

مجلة جامعة البعث

سلسلة العلوم الهندسية الميكانيكية
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 46 . العدد 1

1445 هـ - 2024 م

الأستاذ الدكتور عبد الباسط الخطيب
رئيس جامعة البعث
المدير المسؤول عن المجلة

رئيس هيئة التحرير	أ. د. محمود حديد
رئيس التحرير	أ. د. درغام سلوم

مديرة مكتب مجلة جامعة البعث
م. هلا معروف

عضو هيئة التحرير	د. محمد هلال
عضو هيئة التحرير	د. فهد شريباتي
عضو هيئة التحرير	د. معن سلامة
عضو هيئة التحرير	د. جمال العلي
عضو هيئة التحرير	د. عباد كاسوحة
عضو هيئة التحرير	د. محمود عامر
عضو هيئة التحرير	د. أحمد الحسن
عضو هيئة التحرير	د. سونيا عطية
عضو هيئة التحرير	د. ريم ديب
عضو هيئة التحرير	د. حسن مشرفي
عضو هيئة التحرير	د. هيثم حسن
عضو هيئة التحرير	د. نزار عبشي

تهدف المجلة إلى نشر البحوث العلمية الأصيلة، ويمكن للراغبين في طلبها

الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة البعث

سورية . حمص . جامعة البعث . الإدارة المركزية . ص . ب (77)

. هاتف / فاكس : 963 31 2138071 ++

. موقع الإنترنت : www.albaath-univ.edu.sy

. البريد الإلكتروني : [magazine@ albaath-univ.edu.sy](mailto:magazine@albaath-univ.edu.sy)

ISSN: 1022-467X

شروط النشر في مجلة جامعة البعث

الأوراق المطلوبة:

- 2 نسخة ورقية من البحث بدون اسم الباحث / الكلية / الجامعة) + CD / word من البحث منسق حسب شروط المجلة.
 - طابع بحث علمي + طابع نقابة معلمين.
 - إذا كان الباحث طالب دراسات عليا:
يجب إرفاق قرار تسجيل الدكتوراه / ماجستير + كتاب من الدكتور المشرف بموافقة على النشر في المجلة.
 - إذا كان الباحث عضو هيئة تدريسية:
يجب إرفاق قرار المجلس المختص بإنجاز البحث أو قرار قسم بالموافقة على اعتماده حسب الحال.
 - إذا كان الباحث عضو هيئة تدريسية من خارج جامعة البعث :
يجب إحضار كتاب من عمادة كليته تثبت أنه عضو بالهيئة التدريسية و على رأس عمله حتى تاريخه.
 - إذا كان الباحث عضواً في الهيئة الفنية :
يجب إرفاق كتاب يحدد فيه مكان و زمان إجراء البحث ، وما يثبت صفته وأنه على رأس عمله.
 - يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (العلوم الطبية والهندسية والأساسية والتطبيقية):
عنوان البحث .. ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).
- 1- مقدمة
 - 2- هدف البحث
 - 3- مواد وطرق البحث
 - 4- النتائج ومناقشتها .
 - 5- الاستنتاجات والتوصيات .
 - 6- المراجع.

- يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (الآداب - الاقتصاد - التربية - الحقوق - السياحة - التربية الموسيقية وجميع العلوم الإنسانية):
- عنوان البحث .. ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).
- 1. مقدمة.
- 2. مشكلة البحث وأهميته والجديد فيه.
- 3. أهداف البحث و أسئلته.
- 4. فرضيات البحث و حدوده.
- 5. مصطلحات البحث و تعريفاته الإجرائية.
- 6. الإطار النظري و الدراسات السابقة.
- 7. منهج البحث و إجراءاته.
- 8. عرض البحث و المناقشة والتحليل
- 9. نتائج البحث.
- 10. مقترحات البحث إن وجدت.
- 11. قائمة المصادر والمراجع.
- 7- يجب اعتماد الإعدادات الآتية أثناء طباعة البحث على الكمبيوتر:
 - أ- قياس الورق 25×17.5 B5.
 - ب- هوامش الصفحة: أعلى 2.54- أسفل 2.54 - يمين 2.5- يسار 2.5 سم
 - ت- رأس الصفحة 1.6 / تذييل الصفحة 1.8
 - ث- نوع الخط وقياسه: العنوان . Monotype Koufi قياس 20
- . كتابة النص Simplified Arabic قياس 13 عادي . العناوين الفرعية Simplified Arabic قياس 13 عريض.
- ج . يجب مراعاة أن يكون قياس الصور والجداول المدرجة في البحث لا يتعدى 12سم.
- 8- في حال عدم إجراء البحث وفقاً لما ورد أعلاه من إشارات فإن البحث سيهمل ولا يرد البحث إلى صاحبه.
- 9- تقديم أي بحث للنشر في المجلة يدل ضمناً على عدم نشره في أي مكان آخر، وفي حال قبول البحث للنشر في مجلة جامعة البعث يجب عدم نشره في أي مجلة أخرى.
- 10- الناشر غير مسؤول عن محتوى ما ينشر من مادة الموضوعات التي تنشر في المجلة

11- تكتب المراجع ضمن النص على الشكل التالي: [1] ثم رقم الصفحة ويفضل استخدام التهميش الإلكتروني المعمول به في نظام وورد WORD حيث يشير الرقم إلى رقم المرجع الوارد في قائمة المراجع.

تكتب جميع المراجع باللغة الانكليزية (الأحرف الرومانية) وفق التالي:
آ . إذا كان المرجع أجنبياً:

الكنية بالأحرف الكبيرة . الحرف الأول من الاسم تتبعه فاصلة . سنة النشر . وتتبعها معترضة (-) عنوان الكتاب ويوضع تحته خط وتتبعه نقطة . دار النشر وتتبعها فاصلة . الطبعة (ثانية . ثالثة) . بلد النشر وتتبعها فاصلة . عدد صفحات الكتاب وتتبعها نقطة . وفيما يلي مثال على ذلك:

-MAVRODEANUS, R1986- Flame Spectroscopy. Willy, New York, 373p.

ب . إذا كان المرجع بحثاً منشوراً في مجلة باللغة الأجنبية:

. بعد الكنية والاسم وسنة النشر يضاف عنوان البحث وتتبعه فاصلة، اسم المجلد ويوضع تحته خط وتتبعه فاصلة . المجلد والعدد (كتابة مختزلة) وبعدها فاصلة . أرقام الصفحات الخاصة بالبحث ضمن المجلة . مثال على ذلك:

BUSSE,E 1980 Organic Brain Diseases Clinical Psychiatry News , Vol. 4. 20 – 60

ج . إذا كان المرجع أو البحث منشوراً باللغة العربية فيجب تحويله إلى اللغة الإنكليزية و التقيد

بالبنود (أ و ب) ويكتب في نهاية المراجع العربية: (المراجع In Arabic)

رسوم النشر في مجلة جامعة البعث

1. دفع رسم نشر (40000) ل.س أربعون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.
2. دفع رسم نشر (100000) ل.س مئة الف ليرة سورية عن كل بحث للباحثين من الجامعة الخاصة والافتراضية .
3. دفع رسم نشر (200) مئتا دولار أمريكي فقط للباحثين من خارج القطر العربي السوري .
4. دفع مبلغ (6000) ل.س ستة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.

المحتوى

الصفحة	اسم الباحث	اسم البحث
48-11	م. محمد أسامة عبيد	تصميم نظام قيادة ونظام تحكّم لتمكين طائرة رباعية المراوح مزودة بنظام استقرار من ملاحقة مسار
68-49	د. محمد الخطيب	التحديد الكمي لبينوكسينات هيدروكلورايد في المستحضرات الصيدلانية بالطريقة الطيفية الضوئية الحركية
92-69	م. علي محرز د. وسيم رمضان د. ناصر أبو صالح	تقييم أداء نماذج التعلم العميق على مهمة الكشف عن الموقف تجاه الأخبار الزائفة
132-93	م. عمار مقصود د. ماهر عباس	تحسين دقة التعرف على النشاط البشري باستخدام المشفر التلقائي، الذاكرة طويلة المدى والمحول
168-133	م. ربيع الكردي د. كمال السلوم د. وسيم رمضان	تفسير نموذج مُدرّب لاكتشاف خطاب الكراهية في التغريدات العربية

تصميم نظام قيادة ونظام تحكّم لتمكين طائرةٍ

رباعية المراوح مزودة بنظام استقرارٍ

من ملاحقة مسارٍ

¹ الباحث: م. محمد أسامة عبيد

الملخص

يهدف هذا البحث إلى بناء نظام قيادة ونظام تحكّم لتمكين طائرةٍ رباعية المراوح مزودة بنظام استقرارٍ من ملاحقة مسارٍ، يتكوّن المسار من نقط مسارٍ والقطع المستقيمة الواصلة بينها. يستفيد العمل من ميزات هذا النوع من الطائرات في القدرة على التعليق والقدرة على الحركة في أي اتجاهٍ في المستوي الأفقي لتحقيق الطيران وفق خطوط المسار والوصول إلى النقط المحددة له دون تجاوزٍ، وتعديل توجه الطائرة أثناء الملاحقة بحيث يصبح موافقاً لاتجاه قطع المسار التي تسير وفقها، وهذا الأمر أساسي لتطبيقات مختلفة كتصوير الأفلام وملاحقة الأهداف، ونقل الأغراض، والاستطلاع والمراقبة، الخ. تمت الاستفادة من متحكّمات الاستقرار (الوضع الزاوي والتوجه) والارتفاع التي تقدمها المكتبة rvctools، وبدلاً عن متحكّم الموضع في هذه المكتبة جرى تصميم متحكّم لملاحقة المسار يؤمّن إلغاء خطأ الملاحقة العرضي، وإلغاء المسافة المتبقية، كما يؤمّن التحكّم بسرعة الطائرة أثناء الملاحقة. كما تمّ بناء نظام قيادة يعتمد على حلقات التحكم السابقة، ويؤمن قيادة الطائرة على المسار بالشكل المرغوب، وقد جرى اختبار أداء نظام القيادة والتحكّم بواسطة المحاكاة على نموذج الطائرة X-4 Flyer II، حيث بينت النتائج أن النظام المُصمّم يحقق المطلوب بكلفة حسابية منخفضة.

الكلمات المفتاحية: الطائرة المسيّرة رباعية المراوح، القيادة التلقائية للطائرة، زوايا أولر، مصفوفات الدوران، ملاحقة مسارٍ، نظام قيادة، نظام تحكّم.

¹ مهندس نظم إلكترونية في المعهد العالي للعلوم التطبيقية والتكنولوجيا - دمشق - سورية

Design of a Guidance and a Control System to Enable a Quadrotor equipped with a stability system to Follow a Path

Mohammed Osama Obeid

Electronic Systems Engineer at HIAST – Damascus – Syria

Abstract

This paper aims to build a guidance and a control system to enable a Quadrotor UAV provided with a stability control system to follow a path. The path consists of waypoints and straight segments connecting them. This work exploits the quadrotor ability of hovering and moving in any horizontal direction to fly along path segments and reach waypoints without exceeding, and to align the UAV heading with the followed path segment during path following. This behavior is important for a variety of applications such as surveillance and observation, filming, photographing, target tracking, etc. The stability controllers (attitude and heading) and altitude provided by the rvctools (robotics, vision and control toolboxes) library have been utilized. Instead of using the position controller in this library, a path following controller has been

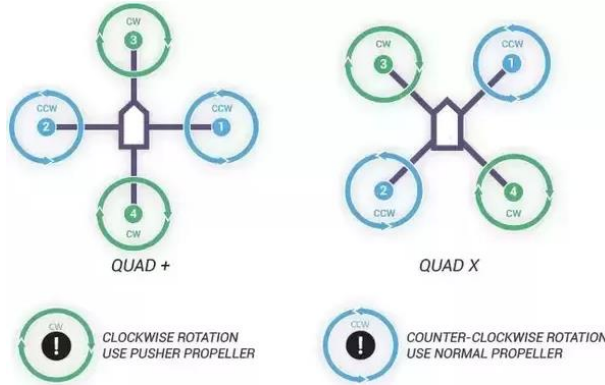
designed to cancel lateral tracking error and remained distance, while controlling the UAV velocity during path following.

A guidance system has been also built based on the previous control loops, ensuring desired flight along the path. The performance of the guidance and control system has been tested through simulation on the nonlinear dynamical model of X-4 Flyer II UAV, where the results showed that the designed system achieves the desired objectives at low computational cost .

Keywords: Quadrotor, UAV, Autonomous UAV Quadrotor, Euler Angles, Rotation Matrices, Path Following, Guidance System, Control System.

1- مقدّمة

الطائرة المسيّرة رباعية المراوح عبارة عن طائرة مسيّرة ذات إقلاع وهبوطٍ شاقوليين (VTOL: Vertical Take-off and Landing) مزوّدة بأربع مراوح يقودها أربعة محرّكات، تتمتع بُنيته بسهولة التّصميم، كما تتميز بسهولة تحريكها والمناورة بها، حيث أنّها قادرة على التعلّق في موقع ثابت (hovering) والحركة الأمامية والخلفية والجانبية، الأمر الذي يمكنها من الطيران حتّى في الأماكن الضيّقة وضمن الأبنية (Indoor). يُستخدم هذا النوع من الطائرات للمراقبة والاستطلاع وإيصال المساعدات عند وقوع الكوارث، السّيئة الأساسيّة لهذا النوع من الطائرات هو قصر مدّة الطيران. تتألّف الطائرة من هيكلٍ مئتين خفيف الوزن مُصنّع عادةً من البلاستيك أو الألمنيوم أو ألياف الكربون، للهيكل عدّة نماذج أهمّها: التّمودج (+) والتّمودج (x)، يتألّف الهيكل في هذين التّمودجين من حجرة مركزيّة وأربع أذرعٍ متماثلةٍ بحيث يتحقّق التناظر بالنسبة للمركز الهندسيّ للهيكل، تضمّ الحجرة المركزيّة الحساسات وبطاقة التّحكّم ووحدة الاتّصالات ووحدة التّغذية، أما الأذرع فتحمّل كلّ منها في نهايتها محرّكاً مزوّداً بمروحة. يدور كل زوجٍ متقابلٍ من المراوح بالاتّجاه نفسه، ويدور الزوج الآخر بالاتّجاه المعاكس وذلك لإلغاء محصلة عزم الفتل المتولد عن دوران المراوح.

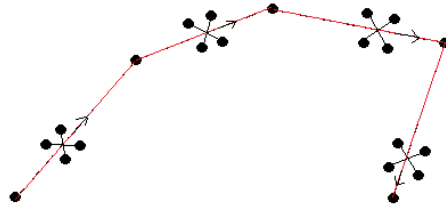


الشكل (1): دوران المراوح في طائرتين رباعية من التّمودج x أو + [1].

نظراً للحاجة المتزايدة إلى وجود طائراتٍ مسيّرةٍ تتمتعُ بخاصيةِ التعلّيقِ وسهولةِ التّحركِ، فقد انتشرت بكثرةٍ في السّنوات الأخيرة نماذج تجاريةً جاهزةً من الطائرات رباعية المرواح مزودةً بنظام استقرارٍ يؤمّن التّحكّم بالوضع الزاويّ للطائرة حول محاورها الأساسية، وذلك كما في الطّائرات (DJI Phantom, DJI Inspire, Parrot AR.Drone, Yuneec) كما في الطّائرات (Typhoon, 3D Robotics Solo, Syma Cheer wing, Parrot Bebop, etc.) وبسبب كلفتها المنخفضة وسهولة استخدامها أصبحت الطائرة رباعية المرواح منصةً هامةً لاختبار خوارزميات القيادة وطرق التّحكّم المختلفة.

لجعل الطائرة الرباعية فعالةً في أداء مهامّ تطبيقيةٍ كأعمال المراقبة والبحث وإيصال المساعدات لأبدّ من تزويدها بمتحكّمٍ لملاحقة مسارٍ معطى، بحيث يمكنها من الطيران بشكلٍ تلقائيّ (Autonomous) وفق هذا المسار لتحقيق الهدف المطلوب.

على خلاف طائرات الجناح الثابت تتمتع الطائرات رباعية المرواح بالقدرة على التعلّيق، الأمر الذي يمكنها من تنفيذ مسارٍ يضمن الطيران وفق خطوط المسار والوصول إلى النّقط المحدّدة له دون تجاوز، وتعديل توجّحها ليصبح موافقاً لاتّجاه قطع المسار التي تسير وفقها (كما يبيّن الشّكل (2))، الأمر الذي يسهّل تنفيذ تطبيقاتٍ مختلفةٍ كأعمال التّصوير والمراقبة والاستطلاع وتصوير الأفلام وإيصال الأغراض، الخ.



الشكل (2): المسار المقترح في حالة طائرة مسيّرة رباعية المرواح.

2- دراسة مرجعية

مع بداية القرن الحادي والعشرين بدأت نماذج تجاريةً جاهزةً من الطائرات رباعية المرواح بالانتشار، وأصبح استخدامها شائعاً بين الهواة بسبب كلفتها المنخفضة وسهولة وأمان

استخدامها مقارنةً بالطائرات المسيّرة الأخرى، كما جذبت هذه الطائرة اهتمام مراكز البحث والجامعات فأصبحت منصّة هامةً لاختبار خوارزميات القيادة وطرق التّحكّم المختلفة.

تُعتبر الأعمال [4] و[5] و[6] من أهمّ الأعمال التي درست أساسيات الطائرة رباعية المراوح على الصّعيد الأكاديمي، إذ تمّ إيجاد النّمودج الرّياضي لها، ودراسة القوى والعزوم الأيروديناميكية المؤثّرة عليها، وتصميم خوارزمية لاختيار مكوّناتها، كما تمّ تصميم متحكّماتٍ للتّحكّم بكلّ من الوضع الزّاويّ والتوجّه والموضع للطائرة باستخدام عدّة طرق، وتمّ التّحقّق من أداء المتحكّمات من خلال المحاكاة تمّ عملياً.

أمّا فيما يتعلّق بموضوع المسار بشكلٍ خاصّ، فقد قام [7] (على سبيل المثال) بكتابة الطّاقة الكهربائية المقدّمة للطائرة كتابع كلفةٍ بدلالة سرعة دوران المحرّكات، ثمّ قام بحل مسألة أمثلةٍ لإيجاد سرعة دوران المحرّكات التي تؤمّن تصغير تابع الكلفة بين موضعين ابتدائيّ ونهائيّ للطائرة، وبالتالي جعل الطائرة تسلك مساراً بين الموضعين تكون فيه الطّاقة الكهربائية التّيّس تجرّها المحرّكات أصغر ما يمكن، واكتفى بالاختبار بواسطة المحاكاة فقط.

في [8] تمّ تصميم خوارزمية قيادةٍ بالاعتماد على الخوارزمية (PG: Pursuit Guidance) لتمكين طائرة مزوّدة بنظام استقرار (PX4 Autopilot) من ملاحقة مسار، وذلك من خلال جعل الطائرة تلاحق هدفاً افتراضياً موضعه وسرعته معروفان خلال الزّمن. تولّد خوارزمية القيادة أمرى توجّه (ψ_c) وانقضاض (θ_c) لشعاع سرعة الطائرة، يُشتقّ منهما التّسارعان الأفقيّ والشاقوليّ المطلوبان من الطائرة، ثمّ يتمّ حساب المركّبات الديكارتيّة لشعاع التّسارع الذي يُستخدم بعد ذلك لحساب أوامر الوضع الزّاويّ وقوّة الرّفع اللاّزمة لجعل الطائرة متوجّهةً باتجاه الهدف. تمّ التّحقّق من الخوارزمية من خلال المحاكاة ثمّ بالتّطبيق العمليّ في بيئةٍ مخبريّة، ونجد في [9] عملاً مشابهاً لكن مع تصميم خوارزمية قيادةٍ بالاعتماد على خوارزمية (TS: Trajectory Shaping).

في [10] جرى تصميم خوارزميةٍ لملاحقة مسارٍ مؤلّفٍ من نقط مسارٍ (waypoints) وقطعٍ مستقيمةٍ تصل بينها مع تجنّب للعوائق الموجودة على المسار، وذلك عن طريق

تعديل زاوية توجه الطائرة باتجاه المسار وتحريكها للأمام فقط بسرعة ثابتة (أي تم التعامل مع الطائرة الرباعية كما يتم التعامل مع طائرات الجناح الثابت)، وعند ملاحظة عائق يتم تعديل زاوية توجه الطائرة للابتعاد عنه، وبعد اجتيازه تعود الطائرة لملاحقة المسار، وجرى التحقق من خوارزمية الملاحقة من خلال المحاكاة فقط.

قام [11] بتدريب شبكة عصبونية لتمكين طائرة رباعية المرواح مزودة بنظام استقرار من ملاحقة مسار مع تجنب العوائق خلال الملاحقة، يتألف المسار من نقط محدّدة وخطوط منحنية مُشكّلة من كثيرات حدود استيفائية من الدرجة الثالثة تصل بين هذه النقط. خوارزمية الإشراف المستخدمة في تدريب الشبكة (MPPC: Model Predictive Contouring Controller)، وهي مطوّرة في المعهد الفيدرالي السويسري للتقانة (ETH Zurich). دخل الخوارزمية هو الشعاع $o_t = (d_t, v_t, l_t)^T$ حيث d_t هو الخطأ العرضي بين الطائرة والمسار، v_t سرعة الطائرة في المستوي الأفقي، l_t ارتفاع الطائرة، أما خرج الخوارزمية فهو شعاع تحكّم $u_\pi = (v_z, \phi_d, \theta_d)^T$ يضمّ السرعة المرغوبة من الطائرة وفق المحور z للجملة المرجعية، وزاويتي الميل والانقضاض المرغوبتين من الطائرة، دخل الشبكة العصبونية وخرجها متوافقان مع دخل وخرج خوارزمية الإشراف. يهدف شعاع التحكّم في خوارزمية الإشراف إلى جعل الخطأ العرضي بين الطائرة والمسار أصغر ما يمكن، وجعل مسافة التقدّم على المسار أكبر ما يمكن وذلك لتأمين قطع المسار بأكثر سرعة ممكنة (أقصر زمن ممكن). يتمّ التحكّم بتوجه الطائرة خلال الملاحقة بشكل مستقلّ عن خوارزمية القيادة بحيث تبقى // متوافقة مع المسار. تمّ التحقق من أداء الخوارزمية من خلال المحاكاة على نموذج جاهز في البيئة Gazebo، ثمّ عملياً على الطائرة Parrot Bebop2.

احتلت خوارزميات معالجة الصورة دوراً في تصميم المسارات وتجنب العوائق للطائرات رباعية المرواح، ونجد في [12] مثلاً على ذلك، حيث قام الباحثون في هذا العمل ببناء مسارٍ ومنحكّمٍ لملاحقة هذا المسار باستخدام تقنيات معالجة الصورة، وتمّ التحقق من أداء الخوارزمية عملياً باستخدام الطائرة Parrot Bebop، نجد كذلك عملاً آخر في هذا

المجال في [16]. في كلّ من ([13]، [14]، [15]، [17]، [18]، [19]، [20]) نجد خوارزمياتٍ مختلفةً لتصميم المسارات وملاحقتها، وتطبيق هذه الخوارزميات على طائراتٍ مجهزةٍ بنظام استقرارٍ بعد تزويدها بالحساسات المناسبة وبنية إلكترونية تسمح بتنفيذ هذه الخوارزميات.

مناقشة الدراسة المرجعية

إنّ الخوارزميات السابقة مختلفةً في طريقة معالجتها لمسائل تصميم المسار وملاحقتها، وهي كذلك مختلفةً في المجالات التطبيقية التي يمكن توظيفها فيها.

مثلاً، لا يهتم [7] بشكل الطّريق المسلك بين نقطتي مسارٍ، ولذلك لا يمكن توظيف نتائج هذه الدراسة في أعمال الاستطلاع والمراقبة، لكن يمكن الاستفادة منها في نقل الأغراض بين مكانين مثلاً.

يمكن الاستفادة من كلّ من [8] و[9] في تطبيقات ملاحقة الأهداف المتحرّكة، لكن استخدامها بهدف مراقبة موضع ما والتعليق عنده (hovering) قد لا يكون متاحاً لأنّ العمليّات الحسّابية في كلّ من هاتين الدّراستين قد تقود إلى حالات عدم تعيينٍ.

تلغي الدّراسات [10] و[13] و[14] إمكانيّة الحركة العرضية للطائرة رباعية المرواح، وهي بذلك تفرض قيوداً غير ضرورية على كيفية تحريك الطائرة، الأمر الذي يمنع الاستفادة من إمكانيّاتها، لذلك ربّما من الأفضل استثمار هذه الدّراسات في التّطبيقات المطلوبة من طائرات الجناح الثّابت خصوصاً أنّها تتمتع بزمن طيرانٍ أكبر.

استخدم [11] أسلوب التعليم العميق لتحقيق ملاحقة بزمنٍ أصغريٍّ لمسارٍ في بيئةٍ مخبريةٍ مع تجنّب للعوائق . إنّ الملاحقة بزمنٍ أصغريٍّ تقتضي استنزاف المدخّرة الكهربائية الموجودة في الطائرة، وإجهاد محرّكات الطائرة ومتحكّمات السرعة الخاصّة بها (ESCs: Electronic Speed Controllers)، ولذلك تُعتبر هذه الخوارزمية غير مناسبة لتطبيقات المراقبة والاستطلاع، ولكنّها فعّالة من أجل السّباقات.

الدراسات التي تعتمد على تقنيات معالجة الصورة (مثل [12] و[16]) مفيدة لتطبيقات التعرف على الأشكال وملاحقة الأهداف، لكن استخدامها في خوارزميات تصميم المسارات وملاحقتها يُشغل الكاميرا ويحد من إمكانية الاستفادة منها في أعمال الاستطلاع والمراقبة، أو يجب تزويد الطائرة بعدة كاميرات تُستخدم إحداها لأعمال الاستطلاع والمراقبة، وتُستخدم بقية الكاميرات لملاحقة المسار وتجنب العوائق والأغراض التطبيقية الأخرى.

3- مسألة البحث

تتلخص مسألة البحث في بناء نظام قيادة وتحكم لتمكين طائرة مسيرة رباعية المرواح مزودة بنظام استقرار من ملاحقة مسار مؤلف من نقط مرتبة Waypoints تتوضع في مستوى أفقي واحد، ومن القطع المستقيمة الواصلة بينها (2D Path)، وذلك دون تجاوز للنقط المحددة للمسار، مع تعديل توجه الطائرة أثناء المسير (وليس عند محددات المسار) ليصبح موافقاً للقطعة المستقيمة التي تسير الطائرة وفقها كما يبين الشكل (2).

4- أهمية البحث

إن الخط المستقيم هو أقصر طريق يصل بين نقطتين، لذلك يؤمن حل مسألة البحث بشكلٍ حدسيّ تخفيف استهلاك الوقود خلال ملاحقة المسار، كما يؤمن تجنب العوائق الثابتة ذات الإحداثيات المعروفة بالنسبة للمستثمر إذ أن الملاحقة تتم دون تجاوز لنقط المسار.

إن تعديل توجه الطائرة ليصبح موافقاً لقطعة المسار التي تسير وفقها يُسهل التعرف على أماكن الأهداف بالنسبة للمسار وبالتالي تسهيل أعمال الاستطلاع والمراقبة، كما يسهل تنفيذ أعمال التصوير في حال استخدام كاميرا أمامية، ويضمن تعديل توجه الطائرة خلال ملاحقة قطعة المسار عدم إضاعة وقت في انتظار تعديل التوجه عند الانتقال بين قطعتي مسار متتاليتين.

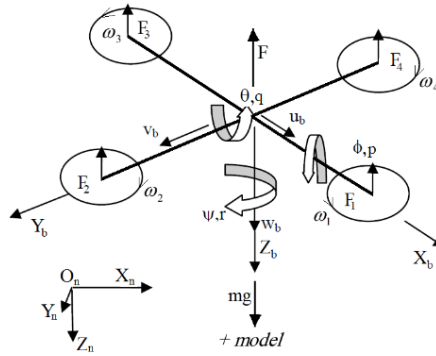
إنّ ملاحقة المسار بالشكل المطروح يمكن أن يستثمر في تنفيذ تطبيقات مختلفة كأعمال الاستطلاع والمراقبة أو توصيل الأغراض بين موضعين أو أكثر أو تصوير الأفلام.

5- حدود البحث

في هذا البحث يتم اعتماد الفرضيات التالية:

- هيكل الطائرة قاسٍ (قوى المرونة مهملة)، ومتناظرٌ بالنسبة للمركز الهندسي.
- مركز ثقل الطائرة منطبقٌ على مركزها الهندسي في المستوي الأفقيّ.
- أثر الرّياح مهملاً على حركة الطائرة.

6- النموذج الديناميكي للطائرة رباعية المراوح



الشكل (3): العناصر الحركية والتحريرية الأساسية للطائرة الرباعية [23] (مع تعديل).

يُعطى النموذج الديناميكي للطائرة رباعية المراوح بالعلاقتين [3]:

$$(1) \quad m \begin{pmatrix} \ddot{x}_n \\ \ddot{y}_n \\ \ddot{z}_n \end{pmatrix} = m \begin{pmatrix} 0 \\ 0 \\ g \end{pmatrix} - \mathbf{R}_{n/b} \begin{pmatrix} 0 \\ 0 \\ F \end{pmatrix}$$

$$(2) \quad \mathbf{I}_b \cdot \dot{\boldsymbol{\Omega}}_b = -\boldsymbol{\Omega}_b \times \mathbf{I}_b \cdot \boldsymbol{\Omega}_b + \boldsymbol{\tau}_b$$

حيث:

- الجملة $O_n X_n Y_n Z_n$ هي جملة أرضية موضعية نعتبرها جملة مرجعية لحساب موضع الطائرة $\mathbf{P}_n(x_n, y_n, z_n)^T$ وبقية العناصر الحركية لها (كالسرعة مثلاً)، محاور الجملة موجهة باتجاه الشمال الجغرافي والشرق الجغرافي والأسفل على الترتيب.
- الجملة $O_b X_b Y_b Z_b$ هي جملة مرتبطة بجسم الطائرة، محاورها هي المحاور الأساسية للطائرة، موجهة باتجاه الأمام واليمين والأسفل على الترتيب.
- ω_i سرعة دوران المروحة ذات الترتيب $(1 \leq i \leq 4)$ ، F_i قوة رفعها، F قوة الرفع الكلية للمراوح.
- g القيمة المطلقة لتسارع الجاذبية الأرضي.

$$R_{n/b} = \begin{pmatrix} C_\theta C_\psi & S_\phi S_\theta C_\psi - C_\phi S_\psi & C_\phi S_\theta C_\psi + S_\phi S_\psi \\ C_\theta S_\psi & S_\phi S_\theta S_\psi + C_\phi C_\psi & C_\phi S_\theta S_\psi - S_\phi C_\psi \\ -S_\theta & S_\phi C_\theta & C_\phi C_\theta \end{pmatrix} \quad \bullet$$

هيمصفوفة دوران من الجملة المرتبطة بالجسم إلى الجملة المرجعية وفق زوايا أولر (ϕ, θ, ψ) ، الرمز C يدل على التابع $\cos()$ ، والرمز S يدل على التابع $\sin()$.

- m كتلة الطائرة، العزم الكلي المؤثر على الطائرة حول مركز ثقلها مكتوباً بدلالة الجملة المرتبطة بالجسم، $\mathbf{\Omega}_b = (p, q, r)^T$ شعاع السرعة الزاوية للطائرة بالنسبة للجملة المرجعية بدلالة الجملة المرتبطة بالجسم، $\mathbf{V}_b(t) = (u, v, w)^T$ سرعة مركز ثقل الطائرة بالنسبة للجملة المرجعية بدلالة الجملة المرتبطة بالجسم، \mathbf{I}_b مصفوفة العطالة للطائرة عند مركز ثقلها مكتوبةً بدلالة الجملة المرتبطة بالجسم.

لتصميم نظام تحكّم لملاحقة مسارنهم بجملة المعادلات (1)، عادةً يجري تصميم المتحكّمات بالنظم بعد تقريب النموذج اللاخطّي لهذه النظم إلى نموذج خطّي، وذلك لتسهيل عمليّة التصميم وتبسيط المتحكّم ما أمكن، ثمّ يجري اختبار أداء المتحكّمات على النموذج اللاخطّي، لذلك سيتمّ تقريب نموذج الطائرة حول نقطة توازن، نقطة التوازن البديهية للطائرة الرباعية هي نقطة التعليق (hovering)، عند هذه النقطة تكون قيمة كلّ من θ و ϕ صغيرة، ويكون $F \approx m.g$ ، وبالتالي يمكن إجراء التقريب ([3]، [24]):

$$(3) \quad \begin{cases} C_\theta = 1, C_\phi = 1 \\ S_\theta = \theta, S_\phi = \phi \end{cases}$$

وذلك بشرط أخذ قيم θ و ϕ بالراديان ([3]، [24])، وتصبح جملة المعادلات (1):

$$(4) \quad \begin{pmatrix} \ddot{x}_n \\ \ddot{y}_n \\ \ddot{z}_n \end{pmatrix} = \begin{pmatrix} -(\theta.C_\psi + \phi.S_\psi).g \\ -(\theta.S_\psi - \phi.C_\psi).g \\ g - F/m \end{pmatrix}$$

7- بناء نظام قيادة للطائرة لملاحقة مسار

قبل البدء بتصميم نظام القيادة ونظام التّحكّم لابدّ من تعريف بعض المفاهيم الأساسية:

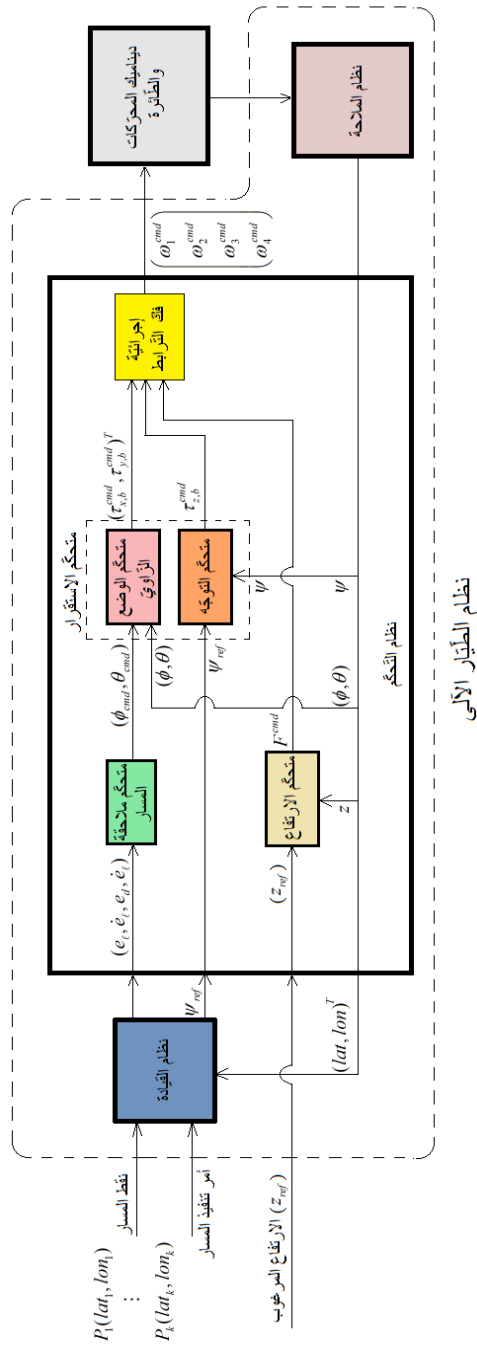
الطّيّار الآلي (Autopilot or Automatic Pilot): هو نظام يتحكّم بمركباتٍ جويّةٍ أو أرضيّةٍ أو بحريّةٍ بدون تدخّلٍ بشريٍّ دائمٍ. فيما مضى كانت نظم الطّيّار الآلي بسيطةً تهدف لتحقيق الاستقرار وثبّيت زاوية التوجّه Heading أو زاوية الاتجاه Course للمركبة، أمّا في الوقت الرّاهن فهي معقّدةٌ وتتألّف من ثلاثة نظمٍ جزئيّةٍ Subsystems مستقلّةٍ هي:

- نظام القيادة Guidance System: هو نظامٌ يقوم بشكلٍ مستمرٍّ بحساب قيم الموضع والسرعة والتسارع اللازمة للمركبة لتنفيذ المهمة المطلوبة، كما يقوم بحساب المسار المطلوب من المركبة ملاحظته، وهذا يمكن أن يتم باستخدام طرائق أمثلة متقدمة كأمثلة استهلاك الوقود أو الحركة بزمنٍ أصغريٍّ أو تجنب العوائق، أو يقوم بحساب المسار انطلاقاً من نقط مسارٍ مرتبةٍ Waypoints.

- نظام الملاحة Navigation System: هو نظامٌ يهدف لقياس (أو تقدير) العناصر الحركية لمركبةٍ من موضعٍ ووضعٍ زاويٍّ وزاوية اتجاهٍ (الزاوية بين شعاع السرعة والشمال الجغرافي) وسرعةٍ زاويةٍ وسرعةٍ وتسارعٍ خطيين، يتم ذلك عادةً باستخدام نظام ملاحية عالميٍّ يعتمد على الأقمار الصناعية (GNSS: Global Navigation Satellite System) مثل GPS أو GLONASS أو GALILEO بالإضافة إلى حساسات الحركة من قائلات تسارعٍ خطيٍّ وسرعةٍ زاويةٍ، أو باستخدام نظم ملاحية عطاليةٍ (INS: Inertial Navigation System).

- نظام التحكم Control System: هو نظامٌ يقوم بحساب القوى والعزوم التي يلزم تزويد المركبة بها لتحقيق أهدافٍ تحكّميةٍ معينةٍ مثل التنظيم حول نقطة عملٍ معينةٍ أو ملاحقة مسارٍ أو تحقيق مناورةٍ أو تصغير طاقة التحكم.

يُشار إلى النظم الثلاث السابقة بنظام (Guidance, Navigation and Control) وتُشكلُ مجموعها نظام الطيار الآلي [2]، يُبين الشكل (4) نظام الطيار الآلي المستخدم في حالتنا. سنعتمد بدايةً منهجيةً بسيطةً ومنخفضة الكلفة الحاسوبية لتصميم كلٍ من نظامي

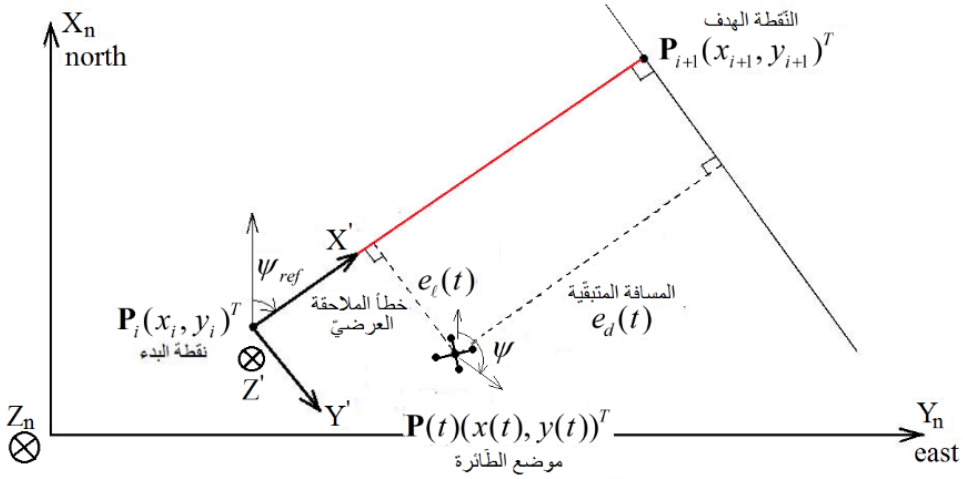


الشكل (4): نظام الطيار الآلي المستخدم في هذا البحث.

القيادة والتحكّم، وإذا كان الحلّ قاصراً أو غير ممكنٍ عندئذٍ يمكن الاستفادة من الخوارزميات المختلفة المطروحة في الدراسة المرجعية لحلّ مسألة ملاحقة المسار.

يهدف هذا البحث إلى بناء نظام قيادة للطائرة الرباعية يمكنها من ملاحقة مسارٍ يتألف من نقطٍ مرتبة متوضّعة على ارتفاعٍ واحدٍ (2D path)، والقطع المستقيمة الواصلة بينها، بحيث يكون توجّه الطائرة موافقاً للقطعة المستقيمة التي تسير وفقها، أي يهدف نظام القيادة في حالتنا إلى:

- تحديد القطعة المستقيمة المرجعية (المسار الجزئي المرجعي) التي ينبغي السير وفقها.
- حساب بُعد موضع الطائرة (موضع مركز ثقلها) عن هذه القطعة (بفرض أنّ الطائرة والمسار يقعان في مستويٍ أفقيٍّ واحدٍ)، وهو ما يُدعى بالخطأ العرضي (lateral error).
- حساب بُعد موضع الطائرة عن المستقيم العمود على هذه القطعة عند النقطة الهدف، وهو ما يُدعى بالمسافة المتبقية (remaining distance).
- حساب التوجّه المرجعيّ V_{ref} للطائرة ليصبح توجّهها موافقاً للقطعة المستقيمة التي تسير وفقها.
- الانتقال إلى قطعة المسار التالية عند تحقّق شرط الوصول إلى النقطة الهدف.



