كلية هندسة الذكاء الاصطناعي
FACULTY OF ARTIFICIAL INTELLIGENCE ENGINEERING

الجامعة السورية الخاصة
SYRIAN PRIVATE UNIVERSITY

English Sign Language Recognition

Authors:

Zain Mahfouz , Ahmad al tayar , Raneem Omran

Supervised by:

Dr.Majida Albakour , Eng.Farah Fares

Damascus: 1447–1448هـ / 2025–2026م

SUPERVISOR CERTIFICATION

I certify that the preparation of this project, entitled
“English Sign Language Recognition”
prepared by
“Zain Mahfouz , Raneem Omran , Ahmad al Tayar”
was carried out under my supervision at the Faculty of Artificial Intelligence Engineering in partial fulfillment of the requirements for the degree of **B.Sc.** in the Department of Artificial Intelligence Engineering and Data Science.

Name:

Date:

Signature:

كلمة شكر وتقدير

جرى هذا المشروع في كلية هندسة الذكاء الصنعي، الجامعةالسورية الخاصة. أود أن أشكر ...

كما أتقدم بالشكر والثناء إلى الكادر التدريسي في كلية هندسة الذكاء الصنعي ولكل من مد لي يد العون لإنجاز هذا البحث.

إهداء

أود أن أعرب عن امتناني ..

ملخص المشروع:

يستخدم الأشخاص الذين يعانون من مشكلات في السمع والنطق لغة الإشارة كوزيلة أساسية للتواصل فيما بينهم ومع الآخرين داخل المجتمع. وتعتمد هذه اللغة على مجموعة من الحركات اليدوية، والإيماءات، وتعبيرات الوجه، ووضعية الجسد التي تشكل نظاماً لغوياً متكاملاً يعبر عن الكلمات والأفكار. وعلى الرغم من أهمية هذه اللغة، إلا أنها ليست معروفة على نطاق واسع، مما يؤدي إلى فجوة تواصلية تحد من اندماج الصم في الحياة الاجتماعية والمهنية.

ومع التطور السريع في تقنيات الذكاء الاصطناعي والرؤية الحاسوبية، أصبح من الممكن تطوير أنظمة ذكية قادرة على تفسير لغة الإشارة وترجمتها. وتعد أنظمة التعرف على لغة الإشارة (Sign Language Recognition Systems) من أهم تطبيقات التفاعل بين الإنسان والحاسوب، حيث انتقلت من تحليل الصور الثابتة إلى تحليل مقاطع الفيديو المتسلسلة باستخدام تقنيات التعلم العميق.(Deep Learning)

وفي هذا السياق، تبرز لغة الإشارة الأمريكية (ASL) كواحدة من أكثر لغات الإشارة دراسة واستخداماً عالمياً. ومع ذلك، يواجه الباحثون تحديات كبيرة في بناء نماذج قادرة على فهم الإشارات بمستوى الكلمات (Word-level) نظراً للتعقيد العالي في الحركات وتداخل الإيماءات.

قاعدة بيانات WLASL-100 المعتمدة :تم الاعتماد على نسخة مخصصة من قاعدة بيانات **WLASL (World-Level American Sign Language)**المتاحة عبر منصة"Kaggle" ، وبالتحديد مجموعة الـ **100فئة .(WLASL-100)** وتعتبر هذه القاعدة المعيار الذهبي الحالي للأبحاث، حيث تحتوي على مقاطع فيديو واقعية تم جمعها من مصادر متنوعة على الإنترنت لمؤدين مختلفين، مما يزيد من صعوبة المهمة وقيمة النتائج. تضم هذه المجموعة 100 كلمة (فئة) شائعة في لغة الإشارة الأمريكية، مما يوفر بيئة غنية لتدريب النماذج على التعرف على المفردات بدقة .

يساهم هذا العمل في دعم الجهود الرامية إلى تطوير أنظمة ترجمة تلقائية ذكية، مما يساعد على تحسين التواصل بين الأشخاص الصم وبقية أفراد المجتمع، ويعزز من فرص دمجهم في مختلف مجالات الحياة من خلال تقليل الاعتماد على المترجمين البشريين وتوفير حلول تقنية فورية.

Abstract:

People with hearing and speech impairments use sign language as their primary means of communication with each other and with others in their communities. This language relies on a set of hand movements, gestures, facial expressions, and body postures that form a complete linguistic system expressing words and ideas. Despite its importance, sign language is not widely known, leading to a communication gap that limits the social and professional integration of deaf individuals. With the rapid development of artificial intelligence and computer vision technologies, it has become possible to develop intelligent systems capable of interpreting and translating sign language. Sign language recognition systems are among the most important applications of human-computer interaction, having evolved from analyzing still images to analyzing sequential video clips using deep learning techniques. In this context, American Sign Language (ASL) stands out as one of the most studied and used sign languages globally. However, researchers face significant challenges in building models capable of understanding word-level signs due to the high complexity of the movements and the overlap of gestures. The WLASL-100 Database:

In this work, a customized version of the WLASL (World-Level American Sign Language) database available through the Kaggle platform was used, specifically the 100-category set (WLASL-100). This database is considered the current gold standard for research, as it contains real-life videos collected from various online sources featuring different performers, thus increasing the difficulty of the task and the value of the results. This set includes 100 common words (categories) in American Sign Language, providing a rich environment for training models to accurately recognize vocabulary. This work contributes to efforts to develop intelligent automatic translation systems, which help improve communication between deaf people and the rest of society and enhance their inclusion in various aspects of life by reducing reliance on human interpreters and providing immediate technological solutions.

