدستوري	محمد أحمد المهدي محمد	202-163
6	دور التكنولوجيا الرقمية في مواجهة تحديات المياه في مصر (دراسة تحليلية مقارنة في ضوء تجارب دولية مختارة)	محمد عبد الفتاح أحمد البارودي	308-203
	دراسات اللغة العربية		
7	صناعة الذاكرة في اللوحة الإشهارية لرواية الزيني بركات (دراسة سيميائية ثقافية وبيولوجية)	فاطمة حسن حامد	346-309
	دراسات الصوتيات		
8	استخدام الذكاء الاصطناعي في تحديد هوية المتحدث في اللغة العربية: نظرة عامة على الأسس النظرية	ايمان محمد محمد قاسم	376-347
	الدراسات الفنية		
9	البلاغة السردية في الخطاب الاتصالي لتصاميم الملصقات الإرشادية	وفاء جاسم محمد	418-377
	الدراسات اليونانية واللاتينية		

10	قصر التيه (اللابيرينثوس) ومعبد اللابيرنت بين الرؤية الأسطورية والحقيقة التاريخية "دراسة أسطورية أثرية مقارنة"	نسرين أمير سيد	510-419
	دراسات اللغة الانجليزية		
11	جيبوليتكية الوطن في سير غادة كرمي البحث عن فاطمة والعودة	زينب سعيد مصطفى	534-511

افتتاحية العدد (115)

يسعد مجلة بحوث الشرق الأوسط أن تصدر عددها الجديد (115) الذي يضم مجموعة من الدراسات والبحوث العلمية المتنوعة، والتي تسعى إلى إثراء الفكر الأكاديمي وتقديم رؤى تحليلية ومعرفية في مجالات القانون والاقتصاد والسياسة والتكنولوجيا والعلوم الإنسانية.

يُفتتح هذا العدد ببحث يتناول تاريخ التنظيم القضائي في مصر، مسلطًا الضوء على المراحل المتعاقبة لتطور النظام القضائي ودوره في ترسيخ مبادئ العدالة وسيادة القانون. ويتبعه عرض تحليلي لـ قوانين الاستثمار المصرية المتعاقبة، وما أفرزته من تحديات وفرص في البيئة الاقتصادية الوطنية، وصولًا إلى دراسة متخصصة بعنوان: ماذا أضاف قانون الاستثمار رقم 72 لسنة 2017؟، حيث يتم تقييم ما جاء به من مستجدات تشريعية لدعم بيئة الأعمال وتحفيز الاستثمار.

كما يتضمن العدد مقالة علمية حول مفاهيم ومعايير الحوكمة في المشروعات الاقتصادية، باعتبارها ركيزة أساسية لتحقيق التنمية المستدامة، بالإضافة إلى دراسة تبحث في الذكاء الاصطناعي ومشكلة البطالة في مصر، حيث تناقش الواقع الراهن والتحديات المستقبلية أمام صانع القرار. وفي السياق القانوني أيضًا، يقدم العدد مثالًا عمليًا لتطبيق منازعة تنفيذ أمام القضاء الدستوري، موضحًا الأبعاد الإجرائية والدستورية لهذا النوع من المنازعات.

وفي مجال قضايا التنمية، يضم العدد دراسة مقارنة بعنوان: دور التكنولوجيا الرقمية في مواجهة تحديات المياه في مصر، والتي تعرض خبرات وتجارب دولية مختارة يمكن الاستفادة منها في مواجهة التحديات المائية. أما في المجال الثقافي والأدبي، فنجد دراسة سيميائية معمقة بعنوان: صناعة الذاكرة في اللوحة الإشهارية لرواية "الزيني بركات"، تكشف عن أبعاد ثقافية وبيولوجية في تشكيل الذاكرة الجمعية. كما يضم العدد بحث حول استخدام الذكاء الاصطناعي في تحديد هوية المتحدث في اللغة العربية، تقدم نظرة عامة على الأسس

النظرية لهذا المجال الواعد، إلى جانب دراسة في البلاغة السردية في الخطاب الاتصالي لتصاميم الملصقات الإرشادية، تبرز العلاقة بين الرسالة البصرية والبعد البلاغي.

وفي ميدان الدراسات الأثرية والأسطورية، يتناول بحث بعنوان: قصر التيه (اللابيرينثوس) ومعبد اللابيرنت بين الرؤية الأسطورية والحقيقة التاريخية، مقدّمًا قراءة مقارنة بين الأسطورة والواقع الأثري. ويختتم العدد بدراسة تحمل بعدًا جيو-ثقافي بعنوان: جيو-بولتيكية الوطن في سير غادة كرمي "البحث عن فاطمة" و"العودة"، حيث تتم قراءة التجربة الذاتية في سياقها الوطني والسياسي.

إننا في هذا العدد نؤكد التزام مجلة بحوث الشرق الأوسط بمواصلة رسالتها العلمية، وتقديم بحوث متجددة تسعى إلى فتح آفاق جديدة للباحثين والمتخصصين، بما يخدم المعرفة ويثري النقاش الأكاديمي في مختلف الحقول

والله ولي التوفيق،

رئيس التحرير

د. حاتم العبد

دراسات الصوتيات

www.mercj.journals.ekb.eg

استخدام الذكاء الاصطناعي

في تحديد هوية المتحدث في اللغة العربية:

نظرة عامة على الأسس النظرية

Using Artificial Intelligence to Identify Speakers in Arabic:

An Overview of the theoretical basics

ايمان محمد محمد قاسم

eman Mohamed Mohamed kassem

مدرس بقسم الصوتيات واللسانيات ـ كلية الآداب

جامعة الاسكندرية

Lecturer of phonetics and linguistics department

faculty of arts – Alexandria Universit

Eman.qasem@alexu.edu.eg

e.kassem10@gmail.com

www.mercj.journals.ekb.eg

الملخص

يُعد تحديد هوية المتحدث (Speaker Identification) فرعًا حيويًا من فروع التعرف على المتحدث (Speaker Recognition)، ويهدف إلى تحديد هوية متحدث غير معروف عن طريق مقارنة خصائص صوته بخصائص أصوات متحدثين معروفين مسبقًا. يُعرف هذا المجال أيضًا بالبحث عن المتحدث أو رصد المتحدث. يختلف تحديد هوية المتحدث عن التحقق من المتحدث (Speaker Verification) في أن الأول يقوم بمطابقة عينة صوتية مجهولة مع مجموعة من الهويات الصوتية المعروفة (مطابقة واحد–لكثير)، بينما يهدف الثاني إلى تأكيد هوية يدعيها المتحدث (مطابقة صوت متحدث–لمتحدث آخر).

تكتسب تقنيات تحديد هوية المتحدث أهمية متزايدة في العديد من التطبيقات العملية، مثل الرعاية الصحية حيث تُستخدم لتحديد المرضى وتخزين معلوماتهم في السجلات الطبية، وفي إنفاذ القانون لتحديد المشتبه بهم من خلال مقارنة التسجيلات الصوتية بملفات صوتية معروفة، وفي خدمة العملاء لتقديم تجارب مخصصة وضبط التفاعلات وفق احتياجات الأفراد. كما تمتد تطبيقاتها لتشمل البحث الصوتي، وتحويل الكلام إلى نص، والأوامر الصوتية للأجهزة المنزلية الذكية، والقياسات الحيوية الصوتية Biometric[1] للأغراض الأمنية وللأغراض التجارية.

