المجلة المصرية لعلوم المعلومات

دور شبكات التعلم العميق في ترجمة لغة الإشارة للمستفيدين ذوي الهمم في المكتبات العامة المصرية: دراسة استكشافية

مؤمن سيد النشرتي النشرتيي

مدرس المكتبات والمعلومات

كلية الآداب – جامعة المنصورة

Exploring the Role of AI Deep Learning Networks for Translate Sign Language for Users with Disabilities in Egyptian Public Libraries: A Pilot Study

By

Moamen Sayed El-Nasharty El-Nasharty
Lecturer in Library and Information Science
Faculty of Arts – Mansoura University

تاريخ الإرسال **2025/4/3**

تاريخ القبول **2025/4/29**

المستخلص:

تأتي لغة الإشارة كاحد أبرز طرق التواصل الأكثر تعبيرًا في مجتمعات ذوي الهمم خاصة من ذوي إعاقة الصمم وضعاف السمع، والتي تعتمد بشكل رئيسي على التواصل ونقل المعلومات من خلال إشارات اليد، وعلى رغم ذلك، فإن إحدى أكثر المشكلات التي تواجه الافراد والمجالات العامة خارج مجتمع الصم هي فهم لغة الإشارة، حيث تبرو ضرورة وجود مترجم لفهم هذه اللغة، وهو الامر الذي قد لا يتوافر بشكل دائم، ومن هنا تبرز دور تقنيات الذكاء الاصطناعي، وخاصة تقنيات الرؤية الحاسوبية وما تعتمد عليه من شبكات التعلم العميق كأحد أبرز الحلول العلمية

والعملية لمشكلة توافر الفهم للغة الاشارة أو لعدم وجود مترجم لها، وتأتي هذه الدراسة لتقديم نموذجًا لترجمة لغة الإشارة في الوقت الفعلي معتمدة على شبكات التعلم العميق TensorFlow و OpenCV، ومن ثم يعد الهدف الرئيسي من هذه الدراسة هو تطوير نموذج لديه القدرة على ترجمة لغة الإشارة، ويمكن استخدامه دون اتصال بالإنترنت، ويكفل هذا النموذج القدرة على تحديد وترجمة الرموز الشائعة التي يتم التعبير عنها بإشارات اليد، والرموز المشتركة بين لغات الإشارة العالمية كالأمريكية وكالعربية مثل جيد (إبهام لأعلى)، سيئ (إبهام لأسفل)، شكرًا، وعيش طويلًا، وقد اعتمدت هذه الدراسة على منهج البحث التجريبي والتطبيقي، ويعد أبرز نتائج هذه الدراسة أن النموذج نجح في اكتشاف إشارات اليد والتعرف عليها بدقة 100٪.

الكلمات الدالة:

لغة الإشارة – الذكاء الاصطناعي – المكتبات العامة – التعلم العميق– ذوي الهمم.

Abstract

Sign Language is the most expressive way for communication between deaf and hearing-impaired people, where information is majorly conveyed through the hand gestures. Despite this, one of the most issues that facing people and public domains that are outside the deaf community is understanding sign language, they are need an interpreter to understand this language. Technologies including deep learning and computer vision aid this issue by providing several technical apps and other platforms. In this paper, the study proposes a real-time sign language detection model using TensorFlow and OpenCV. The main aim of this study is developing model that recognize sign language, and that can be used offline. Through this model, the study can detect a common words and sentences that are expressed by hand gestures, like among American and Arabic sign languages like good (thumbs up), bad (thumbs down), thanks, and live long. This study is belonging to the experimental and applicable research. The results of this study were that the model succeed to detect the hand gestures and recognize it with accuracy 100%.