فهرس المحتويات

Contents

فهرس المحتويات..... 7
فهرس الجداول..... 10
فهرس الأشكال..... 11
فهرس المعادلات..... 12
فهرس المصطلحات والرموز..... 13
الفصل الأول: تمهيد..... 15
1- مقدمة.....16
1-2- المشكلة العلمية في مشروع البحث.....16
1-3- هدف البحث.....16
1-4- مساهمة البحث.....16
1-5- الجهات المستفيدة من البحث.....17
1-6- تنظيم الأطروحة.....17
الفصل الثاني: الدراسة النظرية..... 18
2-1مقدمة.....18
2-2**التعرف على لغة الإشارة(Sign Language Recognition)**.....18
2-3رؤية الحاسوب(Computer Vision).....18
2-4النقاط الرئيسية(Keypoints Extraction).....19
2-5التعلم العميق(Deep Learning).....19
2-6الشبكات العصبية المتكررة(Recurrent Neural Networks - RNN).....19
2-7شبكات(LSTM (Long Short-Term Memory.....20
2-8نموذج(Bi-LSTM (Bidirectional LSTM.....20
2-9معالجة البيانات الزمنية(Time-Series Processing).....20
2-10خاتمة الفصل.....21
الفصل الثالث: الدراسة المرجعية..... 22
3-1مقدمة.....22
3-2الدراسات السابقة في المجالات المشابهة.....22
الدراسة الأولى(2023).....22
الدراسة الثانية(2022).....22

الدراسة الثالثة(2024) 23
الدراسة الرابعة(2023) 23
الدراسة الخامسة(2022) 23
3-3التقنيات والأطر الحديثة في المجال 23
3-4التحليل المقارن للأنظمة 24
3-5الفجوة البحثية والابتكار المقترح 24
الحل المقترح في هذا البحث 24
3-6خاتمة الفصل 25
الفصل الرابع: منهج البحث 26
4-1مقدمة 27
4-2الطرق التقليدية لحل المسألة المطروحة 27
4-2-1الطرق العددية(Numerical Methods) 27
4-2-2طرق تعلم الآلة(Machine Learning) 27
4-3النموذج المقترح 27
4-3-1وصف النموذج 27
4-3-2تابع الخسارة وخوارزمية التدريب 27
4-3-3شرح الPipeline 27
4-4مجموعات البيانات المستخدمة 27
خطوات المعالجة المسبقة: 28
الأدوات المستخدمة: 28
4-5خاتمة الفصل والمساهمة البحثية 28
الفصل الخامس: نتائج البحث 29
5-1مقدمة 30
5-2الأدوات المستعملة وأساليب التقييم المتبعة 30
الأدوات المستخدمة 30
معايير التقييم 30
5-3نتائج النموذج المقترح 31
5-4مقارنة النتائج 31
5-4-1مقارنة نتائج النماذج المقترحة 31
5-4-2مقارنة النتائج مع الأعمال السابقة 32
5-5تحليل النتائج 32
نقاط القوة 32
نقاط الضعف 32

تحسينات ممكنة 32
6-5خاتمة الفصل 32
الفصل السادس: الخاتمة 35
6-1الاستنتاجات 36
6-2التحديات والآفاق المستقبلية 36
التحديات..... 36
الآفاق المستقبلية..... 36

فهرس الجداول

ترد الجداول في ا لأطروحة بحيث يكون لكل جدول عنوان ورقم متسلسل مرتبط بالفصل الذي يحتويه ويكتب العنوان في أعلى الجدول

جدول 3.1:عنوان الجدول .....رقم صفحة ورود الجدول

جدول 3.2: عنوان الجدول.....**Error! Bookmark not defined.**

جدول 3.3: عنوان الجدول.....**Error! Bookmark not defined.**

فهرس الأشكال

ترد الجداول في الأطروحة بحيث يكون لكل شكل عنوان ورقم متسلسل مرتبط بالفصل الذي يحتويه ويكتب العنوان في أسفل الشكل

شكل 2.1: عنوان الشكل ..... رقم الصفحة التي ورد فيها
شكل 2.2: عنوان الشكل .....**Error! Bookmark not defined.**
شكل 2.3: عنوان الشكل. .....**Error! Bookmark not defined.**
شكل 2.4: عنوان الشكل .....**Error! Bookmark not defined.**
شكل 2.5: .....**Error! Bookmark not defined.**

فهرس المعادلات

ترد المعادلات في ا لأطروحة بحيث يكون لكل معادلة رقم يوضع بين قوسين يربطها بالفصل الذي يحتويها

المعادلة [2.1]:نص المعادلة ..... رقم الصفحةالتي ورد فيها
المعادلة [2.2]:نص المعادلة .....**Error! Bookmark not defined.**
المعادلة [2.2]:نص المعادلة .....**Error! Bookmark not defined.**
المعادلة [2.2]:نص المعادلة .....**Error! Bookmark not defined.**

فهرس المصطلحات والرموز

رمز/ اختصار	المقابل بالإنكليزية	المصطلح
ML	Machine Learning	تعلم الآلة
-	Supervised Learning	التعلم الخاضع للإشراف
-	Unsupervised Learning	التعلم غير الخاضع للإشراف
-	Overfitting	الإفراط في التعلمُ
-	Underfitting	ضعف التعلمُ
-	Epoch	تكرار التدريب
-	Learning Rate	معدل التعلمّ
-	Loss Function	دالة خسارة
-	Optimizer	المُحسِّ ّن
NN	Neural Network	الشبكة العصبية
DL	Deep Learning	التعلمُ العميق
-	Classification	التصنيف
-	Regression	الانحدار
-	Accuracy	الدقة
-	Pandas	باندا س لمعالجة البيانات
-	NumBy	الحوسبة العددية
-	matplotlib	مكتبة الرسوم البيانية
-	Seaborn	التصوير الإحصائي للبيانات
-	Scikit-learn	-
-	TensorFlow	-
SVM	Support Vector Machine	الة دعم المتجهات
RF	Random Forest	الغابة العشوائية
DT	Decision Tree	شجرة القرار
R^2	Coefficient of Determination	معامل التحديد
RMSE	Root Mean Squared Error	الجذر التربيعي لمتوسط مربع الخطأ
SD	Standard Deviation	الانحراف المعياري