الشكل (5): المقادير المحسوبة من قبل نظام القيادة.

نرمز للنقطة ذات الترتيب i من المسار بـ $\mathbf{P}_{i,n}(x_{i,n}, y_{i,n})^T$ ، الدليل n يشير إلى جملة الإحداثيات المرجعية (يمكن عدم كتابته عند ضمان عدم الالتباس وهو ما سنفعله)، ونرمز بـ $e_l(t)$ للخط العرضي (البعد بين موضع الطائرة والقطعة $[\mathbf{P}_i, \mathbf{P}_{i+1}]$)، ونرمز بـ $e_d(t)$ للمسافة المتبقية (البعد بين موضع الطائرة والمستقيم العمود على القطعة المستقيمة $[\mathbf{P}_i, \mathbf{P}_{i+1}]$ عند النقطة \mathbf{P}_{i+1} (التي تُدعى بالنقطة الهدف))، ونرمز بـ $\mathbf{P}(t)$ لموضع الطائرة عند اللحظة t ، المقادير السابقة مبينة في الشكل (5).

أما هدف نظام التحكم فسيكون جعل كل من $e_l(t)$ و $e_d(t)$ ينتهي إلى الصفر خلال تنفيذ المسار. نربط بقطعة المسار $[\mathbf{P}_i, \mathbf{P}_{i+1}]$ جملة الإحداثيات $\mathbf{P}_i X' Y' Z'$ بحيث يكون X' منطبقاً على $[\mathbf{P}_i, \mathbf{P}_{i+1}]$ وموجهاً باتجاه \mathbf{P}_{i+1} ، ويكون Z' ممحوماً على الشاقول وموجهاً باتجاه الأسفل، Y' يحقق ثلاثية مباشرة مع X' و Z' . تنتج الجملة الجديدة عن الجملة المرجعية بانسحاب شعاعه \mathbf{P}_i ثم بدوران بالزاوية ψ_{ref} ، وبالتالي:

إحداثيات \mathbf{P}_{i+1} و $\mathbf{P}(t)$ بدلالة الجملة الجديدة:

$$(5) \quad \left\{ \begin{array}{l} \mathbf{P}'_{i+1} = \begin{pmatrix} x'_{i+1} \\ y'_{i+1} \end{pmatrix} = \mathbf{R}_{\psi_{ref}} \cdot (\mathbf{P}_{i+1} - \mathbf{P}_i) = \begin{pmatrix} C_{\psi_{ref}} & S_{\psi_{ref}} \\ -S_{\psi_{ref}} & C_{\psi_{ref}} \end{pmatrix} (\mathbf{P}_{i+1} - \mathbf{P}_i) \\ = \begin{pmatrix} (x_{i+1} - x_i)C_{\psi_{ref}} + (y_{i+1} - y_i)S_{\psi_{ref}} \\ -(x_{i+1} - x_i)S_{\psi_{ref}} + (y_{i+1} - y_i)C_{\psi_{ref}} \end{pmatrix} \\ \mathbf{P}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \mathbf{R}_{\psi_{ref}} \cdot (\mathbf{P}_{i+1} - \mathbf{P}(t)) \\ = \begin{pmatrix} (x(t) - x_i)C_{\psi_{ref}} + (y(t) - y_i)S_{\psi_{ref}} \\ -(x(t) - x_i)S_{\psi_{ref}} + (y(t) - y_i)C_{\psi_{ref}} \end{pmatrix} \end{array} \right.$$

يُعطى كلُّ من الخطأ العرضيَّ والمسافة المتبقية باستخدام جملة الإحداثيات الجديدة:

$$(6) \quad \begin{cases} e_{\ell}(t) = y'_{i+1} - y'(t) \\ e_d(t) = x'_{i+1} - x'(t) \end{cases}$$

ويُعطى الخطأ العرضيَّ والمسافة المتبقية باستخدام جملة الإحداثيات المرجعية بعد التعويض من (5) في (6):

$$(7) \quad \begin{cases} e_{\ell}(t) = -(x_{i+1} - x(t))S_{\psi_{ref}} + (y_{i+1} - y(t))C_{\psi_{ref}} \\ e_d(t) = (x_{i+1} - x(t))C_{\psi_{ref}} + (y_{i+1} - y(t))S_{\psi_{ref}} \end{cases}$$

سيحتاج متحكم ملاحقة المسار للمشتق الزمني من المرتبتين الأولى والثانية لكل من الخطأ العرضيَّ والمسافة المتبقية، باشتقاق جملة المعادلتين (7) بالنسبة للزمن نجد:

$$(8) \quad \begin{cases} \dot{e}_\ell(t) = \dot{x}(t) \cdot S_{\psi_{ref}} - \dot{y}(t) \cdot C_{\psi_{ref}} \\ \dot{e}_d(t) = -\dot{x}(t) \cdot C_{\psi_{ref}} - \dot{y}(t) \cdot S_{\psi_{ref}} \end{cases}$$

بالاشتقاق مرّةً أخرى بالنسبة للزّمن:

$$(9) \quad \begin{cases} \ddot{e}_\ell(t) = \ddot{x}(t) \cdot S_{\psi_{ref}} - \ddot{y}(t) \cdot C_{\psi_{ref}} \\ \ddot{e}_d(t) = -\ddot{x}(t) \cdot C_{\psi_{ref}} - \ddot{y}(t) \cdot S_{\psi_{ref}} \end{cases}$$

بالتعويض من (4) في (9) نجد ديناميك الخطأ العرضيّ والمسافة المتبقّية:

$$(10) \quad \begin{cases} \ddot{e}_\ell(t) = -[\theta S_{e_\psi} + \phi C_{e_\psi}] \cdot g = -u_\ell \cdot g \\ \ddot{e}_d(t) = [\theta C_{e_\psi} - \phi S_{e_\psi}] \cdot g = u_d \cdot g \\ e_\psi = \psi_{ref} - \psi \end{cases}$$

نُحسب زاوية التوجّه المرجعيّة ψ_{ref} لقطعة المسار باستخدام التابع $atan2(y, x)$ [2]

$$(11) \quad \psi_{ref} = atan2(y_{i+1} - y_i, x_{i+1} - x_i) \text{ : كما يلي}$$

حيث $atan2(y, x)$ تابع التقابل العكسي للتابع tan ، يعطي الزاوية بالراديان ضمن المجال $[-\pi, \pi]$ ويعالج الحالات المختلفة لكلّ من x و y ، له الصيغة [21]:

$$(12) \quad atan2(y, x) = \begin{cases} \arctan(y/x) & \text{if } (x > 0) \\ \arctan(y/x) + \pi & \text{if } (x < 0) \& (y \geq 0) \\ \arctan(y/x) - \pi & \text{if } (x < 0) \& (y < 0) \\ \pi/2 & \text{if } (x = 0) \& (y > 0) \\ -\pi/2 & \text{if } (x = 0) \& (y < 0) \\ \text{undefined} & \text{if } (x = 0) \& (y = 0) \end{cases}$$

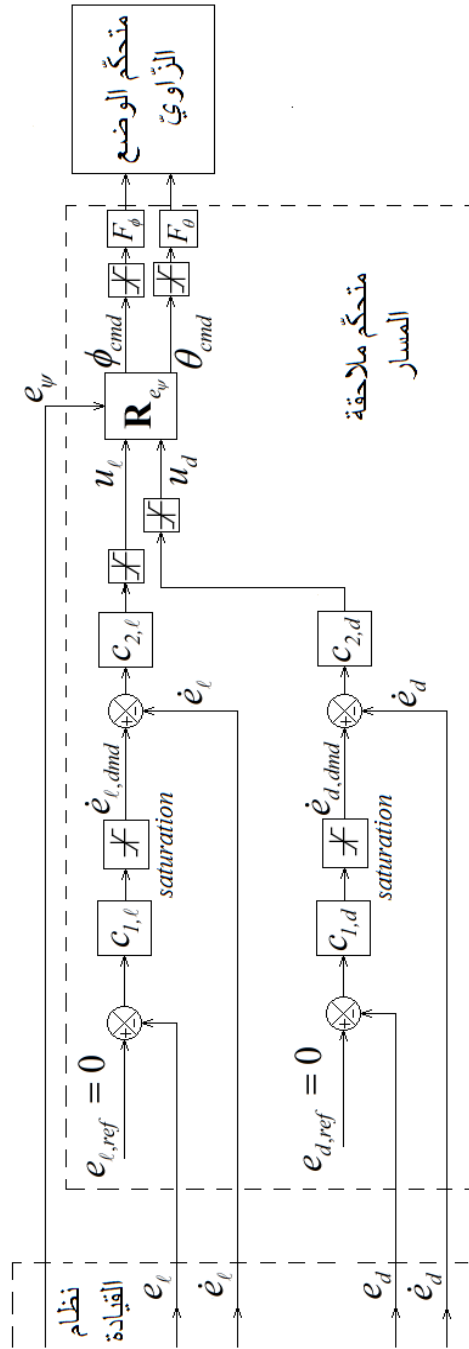
عند الحركة على مسارٍ لا بدّ من إيجاد آليّة يقرّر نظام القيادة وفقها الانتهاء من الحركة على قطعة المسار الحاليّة والانتقال إلى قطعة المسار التّالية، يتمّ ذلك بتعريف دائرة الوصول Circle of Acceptance التي هي دائرة مركزها النّقطة الهدف، ونصف قطرها R_i يُحدّد من قبل المشغّل حسب الخطأ المسموح للقيادة على المسار، كما يمكن استخدام شروط وصولٍ أخرى [2]. في هذا العمل أُعتمد الشرط:

$$(13) \quad \begin{cases} |e_l(t)| \leq e_{l,\max} \\ |e_d(t)| \leq e_{d,\max} \end{cases}$$

وفي حال تحقّقه سينتقل تنفيذ المسار إلى القطعة التّالية مباشرةً.

8- بناء نظام تحكّم للطائرة لملاحقة المسار

يهدف نظام التّحكّم المراد تصميمه إلى إلغاء كلّ من الخطأ العرضيّ والمسافة المتبقّيّة المحسوبين في نظام القيادة السّابق، يتمّ ذلك من خلال توليد أمر ميلٍ للطائرة باتجاه قطعة المسار لإلغاء الخطأ العرضيّ، وأمر ميلٍ باتجاه العمود على قطعة المسار عند النّقطة الهدف لإلغاء المسافة المتبقّيّة، ومن ثمّ يتمّ توليد أمرٍ بالزاويتين ϕ و θ للطائرة من أمرٍ الميل السّابقين باستخدام مصفوفة دورانٍ بالزاوية $(e_\psi = \psi_{ref} - \psi)$.



الشكل (6): نظام التحكم لملاحقة المسار.

يتكفل متحكّم وضع زوايٍ مستقلّ (Attitude Controller) بتنفيذ الأمرين ϕ و θ ، كما يتكفل متحكّم مستقلّ للتوجّه (Heading Controller) بتنفيذ أمر التوجّه ψ_{ref} المحسوب في نظام القيادة لجعل توجّه الطائّرة موافقاً لقطعة المسار التي تسير وفقها. ويتكفل نظام تحكّم مستقلّ للارتفاع (Altitude Controller) بالتحكّم بارتفاع الطائّرة (الشكل (4)).

يبيّن الشكل (6) مخطّط نظام التّحكّم لملاحقة المسار، في هذا المخطّط $\dot{e}_{\ell,dmd}$ هو القيمة المرغوبة لـ \dot{e}_{ℓ} الناتجة عن المتحكّم $c_{1,\ell}$ وليس المشتق الزمني لـ $e_{\ell,ref}$ لذلك تمّ استخدام الرّمز dmd بدلاً من ref لمنع الالتباس، وكذلك الأمر بالنسبة لـ $\dot{e}_{d,dmd}$.

إنّ لأيّ جسمٍ متحرّكٍ سرعةً انسحابيّةً قصوى لا يُسمح بتجاوزها سواءً في المستوي الأفقيّ أو وفق المحور الشاقوليّ، لذلك تمّ استخدام محدّدٍ في الحلقة الأساسيّة لتحديد القيمة العظمى المرغوبة لكلّ من \dot{e}_{ℓ} و \dot{e}_d ، إذ أنّ المقدار \dot{e}_{ℓ} يساوي السرعة الانسحابيّة للطائّرة باتّجاه قطعة المسار ويعاكسها بالإشارة، والمقدار \dot{e}_d يساوي السرعة الانسحابيّة للطائّرة باتّجاه العمود على قطعة المسار عند النّقطة الهدف ويعاكسها بالإشارة. نفترض في حالتنا:

$$|(\dot{e}_{\ell,dmd})_{\max}| = |(\dot{e}_{d,dmd})_{\max}| = 15m.sec^{-1}$$

u_d و u_{ℓ} إشارتا التّحكّم الناتجتان عن المتحكّمين $c_{2,\ell}$ و $c_{2,d}$ ، واحدهما هي الرّاديان وثمّتلان أمر ميل الطائّرة باتّجاه قطعة المسار وباتّجاه العمود على قطعة المسار عند النّقطة الهدف على التّرتيب.

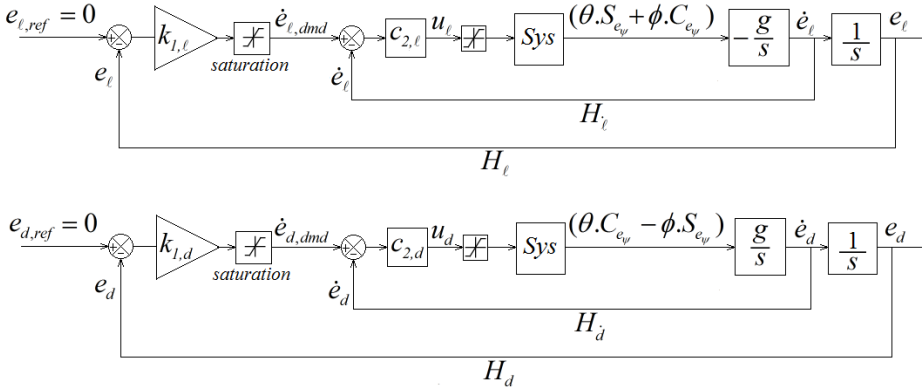
ϕ_{cmd} و θ_{cmd} إشارتا التحكّم الناتجتان عن تدوير u_ℓ و u_d بالمصفوفة \mathbf{R}_{e_ψ} ، يتمّ تمريرهما إلى متحكّم الوضع الزاويّ بعد تحديدهما ثم ترشيحهما بمرشّح تمرير تردّد منخفض (F_ϕ, F_θ) بهدف تنعيم الأوامر، نختار 15° $|(\phi_{cmd})_{\max}| = |(\theta_{cmd})_{\max}|$ لأنّ التحكّم بالطائرة يتمّ من أجل قيمٍ صغيرة لـ (ϕ, θ) ، لدينا من جملة المعادلات (10):

$$\begin{cases} u_\ell = \phi C_{e_\psi} + \theta S_{e_\psi} \\ u_d = -\phi S_{e_\psi} + \theta C_{e_\psi} \end{cases} \Leftrightarrow \begin{pmatrix} u_\ell \\ u_d \end{pmatrix} = \underbrace{\begin{pmatrix} C_{e_\psi} & S_{e_\psi} \\ -S_{e_\psi} & C_{e_\psi} \end{pmatrix}}_{\mathbf{R}_{e_\psi}^T} \begin{pmatrix} \phi \\ \theta \end{pmatrix}$$

وبالتالي تكون أوامر الزوايا لمتحكّم الوضع الزاويّ:

$$\begin{pmatrix} \phi_{cmd} \\ \theta_{cmd} \end{pmatrix} = (\mathbf{R}_{e_\psi}^T)^T \begin{pmatrix} u_\ell \\ u_d \end{pmatrix} = \mathbf{R}_{e_\psi} \begin{pmatrix} u_\ell \\ u_d \end{pmatrix} = \begin{pmatrix} C_{e_\psi} & -S_{e_\psi} \\ S_{e_\psi} & C_{e_\psi} \end{pmatrix} \begin{pmatrix} u_\ell \\ u_d \end{pmatrix}$$

يُبين الشكل (7) مخطّطيّ التحكّم لإلغاء كلّ من الخطأ العرضيّ والمسافة المتبقية بشكلٍ منفصلٍ عن بعضهما، حيث جرى تضمين مصفوفة الدوران \mathbf{R}_{e_ψ} في النظام Sys في كلّ من المخطّطين، يضمّ Sys إلى جانب \mathbf{R}_{e_ψ} كلاً من متحكّم الوضع الزاويّ وآليّة فكّ الترابط ومتحكّمات سرعة المحرّكات وديناميك المحرّكات والطائرة، ويقوم بتأمين إشارة التحكّم u_ℓ الناتجة عن المتحكّم $C_{2,\ell}$ بشكلٍ صريحٍ في H_ℓ ، وتأمين إشارة التحكّم u_d الناتجة عن المتحكّم $C_{2,d}$ بشكلٍ صريحٍ في H_d أي يُعتبر مفعلاً.



الشكل (7): حلقتنا التحكم لإلغاء الخطأ العرضي والمسافة المتبقية.

في الشكل (7) لدينا H_l هي حلقة التحكم بـ \dot{e}_l و H_d هي حلقة التحكم بـ e_l ، و H_d هي حلقة التحكم بـ \dot{e}_d و H_l هي حلقة التحكم بـ e_d . بافتراض أن Sys نظامً واحدٍ بالنسبة لكل من H_d و H_l ، أي هو حلقةً داخليةً بالنسبة لكل من H_d و H_l تنفذ الأمر الوارد إليها وتصل إلى الحالة المستقرة $steady state$ خلال زمنٍ غير محسوسٍ بالنسبة لـ H_d و H_l ، عملياً تُعتبر الحلقة الداخليةً واحدةً ضمن حلقةٍ خارجيةٍ إذا كانت استجابة الحلقة الداخليةً أسرع بثلاث مراتٍ على الأقل من استجابة الحلقة الخارجية، أي الحلقة الداخليةً تصل إلى حالتها المستقرة بثلاث الزمن على الأكثر الذي تتطلبه الحلقة الخارجية لتصل إلى حالتها المستقرة [22]. بالنظر إلى جملة المعادلات (10)، نجد:

$$(14) \quad \begin{cases} \ddot{e}_l(t) = -u_l \cdot g \Rightarrow s \dot{e}_l = -u_l \cdot g \Rightarrow \frac{\dot{e}_l}{u_l} = \frac{-g}{s} \\ \ddot{e}_d(t) = u_d \cdot g \Rightarrow s \dot{e}_d = u_d \cdot g \Rightarrow \frac{\dot{e}_d}{u_d} = \frac{g}{s} \end{cases}$$

أي أنّ النظام المتحكّم به في كلّ من الحلقتين H_i و H_d هو نظام من المرتبة الأولى لتأمين التحكّم به يكفي وضع متحكّم تناسبيّ، أي يكفي اختيار $c_{2,\ell} = k_{2,\ell} < 0$ و $c_{2,d} = k_{2,d} > 0$.

بافتراض أنّ H_i هي حلقةٌ واحدةٌ بالنسبة لـ H_ℓ ، عندئذٍ يكون:

$$(15) \quad H_\ell = \frac{e_\ell}{e_{\ell,ref}} = \frac{c_{1,\ell}(1/s)}{1+c_{1,\ell}(1/s)} = \frac{c_{1,\ell}}{s+c_{1,\ell}}$$

حيث s هو متحوّل لابلاس، لتكون الحلقة H_ℓ مستقرّةً يمكن أن يتمّ اختيار $c_{1,\ell} = k_{1,\ell} > 0$ (مصحّح تناسبيّ)، مع الانتباه إلى أنّ قيمة $k_{1,\ell}$ يجب أن تكون أصغر من حدّ معيّليكون زمن استجابة H_ℓ أبداً بثلاث مرّاتٍ على الأقل من زمن استجابة H_i ، وبالتالي ضمان تصرّف الحلقة H_i كحلقةٍ واحدةٍ بالنسبة لـ H_ℓ . عند دراسة الحلقة H_d نجد من أجل $c_{1,d}$ نتيجةً مشابهةً لما وجدناه من أجل $c_{1,\ell}$.

الشروط على أداء متحكّم ملاحقة المسار

إنّ حلقة التّحكّم بالوضع الزاويّ هي حلقةٌ داخليةٌ بالنسبة لكلّ من H_i و H_d ، زمن استجابتها $t_{r,attitude}$ ، لذلك ليكون نظام التّحكّم لملاحقة المسار مستقرّاً يجب أن يتحقّق $t_{r,\ell}, t_{r,d} > 3 * t_{r,attitude}$ ، حيث $t_{r,\ell}$ و $t_{r,d}$ هما زمنا استجابة حلقتي التّحكّم H_i و H_d على الترتيب، أي أنّ قيم الأرباح $(c_{1,\ell}, c_{2,\ell}, c_{1,d}, c_{2,d})$ متعلّقةٌ بـ $t_{r,attitude}$ زمن استجابة حلقة التّحكّم بالوضع الزاويّ.

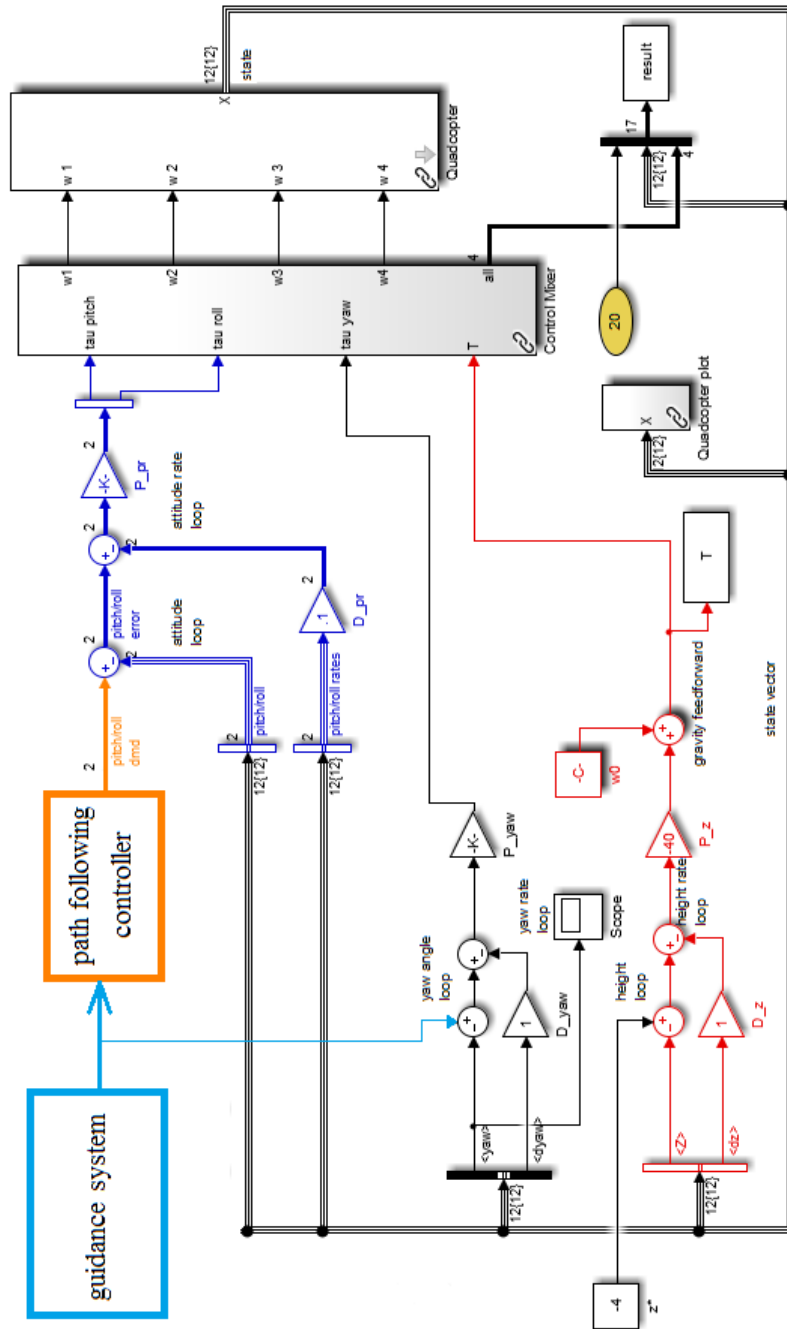
- ينبغي أن يكون كلّ من e_ℓ و e_d ضمن المجال $\pm 15m.sec^{-1}$.

- ينبغي أن يكون كلٌّ من u_d و u_ℓ و ϕ_{cmd} و θ_{cmd} ضمن المجال $\pm 15^\circ$ ، ولا يحتوي أيٌّ منها على اهتزازات ذات تردداتٍ عاليةٍ.
- استجابة دون تجاوزٍ لكلٍّ من H_d و H_ℓ .

9- ضبط قيم المصحّحات في نظام التّحكّم واختباره بالمحاكاة مع نظام القيادة

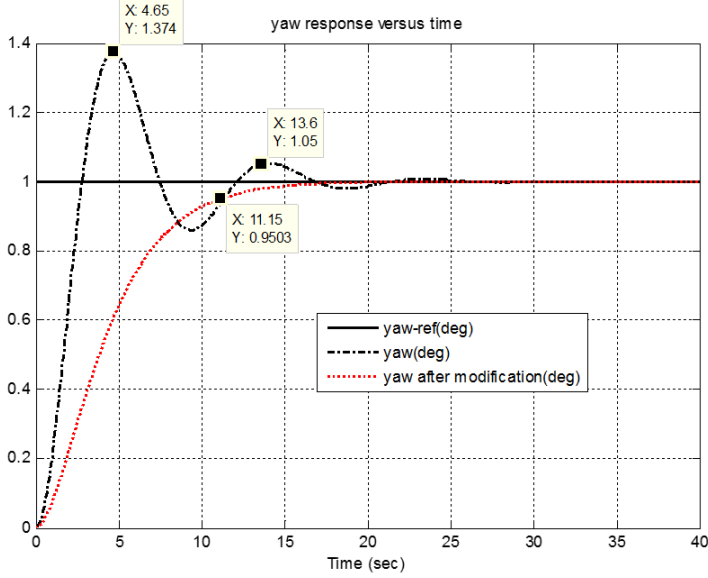
سنختبر نظامي القيادة والتّحكّم لملاحقة المسار في البيئة Matlab/Simulink باستخدام المكتبة (RVCtools: Robotics, Vision & Control Toolboxes) الإصدار 9.1. هذه المكتبة مبنية في جامعة Queensland University of Technology ابتداءً من العام 1992، وهي معتمدة لدراسة ومحاكاة الروبوتات في العديد من الجامعات والمراكز البحثية، حيث تؤمّن مجموعة كبيرة من التّوابع الهامة لدراسة ومحاكاة الروبوتات بنوعها التقليديّة (ذات الأذرع) والمتحرّكة، كما تحوي المكتبة على نموذج مفصلٍ لطائرة رباعيّة المرواح (الطائرة X-4 Flyer II) مع نظام تحكّم للوضع الزاويّ والتوجّه والموضع، وهي مكتبة مجانيةّ متاحة على الموقع : <http://www.petercorke.com/robot>.

النّموذج مبيّن في الشّكل (8)، وهو مبنيّ باستخدام S-Function، دخل النّموذج هو شعاع الأوامر $\mathbf{u} = [\omega_1, \omega_2, \omega_3, \omega_4]^T$ وهو عبارة أوامر سرعة دوران المحرّكات مقدّرة بـ rad/sec، خرج النّموذج هو الشعاع $\mathbf{X} = [x, y, z, \psi, \theta, \phi, u, v, w, \dot{\psi}, \dot{\theta}, \dot{\phi}]^T$ ليصبح $\mathbf{X} = [x, y, z, \psi, \theta, \phi, \dot{x}, \dot{y}, \dot{z}, \dot{\psi}, \dot{\theta}, \dot{\phi}, u, v, w]^T$ ، حيث:



الشكل (8): الطائرة المُنمذجة مع نظام التَّحكُّم في المكتبة RVCtools بعد حذف متحكِّم الموضع وإضافة نظامي القيادة والتَّحكُّم لملاحقة مسارٍ.

الجملة المرجعية إلى الجملة المرتبطة بجسم الطائرة، $\mathbf{V}_b(t) = (u, v, w)^T$ سرعة مركز ثقل الطائرة بالنسبة للجملة المرجعية بدلالة الجملة المرتبطة بالجسم. كما تم تعديل معاملات حلقة التحكم بالتوجه لتصبح دون تجاوز كما يبين الشكل (9)، وأصبح زمن الاستجابة 11.15 sec بعد أن كان 13.6 sec.



الشكل (9): استجابة حلقة التحكم بتوجه الطائرة في المكتبة RVCtools قبل وبعد التعديل. أخيراً جرى حذف متحكم الموضع من المكتبة، وأضيف نظاما القيادة والتحكم لملاحقة مسار (الشكل (8)).

إن زمن الاستجابة لحلقة التحكم بالموضع الزاوي $t_{r,attitude} = 0.862 \text{ sec}$ ، لنضع

$$t_{r,i} = t_{r,d} = 3 \text{ sec} > 3 * t_{r,attitude} = 2.586 \text{ sec}$$

$$(16) \quad \begin{cases} H_\ell = \frac{\dot{e}_\ell}{\dot{e}_{\ell,dmd}} = \frac{c_{2,\ell} \cdot (-g/s)}{1 + c_{2,\ell} \cdot (-g/s)} = \frac{1}{T_\ell \cdot s + 1} \\ H_d = \frac{\dot{e}_d}{\dot{e}_{d,dmd}} = \frac{c_{2,d} \cdot (g/s)}{1 + c_{2,d} \cdot (g/s)} = \frac{1}{T_d \cdot s + 1} \end{cases}$$

تابعاً تحويلاً من المرتبة الأولى والثابتاهما الزمانيان:

$$(17) \quad \begin{cases} T_\ell = \frac{1}{-c_{2,\ell} \cdot g} = \frac{t_{r,\ell}}{3} = 1 \text{ sec} \Rightarrow c_{2,\ell} = -\frac{1}{g} \approx -0.1 \\ T_d = \frac{1}{c_{2,d} \cdot g} = \frac{t_{r,d}}{3} = 1 \text{ sec} \Rightarrow c_{2,d} = \frac{1}{g} \approx 0.1 \end{cases}$$

حيث زمن الاستجابة لتابع التحويل من المرتبة الأولى يساوي ثلاثة أضعاف قيمة الثابت الزمني.

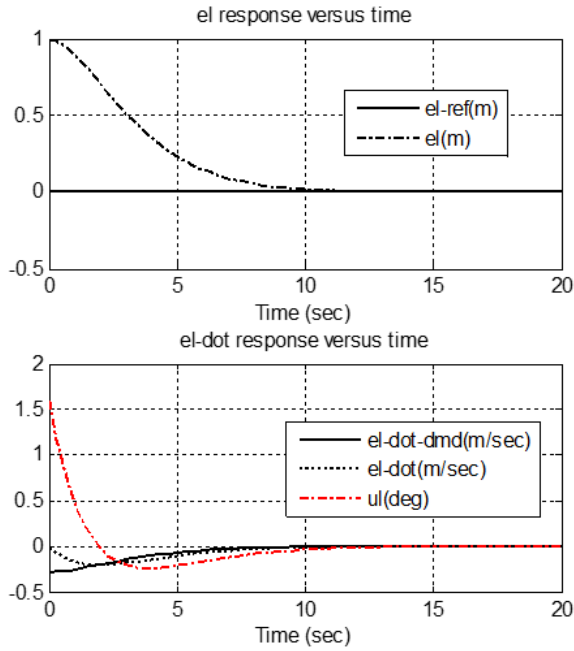
تابعاً تحويل حقتي التحكم الكليتين لإلغاء الخطأ العرضي والمسافة المتبقية (بإهمال المحدد على $\dot{e}_{\ell,dmd}$ و $\dot{e}_{d,dmd}$) هما:

$$(18) \quad \begin{cases} H_\ell = \frac{e_\ell}{e_{\ell,ref}} = \frac{c_{1,\ell} \cdot H_\ell / s}{1 + c_{1,\ell} \cdot H_\ell / s} = \frac{c_{1,\ell}}{s^2 + s + c_{1,\ell}} = \frac{\omega_n^2}{s^2 + 2 \cdot \xi \cdot \omega_n \cdot s + \omega_n^2} \\ H_d = \frac{e_d}{e_{d,ref}} = \frac{c_{1,d} \cdot H_d / s}{1 + c_{1,d} \cdot H_d / s} = \frac{c_{1,d}}{s^2 + s + c_{1,d}} = \frac{\omega_n^2}{s^2 + 2 \cdot \xi \cdot \omega_n \cdot s + \omega_n^2} \end{cases}$$

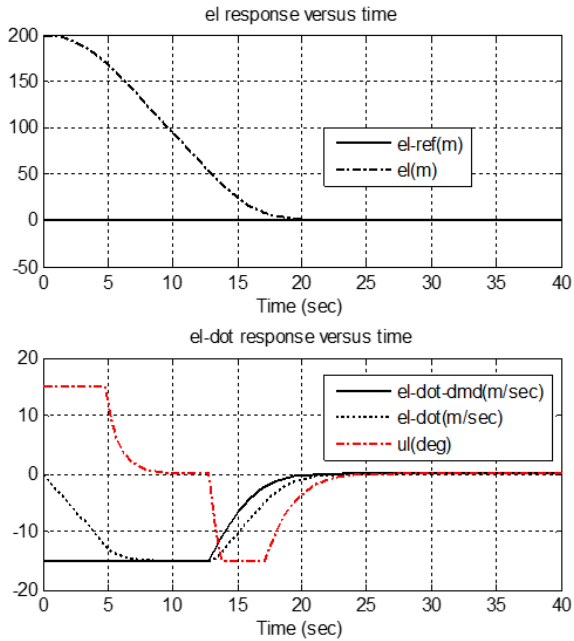
وهي استجابة نظام من المرتبة الثانية نبضه الذاتي ω_n ومعامل التخميد له ξ ، لدينا:

$$(19) \quad 2 \cdot \xi \cdot \omega_n = 1 \Rightarrow \omega_n = \frac{1}{2 \xi}, \quad c_{1,\ell} = c_{1,d} = \omega_n^2$$

من أجل نظم المرتبة الثانية لتحقيق نسبة تجاوز $D\% = 0.01$ (التجاوز شبه معدوم) يجب أن يكون $\xi = 0.95$ ، وبالتالي $\omega_n = 0.53 \text{ rad} \cdot \text{sec}^{-1}$ ، $c_{1,\ell} = c_{1,d} = 0.28$.



الشكل (10): استجابة الحلقة H_l من أجل $e_l(0) = 1m$.

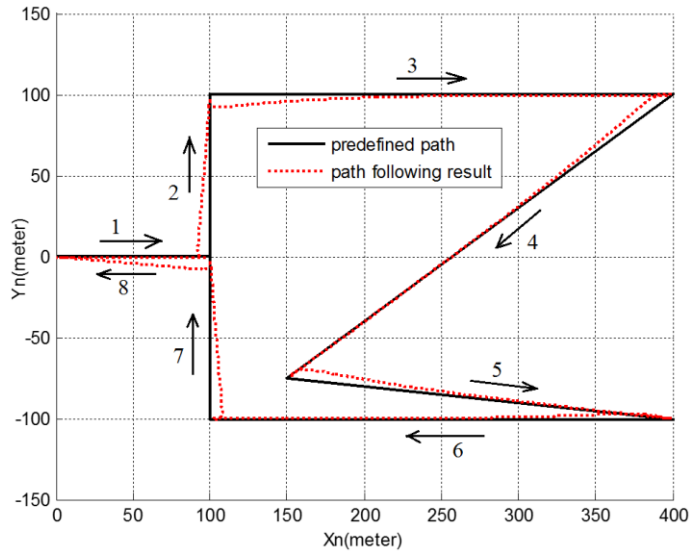


الشكل (11): استجابة الحلقة H_l من أجل $e_l(0) = 200m$.

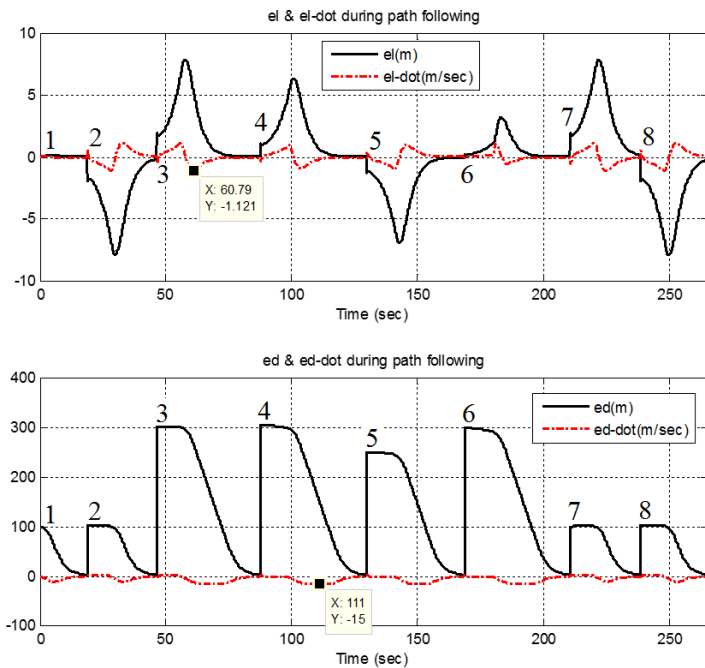
نقوم بإجراء محاكاة في البيئة Matlab/Simulink لحلقة التحكّم H_ℓ ، من أجل قيمة ابتدائية لخطأ الملاحقة العرضي $e_\ell(0) = 1m$ نلاحظ الحصول على استجابة تحقّق الشّروط على أداء متحكّم ملاحقة المسار، حيث أنّ قيم كلّ من e_ℓ و u_ℓ تقع ضمن مجال التّحديد لكُلّ منهما، ولا يوجد تجاوز في استجابة H_ℓ كما يُبيّن الشكل (10).

من أجل $e_\ell(0) = 200m$ (قيمة ابتدائية كبيرة لخطأ الملاحقة العرضي تحقّق استخدام قيما لإشباع للسرعة) نلاحظ أيضاً الحصول على استجابة تحقّق الشّروط على أداء متحكّم ملاحقة المسار، كما نلاحظ أنّ كلاً من u_ℓ و $e_{\ell,dmd}$ قد وصلنا إلى حدود الإشباع، وأستفيد من سرعة الطّائرة الأعظميّة، الاستجابة موضّحة في الشكل (11)، ونحصل على نتائج مماثلة من أجل حلقة التحكّم لإلغاء المسافة المتبقية.

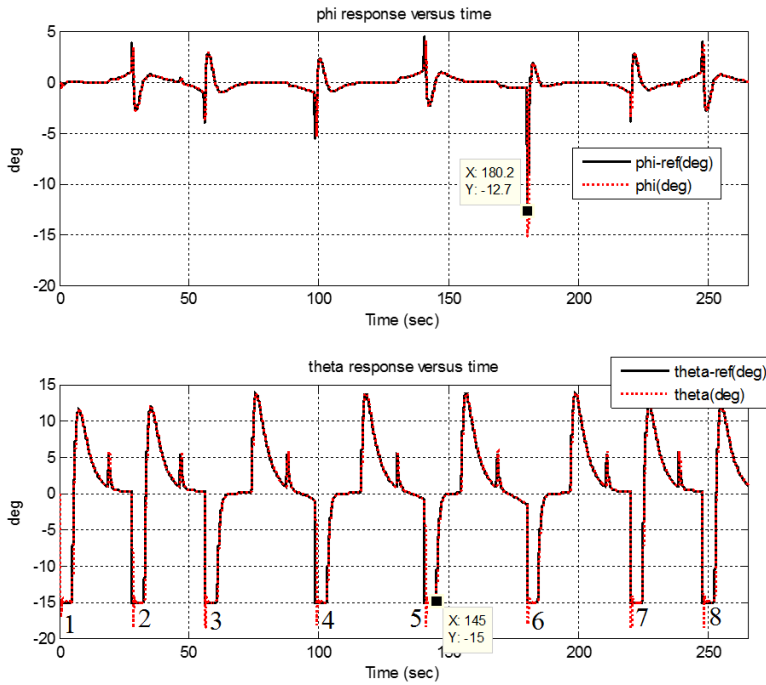
بعد ضبط قيم معاملات حلقة التحكّم لإلغاء الخطأ العرضي والمسافة المتبقية، تمّت كتابة خوارزمية القيادة والتحكّم لملاحقة مسار بلغة C في البيئة Matlab/Simulink باستخدام تابعي S-Function، وأضيفت الكتلتان إلى الملف الخاص بالطائرة رباعية المرواح في المكتبة RVCtools بعد حذف متحكّم الموضع من هذا الملف. دخل نظام القيادة هو نقط مسار $P_i(x_i, y_i)$ ، أمّا خرجة فهو $(e_\ell, \dot{e}_\ell, e_d, \dot{e}_d, \psi_{ref})$. تمّ أثناء الاختبار بواسطة المحاكاة افتراض نقط المسار التالية: $P_1(0,0)$, $P_2(100,0)$, $P_3(100,100)$, $P_4(400,100)$, $P_5(150,-75)$, $P_6(400,-100)$, $P_7(100,-100)$, $P_8(100,0)$, $P_9(0,0)$ ، وأعتد شرط الوصول $|e_\ell(t)| \leq 2m$ و $|\dot{e}_d(t)| \leq 2m$.