Keywords: **Sign language - Artificial intelligence - Public libraries - Deep learning - Handicap people.**

تمهيد

يعد توفير الوصـــول إلى المعلومات هو جوهر مهمة المكتبات ومراكز المعلومات في جميع أنحاء العالم، وهو ما تؤكد عليه العديد من جمعيات المكتبات في جميع أنحاء العالم كمنظمة الإفلا IFLA وجمعية المكتبات الامريكية ALA وجمعية المكتبات البريطانية BL بإن إتاحة المعلومات هو مبدء أســاســي لهذه المؤســســات للعديد من الاغراض كدعم التطوير والتعلم والإبداع والابتكار من واقع ما يتم تقديمه من خدمات مكتبية تتســم بالجودة، وضرورة أن تتوافر هذه الخدمات لكافة فئات المستفيدين من الأطفال والشباب والبالغين والمستخدمين ذوي الاحتياجات الخاصة.

أشار الاتحاد الدولي لجمعيات ومؤسسات المكتبات (IFLA) إلى وجود العديد من الأسباب التي حالت دون اهتمام معظم المكتبات بتقديم خدمات خاصــة لذوي الهمم من فئات ذوي الإعاقة من الصـم وضـعاف السـمع، فالأشخاص ذوي إعاقة الصمم منذ الولادة أو في سن مبكرة غالبًا ما يواجهون صعوبة في القراءة، ويميلون إلى عدم اســتخدام المكتبات، ونتيجةً لذلك، ومن ثم ينأى كل من مجتمع إعاقة الصــمم والمكتبات عن بعضــهم البعض **[Day, 2000]**.

إلا أنه في الســنوات الأخيرة، أصــبح الأشــخاص ذوو الاحتياجات الخاصــة من أهم الفئات التي حظيت بالاهتمام كمســتخدمين للمكتبات، ففي عام 2017 أجرت منظمة الإفلا IFLA دراســة مســحية لعينة من المكتبات العامة على صعيد العالم بلغت نحو 470 مكتبة من 92 دولة، وقد بلغ عدد المكتبات العامة في هذا المسح نحو 167 وذلك بهدف الحصول على لمحة عامة عن السياسات الرسمية والمساعدة العملية في دعم الوصول إلى المعلومات – لا سيما من خلال التكنولوجيا – للأشخاص ذوي الاحتياجات الخاصة، وقد أفضى

هذا التقرير الرسمي عن هذة الدراسة المسحية ان 30 أن حوالي 30٪ من هذه المكتبات العامة لديها سياسات رسمية في توفير الوصول والإتاحة لمصادر المعلومات للأشخاص ذوي الاحتياجات الخاصة، وأن نحو 40٪ من المكتبات العامة (82 مكتبة عامة) لديها خدمات تم تكييفها لمساعدة مجتمعات الصم من المستفيدين **[Bolt & Wyber, 2017].**

تعد المشكلة الرئيسية المتعلقة بتقديم الخدمات المعلوماتية لمجتمع الصم هي أن التواصل غالبًا ما يتطلب جهدًا إضافيًا ومعرفة وصبرًا ومساعدات تكنولوجية إذا توافرت، فضلا عن الاختلافات الفردية **[Day, 2004].**

ووفقا لدراسة **[مصطفى، 2017]** فإن العديد من المكتبات في مصر تفتقر إلى الأساليب والتكنولوجيا والعاملين المحترفين للتواصل مع الأشخاص ذوي الصمم.

وفقًا لإحصاءات الاتحاد العالمي للصم، يوجد نحو 72 مليون شخص أصم حول العالم، ويعيش نحو 80% من هذه النسبة في الدول النامية، ويوجد ما بين 138 و300 نوع مختلف من لغات الإشارة حول العالم اليوم **[Rastgoo. 2021].**

وتأتي لغة الإشارة كأبرز وسائل التواصل بين أفراد مجتمع الصمم وضعاف السمع، حيث تعتمد عملية نقل المعلومات بشكل رئيسي عبر إشارات اليد **(Sense, 2023)**، ورغم ذلك، تُعدّ فهم لغة الإشارة من أبرز التحديات التي تواجه الأفراد والمجتمعات خارج مجتمع الصم، وتبرز الحاجة إلى مترجم لفهمها.