يركز هذا البحث على استكشاف استخدام الذكاء الاصطناعي في تحديد هوية المتحدث في اللغة العربية، مع الأخذ في الاعتبار التعقيدات اللغوية والصوتية الفريدة لهذه اللغة. يتناول البحث الجوانب النظرية والأكاديمية التي تشمل مراحل عمل الأنظمة، وتقنيات استخلاص السمات الصوتية، ونماذج التعلم الآلي والتعلم العميق الحديثة. كما يستعرض التحديات الجوهرية التي تواجه هذا المجال في السياق العربي، ويقدم أمثلة عملية لأدوات وحلول معاصرة، ويُسلط الضوء على دراسات الحالة والتطبيقات الواقعية في القطاعات المختلفة.

Abstract

Speaker identification is a vital branch of speaker recognition that aims to identify an unknown speaker by comparing their voice characteristics with those of previously known speakers. This field is also known as speaker searching or speaker detection. Speaker identification differs from speaker verification in that the former matches an unknown voice sample with a set of known voice identities (one-to-many matching), while the latter aims to confirm a speaker's claimed identity (speaker-to-speaker matching).

Speaker identification technologies are gaining increasing importance in many practical applications, such as healthcare, where they are used to identify patients and store their information in medical records; in law enforcement, where they identify suspects by comparing voice recordings to known audio files; and in customer service, where they provide personalized experiences and tailor interactions with individuals. Their applications also extend to voice search, speech-to-text conversion, voice commands for smart home devices, and voice biometrics for security and commercial purposes

This research focuses on exploring the use of artificial intelligence in speaker identification in Arabic, taking into account the language's unique linguistic and phonetic complexities. The research addresses theoretical and academic aspects, including the stages of system operation, acoustic feature extraction techniques, and modern machine learning and deep learning models. It also reviews the fundamental challenges facing this field in the Arabic context, provides practical examples of contemporary tools and solutions, and highlights case studies and real-world applications in various sectors.

مقدمة

١.١. تعريف التعرف على المتحدث وأهميته

يُمثل التعرف على المتحدث مجالًا بحثيًا متطورًا يهدف إلى تحديد الأفراد بناءً على خصائص أصواتهم الفريدة. يندرج ضمن هذا المجال فرعان رئيسيان: تحديد هوية المتحدث (*Speaker Identification*) والتحقق من المتحدث (*Speaker Verification*). في سياق تحديد هوية المتحدث، تُحلل عينة صوتية مجهولة لتحديد هويتها ضمن مجموعة من المتحدثين المعروفين مسبقًا. يمكن تشبيه هذه العملية بالبحث عن تطابق في قاعدة بيانات واسعة من البصمات الصوتية، ولهذا يُشار إليها أحيانًا بالبحث عن المتحدث أو رصد المتحدث. على النقيض من ذلك، يركز التحقق من المتحدث على تأكيد هوية يدعيها الفرد؛ فإذا ادعى متحدث هوية معينة، تُستخدم عينة صوته لتأكيد هذا الادعاء عبر مطابقة فردية. وبالتالي، يقوم نظام التعرف على المتحدث بمطابقة واحد–لواحد للتحقق من المتحدث.

تتجلى أهمية هذه التقنيات في قدرتها على توفير مستوى عالٍ من الراحة والأمان في التفاعلات اليومية. في القطاع المصرفي، على سبيل المثال، يمكن لأنظمة التعرف على الصوت أن تتيح للعملاء الوصول الآمن إلى حساباتهم دون الحاجة إلى إدخال أرقام تعريف شخصية أو كلمات مرور تقليدية، مما يُعزز تجربة المستخدم ويُقلل من الاحتكاك. في مجال الرعاية الصحية، يمكن لتطبيقات الرعاية الصحية عن بُعد استخدام تحديد المتحدث لتحديد المرضى الذين يعانون من حالات مزمنة وتخزين معلوماتهم الصحية تلقائيًا، مما يُحسن دقة السجلات الطبية. أما في إنفاذ القانون، فتُعد هذه التقنيات أداة حاسمة لتحديد المشتبه بهم من خلال مقارنة التسجيلات الصوتية التي تُجمع من مسارح الجريمة بعينات صوتية معروفة، مما يُسرع من عمليات التحقيق.

إضافة إلى ذلك، تُسهم أنظمة تحديد هوية المتحدث في تقديم تجارب

شخصية في خدمة العملاء، حيث يمكن للوكلاء أو أنظمة الرد الصوتي التفاعلي (IVR)[٢] التعرف على العملاء من أصواتهم وتكييف التفاعلات لتلبية احتياجاتهم الفردية. كما تُستخدم هذه التقنيات على نطاق واسع في الأوامر الصوتية للأجهزة المنزلية الذكية، والبحث الصوتي، وتحويل الكلام إلى نص، مما يُمكن المستخدمين من التفاعل مع التكنولوجيا بطرق أكثر طبيعية وبدون استخدام اليدين. هذه القدرات المتعددة تُبرز الدور المحوري للذكاء الاصطناعي في تطوير أنظمة تحديد هوية المتحدث، وتُشير إلى إمكاناتها التحويلية في مختلف الصناعات والقطاعات.

١.٢. نطاق البحث وهيكله

يهدف هذا البحث إلى تقديم تحليل معمق لاستخدام الذكاء الاصطناعي في تحديد هوية المتحدث، مع تركيز خاص على تطبيقاته في اللغة العربية. يتناول البحث الجوانب النظرية والعملية لهذا المجال، بدءًا من المبادئ الأساسية التي تقوم عليها أنظمة التعرف على المتحدث وصولًا إلى أحدث التطورات في نماذج التعلم العميق والأدوات المتاحة.

يُقسم هيكل البحث إلى عدة أقسام رئيسية. يبدأ القسم التالي بتقديم الخلفية النظرية الشاملة، مع تفصيل مراحل عمل أنظمة التعرف على المتحدث، وتقنيات استخلاص الميزات الصوتية، والنماذج الإحصائية والشبكات العصبية العميقة المستخدمة في هذا المجال. ينتقل القسم الثالث إلى استعراض التحديات الفريدة التي تواجه تحديد هوية المتحدث في اللغة العربية، بما في ذلك التنوع الصوتي واللهجي، وندرة البيانات عالية الجودة، وتأثير العوامل البيئية وقصر المدة الصوتية، بالإضافة إلى التهديدات الأمنية المتمثلة في الهجمات العدائية.

يُخصص القسم الرابع لمناقشة الحلول المبتكرة والتطورات الحديثة التي تسعى للتغلب على هذه التحديات، مع التركيز على نماذج التعلم العميق المتقدمة

وتضمينات المتحدث، وكيفية معالجة التنوع اللهجي والضوضاء. أما القسم الخامس فيُقدم استعراضًا للأدوات والمكتبات مفتوحة المصدر والحلول التجارية وواجهات برمجة التطبيقات [٣](*APIs*) المتاحة حاليًا لدعم تحديد هوية المتحدث باللغة العربية، مع ذكر أمثلة عملية ودراسات حالة من القطاعات المختلفة.

٢. الخلفية النظرية للتعرف على المتحدث

تعتمد أنظمة التعرف على المتحدث على مجموعة من العمليات المعقدة لتحليل الإشارات الصوتية واستخلاص السمات الفريدة التي تميز كل متحدث. تتضمن هذه العمليات عادةً مرحلتين رئيسيتين: المعالجة الأمامية ونمذجة المتحدث والمطابقة.

٢.١. مراحل عمل أنظمة التعرف على المتحدث

تتطلب أنظمة التعرف على المتحدث، سواء لتحديد الهوية أو التحقق، معالجة دقيقة للإشارة الصوتية لضمان استخلاص المعلومات الأكثر تميزًا للمتحدث. تُقسم هذه العملية إلى مراحل متسلسلة تهدف إلى تحويل الإشارة الصوتية الخام إلى تمثيل رقمي يمكن للنماذج الحسابية التعامل معه بفعالية.