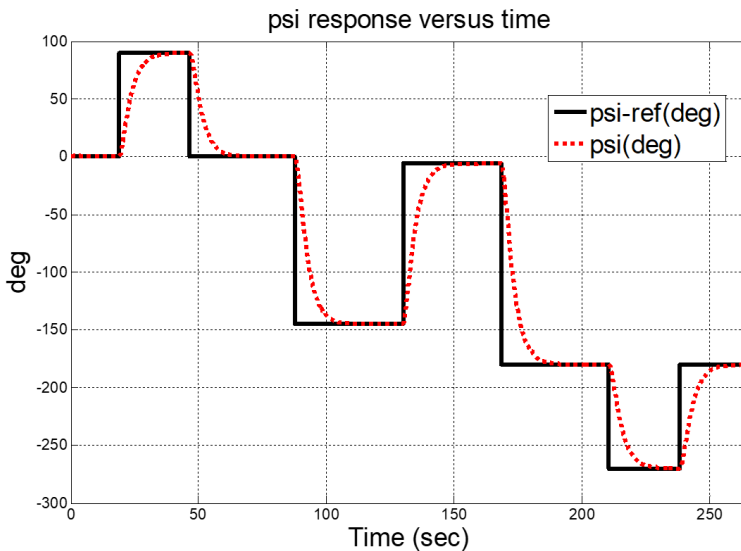
الشكل (12): المسار الاختباري المُفترض المطلوب من الطائرة ملاحقته مع نتيجة التنفيذ.



الشكل (13): الخطأ العرضي والمسافة المتبقية وسرعة حركة الطائرة أثناء تنفيذ المسار.



الشكل (14): أمر الوضع الزاوي الناتج عن متحكم ملاحقة المسار مع استجابة الطائرة على كلٍّ منهما.



الشكل (15): أمر زاوية التوجّه الناتج عن نظام القيادة مع استجابة الطائرة عليه.

نجد من الشكل (12) أنّ الطائرة تمكّنت من ملاحقة المسار المُفترض بعد تزويدها بنظام القيادة والتحكّم المُصمّمين، ونجد من الشكلين (13) و(14) أنّ إلغاء الخطأ العرضيّ والمسافة المتبقّية يكون عند كلّ قطعة مسارٍ بدون تجاوزٍ، ويتمّ الانتقال إلى قطعة المسار التّالية فور إلغائها، كما نجد أنّ أوامر الوضع الزّاويّ ضمن المجال المطلوب $\pm 15^\circ$ ، وسرعة الطائرة خلال تنفيذ المسار ضمن المجال المطلوب $\pm 15m.sec^{-1}$ ، ويبيّن الشكل (15) أنّ الطائرة تقوم بتعديل توجّهها أثناء ملاحقة المسار ليصبح موافقاً لقطعة المسار التي تسير وفقها.

10- الخاتمة والآفاق المستقبلية

تمّ في هذا البحث تصميم نظام قيادة وتحكّم لتمكين طائرة رباعيّة المراوح مزوّدة بنظام استقرارٍ من ملاحقة مسارٍ مكوّنٍ من نقطٍ مسارٍ والقطع المستقيمة الواصلة بينها، وذلك لجعل الطائرة مستقلةً عن المُشغّل (Autonomous) بهدف تسهيل أعمال الاستطلاع والمراقبة، وتمّ اختبار النظام المُصمّم بواسطة المحاكاة باستخدام مكتبةٍ معتمدةٍ لدراسة ومحاكاة الروبوتات (RVCtools)، تحوي هذه المكتبة نموذجاً لاخطياً لطائرة رباعيّة المراوح مع نظام تحكّم للوضع الزاويّ والتوجّه والارتفاع والموضع، وقد بيّنت نتائج المحاكاة فعالية نظام القيادة والتحكّم. في الخطوة التّالية من هذا البحث سيجري التطبيق العملي لنظام القيادة والتحكّم المُطوّر على طائرٍ حقيقيّة، كما سيتمّ تزويد الطائرة بخوارزميةٍ مناسبةٍ (كخوارزمية الحقل الكموني الصناعي Artificial Potential Field) لتمكينها من تجنّب العوائق غير المعروفة وذلك بعد تزويدها بالحساسات المناسبة.

-11 المراجع

- [1] <https://code.google.com/p/arducopter/wiki/PX4RC>
- [2] FOSSEN T, 2011– Handbook of Marine Craft Hydrodynamics and Motion Control. John Wiley & Sons Ltd, First Edition, 575p.
- [3] CORKE P, 2013 – Robotics, Vision and Control. Fundamental Algorithms in MATLAB®. Springer, Corrected Second Printing, Berlin Heidelberg, 570p.
- [4] BOUABDALLAH S, 2007 – Design and Control of Quadrotors with Application to Autonomous Flying. PhD Thesis, École Polytechnique Fédérale de Lausanne.
- [5] POUNDS P, 2007 – Design, Construction and Control of a Large Quadrotor Micro Air Vehicle. PhD Thesis, The Australian National University.
- [6] HOFFMANN G, 2008 – Autonomy for Sensor-Rich Vehicles: Interaction Between Sensing and Control Actions. PhD Thesis, Stanford University.

- [7] MORBIDI F, CANO R, LARA D, 2016– Minimum–Energy Path Generation for a Quadrotor UAV. IEEE International Conference on Robotics and Automation, Sweden, Stockholm.
- [8] MANJUNATH A, 2016 – Path Following by a Quadrotor Using Virtual Target Pursuit Guidance. Master Thesis, Utah State University.
- [9] MEHROK P, 2016 – Quadrotor UAV Path Following Using Trajectory Shaping, Master Thesis, Utah State University.
- [10] LIU Z, CIARLETTA L, YUAN C, ZHANG Y, and D. THEILLIOL, Jun 2017 – Path following control of unmanned quadrotor helicopter with obstacle avoidance capability, International Conference on Unmanned Aircraft Systems, ICUAS'17, United States, Florida, Miami.
- [11] STEVŠIĆ S, NÄGELI T, ALONSO–MORA J, and HILLIGES O, October 2018 – Sample efficient learning of path following and obstacle avoidance behavior for Quadrotors. IEEE Robotics and Automation Letters, Vol. 3, No. 4, pp. 3852–3859.
- [12] DO T, CARRILLO–ARCE L, and S. ROUMELIOTIS, 2019 – High–speed autonomous quadrotor navigation through visual

- and inertial paths.The International Journal of Robotics Research, Vol. 38, No. 4,pp. 486–504.
- [13] DING Y, WANG X, BAI Y, and CUI N,2020 –High–precision prescribed–time path following for quadrotor.International journal of Advanced Robotic Systems, Vol. 17, No. 2.
- [14] BOUZID Y, BESTAOUY Y, SIGUERDIDJANE H, and ZAREB M, Jun 2018 – Quadrotor Guidance–Control for flight like nonholonomic vehicle system.International Conference on Unmanned Aircraft Systems, ICUAS’18, United States, Dallas.
- [15] HONG Y, KIM S, KIM Y, andCHA J,2021 – Quadrotor path planning using A* search algorithm and minimum snap trajectory generation,ETRI Journal, Vol. 43, Issue6,pp. 1013–1023.
- [16] LIN H, and PENG X, 2021 – Autonomous Quadrotor Navigation with Vision Based Obstacle Avoidance and Path Planning,IEEE Access, Vol. 9, pp. 102450–102459.

- [17] FOEHN P, ROMERO A, and SCARAMUZZA D, 2021 – Time-Optimal Planning for Quadrotor Waypoint Flight, Science Robotics, Vol.6, Issue 56, pp. eabh1221.
- [18] KHAKBAZ M, and ANJIDAN M, 2022 – Global Path Planning of Quadrotor Using Reinforcement Learning, Journal of Computer & Robotics, Vol. 15, Issue 1, pp. 65–79.
- [19] MOKRANE A, BENALLEGUE A, CHOUKCHOU-BRAHAM A, EI-HADRI A, and CHERKI B, 2022 – Guidance, Navigation and Control for Autonomous Quadrotor Flight in an Agricultural Field: The Case of Vineyards, Journal of Sensors, Vol. 22, No. 22.
- [20] REN Y, ZHU F, LIU W, WANG Z, LIN Y, GAO F, and ZHANG F, 2022 – Bubble Planner: Planning High-speed Smooth Quadrotor Trajectories using Receding Corridors. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Japan, Kyoto, pp. 6332–6339.
- [21] <https://en.wikipedia.org/wiki/Atan2/>
- [22] MARLIN T, 2015 – Process Control, Designing Processes and Control Systems for Dynamic Performance. Published by

McGraw–Hill from 1995–2014, The copyright was returned to the author after publication ceased, 2nd Edition, 1017p.

[23] KIVRAK A,2006 – Design of Control Systems for a Quadrotor Flight Vehicle Equipped with Inertial Sensors,Master Thesis, Atilim University.

[24] HOFFMANN G, HUANG H, WASLANDER S, and TOMLIN C,2011 –Precision Flight Control for a Multi–Vehicle Quadrotor Helicopter Testbed.Control Engineering Practice, Vol. 19, Issue9,pp. 1023–1036.

التحديد الكمي لبينوكسينات هيدروكلورايد في المستحضرات الصيدلانية بالطريقة الطيفية الضوئية الحركية

الدكتور محمد خطيب - أستاذ مساعد في الكيمياء التحليلية

قسم العلوم الأساسية ، كلية الهندسة الميكانيكية والكهربائية ، جامعة البعث

الملخص

تم تطوير طريقة طيفية ضوئية حركية جديدة لتحديد بينوكسينات هيدروكلورايد Oxybuprocaine Hydrochloride (BNX) والذي يسمى أيضاً Oxybuprocaine Hydrochloride في الحالة النقية وفي المستحضرات الصيدلانية. تعتمد الطريقة على أكسدة BNX في وسط من حمض الكبريت بأيونات الحديد الثلاثي والتي بدورها تُرجع إلى أيونات الحديد الثنائي التي تتفاعل لاحقاً مع فري سيانيد البوتاسيوم لتشكيل ناتج أخضر مزرق منحل في الماء يبدي قمة امتصاص عظمى عند طول الموجة 765nm مقابل المحلول المقارن عند درجة حرارة المختبر $25 \pm 0.5^\circ\text{C}$ ، تمت دراسة التفاعل حركياً وذلك بحساب التزايد في الامتصاصية كتابع للزمن عند 765nm. رسمنا الخط البياني العياري وذلك اعتماداً على طريقة الزمن الثابت (15min). تحقق قانون بير ضمن مجال خطي من تركيز BNX يتراوح ما بين $5.0-25.0 \mu\text{g mL}^{-1}$. كانت قيمة معامل الارتباط 0.9984 وقيمة حد الكشف $0.227 \mu\text{g mL}^{-1}$ بالنسبة للطريقة المقترحة. اعتمدنا النتائج التحليلية بعد معالجتها إحصائياً. طبقنا الطريقة المقترحة بنجاح من أجل تحديد BNX في المستحضرات الصيدلانية دون حصول إعاقة من المواد المضافة.

الكلمات المفتاحية: بينوكسينات هيدروكلورايد، التحليل الطيفي الضوئي الحركي ، فري سيانيد البوتاسيوم، المستحضرات الصيدلانية.

Quantitative Determination of Benoxinate Hydrochloride in Pharmaceutical Preparations by Kinetic Spectrophotometric Method

Mouhammed Khateeb – Professor Assistant

Department of Basic Science, Faculty of Mechanical and Electrical Engineering, University of Albaath

Abstract

Simple, accurate and reliable kinetic spectrophotometric method for the quantitative determination of Benoxinate Hydrochloride (BNX) in pure and pharmaceutical formulations has been developed. The method is based on the oxidation of BNX by use Fe(III) in sulfuric acid medium. Fe(III) subsequently reduces to Fe(II) which is coupled with potassium ferricyanide to form bluish green product. The reaction is followed spectrophotometrically by measuring the increase in absorbance at λ_{\max} 765 nm. The fixed time method (15 min) was adopted for constructing the calibration curve. The linearity range was found to be 5.0-25.0 $\mu\text{g mL}^{-1}$. The correlation coefficient and detection limit for fixed time method were 0.9984 and 0.227 $\mu\text{g mL}^{-1}$ respectively. The proposed method has been successfully applied to the determination of BNX in pharmaceutical dosage forms with no interference from the additives.

Keywords: Benoxinate Hydrochloride, kinetic spectrophotometry, potassium ferricyanide, pharmaceutical formulations.

المقدمة

بينوكسينات هيدروكلورايد هو عبارة عن مسحوق بلوري أبيض أو بلورات عديمة اللون، ينحل في الماء والكحول، تتراوح قيمة pH محلول منه 10% ما بين 4.5-6. اسمه العلمي: 2-Diethylaminoethyl 4-amino-3-butoxybenzoate hydrochloride. صيغته المجرى: $(C_{17}H_{28}N_2O_3, HCl)$ (M = 344.9 g mol⁻¹). يستخدم كمخدر موضعي في طب العيون والأنف والأذن والحنجرة، تتمثل آلية عمله بارتباطه بقنوات الصوديوم الموجودة في المحاور العصبية مما يؤدي إلى استقرار الغشاء العصبي ومنع توصيل السيالات العصبية. كما يستخدم في العين كمخدر موضعي لأغراض تشخيصية أو علاجية مثل إجراء عمليات صغرى، إزالة الأجسام الغريبة من الملتحمة والقرنية، وقياس ضغط العين وفحصها، وتركيب العدسات اللاصقة [1].

أجريت دراسات عديدة لتحديد بينوكسينات BNX في الحالة النقية وفي المستحضرات الصيدلانية باستخدام طرائق تحليلية مختلفة.

حيث قُدمت دراسة طيفية لتحديد BNX باستخدام طريقة المشتق الأول عند طولي الموجة (268.4; 272.4)nm باستخدام مذيب من فوسفات حمض السيتريك عند pH=7، كما استخدمت طريقة الكروماتوغرافيا السائلة لتحديد BNX باستخدام كاشف طيفي عند طول الموجة 310nm [2]، كما طورت طريقة طيفية لتحديد BNX في المستحضرات الصيدلانية من خلال التفاعل مع الكاشفين التاليين

4-dimethylaminocinnamaldehyde ،4-dimethylaminobenzaldehyde

قيست الامتصاصية لنواتج التفاعل عند الأطوال الموجية 453nm و 549nm على الترتيب. [3]، طبقت طريقة طيفية لتحديد BNX بشكله النقي وفي المستحضرات الصيدلانية من خلال تفاعل انتقال الشحنة بين BNX كمناح n-electron و

p-chloranilic acid, 2, 3 – dichloro 5, 6 – dicyano 1, 4 benzoquinone واليود كمستقبل σ, π [4]، حدد BNX و الميتوكلوبراميد بالطريقة الطيفية من خلال تشكيل معقدات الزوج الأيوني مما بين المادتين الدوائيتين وثيوسيانات Mo(V), Co(II) [5]، طورت طريقة طيفية تفلورية لتقدير BNX في قطرات العين والخلط المائي الاصطناعي. من خلال تفاعل الفلورسكامين مع المجموعة الأمينية الأولية لـ BNX بعد إثارة منتج التفاعل عند 393nm، قيست شدة الإصدار عند طول الموجة 483nm [6]، كما استخدمت تقانة الفلورة لتحديد BNX في القطرات العينية باستخدام مذيب الايتانول حيث أظهرت قمة اثاره عظمى عند طول الموجة 297nm وقمة اصدار عظمى عند طول الموجة 371nm [7]، تمت دراسة الأكسدة الكهروكيميائية لـ BNX في مستحضراته الصيدلانية باستخدام كهليلت داعم من واقي بربتون على مسرى الكربون الزجاجي المعالج كهروكيميائياً (EA-GCE) وفق الطريقة النبضية التفاضلية [8]، كما حُضرت أربع مسار غشائية مختلفة مصنوعة من PVC المعدل بمعقدات الزوج الأيوني باستخدام الطريقة الكمونية لتحديد BNX في محاليله النقية وفي القطرات العينية [9]، تمت دراسة السلوك الحراري لـ BNX في مستحضراته الصيدلانية. باستخدام تقنيات تحليل قياس الوزن الحراري، وقياس الوزن الحراري المشتق وتقنيات التحليل الحراري التفاضلي [10]، حضرت مسار غشائية انتقائية لمادتي بوبيفاكايبين و BNX حيث أبدت المساري المحضرة استجابة لمعادلة نرنست بعمر عشرة أيام للمسرى، كما استخدمت المساري المحضرة في المعايير الكمونية [11]، استخدمت طريقة الكروماتوغرافيا السائلة عالية الأداء HPLC لتحديد كل من الفلوريسين و BNX باستخدام عمود C_{18}

(100 mm × 4.6 mm × 3.5 μ m) وكاشف المصفوفة الضوئية [12]، حدد كل من الفلوريسين و BNX في مزيج منهما باستخدام تقانة الكروماتوغرافيا السائلة عالية الأداء معكوسة الأطوار باستخدام طور متحرك من 40% أسيتونتريل و 60% فوسفات أحادية

البوتاسيوم 50mM وعمود ODS-3 [13]، حدد كل من الديكلوفيناك و BNX في سوائل الحسم الحيوية بتقنية RP-HPLC باستخدام عمود من Regis SPS 100 RP-8 وكاشف كهركيميائي [14]، حدد BNX في مستحضراته الصيدلانية بتقنية HPLC باستخدام عمود من Spherisorb phenyl وطور متحرك مؤلف من أسيتونتريل- وافي pH(3.5) بنسبة (35v/v : 65) وكاسف طيفي عند طول الموجة 308nm [15].

أهمية البحث وأهدافه:

تكمّن أهمية هذا البحث باعتبار أنه حتى الآن الطرائق المنشورة لتحديد بينوكسينات هيدروكلورايد هي طرائق قليلة ومحدودة ، ولم تنتشر أي طريقة لتحليل بينوكسينات هيدروكلورايد في المستحضرات الدوائية بطريقة التحليل الطيفي الضوئي الحركي. يهدف هذا البحث إلى تطوير طريقة جديدة من أجل تحديد بينوكسينات هيدروكلورايد كمادة عيارية باستخدام التحليل الطيفي الضوئي الحركي ومن ثم اختبار صلاحية هذه الطريقة لاستخدامها في الرقابة الدوائية اليومية من أجل تحليل المستحضرات الدوائية التي تحوي بينوكسينات هيدروكلورايد.

القسم العملي

الأجهزة

استخدمنا جهاز طيفي ضوئي مرئي وما فوق بنفسجي Optima ياباني الصنع طراز SP-3000 plus مزود بخلايا كوارتز 1cm . يسمح الطيف آلياً بوساطة برمجيات خاصة بالجهاز، اخترنا المجال 500-1100nm لإجراء المسح الطيفي.

المواد والمحاليل

جميع المواد المستخدمة هي من النوع النقي تحليلياً. والماء المستخدم ثنائي التقطير. حضرنا محلول فري سيانيد البوتاسيوم $K_3[Fe(CN)_6]$ إنتاج شركة (Sigma-Aldrich) بتركيز $2.5 \times 10^{-3} M$ في الماء. حضرنا محلول كبريتات الأمونيوم والحديد الثلاثي $2.3.0 \times 10^{-3} M$ إنتاج شركة (Sigma-Aldrich) بتركيز $3.0 \times 10^{-3} M$ في حمض الكبريت تركيزه $3.0 \times 10^{-3} M$. استخدمنا المركب الدوائي بينوكسينات هيدروكلورايد إنتاج شركة Shandong Haiyou Freda الهولندية كمادة عيارية $C_{17}H_{28}N_2O_3.HCl = 344.9 g mol^{-1}$ ، نقاوته هي 99.68%. حضرنا المحلول العياري الأصلي بتركيز $0.5 mg mL^{-1}$ بحل الكمية المناسبة من BNX في الماء. تم تحليل مستحضرات صيدلانية (من إنتاج شركات محلية) تحوي BNX وهو قطرة ميديكائين من إنتاج شركة ميديكو للصناعات الدوائية وقطرة دلتاكائين من إنتاج شركة دلتا لصناعة الأدوية وتركيز كل منهما 0.4% ($4 mg/1 mL$).

طريقة العمل من أجل المركب الدوائي (طريقة الزمن الثابت)

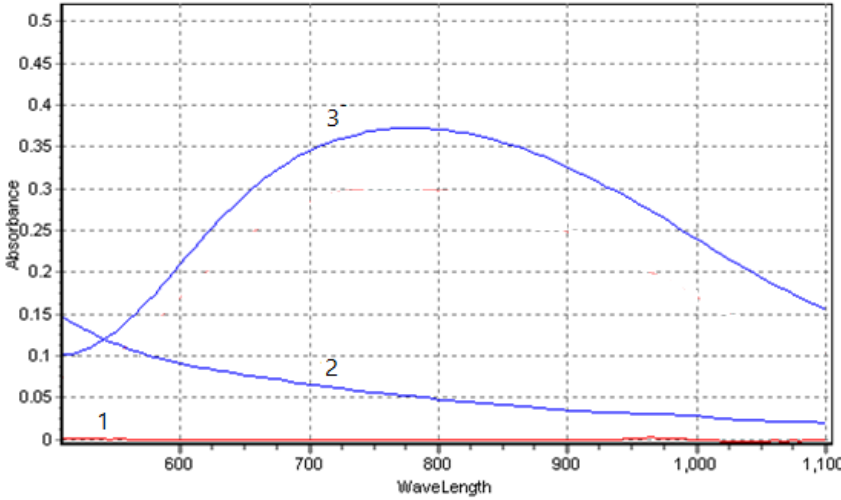
وضعنا في سلسلة دوارق حجمية سعة 10mL حجوماً مناسبة من محلول BNX العياري $0.5 mg mL^{-1}$ (0.10-0.50mL) و 2.0mL من محلول $K_3[Fe(CN)_6]$ و 1.5mL من محلول $NH_4 Fe(SO_4)_2 \cdot 12H_2O$ ، أكملنا الحجم حتى إشارة الدورق بالماء المقطر، ومزجنا جيداً. بعد 15 دقيقة قسنا امتصاصية المحلول الملون الناتج عند 765nm مقابل محلول مقارن محضر بالطريقة نفسها بدون المركب الدوائي. رسمنا بيانياً الخط البياني العياري، الممثل لتغيرات الامتصاصية بدلالة تركيز BNX.

طريقة العمل من أجل المستحضر الصيدلاني

أخذنا من القطرة العينية حجم مقداره 1.25mL في دورق حجمي سعة 10mL، مددنا هذا الحجم حتى إشارة الدورق بالماء ثنائي التقطير، وبالتالي نكون قد حصلنا على محلول تركيز BNX فيه هو 0.5mg mL^{-1} تقريباً. أخذنا حجوماً مناسبة من هذا المحلول لتحديد BNX كما ورد سابقاً.

النتائج والمناقشة

تتم أكسدة BNX في وسط من حمض الكبريت بأيونات الحديد الثلاثي والتي بدورها تُرجع إلى أيونات الحديد الثنائي التي تتفاعل لاحقاً مع فري سيانيد البوتاسيوم لتشكيل ناتج أخضر مزرق منحل في الماء يبدي قمة امتصاص عظمى عند طول الموجة 765nm مقابل المحلول المقارن عند درجة حرارة المختبر $25\pm 0.5^\circ\text{C}$ ، كما هو مبين في الشكل رقم (1). لوحظ أن شدة الامتصاصية تتناسب طردياً مع تركيز BNX.

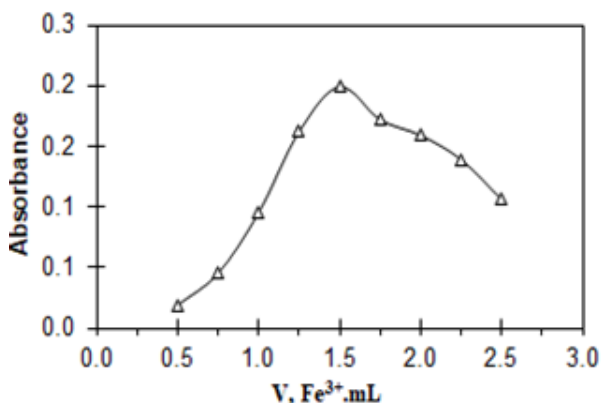


الشكل رقم (1) طيوف امتصاص (1) محلول BNX تركيزه $10\mu\text{g mL}^{-1}$ مقابل الماء، و (2) طيف امتصاص المحلول المقارن مقابل الماء، و (3) طيف امتصاص المحلول الملون مقابل المحلول المقارن، حيث أن: $[\text{BNX}] = 10\mu\text{g mL}^{-1} + 1.5\text{mL Fe}^{3+}$ ($3.0 \times 10^{-3}\text{M}$)، $2.0\text{mL K}_3[\text{Fe}(\text{CN})_6]$ ($2.5 \times 10^{-3}\text{M}$) في حجم نهائي مقداره 10mL. بعد 15 دقيقة.

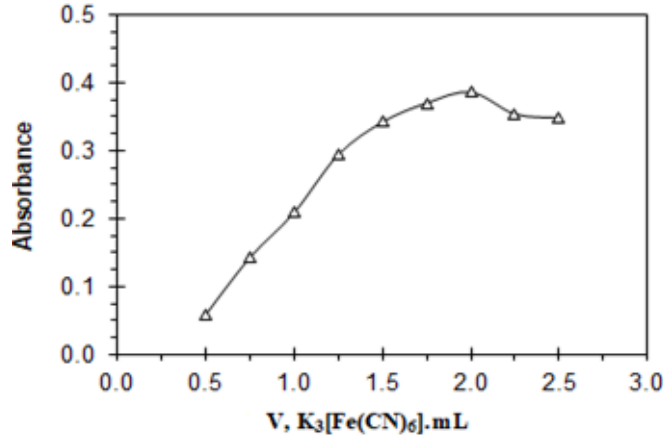
ضبط الشروط المثلى للطريقة

أجرينا عدداً من الدراسات التمهيديّة للوصول إلى الشروط المثلى، لإجراء التفاعل السريع والكمّي بحساسية قصوى، وذلك بتغيير أحد القرائن في كلّ مرة بينما نبقي القرائن الأخرى ثابتة ونلاحظ تأثيره في الامتصاصية عند طول موجة الامتصاص الأعظمي.

درسنا تأثير تركيز Fe^{3+} في تشكيل ناتج التفاعل. تبين أن إضافة 1.5mL من كبريتات الأمونيوم والحديد الثلاثي بتركيز $3.0 \times 10^{-3} M$ هو الحجم الأمثل. الشكل رقم (2). كذلك درسنا تأثير تركيز $K_3[Fe(CN)_6]$. وجدنا أن إضافة 2.0mL منه تركيزه $2.5 \times 10^{-3} M$ هو الحجم الأمثل لتشكيل ناتج التفاعل، الشكل رقم (3).



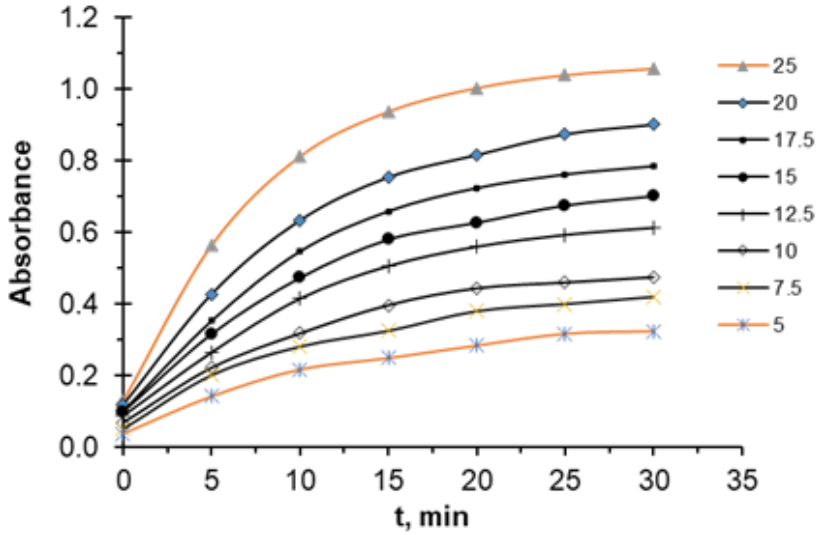
الشكل رقم (2) تأثير تغير تركيز Fe^{3+} ($3 \times 10^{-3} M$) بوجود $K_3 [Fe(CN)_6]$ ($2.5 \times 10^{-3} M$) 1.0mL عند $[BNX]=10.0 \mu g.mL^{-1}$, 765nm, (t=15min).



الشكل رقم (3) تأثير تغير تركيز $K_3[Fe(CN)_6]$ ($2.5 \times 10^{-3} M$) بوجود Fe^{3+} ($3 \times 10^{-3} M$) $1.5 mL$ عند $765 nm$, $[BNX] = 10.0 \mu g \cdot mL^{-1}$, $t = 15 min$.

تقييم الطرائق الحركية التحليلية الكمية

لاحظنا من خلال دراسة ناتج التفاعل الملون في الشروط المثلى، تزايداً في الامتصاصية عند $765 nm$ بمرور الزمن، كما هو موضح في الشكل رقم (4). وبناءً عليه درسنا إمكانية تقدير BNX كميًا بطريقة حركية من خلال تقييم طريقة معطيات السرعة للتفاعل وطريقة ثابت سرعة التفاعل وطريقة الامتصاصية الثابتة (الزمن المتغير) وطريقة الزمن الثابت [16]. ومن ثم اختيار الطريقة الأنسب لتقدير BNX اعتماداً على قيمة معامل الارتباط الخطي لكل طريقة.



الشكل رقم (4) منحنيات تغير امتصاصية المحلول الملون الناتج بدلالة الزمن بوجود 2mL من $5-25\mu\text{g mL}^{-1}$ [BNX] و $3.0 \times 10^{-3}\text{M Fe}^{3+}$ من 1.5mL و $2.5 \times 10^{-3}\text{M K}_3[\text{Fe}(\text{CN})_6]$

طريقة معطيات السرعة

يعبر عن سرعة التفاعل بالعلاقة التالية:

$$\text{Log } v = \text{Log } \Delta A / \Delta t = \text{Log } k + n \text{ Log } C$$

حيث أن: v سرعة التفاعل، ΔA التغير في قيمة الامتصاصية بين الزمن t_1 والزمن t_2 ، k ثابت سرعة التفاعل الافتراضي، C تركيز المادة الدوائية مقدراً بالجزيئي، n مرتبة التفاعل. حسبنا لوغاريتم سرعة التفاعل (حيث $t_1=5\text{min}$, $t_2=20\text{min}$) لتركيز مختلفة من BNX. رسمنا الخط البياني العياري، الممثل لتغيرات لوغاريتم سرعة التفاعل بدلالة لوغاريتم تركيز BNX. لاحظنا وجود ارتباط خطي بين لوغاريتم سرعة التفاعل ولوغاريتم تركيز BNX ضمن مجال من التراكيز ما بين $10.0 - 25.0\mu\text{g mL}^{-1}$ ($2.90 \times 10^{-5} - 7.25 \times 10^{-5}\text{M}$) كما حصلنا من الحاسوب مباشرة على قيمة معامل الارتباط ومعادلة المنحني العياري التالية:

$$\text{Log } \Delta A / \Delta t = 0.733 + 0.9063 \text{ Log } [\text{BNX}] \quad (r = 0.9962)$$

بالمقارنة مع العلاقة السابقة نجد أن قيمة ثابت السرعة $k=5.41S^{-1}$ ، والتفاعل من المرتبة الأولى ($n=0.9063\approx 1$) بالنسبة BNX.

طريقة ثابت سرعة التفاعل

رسمنا الخطوط البيانية لتغير لوغاريتم الامتصاصية بدلالة الزمن لتراكيز من BNX تقع ضمن المجال $5.0-10.0\mu g \text{ mL}^{-1}$ ($1.45\times 10^{-5}-2.90\times 10^{-5}M$). حصلنا من معادلات هذه الخطوط على قيمة ميل كل خط. حسبنا ثابت السرعة k لكل تركيز، وذلك بضرب قيمة الميل بقيمة الثابت 2.303 . رسمنا الخط البياني العياري لتغيرات ثابت سرعة التفاعل بدلالة تركيز BNX. وحصلنا على معادلة الخط البياني العياري التالية:

$$k = -0.001 - 5.96[BNX] \quad (r = 0.9904)$$

طريقة الامتصاصية الثابتة (الزمن المتغير):

ثبتنا الامتصاصية عند القيمة 0.4 ثم قسنا الزمن اللازم للوصول إلى هذه القيمة مقدراً بالثانية من خلال منحنيات تغير الامتصاصية بدلالة الزمن لتراكيز مختلفة من BNX، تتراوح ما بين $12.5-20.0mg.mL^{-1}$ ($3.62\times 10^{-5}-5.80\times 10^{-5}M$). مثلنا بيانياً المنحني العياري لتغيرات مقلوب الزمن بدلالة تركيز BNX. وحصلنا من الحاسوب مباشرة على قيمة معامل الارتباط ومعادلة المنحني العياري التالية:

$$1/t = -0.0022 + 99.252[BNX] \quad (r = 0.9927)$$

طريقة الزمن الثابت

مثلنا بيانياً الخطوط البيانية العيارية لتغيرات الامتصاصية (عند $765nm$) بدلالة تركيز BNX عند أزمنة محددة من $0-30$ دقيقة. حددنا المجال الخطي، وحصلنا على معادلة الخط البياني العياري ومعامل الارتباط وقيمة الميل والتقاطع عند كل زمن، وهذه القيم مبينة في الجدول رقم (1).

الجدول رقم (1) الخطوط البيانية العيارية لتحديد BNX عند أزمنة محددة

المجال الخطي $\mu\text{g.mL}^{-1}$	معامل الارتباط r	معادلة الخط البياني العياري	الزمن (دقيقة)
5.0-25.0	0.9573	$A = 0.0047C + 0.0209$	0
5.0-25.0	0.9774	$A = 0.0202C + 0.0273$	5
5.0-25.0	0.9908	$A = 0.0296C + 0.0463$	10
5.0-25.0	0.9984	$A = 0.0345C + 0.0663$	15
5.0-25.0	0.9983	$A = 0.0358C + 0.1019$	20
5.0-25.0	0.9967	$A = 0.0369C + 0.1206$	25
5.0-25.0	0.9959	$A = 0.0375C + 0.1323$	30

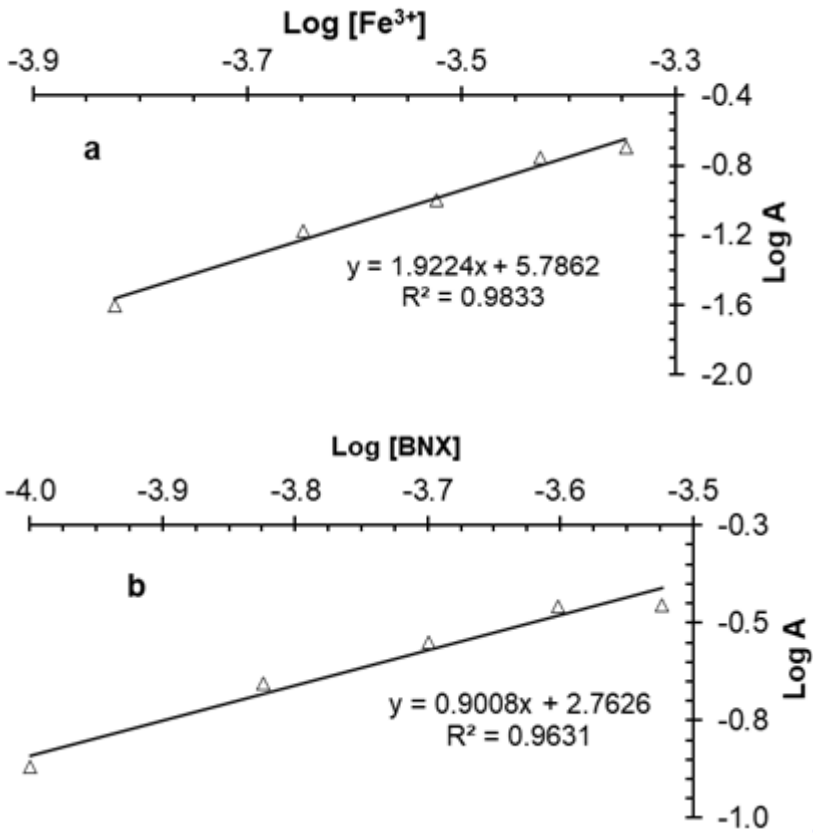
A، الامتصاصية؛ C، التركيز

نستنتج من الدراسة السابقة أن طريقة الزمن الثابت ($t=15\text{min}$) هي الأمثل لتحديد BNX، كونها أبدت أفضل قيمة لمعامل الارتباط (0.9984) بالمقارنة مع معاملات الارتباط للطرائق الأخرى، علاوة على المجال الخطي التحليلي الجيد الذي تمتعت به الطريقة، وتفاوت بين $5.0-25.0\mu\text{g mL}^{-1}$.

تحديد نسبة التفاعل

حددنا نسبة التفاعل بين BNX و Fe^{3+} بطريقة الحد اللوغاريتمي للمجال الخطي [17]، ثبتنا تركيز BNX وأضفنا تركيزاً متزايداً من Fe^{3+} . مثلنا بيانياً الخط البياني لتغيرات لوغاريتم الامتصاصية بدلالة لوغاريتم تركيز Fe^{3+} ، الشكل [5,a]، ثم ثبتنا تركيز Fe^{3+} وأضفنا تركيزاً متزايداً من BNX. مثلنا بيانياً الخط البياني لتغيرات لوغاريتم الامتصاصية بدلالة لوغاريتم تركيز BNX، الشكل [5,b].

نسبنا ميل الخط الأول 1.9224 إلى ميل الخط الثاني 0.9008، فاستنتجنا أن نسبة التفاعل بين Fe^{3+} و BNX هي 1:2.

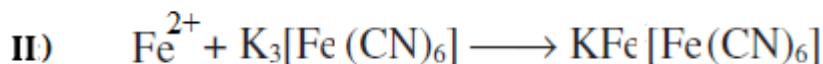
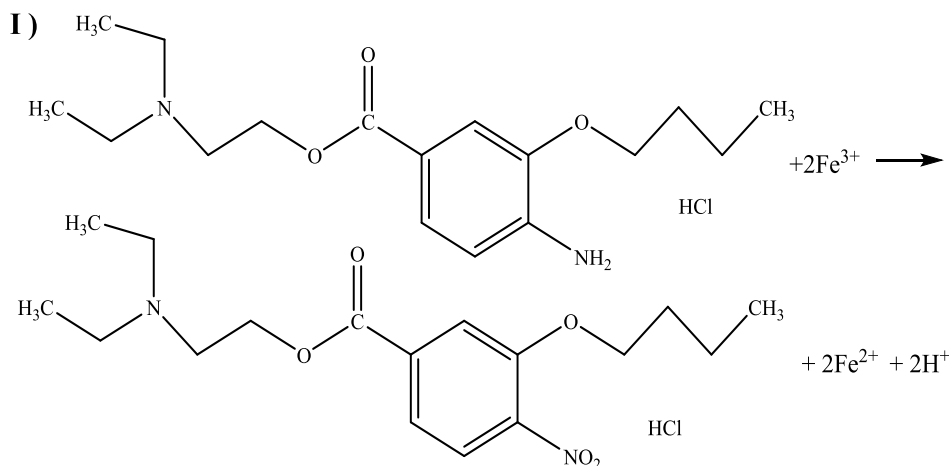


الشكل رقم (5) حساب نسبة التفاعل بطريقة الحد اللوغاريتمي للمجال الخطي

$$[\text{BNX}] = 1 \times 10^{-5} \text{M}, [\text{Fe}^{3+}] = 1.5 \times 10^{-4} - 4.5 \times 10^{-4} \text{M} \quad (\text{a})$$

$$[\text{Fe}^{3+}] = 1.5 \times 10^{-4} \text{M}, [\text{BNX}] = 1.0 \times 10^{-4} - 3.0 \times 10^{-4} \text{M} \quad (\text{b})$$

تم اقتراح آلية التفاعل المحتملة وفق المعادلات الآتية:



حسبنا من المعطيات العيارية لطريقة الزمن الثابت قيم معامل الامتصاص الجزيئي الأعظمي لنواتج التفاعل المتشكل وحساسية ساندل وحد الكشف، وحددنا معطيات الخط البياني العياري للطريقة، مثل معامل الارتباط ونقطة التقاطع مع محور الامتصاصية والميل، قدرنا مجال التركيز الخطي الأمثل للمركب BNX بطريقة Ringbom. كما هو مبين في الجدول رقم (2).