يجدر الإشارة إلى ان لغة الإشارة لها طبيعة خاصة من حيث مفرداتها ومعناها وقواعدها الخاصة، والتي تختلف عن اللغات المنطوقة أو المكتوبة **[Rastgoo. 2021].**

تعد أبرز الفروق بين اللغات المنطوقة ولغة الإشارة فيما يلي:

- اللغات المنطوقة هي لغة تُنتَج من خلال أصوات واضحة مُرتبطة بكلمات مُحددة وتركيبات نحوية لنقل رسائل ذات معنى.

- بينما تستخدم لغة الإشارة الإيماءات البصرية لليد والجسم لنقل رسائل ذات معنى [McInnes, 1993].

أما فيما يتعلق بالإشارات وإيماءات اليد في لغة الإشارة فيمكن تصنيفها إلى فئتين:

- الفئة الأولى الإيماءات الثابتة: تُستخدم الإيماءة الثابتة لتمثيل الحروف الأبجدية والأرقام، وتتكون من إيماءات اليد.

- أما الفئة الثانية فهي الإيماءات المتحركة: ويطلق عليها الإيماءة الديناميكية، فتُستخدم لمفاهيم محددة، وتشمل الكلمات والجمل، وما إلى ذلك.

كذلك يمكن أن ينظر إلى فئات الإيماءات إلى نوعين هما الإيماءات الطبيعية والإشارات الرسمية:

• الإشارة الطبيعية: هي تعبير يدوي (باليد) يتفق عليه المستخدم (تقليديًا)، ويُعرف بأنه محدود في فئة معينة، وهي بديل للكلمات التي يستخدمها الشخص الأصم (على عكس لغة الجسد).

• الإشارة الرسمية هي إشارة تُنشأ عمدًا ولها نفس بنية اللغة المنطوقة في المجتمع [Suthar, 2021].

يوضح **[Kodandaram, 2021]** إن التعرف على إيماءات اليد الثابتة أسهل من التعرف على إيماءات اليد الديناميكية، ولكن كلا الفئتين يعتبرا مهما للمجتمع البشري.

يوضح **[عبد الفتاح، 2005]** أن لغات الإشارة العربية (ARSLs) لا تزال في مراحلها التنموية، وان الوعي بالاهتمام وجود بمجتمعات ذوي الإعاقة برز بقوة في السنوات الأخيرة، لعد أسباب أبرزها أن مجتمعات الصم العربية تعد مجتمعات شبه مغلقة، وان التفاعل بين مجتمع الصم ومجتمع السمع يتسم بالمحدودية، ويتركز بشكل أساسي حول العائلات التي تضم أفرادًا من الصم، وأقارب الصم، وأحيانًا أصدقاء الصم والمهنيين، ووفقًا لوحدة الصم بالقاهرة، يوجد حوالي 1.2 مليون شخص من الصم وضعاف السمع في مصر، ممن تبلغ أعمارهم خمس سنوات فأكثر **[عبد الفتاح، 2005]**.

في هذا السياق يجدر الإشارة إلى أن لدى مصر لغة إشارة خاصة بها، وهي لغة الإشارة المصرية، والتي يستخدمها أفراد مجتمع الصم في مصر، وعلى رغم عدم وجود إحصاءات رسمية حول عدد الصم أو عدد من يستخدمون لغة الإشارة المصرية كلغة رئيسية، فقد أشارت دراسة أجرتها منظمة الصحة العالمية عام ٢٠٠٧ إلى أن معدل انتشار فقدان السمع في مصر بلغ نحو 16.02% بين جميع الفئات العمرية، فضلا عن عدم اعتراف الدولة رسميًا بلغة الإشارة المصرية **[LibGuides, 2023]**.