٢.١.١. المعالجة الأمامية ***Front-end Processing***

تُعد المعالجة الأمامية الخطوة الأولى والحاسمة في أي نظام للتعرف على المتحدث. تهدف هذه المرحلة إلى تجهيز الإشارة الصوتية الخام لاستخلاص الميزات، وتقليل تأثير العوامل غير المرغوب فيها مثل الضوضاء أو اختلاف القنوات. تتكون هذه المرحلة بشكل أساسي من ثلاث مكونات رئيسية:

- اكتشاف نشاط الصوت **:**(*Voice Activity Detection - VAD***)** الهدف الأساسي من اكتشاف نشاط الصوت هو تحديد الأجزاء التي تحتوي على كلام والأجزاء التي لا تحتوي على كلام في الإشارة الصوتية. يُسهم خوارزمية VAD القوية في تحسين أداء نظام التعرف على المتحدث من خلال ضمان أن يتم حساب هوية المتحدث فقط من مناطق الكلام الفعلية، وتجاهل فترات

الصمت أو الضوضاء غير الكلامية. تشمل التقنيات الشائعة لـ *VAD* الأساليب القائمة على الطاقة، والقائمة على النموذج، والأساليب الهجينة.

- استخلاص السمات الصوتية **:***(Feature Extraction)* بعد تحديد أجزاء الكلام، تُستخدم تقنيات استخلاص السمات الصوتية لتحويل إشارات الكلام إلى متجهات سمات صوتية. يجب أن تحمل هذه الميزات المستخرجة الخصائص الأساسية لإشارة الكلام التي تُمكن من التعرف على هوية المتحدث من خلال صوته. الهدف من استخلاص السمات الصوتية هو تقليل أبعاد متجهات السمات الصوتية عن طريق إزالة المعلومات غير المرغوب فيها وتأكيد المعلومات الخاصة بالمتحدث. تُعد معاملات التردد الطيفي ميل [٤](*MFCCs*) من أكثر تقنيات استخلاص السمات الصوتية شيوعًا واستخدامًا في أنظمة التعرف على المتحدث الحديثة.
- تعويض القناة *Channel Compensation):(* تُستخدم تقنيات تعويض القناة لتقليل تأثير عدم تطابق القناة والضوضاء البيئية. يمكن تطبيق تعويض القناة في مراحل مختلفة من عملية التعرف على المتحدث، مثل مجالات السمات والنماذج. تُستخدم العديد من التقنيات مثل طرح متوسط التردد الطيفي (*CMS*)، وتشويه السمات (*feature warping*)، وتطبيع متوسط تباين التردد الطيفي (*CMVN*)، ومعالجة الطيف النسبي [٥](*RASTA*) لتقليل تأثير عدم تطابق القناة أثناء مرحلة استخلاص الميزات. في مجال النموذج، تُستخدم تقنيات مثل تحليل العوامل المشتركة (*JFA*) والمتجهات الآي (i-*vectors*) لمكافحة عدم التطابق بين التسجيل والتحقق.

٢.١.٢. نمذجة المتحدث والمطابقة

بعد المعالجة الأمامية واستخلاص السمات الصوتية، تأتي مرحلة نمذجة المتحدث والمطابقة، حيث تُبنى نماذج تمثل خصائص صوت كل متحدث، ثم تُقارن

هذه النماذج لتحديد الهوية:

- إنشاء الهويات الصوتية *(Enrollment):* في هذه المرحلة، تُجمع عينات صوتية من المتحدثين المعروفين. تُحلل هذه العينات لإنشاء "هويات صوتية" أو "بصمات صوتية" فريدة لكل فرد. تُخزن هذه الهويات في قاعدة بيانات مرجعية.
- مطابقة الهويات *(Matching):* عند تقديم عينة صوتية لمتحدث مجهول، تُحلل هذه العينة وتُستخلص سماتها الصوتية بنفس الطريقة التي تُعالج بها عينات التسجيل. ثم تُقارن الميزات المستخلصة مع الهويات الصوتية المخزنة في قاعدة البيانات. تُقدم أنظمة التعرف على المتحدث بعد ذلك درجات تُشير إلى مدى احتمالية التشابه بين العينة المجهولة وكل هوية صوتية معروفة، أو تُقدم قائمة بالمتطابقات المحتملة.
- منطق القرار *(Decision Logic):* تُتخذ الخطوة النهائية في تحديد هوية المتحدث من خلال منطق القرار، الذي يُقارن متجهات السمات غير المعروفة بجميع النماذج في قاعدة البيانات ويختار النموذج الأفضل مطابقة. في حالة التحقق من المتحدث، يتضمن ذلك قبول Accept أو رفض Reject ادعاء الهوية بناءً على عتبة محددة.

٢.٢. تقنيات استخلاص السمات الصوتية

تُعد عملية استخلاص الميزات الصوتية جوهرية في أنظمة التعرف على المتحدث، حيث تُحول الإشارة الصوتية المعقدة إلى تمثيل رقمي Digitized مُبسط يُبرز الخصائص الفريدة لصوت كل فرد. من بين التقنيات الأكثر شيوعًا وفعالية في هذا المجال هي معاملات التردد الطيفي ميل (*MFCCs*).

٢.٢.١. معاملات التردد الطيفي ميل (MFCCs - Mel-Frequency Cepstral Coefficients)

تُعد معاملات التردد الطيفي ميل (*MFCCs*) من الميزات الصوتية الأكثر استخدامًا وشعبية في التعرف على الكلام وتحديد هوية المتحدث، وذلك لقدرتها على التقاط الخصائص الصوتية المهمة للكلام. تعتمد هذه المعاملات على محاكاة طريقة إدراك الأذن البشرية للترددات، والتي لا تتبع مقياسًا خطيًا، بل مقياسًا لوغاريتميًا يُعرف بمقياس "ميل". تتضمن عملية استخلاص *MFCCs* عدة خطوات رئيسية: (أنظر تمثيل الخطوات فى الشكل الآتى)

١. تقسيم الإشارة إلى إطارات :*(Frame Blocking)* تُقسم إشارة الكلام المستمرة إلى إطارات قصيرة ومتداخلة (مثل ٣٠ مللي ثانية مع تداخل). يُفترض أن تكون خصائص الكلام ثابتة نسبيًا خلال هذه الفترات القصيرة (٥-١٠٠ مللي ثانية)، ولكنها تتغير على فترات أطول.