الجدول رقم (2) الخصائص الطيفية لنتاج أكسدة BNX بوجود Fe^{3+} وفري سيانيد البوتاسيوم

طريقة الزمن الثابت (t=15min)	القرائن
5.0-25.0	المجال الخطي، $\mu\text{g mL}^{-1}$
12.5-25.0	مجال Ringbom، $\mu\text{g mL}^{-1}$
1 : 2	نسبة التفاعل، $BNX- Fe^{3+}$
1.4×10^4	ϵ ، $\text{L mol}^{-1} \text{cm}^{-1}$
0.227	حد الكشف، $\mu\text{g mL}^{-1}$
5.00	حد التحديد الكمي، $\mu\text{g mL}^{-1}$
0.049	حساسية ساندل (SS)، $\mu\text{g cm}^{-2}$
$*(A = m C + b)$	
m= 0.0345	معطيات معادلة الخط البياني العياري
b= 0.0663	
0.9984	معامل الارتباط (r)

*فيما يتعلق بالمعادلة $A = m C + b$ ، حيث أن A الامتصاصية و C التركيز $\mu\text{g mL}^{-1}$ و m ميل الخط العياري و b نقطة تقاطع الخط العياري مع محور الامتصاصية.

اعتماد الطريقة

لتحديد دقة وصحة الطريقة المقترحة من أجل تحديد المركب BNX بالاعتماد. حضرنا خمسة محاليل عيارية من BNX تركيزها يقع ضمن المجال الخطي المحدد للطريقة. حللنا خمسة مكررات من كل واحد منها بتطبيق طريقة العمل المقترحة. أجرينا بعض الحسابات الاحصائية، مثل الانحراف العياري والانحراف العياري النسبي المئوي والاسترجاعية النسبية (الصحة النسبية) وحد الثقة عند سوية ثقة مقدارها 95%، لتراخيص مختلفة محددة من الخط البياني العياري، كما هو وارد في الجدول رقم (3). تميزت نتائج الطريقة الطيفية بدقة عالية من خلال الانحراف العياري النسبي المئوي المنخفض وبصحة ممتازة من خلال الاسترجاعية الجيدة.

الجدول رقم (3) تقدير دقة وصحة طريقة الزمن الثابت المقترحة من أجل تحديد BNX

الاسترجاعية %	حد الثقة	RSD%	$\mu\text{g mL}^{-1}$, BNX		
			S.D.	المحدد*	المأخوذ
100.80	5.04±0.17	2.77	0.14	5.04	5.00
100.90	10.09±0.23	1.88	0.19	10.09	10.00
100.87	15.13±0.33	1.78	0.27	15.13	15.00
101.40	20.28±0.28	1.13	0.23	20.28	20.00
100.36	25.09±0.33	1.07	0.27	25.09	25.00

*متوسط خمس تجارب.

التطبيق العملي

طبقتنا الطريقة المطورة من أجل تقدير BNX في المستحضرات الصيدلانية التجارية، حيث امتازت الطريقة بالدقة العالية والتكرارية الجيدة من خلال قيم الانحراف المعياري المنخفضة، وبالصحة الممتازة من خلال قيم الاسترجاعية الجيدة، وبالبساطة والسرعة وصلاحية الاستخدام في ضبط الجودة اليومية دون حدوث إعاقة من المواد المضافة إلى المستحضر الدوائي. كما تميزت الطريقة بعدم استهلاكها لكواشف ومذيبات مرتفعة الثمن. يبين الجدول رقم (4) النتائج التي حصلنا عليها.

الجدول رقم (4) تقدير BNX في المستحضرات الصيدلانية بالطريقة المقترحة

الاسترجاعية % ± S.D.		المستحضر
الطريقة المرجعية [18]	الطريقة المقترحة	
100.31±0.36 ** <i>t</i> = 1.39	100.66 ± 0.17 ** <i>t</i> = 1.69 ** <i>F</i> = 4.48	ميديكائين 0.4% (4mg/1mL)
100.44±0.56 ** <i>t</i> = 1.79	100.84±0.24 ** <i>t</i> = 1.16 ** <i>F</i> = 5.44	دلتاكائين 0.4% (4mg/1mL)

*متوسط خمس تجارب.

** من أجل أربع درجات حرية وحد ثقة 95% قيمة *t* الجدولة هي 2.776 وقيمة *F* الجدولة هي 6.26.

References

- [1] S. C. Sweetman, Martindale: The complete drug reference”, 36th ed., 2009, The pharmaceutical press, London, pp. 1868.
- [2] A El-Gindy, 2000, First Derivative Spectrophotometric and LC Determination of Benoxinate Hydrochloride and its Degradation Products, **Journal of Pharmaceutical and Biomedical Analysis**, 22(2):215-34.
- [3] Al-Farhan B. S., Khalil M. H., 2011, Sensitive Spectrophotometric Determination of Benoxinate Hydrochloride Using Different Aldehydes Through Condensation Reactions, **Journal of Saudi Chemical Society**, 15, 89–93.
- [4] A. Mery H., A. Mohammed M., A. Morsy F., Salem Y.M., 2015, Validated Spectrophotometric Methods for the Determination of Oxybuprocaine Hydrochloride, **Journal of Progressive Research in Chemistry**, 1 (1), 41-53.
- [5] Abdel-Gawad F. M., El-Guindi N., 1995, Spectrophotometric Determination of Metoclopramide and Oxybuprocaine Through Ion Pair Formation with Thiocyanate and Molybdenum(V) or Cobalt(II), **Analytical Letter**, 28(8), 1437-1447.
- [6] A. El Hamd M., El- Maghrabey M., Magdy G., A. Mahdi W., Alshehri S., K. A. Bass A., A. Batakoushy H., 2023, Application of Quality- by- Design for Adopting an Environmentally Green Fluorogenic Determination of Benoxinate Hydrochloride in Eye Drops and Artificial Aqueous Humour, **Scientific Reports**, 13, 8559.
- [7] Aboshabana R., Abo El Abass S., Elmansi H., 2022, Investigation of the Fluorescence of Benoxinate for Facile Determination in Pure form, Eye Drops and Aqueous Humor, **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, (264).

- [8] A. Attia, 2010, Electrochemical Determination of Anaesthetic Drug Benoxinate Hydrochloride, **Sensing in Electroanalysis**, (5), 209-219.
- [9] Elzanfaly E., Nebsen M., 2013, Ion Selective Membrane Electrodes for Stability Indicating Determination of Benoxinate Hydrochloride in Pure form and in Drug Product, **Research and Reviews in Electrochemistry**, 4(2), 41 – 47.
- [10] A. Mery H., A. Mohammed M., A. Morsy F., Salem Y.M., 2015, Determination of Oxybuprocaine HCL in Pharmaceutical Formulations Using Thermal Techniques, **Analytical Chemiistry An Indian Journal**, 15(9), 379-386.
- [11] Shoukry A. , Issa Y. , El-shiekR. h, Zareh M., 1991, New Ion-Selective Electrodes for Determination of Bupivacaine and Oxybuprocaine, **Analytical Letter**, 24(9), 1581-1590.
- [12] H. Kamal A., A. Habib A., F. Hammad Sh., M. Megahed S., 2023, Quality by Design Paradigm for Optimization of Green Stability Indicating HPLC Method for Concomitant Determination of Fluorescein and Benoxinate, **Scientific Reports**, 13:10471.
- [13] Habib A. A, Mabrouk M. M., Mammad F. Sh., Megahed M. S., 2020, Application of chromatographic Response Function in Deleopment of Stability Indicating HPLC Method for Determination of benoxinate Hydrochloride and Fluorescein Soduim Mixture Using Factorial Design, **Journal of Advanced Medical and Pharmaceutical Reseach**, 1, 1-8.
- [14] Kuhlmann O., Stoldt G., Struck H. G., Krauss G. J., 1998, Simultaneous Determination of Diclofenac and Oxybuprocaine in Human Aqueous Humor with HPLC and Electrochemical Detection, **Journal of Pharmaceutical and Biomedical Analysis**, 17(8), 1351-1356.

- [15] Chorny M., Levy D., Schumacher I., Lichaa Ch., Gruzman B., Livshits O., Lomnický Y., 2003, Development And Validation of A Stability-Indicating High Performance Liquid Chromatographic Assay for Benoxinate, **Journal of Pharmaceutical and Biomedical Analysis.**, 32(1), 189-196
- [16] V. M.KOPANICA, K. SATRA, I .ECHSCHLAGER, Z. RORSAK, K. KODUYS, SANDR (Eds)., 1983, Kinetic methods in chemical analysis, **Elsevier, Amsterdam, The Netherlands, 25-27.**
- [17] J. ROSE, 1964, Advanced Physico-chemical Experiments, **Pittman, London.**, P 67.
- [18] The United States Pharmacopeia: The National Formulary. 24th ed. Rockville: The United States Pharmacopeial Convention – Twinbrook Parkway, 2005.

تقييم أداء نماذج التعلم العميق على مهمة الكشف عن الموقف تجاه الأخبار الزائفة

إعداد: م. علي محرز*

إشراف: دكتور وسيم رمضان* - دكتور ناصر أبو صالح*

الملخص

تلعب مهمة الكشف عن الموقف دورا رئيسيا في منع انتشار الأخبار الزائفة. تركزت معظم الجهود البحثية على المهمة في اللغة الإنكليزية، إلا أن الدراسات على اللغة العربية لا تزال محدودة جدا. يوجد نقص كبير في الدراسات العربية التي تختبر فعالية نماذج التعلم العميق المتاحة. بناء على ذلك، قمنا في هذه الورقة بدراسة لتقييم أداء العديد من نماذج التعلم العميق: MLP، CNN، BiLSTM، CNNBiLSTM، AraStance، BiLSTM CNN على مجموعة بيانات عربية تدعى AraStance. تبين من خلال النتائج وجود صعوبة كبيرة في المهمة حيث تراوح أداء النماذج بين 66 و72 متوسط f1-score. تبين أيضا تفوق النموذج MLP على باقي النماذج محققا 72.29 متوسط f1-score مما يشير إلى فعالية إشارة الكلمات في تصنيف الموقف.

الكلمات المفتاحية: الكشف عن الموقف - الأخبار الزائفة - اللغة العربية - التعلم

العميق - معالجة اللغات الطبيعية

*طالب دكتوراة في قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث

*دكتور في قسم الاقتصاد الزراعي - كلية الزراعة - جامعة البعث

*دكتور في قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث

Evaluating the Performance of Deep Learning Models on the Task of Stance Detection Toward Fake News

By: Eng. Ali Mhrez

Supervised by: Dr. Wassim Ramadan, Dr. Naser Abo Saleh

Abstract

Stance detection plays a key role in preventing the spread of fake news. Most research efforts have focused on the task in English, but studies on Arabic are still very limited. There is a significant lack of Arabic studies that examine the effectiveness of available deep learning models. Accordingly, in this paper we conduct a study to evaluate the performance of several deep learning models: MLP, CNN, BiLSTM, CNNBiLSTM, BiLSTMCNN, and CNN~BiLSTM on an Arabic dataset called AraStance. The achieved results expose the difficulty of the task, as model performances ranged between 66 and 72 macro f1-score. Also, it turns out that the MLP outperforms the remaining models, achieving 72.29 macro f1-score, which indicates the effectiveness of the word signal in classifying the stance.

Keywords: Stance detection, fake news, Arabic language, deep learning, natural language processing

1. مقدمة

ساعد انتشار وسائل التواصل الاجتماعي واستخدامها من قبل الجميع للنشر بدون تكلفة من انتشار الأخبار غير الدقيقة وبشكل خاص الأخبار الزائفة التي تملأ وسائل التواصل الاجتماعي. في الواقع، تحد كمية المنشورات الهائلة التي يتم نشرها بشكل يومي على وسائل التواصل الاجتماعي من إمكانية ضبط انتشار الأخبار الزائفة بشكل يدوي.

يمكن تعريف مهمة الكشف عن الموقف على أنها مهمة التصنيف الآلي لموقف مؤلف قطعة من النص تجاه هدف يثير اهتمامنا وذلك على شكل تسمية فئوية من المجموعة: مع، ضد لا هذا ولا ذاك [1]. كما تم تعريف الأخبار الزائفة على أنها معلومات ملفقة تبدو كمحتوى إخباري حقيقي، لكن يفتقر إنتاجها إلى المعايير والعمليات المطلوبة لضمان دقتها ومصداقيتها [1].

نظرا لخطورة الأخبار الزائفة وتأثيرها على المجتمعات، ونظرا لعجز الأساليب التقليدية في منع انتشارها، بدء المجتمع البحثي بالبحث عن تقنيات وأنظمة قادرة على التخفيف من انتشار هذه الأخبار إن لم يكن هنالك إمكانية لمنعها بالكامل. وقد تم اقتراح مهمة الكشف عن الموقف كخطوة أساسية في أنبوب عمل الكشف عن الأخبار الزائفة [2].

تلعب مهمة الكشف عن الموقف دورين رئيسيين في اكتشاف الأخبار الزائفة. الدور الأول هو اكتشاف الأخبار الزائفة التي تعتمد على عنوان المقال في إيصال الفكرة. يلاحظ في هذه المقالات أن الأخبار في متن النص غير مرتبطة بالعنوان على الإطلاق. فتقوم خوارزمية الكشف عن الموقف باكتشاف عدم الارتباط هذا وبالتالي زيف الخبر.

الدور الثاني هو تصنيف مواقف مختلف وكالات الأخبار والجهات الرسمية تجاه صحة الأخبار المنشورة. فإن وافقت معظم الوكالات على الخبر فمن الأرجح أن يكون الخبر غير زائف. في المقابل، في حال رفضت معظم الوكالات الخبر فمن الأرجح أن يكون الخبر زائفا. أما في حال لم يكن هنالك موقف واضح وصريح تجاه صحة الخبر،

فلا يمكن الحكم في هذه الحالة. لابد من الإشارة هنا إلى ضرورة أخذ مصداقية الجهات التي يتم استقصاء رأيها بعين الاعتبار.

بدء الباحثون بتطوير نماذج تعلم آلة للكشف عن الموقف تجاه الأخبار الزائفة. كانت أوائل النماذج نماذج تقليدية قائمة على الميزات [4], [3]. لكن نظرا لعدم فعاليتها، تم البدء باستخدام نماذج التعلم العميق [11]-[5] التي قدمت أداء أفضل. مؤخرا ومع ظهور المحولات transformers [12] بدء المجتمع البحثي بالبحث في فعالية هذه المحولات transformers [14], [13] على المهمة البحثية.

تركزت الجهود البحثية على مهمة الكشف عن الموقف تجاه الأخبار الزائفة على اللغة الإنكليزية، إلا أن الدراسات على اللغة العربية لا تزال محدودة جدا [17]-[15]. في الواقع، لا يوجد دراسات عربية تختبر الكثير من نماذج التعلم العميق المتاحة في يومنا هذا والتي أثبتت بعض الفعالية على اللغة الإنكليزية.

تهدف هذه الدراسة إلى تقييم أداء العديد من نماذج التعلم العميق وإتاحتها للمجتمع البحثي. تساعد الدراسة على قياس التقدم في المهمة من جهة، بالإضافة إلى تمكين المقارنة من جهة أخرى.

لتحقيق ذلك، قمنا بتدريب النماذج التالية: بيرسبترون متعددة الطبقات - Multi-Layer Perceptron أو MLP، شبكة عصبية ترشيحية Convolutional Neural Network أو CNN، ذاكرة قصيرة المدى مطولة ثنائية الاتجاه Bidirectional Long Short-Term Memory أو BiLSTM، CNNBiLSTM، BiLSTMCNN، CNN~BiLSTM على مجموعة بيانات عربية تدعى AraStance [15] للكشف عن الموقف تجاه الأخبار الزائفة. لدعم هذه النماذج وتمكينها من تحقيق أفضل أداء، قمنا بالاعتماد على التعلم المنقول transfer learning عن طريق استخدام أشعة كلمات AraVec [18] في دخل هذه النماذج.

تشير النتائج التي تم الحصول عليها والتي تراوحت بين الـ 66 والـ 72 متوسط $f1$ - score إلى صعوبة المهمة. على الرغم من ذلك، تفوق النموذج MLP على باقي النماذج محققا متوسط $f1$ -score بلغ 72.29 % مما يشير إلى فعالية إشارة الكلمات في تصنيف الموقف.

الورقة منظمة على الشكل التالي. تمت مراجعة الدراسات السابقة على مهمة الكشف عن الموقف تجاه الأخبار الزائفة في القسم 2. يقدم القسم 3 أدوات وطرائق البحث متناولا مجموعة البيانات والنماذج المختبرة بالإضافة إلى إعدادات التجارب العملية. في القسم 4، تم تقديم نتائج التجارب العملية ومناقشتها بالتفصيل. في النهاية، يلخص القسم 5 الورقة كما يقدم أهم التوجهات المستقبلية.

2. الدراسات السابقة

تدرج خوارزميات الكشف عن الموقف في سياق الأخبار الزائفة تحت مظلة مجال تعلم الآلة، وتحديدًا التعلم العميق مع ظهور محدود لخوارزميات التعلم التقليدية. تعتمد معظم الخوارزميات المقترحة على التعلم المنقول $transfer learning$ من الأشعة المدربة مسبقا بمختلف مستوياتها (أحرف، كلمات، جمل... إلخ) إلى المحولات المدربة مسبقا.

كانت أوائل الدراسات عبارة عن نماذج تعلم آلة تقليدية قائمة على الميزات. فاقترح الباحثون في [3] مجموعة مكدسة من المصنفات مكونة من طبقتين. تحوي الطبقة الأولى على ثلاث نماذج MLP، نموذج انحدار لوجستي Logistic Regression أو LR ونموذج أشجار قرار معززة بالمشق Gradient Boosted Decision Trees أو GBDT. أما الطبقة الثانية فتحتوي على نموذج GBDT. قدم هذا النموذج المكدس أداء ضعيفا محتلا المركز الحادي عشر في منافسة الأخبار الزائفة [2] التي حصلت عام 2017.

أيضا اقترح الباحثون في [4] نظاما مكونا من خطوتين. يتم في الخطوة الأولى تصنيف الارتباط، وفي الثانية تصنيف الموقف باستخدام مصنف LR. وقد قدم هذا النموذج كما السابق أداء ضعيفا محققا المركز التاسع في المنافسة.

اقترح الباحثون في [5] نموذج MLP بطبقة مخفية واحدة يعتمد في دخله على أشعة tf-idf. وعلى الرغم من بساطة هذا النموذج إلا أنه تفوق على النماذج السابقة محققا المركز الثالث في المنافسة.

وبنفس الطريقة اقترح الباحثون في [6] مصنف مجموعة مكون من 5 نماذج MLP. يحوي كل نموذج سبع طبقات مخفية. وقد تفوق هذا النموذج على النموذج السابق بفارق صغير جدا محققا المركز الثاني في المنافسة.

أما بالنسبة للنموذج الحاصل على المركز الأول في المنافسة [7]، فقد اعتمد على مصنف مجموعة مكون من نموذج CNN ونموذج GBDT [19].

في الواقع عجزت جميع الأنظمة المشاركة في المنافسة عن إيجاد حل فعال ومقبول لمشكلة الكشف عن الموقف تجاه الأخبار الزائفة. فوجد الباحثون في [8] أن جميع الأنظمة المشاركة في منافسة الأخبار الزائفة تفتقر إلى القدرة على فهم النص. فقاموا بتكديس نموذجي ذاكرة LSTM لتشكيل تمثيل للدخل باستخدام أشعة GloVe-50. للتصنيف، قاموا بإغناء التمثيل ببعض الميزات المستخرجة يدويا وإدخال التمثيل النهائي إلى نموذج MLP ثلاثي الطبقات للمزيد من المعالجة. وقد تفوق هذا النموذج على جميع أنظمة المنافسة بالاعتماد على مقياس متوسط f1-score.

اقترح الباحثون في [9] دمج نموذج LSTM ونموذج CNN لتشكيل شبكة ذاكرة تقوم بتصنيف الموقف بالإضافة إلى استخراج قطع نصية لتبرير قرار التصنيف. تقوم الشبكة بتمثيل الدخل بالاعتماد على أشعة كلمات GloVe-100. كما تستخدم مصفوفة تشابه لاستخراج أجزاء المقال الأكثر ارتباطا بالعنوان. للتصنيف، يتم الاعتماد على تمثيلات

العنوان والمقال بالإضافة إلى التشابه بينهما. وقد تفوقت شبكة الذاكرة المقترحة على نموذج الذاكرة المكسد السابق [8].

ركز الباحثون في [10] على تصنيف الموقف عن طريق إزالة جميع الأمثلة الغير مرتبطة من مجموعة بيانات المنافسة. قاموا باختبار نموذجين لترميز العلاقة بين العنوان والمقالة. في النموذج الأولي، يتم ترميز العلاقة باستخدام الترميز الشرطي المزوج double conditional encoding باستخدام نموذج BiLSTM شرطي. وفي النموذج الثاني، يتم ترميز العلاقة باستخدام الانتباه المتطابق co-matching attention. للتصنيف، تقوم LSTM عكسية بإنتاج التمثيل النهائي للمقال. وقد تفوق نموذج الانتباه المتطابق على نموذج الترميز الشرطي المزوج، كما تفوق على شبكة الذاكرة السابقة [9] محققا أداء أفضل على مقياس متوسط f1-score.

قام الباحثون في [11] بتحسين نموذج BiLSTM باستخدام تقنية انتباه Attention تسمح للنموذج بالتركيز على أجزاء التمثيل الأكثر أهمية في عملية التصنيف. وقد اختبروا ثلاث تقنيات انتباه. الأولى بسيطة يتم فيها مقارنة جميع كلمات المقال بالتمثيل النهائي للعنوان. الثانية كاملة يتم فيها مقارنة جميع كلمات المقال بجميع كلمات العنوان. الثالثة متعددة يتم فيها مقارنة جميع الكلمات من عدة اتجاهات. تم تقديم النتائج بالاعتماد على مقياس الدقة الموزونة مما لم يسمح بالمقارنة مع الدراسات السابقة.

مؤخرا بدئت الدراسات باختبار أداء المحولات transformers على مهمة الكشف عن الموقف تجاه الأخبار الزائفة. فقام الباحثون في [13] باختبار أداء BERT [20] وXLNet [21] وRoBERTa [22] على مجموعة بيانات المنافسة. وقد وجد الباحثون أن هذه المحولات تتفوق على جميع النماذج السابقة.

وضمن نفس التوجه أيضا قام الباحثون في [14] باختبار أداء BERT و RoBERTa و DistilBERT [23] و ALBERT [24] و XLNet على مجموعة بيانات FNC-1 أيضا. وقد تم تأكيد تفوق هذه المحولات على النماذج السابقة.

على الرغم من الاهتمام الكبير الذي حظيت به مهمة الكشف عن الموقف بشكل عام وتحدي الأخبار الزائفة بشكل خاص، إلا أن دراسات الكشف عن الموقف تجاه الأخبار الزائفة في اللغة العربية لاتزال محدودة جدا. فقام الباحثون في [16] ببناء مجموعة بيانات عربية جديدة. لاختبار جودة وإمكانية استخدام مجموعة البيانات، قام الباحثون باختبار نماذج الـ MLP المقترحة في [6], [5]. كما قاموا باختبار أشجار القرار المقترحة في [19] ونموذج الذاكرة المقترح في [9]. وقد كشفت الاختبارات على مجموعة البيانات الجديدة عن بعض الصعوبات في المهمة في اللغة العربية.

بنفس الطريقة، قام الباحثون في [17] بمقارنة العديد من نماذج الـ LSTM مع نموذج mBERT [25] على مجموعة بيانات عربية قاموا بإنشائها. أظهرت النتائج أن mBERT المدرب مسبقا يتفوق في الأداء على الـ LSTMs المدربين فقط على المهمة الهدف.

وفي النهاية، قام الباحثون في [15] بمقارنة بين mBERT و ArabicBERT [26] و ArBERT و MarBERT [27]، على مجموعة بيانات عربية جديدة قاموا بإنشائها. وأظهرت المقارنة فعالية تلك النماذج مع أفضلية طفيفة لـ ArBERT على النماذج الأخرى.

بعد مراجعة الدراسات السابقة تبين أن الدراسات العربية على المهمة ركزت على إنشاء مجموعات بيانات واختبار جودة هذه المجموعات عن طريق اختبار النماذج المقترحة سابقا عليها. لكن لم يتم أحد باختبار الطيف الواسع من نماذج التعلم العميق على المهمة وإتاحتها لقياس التقدم والمقارنة المستقبلية.

3. أدوات وطرائق البحث

3.1 مجموعة البيانات

تم الاعتماد على مجموعة بيانات AraStance [15] في التجارب العملية. تتميز هذه المجموعة عن مجموعة بيانات ArabicFC [16] ومجموعة بيانات ANS [17] في كونها متعددة المواضيع، أي أنها تتناول مواضيعا سياسية واقتصادية وعسكرية ورياضية من دون التركيز على موضوع واحد فقط، كما تتميز في كونها الأحدث والأكثر توازنا [15]. تحتوي مجموعة البيانات على 910 ادعاءات و4063 وثيقة. تم جمع الادعاءات من مواقع التحقق من الصحة ARAANEWS و DABEGAD و NORUMORS بالإضافة إلى موقع REUTERS الإخباري. أما بالنسبة للوثائق، فقد تم جمعها بالاعتماد على محرك بحث غوغل. تأتي مجموعة البيانات مقسمة إلى ثلاث مجموعات تدريب وتحقق واختبار، وبالتالي لا حاجة لإعادة تقسيمها. يعرض الجدول 1 إحصائيات مجموعة بيانات AraStance.

جدول 1 - إحصائيات مجموعة بيانات AraStance.

الصف	تدريب	تحقق	اختبار
يوافق	739	129	154
لا يوافق	309	76	64
يناقش	247	70	70
غير مرتبط	1553	294	358
الإجمالي	2848	569	646

3.2 النماذج

يعرض الجدول 2 معمارية النماذج المختبرة. فيما يلي شرح مختصر لهذه النماذج:

MLP: هذا النموذج مكون من طبقة كثيفة dense واحدة تحوي أربع خلايا عصبية.

CNN: يتكون هذا النموذج من ثلاث طبقات ترشيح convolutional layers تحوي كل منها 100 مرشح. تستخدم هذه الطبقات مرشحات بأحجام تتراوح بين 2 و 4. لتجميع خرائط الخصائص feature maps المستخرجة، يتم استخدام طبقة تجميع pooling layer لكل طبقة ترشيح. لمنع النموذج من الوقوع في مشكلة فرط التخصيص overfitting، تم استخدام طبقة إسقاط dropout layer. في النهاية، يتم إدخال الخصائص الناتجة إلى طبقة كثيفة تحوي 4 خلايا عصبية.

BiLSTM: طبقتي ذاكرة ثنائي الاتجاه متتاليتين، تقوم الثانية بمعالجة خرج الأولى. يتم معالجة الأشعة السياقية الناتجة من طبقة الذاكرة الثانية باستخدام طبقة كثيفة بأربع خلايا مخفية.

CNNBiLSTM: تقوم ثلاث طبقات ترشيح تحوي كل منها على 100 مرشح، باستخراج خرائط الخصائص من الدخل. يتم معالجة خرائط الخصائص المستخرجة بالاعتماد على طبقة ذاكرة ثنائية الاتجاه. في النهاية، يتم إدخال الأشعة السياقية contextual vectors الناتجة من طبقة الذاكرة إلى طبقة كثيفة بأربع خلايا مخفية.

جدول 2 - الطبقات التي تتكون منها النماذج المختبرة بالإضافة إلى معاملاتها.

الطبقات	النموذج	
كثيفة: عدد الوحدات: 4	MLP	
تضمين: عدد الخصائص: 30000 - طول السلسلة: 300	CNN	
ترشيح: عدد المرشحات: [100، 100، 100] - الأحجام: [2، 3، 4] - التنغيم: relu		
تجميع: نوع: أعظمي عالمي - الأبعاد: أحادي البعد		
إسقاط: معامل الإسقاط: 0.3		
كثيفة: عدد الوحدات: 4	BiLSTM	
تضمين: عدد الخصائص: 20000 - طول السلسلة: 100		
ذاكرة: عدد الوحدات: 32 - الاتجاه: ثنائي - التنغيم: tanh		
ذاكرة: عدد الوحدات: 32 - الاتجاه: ثنائي - التنغيم: tanh		
كثيفة: عدد الوحدات: 4	CNNBiLSTM	
تضمين: عدد الخصائص: 30000 - طول السلسلة: 300		
ترشيح: عدد المرشحات: [100، 100، 100] - الأحجام: [2، 3، 4] - التنغيم: relu		
ذاكرة: عدد الوحدات: 32 - الاتجاه: ثنائي - التنغيم: tanh		
كثيفة: عدد الوحدات: 4	BiLSTM CNN	
تضمين: عدد الخصائص: 20000 - طول السلسلة: 150		
ذاكرة: عدد الوحدات: 32 - الاتجاه: ثنائي - التنغيم: tanh		
ذاكرة: عدد الوحدات: 32 - الاتجاه: ثنائي - التنغيم: tanh		
ترشيح: عدد المرشحات: [50، 50، 50] - الأحجام: [2، 3، 4] - التنغيم: relu		
تجميع: نوع: أعظمي عالمي - الأبعاد: أحادي البعد		
كثيفة: عدد الوحدات: 4	CNN~BiLSTM	
تضمين: عدد الخصائص: 30000 - طول السلسلة: 300		
BiLSTM	CNN	M

BiLSTM CNN: تقوم طبقتي ذاكرة ثنائيتي الاتجاه بمعالجة الدخل وتشكيل تمثيلات سياقية له. فتقوم ثلاث طبقات ترشيح متعددة الأحجام تحوي كل منها على 50 مرشح باستخراج خرائط الخصائص بالاعتماد على التمثيلات السياقية. ومن ثم يتم تجميع خرائط الخصائص باستخدام طبقات تجميع أحادية البعد. وفي النهاية يتم القيام بمعالجة إضافية باستخدام طبقة كثيفة مكونة من أربع خلايا عصبية.

CNN~BiLSTM: في هذا النموذج المجمع، يتم استخراج خرائط الخصائص باستخدام نموذج الـ CNN المستقل السابق. يتم أيضا معالجة الدخل بشكل تسلسلي باستخدام نموذج الـ BiLSTM المستقل السابق. يقوم النموذجين بالتصويت على الصف المناسب، ليتم في النهاية تجميع الاحتمالات وأخذ متوسطها للحصول على القرار النهائي.

3.3 إعدادات التجارب العملية

يعرض الجدول 3 إعدادات التدريب المستخدمة لتدريب جميع النماذج. تم تحقيق هذه النماذج باستخدام مكتبة keras من منصة عمل tensorflow [28]. تم إيجاد القيمة الصغرى لتابع الخسارة cross entropy بالاعتماد على خوارزمية Adam [29]. للتقييم، تم الاعتماد على مقياس متوسط f1-score الذي يمثل المتوسط الحسابي لمجموع قيم الـ f1-scores عبر الصفوف.

جدول 3 - إعدادات التدريب للنماذج المختبرة.

النموذج	حجم الدفعة	معامل التعلم
MLP	512	0.1
CNN	64	0.001
BiLSTM	128	0.0006
CNNBiLSTM	64	0.0006
BiLSTMCNN	128	0.0006
CNN~BiLSTM	128	0.0009

3.4 أشعة كلمات AraVec وتمثيل الدخل

AraVec [18] هو مشروع أشعة كلمات موزعة مدربة مسبقا مفتوح المصدر يهدف إلى تقديم أشعة كلمات قوية لمجتمع معالجة اللغات الطبيعية في اللغة العربية. يقدم المشروع في إصداره الأول 6 أشعة كلمات مبنية على مصادر نصية مختلفة: التويتز وصفحات الويب والويكيبيديا، حيث يوجد شعاعين كلمات لكل مصدر، أحدهما يعتمد في التدريب على نموذج حقائب الكلمات المستمرة Continuous Bag Of Words (CBOW) والآخر على نموذج skip-gram. يقدم المشروع في آخر إصداراته¹ أشعة كلمات مبنية على التويتز والويكيبيديا بالاعتماد على النماذج المذكورة وبأبعاد 100 و300.

¹ <https://github.com/bakrianoo/aravec>

نظرا لأن مجموعة بيانات AraStance لا تحوي على بيانات من مواقع التواصل الاجتماعي، قمنا في تجاربنا العملية بالاعتماد على أشعة AraVec المبنية على الـ word2vec باستخدام نموذج CBOW. تم الاعتماد على أشعة بـ 300 بعد كونها أغنى من الأشعة التي تأتي بـ 100 بعد، وهذا ما أكدته التجارب العملية الأولية التي قمنا بها.

تم استغلال أشعة الكلمات المذكورة عن طريق بناء طبقة تضمين embedding layer كدخل لجميع النماذج (باستثناء نموذج الـ MLP الذي يعتمد في دخله على خصائص مستخرجة يدويا من النص).

أما بالنسبة لنموذج الـ MLP، فيعتمد في دخله على أشعة tf-idf تم تشكيلها بالاعتماد على أكثر 4000 كلمة مستخدمة ضمن مجموعة التدريب. يتم في البداية تحديد أكثر 4000 كلمة مستخدمة. بناء عليها يتم تشكيل تمثيل للعنوان والوثيقة بشكل منفصل، كما يتم حساب التشابه بين هذين التمثيلين بالاعتماد على مقياس تشابه الـ cosine. وفي النهاية، يتم دمج تمثيلي الوثيقة والعنوان مع قيمة التشابه كدخل لهذا النموذج.

3.5 المعالجة المسبقة للبيانات

قبل استخدام أشعة AraVec لا بد من القيام بالمعالجة المسبقة الضرورية التي تم القيام بها لإنتاج هذه الأشعة. فقمنا بتوحيد حالة الأحرف عن طريق إزالة الهمزات عن الألف أ، إ، آ لتصبح ا، تحويل التاء المربوطة إلى هاء ه، تحويل الألف المقصورة إلى ياء ي. تم أيضا إزالة التشكيل (الفتحة، الضمة، الكسرة، السكون، تنوين الفتحة، تنوين الضمة، تنوين الكسرة، الشدة) والتطويل.

بالإضافة إلى ما سبق، قمنا بإزالة علامات الترقيم التي لا تضيف أي معنى للنص، وتحويل الأرقام إلى كلمات للسماح بتمثيلها ومعالجتها، تحويل إشارة % إلى %، وفي النهاية، إزالة الفراغات الإضافية.

أما بالنسبة للنموذج MLP، فقد قمنا بإزالة كلمات التوقف stop words لمنعها من الهيمنة على باقي الكلمات عند البحث عن أكثر الكلمات تكرارا في مجموعة البيانات.

4. النتائج والمناقشة

يعرض الجدولين 4 و5 نتائج النماذج على مجموعتي التحقق والاختبار في مجموعة بيانات AraStance. تفوق النموذج MLP على باقي النماذج في مجموعتي التحقق والاختبار محققا 70.75 متوسط f1-score على مجموعة التحقق و72.29 على مجموعة الاختبار. يأتي ذلك بسبب تفوق هذا النموذج على باقي النماذج عبر الصفوف يوافق ويناقش وغير مرتبط في مجموعة التحقق وعبر الصفوف لا يوافق ويناقش في مجموعة الاختبار، بالإضافة إلى أداءه المنافس جدا على باقي الصفوف.

كما احتل النموذج CNN المركز الثاني عبر المجموعتين أيضا محققا 69.43 متوسط f1-score على مجموعة التحقق و71.59 على مجموعة الاختبار. في الواقع، تفوق النموذج MLP على نموذج الـ CNN بفارق 0.7 نقطة فقط في مجموعة الاختبار مما يشير إلى وجود منافسة بين هذين النموذجين.

جدول 4 - نتائج النماذج على مجموعة التحقق. هذه النتائج هي متوسط ثلاث تكرارات لكل تجربة.

المتوسط f1- score	غير مرتبط	يناقش	لا يوافق	يوافق	الدقة	النموذج (%)
70.75	84.32	50.44	72.61	75.64	77.57	MLP
69.43	83.06	48.16	73.80	72.71	76.92	CNN
63.83	81.27	43.77	65.01	65.28	72.17	BiLSTM
64.13	80.79	44.34	62.54	68.86	72.58	CNNBiLSTM
68.12	82.72	47.47	69.52	72.77	74.75	BiLSTMCNN
66.99	81.59	48.37	69.80	68.20	74.40	CNN~BiLSTM

في المركز الثالث، يأتي النموذج BiLSTMCNN محققا 68.12 و 70.13 متوسط f1-scores على مجموعتي التحقق والاختبار. من الواضح أن فجوة الأداء بين هذا النموذج والنموذج الأول بدئت تتسع حيث يمكن ملاحظة انخفاض يتراوح بين النقطة والنصف والخمس نقاط عبر جميع الصفوف (باستثناء الصف غير مرتبط) مما يشير إلى تفوق تام للنموذج MLP على هذا النموذج، وبالتالي على جميع النماذج التي تليه.

يأتي مصنف المجموعة CNN~BiLSTM في المركز الرابع محققا 66.99 متوسط f1-score على مجموعة التحقق و 69.89 على مجموعة الاختبار. في الواقع، قدم مصنف المجموعة أداء مقاربا جدا للنموذج BiLSTMCNN على مجموعة الاختبار وبفارق حوالي 0.24 نقطة فقط.

بشكل مفاجئ، احتل النموذج BiLSTM المركز ما قبل الأخير مسجلا 63.83 متوسط f1-score على مجموعة التحقق و 68.19 على مجموعة الاختبار. يمكن

بسهولة ملاحظة معاناة هذا النموذج مع الصفوف المرتبطة (يوافق ولا يوافق ويناقدش) عبر مجموعتي التحقق والاختبار.

جدول 5 - نتائج النماذج على مجموعة الاختبار. هذه النتائج هي متوسط ثلاث تكرارات لكل تجربة.

النموذج	الدقة	يوافق	لا يوافق	يناقدش	غير مرتبط	متوسط f1-score
MLP	82.30	82.03	<u>75.90</u>	<u>42.86</u>	88.36	<u>72.29</u>
CNN	<u>82.61</u>	<u>83.57</u>	72.80	41.25	88.73	71.59
BiLSTM	78.64	76.75	70.62	39.33	86.05	68.19
CNNBiLSTM	78.89	77.25	69.21	32.57	86.93	66.49
BiLSTMCNN	80.34	76.85	73.98	40.73	<u>88.98</u>	70.13
CNN~BiLSTM	80.86	78.66	72.43	40.12	88.35	69.89

أما المركز الأخير فكان من نصيب النموذج CNNBiLSTM الذي قدم الأداء الأسوأ مسجلا 64.13 على مجموعة التحقق و66.49 على مجموعة الاختبار. يبدو من خلال النتائج أن عجز هذا النموذج على الصف يناقدش في مجموعة الاختبار كان السبب الرئيسي في تحقيقه للمركز الأخير.

يبدو من خلال ترتيب النماذج والنتائج المحققة أن إشارة الكلمات لوحدها قادرة على تصنيف الموقف بدقة حوالي 72 متوسط f1-score. هذا واضح من خلال النموذج MLP البسيط الذي حقق المركز الأول عن طريق اعتماده في الدخل على أشعة tf-idf فقط. يؤكد ذلك أيضا المركز الثاني الذي حققه النموذج CNN والذي يقوم داخليا باستخراج الـ n-grams من الدخل (في حالتنا بأطوال 2، 3، 4) عن طريق عمليات الترشيح التي يقوم بها.

يبدو أيضا من خلال النتائج أن المعالجة التسلسلية التي يقوم بها نموذج الـ BiLSTM غير مفيدة في اكتشاف الموقف في الإعدادات المدروسة. فقد حقق النموذج BiLSTM المركز ما قبل الأخير. بالإضافة إلى ذلك، لم يستطع أي نموذج يقوم بالمعالجة التسلسلية للدخل باستخدام الـ BiLSTM التفوق على النماذج التي تعتمد على إشارة الكلمات (MLP، CNN) في عملها وبدون أي معالجة تسلسلية.

من الواضح أن الأشعة السياقية الناتجة عن المعالجة التسلسلية غير فعالة في اكتشاف الموقف، حيث أن الاعتماد عليها في التصنيف في النموذجين BiLSTM و CNNBiLSTM أدى إلى المركزين الأخيرين. في المقابل، القيام بترشيح هذه الأشعة عن طريق نموذج الـ CNN يحسن قليلا من الأداء (BiLSTM CNN).

أما بالنسبة للنموذج المجمع CNN~BiLSTM فمن الواضح تأثر الأداء بالمعالجة التسلسلية الغير مفيدة. على الرغم من ذلك، إلا أن هذا النموذج لا يزال أفضل من الـ BiLSTM لوحدها والـ CNNBiLSTM.