رمز/ اختصار	المقابل بالإنكليزية	المصطلح
MAE	Mean Absolute Error	متوسط الخطأ المطلق
MSE	Mean Squared Error	متوسط مربع الخطأ
-	Precision	الدقة
-	Recall	معدل الاسترجاع
-	F1-Score	درجة F1
-	Voting Classifier	مصنف التصويت
-	Ensemble Learning	التعلم التجميعي
PSO	Particle Swarm Optimization	خوارزميات السرب
LSTM	Long Short-Term Memory	الذاكرة طويلة قصيرة المدى ثنائية الاتجاه
Bi-LSTM	Bidirectional Long Short-Term Memory	الذاكرة طويلة قصيرة المدى
CNN	Convolutional Neural Networks	الشبكات العصبية الالتفافية
RNN	Recurrent Neural Networks	الشبكات العصبية التكرارية
-	Autoencoder	المُرمِّز التلقائي
MLP	Multi-Layer Perceptron	الشبكات العصبية متعددة الطبقات
-	Transformers	نماذج المحولات
GRU	Graph Convolutional Network	شبكات التلافيف المخططية

الفصل الأول: تمهيد

1-1- مقدمة 16
1-2- المشكلة العلمية في مشروع البحث 16
1-3- هدف البحث 16
1-4- مساهمة البحث 16
1-5- الجهات المستفيدة من البحث 17
1-6- تنظيم الأطروحة 17

-1- مقدمة

تُعد لغة الإشارة من الوسائل الأساسية التي يعتمد عليها الأشخاص ذوو الإعاقة السمعية في التواصل مع الآخرين، حيث تعتمد على مجموعة من الحركات اليدوية وتعابير الوجه التي تشكل نظامًا لغويًا متكاملًا. وعلى الرغم من أهميتها، إلا أن استخدامها يظل محدودًا ضمن فئة معينة من المجتمع، مما يؤدي إلى وجود فجوة تواصلية واضحة بين مستخدمي هذه اللغة وبقية الأفراد.

مع التطور الكبير في مجالات الذكاء الاصطناعي ورؤية الحاسوب، أصبح من الممكن تطوير أنظمة ذكية قادرة على تحليل الإشارات الحركية وفهمها بشكل آلي. وقد ساهمت تقنيات التعلم العميق بشكل خاص في تحسين قدرة هذه الأنظمة على التعامل مع البيانات المعقدة، مثل البيانات الزمنية الناتجة عن تسلسل الحركات.

في هذا السياق، يركز هذا المشروع على تطوير نموذج يعتمد على التعلم العميق للتعرف على إشارات اليد باستخدام بيانات مستخرجة من مقاطع فيديو، بهدف المساهمة في تقليل الفجوة التواصلية وتحسين دمج الأشخاص الصم في المجتمع.

1-2- المشكلة العلمية في مشروع البحث

تكمن المشكلة الأساسية في صعوبة تفسير الإشارات الحركية بشكل دقيق، نظرًا لطبيعتها الزمنية وتعقيدها، حيث تتداخل الحركات وتتغير بسرعة. بالإضافة إلى ذلك، فإن الأنظمة الحالية تعاني من عدة تحديات، منها:

- محدودية الدقة في التعرف على الإشارات
- صعوبة التعامل مع البيانات الزمنية الطويلة
- الحاجة إلى نماذج قادرة على فهم السياق الكامل للحركة

بناءً على ذلك، تظهر الحاجة إلى تطوير نموذج قادر على تحليل تسلسل الإشارات بشكل أكثر كفاءة ودقة.

1-3- هدف البحث

لهدف الرئيسي:
تطوير نظام ذكي قادر على التعرف على إشارات اليد اعتمادًا على تقنيات التعلم العميق.

الأهداف الفرعية:

- تحسين دقة التعرف على الإشارات
- بناء نموذج قادر على تحليل البيانات الزمنية بكفاءة
- تقليل زمن الاستجابة أثناء التنبؤ
- توفير نموذج قابل للتطوير والتطبيق العملي

1-4- مساهمة البحث

تتمثل مساهمة هذا البحث في:

- استخدام نموذج Bi-LSTM لمعالجة البيانات الزمنية الخاصة بالإشارات
- بناء Pipeline متكامل لمعالجة البيانات بدءًا من استخراج النقاط وانتهاءً بالتصنيف
- تحسين فهم الإشارات من خلال تحليلها في اتجاهين (أمامي وخلفي)

- تقديم نموذج يمكن تطويره لاحقًا لتطبيقات عملية

1-5- الجهات المستفيدة من البحث

يمكن أن تستفيد عدة جهات من هذا المشروع، منها:

- الأشخاص ذوو الإعاقة السمعية
- المؤسسات التعليمية
- مراكز الأبحاث في الذكاء الاصطناعي
- مطورو التطبيقات الذكية

1-6- تنظيم الأطروحة

تم تنظيم هذا البحث على الشكل التالي:

- يتناول الفصل الأول تمهيدًا عامًا حول موضوع البحث
- يعرض الفصل الثاني الأسس النظرية والتقنيات المستخدمة
- يستعرض الفصل الثالث الدراسات السابقة وتحليلها
- يوضح الفصل الرابع منهجية العمل والنموذج المقترح
- يعرض الفصل الخامس النتائج والتقييم
- يختتم الفصل السادس البحث بالاستنتاجات والتوصيات المستقبلية

الفصل الثاني: الدراسة النظرية

2-1مقدمة

يستند مشروعنا على مجموعة من المفاهيم والتقنيات الأساسية في مجالات الذكاء الاصطناعي، والتعلم العميق، ورؤية الحاسوب، حيث يتم توظيف هذه المجالات معًا لبناء نظام قادر على فهم وتحليل الإشارات الحركية. ونظرًا للطبيعة الزمنية المعقدة للبيانات المستخدمة في هذا النوع من التطبيقات، فإن اختيار النماذج والخوارزميات المناسبة يُعد عنصرًا حاسمًا في تحقيق أداء دقيق وموثوق.