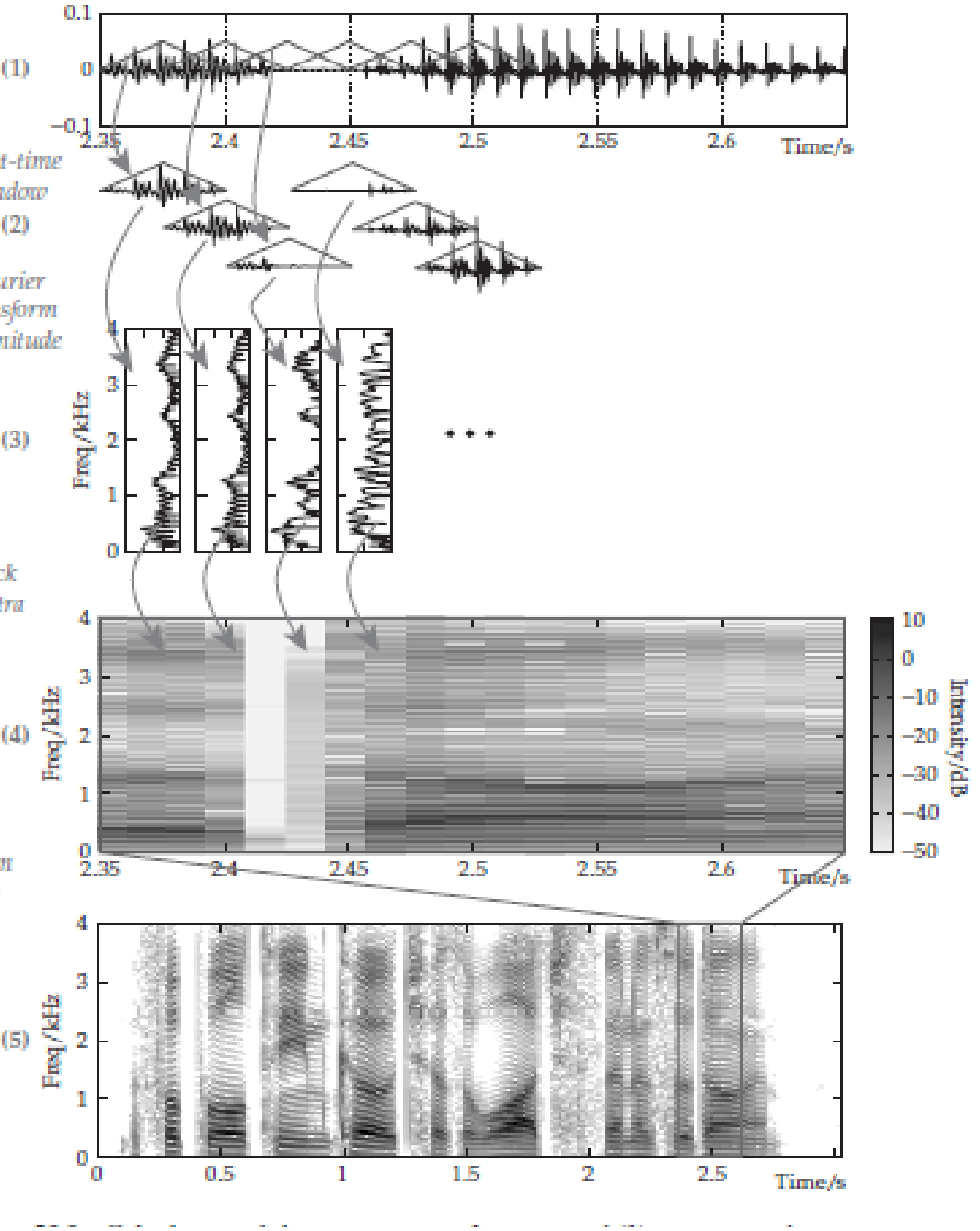
٢. تطبيق نافذة :*(Windowing)* تُطبق دالة نافذة (مثل نافذة *Hamming*) على كل إطار لتقليل التسرب الطيفي (*spectral leakage*) الناتج عن تقسيم الإشارة.

٣. تحويل فورييه السريع :*(FFT - Fast Fourier Transform)* يُستخدم تحويل فورييه السريع لتحويل الإشارة من المجال الزمني إلى المجال الترددي، مما يُوفر المحتوى الطيفي لكل إطار.

٤. تصفية ترددات ميل :*(Mel-Frequency Wrapping)* في هذه الخطوة الحاسمة، تُحول الترددات الفعلية (بالهرتز) إلى مقياس ميل الإدراكي. يتم ذلك عادةً باستخدام بنك من المرشحات المثلثية ذات الاستجابة الترددية النطاقية، حيث تُحدد تباعدها وعرض نطاقها بواسطة فاصل تردد ميل ثابت. يتميز مقياس ميل بكونه خطيًا تحت ١٠٠٠ هرتز ولوغاريتميًا فوق ١٠٠٠ هرتز.

٥. تحويل جيب التمام المنفصل :*(DCT - Discrete Cosine Transform)* في الخطوة الأخيرة، يُحول طيف ميل اللوغاريتمي مرة أخرى إلى مجال زمني

باستخدام تحويل جيب التمام المنفصل، وتُسمى النتيجة بمعاملات التردد الطيفي ميل .(*MFCCs*) عادةً، يُستبعد المكون الأول (c~٠) لأنه يُمثل القيمة المتوسطة للإشارة ويحمل القليل من المعلومات الخاصة بالمتحدث. تُقدم *MFCCs* تمثيلاً جيدًا للخصائص الطيفية المحلية للإشارة الصوتية. كما أنها أقل تأثرًا بالاختلافات في شكل الموجة الكلامية التي قد تحدث بسبب الحالة الفيزيائية للأحبال الصوتية للمتحدثين. بمجرد حسابها، تُستخدم هذه المجموعات من المعاملات (المتجهات الصوتية) لتمثيل خصائص صوت المتحدث والتعرف عليها، حيث تُحول كل نطق مدخل إلى سلسلة من هذه المتجهات الصوتية التي تُستخدم كمدخلات للشبكات العصبية أو النماذج الإحصائية لتحديد هوية المتحدث .



مراحل متسلسلة تهدف إلى تحويل الإشارة الصوتية الخام إلى تمثيل رقمي

يمكن للنماذج الحسابية التعامل معه بفعالية.

٣. التحديات الخاصة باللغة العربية في التعرف على المتحدث

تتسم اللغة العربية بثرائها اللغوي وتنوعها الصوتي، حيث مجموعة فريدة من التحديات لأنظمة التعرف على المتحدث. تتجاوز هذه التحديات مجرد ترجمة التقنيات المُطورة للغات أخرى، وتتطلب فهمًا عميقًا للخصائص الصوتية واللهجية، بالإضافة إلى معالجة قضايا البيانات والأمن.

٣.١. التنوع الصوتي واللهجي

تُعد اللغة العربية من أقدم وأهم اللغات عالميًا، يتحدثها أكثر من ٤٢٢ مليون شخص في ٢٥ دولة، مما يجعلها واحدة من أكثر خمس لغات تحدثًا على مستوى العالم. ومع ذلك، فإن هذا الانتشار الواسع يُصاحبه تنوع صوتي ولهجي كبير يُشكل تحديًا جوهريًا لأنظمة تحديد هوية المتحدث. تتميز اللغة العربية ببنية صوتية معقدة، مع مجموعة واسعة من الفونيمات واللهجات المتنوعة عبر مختلف البلدان، مما يُعيق دقة أنظمة تحديد المتحدث.

الفونيمات واللهجات

- ***التنوع الفونيمي*** *phonemic variation*: تمتلك اللغة العربية العديد من الفونيمات المميزة التي تُسهم في الأصوات الفريدة للكلمات. حتى الأخطاء البسيطة في النطق، مثل استبدال أو حذف الحروف المفخمة أو الحركات، يمكن أن تُغير المعنى الدلالي وتُشكل تحديات كبيرة لأنظمة الكشف الآلي.
- ***التنوع اللهجي*** *dialectal variation* : تختلف اللهجات الإقليمية العربية بشكل كبير من حيث النطق والقواعد والمفردات. على الرغم من وجود مستوى معقول من الفهم المتبادل بين اللهجات، إلا أن مدى فهم الفرد للهجات الأخرى يعتمد بشكل كبير على لهجته الخاصة ومدى تعرضه للثقافة والأدب العربي من خارج بلده. تُستخدم اللهجات العربية بشكل أساسي في الإعدادات

غير الرسمية للتواصل اليومي، بينما تُدرس العربية الفصحى المعاصرة Modern Standard Arabic(MSA) في المدارس وتُستخدم في السياقات الرسمية. هذا التباين اللغوي يُشكل عقبة كبيرة أمام أنظمة التعرف على الكلام الآلي Automatic Speaker Recognition(ASR) التي تُدرب على العربية الفصحى، حيث تفشل في كثير من الأحيان في التعامل بدقة مع الكلام اللهجي.