في النهاية، يشير تفوق النموذج البسيط MLP على باقي النماذج الأكثر كلفة وتعقيدا والتي تعتمد على التعلم المنقول من خلال أشعة كلمات AraVec إلى أن التعلم المنقول لوحده لم يحقق ميزة إضافية ونقطة تفوق للنماذج المعتمدة عليه. من الأكيد أن طول الوثائق هو أحد التحديات الكبيرة في هذه المهمة. بالإضافة إلى ذلك، فإن نوع الدخل المكون من قطعتين نصيتين متفاوتتي الطول، بدلا من قطعة نصية واحدة، يزيد من الصعوبة أيضا. من الممكن في هذه الحالة أن تكون تقنيات الانتباه ذات قيمة كبيرة عن طريق السماح للنماذج بالتركيز على أجزاء الدخل الأكثر أهمية في عملية التصنيف.

5. الخاتمة

قمنا في هذه الورقة باختبار أداء العديد من نماذج التعلم العميق على مهمة الكشف عن الموقف تجاه الأخبار الزائفة في اللغة العربية. تبين من خلال النتائج المحققة مدى صعوبة المهمة حيث تراوحت النتائج بين 66 و72 متوسط f1-score. كما تبين تفوق النموذج MLP على باقي النماذج ويليه النموذج CNN بأداء منافس جدا مما يشير إلى أهمية إشارة الكلمات وفعاليتها في الكشف عن الموقف تجاه الأخبار الزائفة. تبين أيضا أن المعالجة التسلسلية التي يقوم بها نموذج الـ BiLSTM والأشعة السياقية الناتجة غير فعالة في اكتشاف الموقف على مجموعة البيانات المدروسة.

سيتم لاحقا اختبار المزيد من التقنيات، وبشكل خاص تقنيات الانتباه التي نعتقد أن لها دور كبير في هذا النوع من المهام الذي يتضمن قطعتين نصيتين بدلا من قطعة نصية واحدة. كما سنقوم باختبار تقنيات لاستخراج أجزاء الوثيقة الأكثر ارتباطا بالعنوان. نعتقد أن هذا التوجه الأخير من أهم التوجهات التي يمكن العمل عليها في المستقبل.

المراجع

- [1] KÜÇÜK D. *et al.*, 2021 – Stance Detection: A Survey, ACM Computing Surveys, vol. 53, no. 1, pp. 1–37
- [2] POMERLEAU D. *et al.*, 2017 – Fake news challenge stage 1 (FNC-1): Stance detection. <http://www.fakenewschallenge.org/> (accessed Jun. 04, 2022).
- [3] THORNE J. *et al.*, 2017 – Fake news stance detection using stacked ensemble of classifiers, in *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*, pp.80–83
- [4] BOURGONJE P. *et al.*, 2017 – From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles, in *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*, pp.84–89
- [5] RIEDEL B. *et al.*, 2017 – A simple but tough-to-beat baseline for the Fake News Challenge stance detection task, arXiv:1707.03264v2 [cs.CL]
- [6] HANSELOWSKI A. *et al.*, Athene System. https://github.com/hanselowski/athene_system (accessed Apr. 04, 2023).

- [7] PAN Y. *et al.*, 2017 – Team SOLAT IN THE SWEN.
<https://github.com/Cisco-Talos/fnc-1> (accessed Apr. 04, 2023).
- [8] HANSELOWSKI A. *et al.*, 2018 – A Retrospective Analysis of the Fake News Challenge Stance-Detection Task, in *Proceedings of the 27th International Conference on Computational Linguistics*, pp.1859–1874
- [9] MOHTARAMI M. *et al.*, 2018 – Automatic Stance Detection Using End-to-End Memory Networks, in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pp.767–776
- [10] CONFORTI C. *et al.*, 2018 – Towards Automatic Fake News Detection: Cross-Level Stance Detection in News Articles, in *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)*, pp.40–49
- [11] ZENG Q. *et al.*, 2018 – Neural Stance Detectors for Fake News Challenge.
- [12] VASWANI A. *et al.*, 2017 – Attention Is All You Need, [arXiv:1706.03762v7](https://arxiv.org/abs/1706.03762v7) [cs.CL]
- [13] SLOVIKOVSKAYA V. *et al.*, 2020 – Transfer Learning from Transformers to Fake News Challenge Stance Detection

- (FNC-1) Task, in *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pp.1211-1218
- [14] GUDERLEI M. *et al.*, 2020 – Evaluating Unsupervised Representation Learning for Detecting Stances of Fake News, in *Proceedings of the 28th International Conference on Computational Linguistics*, pp.6339-6349
- [15] ALHINDI T. *et al.*, 2021 – AraStance: A Multi-Country and Multi-Domain Dataset of Arabic Stance Detection for Fact Checking, in *Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda*, pp.57-65
- [16] BALY R. *et al.*, 2018 – Integrating Stance Detection and Fact Checking in a Unified Corpus, in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pp.21-27
- [17] KHOUJA J., 2020 – Stance Prediction and Claim Verification: An Arabic Perspective, in *Proceedings of the Third Workshop on Fact Extraction and VERification (FEVER)*, pp.8-

- [18] SOLIMAN A. B. *et al.*, 2017 – AraVec: A set of Arabic Word Embedding Models for use in Arabic NLP, Procedia Computer Science, vol. 117, pp. 256–265
- [19] POMERLEAU D. *et al.*, FNC-1 Baseline System. <https://github.com/FakeNewsChallenge/fnc-1-baseline> (accessed Apr. 04, 2023).
- [20] DEVLIN J. *et al.*, 2018 – BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805v2 [cs.CL]
- [21] YANG Z. *et al.*, 2019 – XLNet: Generalized Autoregressive Pretraining for Language Understanding, arXiv:1906.08237v2 [cs.CL]
- [22] LIU Y. *et al.*, 2019 – RoBERTa: A Robustly Optimized BERT Pretraining Approach, arXiv:1907.11692v1 [cs.CL]
- [23] SANH V. *et al.*, 2019 – DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, arXiv:1910.01108v4 [cs.CL]
- [24] LAN Z. *et al.*, 2019 – ALBERT: A Lite BERT for Self-supervised Learning of Language Representations, arXiv:1909.11942v6 [cs.CL]

- [25] Multilingual BERT. <https://github.com/google-research/bert/blob/master/multilingual.md> (accessed Apr. 04, 2023).
- [26] SAFAYA A. *et al.*, 2020 – KUISAIL at SemEval-2020 Task 12: BERT-CNN for Offensive Speech Identification in Social Media, in *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp.2054–2059
- [27] ABDUL-MAGEED M. *et al.*, 2021 – ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic, in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp.7088–7105
- [28] ABADI M. *et al.*, 2016 – TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, [arXiv:1603.04467v2](https://arxiv.org/abs/1603.04467v2) [cs.DC]
- [29] KINGMA D. P. *et al.*, 2014 – Adam: A Method for Stochastic Optimization, [arXiv:1412.6980v9](https://arxiv.org/abs/1412.6980v9) [cs.LG]

تحسين دقة التعرف على النشاط البشري باستخدام المشفر التلقائي، الذاكرة طويلة المدى والمحول

طالب الدراسات العليا: عمار سهيل مقصود
كلية الهندسة المعلوماتية - جامعة البعث
المشرف العلمي: د. د. ماهر عباس

المخلص

تُعتبر إنترنت الأشياء أحد المصادر الرئيسية لتوليد البيانات الضخمة، لأنها تعتمد على ربط عدد كبير من الأجهزة الذكية بالإنترنت للإبلاغ عن حالة بياناتها التي يتم التقاطها بشكل مستمر. يُعد التعرف على الأنماط ذات المعنى واستخراجها من بيانات المدخلات الخام الهائلة بمثابة الأداة الأساسية لتحليل البيانات الضخمة لأنها تزود بمستويات أعلى من الرؤى لاتخاذ القرارات والتنبؤ بالنزعات والاتجاهات Trends. تُعرف البيانات الضخمة في بيئة إنترنت الأشياء ببيانات لها ميزات الحجم الكبير Volume، السرعة الكبيرة في التوليد والانتقال Velocity، والتنوع Variety. بشكل عام، تأتي البيانات الضخمة بأشكال وتنسيقات مختلفة واردة من مصادر متعددة. يلعب دمج البيانات Data Fusion دوراً حاسماً في تطوير بيئات واسعة الانتشار معتمدة على بيانات إنترنت الأشياء. يعد هذا الدور أكثر أهمية بالنسبة إلى تطبيقات إنترنت الأشياء الحساسة للوقت حيث يلزم دمج البيانات في الوقت المناسب لجلب كل أجزاء البيانات معاً للتحليل وبالتالي توفير رؤى موثوقة ودقيقة وقابلة للتنفيذ. تهدف هذه الدراسة إلى استعراض الأبحاث التي عالجت خاصية التنوع (دمج البيانات) للبيانات الضخمة في تطبيقات التعرف على النشاط البشري كأحد بيئات إنترنت الأشياء، بغرض تقييمها والانطلاق نحو بناء نموذج يلحظ تنوع مصادر البيانات واختلاف تنسيقها، حيث تقوم فكرة النموذج

المقترح على استخلاص الميزات الغنية من بيانات المصدر الأكبر لناحية حجم الأبعاد عبر بناء نموذج مشفر تلقائي Autoencoder وتدريبه بهدف استخدامه لاحقاً كتقنية لنقل التعلم Transfer Learning. واستخراج الارتباطات الزمنية من كل مصدر باستخدام شبكة Lstm وأخيراً دمج الميزات المستخرجة بواسطة تقنية Cross Attention. من ثم تحليل النتائج من وجهة نظر مجموعة من معايير الأداء مثل الدقة Accuracy، F1 Score، مصفوفة التشويش Confusion Matrix. تشير النتائج إلى أن أداء النموذج المقترح حقق دقة أعلى بمقدار 2% من الدراسات التي استخدمت نفس مصادر البيانات من نفس مجموعة البيانات UTD-MHAD.

الكلمات المفتاحية: البيانات الضخمة، التعلم العميق، التعرف على النشاط البشري، انترنت الأشياء، دمج البيانات، المشفر التلقائي، المحول.

Improving the Accuracy of Human Activity Recognition by using Autoencoder, Lstm and Transformer

ABSTRACT

IoT is well-known to be one of the major sources of big data, as it is based on connecting a huge number of smart devices to the Internet to report their frequently captured status of their environments. Recognizing and extracting meaningful patterns from enormous raw input data is the core utility of big data analytics as it results in higher levels of insights for decision-making and trend prediction. The general definition of big data to characterize the IoT big data through the following “3V’s” features: Volume, Velocity and Variety. Generally, big data comes in different forms and types from different sources. Data fusion and sharing play a critical role in developing ubiquitous environments based on IoT data. This role is more critical for time-sensitive IoT applications where a timely fusion of data is needed to bring all pieces of data together for analysis and consequently providing reliable and accurate actionable insights. This study aims to review the research that addressed the variety characteristic (Data Fusion) for Big data in Human Activity Recognition applications as one of the Internet of Things environments. with the aim of evaluating it and moving towards making a model that notes the variety of data

sources and their different formats. The idea of the proposed model is based on extracting rich features from The largest source data in terms of dimensional size by making an *autoencoder* model and training it with the aim of using it later as a transfer learning technique. Extracting the temporal correlations from each source using the *Lstm* network and finally Fusion the extracted features using the *Cross Attention* technique. Then analyze the results from the point of view of a set of performance parameters such as Accuracy, F1 Score, and Confusion Matrix. The results indicate that the performance of the proposed model achieved 2% higher accuracy than studies that used the same data sources from the same UTD-MHAD dataset.

Keywords: Big Data, Deep Learning, Human Activity Recognition, Internet of things, Data Fusion, Autoencoder, Lstm, Cross Attention, Transformer.

1. مقدمة:

يمكن لبيانات انترنت الأشياء أن تتدفق باستمرار Streaming أو أن يتم تجميعها Accumulating لتكون مصدر للبيانات الضخمة. يُقصد بالبيانات المتدفقة البيانات التي يتم إنشاؤها أو التقاطها خلال فترات زمنية صغيرة وتحتاج إلى تحليلها بسرعة لاستخراج رؤى فورية و/أو اتخاذ قرارات سريعة. يُقصد بالبيانات الضخمة مجموعات بيانات هائلة لا تستطيع منصات البرمجيات التقليدية والعتاد الصلب التقليدي تخزينها وإدارتها ومعالجتها وتحليلها. وينبغي التعامل مع هذين النهجين بشكل مختلف لأن متطلباتهما لناحية الاستجابة التحليلية مختلفة. حيث يمكن تسليم الرؤى من تحليلات البيانات الضخمة بعد عدة أيام من توليد البيانات، ولكن الرؤى من تحليلات البيانات المتدفقة يجب أن تكون جاهزة خلال مدة تتراوح بين بضع مئات من الميلي ثانية إلى بضع ثوان. تُعتبر خاصية تنوع تنسيقات ومصادر البيانات أحد أهم صفات البيانات الضخمة، ويلعب دمج هذه البيانات Data Fusion ومشاركتها دوراً حاسماً في تطوير بيئات واسعة الانتشار معتمدة على بيانات انترنت الأشياء المولدة للبيانات الضخمة. يعد هذا الدور أكثر أهمية بالنسبة إلى تطبيقات انترنت الأشياء الحساسة للوقت حيث يلزم دمج البيانات في الوقت المناسب لجلب كل أجزاء البيانات معاً للتحليل وبالتالي توفير رؤى موثوقة ودقيقة وقابلة للتنفيذ.

2. الهدف من البحث:

تهدف هذه الدراسة إلى استعراض الأبحاث التي عالجت خاصية التنوع (دمج البيانات) للبيانات الضخمة في تطبيقات التعرف على النشاط البشري كأحد بيئات انترنت الأشياء، بغرض تقييمها والانطلاق نحو بناء نموذج يلحظ تنوع مصادر البيانات واختلاف تنسيقها، حيث تقوم فكرة النموذج المقترح على استخراج الميزات الغنية من بيانات المصدر الأكبر لناحية حجم الأبعاد عبر بناء نموذج مشفر تلقائي Autoencoder وتدريبه و استخدامه لاحقاً كتقنية لنقل التعلم Transfer Learning في النموذج

الأساسي، واستخراج الارتباطات الزمنية من بيانات كل مصدر باستخدام شبكة Lstm وأخيراً دمج الميزات المستخرجة من مصادر البيانات بواسطة تقنية Cross Attention. من ثم تحليل النتائج من وجهة نظر مجموعة من معايير الأداء مثل الدقة Accuracy، F1 Score، مصفوفة التشويش Confusion Matrix.

3. خصائص بيانات انترنت الأشياء ومتطلبات تحليلها:

1.3 بيانات انترنت الأشياء السريعة والمتدفقة IoT fast and streaming data:

اقترحت العديد من الأبحاث تحليلات البيانات المتدفقة التي يمكن نشرها بشكل أساسي على أنظمة حوسبة عالية الأداء أو المنصات السحابية. تعتمد تحليلات البيانات المتدفقة في مثل هذه الأطر على المعالجة التفرعية والتراكمية [13]. من خلال تفرعية البيانات، يتم تقسيم مجموعة البيانات الكبيرة إلى عدة مجموعات من بيانات أصغر، حيث يتم تنفيذ التحليل على التفرع في وقت واحد. أما مصطلح المعالجة التراكمية يشير إلى جلب مجموعة صغيرة small batch من البيانات لتتم معالجتها بسرعة ضمن سلسلة pipeline من المهام الحسابية. على الرغم من أن هذه التقنيات تقلل من العطالة الزمنية لتحقيق استجابة من إطار تحليل البيانات المتدفقة، إلا أنها ليست الحل الأفضل الممكن لتطبيقات انترنت الأشياء ذات الوقت المحدود. من خلال تقريب مكان تحليل البيانات المتدفقة من مصدر البيانات (أي أجهزة انترنت الأشياء أو الأجهزة الطرفية)، تصبح الحاجة إلى تفرعية البيانات والمعالجة التراكمية أقل منطقية لأن حجم البيانات في المصدر يسمح بمعالجتها بسرعة. ومع ذلك، فإن تقديم تحليلات سريعة على أجهزة انترنت الأشياء يطرح تحدياته الخاصة مثل محدودية موارد الحوسبة والتخزين والطاقة في مصدر البيانات.

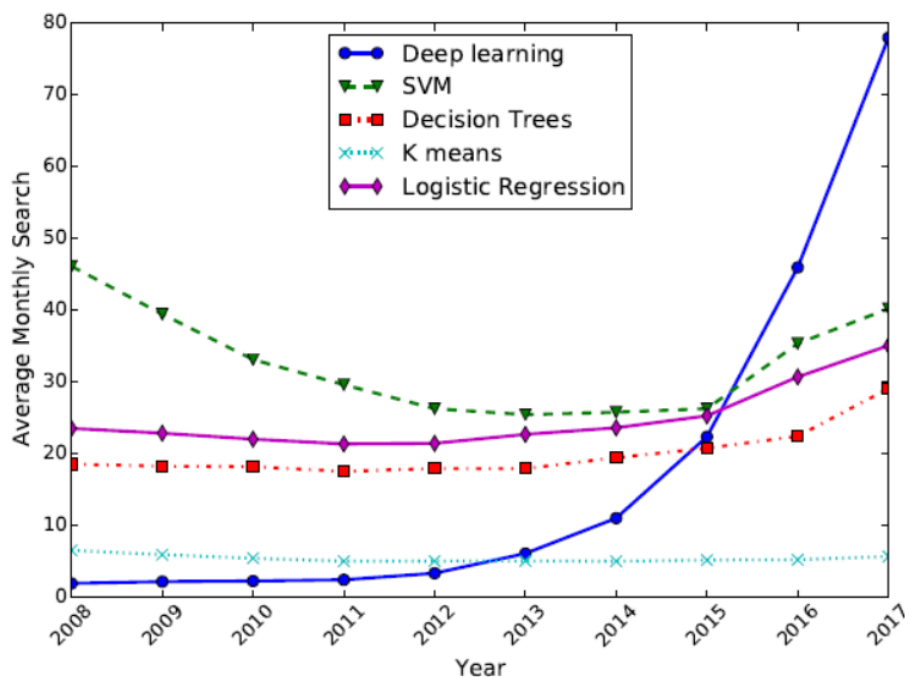
2.3 بيانات انترنت الأشياء الضخمة IoT Big Data:

من المعروف أن انترنت الأشياء هو أحد المصادر الرئيسية للبيانات الضخمة، لأنه يعتمد على ربط عدد كبير من الأجهزة الذكية بالإنترنت للإبلاغ عن حالة بياناتها التي يتم التقاطها بشكل متكرر. يعد التعرف على الأنماط ذات المعنى واستخراجها من بيانات المدخلات الخام الهائلة بمثابة الأداة الأساسية لتحليل البيانات الضخمة لأنها تزود بمستويات أعلى من الرؤى لاتخاذ القرارات والتنبؤ بالنزعات والاتجاهات Trends. ولذلك فإن استخلاص هذه الأفكار والمعرفة من البيانات الضخمة له أهمية بالغة للعديد من الشركات، لأنه يمكنها من الحصول على مزايا تنافسية. وصفت العديد من الأعمال البحثية الصفات العامة للبيانات الضخمة [14-17] من حيث الحجم والسرعة والتنوع، لذلك سنستخدم التعريف العام للبيانات الضخمة لوصف البيانات الضخمة لإنترنت الأشياء من خلال ميزات "3Vs" التالية:

- ✓ الحجم Volume: أصبحت كمية البيانات التي يتم إنشاؤها باستخدام أجهزة انترنت الأشياء أكبر بكثير من ذي قبل وتتناسب بشكل واضح مع هذه الصفة.
- ✓ السرعة Velocity: معدل إنتاج ومعالجة البيانات الضخمة في انترنت الأشياء مرتفع بما يكفي لعدم توفر البيانات الضخمة في الزمن الحقيقي.
- ✓ التنوع Variety: بشكل عام، تأتي البيانات الضخمة بأشكال وأنواع مختلفة. وقد تتكون من بيانات منظمة structures وشبه منظمة semi-structured وغير منظمة unstructured. يمكن أن تولد انترنت الأشياء مجموعة واسعة من أنواع البيانات مثل النصوص والصوت والفيديو وبيانات الحساسات (أرقام سلاسل زمنية) واردة من مصادر مختلفة.

4. التعلم العميق Deep Learning :

يتكون التعلم العميق من تقنيات التعلم الخاضعة للإشراف supervised وغير الخاضعة للإشراف unsupervised والتي تعتمد على طبقات عديدة من الشبكات العصبية الاصطناعية ANN القادرة على تعلم التمثيلات الهرمية في البنى العميقة. تتكون بنى التعلم العميق من طبقات معالجة متعددة. كل طبقة قادرة على توليد استجابات غير خطية بناءً على البيانات الواردة من طبقة الإدخال الخاصة بها. يوضح الشكل 1 اتجاه البحث لخمس خوارزميات شائعة للتعلم الآلي ضمن عمل Google.



الشكل 1: اتجاه البحث لخمس خوارزميات شائعة للتعلم الآلي ضمن عمل Google.

[23]

1.4 بنى التعلم العميق :Deep Learning Architectures

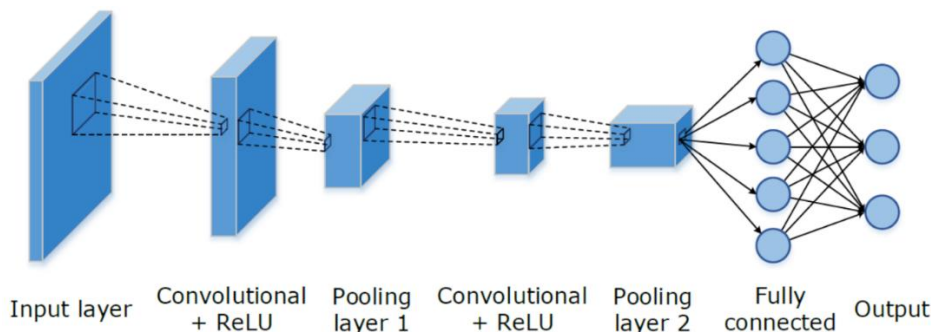
يلخص الجدول التالي هذه البنى وخصائصها ونوع البيانات التي تتعامل معه.

الجدول 1: نماذج التعلم العميق [23]

م	النموذج	الفئة	نمط التعلم	بيانات الدخل
1	AE	توليدي	غير خاضع للإشراف	متنوعة
2	RNN	تمميبيزي	خاضع للإشراف	سلاسل زمنية، بيانات متسلسلة
3	LSTM	تمميبيزي	خاضع للإشراف	سلاسل زمنية، بيانات متسلسلة
4	CNN	تمميبيزي	خاضع للإشراف	صورة، صوت
5	VAE	توليدي	شبه خاضع للإشراف	متنوعة
6	GAN	هجين	شبه خاضع للإشراف	متنوعة
7	Transformer	تمميبيزي	خاضع للإشراف	متنوعة

1.1.4 الشبكات العصبية التلافيفية CNN (Convolutional Neural Network)

بالنسبة للمسائل القائمة على الرؤية Vision، تستقبل شبكة CNN مدخلات ثنائية الأبعاد على سبيل المثال صورة أو صوت وتستخرج ميزات عالية المستوى من خلال سلسلة من الطبقات المخفية. تتكون الطبقات المخفية من طبقات تلافيفية بالإضافة إلى طبقات متصلة بالكامل FC: Full Connected في النهاية . يصور الشكل 2 بنية الشبكة CNN:



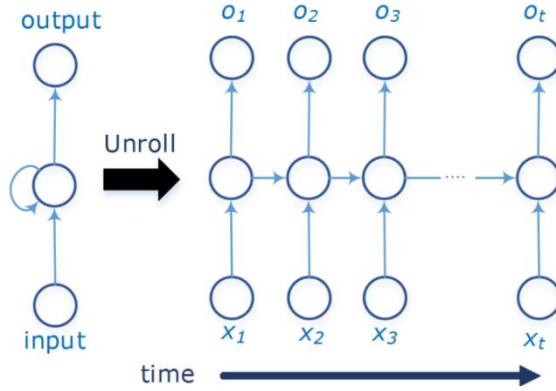
الشكل 2: بنية الشبكة CNN [23]

تم تجهيز العديد من أجهزة إنترنت الأشياء IoT، مثل الطائرات بدون طيار drones والهواتف الذكية والسيارات الذكية المتصلة بكاميرات. تم اختبار بنية CNN وتغيراتها في مجموعة متنوعة من التطبيقات التي تتضمن هذه الأجهزة. تشمل بعض التطبيقات النموذجية التنبؤ بالفيضانات أو الزلازل من خلال صور الطائرات بدون طيار، والكشف عن أمراض النبات باستخدام صور النباتات على الهواتف الذكية، والكشف عن إشارات المرور باستخدام كاميرات المراقبة.

2.1.4 الشبكات العصبية التكرارية (Recurrent Neural Network) RNN :

في العديد من المهام يعتمد التنبؤ على عدة عينات سابقة بحيث نحتاج أيضاً بالإضافة إلى تصنيف العينات الفردية إلى تحليل تسلسل المدخلات. في مثل هذه التطبيقات لا تكون الشبكة العصبية ذات التغذية الأمامية FFN قابلة للتطبيق لأنها لا تفترض أي تبعية بين طبقات الإدخال والإخراج. تم تطوير شبكة RNN لمعالجة هذه المشكلة بشكل تسلسلي (كالكلام أو النص) أو مشكلات السلاسل الزمنية (بيانات الحساسات) بأطوال مختلفة. يعد اكتشاف سلوكيات السائقين في المركبات الذكية، وتحديد أنماط حركة الأشخاص وتقدير استهلاك الطاقة لأسرة من بعض الأمثلة التي يمكن تطبيق شبكات

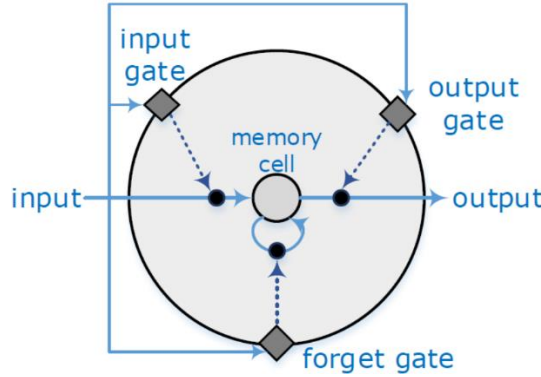
RNN عليها. يتكون الإدخال لشبكة RNN من العينة الحالية والعينة السابقة المشاهدة بمعنى آخر يؤثر خرج RNN في الخطوة الزمنية $t-1$ على الإخراج في الخطوة الزمنية t . يصور الشكل 3 بنية RNN:



الشكل 3: بنية RNN [23]

3.1.4 شبكة الذاكرة طويلة قصيرة المدى (Long Short –Term Memory) (LSTM):

تعتبر LSTM امتداد لشبكات RNN. على الرغم من انه تم اقتراح أشكال مختلفة من LSTM إلا أن معظمها اتبع نفس تصميم الشبكة الأصلية [18]. تستخدم LSTM مفهوم البوابات في وحداتها، حيث تحسب كل منها قيمة بين 0 و 1 بالاعتماد على دخلها. بالإضافة إلى حلقة التغذية الراجعة لتخزين المعلومات، تحتوي كل خلية عصبية في LSTM (وتسمى أيضاً خلية الذاكرة) على بوابة نسيان وبوابة قراءة وبوابة كتابة.



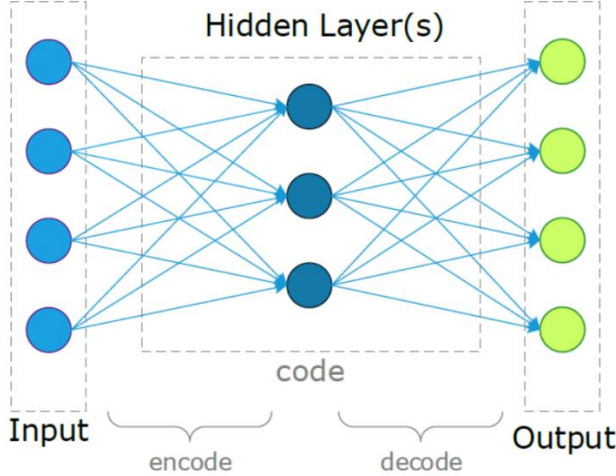
الشكل 4: بنية خلية الذاكرة LSTM. تُظهر الخطوط المستمرة تدفق البيانات، والخطوط المتقطعة الإشارات القادمة من البوابات [23]

أحد الاختلافات المهمة بين شبكة LSTM وشبكة RNN هو أن وحدات LSTM تستخدم بوابات النسيان للتحكم بشكل فعال في حالات الخلية والتأكد من عدم تدهورها. عندما تتميز البيانات بالتبعية الزمنية الطويلة (وجود ترابط زمني)، فإن نماذج LSTM تعمل بشكل أفضل من نماذج RNN [19]. ويمكن ملاحظة طول التبعية الزمنية في تطبيقات انترنت الأشياء مثل التعرف على النشاط البشري Human Activity Recognition، والتنبؤ بالأداء التعليمي في برامج الانترنت والتنبؤ بالكوارث بناءً على المراقبة البيئية، على سبيل المثال لا الحصر.

4.1.4 أجهزة التشفير التلقائي AEs Autoencoders :

تتكون أجهزة التشفير التلقائي AEs من طبقة إدخال وطبقة إخراج متصلتين من خلال طبقة مخفية واحدة أو أكثر. تمتلك AEs نفس عدد وحدات الإدخال والإخراج. تهدف هذه الشبكة إلى إعادة بناء المدخلات عن طريق تحويل المدخلات إلى مخرجات بأبسط طريقة ممكنة، بحيث لا تتشوه المدخلات بشكل كبير. تم استخدام هذه النوع من الشبكات

العصبية بشكل أساسي في حل مشكلات التعلم غير الخاضع للإشراف وكذلك نقل التعلم Transfer Learning [20]. نظراً إلى سلوكها في بناء المدخلات في طبقة الإخراج، تُستخدم AEs بشكل أساسي في مهام التشخيص واكتشاف الأخطاء. وهذا أمر ذو أهمية كبيرة لإنترنت الأشياء الصناعي Industrial IoT لخدمة العديد من التطبيقات مثل تشخيص الأخطاء في الأجهزة والآلات، والكشف عن الشذوذ في أداء خطوط التجميع. تتضمن AEs مكونين رئيسيين: جهاز التشفير Encoder وجهاز فك التشفير Decoder. يستقبل المشفر المدخلات ويحولها إلى تمثيل جديد، والذي يسمى عادة بالـ Code أو المتغير الكامن Latent variable. يستقبل جهاز فك التشفير الكود الذي يتم إنشاؤه في جهاز التشفير، ويحوّله إلى شكل معاد بناؤه من المدخلات الأصلية. تتضمن عملية التدريب في AEs تقليل أخطاء إعادة البناء، أي أن المخرجات والمدخلات تظهر الحد الأدنى من الاختلاف. يوضح الشكل 5 بنية المشفر التلقائي AE النموذجية.



الشكل 5 : بنية المشفر التلقائي النموذجية AE [23]

يوجد العديد من الاختلافات والامتدادات لـ AEs مثل:

- Stacked AE

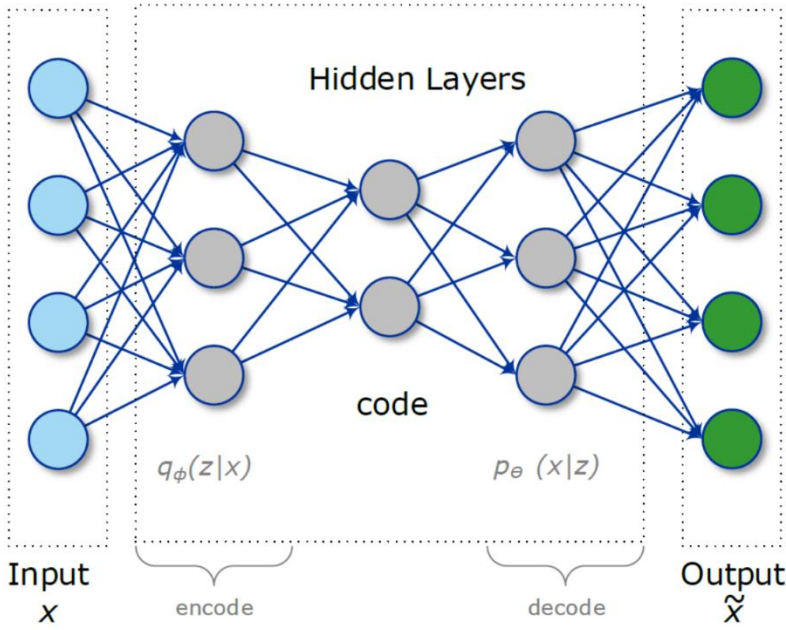
- Sprase AE

- Variational AE

5.1.4 أجهزة التشفير التلقائي المتغيرة Variational AE :

يعد نموذج VAE، والذي تم تقديمه عام 2013، إطاراً توليدياً شائعاً ذي افتراضات ضعيفة حول بنية البيانات، بينما تتميز بعملية تدريب سريعة من خلال الانتشار العكسي [21]. علاوة على ذلك تم استخدام هذا النموذج للتعلم شبه الخاضع للإشراف [22]. ولذلك فهو مناسب تماماً لحلول انترنت الأشياء IoT التي تتعامل مع البيانات المتنوعة وندرة البيانات المصنفة. وتتضمن هذه التطبيقات اكتشاف الفشل في مستويات الاستشعار أو التشغيل وكشف الاختراق في أنظمة الحماية. لكل نقطة بيانات X ، يوجد شعاع للمتغيرات الكامنة المقابلة التي يشار إليها بـ Z .

تتكون بنية التدريب الخاصة بالمشفر التلقائي المتغير VAES من جهاز تشفير وجهاز فك تشفير مع وسطاء \square و Θ على الترتيب. يساعد النموذج ذو الشكل الثابت $q(Z|X)$ المشفر في تقدير التوزيع السابق $p(Z|X)$. يتكون النموذج من شبكتين: إحداها تولد العينات والأخرى تقوم بالاستدلال التقريبي. يمثل الشكل 6 بنية المشفر التلقائي المتغير VAE.

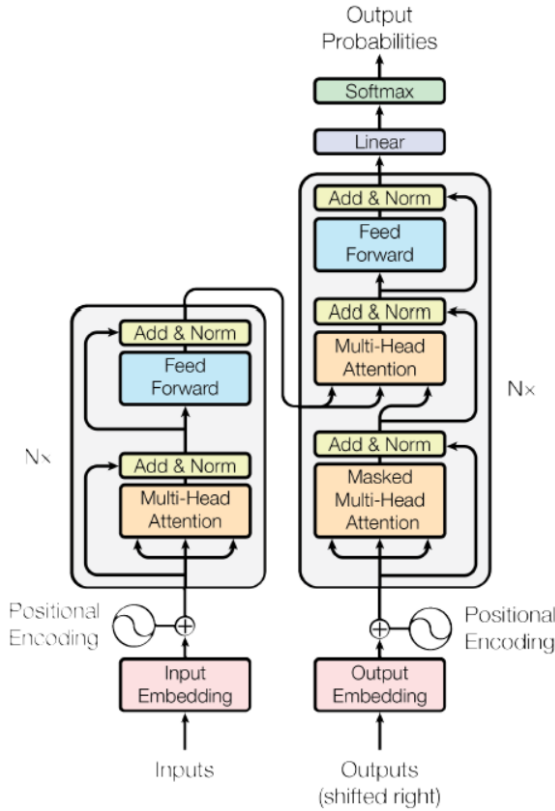


الشكل 6 : بنية المشفر التلقائي المتغير VAE [23]

6.1.4 المحول Transformer :

المحول هو نموذج للتعلم العميق يستخدم بشكل أساسي آلية الاهتمام الذاتي Self-attention [26]. تم تصميم المحولات للتعامل مع بيانات الإدخال (مثل النص والصور) لمهام مثل الترجمة والتلخيص والتنبؤ والتصنيف. على عكس شبكات LSTM و RNN لا تقوم المحولات بمعالجة البيانات بشكل تسلسلي في مرحلة الإدخال. بدلاً من ذلك، يقوم المحول بمعالجة المدخلات كلها في سلسلة الإدخال (سلسلة كلمات مكونة لجملة مثلاً أو سلسلة صور مكونة لمقطع فيديو) بنفس الوقت وعلى التفرع، ويحدد علاقة كل جزء من السلسلة (في خطوة زمنية ما) بباقي عناصر السلسلة (باقي الخطوات الزمنية)، مما يمنح بعد ذلك معنى لكل كلمة في الجملة

[27]. تتبنى المحولات بنية التشفير وفك التشفير. تتمثل وظيفة طبقات التشفير في إنشاء ترميزات encodings تحتوي على معلومات حول أجزاء المدخلات ذات الصلة ببعضها البعض. تقوم طبقة فك التشفير بعمل معاكس، حيث تأخذ جميع الترميزات وتستخدم معلوماتها السياقية لإنشاء تسلسل خرج منظم. ولتحقيق ذلك، تستخدم طبقة التشفير آلية الاهتمام الذاتي Self-attention ووحدة فك تشفير آلية الاهتمام المتبادل Cross-Attention الذي يتم تطبيقه بين سلسلة الدخل بعد خضوعها للاهتمام الذاتي وبين سلسلة الخرج.



الشكل 7 : بنية المحول Transformer

5. التعرف على النشاط البشري كأحد تطبيقات التعلم العميق على البيانات الضخمة المولدة في بيئة انترنت الأشياء:

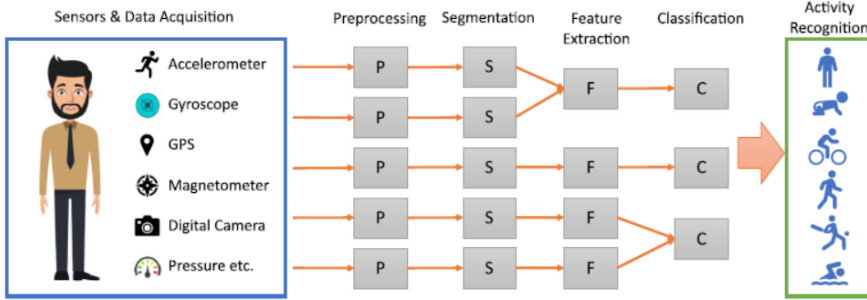
1.5 التعلم العميق في البيانات الضخمة Deep Learning for Big Data :

يتعامل التعلم العميق بشكل أساسي مع اثنتين من خصائص البيانات الضخمة هما الحجم Volume والتنوع Variety أي أن خوارزميات التعلم العميق مناسبة لاستخراج المعرفة المفيدة من كميات هائلة من البيانات ومن البيانات التي تم جمعها من مصادر مختلفة بتنسيقات مختلفة [2] .

2.5 التعرف على النشاط البشري:

يلعب التعرف على النشاط البشري دوراً مهماً في الحياة اليومية للأفراد، نظراً لكفاءته في تكوين

معرفة عميقة وعالية المستوى حول النشاط البشري باستخدام البيانات الخام الواردة من الحساسات [24] . تشمل تطبيقات التعرف على النشاط البشري تحليل السلوك المنزلي والتعرف على الإيماءات.. الخ.



الشكل 8: مراحل عمل نظام التعرف المعتمد على عدة مصادر للبيانات [6]

سهل التطور السريع في أجهزة الموبايل والأجهزة القابلة للارتداء تطبيق تقنيات انترنت الأشياء (IoT) في مجال الرعاية الصحية من خلال مراقبة النشاط البشري (Human Activity Recognition)، عبر قيام هذه الحساسات بجمع وترحيل أنواع مختلفة من

البيانات (Variety) بكميات كبيرة (Volume) وفي الزمن الحقيقي (Velocity). إذاً يمكن اعتبار تطبيقات التعرف على النشاط البشري HAR أحد مجالات توليد البيانات الضخمة [5]. يمكن تقسيم أشكال الحساسات المستخدمة في نظم التعرف إلى أربعة أنواع [3]:

1. قابلة للارتداء Wearable Sensors: مقياس التسارع، الجيروسكوب، مقياس المغنطة.
2. حساسات محيطية Ambient Sensors: مثل Camera، WiFi، RFID.
3. حساسات الأغراض Object Sensors: تزود بمعلومات عن استخدام الأشخاص للأغراض مثل كوب مياه، صنوبر.. الخ.
4. حساسات أخرى: مثل مسجلات الصوت Audio Sensors الموجودة في هواتف المستخدمين.