في هذا الفصل، سيتم عرض الأسس النظرية التي يعتمد عليها المشروع، بدءًا من التعرف على لغة الإشارة، مرورًا بمفاهيم التعلم العميق، وصولًا إلى النماذج المستخدمة في معالجة البيانات الزمنية.

2-2التعرف على لغة الإشارة(Sign Language Recognition)

يُعرف التعرف على لغة الإشارة بأنه عملية تحليل الإشارات الحركية وتحويلها إلى تمثيل مفهوم، مثل نص أو صوت. وتعتمد هذه العملية على تفسير مجموعة من العناصر، منها:

- حركات اليد
- وضعية الجسم
- تعابير الوجه

وتنقسم أنظمة التعرف على لغة الإشارة إلى نوعين رئيسيين:

- **أنظمة تعتمد على الصور الثابتة** :تقوم بتحليل كل صورة بشكل منفصل، وتناسب الحالات البسيطة .
- **أنظمة تعتمد على الفيديو (التسلسل الزمني)** :تقوم بتحليل سلسلة من الإطارات، وهي الأكثر دقة نظرًا لقدرتها على فهم تطور الحركة عبر الزمن .

ويُعد هذا المشروع من النوع الثاني، حيث يعتمد على تحليل تسلسل زمني للحركات.

2-3رؤية الحاسوب(Computer Vision)

تُعد رؤية الحاسوب أحد الفروع الأساسية في الذكاء الاصطناعي، وتهدف إلى تمكين الحاسوب من فهم وتحليل الصور والفيديوهات بطريقة مشابهة للرؤية البشرية. وفي سياق هذا المشروع، يتم استخدام تقنيات رؤية الحاسوب لاستخراج المعلومات المهمة من الفيديو، مثل:

- تحديد مواقع اليد
- استخراج النقاط الرئيسية (Keypoints)
- تتبع الحركة عبر الزمن

وتُحول هذه البيانات لاحقًا إلى تمثيل رقمي يمكن استخدامه في تدريب النماذج.

2-4النقاط الرئيسية(Keypoints Extraction)

تعتمد العديد من أنظمة التعرف على الإشارات على استخراج النقاط الرئيسية من جسم الإنسان، وخاصة اليدين. وتمثل هذه النقاط إحداثيات مفاصل معينة، مثل:

- مفاصل الأصابع
- المعصم
- أجزاء من الذراع

ويتم تمثيل كل نقطة بإحداثيات (x, y) أو(x, y, z) ، مما يسمح بتحويل الحركة إلى بيانات رقمية منظمة.

تُستخدم هذه الطريقة لأنها:

- تقلل من حجم البيانات مقارنة بالفيديو الكامل
- تركز على المعلومات المهمة فقط
- تسهل عملية التدريب للنماذج

2-5التعلم العميق(Deep Learning)

يُعتبر التعلم العميق أحد فروع تعلم الآلة، ويعتمد على الشبكات العصبية الاصطناعية ذات الطبقات المتعددة. ويتميز بقدرته على استخراج الأنماط المعقدة من البيانات دون الحاجة إلى تدخل بشري كبير.

في هذا المشروع، يتم استخدام التعلم العميق لتحليل البيانات الزمنية الناتجة عن تسلسل الإشارات، حيث يقوم النموذج بتعلم العلاقات بين الحركات المختلفة من أجل تصنيفها بشكل صحيح.

2-6الشبكات العصبية المتكررة(Recurrent Neural Networks - RNN)

تم تصميم الشبكات العصبية المتكررة للتعامل مع البيانات المتسلسلة، حيث تأخذ بعين الاعتبار ترتيب البيانات واعتماد كل عنصر على العناصر السابقة.

تتميز هذه الشبكات بوجود ذاكرة داخلية تسمح لها بالاحتفاظ بالمعلومات السابقة، مما يجعلها مناسبة لمعالجة:

- النصوص
- الإشارات الزمنية
- الفيديوهات

إلا أنها تعاني من مشكلة تُعرف بتلاشي التدرج(Vanishing Gradient) ، مما يؤثر على قدرتها في تعلم العلاقات طويلة المدى.

2-7شبكات(LSTM (Long Short-Term Memory

تم تطوير شبكات LSTM كحل لمشاكل الشبكات المتكررة التقليدية، حيث تتميز بقدرتها على الاحتفاظ بالمعلومات لفترات زمنية طويلة.

تعتمد LSTM على بنية داخلية تتكون من ثلاث بوابات رئيسية:

- **بوابة النسيان :(Forget Gate)** تحدد المعلومات التي يجب تجاهلها
- **بوابة الإدخال :(Input Gate)** تحدد المعلومات التي يجب إضافتها
- **بوابة الإخراج :(Output Gate)** تحدد المعلومات التي سيتم تمريرها للخطوة التالية

تسمح هذه البنية للنموذج بالتحكم في تدفق المعلومات، مما يحسن من أدائه في معالجة البيانات الزمنية المعقدة.