- ***الازدواجية اللغوية* :(*Diglossia*)** تُشير الازدواجية اللغوية في السياق العربي إلى التعايش بين العربية الفصحى الحديثة (*MSA*) واللهجات العامية المنطوقة داخل نفس المجتمع. تُستخدم العربية الفصحى في السياقات الرسمية والتعليمية والكتابية، بينما تُستخدم اللهجات العامية في التواصل اليومي غير الرسمي. هذا الوضع اللغوي الفريد يُقدم تحديات كبيرة لأنظمة التعرف على المتحدث، حيث يمكن للمتحدث أن يُغير بين الفصحى واللهجة في نفس المحادثة، مما يُؤثر على اتساق الخصائص الصوتية. تُظهر الأبحاث أن المسافات الصوتية والمعجمية بين العربية الفصحى واللهجات العامية يمكن أن تُعقد مهام الوعي الصوتي وتُؤثر على مهارات القراءة.

الخصائص الصوتية الدقيقة للغة العربية:

تُساهم الخصائص الصوتية الدقيقة للغة العربية في تعقيد مهمة تحديد هوية المتحدث:

- الأصوات الحلقية *(pharyngeal/Laryngeal Sounds):* تُنتج الأصوات الحلقية (مثل /h/ ,/ʕ/ ,/ħ/ ,/ʁ/ ,/χ/) كالحاء و الغين والخاء و الهاء و العين – في المنطقة الخلفية من الجهاز الصوتي (اللهاة، البلعوم، الحنجرة). تُعد هذه الأصوات نادرة عالميًا وتتطلب وضعًا معقدًا للسان والقناة الصوتية، مما يُشكل تحديًا خاصًا لأنظمة تحليل الكلام. تُظهر الأبحاث أن هذه

الأصوات لها خصائص صوتية مميزة، مثل مركز الثقل والانحراف المعياري لضوضاء الاحتكاك، والتي تختلف بين الأصوات اللهاوية والبلعومية.

- الحروف المفخمة *(Emphatic Consonants):* تُعد الحروف المفخمة (مثل /ð /, /ḍ /, /ṣ /, /ṭ/) كالظاء والضاد والصاد والطاء– سمة مميزة للغات السامية مثل العربية. تُنتج هذه الحروف بتراجع جذر اللسان و/أو تضييق البلعوم، مما يُؤثر على الحروف المتحركة المجاورة لها، حيث تُصبح أكثر تراجعًا في النطق. تُظهر التحليلات الصوتية أن الحروف المتحركة في السياق المفخم تُظهر ترددات F1 و F3 أعلى و F2 أقل. هذا التأثير يمكن أن يمتد ليشمل جميع المقاطع داخل الكلمة الواحدة (*emphasis spread*)، مما يُضيف تعقيدًا إلى عملية استخلاص الميزات الصوتية الخاصة بالمتحدث.
- طول الحركات *(Vowel Length):* تُفرق اللغة العربية بين الحركات القصيرة والطويلة، وهذا التمييز له دلالة صوتية ومعنوية. تُشير الدراسات الصوتية إلى أن الحركات الطويلة تكون تقريبًا ضعف طول نظيراتها القصيرة، وتختلف أيضًا طيفيًا، حيث تكون الحركات القصيرة أكثر تمركزًا بشكل ملحوظ. هذا التباين في المدة والنوعية بين الحركات يُعد خاصية صوتية مهمة يمكن أن تُسهم في تحديد هوية المتحدث، لكنها تتطلب نماذج قادرة على التقاط هذه الفروق الدقيقة.

السمات الفيزيائية المميزة لهوية المتحدث

- يشير التردد الأساسي (F0) إلى أحد أهم السمات الصوتية الرئيسية في الكلام، وهو يمثل معدل اهتزاز الأحبال الصوتية أثناء الأصوات المجهورة voiced sounds، مما يؤثر على درجة الصوت كما يُدركها السامع. يُقاس F0 عادة بوحدة الهرتز (Hz)، ويختلف بشكل ملحوظ بين الأفراد،

ويرتبط ارتباطًا وثيقًا بخصائص مثل العمر والجنس، حيث تكون أصوات الذكور عادةً ذات ترددات أقل مقارنة بأصوات الإناث. وبالإضافة إلى دوره في تحديد طبقة الصوت pitch، يُعد F0 من الخصائص الصوتية المهمة في العديد من التطبيقات مثل التعرف على الكلام وتحديد هوية المتحدث، مما يجعله عنصرًا أساسيًا في التحليل الصوتي في مجالي اللسانيات والتكنولوجيا. ويُعد تحديد هوية المتحدث جانبًا أساسيًا من تقنيات التعرف على المتحدث. وتتضمن هذه العملية استخراج عدد من السمات الصوتية من الكلام، بما في ذلك F0، ولتحسين دقة أنظمة التعرف.تم وصف طريقة لتحديد هوية المتحدث الجنائية تعتمد على التحليل الإحصائي لتردد الأساس (f0). وقد تم اختبار موثوقية هذه الطريقة باستخدام عينة كبيرة من المواد الكلامية الحقيقية المتمثلة في مكالمات هاتفية.

- تُعد المكونات الترددية Formant Frequencies، F1, F2, F3 &F4وهي المكونات الترددية الرنينية داخل الجهاز الصوتي، من العناصر الحيوية في إدراك الكلام، حيث يعكس كل تردد الرنين الحادث في كل تجويف للجهاز الصوتي، التجويف البلعومي والتجويف الفموي والتجويف الأنفي وحجم الجمجمة؛ حيث تميز بين الأصوات المتحركة (الصوائت) وتعزز من وضوح التواصل، خصوصًا في البيئات الصاخبة. تلعب هذه الترددات دورًا مهمًا في تحديد هوية المتحدث والتحليل الجنائي من خلال تحليل الأنماط الطيفية للتحقق من الهوية، وقد ساهم التقدم التكنولوجي، مثل الشبكات العصبية العميقة، في تحسين دقة هذا التحقق حتى في الظروف الصعبة.
- ومع ذلك، لا تزال هناك تحديات قائمة ناتجة عن الفروق الفردية التي تتأثر بعوامل مثل ضعف السمع، واللهجات، والجنس، والعمر، والتي يمكن أن تؤثر

على إدراك الفورمنت وعمومية النتائج. وتسعى الأبحاث المستمرة إلى تطوير أساليب التحليل من خلال تقنيات التعلم الآلي، إلى جانب معالجة القضايا الأخلاقية، وخصوصًا فيما يتعلق بمواجهة تقنيات التلاعب الصوتي مثل "التزييف العميق" (Deepfakes) التي تهدد مصداقية الأدلة الصوتية في المجالات السريرية والجنائية.