3.5 خاصية التنوع Variety للبيانات الضخمة في نظم التعرف على النشاط HAR وطرق الدمج المستخدمة:

سنقوم باستعراض الأبحاث التي عالجت أحد خواص البيانات الضخمة وهي التنوع Variety في نظم التعرف على النشاط (والتي تعتبر بيئة انترنت أشياء تولد بيانات تتسم بصفات البيانات الضخمة)، تم التطرق بشكل عام في الأدبيات إلى ثلاثة أصناف للتعرف على النشاط:

- 1- باستخدام الحساسات القابلة للارتداء (البيانات الخام المولدة على شكل أرقام).
 - 2- باستخدام حساسات خارجية مرتبطة بالبيئة المحيطة (البيانات الخام على شكل صور مثلاً).
 - 3- باستخدام الجمع بين النوعين السابقين: وهو ما سنركز عليه باعتباره يتضمن توليد أشكال مختلفة من البيانات (أرقام و صور و صوت ..). حيث تم استخدام تقنيتين في دمج البيانات مختلفة التنسيق والمولدة من عدة مصادر كما في الشكل 9:
- دمج على مستوى الميزات (Early Fusion): حيث يتم دمج الميزات بعد استخراجها من البيانات الخام، وتدريبها معاً باستخدام مصنف واحد [8]. تتضمن طرق الدمج:

التكديس/التسلسل Concatenation، الجمع Sum، الارتباط القويم Canonical Correlation. يساعد هذا النوع في جمع ميزات متزامنة من الحساسات المختلفة ودمجها لتوليد معلومات كافية لاتخاذ قرار قوي.

دمج على مستوى القرار (*Late Fusion*): يتم عبر دمج القرارات الناتجة عن المصنفات المدربة بشكل مستقل. على الرغم من أن هذا النوع من الدمج يضمن عدم سيطرة بيانات حساس على حساب حساس آخر، لكن تكمن سلبيته في عدم أخذ الارتباطات بين بيانات المستشعرات بعين الاعتبار. يترافق هذا النوع بزمن تدريب طويل وكلف حسابية عالية [8].

ويتم عبر طرق الجمع، الجداء، القيمة العظمى، كما يلي:

- *Sum rule:*

$$c = \operatorname{argmax}_j \sum_{i=1}^n P(\omega_j | \mathbf{x}_i).$$

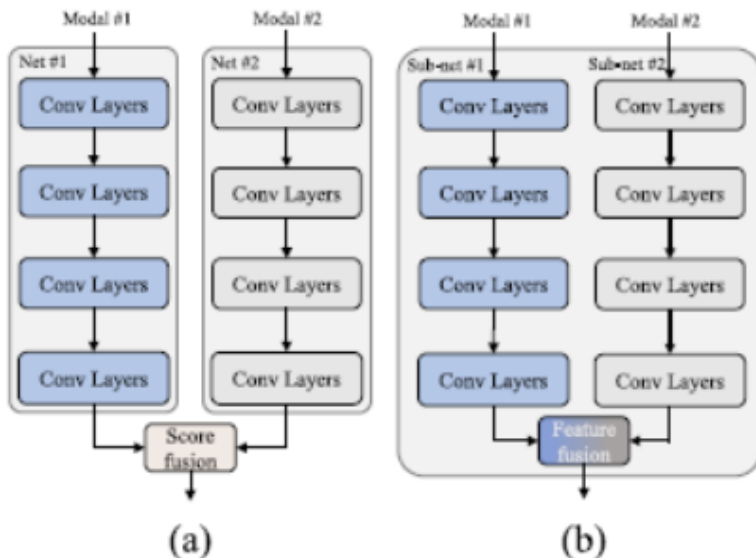
- *Product rule:*

$$c = \operatorname{argmax}_j \prod_{i=1}^n P(\omega_j | \mathbf{x}_i).$$

- *Max rule:*

$$c = \operatorname{argmax}_j \max_i P(\omega_j | \mathbf{x}_i),$$

where ω_j denotes the j -th class.

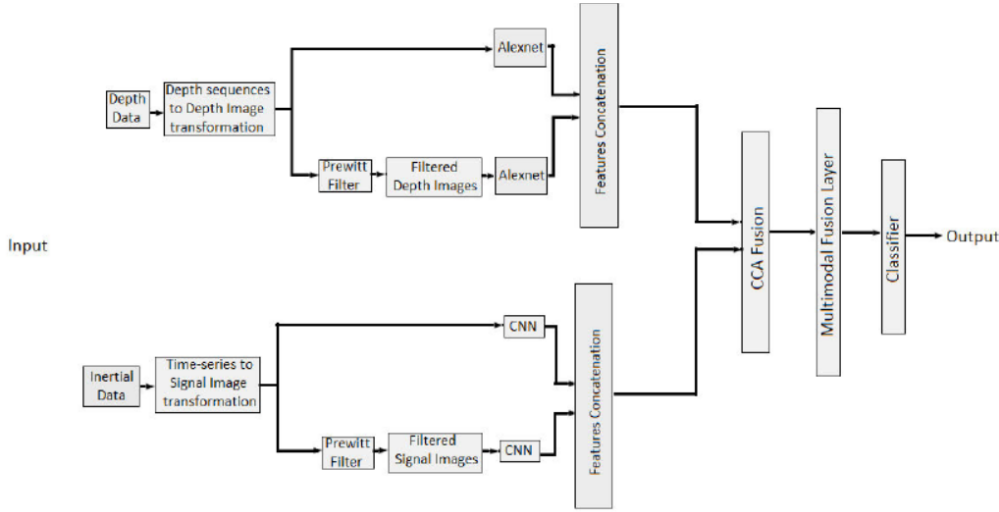


الشكل 9: الدمج المتأخر (a)، ودمج الميزات (b) [4]

4.5 دمج البيانات في أدبيات نظم التعرف على النشاط البشري HAR:

تم التطرق إلى دمج البيانات الواردة من عدة مصادر للتعرف على النشاط البشري- كبيئة انترنت أشياء مولدة للبيانات الضخمة- في عدة أبحاث، تم تلخيصها كما يلي:

استخدم الباحثون في الدراسة [7] كاميرات Kinect والهواتف الذكية لتوفير قراءات حساسات مقياس التسارع والجايروسكوب، حيث أوضحوا أن استخدام طريقة دمج الميزات لوحدها أو الدمج المتأخر لوحده ينطوي على عدة سلبيات واقترحوا استخدام كلا النوعين معاً عبر تطوير وتحقيق عدة بنى بالاستفادة والاعتماد كلياً على بنية CNN.



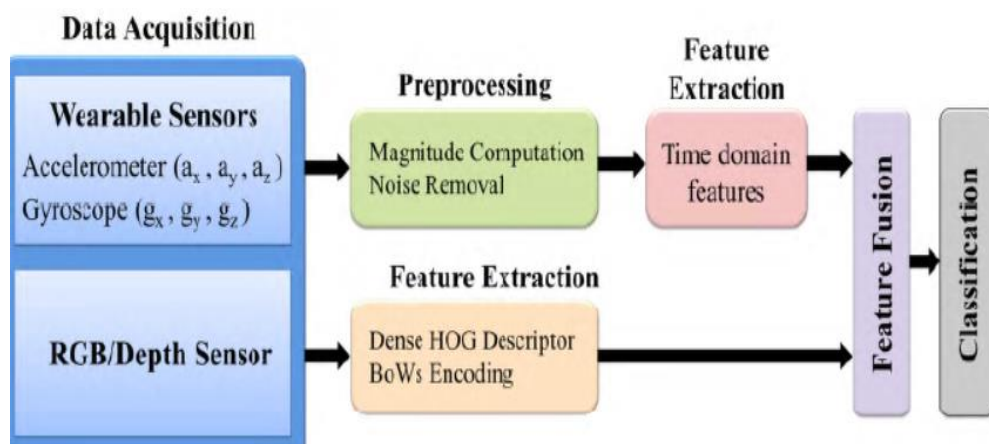
الشكل 10: دمج ميزات عميق متعدد المراحل [7]

نقد الدراسة [7]: تحويل البيانات القادمة على شكل أرقام في سلاسل زمنية إلى صور، يكسر الحدود الطبيعية بين العينات (أحادية البعد) المتتالية ويجعلها تضيق في الصور (ثنائية البعد). حيث أوضح الباحثون في [12] أثناء تجربتهم نقل المتجهات المتعاقبة من شكل أحادي البعد 1D إلى صور 2D إلى تدهور أداء النموذج الخاص بهم. تم حصر البنى الثلاث المقترحة باستخدام الخوارزمية CNN وهي مناسبة لبيانات الصور فقط.

في الدراسة [8] ركز الباحثون على الدمج على مستوى الميزة لطرائق الاستشعار المتعددة (الحساسات) حيث استخدموا بيانات الحساسات العطالية inertial sensor مع صور الكاميرات Depth، RGB من مجموعة البيانات UTD-MHAD. واعتمدوا طريقة التسلسل concatenation في دمج الميزات المستخرجة من كل من طرائق الاستشعار، حيث أن الميزات مستخرجة في نفس الوقت وتمثل نفس النشاط وحصلوا على شعاع

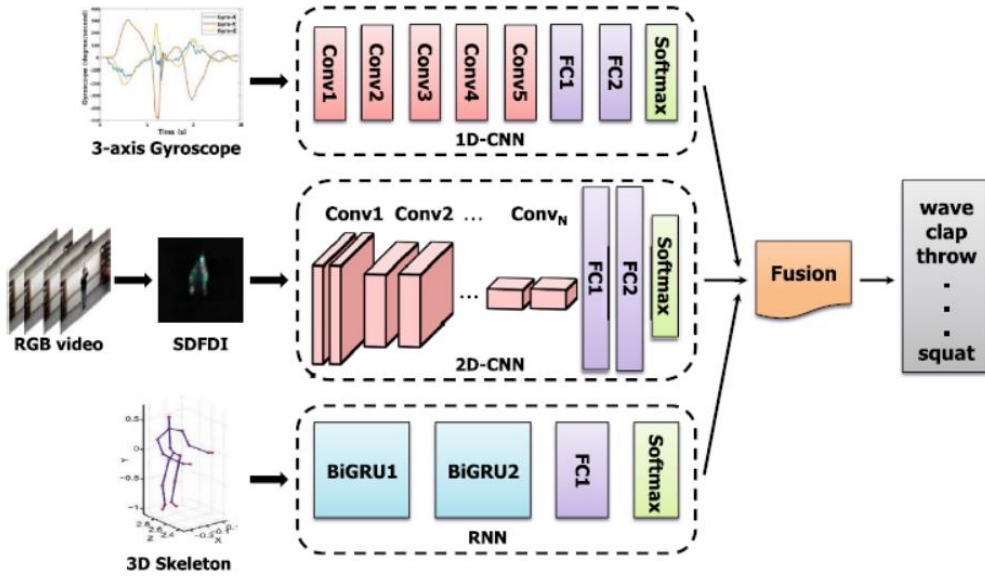
مميزات جديد بأبعاد عالية بحيث إذا كان المتجه $IMU[1*N]$ والصور $[1*k]$ يكون الناتج $[1*(N+k)]$

نقد الدراسة [8]: تتطوي طريقة Concatenation على زيادة أبعاد شعاع المميزات الداخل في عملية التدريب، كما تتضمن تكرار محتمل redundant بين متجهات المميزات.



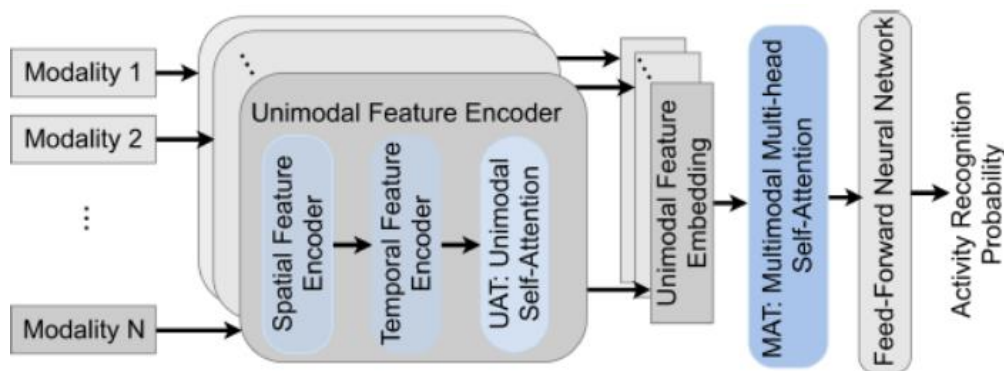
الشكل 11: بنية الدمج المقترحة [8]

في الدراسة [9] اقترح الباحثون بنية تتعامل مع ثلاثة تدفقات دخل للتعرف على النشاط، قارنوا من خلالها طرق دمج البيانات وهي الدمج المتأخر والدمج على مستوى الميزة، ولكنهم استبدلوا طريقة الدمج بالتكديس الأفقي Concatenation بطريقة إحصائية في دمج المميزات Canonical Correlation، والتي تقوم على فكرة استخراج المميزات الأكثر ارتباطاً بين شعاعين وتوليد شعاع واحد يتضمن الأزواج المرتبطة مع بعضها من شعاعي الدخل.

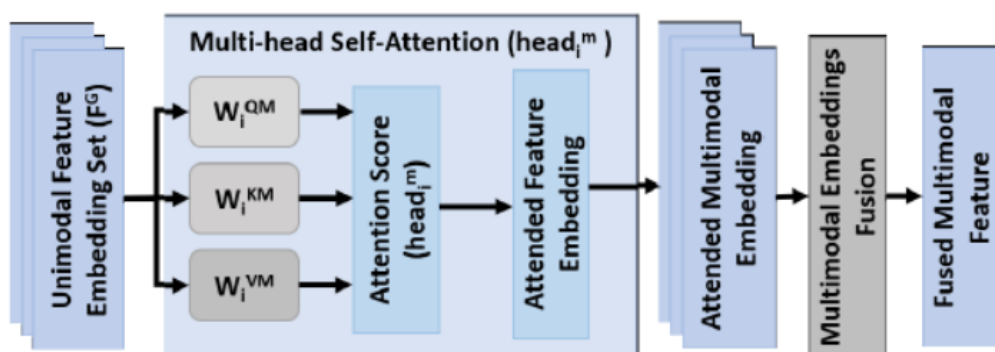


الشكل 12: بنية الدمج المقترحة [9]

أشار الباحثون في الدراسة [10] أن كل أساليب الدمج السابقة توفر وزناً موحداً للبيانات من جميع طرق الاستشعار المشاركة في التعرف على النشاط، بالرغم من أنه قد توفر إحدى طرق الاستشعار معلومات محسنة أكثر من طريقة أخرى. تقوم خوارزمية HAMLET المقترحة أولاً على استخراج السمات البارزة المكانية والزمانية لكل طريقة استشعار على حدى، ثم بعد ذلك توظف آلية انتباه جديدة تسمى Attention based Feature Fusion بهدف دمج الميزات الواردة من طرق الاستشعار المختلفة.



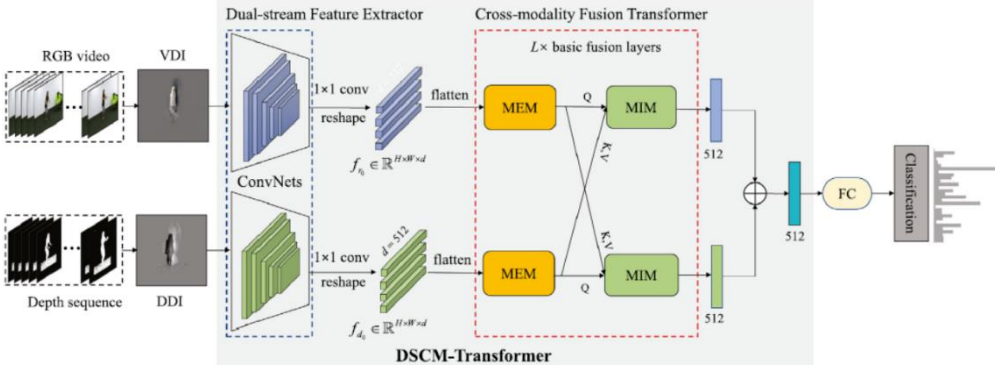
الشكل 13: بنية استخراج الميزات في HAMLET [10]



الشكل 14: بنية الدمج MAT [10]

نقد الدراسة [10]: تفترض بنية استخراج الميزات لكل طريقة استشعار، مرور البيانات في ممر pipeline حتمي يتضمن استخراج الميزات المكانية ثم استخراج الميزات الزمانية وتطبيق آلية الانتباه Attention. علماً أن بعض أنواع البيانات لا يحتاج استخراج الميزات منه تطبيق آلية التقاط الارتباطات المكانية مثل بيانات الحساسات العطالية Inertial sensors وبالتالي كلفة حسابية وزمنية بلا جدوى.

أشار الباحثون في الدراسة [11] إلى كثرة الأعمال التي استخدمت المحول Transformer في استخراج السمات وقلة الأعمال التي استخدمته في دمج الميزات. وبناءً عليه اقترحوا بنية دمج جديدة للسمات الواردة من عدة طرق استشعار بالاعتماد على المحول Transformer.



الشكل 15: بنية DSCMT [11]

نقد الدراسة [11]: محدودية عدد طرق الاستشعار الممكن الاستفادة منها حيث يتيح المحول Transformer عبر تقنيته cross attention دمج سلسلتي دخل فقط (مصدري دخل). بالإضافة إلى التقييد بنوع بيانات الدخل كصور عبر الاعتماد على بنية CNN في كلا طرفي التدفق.

نقاط الضعف في المحاولات السابقة لدمج بيانات من مصادر متعددة:

- استخدام الدمج المتأخر البسيط LF الذي يتجاهل العلاقات والتفاعلات بين بيانات الحساسات المختلفة.
- استخدام دمج الميزات في الطبقة الأخيرة وذلك باستخدام طرق بسيطة مثل التكديس Concatenation والمعدل AVG ثم تدريب ناتج هذه العمليات بشكل مشترك، يفتقر إلى القدرة على التقاط كامل الارتباطات بين بيانات طرق الاستشعار المختلفة.
- نلاحظ أن عدد محدود من الأبحاث تعامل مع تنسيقات هجينة لبيانات الدخل مثل السلاسل الزمنية (الأرقام) مع الفيديو أو السلاسل الزمنية (الأرقام) مع الصوت. والأبحاث التي تناولتها قامت بتحويل السلاسل الزمنية إلى صورة أولاً ثم استخراج الميزات من الصور وهذا يقود إلى انخفاض أداء النموذج بحسب [12].

- لم يتم التطرق إلى تعلم الميزات المضغوطة Compact والتمثيلية (المعبرة) representative وإدخالها لآلية الانتباه attention أو للمحول Transformer بشكل عام.

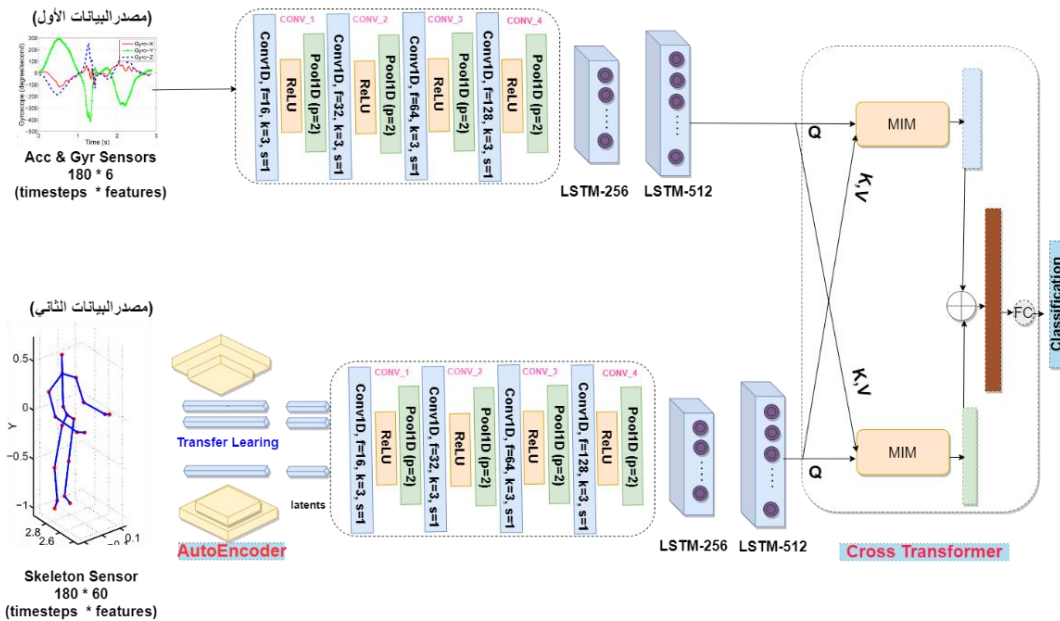
6. البنية المقترحة Autoencoder Transformer fusion (AETF) :

نقترح تعديل البنية الواردة في البحث [9] عبر إضافة المكونات التالية:

1- شبكة مشفر تلقائي Auto encoder [23]، وتطبيقها على مصدر البيانات الذي يتمتع بعدد ميزات أكبر، مما يمكننا من الاستفادة مع الميزات الغنية والأكثر تمثيلاً لهذا المصدر.

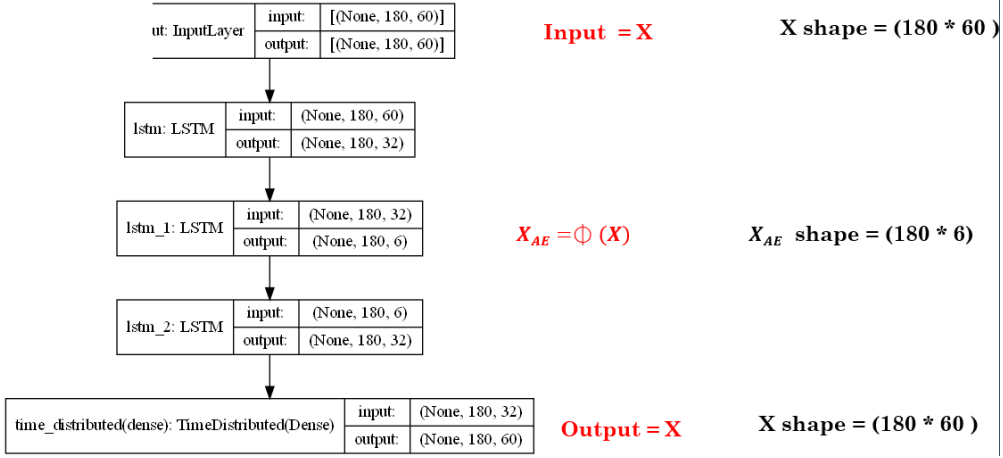
2- شبكة Lstm، وتطبيقها على مصدري البيانات. حيث تتميز هذه الشبكة بقدرتها العالية على التقاط الترابطات الزمنية بشكل فعال ضمن كل مصدر بيانات في حال وجودها [23].

3- استبدال عملية الجمع Concatenation المستخدمة في دمج الميزات بآلية Cross Attention التي أثبتت فعاليتها في استخراج الميزات [11]. يوضح المخطط التالي النموذج المقترح:



الشكل 16: بنية النموذج المقترح

تم تزويد النموذج بمصدرين للبيانات، الأول بيانات الحساس العطالي Inertial (التسارع والجيروسكوب)، والثاني بيانات حساس مفاصل الهيكل Skeleton. قمنا بضغط بيانات المصدر Skeleton ذي الميزات الأكبر عدداً ليصبح مساوي لعدد بيانات المصدر الثاني. تم ذلك عبر تدريب نموذج المشفر التلقائي LSTM Autoencoder ذي البنية التالية:



الشكل 17: نموذج المشفر التلقائي المدرب

حيث يتعبر ذلك توظيف لتقنية نقل التعلم Transfer Learning، عبر استخدام طبقة Latent من النموذج المدرب، كما يلي:

```

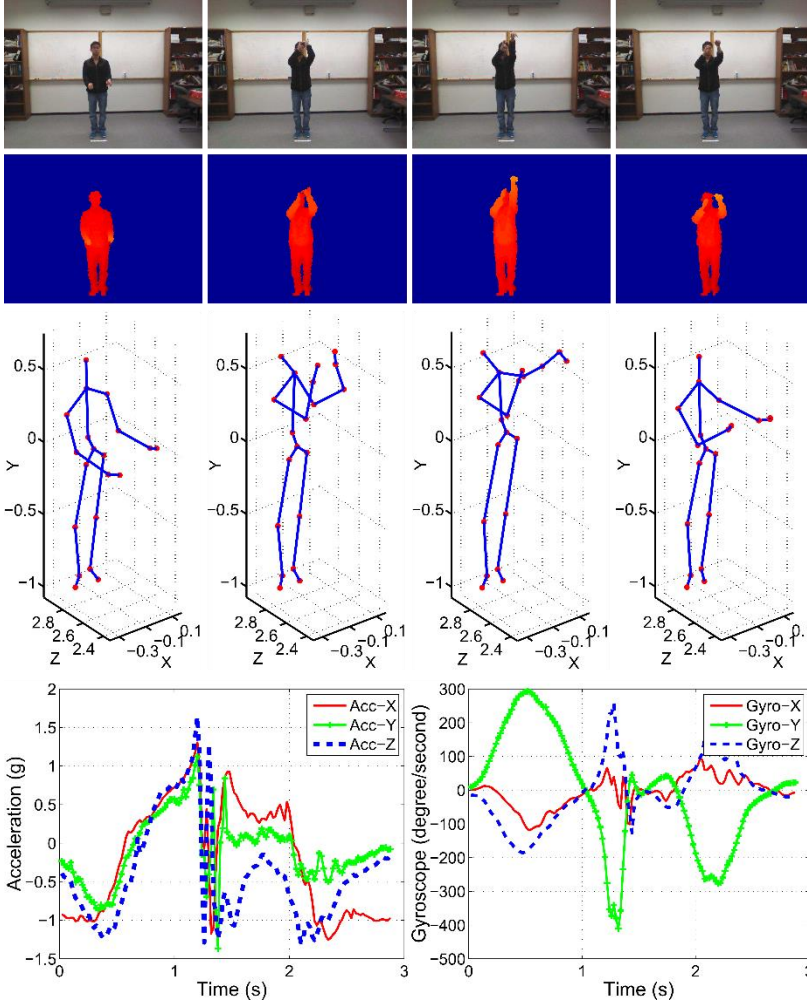
json_file = open('ae_skel_5_9_23.json', 'r')
loaded_model_json = json_file.read()
loaded_model = model_from_json(loaded_model_json)
 $X_{AE}$  = Model(inputs=loaded_model.inputs,
outputs=loaded_model.layers[1].output)

```

1.6 مجموعة البيانات المستخدمة:

تم تدريب النموذج المقترح باستخدام مجموعة بيانات UTD-MHAD [25] التي تم جمع بياناتها باستخدام جهاز Microsoft Kinect وحساس Acc & Inertail (Gyr).

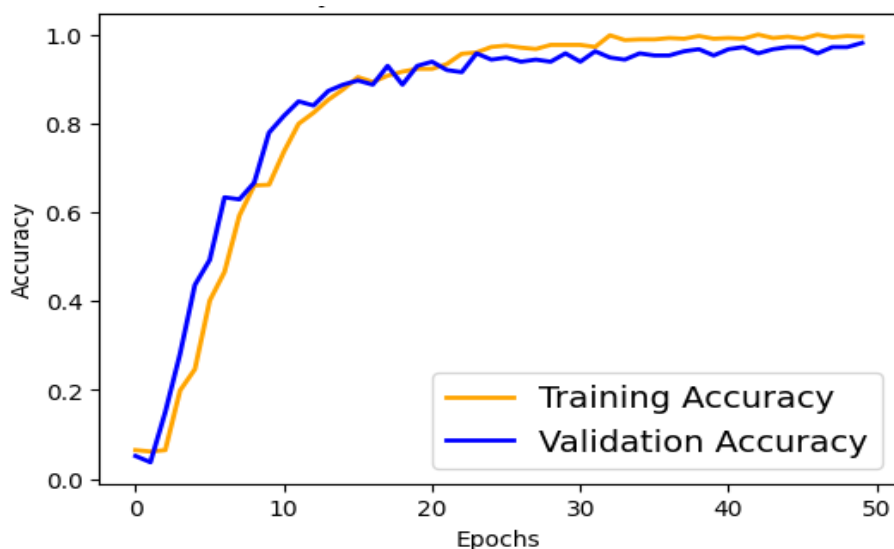
تتضمن المجموعة UTD-MHAD 27 نشاط مختلف قام به 8 أشخاص (4 ذكور و4 إناث)، حيث كررو كل نشاط 4 مرات (تجارب Trials). تتضمن هذه المجموعة 861 تسلسل بيانات. تم تسجيل أربعة أنواع بيانات Modalities هي مقاطع فيديو RGB ومقاطع فيديو العمق Depth ومواضع مفاصل الهيكل العظمي Skeleton وإشارات الحساس Inertail (مقياس التسارع ومقياس الجيروسكوب). ولمزامنة البيانات تم تسجيل طابع زمني لكل عينة. بالنسبة لكل تجربة Trial، تم تخزين بيانات الألوان في ملفات فيديو (.avi)، وتم تخزين بيانات حساس صور العمق والهيكل العظمي والحساس العطالي inertial باستخدام بيئة ماتلاب كثلاثة ملفات (.mat). ونتيجة لذلك تم تضمين أربعة ملفات لإجراء تجربة واحدة في مجموعة البيانات. تم تسمية الملفات بالصيغة التالية: "a_i_s_j_t_k_modality"، حيث يرمز a_i إلى النشاط رقم i (- 1: i 27)، و s_j إلى الشخص رقم j، (j: 1-8)، و t_k إلى التجربة رقم k (k: 1-4). تم اعتماد هذه الصيغة مع الأنواع الأربعة للبيانات (RGB, Depth, Skeleton, Inertail). قمنا بتقسيم مجموعة البيانات إلى مجموعة تدريب (80%) ومجموعة اختبار (20%). يوضح الشكل 18 جزء من مجموعة البيانات UTD-MHAD:



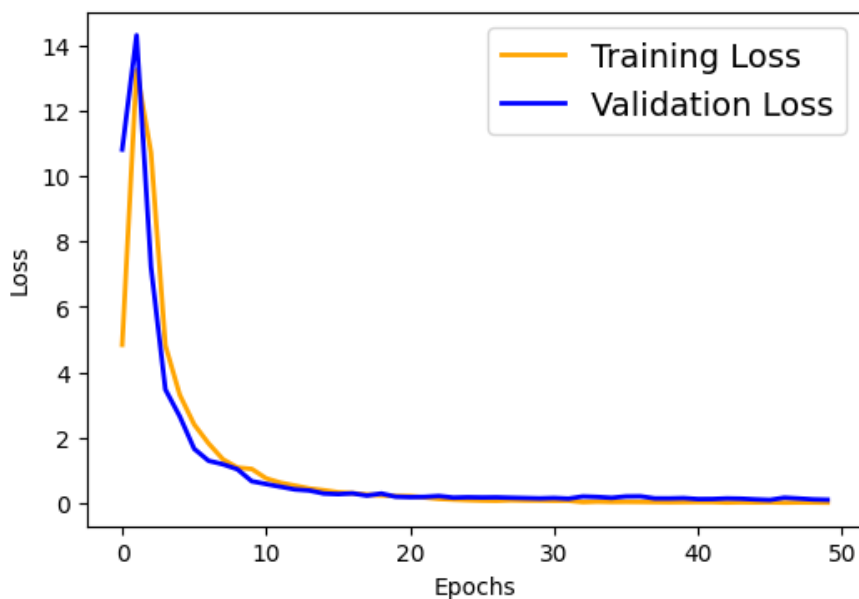
الشكل 18: مثال على بيانات الوسائط المتعددة المقابلة لنشاط كرة السلة. يعرض الصف الأول الصور الملونة، والصف الثاني صور العمق (تمت إزالة خلفية كل إطار عمق)، والصف الثالث إطارات مفاصل الهيكل العظمي، والصف الأخير بيانات مستشعر Inertial (إشارات التسارع والجيروسكوب) [25].

2.6 نتائج الاختبارات ومناقشتها:

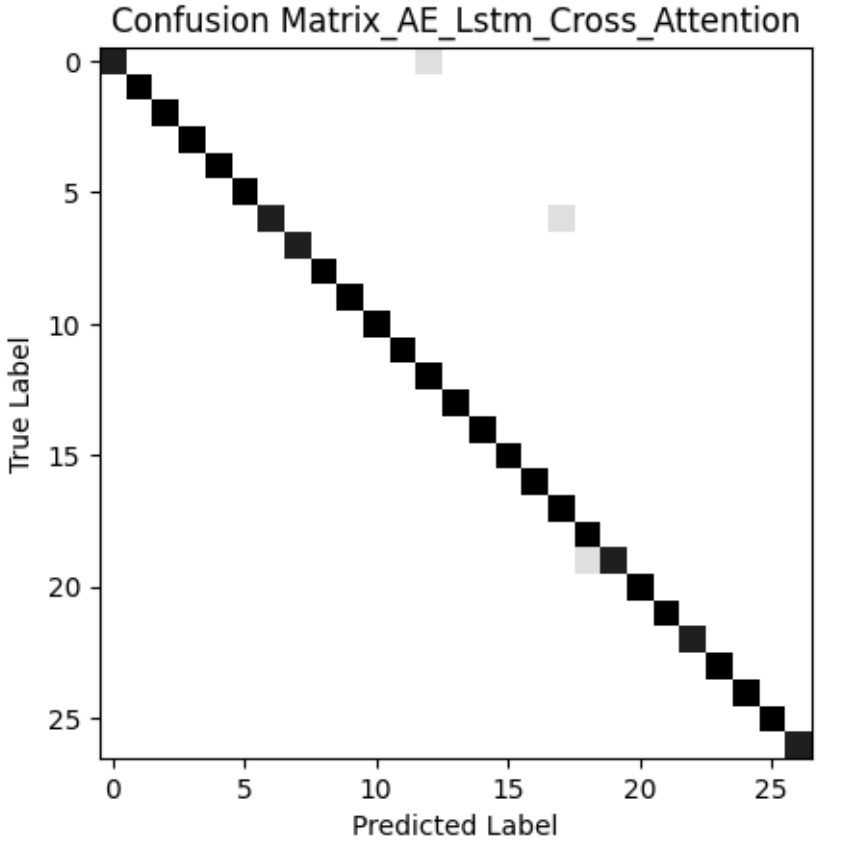
بلغت دقة التحقق للنموذج المقترح باستخدام AE-Lstm-Cross Attention **98.59%**.



الشكل 19: منحنى دقة النموذج المقترح باستخدام AE-Lstm-Cross Attention



الشكل 20: تابع خسارة النموذج المقترح باستخدام AE-Lstm-CrossAttention



n

الشكل 21: مصفوفة التشويش للنموذج المقترح. دقة النموذج 98.59%

Accuracy: 0.99

Micro Precision: 0.99

Micro Recall: 0.99

Micro F1-score: 0.99

Macro Precision: 0.99

Macro Recall: 0.99

Macro F1-score: 0.99

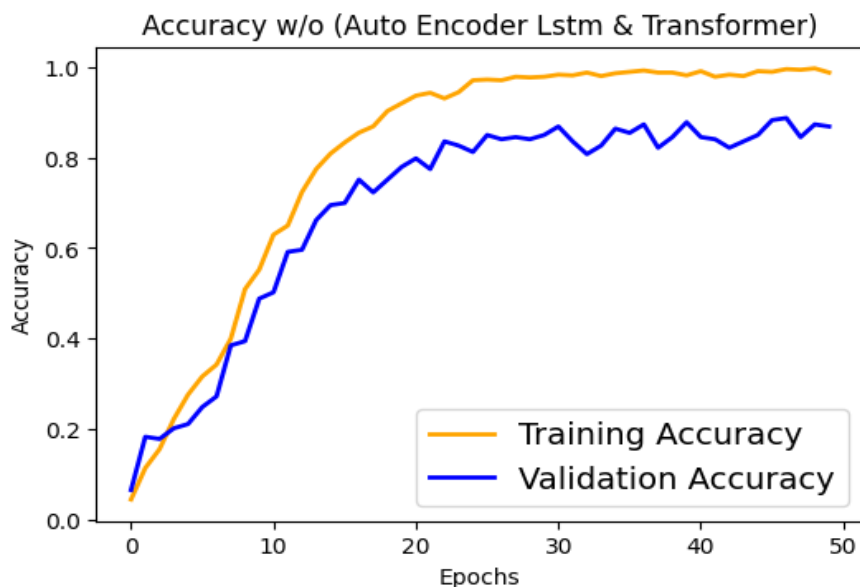
Weighted Precision: 0.99

Weighted Recall: 0.99

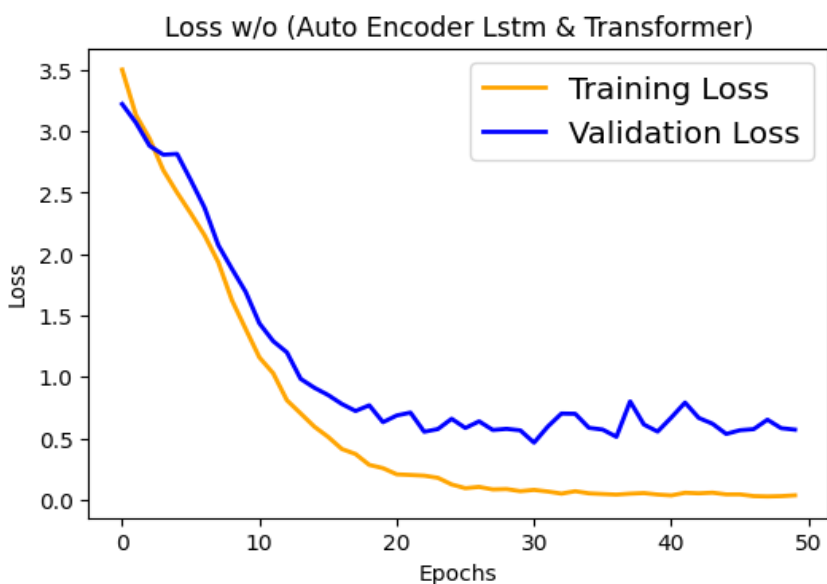
Weighted F1-score: 0.99

مقاييس أداء النموذج المقترح (Recall، Precision، F1 Score).

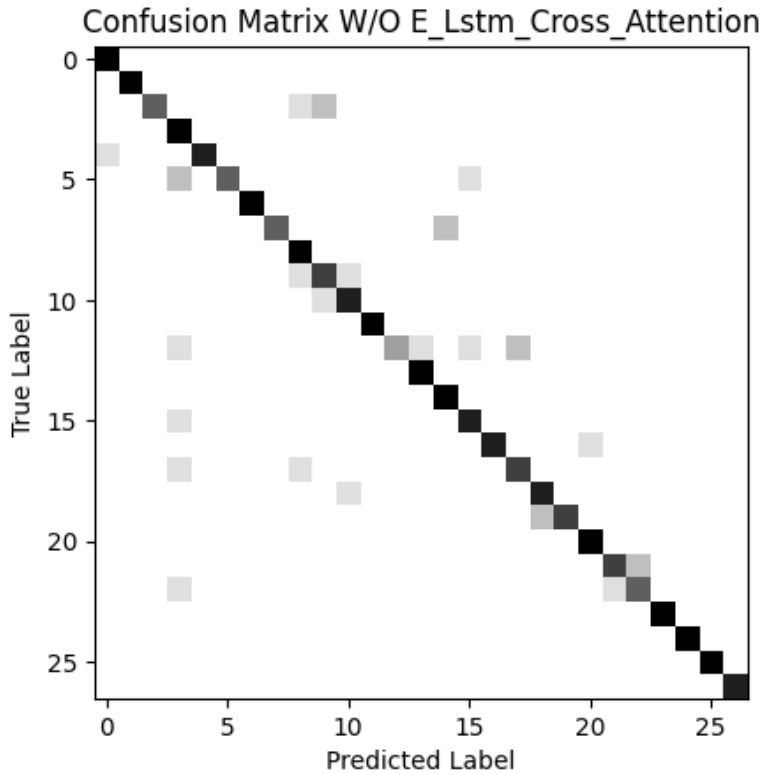
بلغت دقة التحقق للنموذج المقترح بدون استخدام AE-Lstm-Cross Attention **86.85%**.



الشكل 22: منحنى دقة النموذج بدون استخدام AE -Lstm-Cross Attention



الشكل 23: تابع خسارة النموذج بدون استخدام AutoEncoder-Lstm-Cross Attention



الشكل 24: مصفوفة التشويش للنموذج بدون التعديلات المقترحة. دقة النموذج %86.85

Accuracy: 0.87

Micro Precision: 0.87

Micro Recall: 0.87

Micro F1-score: 0.87

Macro Precision: 0.89

Macro Recall: 0.87

Macro F1-score: 0.87

Weighted Precision: 0.89

Weighted Recall: 0.87

Weighted F1-score: 0.87

مقاييس أداء النموذج (Recall، Precision، F1 Score) بدون التعديلات المقترحة.