2-8نموذج(Bi-LSTM (Bidirectional LSTM

يُعد نموذج Bi-LSTM تطويرًا لنموذجLSTM ، حيث يقوم بمعالجة البيانات في اتجاهين:

- من البداية إلى النهاية
- من النهاية إلى البداية

ثم يتم دمج النتائج للحصول على تمثيل أكثر شمولًا للسلسلة الزمنية.

تكمن أهمية هذا النموذج في أنه:

- يأخذ بعين الاعتبار الماضي والمستقبل
- يحسن دقة التنبؤ
- مناسب جدًا لتحليل الإشارات الحركية

ولذلك تم اعتماده كنموذج أساسي في هذا المشروع.

2-9معالجة البيانات الزمنية(Time-Series Processing)

تُعد البيانات المستخدمة في هذا المشروع من نوع البيانات الزمنية، حيث تمثل كل إشارة سلسلة من القيم المرتبة زمنيًا. ويتطلب التعامل مع هذا النوع من البيانات:

- الحفاظ على ترتيب القيم
- فهم العلاقة بين النقاط المتتالية
- التعامل مع اختلاف طول التسلسلات

ويتم تجهيز هذه البيانات قبل إدخالها إلى النموذج من خلال:

- توحيد طول التسلسلات

- تنظيف البيانات
- تنظيمها ضمن مصفوفات مناسبة للتدريب

2-10خاتمة الفصل

تم في هذا الفصل استعراض المفاهيم النظرية الأساسية التي يعتمد عليها هذا البحث، بدءًا من التعرف على لغة الإشارة، مرورًا برؤية الحاسوب واستخراج النقاط الرئيسية، وصولًا إلى نماذج التعلم العميق المستخدمة في تحليل البيانات الزمنية.

تشكل هذه الأسس قاعدة علمية متينة سيتم الاعتماد عليها في الفصل التالي، الذي يتناول الدراسات السابقة وتحليل الأعمال المتعلقة بهذا المجال، بهدف تحديد الفجوة البحثية وبناء الحل المقتر

الفصل الثالث: الدراسة المرجعية

3-1مقدمة

يهدف هذا الفصل إلى استعراض وتحليل مجموعة من الدراسات الحديثة في مجال التعرف على لغة الإشارة باستخدام تقنيات الذكاء الاصطناعي، وذلك لإبراز التطورات التي شهدها هذا المجال خلال السنوات الأخيرة. كما يهدف إلى تحديد نقاط القوة والضعف في الأعمال السابقة، واستخلاص الفجوة البحثية التي يسعى هذا المشروع إلى معالجتها.

تم التركيز على الدراسات التي تعتمد على التعلم العميق وتحليل البيانات الزمنية، نظرًا لارتباطها المباشر بموضوع هذا البحث.

3-2الدراسات السابقة في المجالات المشابهة

الدراسة الأولى(2023)

قدّمت مجموعة بحثية في عام 2023 نموذجًا يعتمد على الشبكات العصبية الالتفافية (CNN) لاستخراج الميزات من صور الإشارات، متبوعًا بنموذج LSTM لتحليل التسلسل الزمني للحركة.

- **المساهمة :**دمج بين تحليل الصور والتسلسل الزمني
- **النموذج المستخدم** CNN + LSTM
- **المعايير :**الدقة(Accuracy) ، الاستدعاء (Recall)
- **النتائج :**حقق النموذج دقة تجاوزت 88 %
- **نقاط القوة :**قدرة جيدة على استخراج الميزات
- **نقاط الضعف :**الأداء يتأثر بجودة الصور وعدم الاستقرار في الفيديو

الدراسة الثانية(2022)

اقترحت دراسة أخرى استخدام نموذج ثلاثي الأبعاد (3D-CNN) لتحليل مقاطع الفيديو مباشرة دون الحاجة إلى فصل الإطارات.

- **المساهمة :**معالجة الفيديو بشكل متكامل
- **النموذج المستخدم** 3D-CNN
- **المعايير :**الدقة وزمن التنفيذ
- **النتائج :**أداء جيد مع دقة تقارب 85 %
- **نقاط القوة :**التقاط المعلومات الزمنية والمكانية معًا
- **نقاط الضعف :**استهلاك عالٍ للموارد الحاسوبية

الدراسة الثالثة(2024)

استخدمت هذه الدراسة نموذج Transformer لمعالجة تسلسل الإشارات، مستفيدة من آلية الانتباه (Attention Mechanism).

- **المساهمة :**إدخال تقنيات حديثة لمعالجة التسلسل
- **النموذج المستخدم** Transformer
- **المعايير :**الدقة، F1-Score
- **النتائج :**دقة وصلت إلى 91 %
- **نقاط القوة :**فهم العلاقات طويلة المدى
- **نقاط الضعف :**تعقيد النموذج والحاجة إلى بيانات كبيرة

الدراسة الرابعة(2023)

اعتمدت هذه الدراسة على استخراج النقاط الرئيسية (Keypoints) من الجسم، واستخدام نموذج LSTM لتحليلها.

- **المساهمة :**تقليل حجم البيانات والتركيز على المعلومات المهمة
- **النموذج المستخدم** LSTM
- **المعايير :**الدقة
- **النتائج :**دقة حوالي 87 %
- **نقاط القوة :**كفاءة في المعالجة
- **نقاط الضعف :**عدم الاستفادة من السياق المستقبلي

الدراسة الخامسة(2022)

تم في هذه الدراسة استخدام نموذج Bi-LSTM لتحليل البيانات الزمنية المستخرجة من الفيديو.