- الجيتر (Jitter) والشيمر (Shimmer) هما من المعايير الصوتية المهمة في تحليل الصوت، خاصة في مجال تحديد هوية المتحدث. يشير الجيتر إلى التغير في مدة الذبذبات الصوتية بين كل دورة وأخرى، مما يعكس استقرار وجودة الصوت. أما الشيمر فيقيس التغير في شدة الصوت بين الدورات الصوتية، ما يدل على تقلبات السعة الصوتية. وتُستخدم هاتان الميزتان في التحقق من هوية المتحدث، حيث تساعدان في تمييز الأصوات الفردية. وقد أثبتت التجارب نجاح استخدام الجيتر والشيمر في التحقق من المتحدث، خاصة عند دمجهما مع السمات البروزودية والتحليل الطيفي باستخدام تقنيات دمج ومعالجة مختلفة، مما يُحسن من دقة التحقق. وتبرز أهمية هذه الخصائص بشكل خاص في التطبيقات الأمنية والتحليل الجنائي، حيث تكون دقة التحديد ضرورية.
- نسبة المكونات التوافقية إلى الضوضاء (Harmonics-to-Noise Ratio - HNR) هي مقياس مهم في مجال تحديد هوية المتحدث، حيث تُحدد نسبة المكونات التوافقية (الدورية) في الإشارة الصوتية مقارنة بالمكونات غير الدورية (الضوضاء). يلعب هذا القياس دورًا حاسمًا في التمييز بين المتحدثين استنادًا إلى خصائصهم الصوتية الفريدة، وهو ما يُعد ذا أهمية في سياقات متعددة مثل التحليل الجنائي، وتفريغ النصوص القانونية، وأنظمة التعرف الآلي على الكلام. وقد حظي HNR

باهتمام متزايد في التطبيقات العلمية والعملية نظرًا لقدرته على تحسين دقة تحديد هوية المتحدث، رغم التحديات التي يواجهها، مثل التباين في عينات الكلام والعوامل البيئية كضوضاء الخلفية.

- السمات الفوقطعية **suprasegmental features** أو السمات البروذودية **prosodic features** كالتنغيم والإيقاع والنبر *(Intonation, Rhythm, and Stress):* تُشكل هذه الخصائص الفوقطعية جزءًا لا يتجزأ من هوية المتحدث، ويمكن أن تُسهم في تمييز اللهجات وتحديد الأفراد. تُظهر الأبحاث أن السمات البروذودية مثل التنغيم والإيقاع يمكن أن تُحسن بشكل كبير من تحديد اللهجة العربية. على سبيل المثال، يتميز المتحدثون السوريون باستخدام "تنغيم غنائي" في نطقهم، بينما تُظهر اللهجات العربية الشرقية اختلافات ملحوظة في التنغيم. تُشير الدراسات أيضًا إلى أن النبر (*stress*) في اللغة العربية يتبع أنماطًا معينة بناءً على بنية المقاطع (خفيفة، ثقيلة، فائقة الثقل)، وأن مدة المقطع وكثافته وتردده الأساسي (*F0*) تُعد مؤشرات صوتية للنبر. تُشكل هذه التباينات تحديًا وفرصة في آن واحد لتطوير أنظمة أكثر دقة لتحديد هوية المتحدث باللغة العربية.

٣.٢. ندرة البيانات عالية الجودة

تُعد ندرة البيانات عالية الجودة والمُعنونة بشكل كافٍ أحد أبرز التحديات التي تواجه تطوير أنظمة التعرف على المتحدث باللغة العربية. على الرغم من أن اللغة العربية هي رابع أكثر اللغات تحدثًا في العالم. هذا النقص الحاد في الموارد الرقمية يُؤثر سلبًا على قدرة الباحثين والمطورين على تدريب نماذج الذكاء الاصطناعي القوية التي تتطلب آلاف الساعات من البيانات المُعنونة لتعلم الأنماط اللغوية والصوتية المعقدة.

تُعاني معظم نماذج اللغة العربية الكبيرة الموجودة حاليًا من اعتمادها الكبير

على البيانات الإنجليزية، مما يُقلل من فعاليتها في التعامل مع التفكير العربي المعقد واللهجات المختلفة. يُؤدي هذا الاعتماد إلى تراجع قدرات الأنظمة في الحفاظ على التقاليد اللغوية التي تحفظ سمات الأسلوب العربي، ويُعيق تطوير حلول مُخصصة للسياق الثقافي واللغوي العربي.

علاوة على ذلك، يُشكل إنشاء قواميس صوتية عربية تحديًا بحد ذاته. على عكس اللغات التي تحتوي على عدد كبير من استثناءات النطق (مثل الإنجليزية)، يتبع نطق النص العربي قواعد محددة عندما يكون النص مُشكلًا بالكامل. ومع ذلك، فإن عدم وجود تشكيل في معظم النصوص العربية يُصعب التمييز بين النطق والمعاني المختلفة لنفس الكلمة. هذا يتطلب جهودًا كبيرة لإنشاء قواميس نطق دقيقة ومُفصلة للهجات العربية المتنوعة، والتي تُعد مكونًا حيويًا لأنظمة التعرف على الكلام.

يُعيق نقص البيانات المُعنونة أيضًا تطوير نماذج التعلم العميق المُشرفة، التي تُعد أساس الأنظمة الحديثة عالية الأداء. هذا النقص يُجبر الباحثين على استكشاف أساليب مثل التعلم شبه المُشرف والتعلم باللقطات القليلة، والتي تُمكن من الاستفادة من البيانات غير المُعنونة أو المحدودة، ولكنها لا تزال تتطلب جهودًا بحثية مكثفة لتحقيق الأداء الأمثل في بيئات العالم الحقيقي.

٣.٣. التحديات البيئية وقصر المدة الصوتية

تُشكل العوامل البيئية وقصر مدة النطق تحديات كبيرة لأنظمة تحديد هوية المتحدث، خاصة في بيئات العالم الحقيقي حيث تكون الظروف غير مثالية.

- الضوضاء البيئية وظروف التسجيل: تُعد الضوضاء الخلفية (*babble noise*)، مثل أحاديث متعددة أو ضوضاء حركة المرور في المدن (traffic noise)، من أبرز العوائق التي تُقلل من دقة أنظمة التعرف على المتحدث.

تُدخل هذه الضوضاء تباينًا في الإشارة الصوتية وتُصعب على النظام استخلاص الميزات الصوتية النقية الخاصة بالمتحدث. بالإضافة إلى ذلك، تُؤثر ظروف التسجيل المختلفة، مثل جودة الميكروفون ونوع قناة الإرسال (خط أرضي، هاتف محمول، VoIP[٦])، على خصائص الصوت وتُقلل من أداء النظام. تُشير الدراسات إلى أن أداء أنظمة تحديد المتحدث يتدهور بشكل كبير في وجود الضوضاء، وأن الضوضاء الكلامية تُعد من أصعب التداخلات بسبب خصائصها الشبيهة بالكلام.

- قصر المدة الصوتية :*(Short Utterances)* تُعد القدرة على تحديد المتحدث بدقة من مقاطع صوتية قصيرة تحديًا بحثيًا مستمرًا. في العديد من التطبيقات العملية، قد لا تتوفر تسجيلات طويلة للمتحدث، مما يُصعب على النظام استخلاص معلومات كافية لإنشاء بصمة صوتية موثوقة. على سبيل المثال، تُظهر بعض الدراسات أن تحديد المتحدث باللغة العربية باستخدام نطق قصير يُقدم رؤى جيدة، ولكنه لا يزال يتطلب مزيدًا من التحسينات. تُشير دراسات أخرى إلى أن أنظمة التعرف على المتحدث القائمة على التعلم العميق تُقدم نتائج واعدة حتى مع كميات قليلة من بيانات التدريب، ولكن التحديات الكبيرة لا تزال قائمة، بما في ذلك البيانات المحدودة، وتُعد هذه المشكلة حاسمة في سيناريوهات مثل مراكز الاتصال حيث تكون نسبة الإشارة إلى الضوضاء منخفضة وقد تكون مدة النطق قصيرة.