3.6 مقارنة نتائج النموذج المقترح بالقيم الواردة في الأبحاث المشابهة والمطبقة على نفس مجموعة البيانات UTD_MHAD:

الدقة % (Accuracy)	مصادر البيانات	تاريخ ومكان النشر	اسم الدراسة
95.12	Video + Skeleton + Sensor	،IEEE Publisher 2020	HAMLET [10]
97.6	Video + Depth + Skeleton + Sensor	AAAI Conference on Artificial 2022،Intelligence	MuMu [14]
94.87	Video + Sensor	،IEEE Publisher 2022	VSKD + DASK [15]
96.04	Inertial + skeleton	AIHC (Springer) 2020	[9]
86.85	Inertial + skeleton		النموذج (بدون التعديلات المقترحة)
98.59	Inertial + skeleton		النموذج المقترح (مع التعديلات المقترحة)

7. الخاتمة والتوصيات:

تأتي البيانات الضخمة بأشكال وأنواع مختلفة. وقد تتكون من بيانات منظمة وشبه منظمة وغير منظمة. كما يمكن أن تولد انترنت الأشياء مجموعة واسعة من أنواع البيانات مثل النصوص والصوت والفيديو وبيانات الحساسات (أرقام سلاسل زمنية) واردة من مصادر مختلفة. يلعب دمج البيانات Data Fusion دوراً حاسماً في تطوير بيئات واسعة الانتشار معتمدة على بيانات انترنت الأشياء، يعد هذا الدور أكثر أهمية بالنسبة إلى تطبيقات انترنت الأشياء الحساسة للوقت حيث يلزم دمج البيانات في الوقت المناسب لجلب كل أجزاء البيانات معاً للتحليل وبالتالي توفير رؤى موثوقة ودقيقة وقابلة للتنفيذ. قامت هذه الدراسة باستعراض الأبحاث التي عالجت خاصية التنوع (دمج البيانات) للبيانات الضخمة في تطبيقات التعرف على النشاط البشري كأحد بيئات انترنت الأشياء، بغرض تقييمها والانطلاق نحو بناء نموذج يلحظ تنوع مصادر البيانات واختلاف تنسيقها.

تم تدريب النموذج باستخدام مجموعة البيانات الشهيرة UTD-MHAD [25] والتي تتضمن أنواع مختلفة من البيانات وهي مقاطع فيديو RGB، صور Depth، سلاسل زمنية لحساسات التسارع والجيروسكوب Inertial، سلاسل زمنية لحساسات الهيكل العظمي Skeleton. أظهرت نتائج تدريب النموذج على دمج بيانات حساسات Inertial مع بيانات حساسات الهيكل العظمي Skeleton دقة تعرف عالية على النشاط البشري باستخدام النموذج المقترح.

حيث وجدنا أنه بإضافة طبقة استخراج الميزات الغنية والتمثيلية من أحد مصادر البيانات الذي يتمتع بعدد ميزات كافي عبر توظيف شبكة مشفر تلقائي Auto encoder، وإضافة شبكة Lstm التي تتيح استخراج الارتباطات الزمنية لبيانات الحساسات،

وإستبدال الدمج البسيط Concatenation بالدمج عبر تقنية Cross Attention، هي مجموعة عوامل محسنة، حيث أنها رفعت دقة نموذج التعرف على النشاط البشري الذي تمت إضافتها له.

8. التطلعات المستقبلية:

إن استخدام الإضافات السابقة أدى إلى نتائج أفضل، لكن نلاحظ أن توظيف تقنية Cross Attention في عملية الدمج لا يتيح سوى استخدام مصدرين للبيانات فقط، مما يحد من إمكانية الاستفادة من وفرة مصادر البيانات. يمكن التغلب على هذه المشكلة مع الاستمرار بالاستفادة من تقنية Attention عبر توظيف تقنية Self-Attention وتطبيقها عدة مصادر بيانات. بعد إضافة التعديلات لاحظنا أن عملية التدريب أصبحت تستهلك زمناً أطول، وبعد التدقيق تبين أن شبكة Lstm هي التي تتطلب الكلفة الحسابية الأعلى بسبب بنيتها التي تتضمن معالجة تسلسلية للبيانات (كونها تأخذ الخرج السابق للخلية مع الدخل الحالي)، يمكن العمل مستقبلاً على وضع شبكة Lstm واحدة على ناتج عملية دمج مصدري البيانات عوضاً عن تخصيص شبكة Lstm لكل مصدر بيانات.

9. جدول المختصرات:

HAR	Human Activity Recognition
DL	Deep Learning
ML	Machine Learning
CNN	Convolution Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short Term
AE	Autoencoders
VAE	Variational Autoencoders
CA	Cross Attention
SA	Self Attention
FC	Fully Connected
MIM	Modal Enhancement Module
Conv	Convolution
MEM	Modal Interaction Module
EF	Early Fusion
LF	Late Fusion
GAN	Generative Adversarial Networks

- [1] R. Kitchin and G. McArdle, "What makes Big Data, Big Data Exploring the ontological characteristics of 26 datasets", Big Data & Society, 2016.
- [2] M. Gheisari, G. Wang, M. Bhuiyan, "A Survey on Deep Learning in Big Data", IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017.
- [3] k. chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep Learning for Sensor-based Human Activity Recognition: Overview, Challenges, and Opportunities", ACM, 2021.
- [4] Z. Liu, Q. Cheng, C. Song, and J. Cheng, "Cross-scale cascade transformer for multimodal human action recognition", Pattern Recognition Letters, Elsevier, 2023.
- [4] Z. Liu, Q. Cheng, C. Song, and J. Cheng, "Cross-scale cascade transformer for multimodal human action recognition", Pattern Recognition Letters, Elsevier, 2023.
- [5] X. Zhou, W. Liang, K. Wang, H. Wang, L. Yang, and Q. Jin, "Deep Learning Enhanced Human Activity Recognition for Internet of Healthcare Things" IEEE Internet of Things Journal, 2020.
- [6] SK Yadav, K Tiwari, HM Pandey, and SA Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions" Knowledge-Based Systems, Elsevier, 2021.
- [6] SK Yadav, K Tiwari, HM Pandey, and SA Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions" Knowledge-Based Systems, Elsevier, 2021.
- [7] Z. Ahmad, and N. Khan, "Human Action Recognition Using Deep Multilevel Multimodal (M2) Fusion of Depth and Inertial Sensors", IEEE, 2019.
- [8] M. EHATISHAM-UL- HAQ and A. JAVED, "Robust Human Activity Recognition Using Multimodal Feature-Level Fusion", IEEE, 2019.

- [9] J. Imran and B.Raman, "Evaluating fusion of RGB-D and inertial sensors for multimodal human action recognition[9]"، Ambient Intelligence and Humanized Computing (Springer), 2020.
- [10] MM. Islam and T. Iqbal, "Hamlet :A Hierarchical multimodal attention-based human activity recognition algorithm"، IEEE, 2020.
- [11] E. A. Barbosa, "Dual stream cross-modality fusion transformer for RGB-D action recognition"، Knowledge-Based Systems (Elsevier), 2022.
- [12] A. Raza, K. Tran, L. Koehl, S. Li, X. Zeng, K. Benzaidi, "Lightweight Transformer in Federated Setting for Human Activity Recognition in Home Healthcare Applications"، Elsevier, 2022.
- [13] B. Li, Y. Diao, and P. Shenoy, "Supporting scalable analytics with latency constraints," Proceedings of the VLDB Endowment, vol. 8, no. 11, pp. 1166–1177, 2015.
- [14] M. Hilbert, "Big data for development: a review of promises and challenges," Development Policy Review, vol. 34, no. 1, pp. 135–174, 2016.
- [15] W. Fan and A. Bifet, "Mining big data: current status, and forecast to the future," ACM SIGKDD Explorations Newsletter, vol. 14, no. 2, pp. 1–5, 2013.
- [16] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," IEEE Access, vol. 2, pp. 652–687, 2014.
- [17] Y. Demchenko, P. Grosso, C. De Laat, and P. Membrey, "Addressing big data issues in scientific data infrastructure," in Collaboration Technologies and Systems (CTS), 2013 International Conference on. IEEE, 2013, pp. 48–55.
- [18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [19] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555v1 [cs.NE], 2014.
- [20] P. Baldi, "Autoencoders, unsupervised learning, and deep architectures." ICML unsupervised and transfer learning, vol. 27, no. 37-50, p. 1, 2012.

- [21] C. Doersch, "Tutorial on variational autoencoders," arXiv preprint arXiv:1606.05908v2 [stat.ML], 2016.
- [22] D. P. Kingma, S. Mohamed, D. J. Rezende, and M. Welling, "Semisupervised learning with deep generative models," in Advances in Neural Information Processing Systems, 2014, pp. 3581–3589.
- [23] M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey", IEEE, 2018.
- [24] J Wang, Y Chen, S Hao, X Peng, L Hu (2018): Deep learning for sensor-based activity recognition: A survey.
- [25] C. Chen, R. Jafari, and N. Kehtarnavaz, "Utd-mhad: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor," in 2015 IEEE International conference on image processing (ICIP). IEEE, 2015, pp. 168–172.
- [26] A Katrompas, T Ntakouris, V Metsis, " Recurrence and self-attention vs the transformer for time-series classification: a comparative study", International Conference on Artificial Intelligence in Medicine, Springer, 2022.
- "
- [27] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I., "Attention is all you need". 31st Conference on Neural Information Processing Systems (NIPS 2017) (06 2017)

تفسير نموذج مدرب لاكتشاف خطاب الكراهية في التغريدات العربية

المهندس ربيع محي الدين الكردي¹

إشراف: الأستاذ الدكتور كمال السلوم² - الدكتور وسيم رمضان³

الملخص

يُعتبر انتشار خطاب الكراهية أحد أكبر المشكلات التي تقض مضجع شبكات التواصل الاجتماعي وتؤثر بشكل كبير على مستخدميه، ما دفع الباحثين إلى توظيف نماذج تعلم الآلة المتنوعة في اكتشافه ومواجهته. إلا أن هذه النماذج تفتقر إلى الشفافية والوضوح فيما يخص آلية اتخاذها لقرار التصنيف النهائي، ما يؤثر سلباً على فهم مُخرجاتها واكتشاف نقاط ضعفها وتحسين أدائها.

يهدف البحث إلى تفسير مُخرجات النموذج ماربيرت MARBERT وذلك بعد الضبط الدقيق لبارامتراته في مهمة اكتشاف خطاب الكراهية في اللهجات العربية العامية. استُخدمت خوارزمية المشتقات المتكاملة Integrated Gradients لتفسير مُخرجات النموذج فيما يخص صنف خطاب الكراهية ودُرست العوامل التي تدخل في اتخاذ قراراته من خلال تحليل تأثير مساهمات كلمات التغريدة في تصنيفها. سلط الضوء أيضاً على مواطن الضعف التي تقود النموذج إلى تصنيفات خاطئة. كما سُكّلت قوائم من الكلمات الأكثر تأثيراً في اختيار صنف خطاب الكراهية.

توصل البحث إلى إثبات تمكن النموذج من فهم أنماط لغوية معقدة وربطها مع خطاب الكراهية، واكتشف المطبات التي يقع فيها والتي تتمثل بالاستعارات والسخرية وغيرها ما يفيد في معرفة أماكن التحسين وتركيز الجهود فيها.

الكلمات المفتاحية: تعلم آلة، خطاب كراهية، نماذج مبنية على المحولات، ذكاء صناعي قابل للتفسير، ماربيرت، مشتقات متكاملة، تغريدة.

¹ طالب دكتوراه - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سورية

² أستاذ دكتور - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سورية

³ دكتور مشارك - قسم الاقتصاد الزراعي - كلية الزراعة - جامعة البعث - حمص - سورية

Interpreting a Model Trained for Detecting Hate Speech in Arabic Tweets

Eng. Rabee Mehi Addin AlKourdi

Dr. Kamal Al-Salloum – Dr. Wassim Ramadan

Abstract

The spread of hate speech is considered one of the biggest problems plaguing social media networks, greatly affecting its users. This has led researchers to employ various machine learning models for its detection and mitigation. However, these models lack transparency and clarity regarding the mechanisms used for its final classification decisions, which negatively impacts the understandability of its outputs, finding its weaknesses, and overall improve its performance.

This research aims to interpret the outputs of the MARBERT model after fine tuning its parameters for the task of hate speech detection in Arabic dialects. The Integrated Gradients algorithm was used to interpret the model's outputs regarding the hate speech class and the factors influencing its decisions were studied by analyzing the impact of a tweet words' contributions on its classification. The research also highlighted the weaknesses that lead the model making incorrect classifications. Lists of the most influential words in hate speech classification were also compiled.

The research demonstrated the model's ability to comprehend complex linguistic patterns and link them to hate speech. It identified obstacles such as metaphors and sarcasm, among others, which could help in pinpointing areas of improvement to focusing efforts on.

Keywords: Machine Learning, Hate Speech, Transformers-based Models, Explainable Artificial Intelligence, MARBERT, Integrated Gradients, Tweet.

1. مقدمة

أحدثت شبكات التواصل الاجتماعي تطوراً ثورياً في طرق التواصل بين البشر، ليصبح بإمكان أي شخصٍ يمتلك جهازاً مناسباً متصلاً بالإنترنت أن يلج إليها ويتواصل مع أي شخصٍ آخر ويتبادلان الرسائل والصور ومقاطع الفيديو لحظياً. حظي هذا التطور بالترحيب لما يُقدمه من سهولةٍ في طرح الآراء وتبادلها والخوض في نقاشاتٍ تخص جوانب الحياة كلها من سياسيةٍ واجتماعيةٍ ودينيةٍ واقتصاديةٍ وثقافيةٍ وغيرها. تتبع جاذبية استخدامها من عدم حاجة المستخدم إلى الكشف عن هويته الحقيقية، على الرغم من حث معظم شبكات التواصل الاجتماعي مستخدميها على كشفها لإخفائها. أعطى ذلك الأمر المستخدم حريةً كبيرةً في التعبير عن نفسه وآرائه ومعتقداته وجميع ما يخطر بباله مفيداً كان أو لم يكن على حد سواء، دون خوفٍ من أي آثارٍ قد تترتب على حياته الحقيقية. وعليه، أخذت هذه الشبكات بالانتشار على نحوٍ متسارعٍ واستقطبت أعداداً متزايدةً من المستخدمين منذ انطلاقتها وحتى الآن. فعلى سبيل المثال، ازداد عدد المستخدمين النشطين يومياً لتطبيقات شركة ميتا (فيسبوك Facebook) وانستغرام Instagram ومسنجر Messenger وواتساب (WhatsApp) من 2.10 مليار مستخدم في الربع الأول من 2019 إلى 2.93 مليار مستخدم في الربع الثالث من 2022 (وفقاً لتقارير ربع سنوية تنشرها الشركة [1]).

لكن، وكأي تطورٍ آخر على مر تاريخ البشرية، لهذه الشبكات جانبٌ مظلمٌ لا يُخفي نفسه. إذ أن امتيازات الحرية والسرية التي تقدمها جعلت منها أرضاً خصبةً لنقاشاتٍ حادةٍ عادةً ما تنتهي بتوجيه بعض المستخدمين الإساءات والإهانات للمستخدمين الآخرين الذين يخالفونهم الرأي، وليس من الغريب أن يلجأ بعضهم إلى التحقير والشتم وقد يصل به الأمر إلى التهديد بالعنف حتى. لتصبح أحد أهم ميزات هذه الشبكات سلاحاً يستخدمه ضعاف النفوس في نشر خطاب الكراهية دون حسيبٍ أو رقيب.

لكن ما هو خطاب الكراهية؟ لا يوجد تعريفٌ قانونيٌّ واضحٌ ومتفقٌ عليه عالمياً لخطاب الكراهية [2]، إلا أن ذلك لا ينفى وجود صفاتٍ مشتركةٍ بين التعاريف المتنوعة له في أدبيات الدراسة، مثل استهدافه لمجموعاتٍ محددةٍ من الأشخاص أو أفرادٍ من

جماعاتٍ ما ووجود محتوى يعبر عن الكراهية ويُسبب الأذى ويحرض على أفعالٍ سيئةٍ تتجاوز نطاق الخطاب نفسه، بالإضافة إلى وجود نية الأذى أو القيام بفعلٍ سيءٍ. أدى الانتشار الواضح والمتزايد لخطاب الكراهية على شبكات التواصل الاجتماعي إلى ازدياد الجهود الرامية إلى دراسته ومحاربته بشتى الوسائل [3]. تتجلى أهم هذه الوسائل باكتشافه آلياً باستخدام طرائق تعلم الآلة Machine Learning لقدرتها الكبيرة على التعامل حجومٍ ضخمةٍ من البيانات واكتشاف الأنماط المعقدة ضمنها بدقةٍ وسرعةٍ [4].

وعليه، يسعى الباحثون إلى تطوير نماذج تستغل التقدم الكبير المستمر في مجال تعلم الآلة عموماً ومعالجة اللغات الطبيعية Natural Language Processing خصوصاً (مثل التعلم المنقول Transfer Learning والمحولات Transformers [5])، لتتمكن هذه النماذج من اكتشاف الأنماط المعقدة في خطاب الكراهية. إلا أن هذه النماذج أخذت تنمو نموّاً كبيراً بفضل التطور في القدرات الحسابية للحواسيب وحجم البيانات الضخم المتوفر، ما جعلها معقدةً بدرجةٍ كبيرةٍ وأخذت طبيعة «صندوقٍ أسودٍ» [6]، إذ أنها لا تُفصح عن طريقة قيامها بالتنبؤ بشكلٍ يمكن للإنسان فهمه، ما أفقدها الشفافية وجعل تفسير نتائجها وتحسينها أمراً صعباً. يُضاف إلى ما سبق الطبيعة الغامضة لنماذج التنبؤ التي أعاقت تطبيقها بشكلٍ فعليٍّ على أرض الواقع [7] كونها غير جديرةٍ بالثقةٍ من وجهة نظر المستخدمين لعدم قدرتهم على فهم قراراتها التي قد تكون منحازةً (وهو أمرٌ صحيحٌ في كثيرٍ من الحالات). نتيجةً لذلك، أصبحت الحاجة إلى تفسير نماذج تعلم الآلة أمراً ملحاً وازداد اهتمام الباحثين بهذا المجال الذي يُعرف بالذكاء الصناعي القابل للتفسير eXplainable Artificial Intelligence (أو XAI اختصاراً). وبالرغم من أهمية ذلك الأمر، ما زالت الأبحاث التي تسعى إلى فتح ذلك الصندوق الأسود أقل بكثيرٍ من تلك التي تسعى إلى تحسين أداء النماذج فحسب. يبدو هذا النقص واضحاً بشكلٍ كبيرٍ في تفسير نماذج اللغة العربية، التي تعاني أساساً فيما يخص نماذج اكتشاف خطاب الكراهية مقارنةً بقريبتها الإنكليزية على الرغم من عدد المستخدمين الكبير الذين يتحدثون اللغة العربية على شبكات التواصل الاجتماعي [8].

ليس هذا النقص أمراً مفاجئاً لما تضيفه اللغة العربية من تحدياتٍ على تدريب نماذج تعلم الآلة عموماً وعلى تفسيرها خصوصاً. تعود تلك التحديات إلى طبيعة اللغة العربية الخاصة وتعقيدها وتنوع لهجاتها المُستخدمة على أرض الواقع وعلى شبكات التواصل الاجتماعي [8] عوضاً عن الفصحى.

على الرغم من قلتها، بُذلت جهودٌ عدةٌ في تفسير نماذج تعلم الآلة، إذ استخدم بعضها آلياتٍ لحساب أهمية ميزات المُدخلات Input Features من أجل تفسير قرارات النموذج ثم اقترح طريقةً لاستخدامها في تحسين مُخرجاته [9]، واستخدم البعض الآخر الذكاء الصناعي القابل للتفسير من أجل تفسير عدة نماذج لاكتشاف خطاب الكراهية في التغريدات الإنكليزية [10]، في حين دُرِس تأثير ميزات المُدخلات باستخدام طرائق التفسير لتبرير قرارات النموذج فيما يخص تحليل المشاعر باللغة العربية [11].

ومن هذا المنطلق، وحرصاً على زيادة شفافية نماذج اكتشاف خطاب الكراهية في اللغة العربية، يسعى هذا البحث إلى تفسير الميزات التي تؤثر على قرارات النموذج في تصنيف Classification نصٍ ما على أنه حاوٍ على خطاب كراهية أم لا، ودراسة مواضع فشله في ذلك، على أمل أن تفيد هذه المعرفة في زيادة فهم خطاب الكراهية وتحسين طرق اكتشافه الآلية.

أهداف وأهمية البحث

يهدف البحث إلى تحقيق ما يلي:

- 1- تفسير مُخرجات Output نموذجٍ ضُبِطت بارامتراتُه على مسألة اكتشاف خطاب الكراهية ضمن تغريداتٍ عربيةٍ وذلك من خلال توضيح مقدار مساهمة ميزات الإدخال في التنبؤ النهائي لكل تغريدة.
- 2- تحليل الأسباب التي قد تُسهم في التصنيف الخاطئ للتغريدات.
- 3- تشكيل قائمةٍ من الميزات الأكثر تأثيراً والتي تُساعد في تجميع تغريداتٍ جديدةٍ لدراسة خطاب الكراهية على نحوٍ أفضل.

تتبع أهمية هذا البحث من الحاجة إلى تفسير قرارات نماذج اكتشاف خطاب الكراهية في اللغة العربية من خلال تحليل ميزات المُدخلات (عادةً ما تكون على شكل كلمات)

لهذه النماذج وتحديد مقدار تأثير كلٍ منها على التنبؤ النهائي للنموذج (التصنيف)، ما يُساعد في زيادة فهم قرارات النموذج من جهةٍ مما قد يُساعد في تقوية مواضع ضعفه، ورفد المجتمع العلمي بمعلوماتٍ قد تكون مفيدةً في تحسين أداء نماذج التصنيف من جهةٍ أخرى.

تُلخّص مساهمات البحث بالنقاط التالية:

- 1- تدريب نموذج لاكتشاف خطاب الكراهية في اللغة العربية وتفسير مُخرجاته.
- 2- تحليل واستخراج أهم الميزات التي تعبر عن خطاب الكراهية في اللغة العربية.
- 3- تحديد أهم أسباب فشل المحولات وخصوصاً ماربيرت في اكتشاف خطاب الكراهية.
- 4- تشكيل معجم من أهم الكلمات المؤثرة على قرار النموذج في التصنيف إلى خطاب الكراهية مع مقدار أهمية كل منها.

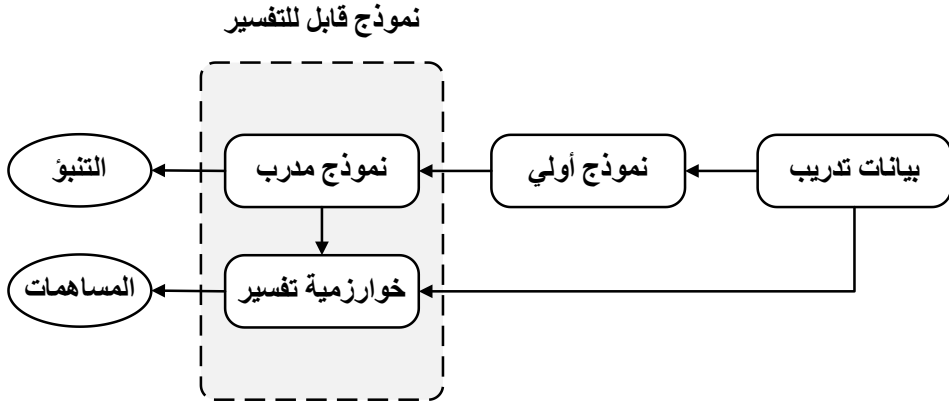
تُرتب فصول البحث على الشكل التالي: يتضمن **الفصل الأول** مقدمةً عن خطاب الكراهية وتفسير نماذج تعلم الآلة. يُقدم **الفصل الثاني** أهم الأبحاث فيما يتعلق بطرائق تفسير نماذج تعلم الآلة. يُوضح **الفصل الثالث** نموذج ماربيرت وخوارزمية المُشتقات المتكاملة. يعرض **الفصل الرابع** تفاصيل مجموعة البيانات وخطوات المعالجة المسبقة ومقاييس التقييم المستخدمة في البحث بالإضافة إلى تفاصيل بيئة التجربة. يتناول **الفصل الخامس** تفاصيل التجارب بدءاً من الضبط الدقيق للنموذج وصولاً إلى استخدام المُشتقات المتكاملة في تفسير مُخرجاته. أخيراً، يُلخّص **الفصل السادس** البحث ويعرض بعض التوجهات المستقبلية.

2. دراسة مرجعية

يُلاحظ نقصٌ واضحٌ في الدراسات التي تتناول تفسير نماذج تعلم الآلة وخاصةً في اللغة العربية. إذ تُركّز دراسات التفسير على اللغة الإنكليزية ويتناول بعضها تفسير خطاب الكراهية ضمن التغريدات، في حين أن الدراسات التي تتناول اللغة العربية لم

تُفسر سوى النماذج الخاصة بالخطاب المسيء Offensive Speech أو تحليل المشاعر Sentiment Analysis عموماً، ولم تُركز إطلاقاً على خطاب الكراهية.

تبدأ رحلة التفسير ببناء نموذج وتدريبه على البيانات الخاصة بالمسألة المراد تناولها (اكتشاف خطاب الكراهية على سبيل المثال)، يمكن اختيار نماذج تقليدية أو عميقة أو مبنية على المحولات. يلي ذلك عملية التفسير التي تبدأ بتمرير بيانات التدريب على هيئة مُدخلات ثم اختيار طريقة التفسير، لتكون مُخرجات التفسير النهائية مؤلفة من التنبؤات بالإضافة إلى التفسيرات (مساهمات ميزات مُدخلات النموذج بمُخرجاته النهائية). يوضح الشكل 1 الخطوات الأساسية في عملية تفسير نموذج ما:



الشكل 1 - خطوات عملية تفسير نموذج تعلم آلة

لا بد من اختيار أحد خوارزميات التفسير لربطها مع النموذج الأولي للحصول على النموذج المُدرَّب القابل للتفسير. هناك الكثير من الخوارزميات، إلا أن أدبيات الدراسة تُركز بشكلٍ واضحٍ على ثلاث خوارزمياتٍ وهي: خوارزمية التفسيرات المحلية القابلة للفهم غير التابعة للنموذج Local Interpretable Model-agnostic Explanations (LIME) وخوارزمية تفسيرات شابلي المُضافة SHapley Additive exPlanations (SHAP) وخوارزمية المُشتقات المتكاملة (IG) Integrated Gradients.

1.2. خوارزمية LIME

أحد طرائق تفسير النماذج التي تُطبق على النموذج المُدرّب بحد ذاته عوضاً عن تضمينها داخله على هيئة ميزةٍ من ميزاته، ما يجعلها غير متعلقةٍ ببنية النموذج المُراد تفسيره [12]. تُؤدّ الخوارزمية تفسيراتٍ محليةٍ (متعلقةٍ بالمثال المُراد تفسيره) وبسيطة. تحسب هذه الخوارزمية ميزات المُدخلات المهمة بإعطاء قيمٍ لعناصرها تمثل الأهمية وذلك باستخدام طرائق اختيار الميزات Feature Selection.

استُخدمت هذه الخوارزمية في تفسير عددٍ من النماذج الخطية غير المعقدة (مثل الغابة العشوائية Random Forest وشجرة القرار Decision Tree) لاكتشاف خطاب الكراهية في اللغة الإنكليزية من أجل تغريدةٍ محددةٍ حاويةٍ على خطاب كراهيةٍ (نص التغريدة قبل المعالجة المسبقة: « *@ComedyPosts: Harlem shake is just an excus to go full retard for 30 s الهارلم شيك هي مجرد عذر لتكون معاقاً بالكامل لمدة ثلاثين ثانية* ». نص التغريدة بعد المعالجة المسبقة: « *comedypost harlem shake excus go full retard* » [10]). توصلت التجارب إلى النتائج التالية:

- عند تفسير نموذج غابةٍ عشوائيةٍ باستخدام LIME فيما يخص التغريدة السابقة، صنّف النموذج المُدرّب التغريدة على أنها حاويةٌ على خطاب كراهيةٍ بثقةٍ بلغت 90%، وكانت نتيجة التفسير هي أن الكلمات *excus* (عذر) و *retard* (معاق) هي الأكثر تأثيراً في هذا القرار بقيم أهميةٍ بلغت 0.01 و 0.07 على التوالي (كلما ازدادت القيمة، ازداد تأثير الكلمة في قرار النموذج)، في حين أثرت الكلمة *full* (بالكامل) بشكلٍ معاكسٍ على القرار ودفعته نحو الصنف الآخر.
- على نحوٍ مشابه، كان للكلمات *full* و *excus* و *retard* تأثيراً موجباً في تصنيف التغريدة على أنها خطاب كراهيةٍ عند استخدام نموذج الشجرة العشوائية. لتكون قيم أهمية تلك الكلمات 0.07 للكلمة *full* و 0.06 للكلمتين *excus* و *retard*.

أما عند تحليل المشاعر في اللغة العربية، استُخدمت خوارزمية LIME مع شبكة ذاكرة قصيرة المدى مطولة Long Short-Term Memory (أو LSTM اختصاراً) مبنية على الانتباه Attention لتصنيف تعليقات حول مشاعر الأشخاص نحو عمليات ليزك الجراحية. تأخذ هذه التعليقات أحد الأصناف Classes الثلاثة التالية: إيجابية أو سلبية أو حيادية، وكانت النتائج التالية [11]:

- صنّف النموذج جملة «نظره ضعف هيعمل شاء الله يرجع يشوف» (المعنى الأصلي للجملة قبل المعالجة: كان نظره ضعيفاً وسوف يستعيد نظره بالكامل قريباً إن شاء الله)، على أنها إيجابية. وذلك بالرغم من أن كلمة «ليزك» ليست مذكورة في النص الأصلي، أي أن التعليق مربوط بفكرة عملية الليزك بشكلٍ ضمني وليس بشكلٍ صريح. ومع ذلك، تمكن النموذج من فهم الشعور الإيجابي بغض النظر عن غياب كلمات تُبين له فكرة الجملة بوضوح. أعطت LIME قيم أهمية موجبة كبيرة للكلمات **نظره** (0.08) و**شاء** و**يشوف** (0.04 لكلٍ منهما) متوافقة مع قرار النموذج بالتصنيف الإيجابي، في حين أخذت الكلمات **هيعمل** و**يرجع** و**ضعف** قيم أهمية سالبة (0.03 لكلٍ منها) علماً أن هذه الكلمات تحتمل أن تأخذ معانٍ إيجابية أو سلبية وفق السياق الذي ترد ضمنه.
- بينما صنّف النموذج جملة «لازم ارجع للنضاره ثاني» على أنها سلبية. حددت LIME الكلمات جميعها على أنها سلبية وأعطت **ارجع** أعلى قيمة أهمية وهي 0.29، أي تمكن النموذج من فهم أن المريض مر بتجربة سلبية عند إجرائه عملية ليزك واحتاج أن يعود لاستخدام النظارات الطبية. بناءً عليه، أكدت خوارزمية التفسير على ذلك وحددت الكلمات على أنها مرتبطة بشعورٍ سلبي.

2.2. خوارزمية SHAP

تتميز خوارزمية SHAP بأنها تُفسر مُخرجات النموذج من خلال تفكيكها إلى مجموع من قيم تمثل تأثير كل ميزة من ميزات المُدخلات على المُخرجات وذلك من خلال حساب قيمة المساهمة التي تقدمها كل ميزة (تعتمد على مفهوم نَسب الميزة المُضافة Additive Feature Attribution)، والتي تبيّن للمستخدم في نهاية الأمر مدى أهمية تلك الميزة في قرار المُصنّف [12].

طبقت خوارزمية SHAP من أجل تفسير نموذج شبكة عصبية ترشيحية Convolutional Neural Network (أو CNN اختصاراً) مدرّب على اكتشاف خطاب الكراهية في تغريدات إنكليزية عن وباء كوفيد-19 [9]. حسبت الخوارزمية متوسط أهمية كل ميزة (كل كلمة ضمن مجموعة البيانات) من أجل كل صنف على مستوى مجموعة البيانات الكليّة.

على نحوٍ مشابه لما سبق، حُسبت قيم SHAP لكل ميزة في تغريدات حاوية على خطاب كراهية ورُتبت هذه الميزات وفق الأهمية وأختير أكثرها تأثيراً على مُخرجات النموذج بهدف تحسين أدائه [13].

3.2. خوارزمية IG

أحد طرائق تفسير النماذج التي تعتمد على المشتقات في حساب مساهمة كل ميزة في مُخرجات النموذج. تتميز خوارزمية IG عن الخوارزميتين السابقتين بارتباطها بالنموذج المراد تفسيره، وتُستخدم عادةً من أجل نماذج الانحدار Regression والتصنيف. تستغل هذه الخوارزمية معرفتها بتفاصيل عمل النموذج (كونها مرتبطة به) لحساب المشتقات في كل طبقة من طبقاته [14].

استُخدمت خوارزمية المشتقات المتكاملة لإيجاد الميزات الأكثر تأثيراً في قرار نموذج شبكة LSTM ثنائية الاتجاه (Bi-LSTM) مدرّب على اكتشاف خطاب الكراهية في تغريدات باللغة الإنكليزية، والاستفادة من هذا التفسير لمعرفة سبب ارتكاب النموذج أخطاءً في تصنيف بعض التغريدات، على عكس نموذج آخر مبني على المحولات Transformers-based Model دُرب لنفس الهدف [15]. على سبيل المثال، تمكن النموذج من تصنيف تغريدة على أنها حاوية على خطابٍ مسيءٍ على الرغم من أنها مُسمّاة في بيانات التدريب خطأً على أنها غير مسيئة. وضحت خوارزمية المشتقات المتكاملة قرار النموذج. إذ كان للكلمات المُسيئة قيم أهمية كبيرة، ما يعني أن النموذج تمكن من فهم السياق وتحديد الكلمات المرتبطة مع الخطاب المسيء وخطاب الكراهية.

على نحوٍ مشابه، تمكنت الخوارزمية من توضيح قرارات نموذج ماربيرت في تصنيف التغريدات العربية الحاوية على خطابٍ مسيءٍ [16]. فعلى سبيل المثال، من أجل

التغريدة «**الله يلغهم**»، أشارت الخوارزمية إلى أن مساهمة كلمة **يلغهم** كبيرةً وتشد التصنيف باتجاه صنف الإساءة ما يتوافق مع قرار النموذج بأن هذه التغريدة مسيئةٌ فعلاً. وكذلك الأمر بالنسبة لتغريدةٍ أخرى «**عشان تكون ناجح بحياتك لازم يكون عندك إصرار**»، إذ صنف النموذج هذه التغريدة بشكلٍ صحيحٍ إلى صنف التغريدات غير المسيئة، وأكدت الخوارزمية صحة قراره من خلال حساب أهمية كلمات التغريدة في التصنيف لتدفع الكلمتان **ناجح** و**إصرار** تصنيف التغريدة نحو صنف عدم الإساءة، وهما بالفعل كلمتان إيجابيتان وغير مسيئتين.

3. النماذج والخوارزميات المستخدمة

يستخدم هذا البحث نموذجاً وخوارزميةً أثبتنا فعاليتها في الأبحاث السابقة وهما ماربيرت والمُشتقات المتكاملة. استُخدم ماربيرت في تكوين نموذجٍ قادرٍ على اكتشاف خطاب الكراهية في اللغة العربية العامية، واستُخدمت المُشتقات المتكاملة في تفسير قرارات التصنيف التي يتخذها نموذج ماربيرت.

1.3. نموذج ماربيرت

نموذج لغويّ مبنيّ على المحولات Transformer Language Model ومُدرّبٍ على اللغة العربية، وبالتحديد اللهجات المحلية باستخدام تغريداتٍ من تويتر Twitter (المُسمى X حالياً) [17].

بُني هذا النموذج وفق معمارية نموذج بيرت (BERT Bidirectional Encoder Representations from Transformers) [18]. دُرّب على شاكلة تدريب بيرت باستخدام أسلوب النموذج اللغوي المُقنع Masked Language Model الذي يعتمد على مفهوم «ملء الفراغات»، إذ غُطيت 15% من الوحدات اللغوية Tokens المُدخلة ودُرّب النموذج على تنبؤ هذه الوحدات على نحوٍ صحيحٍ وفقاً للسياق الذي ترد ضمنه (أي الكلمات التي تسبقها والتي تليها، ومن هنا تأتي ثنائية الاتجاه في تمثيلاته) لتكون نتيجة التدريب أشعة تمثيلٍ للوحدات اللغوية، والتي تشكل بمجموعها فهماً عميقاً من النموذج للغة التي دُرّب عليها.

يستخدم ماربيرت معمارية بيرت نفسها المؤلفة من 12 طبقةً و768 وحدةٍ مخفيةٍ (عدد العصبونات في كل طبقة) و12 رأس انتباهٍ ذاتيٍ Self-Attention Head

(طبقاتٌ خاصةٌ تحقق آلية الانتباه التي تُعطي النموذج القدرة على تركيز «انتباهه» انتقائياً على أجزاء مهمةٍ من عناصر المُدخلات ما يُحسن بدوره من دقة التنبؤ وفعالية استخدام الموارد، وهذا هو الأساس الذي بُنيت عليه المحولات [5]). يبلغ عدد بارامترات نموذج ماريبرت نحو 163 مليون بارامتر.

دُرّب النموذج باستخدام مليار تغريدةٍ عربيةٍ أُختيرت عشوائياً من أصل ست مليارات تغريدةٍ موجودة لدى الباحثين الذين أنشأوه، ليصل حجم مجموعة بيانات التدريب إلى نحو 128 غيغابايت (15.6 مليار وحدة لغوية). أنشئ هذا النموذج لسد الفجوة الموجودة في النماذج متعددة اللغات مثل mBERT [18] المدربة على بياناتٍ باللغة العربية الفصحى فقط، ما يجعلها غير مناسبة للمهام التي تتطلب التعامل مع اللهجات العربية المتنوعة.

2.3. المشتقات المتكاملة

تُعتبر خوارزمية المشتقات المتكاملة أحد أهم خوارزميات تفسير نماذج تعلم الآلة [19]، تقوم هذه الخوارزمية بـ «نَسَب» Attribute مُخرجات الشبكة العصبية Neural Network (التمثلة بالتنبؤ أو التصنيف) إلى ميزات المُدخلات خاصتها على شكلٍ رقمي قابلٍ للفهم البشري [20]. يُمكن تعريف عملية النَسَب رياضياً على النحو الآتي:

ليكن التابع $F : R^n \rightarrow [0, 1]$ يمثل شبكةً عصبيةً عميقةً ما، وليكن مُدخَل الشبكة المراد فهم ميزاته $x = (x_1, \dots, x_n) \in R^n$. يُعرّف نَسَب المُخرَج $F(x)$ لميزات

المُدخَل x على أنه مُنَجَّه من «قيم أهمية» $A(F, x) = (a_1, \dots, a_n) \in R^n$ إذ تُبين القيمة a_i مقدار مساهمة الميزة x_i في مُخرَج الشبكة $F(x)$ (زيادةً أو نقصاناً). تُشير قيمة الأهمية الموجبة إلى أن الميزة تسهم في زيادة قيمة مُخرَج الشبكة (تزيد من قيمة التنبؤ) في حين أن القيمة السالبة تعني أنها تساهم في إنقاصه. بينما تُفسر قيمة الصفر على أن الميزة لا تؤثر على المُخرَج إطلاقاً.

تعمل خوارزمية المشتقات المتكاملة على إيجاد مساهمة ميزات المُدخلات بمُخرجات النموذج انطلاقاً من مُدخَلٍ مرجعيّ Baseline Input وصولاً إلى المُدخَل الحقيقي [20]. يُعرّف المُدخَل المرجعي x' على أنه مُدخَلٌ خالٍ من الميزات وينتج عنه

تنبؤً حياديًا، يمكن أن يكون عبارةً عن صورةٍ سوداءٍ (أو صورةٍ ذات بكسلاتٍ عشوائيةٍ) في مسائل معالجة الصور، أو يكون مُتَّجِهَ تضميناتٍ صفريٍّ في مسائل معالجة النصوص.

1.2.3 التوصيف الرياضي للخوارزمية

من أجل تابعٍ $F : R^n \rightarrow [0, 1]$ الذي يمثل شبكةً عصبيةً عميقةً ومُدخلاً حقيقياً $x \in R^n$ ومُدخلاً مرجعياً $x' \in R^n$ ، يُؤخذ المسار المباشر من المُدخَل المرجعي وصولاً إلى المُدخَل الحقيقي وتُجمع المُشتقات على طول ذلك المسار. تحدد الخوارزمية قيمة الأهمية لميزة المُدخَل ذات الترتيب i وفق المعادلة (1):

$$(1) \quad IG_i(x) ::= (x_i - x'_i) \times \int_{\alpha=0}^1 \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$$

يُمثل الحد $(x_i - x'_i)$ الفرق بين المُدخَل الحقيقي والمُدخَل المرجعي للميزة i ، ويُمثل التكامل الانتقال من المُدخَل المرجعي إلى المُدخَل الحقيقي، بينما يُمثل ما بعد التكامل عملية تجميع المُشتقات المحلية.

إن حساب تكاملٍ محددٍ قد يكون مكلفاً حسابياً، لذا يُستخدم تقريب ريمان Riemann

Approximation لحسابه، وذلك وفقاً للمعادلة (2):

$$(2) \quad IG_i^{approx}(x) ::= (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F\left(x' + \frac{k}{m} \times (x - x')\right)}{\partial x_i} \times \frac{1}{m}$$

تُمثل m عدد خطوات تقريب ريمان.