- **المساهمة :**تحليل الإشارات في اتجاهين
- **النموذج المستخدم** Bi-LSTM :
- **المعايير** Accuracy :و Precision
- **النتائج :**تحسن ملحوظ في الدقة مقارنة بـ LSTM
- **نقاط القوة :**فهم أفضل للتسلسل
- **نقاط الضعف :**زيادة التعقيد الحسابي

3-3التقنيات والأطر الحديثة في المجال

شهد مجال التعرف على لغة الإشارة تطورًا ملحوظًا بفضل ظهور تقنيات حديثة، من أبرزها:

- **Transformers:**
تعتمد على آلية الانتباه، وتتميز بقدرتها على معالجة العلاقات طويلة المدى دون الحاجة إلى تسلسل زمني صارم .

- **Graph Neural Networks (GNN):**
تُستخدم لتمثيل العلاقات بين النقاط الرئيسية (Keypoints) كشبكة مترابطة، مما يحسن فهم البنية الحركية .
- **نماذج التعلم العميق الهجينة:**
تجمع بين أكثر من نموذج (مثل CNN و (LSTMللاستفادة من مزايا كل منها .

تُظهر هذه التقنيات توجهًا نحو تحسين دقة النماذج مع تقليل الاعتماد على المعالجة اليدوية للبيانات.

3-4التحليل المقارن للأنظمة

فيما يلي مقارنة بين بعض الأنظمة المستخدمة في التعرف على لغة الإشارة:

النظام	التقنيات المستخدمة	نقاط القوة	نقاط الضعف
CNN + LSTMنظام يعتمد على	استخراج ميزات + تحليل زمني	دقة جيدة	حساس للضوضاء
نظام D-CNN3	تحليل فيديو مباشر	فهم مكاني وزمني	استهلاك موارد عالي
Transformer	Attention Mechanism	دقة عالية	يحتاج بيانات كبيرة
Bi-LSTM	فهم شامل للتسلسل تحليل ثنائي الاتجاه		تعقيد أعلى

3-5الفجوة البحثية والابتكار المقترح

من خلال تحليل الدراسات السابقة، يمكن ملاحظة ما يلي:

- العديد من النماذج تعاني من صعوبة في فهم التسلسل الزمني بشكل كامل
- بعض النماذج تستهلك موارد حاسوبية كبيرة
- قلة الدراسات التي توازن بين الدقة والكفاءة

بناءً على ذلك، تتمثل الفجوة البحثية في الحاجة إلى نموذج:

- قادر على فهم التسلسل الزمني بشكل شامل
- لا يتطلب موارد حاسوبية كبيرة
- يحقق توازنًا بين الدقة وسرعة التنفيذ

الحل المقترح في هذا البحث

يقترح هذا المشروع استخدام نموذج **Bi-LSTM**لمعالجة البيانات الزمنية المستخرجة من النقاط الرئيسية(Keypoints) ، حيث يتيح هذا النموذج:

- تحليل الإشارات في اتجاهين
- تحسين دقة التعرف
- تقليل التعقيد مقارنة بالنماذج الأكثر حداثة مثل Transformers

كما يعتمد المشروع على معالجة مسبقة فعالة للبيانات، مما يسهم في تحسين جودة المدخلات وبالتالي رفع كفاءة النموذج.

3-6خاتمة الفصل

تم في هذا الفصل استعراض مجموعة من الدراسات الحديثة في مجال التعرف على لغة الإشارة، وتحليل النماذج المستخدمة فيها من حيث الأداء ونقاط القوة والضعف. كما تم عرض أبرز التقنيات الحديثة المستخدمة في هذا المجال.

وبناءً على هذا التحليل، تم تحديد الفجوة البحثية التي يسعى هذا المشروع إلى معالجتها، والتي تتمثل في تحقيق توازن بين الدقة والكفاءة في معالجة البيانات الزمنية.

سيتم في الفصل التالي عرض منهجية البحث والنموذج المقترح بشكل تفصيلي، مع شرح مراحل العمل بدءًا من معالجة البيانات وصولًا إلى تدريب النموذج وتقييمه.

الفصل الرابع: منهج البحث

4-1مقدمة..... 27
4-2الطرق التقليدية لحل المسألة المطروحة..... 27
4-2-1الطرق العددية(Numerical Methods) 27
4-2-2طرق تعلم الآلة(Machine Learning) 27
4-3النموذج المقترح..... 27
4-3-1وصف النموذج..... 27
4-3-2تابع الخسارة وخوارزمية التدريب 27
4-3-3شرح الـPipeline 27
4-4مجموعات البيانات المستخدمة..... 27
خطوات المعالجة المسبقة:..... 28
الأدوات المستخدمة:..... 28
4-5خاتمة الفصل والمساهمة البحثية..... 28

4-1مقدمة

يهدف هذا الفصل إلى عرض المنهجية المتبعة في تطوير النظام المقترح للتعرف على لغة الإشارة، بدءًا من دراسة الطرق التقليدية، مرورًا باختيار النموذج المناسب، وصولًا إلى شرح مراحل بناء النظام بشكل متكامل.

يعتمد هذا البحث على استخدام تقنيات التعلم العميق لمعالجة بيانات زمنية مستخرجة من إشارات اليد، حيث تم تصميم Pipeline متكامل يشمل استخراج البيانات، معالجتها، تدريب النموذج، ومن ثم تقييم أدائه.

4-2الطرق التقليدية لحل المسألة المطروحة

4-2-1الطرق العددية(Numerical Methods)

تعتمد الطرق التقليدية في تحليل الإشارات الحركية على تمثيل البيانات بشكل رياضي واستخدام خوارزميات عددية لاستخلاص الأنماط. ومن هذه الطرق:

- تحليل الإشارات باستخدام الاشتقاق والتكامل
- استخدام تقنيات تصفية البيانات (Filtering)
- استخراج الخصائص الإحصائية

إلا أن هذه الطرق تعاني من محدودية في التعامل مع البيانات المعقدة وغير الخطية، مثل الإشارات الحركية، مما يقلل من فعاليتها في هذا النوع من التطبيقات.