تتطلب معالجة هذه التحديات دمج تقنيات متقدمة في المعالجة الأمامية، مثل ترشيح الضوضاء *(noise filtering)* وكشف نشاط الكلام Voice Activity Detection (*VAD*) الفعال، بالإضافة إلى تطوير نماذج تعلم عميق قادرة على استخلاص سمات متينة من الإشارات الصوتية المتدهورة وتكييفها مع البيانات المحدودة.

٤. الحلول المبتكرة والتطورات الحديثة

شهد مجال تحديد هوية المتحدث بالذكاء الاصطناعي تطورات سريعة،

مدفوعة بالتقدم في التعلم العميق والحاجة المتزايدة لأنظمة أكثر دقة ومتانة. تهدف الحلول المبتكرة إلى معالجة التحديات الجوهرية، لا سيما تلك المتعلقة باللغات ذات الخصائص الصوتية المعقدة مثل اللغة العربية.

٤.١. التعلم العميق المتقدم ودلائل المتحدث

تُشكل نماذج التعلم العميق المتقدمة ودلائل المتحدث حجر الزاوية في أنظمة تحديد الهوية الحديثة، حيث تُمكن من استخلاص تمثيلات صوتية غنية بالمعلومات الخاصة بالمتحدث. وسوف نعرض بعض الأمثلة باختصار:

- يُعد نموذج ***WavLM***[٧] من النماذج المُدربة مسبقًا الرائدة في مجال التعلم شبه المُشرف، وقد أظهرت قدرات استثنائية في فهم تمثيلات الكلام. تُدرب هذه النماذج على كميات هائلة من البيانات الصوتية غير المُعنونة، مما يُمكنها من تعلم ميزات صوتية قوية يمكن بعد ذلك ضبطها بدقة لمهام محددة مثل تحديد المتحدث أو التعرف على الكلام. تُظهر ***WavLM***، على سبيل المثال، قدرة على التعامل مع الكلام المختلط (*mixture speech*) من خلال أداء إزالة الضوضاء والتنبؤ بالكلام المُقنّع ***Masked speech***.
- دمج الأنماط المتعددة (*Multimodal Integration*) اتجاهًا بحثيًا واعدًا في الذكاء الاصطناعي، حيث تسعى النماذج إلى معالجة وفهم المعلومات من مصادر متعددة مثل الصوت والنص والصورة والفيديو. في سياق التعرف على المتحدث، يمكن أن يُعزز دمج الأنماط المتعددة من دقة وموثوقية الأنظمة، خاصة في البيئات المعقدة أو عند وجود معلومات غير كاملة من نمط واحد.

تُعد هذه النماذج حيوية في اللغة العربية، بشكل خاص نظرًا لتعقيداتها اللغوية والثقافية. على سبيل المثال، تُشير الأبحاث الحديثة إلى تطوير معايير لتقييم الاستدلال متعدد الأنماط خطوة بخطوة باللغة العربية، مما يُسلط الضوء على أوجه القصور في النماذج الحالية فيما يتعلق بالتماسك اللغوي والتأصيل الثقافي. تُظهر

هذه الدراسات أن دمج الميزات الصوتية والنصية يمكن أن يُعزز الأداء في مهام معينة مثل تحديد العلاقات الخطابية الضمنية في الكلام العربي المصري. يُمكن أن يُسهم هذا الدمج في بناء أنظمة أكثر شمولية وقادرة على التعامل مع التحديات الفريدة للغة العربية من منظور متعدد الأبعاد.

٤.٢. معالجة تحديات اللغة العربية المحددة

تُركز التطورات الحديثة في التعرف على المتحدث باللغة العربية على تطوير حلول مُخصصة لمعالجة التحديات الجوهرية التي تُقدمها هذه اللغة، مثل التنوع اللهجي والضوضاء والبيانات المحدودة.

- تخفيف تأثير التنوع اللهجي: يُعد التنوع اللهجي في اللغة العربية أحد أكبر التحديات أمام أنظمة التعرف على المتحدث والكلام. لمعالجة هذه المشكلة، تُستخدم عدة استراتيجيات:
- التعامل مع الضوضاء والمدخلات القصيرة: تُعد الضوضاء في التسجيلات الصوتية وقصر مدة النطق من العوامل التي تُعيق أداء أنظمة تحديد هوية المتحدث لمعالجة هذه المشكلات

تُظهر هذه الحلول مجتمعة التزامًا بحثيًا قويًا بتطوير أنظمة تحديد هوية متحدث عربية أكثر متانة ودقة، قادرة على التعامل مع التعقيدات اللغوية والبيئية الفريدة للغة العربية.

الخلاصة: يوفر الذكاء الاصطناعي إمكانات هائلة لتحديد هوية المتحدث بالعربية، لكن نجاحه مرهون بمعالجة التحديات الصوتية و اللغوية والبيئية الفريدة. إن التطورات في نماذج التعلم العميق والأدوات المتخصصة تُمهد الطريق لأنظمة أكثر دقة، خاصة مع تعزيز التعاون بين الباحثيين والمطوريين في العالم العربي.

المراجع

1. Al-Ani, S. H. (1969). Acoustic correlates of Arabic emphatic consonants. Phonetica, 20(1), 1–10.
2. Al-Ani, S. H. (1970). Arabic phonology: An acoustical and physiological investigation. The Hague: Mouton.
3. Al-Ani, S. H. (1978). Pharyngeal and emphatic sounds in Arabic. Linguistics, 16(206), 67–76.
4. Al-Tamimi, J., & Khattab, G. (2015). Acoustic and phonological features of pharyngeals and emphatics in Lebanese Arabic. Journal of Phonetics, 50, 52–64. https://doi.org/10.1016/j.wocn.2015.02.003
5. Alansary, I. (2023). Artificial intelligence model to detect and classify Arabic dialects. Journal of Software Engineering and Applications, 16(7), 263–276. https://www.scirp.org/journal/paperinformation?paperid=126690
6. arrús, M., & Hernando, J. (2008). Using jitter and shimmer in speaker verification. IET Signal Processing, Special Issue on Biometric Recognition. https://scispace.com/pdf/using-jitter-and-shimmer-in-speaker-verification-x6joy6413s.pdf
7. Bai, Y., Zhang, J., Li, J., & Zhang, B. (2021). Speaker recognition based on deep learning: An overview. Neural Networks, 140, 65–82.
8. Becker, T., Kießling, A., & Groth, K. (2008). Forensic speaker recognition using formant features and Gaussian mixture models. In 2008 Speaker and Language Recognition Workshop (pp. 1–4). IEEE.
9. Boss, D. (1996). The problem of F0 and real-life speaker identification: A case study. Forensic Linguistics, 3(1), 155–169. https://www.researchgate.net/publication/276914683_The_problem_of_F0_and_real-life_speaker_identification_a_case_study
10. Bush, C. N. (1967). Some Acoustic parameters of speech and their relationships to the perception of dialect differences. TESOL Quarterly, 1, 20–30.