2.2.3 المبادئ الأساسية لخوارزمية المشتقات المتكاملة

تُحقق الخوارزمية عدة مبادئ أساسية مرغوبةٍ في خوارزميات النَّسَب وهي [20]:

1- مبدأ الحساسية Sensitivity: من أجل كل مُدخَلٍ حقيقيٍّ ومُدخَلٍ مرجعيٍّ

يختلفان بميزةٍ واحدةٍ وينتج عنهما تنبؤان مختلفان، يجب إعطاء تلك الميزة

المسببة للاختلاف قيمة أهمية غير صفرية.

2- مبدأ انعدام تأثير التحقيق Implementation Invariance: يجب أن تكون

قيم الأهمية متطابقة دوماً من أجل الشبكات العصبية المتكافئة وظيفياً (تكون

الشبكتان متكافئتين وظيفياً إذا كانتا تولدان المُخرجات نفسها من أجل المُدخلات نفسها بغض النظر عن الاختلاف في بنيتهما).

3- مبدأ الاكتمال Completeness: مجموع كل قيم الأهمية مساوٍ للفرق بين مُخرَج الشبكة من أجل المُدخَل الحقيقي ومُخرَجها من أجل المُدخَل المرجعي. يُعبر عما سبق رياضياً بالمعادلة (3):

$$(3) \quad \sum_{i=1}^n IG_i(x) = F(x) - F(x')$$

4- مبدأ التناظر Symmetry: يكون متغيراً إدخال متناظرين بالنسبة لتابع F إذا كان التبدل بينهما لا يغير من التابع، أي $F(x, y) = F(y, x)$. بالتالي تكون خوارزمية النَّسَب محافظةً على التناظر إذا أعطت قيم الأهمية نفسها من أجل كل من المُدخَل الحقيقي والميزات المتناظرة في المُدخَل المرجعي.

5- مبدأ الخطية Linearity: لتكن الشبكتان F_1 و F_2 وشكلت شبكةً ثالثةً عن طريق الدمج الخطي لهما وكان تابعها $F_3 = aF_1 + bF_2$ علماً أن a و b تمثلان الأوزان. كي تحافظ الخوارزمية على هذه الخطية، يجب أن تحقق قيم أهمية الشبكة الجديدة المعادلة التالية:

$$(4) \quad IG_{F_3} = a \times IG_{F_1} + b \times IG_{F_2}$$

4. مواد وطرائق البحث

- تشمل دراسة الميزات المؤثرة على قرار المُصنّف عدة خطواتٍ تُلخص وفق الآتي:
- 1- تحديد مجموعة بياناتِ Dataset مناسبةٍ تتضمن أشكالاً متنوعةً من خطاب الكراهية.
 - 2- إجراء عمليات المعالجة المسبقة Preprocessing المناسبة على البيانات.
 - 3- استخدام التعلم المنقول [21] لضبط بارامترات ماريبرت على مجموعة البيانات لبناء النموذج المراد تفسيره.
 - 4- استخدام خوارزمية المُشتقات المتكاملة لتفسير مُخرجات نموذج ماريبرت المُدرب.

يستعرض القسم الحالي توصيف مجموعة البيانات وخطوات المعالجة المُسبقة لها ومقاييس التقييم المستخدمة وتفاصيل بيئة التجربة.

1.4. توصيف مجموعة بيانات البحث

يستخدم هذا البحث مجموعة بيانات مؤلفة من تغريدات بلهجات عربية عامية متنوعة تتضمن خطاباً مسيئاً وخطاب كراهية (بيانات ورشة عمل OSACT5 [22]). استخدم الباحثون الذين أنشأوها الرموز التعبيرية Emojis التي تشيع رؤيتها في العبارات الحاوية على إساءة ما (استخرجوا هذه الرموز من مجموعات بيانات سابقة خاصة بالتغريدات المسيئة)، وذلك لتجميع التغريدات من تويتر ثم فلترتها وتصنيفها يدوياً إلى عدة أصناف هي تغريدات نظيفة ومسيئة ونابية وعنيفة وكارهة. كما تُصنف تغريدات الكراهية إلى 6 أنواع فرعية هي خطاب كراهية ضد عرق أو دين أو انتماء أو إعاقة جسدية أو طبقة اجتماعية أو خطاب كراهية جنساني (ضد النساء أو الرجال).

تتألف مجموعة البيانات من 12698 تغريدة مقسمة بشكل مسبق إلى ثلاث مجموعات وهي مجموعة التدريب Training (70% من مجموعة البيانات الكلية) ومجموعة التحقق Validation (10%) ومجموعة الاختبار Testing (20%). ينظراً لتركيز هذا البحث على خطاب الكراهية بالتحديد وليس على الخطاب المسيء أو غيره، فقد دُمجت أصناف خطاب الكراهية ضمن صنف واحد (HS) ليكون توزيع تغريدات مجموعة البيانات فيما يخص هذا الصنف على النحو التالي:

المجموع	الاختبار	التحقق	التدريب	
1339	271	109	959	خطاب كراهية (HS)
11359	2270	1160	7929	لاكراهية (NOT_HS)
12698	2541	1269	8888	المجموع

الجدول 1 - توزيع التغريدات بالنسبة لصنف خطاب الكراهية

2.4. المعالجة المسبقة للبيانات

أجري عددٌ من خطوات المعالجة المسبقة البسيطة للبيانات قبل تدريب النموذج وذلك بهدف تقليل الضجيج فيها. لم تُستخدم طرائق المعالجة التي تُعدل البيانات على نحوٍ

كبيرٍ (مثل إزالة مُستبعدة الفهرسة Stopwords والتشذيب Stemming) وذلك لعدم الحاجة إلى إجراء تغييراتٍ كبيرةٍ في البيانات عند التعامل مع النماذج المبنية على المحولات [23]. فيما يلي خطوات المعالجة المسبقة التي طُبقت على مجموعتي التدريب والتحقق (لم تُستخدم مجموعة الاختبار في هذا البحث واستُعيض عنها بمجموعة التحقق):

- إزالة الروابط وإعادة التغريد Retweet.
- إزالة الإشارات للمستخدمين User Mention (مثال: @rabee).
- إزالة الرموز التعبيرية.
- تفكيك الهاشتاغ Hashtag إلى الكلمات التي تُولفه. (مثال: #بائع_الكليجا تصبح بايع الكليجا).
- تحويل الأرقام من الصيغة الهندية (٠١٢٣٤٥٦٧٨٩) إلى الصيغة العربية (0123456789).
- إزالة تكرار المحارف من الكلمات، أسلوبٌ يستخدمه المتحدثون العرب للتشديد على أمرٍ ما. (مثال: هههههههههههه تصبح هه)
- إزالة المحارف الخاصة ومعظم علامات الترقيم (مع الإبقاء على إشارات التعجب والاستفهام لارتباطها بمشاعر الاستغراب والاستنكار والغضب والارتباك) والمحارف غير العربية.

3.4. مقاييس التقييم

يستخدم البحث عدداً من المقاييس وهي: الدقة Accuracy والإحكام Precision والاستدعاء Recall ومقياس إف 1 F1-Score. تُحسب هذه المقاييس وفق المعادلات التالية، مع التنويه إلى أن TP تعني تصنيفاً إيجابياً صحيحاً وFP تعني تصنيفاً إيجابياً خاطئاً و TN تعني تصنيفاً سلبياً صحيحاً و FN تعني تصنيفاً سلبياً خاطئاً:

$$(5) \quad \text{الدقة} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$(6) \quad \text{الإحكام} = \frac{TP}{TP + FP}$$

$$(7) \quad \frac{TP}{TP + FN} = \text{الاستدعاء}$$

$$(8) \quad 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \text{مقياس إف 1}$$

استُخدم تنقيح ماكرو Macro عند حساب الإحكام والاستدعاء ومقياس إف 1 كون البيانات غير متوازنة.

4.4. بيئة التجربة

استُخدم عددٌ من مكتبات لغة بايثون Python في تجارب البحث، أهمها:

- **transformers**: مكتبة بايثون تقدمها منصة huggingface¹ وتتضمن آلاف النماذج المبنية على المحولات المدربة مسبقاً على مهماتٍ متنوعةٍ في عدة مجالات منها معالجة النصوص (تصنيف النصوص Text Classification وغيرها).
 - **captum**: مكتبة بايثون مُطورةٌ من قبل شركة فيسبوك وتقدم مجموعةً كبيرةً من طرائق تفسير نماذج تعلم الآلة المبنية باستخدام مكتبة PyTorch [19]. تهدف captum إلى توفير طرائق التفسير بشكلٍ بسيطٍ ومرنٍ وسهل الاستخدام. من أهم الطرائق التي تُحققها المكتبة: Integrated Gradients و LIME و SHAP المذكورة سابقاً.
- أُجريت جميع التجارب ضمن بيئة Google Colaboratory [24] باستخدام Python 3 وعلى وحدة معالجة رسومية (GPU) من نوع Nvidia Tesla T4 15GB ووحدة معالجة مركزية (CPU) من نوع Intel Xeon 2.30GHz 2-Cores وذاكرة وصولٍ عشوائٍي (RAM) بحجم 12.7GB.

¹ <https://huggingface.co>

5. التجارب والنتائج ومناقشتها

1.5. تدريب النموذج

تماشياً مع التوجهات الحالية التي وضحتها الدراسات السابقة، أُجري الضبط الدقيق لبارامترات النموذج ماربيرت عوضاً عن تدريب نموذج من الصفر، وذلك للاستفادة من التعلم المنقول. أُختير نموذج ماربيرت المُدرّب على بياناتٍ عربيةٍ بلهجاتٍ عاميةٍ متنوعة، ما يتناسب مع طبيعة بيانات هذا البحث. بالتالي فإن ضبط نموذج كهذا يُعتبر انتقالاً طبيعياً من معرفةٍ عامةٍ باللهجات العربية العامية نحو معرفةٍ مخصصةٍ بخطاب الكراهية فيها.

1.1.5 تجهيز المُدخلات

بعد أن طُبقت خطوات المعالجة المسبقة على بيانات التدريب، حُولت إلى شكلٍ يستطيع نموذج ماربيرت فهمه عن طريق استخدام المُقسّم Tokenizer الخاص به. يقوم المُقسّم بتقسيم النص المستمر إلى مجموعة مُتجهاتٍ عديدةٍ يستطيع النموذج التعامل معها. يوضح المثال التالي خطوات تحويل تغريدةٍ تابعةٍ لصف الكراهية من صيغتها النصية الخام إلى صيغةٍ عدديةٍ جاهزةٍ لإدخالها إلى النموذج.

التغريدة الأساسية: «الهندي قاعد يثبت للمغرب إن الخروف صار صديقه حتى ما

ينبجه»

الخطوة 1: التحويل من نص إلى وحدات لغوية:

يُحول المُقسّم نص الإدخال إلى وحداتٍ لغويةٍ متقطعةٍ وفقاً لخوارزمية التقسيم التي دُرّب عليها. دُرّب المُقسّم في ماربيرت باستخدام خوارزمية WordPiece [18] (وذلك عندما أُجري التدريب المسبق للنموذج [17]) وشُكل معجمٌ ضخمٌ من جميع الوحدات اللغوية الناتجة (بلغ عددها نحو 100 ألف وحدة).

كي تُحول سلسلة إدخالٍ نصيةٍ إلى سلسلة وحداتٍ لغويةٍ خاصةٍ بالنموذج، تُقسّم أولاً وفق استراتيجية التقسيم الخاصة بالمُقسّم إلى وحدات. يلي ذلك مطابقة هذه الوحدات مع تلك الموجودة في المعجم. إذا كانت موجودةً ضمنه، يُبقى عليها في السلسلة الناتجة. أما إذ كانت غير موجودة، فإنها تُستبدل بوحدةٍ خاصةٍ هي الوحدة المجهولة [UNK].

يُضاف أيضاً لبداية السلسلة الناتجة وحدةً خاصةً [CLS] تمثل بداية سلسلة المُدخَل، ويُضاف في نهايتها الوحدة [SEP] التي تمثل نهاية المُدخَل، يليها عدة وحدات [PAD] تمثل التبطين Padding المستخدم لزيادة طول السلسلة (بهدف توحيد طول جميع السلاسل المُدخلات). في هذا المثال، إذا افترضنا قيمة التبطين مساويةً للقيمة 20، تكون السلسلة التالية هي نتيجة هذه الخطوة: [CLS]، الهندي، قاعد، يثبت، للمع، ##رب، ان، الخروف، صار، صديقه، حتى، ما، يذبح، ##ه، [SEP]، [PAD]، [PAD]، [PAD]، [PAD]، [PAD]، [PAD].

يُلاحظ الرمز ## (مصدره خوارزمية التقسيم WordPiece) ويدل على أن هذه الوحدة مُتجزأة من الوحدة التي تسبقها.

الخطوة 2: تحويل الوحدات اللغوية إلى قيم عددية

لا تستطيع نماذج تعلم الآلة التعامل مع النصوص بشكلها الخام، لذا لا بد من تحويلها إلى صيغةٍ عدديةٍ مناسبة. يُحول المُقسَم الوحدات اللغوية إلى شكلٍ عدديٍ باستبدال كل وحدةٍ بفهرسها الموافق من معجم الوحدات الخاص بالمُقسَم. تكون نتيجة هذه العملية على المثال السابق السلسلة التالية: [2]، 19375، 3604، 9255، 23174، 6364، 1946، 33047، 3025، 16221، 2241، 1962، 16728، 1021، 3، 0، 0، 0، 0.

يُلاحظ أن فهارس الوحدات الخاصة [CLS] و [SEP] و [PAD] هي 2 و 3 و 0 على التوالي.

الخطوة الثالثة: تشكيل قناع الانتباه Attention Mask

نظراً لتبطين المُدخلات، يُستخدم قناعٌ يُسمى قناع الانتباه، مؤلفٌ من أصفارٍ ووحدات. يهدف القناع إلى إبلاغ النموذج أي الوحدات اللغوية التي يجب أن «ينتبه» لها (وحدات)، وأي الوحدات التي ينبغي عليه أن يتجاهلها (وهي عادةً وحدات التبطين، والتي تأخذ قيم أصفار في القناع). يكون قناع الانتباه للمثال السابق كما يلي: [1]، 1، 1، 1، 1، 1، 1، 1، 1، 1، 1، 1، 1، 0، 0، 0، 0، 0، 0، 1.

2.1.5 الضبط الدقيق

بعد تحويل جميع تغريدات مجموعة التدريب إلى الصيغة المناسبة وفقاً لما سبق، ضُبط نموذج ماربيرت عليها بتمريرها على شكل حُزْم Batches مؤلفة كلٍ منها من 64 تغريدة (تُحدد قيمة التبطين لكل حُزْم بطول أطول تغريدة فيها)، وبمعدل تعلم Learning Rate بلغ 5×10^{-5} ومعدل اضمحلال أوزان Weight Decay قيمته 0.01 ودُرب النموذج 15 دورة Epoch، التي كانت كافيةً للحصول على أفضل قيمة إف1 وفق الإعدادات الحالية.

يبين الجدول 2 مقاييس التقييم للنموذج الناتج على مجموعتي التدريب والتحقق. لم تُستخدم مجموعة الاختبار في هذا البحث واكتُفي بمجموعة التحقق لاختبار أداء النموذج على بيانات خارجية لم يرها من قبل:

مقياس إف1	الاستدعاء	الإحكام	الدقة	
98.94%	99.17%	98.73%	99.57%	مجموعة التدريب
83.02%	80.29%	86.45%	95.12%	مجموعة التحقق

الجدول 2 - مقاييس التقييم لنموذج ماربيرت المدرب على اكتشاف الكراهية

2.5. تفسير النموذج

استُخدمت خوارزمية المُشغقات المتكاملة لتفسير نموذج ماربيرت السابق المُدرب على اكتشاف خطاب الكراهية في التغريدات العربية وذلك لدراسة قراراته وعلاقة ميزات المُدخلات (الوحدات اللغوية) بهذه القرارات. مرت التجربة بعددٍ من الخطوات وصولاً إلى التفسير النهائي.

1.2.5 خطوات الوصول إلى التفسير

يتلخص سير عمل تفسير النموذج بست خطواتٍ رئيسيةٍ هي تشكيل المُدخلات وحساب قيم الأهمية ودمج الوحدات اللغوية المُقسّمة وأضيف بنهاية التجربة معجمٌ للوحدات اللغوية الأكثر أهمية في اكتشاف خطاب الكراهية ومقدار أهمية كلٍ منها.

وذلك على جميع أبعاد التضمين. إذا كان كانت p قيمة التبطين و m عدد الوحدات اللغوية في التغريدة، تكون أبعاد مصفوفة مُتجهات الأهمية الناتجة A هي $[128 \times (m + p + 2) \times 768]$. (+2 من أجل وحدتي البداية والنهاية).

تبسيطاً لحساب قيم الأهمية النهائية، اختزلت جميع أبعاد التضمين إلى بُعد واحد وذلك بجمع قيم الأهمية لكل بُعد تضمين من أجل كل وحدة لغوية في التغريدة. ليكن $A_d(tok)$ شعاع الأهمية الخاص بالوحدة اللغوية tok في البعد d ، يُحسب الشعاع المُختزل $A_{sum}(tok)$ وفق المعادلة:

$$(9) \quad A_{sum}(tok) = \sum_{d=1}^{768} A_d(tok)$$

تكون أبعاد المصفوفة A_{sum} من أجل الحزمة الواحدة: $[128 \times (m + p + 2)]$

الخطوة 3: دمج الوحدات اللغوية المُقسّمة

أعيد تشكيل كلمات المُدخلات التي قُسمت إلى عدة وحدات لغوية بدمج الوحدات المتتالية التي تبدأ بالرمز $##$ وجمع قيم أهمية الوحدات الفرعية ثم إسنادها إلى الكلمة الناتجة.

على سبيل المثال: لتكن قيم الأهمية للوحدات اللغوية [يذبح، ##هـ] 0.02 و 0.01 على التوالي. تكون نتيجة الدمج تشكيل الوحدة اللغوية [يذبحه] ذات الأهمية 0.03.

الخطوة 4: تشكيل معجم من الوحدات وقيم أهميتها

بعد حساب قيم الأهمية لكل الوحدات اللغوية، شُكل معجم من الوحدات اللغوية الفريدة وقيمة أهمية كل منها مع تكرار ورودها ضمن مجموعة البيانات. احتفظ بقيمة الأهمية الأكبر في حالة الوحدات المكررة.

2.2.5 نتائج تفسير النموذج

قبل البدء باستعراض نتائج التفسير، يوضح الجدول 3 التمييز البصري المُتبع في تمييز أهمية (مساهمات) الكلمات في تصنيف التغريدة:

المعنى	تمييز شكلي	تمييز لوني
تساهم الكلمة في اختيار صنف الكراهية (أي أن هذه الكلمة تشد النموذج نحو تصنيف التفريدة إلى الصنف HS)	تحتها خط غامق (___)	خلفية حمراء
تساهم الكلمة في اختيار صنف اللاكراهية (أي أن هذه الكلمة تشد النموذج نحو تصنيف التفريدة إلى الصنف NOT_HS)	تحتها خط منقط (.....)	خلفية خضراء

الجدول 3 - آلية التمييز البصرية في أمثلة البحث

في حين تشير درجة اللون إلى مقدار قيمة الأهمية الخاصة بالكلمة (كلما كانت الشدة أكبر كانت القيمة أعلى).

يُبين الجدول 4 تأثير الكلمات في عددٍ من التفريدات التابعة لصنف الكراهية في مجموعة التدريب والتي نجح النموذج بتصنيفها على نحوٍ صحيح.

التفريدة
[CLS] غالبيتكم عملاء لذا الاحتلال جائم فوق روكم [SEP]
[CLS] حركات كفار [SEP]
[CLS] مياشافيت عني مثل همجية وملحجة بنات الرياض إوثيرر ميعيشيش [SEP]
[CLS] حنا بدو يا جاهلين البداهة تصويري [SEP]
[CLS] ويلوموني بكرهه الامارات [SEP]
[CLS] ججد الضفدع الناطق [SEP]
[CLS] هولاء مرتزقة ابواق مفرغه يقولون ما لا يفعلون تنتظر منهم إيه الا الخزي والعار [SEP]

الجدول 4 - أمثلة على تفريدات كراهية من مجموعة التدريب ومساهمات كلماتها في التصنيف النهائي

عند الغوص في تفاصيل قيم الأهمية لأحد التغريدات المُصنفة على أنها خطاب كراهية (الشكل 2)، يبدو واضحاً أن تأثير الكلمات بنات والرياض وهمجية هو الأكبر في تصنيف النموذج للتغريدة على أنها تغريدة كراهية وذلك وفقاً لخوارزمية التفسير. يُمكن تفسير قرار النموذج بإمكانه من فهم معنى التغريدة المُضمن، إذ أن كلمة بنات (التي لا تحمل معنى كراهية بمفردها) قد ساهمت في تصنيف التغريدة إلى صنف الكراهية نتيجة السياق السلبي الذي وردت ضمنه، إذ لحقتها كلمة الرياض لتشكل العبارة بنات الرياض التي فهمها النموذج على أنها ضمن سياقٍ يخص جنساً محدداً في بقعةٍ جغرافيةٍ محددة، وهذا النوع من الوصف يميل إلى أن يرد عادةً ضمن سياقٍ حاوٍ على خطاب كراهية.

يُضاف إلى ما سبق ورود كلمتي همجية وملحجة المؤنثتين وذوات معانٍ سلبية، ليتمكن النموذج بنتيجة ذلك من ربط العبارات ضمن سياقٍ واحدٍ وفهم أن التغريدة تصف النساء اللواتي يقطنن في مدينةٍ معينةٍ بصفاتٍ سلبية، وهذا ما يقع تحت لواء خطاب الكراهية الجنساني. لذا تمكن النموذج من تصنيف التغريدة بشكلٍ صحيحٍ إلى صنف الكراهية.

أما فيما يخص الكلمات التي ساهمت في التصنيف إلى صنف اللاكراهية، فإن الكلمة **معلّيشش** لا تأتي معظم الأحيان ضمن سياقٍ سلبي، بل تعني عادةً طلب الصبر، لذا كانت قيمة أهميتها سالبةً وساهمت في شد التصنيف نحو الصنف NOT_HS.

وفيما يلي جدولٌ بقيم الأهمية والتكرار للكلمات العشر الأكثر تأثيراً:

الكلمة	قيمة الأهمية	التكرار
يابنات	1.537	4
حريمكم	1.286	1
البنات	1.187	7
الزهارنه	1.105	1
هالبنات	1.087	3
بنات	1.065	3
الضفدع	0.958	3
هالنظام	0.943	1
متخلفات	0.935	1
ضفادع	0.926	1

الجدول 5 - معلومات أكثر 10 كلماتٍ تأثيراً في خطاب الكراهية في بيانات التدريب

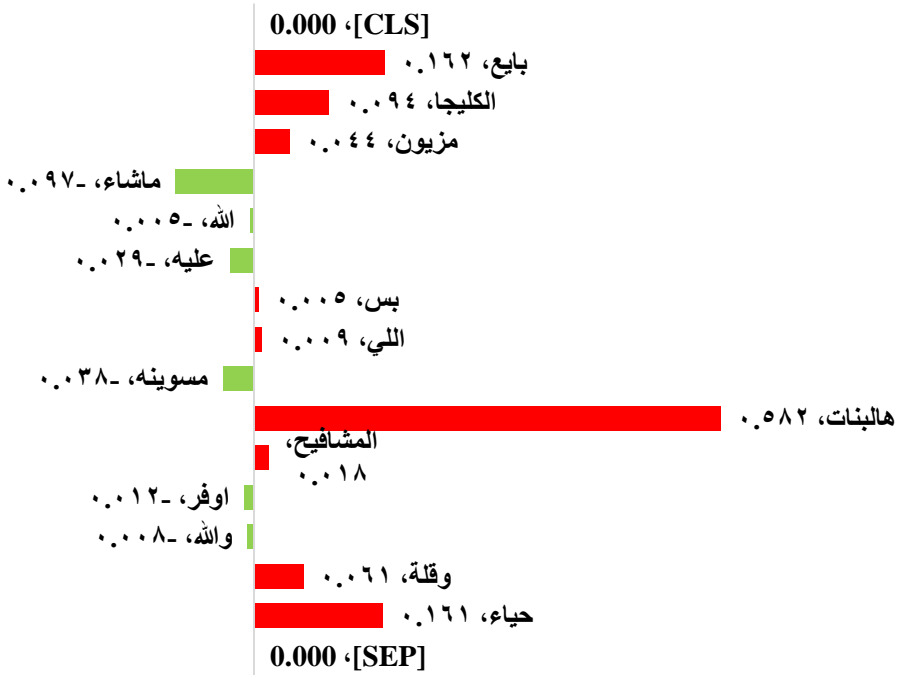
نُلاحظ سيطرةً واضحةً للكلمات التي تشير للنساء، وهو ليس بالأمر الغريب كون مجموعة البيانات تميل إلى خطاب الكراهية الجنساني (نحو 47% من تغريدات الكراهية في بيانات التدريب تابعة لذلك الصنف من خطاب الكراهية [22]). ما يعني أن وجود كلماتٍ من هذه القائمة يزيد من احتمال كون التغريدة حاويةً على خطاب كراهية إذا وردت ضمن سياقٍ يُشدد على ذلك. إن هذه النتائج متوافقةً أيضاً مع الكلمات المفتاحية المُستخدمة عند تجميع تغريدات مجموعة البيانات [22]. نُلاحظ أيضاً أن الكلمات الأكثر تأثيراً ليست نابيةً (باستثناء **متخلفات** و**ضفادع**. إذ تحمل الأولى معنىً سلبياً مهيئاً للنساء في حين تُشبه الثانية الآخرين بالحيوانات، وهو أسلوبٌ شائعٌ من الإساءة في المجتمعات العربية).

لكن ماذا بشأن البيانات التي لم يرها النموذج من قبل؟ يبين الجدول 6 عدة تغريداتٍ من مجموعة التحقق وتنبؤ النموذج الموافق لكلٍ لها، إلى جانب صنفها الحقيقي. يُبين الجدول أيضاً تأثير كلمات كل تغريدةٍ على قرار النموذج النهائي:

الصف الحقيقي	النتبؤ	التغريدة
HS	HS	[CLS] <u>بايع الكليجا مزبون ماشاء الله عليه بس اللي مسويينه</u> <u>هالبنات المشافيح اوفر والله وقلة حياء</u> [SEP]
HS	HS	[CLS] <u>هالبحرين بالذات لاتاريخ ولا كرامه كل حياتهم نهب</u> <u>ويوق</u> [SEP]
HS	HS	[CLS] <u>تسريب خطاب الترشيح خطابين للتاريخ سيطلان وصمة</u> <u>عار في جبين اعلام وجماهير اقزام اسيا الاهلى و النصر</u> <u>الحمد لله على نعمة سفير ...</u> [SEP]
HS	Not HS	[CLS] <u>اهم شي نكور الشعب كله يستشرف وينزل فديوها</u> <u>عين الاخلاق والحشمة بايع الكليجا</u> [SEP]
HS	Not HS	[CLS] <u>اللي بصراحه لا تخلين الذكور واشباه الذكور وما الى</u> <u>ذلك يكسروك حسيتي عليه</u> [SEP]
HS	Not HS	[CLS] <u>هل تعلم ان انبي النسير تموت بعد وفاة زوجها حزنا</u> <u>عليه اما النسيه الي بالخليج تروح تراجع دايرة التقاعد وتاخذ</u> <u>راتيه وتقول ماترك لنا الا الفقر</u> [SEP]

الجدول 6 - أمثلة على تغريدات من مجموعة التحقق ونتائج التصنيف والتفسير

بالدراسة التفصيلية لأحد التغريدات المُصنفة بشكل صحيح، نلاحظ الكلمات المؤثرة على التصنيف وقيم أهميتها في الشكل 4:



الشكل 4 - قيم الأهمية لكلمات لتغريدة كراهية صنفتها النموذج بشكل صحيح

على نحوٍ مشابهٍ لما سبق، تأخذ كلمة **هالبنات** أكبر قيمة أهميةٍ في التغريدة، وتلتها كلمة **المشافيح** ذات المعنى السلبي (صفة سلبية تُقال عن الشخص الذي «لا تشبع عينه»)، ما ساعد النموذج على فهم أن سياق ذكر كلمة **هالبنات** سياقٌ سلبيٌّ. احتوت التغريدة أيضاً على تسلسل كلمتي **وقلة** و**حياء** ذو المعنى السلبي. لذا قام النموذج بربط المعاني ليتوصل إلى أن التغريدة تصف جنساً معيناً بصفةٍ سيئةٍ مما أدى إلى قراره بتصنيفها إلى صنف خطاب الكراهية. في حين أخذت كلمة **اوفر** قيمة أهميةٍ سلبيةٍ (تساهم بالاتجاه المعاكس) وهو مشابهٌ للمثال السابق (الشكل 3). أي أن الميزات التي تعلمها المصنف بتدريبه ساعدته على التصنيف الصحيح في هذه التغريدة، وهذه الميزات متوافقةٌ مع ما ورد حتى الآن.

وقع النموذج أيضاً ببعض الأخطاء أثناء التصنيف (كما يبين الجدول 6). وبالنظر عن قربٍ إلى أحد تغريدات الكراهية هذه التي صنفت خطأً (صُنفت على أنها لا تحتوي

كراهية، الشكل 5) نلاحظ أن النموذج أعطى كلماتٍ مثل **يستشرف** و**الحشمه** قيم أهميةٍ تشد التصنيف نحو صنف اللاكراهية NOT_HS بالرغم من المعنى السلبي لكلمة **يستشرف** في هذا السياق. تمكن النموذج ربط الكلمات **ذكور** و**الشعب** و**بايع** المذكرة بسياقٍ واحدٍ استنتج أنه سلبيّ وسيء للرجال، لذا أعطاهما قيم أهميةٍ موجبةٍ، إلا أن قراره النهائي كان باختيار الصنف NOT_HS، والذي يُمكن تفسيره بأن الكلمات السياقية الخاصة بخطاب الكراهية لم تكن كافيةً للنموذج ليُصنف التغيرية تصنيفاً صحيحاً.

0.000، [CLS]



0.000، [SEP]

الشكل 5 - قيم الأهمية لكلمات لتغيرية كراهية فشل النموذج في تصنيفها

عند تحليل أحد التغيريات الأخرى التي صُنفت خطأً (على سبيل المثال: *هل تعلم ان انثى النسر تموت بعد وفاة زوجها حزنا عليه اما النسره الي بالخليج تروح تراجع دايرة التقاعد وتأخذ راتبه وتقول ماترك لنا الا الفقر*) نلاحظ أن التغيرية تتضمن استعارةً فشل النموذج في فهمها وفهم التشبيه المراد منها (أي لم يتمكن من فهم إشارة النسر هنا إلى المرأة الخليجية)، لذا صنّفها على أنها ليست خطاب كراهيةٍ في حين أنها في الحقيقة تغيرية كراهيةٍ جنسانية.

على نحوٍ مشابهٍ لما سبق، يبين الشكل 6 سحابة كلماتٍ لتكرارات الكلمات الخمسين الأكثر تأثيراً في اختيار صنف خطاب الكراهية وذلك في مجموعة التحقق:



الشكل 6 - سحابة الكلمات الأكثر تأثيراً في خطاب الكراهية في مجموعة التحقق
وفيما يلي جدولٌ يوضح الكلمات العشرة الأكثر تأثيراً منها مع تكرارات ورودها:

التكرار	قيمة الأهمية	الكلمة
1	1.59944	يهودي
1	1.01502	حريم
1	0.93122	كويتي
3	0.87833	الشباب
1	0.76018	ذكور
1	0.67762	هالمخفات
2	0.67136	خرفان
2	0.65710	البنات
1	0.65629	همجية
1	0.59632	الطلبيان

الجدول 7 - معلومات أكثر 10 كلمات في خطاب الكراهية في بيانات التحقق

يبدو من الجدول السابق ومن السحابة وجود كلمات تُشير إلى الرجال، مثل الشباب وذكور وخرقان، لم تكن بين الكلمات الأكثر تأثيراً عند تفسير بيانات التدريب، والتي كانت غالبيتها موجهةً ضد النساء، ما يعني أن النموذج تمكن من اكتشاف خطاب الكراهية الجنساني بشكلٍ عام، وليس فقط الموجه ضد النساء. كما يُلاحظ أن عدد الكلمات التي تحمل معنىً سلبياً بمفردها لا يُشكل سوى نحو 20% من الجدول (هالمخفات وهمجية) وهو أمرٌ إن دل على شيء فهو يدل على قدرة النموذج على فهم السياق العام وليس فقط الكلمات منفصلةً.

تناول هذا القسم تفسير مُخرجات النموذج المُدرّب الذي بُني بعد ضبط بارامترات ماربيرت على مجموعة بياناتٍ عربيةٍ من تويتر، وقد بينت نتائج التفسيرات انتباه النموذج إلى السياق الذي ترد ضمنه الكلمات واستخدام هذه المعرفة في التصنيف. إذ لم يعتمد على المعنى المباشر للكلمة كما هو متوقع، بل تمكن من فهم معنى التغريدة العام وربط الكلمات الأكثر تأثيراً معها. عند دراسة نقاط فشل النموذج، تبين أنه يفشل بشكلٍ أساسي عند وجود المعاني المضمنة العميقة مثل السخرية والاستعارات التي تميز اللغة العربية. ومع ذلك، تُفيد هذه المعلومة في إعطاء رؤيةٍ تفصيليةٍ تبين مواطن الضعف يمكن الاستفادة منها في تحسين دقته.

6. الاستنتاجات والتوصيات

ركز هذا البحث على استخراج أهم الميزات وأكثرها تأثيراً في اكتشاف خطاب الكراهية ضمن التغريدات العربية في اللهجات العامية.

كان استخراج الميزات ممكناً بفضل خوارزمية المُشتقات المتكاملة، التي ساعدت على تفسير قرارات النموذج عند تصنيف تغريدات الكراهية في مجموعتي التدريب والتحقق. ظهر التأثير الكبير لعددٍ من الكلمات جلياً على تصنيف التغريدة إلى صنف الكراهية. إذ كانت الكلمات المرتبطة بالجنس (ذكر أو أنثى) الأكثر تأثيراً وذات قيم الأهمية الأعلى. وهو أمرٌ متوافقٌ مع عدد التغريدات الخاصة بخطاب الكراهية الجنساني والتي تشكل نحو

نصف تغريدات صنف الكراهية في مجموعتي البيانات. من أكثر هذه الكلمات تكراراً: بنات، حريمكم، متخلفات، الشباب، خرفان وغيرها من الكلمات المرتبطة بالجنس. على نحوٍ مشابهٍ، ظهرت بعض الكلمات المرتبطة بالجنسية مثل الطليان وكويتي والرياض ذات التأثير الهام في اختيار صنف الكراهية ضمن قائمة الكلمات المؤثرة.

يدل كل ما سبق على أن النموذج تمكن من ربط كلماتٍ غير مسيئةٍ أو نابيةٍ مع مشاعر الكراهية، ما يدل على تمكنه من تشكيل فهمٍ عميقٍ للمشاعر المضمنة في خطاب الكراهية ما سمح له باكتشافه، ما يعني أنه استطاع اكتشاف خطاب الكراهية المضمن ولو بشكلٍ جزئي.

لكن، وبالرغم من نجاحات النموذج، فإنه فشل في تصنيف تغريدات الكراهية التي تحتوي صيغةً سخريةً أو استعاراتٍ لغوية. ما يعني وجود مجالٍ لتحسين قدرة النموذج على اكتشاف الأنماط الأكثر تعقيداً في خطاب الكراهية. بناءً عليه، يمكن تقديم بعض المقترحات من أجل الخطوات القادمة في هذا المجال:

- دراسة قيم الأهمية على مستوى طبقات النموذج بهدف تحديد أكثرها تأثيراً ما قد يساعد في تبسيط معمارية النماذج المعقدة دون خسارة في الأداء.
- استخدام قائمة الميزات الأكثر تأثيراً المُستخرجة على هيئة كلماتٍ مفتاحيةٍ للبحث عن تغريداتٍ قد تحتوي خطاب كراهيةٍ وبناء مجموعة بيانات ضخمةٍ تسمح بتدريب نماذجٍ أضخم وأكثر دقة.
- توظيف الميزات المؤثرة المُستخرجة من التفسير في تحسين نتائج النموذج لرفع دقة الاكتشاف وتقليل الأخطاء.

7. المراجع

- [1] Meta Reports Third Quarter 2022 Results. <https://investor.fb.com/investor-news/press-release-details/2022/Meta-Reports-Third-Quarter-2022-Results/default.aspx> (accessed Dec. 27, 2022).
- [2] HIETANEN M. *et al.*, 2023 - Towards a Definition of Hate Speech—With a Focus on Online Contexts, Journal of Communication Inquiry, vol. 47, no. 4, pp. 440–458
- [3] TONTODIMAMMA A. *et al.*, 2021 - Thirty years of research into hate speech: topics of interest and their evolution, Scientometrics, vol. 126, no. 1, pp. 157–179
- [4] JAHAN M. S. *et al.*, 2021 - A systematic review of Hate Speech automatic detection using Natural Language Processing, [arXiv:2106.00742](https://arxiv.org/abs/2106.00742) [cs]
- [5] VASWANI A. *et al.*, 2017 - Attention is All you Need, in *Advances in Neural Information Processing Systems*,
- [6] MINH D. *et al.*, 2022 - Explainable artificial intelligence: a comprehensive review, Artificial Intelligence Review, vol. 55, no. 5, pp. 3503–3568
- [7] WEBER L. *et al.*, 2023 - Beyond explaining: Opportunities and challenges of XAI-based model improvement, Information Fusion, vol. 92, pp. 154–176
- [8] ABDELHAMID M. *et al.*, 2022 - Levantine hate speech detection in twitter, Social Network Analysis and Mining, vol. 12, no. 1, p. 121
- [9] HARDAGE D. *et al.*, 2020 - Hate and Toxic Speech Detection in the Context of Covid-19 Pandemic using XAI:

Ongoing Applied Research, presented at the Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020,

- [10] MEHTA H. *et al.*, 2022 - Social Media Hate Speech Detection Using Explainable Artificial Intelligence (XAI), Algorithms, vol. 15, no. 8, p. 291
- [11] ABDELWAHAB Y. *et al.*, 2022 - Justifying Arabic Text Sentiment Analysis Using Explainable AI (XAI): LASIK Surgeries Case Study, Information, vol. 13, no. 11, p. 536
- [12] BUDUMA N. *et al.*, 2022 - Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. O'Reilly Media, 2nd edition, Beijing Boston Farnham Sebastopol Tokyo, 387 pp.
- [13] MITTAL D. *et al.*, 2023 - Enhancing Hate Speech Detection through Explainable AI, in *2023 3rd International Conference on Smart Data Intelligence (ICSMDI)*, pp.118–123
- [14] ANSARI G. *et al.*, 2023 - Data Augmentation for Improving Explainability of Hate Speech Detection, Arabian Journal for Science and Engineering
- [15] SABAH SABRY S. *et al.*, 2022 - HaT5: Hate Language Identification using Text-to-Text Transfer Transformer, *arXiv e-prints*, 10.48550/arXiv.2202.05690.
- [16] MOSTAFA A. *et al.*, 2022 - GOF at Arabic Hate Speech 2022: Breaking The Loss Function Convention For Data-Imbalanced Arabic Offensive Text Detection, in *Proceedings of the 5th Workshop on Open-Source Arabic Corpora and Processing Tools with Shared Tasks on*

- Qur'an QA and Fine-Grained Hate Speech Detection*, pp.167–175
- [17] ABDUL-MAGEED M. *et al.*, 2021 - ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic, in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp.7088–7105
- [18] DEVLIN J. *et al.*, 2019 - BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, [arXiv:1810.04805 \[cs\]](https://arxiv.org/abs/1810.04805)
- [19] KOKHLIKYAN N. *et al.*, 2020 - Captum: A unified and generic model interpretability library for PyTorch, arXiv, 10.48550/arXiv.2009.07896.
- [20] SUNDARARAJAN M. *et al.*, 2017 - Axiomatic Attribution for Deep Networks, in *Proceedings of the 34th International Conference on Machine Learning*, pp.3319–3328
- [21] ALPAYDIN E., 2021 - Machine Learning, revised and updated edition. The MIT Press, Revised, Updated edition, Cambridge, Massachusetts, 280 pp.
- [22] MUBARAK H. *et al.*, 2022 - Emojis as Anchors to Detect Arabic Offensive Language and Hate Speech, arXiv, 10.48550/arXiv.2201.06723.
- [23] BENNESSIR M. A. *et al.*, 2022 - iCompass at Arabic Hate Speech 2022: Detect Hate Speech Using QRNN and Transformers, in *Proceedings of the 5th Workshop on Open-Source Arabic Corpora and Processing Tools with*

Shared Tasks on Qur'an QA and Fine-Grained Hate Speech Detection, pp.176–180

- [24] BISONG E., 2019 - Google Colaboratory, *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, pp. 59–64. Apress, Berkeley, CA, 2019 - .