4-2-2طرق تعلم الآلة(Machine Learning)

تم استخدام خوارزميات تعلم الآلة التقليدية مثل:

- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)
- Decision Trees

تعتمد هذه الطرق على استخراج ميزات يدوية(Handcrafted Features) ، مما يتطلب خبرة مسبقة ويحد من قدرة النموذج على التعميم.

بالمقارنة مع التعلم العميق، فإن هذه الطرق أقل كفاءة في التعامل مع البيانات الزمنية المعقدة.

4-3النموذج المقترح

4-4مجموعات البيانات المستخدمة

تم استخدام بيانات تعتمد على تمثيل الإشارات من خلال النقاط الرئيسية، حيث تتضمن:

- إحداثيات مفاصل اليد والجسم
- تمثيل زمني للحركة
- بيانات مخزنة بصيغة npy.

خطوات المعالجة المسبقة:

- استخراج Keypoints من الفيديو
- تحويل البيانات إلى سلاسل زمنية
- إزالة القيم غير الصالحة
- توحيد طول التسلسلات

الأدوات المستخدمة:

تم استخدام MediaPipe لاستخراج الKeypoint من الفيديو بلاضافة لاستخدام نماذج Bidirectional Long Short Term Memory (BiLSTM) و Graph Convolutional Networks (GRU) و Convolutional Neural Networks (CNN) .

ساهمت هذه النماذج في تحسين جودة البيانات المدخلة للنموذج، مما انعكس بشكل إيجابي على الأداء النهائي.

4-5خاتمة الفصل والمساهمة البحثية

تم في هذا الفصل عرض المنهجية الكاملة المتبعة في هذا البحث، بدءًا من استعراض الطرق التقليدية، وصولًا إلى بناء النموذج المقترح وشرح مراحل العمل بشكل تفصيلي.

تكمن المساهمة الأساسية في هذا الفصل في:

- تصميم نموذج يعتمد علىBi-LSTM و GRU و CNN لمعالجة البيانات الزمنية
- بناء Pipeline متكامل لمعالجة البيانات
- تحسين تمثيل الإشارات باستخدام النقاط الرئيسية

تشكل هذه المنهجية الأساس الذي سيتم من خلاله تقييم أداء النموذج في الفصل التالي، والذي يتناول النتائج والتحليل المقارن.

الفصل الخامس: نتائج البحث

5-2الأدوات المستعملة وأساليب التقييم المتبعة..... 30
الأدوات المستخدمة..... 30
معايير التقييم..... 30
5-3نتائج النموذج المقترح 31
5-4مقارنة النتائج 31
5-4-1مقارنة نتائج النماذج المقترحة 31
5-4-2مقارنة النتائج مع الأعمال السابقة..... 32
5-5تحليل النتائج..... 32
نقاط القوة 32
نقاط الضعف..... 32
تحسينات ممكنة..... 32
5-6خاتمة الفصل 32

5-1مقدمة

يهدف هذا الفصل إلى عرض نتائج النموذج المقترح وتقييم أدائه باستخدام مجموعة من المعايير المناسبة لطبيعة المسألة. كما يتضمن مقارنة بين أداء النموذج المقترح وبعض النماذج الأخرى، بالإضافة إلى تحليل نقاط القوة والضعف.

تم إجراء عملية التقييم اعتمادًا على بيانات اختبار مستقلة، وذلك لضمان موضوعية النتائج وعدم تأثرها ببيانات التدريب.

5-2الأدوات المستعملة وأساليب التقييم المتبعة

الأدوات المستخدمة

تم تنفيذ النموذج باستخدام بيئة برمجية تعتمد على لغةPython ، مع استخدام مكتبات متخصصة في التعلم العميق ومعالجة البيانات، مثل:

- TensorFlow / Kerasلبناء النموذج
- NumPyلمعالجة البيانات
- Label Encoder
-

معايير التقييم

تم استخدام مجموعة من المعايير الشائعة في مسائل التصنيف، وهي:

- **الدقة:(Accuracy)**
  تمثل نسبة التنبؤات الصحيحة إلى إجمالي التنبؤات و تعطى بلعلاقة:

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})}$$

- **الاستدعاء:(Recall)**
  يقيس قدرة النموذج على اكتشاف جميع الحالات الصحيحة و تعطى بلعلاقة:

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

الدقة النوعية:(Precision)
تعكس نسبة التنبؤات الصحيحة من بين التنبؤات الإيجابية و تعطى بلعلاقة:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}$$

مقياس:F1
وهو المتوسط التوافقي بين Precision وRecall ، ويعطي تقييمًا متوازنًا للأداء و يعطى بلعلاقة:

$$\text{F1} - \text{Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

تُعد هذه المعايير مناسبة لهذا النوع من المسائل، لأنها توفر فهمًا شاملًا لأداء النموذج، خاصة في حال وجود تباين بين الفئات.

5-3نتائج النموذج المقترح

أظهرو النماذج المستخدمة أداءً جيدًا في التعرف على الإشارات، حيث تمكن من:

- التقاط الأنماط الزمنية للحركات بشكل فعال
- التمييز بين الإشارات المختلفة بدرجة مقبولة من الدقة
- التعامل مع تسلسلات بيانات ذات أطوال مختلفة

كما ان استخدام طبقات CNNوGRU و Bi-LSTMساهم في تحسين جودة التنبؤات، نتيجة قدرته على تحليل التسلسل في اتجاهين، مما وفر فهمًا أعمق للسياق الحركي.