11. Campbell, J. P. (1997). Speaker recognition: A tutorial. Proceedings of the IEEE, 85(9), 1437–1462.
12. Cedergren, H., & Perreault, H. (1995). On the Analysis of Syllable-timing in everyday speech. Proceedings of the 13th International Congress of Phonetic Sciences, Stockholm, 4, 232–5.
13. Ellis, D. P. W. (2010). An Introduction to Signal Processing for Speech. In The Handbook of Phonetic Sciences: Second Edition, p. 756.
14. Farahani, F., Georgiou, P. G., & Narayanan, S. S. (2004). Speaker identification using supra-segmental pitch pattern dynamics. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004) (Vol. 1, pp. 89–92). IEEE. https://www.researchgate.net/publication/4087515_Speaker_identification_using_supra-segmental_pitch_pattern_dynamics
15. Ferguson, J. D. (1980). Hidden Markov Analysis: An Introduction. In J. D. Ferguson (Ed.), Hidden Markov Models for Speech. Princeton, NJ: IDA-CRD.
16. Fletcher, J. (2010). The Prosody of Speech: Timing and Rhythm. In The Handbook of Phonetic Sciences: Second Edition, p. 523.
17. Gelfer, M. P., & Fendel, D. M. (1995). Comparisons of jitter, shimmer, and signal-to-noise ratio from directly digitized versus taped voice samples. Journal of Voice, 9(4), 378–382. https://doi.org/10.1016/S0892-1997(05)80201-5
18. Ghazeli, S. (1977). Back consonants and backing coarticulation in Arabic. University of Texas at Austin. (Doctoral dissertation)
19. Gold, E., & French, P. (2011). International practices in forensic speaker comparison. International Journal of Speech, Language and the Law, 18(2), 293–307. https://doi.org/10.1558/ijsll.v18i2.293
20. Hamsa, S., & others. (2023). Speaker identification from emotional and noisy

speech using learned voice segregation and speech VGG. Expert Systems with Applications, 224. https://www.sciencedirect.com/science/article/abs/pii/S095741742300372X

21. Hansen, J. H. L., & Hasan, T. (2015). Speaker recognition by machines and humans. In J. Benesty, M. M. Sondhi, & Y. Huang (Eds.), Springer Handbook of Speech Processing (pp. 743–762). Springer.
22. Hayward, K. (2000). Experimental phonetics: An introduction (2nd ed.). London: Longman.
23. Hollien, H. (2001). Forensic voice identification. Academic Press.
24. Jessen, M. (2007). Speaker Classification in Forensic Phonetics and Acoustics. In C. Müller (Ed.), Speaker Classification I, LNAI 4343, pp. 180–204.
25. Kamruzzaman, S. M., Karim, A. N. M. R., Islam, M. S., & Haque, M. E. (2010). Speaker Identification using MFCC-Domain Support Vector ... - arXiv. https://arxiv.org/pdf/1009.4972
26. Künzel, H. J. (2000). Effects of voice disguise on forensic speaker identification. In C. Müller (Ed.), Speaker Classification I (pp. 185–196). Springer.
27. Künzel, H. J. (2001). Automatic speaker recognition by analyzing formant contours of the natural speech signal. Forensic Linguistics, 8(1), 45–54.
28. Künzel, H. J. (2001). Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies. Forensic Linguistics, 8(1).
29. Labutin, P., Koval, S., & Raev, A. (2007). Speaker identification based on the statistical analysis of F0. In Proceedings of IAFPA 2007, The College of St Mark & St John, Plymouth, UK. Speech Technology Center, Saint Petersburg, Russia. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=bd592bdcb

1ecdc6dbbfa21959fc54ff2d2433758

30. Ladefoged, P. (2003). Phonetic data analysis: An introduction to fieldwork and instrumental techniques. Oxford: Blackwell.
31. McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (Ed.), Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form (pp. 191–233). Cambridge University Press.
32. Morgan, N., & Boerlard, H. A. (1995). Neural Networks for Statistical Recognition of continuous Speech. Proceedings of the IEEE, 83, 742–72.
33. Morrison, G. S., & Neumann, C. (2013). Likelihood-ratio calibration in forensic voice comparison. Science & Justice, 53(3), 238–241.
34. Nagrani, A., Chung, J. S., & Zisserman, A. (2017). VoxCeleb: A large-scale speaker identification dataset. In Proceedings of Interspeech 2017 (pp. 2616–2620).
35. Nolan, F. (1983). The phonetic bases of speaker recognition. Cambridge University Press.
36. Obrecht, D. H. (1968). Effects of the second formant on the perception of stop consonants. The Hague: Mouton.
37. Rose, P. (2003). Forensic Speaker Identification. London: Taylor and Francis
38. Rose, P. (2003). Forensic speaker identification. Taylor & Francis.
39. Salhab, M., Elghitany, M., Sait, S., Ullah, S. S., Abusheikh, M., & Abusheikh, H. (2025). Advancing Arabic speech recognition through large-scale weakly supervised learning. arXiv.
https://arxiv.org/abs/2504.12254
40. Speech Recognition Accuracy Test 2024 – Arabic Edition - SESTEK. https://www.sestek.com/speech-recognition-accuracy-test-2024-arabic-edition-blog
41. Stieve, R., & Simon, K. (2010). Automatic speech recognition. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), The handbook of phonetic

sciences (2nd ed., p. 804). Wiley-Blackwell.

42. Stolcke, A., Shriberg, E., Hakkani-Tür, E., & Tür, G. (1999). Modelling the prosody of hidden events for improved word recognition. In Proceedings of Eurospeech '99 (pp. 311–314). Budapest, Hungary.
43. Tiwari, V. (2010). MFCC and its applications in speaker recognition. International Journal on Emerging Technologies, 1(1), 19–22. https://www.researchgate.net/publication/265809116_MFCC_and_its_applications_in_speaker_recognition
44. Variani, E., Lei, X., McDermott, E., Moreno, I. L., & Gonzalez-Dominguez, J. (2014). Deep neural networks for small footprint text-dependent speaker verification. In 2014 IEEE ICASSP (pp. 4052–4056). IEEE.
45. Watson, J. C. E. (2002). The phonology and morphology of Arabic. Oxford University Press.
46. Weggman, S., MacAllaster, D., Orlo, J., & Peskin, B. (1996). Speaker Normalisation on Conversational Telephone Speech. In Proceedings of IEEE ICASSP'96, 1, 339–41.
47. Zaidan, O. F., & Callison-Burch, C. (2014). Arabic dialect identification. Computational Linguistics, 40(1), 171–202. https://direct.mit.edu/coli/article/40/1/171/1458/Arabic-Dialect-Identification

Bio- حيوي Metric , قياس : Biometric
تستخدم للاشارة إلى القياسات أو الخصائص الجسدية أو السلوكية الفريدة التي تستخدم للتعرف على الأشخاص مثل : بصمة الإصبع، ملامح الوجه، الصوت، قزحية العين.

[2] IVR: Interactive Voice Response
هو الاستجابة الصوتية التفاعلية كما في البنوك الرد الآلي التفاعلي للعميل، للحصول على خدمات مثل: الاستعلام عن الرصيد – تحويل الأموال- دفع الفواتير- معرفة أخر الحركات. بدون التعامل مع الموظف

[3] APIs: Application programming interfaces

وجهات برمجة التطبيقات: هي عبارة عن مجموعة من الأدوات والتعريفات التي تمكن التطبيقات والبرامج من التواصل والتكامل مع بعضها لبعض بسهولة.

[4] (MFCCs) Mel – Frequency Cepstural Coefficients.

حساب لوغاريتمي يحاكي الطريقة التي تستجيب بها الأذن البشرية لتردد الصوت، لأن الاحساس البشري بالصوت يكون أقرب إلى اللوغاريتمي من الحساب الخطي.

[5] RASTA:

تقنية تستخدم في معالجة الاشارات الصوتية؛ لتحسين استقرار وموثوقية سمات الصوت للمتكلم

[6] VOIP: voice over internet protocol

[7] WavLM: Wave-based Language Model

هو نموذج متقدم للتعلم الذاتي مخصص لمعالجة الكلام والصوتيات، وهو من تطوير مايكروسوفت

Middle East Research Journal

Refereed Scientific Journal
(Accredited) Monthly

Vol.115
September 2025

Issued by
Middle East
Research Center

Fifty year
Founded in 1974

Issn: 2536 - 9504
Online Issn: 2735 - 5233