بشكل عام، حققو النماذج استقرارًا جيدًا أثناء التدريب، مع انخفاض تدريجي في قيمة تابع الخسارة، مما يدل على فعالية عملية التعلم.

5-4مقارنة النتائج

5-4-1مقارنة نتائج النماذج المقترحة

تمت مقارنة أداء نموذج Bi-LSTM مع نموذج LSTM التقليدي، حيث أظهرت النتائج أن:

- نموذج Bi-LSTM حقق دقة أعلى في التعرف على الإشارات
- كان أكثر قدرة على التعامل مع التعقيد الزمني للحركات
- قدم نتائج أكثر استقرارًا عبر بيانات الاختبار

كما ان تمت مقارنة استخدام طبقة Bi-LSTM فقط مع استخدام طبقات CNN و GRU و Bi-LSTM حيث اظهرت النتائج أن:

- استخدام طبقات من نماذج مختلفة اعط

ويعود هذا التحسن إلى قدرة النموذج على الاستفادة من المعلومات المستقبلية بالإضافة إلى المعلومات السابقة.

5-4-2مقارنة النتائج مع الأعمال السابقة

عند مقارنة النموذج المقترح مع بعض الأعمال السابقة، يمكن ملاحظة ما يلي:

- يحقق النموذج أداءً مقاربًا للنماذج الحديثة المعتمدة على التعلم العميق
- يتميز بكفاءة أعلى من حيث استهلاك الموارد مقارنة بالنماذج المعقدة مثل 3 D-CNNو Transformers
- يقدم توازنًا جيدًا بين الدقة وسرعة التنفيذ

5-5تحليل النتائج

من خلال النتائج التي تم الحصول عليها، يمكن استخلاص عدد من الملاحظات:

نقاط القوة

- قدرة جيدة على تحليل البيانات الزمنية
- أداء مستقر أثناء التدريب والاختبار
- كفاءة في استخدام الموارد الحاسوبية
- بساطة نسبية في البنية مقارنة بالنماذج الحديثة

نقاط الضعف

- يعتمد الأداء بشكل كبير على جودة البيانات المدخلة
- قد يواجه صعوبة في التمييز بين الإشارات المتشابهة
- محدودية في التعامل مع عدد كبير جدًا من الفئات

تحسينات ممكنة

- زيادة حجم البيانات المستخدمة في التدريب
- استخدام تقنيات تحسين مثل Data Augmentation
- دمج النموذج مع تقنيات أحدث مثل Attention Mechanism

5-6خاتمة الفصل

تم في هذا الفصل عرض نتائج النموذج المقترح وتحليل أدائه باستخدام مجموعة من المعايير المناسبة. كما تمت مقارنة النتائج مع نماذج أخرى ومع بعض الأعمال السابقة، مما أظهر أن النموذج يحقق أداءً جيدًا مع كفاءة مقبولة.

تشير النتائج إلى أن استخدام نموذج Bi-LSTM يمثل خيارًا مناسبًا لمعالجة البيانات الزمنية في تطبيقات التعرف على لغة الإشارة.

سيتم في الفصل التالي عرض الاستنتاجات النهائية للبحث، بالإضافة إلى التحديات التي واجهت العمل والآفاق المستقبلية للتطوير.

الفصل السادس: الخاتمة

6-1- الاستنتاجات ..... **Error! Bookmark not defined.**

6-2- التحديات والآفاق المستقبلية ..... **Error! Bookmark not defined.**

6-1الاستنتاجات

تناول هذا البحث تطوير نظام ذكي للتعرف على لغة الإشارة اعتمادًا على تقنيات التعلم العميق، وذلك بهدف تحسين التواصل مع الأشخاص ذوي الإعاقة السمعية والمساهمة في دمجهم في المجتمع.

تم بناء نموذج يعتمد على Bi-LSTM لمعالجة البيانات الزمنية المستخرجة من النقاط الرئيسية(Keypoints) ، حيث أظهر النموذج قدرة جيدة على تحليل تسلسل الحركات والتعرف على الإشارات المختلفة. كما تم تصميم Pipeline متكامل يشمل مراحل معالجة البيانات، التدريب، والتقييم.

أظهرت النتائج أن النموذج المقترح قادر على تحقيق توازن مقبول بين الدقة والكفاءة، حيث استفاد من تحليل البيانات في اتجاهين، مما ساهم في تحسين جودة التنبؤات مقارنة بالنماذج التقليدية.

بشكل عام، يمكن اعتبار هذا العمل خطوة عملية نحو تطوير أنظمة أكثر تطورًا في مجال التعرف على لغة الإشارة.

6-2التحديات والآفاق المستقبلية

التحديات

واجه هذا البحث عددًا من التحديات، من أبرزها:

- محدودية البيانات المتاحة للتدريب
- تعقيد الإشارات الحركية وتداخلها
- صعوبة التمييز بين بعض الإشارات المتشابهة
- الحاجة إلى موارد حاسوبية مناسبة لتدريب النموذج

الآفاق المستقبلية

يمكن تطوير هذا العمل في عدة اتجاهات مستقبلية، منها:

- استخدام نماذج أكثر تطورًا مثل Transformers
- توسيع قاعدة البيانات لتشمل عددًا أكبر من الإشارات
- دعم لغات إشارة مختلفة
- تطوير تطبيق عملي يعمل في الزمن الحقيقي (Real-Time)
- دمج تقنيات إضافية مثل التعرف على تعابير الوجه

تشير هذه الآفاق إلى أن المجال لا يزال مفتوحًا للتطوير، وأن تحسين الأنظمة الحالية يمكن أن يسهم بشكل كبير في تعزيز التواصل الإنساني باستخدام التكنولوجيا.