2-مشكلة الدراسة:

مشكلة الدراسة هي قضية واقعية تواجه المنظمات المجتمعية مثل المكتبات ومراكز المعلومات بل والمؤسسات الأخرى والتي تتعامل مع قطاعات مختلفة من المستفيدين بما فيهم المستفيدين من ذوي الاحتياجات الخاصة، وفي ذات الوقت قد لا تملك هذه المكتبات او المؤسسات في كثير من الاحيان الكوادر البشرية المؤهلة للتواصل مع هذه الفئة من المستفيدين، ومن ثم صعوبة إدراك أو معرفة احتياجاتهم أو الإجابة على استفساراتهم، نظرا لعدم فهم إيماءات لغة الإشارة. وهو ما اوضحه **[عبد الفتاح، 2005]** أن المجتمع العربي لا تزال النظرة إلى الإعاقة هي نظرة استيعاب وليس اندماج.

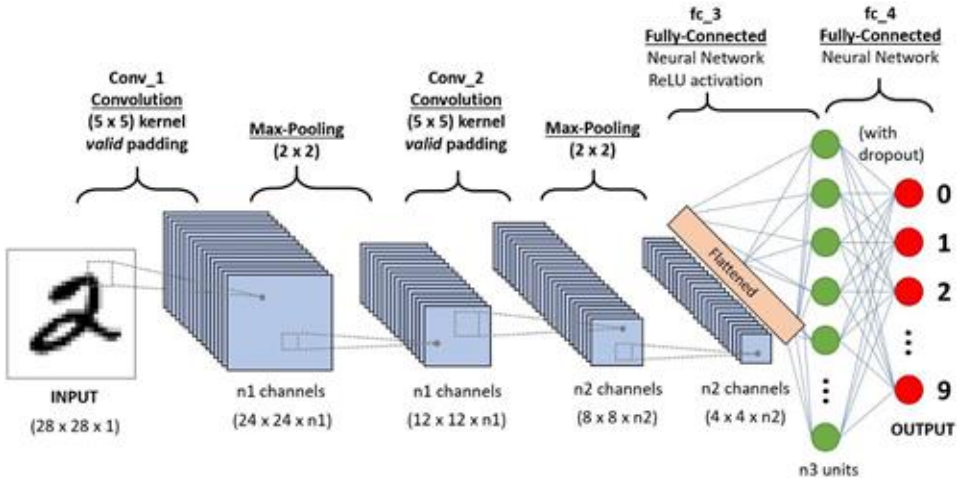
كذلك تبرز قضية أخرى وراء هذه الدراسة وهي أن لغات الإشارة ليست لغة موحدة عالمية، فكل بلد لديها لغة الإشارة الخاصة بها، ففي اللغة العربية، يواجه متعلمو لغة الإشارة العامية صعوبة كبيرة في استيعاب فكرة عدم الإشارة لكل كلمة في النطق، فضلا عن أن عملية تعليم لغة الإشارة للصم قد تعد مشكلة نظرًا لوجود عدد محدود من مترجمي لغة الإشارة في الوقت الراهن **[عبد الفتاح، 2005]**.

أما الواقع المصري، فيوجد في مصر 113 مدرسة للصم، إلا أن مستوى التعليم فيها متدنٍ، فضلا عن قيام بعض هذه المدارس باستخدام لغة الإشارة الإنجليزية، والبعض الآخر يعتمد على اللغة الشفهية، ويجدر الإشارة إلى أن المعلومات المتوفرة على الإنترنت حول هذه المدارس شحيحة؛ فقد أفادت وحدة الصم في القاهرة أنها تستخدم لغة الإشارة الإنجليزية واللغة العربية وفي الوقت نفسه، لا توجد دراسة استقصائية حول مترجمي لغة الإشارة المعتمدين لفئة الصم في مصر **[LibGuides, 2023]**.

هذه القضايا والمشكلات تستدعي ان تتوافر لها حلولا غير تقليدية، تتعاظم فيها الاعتماد على التقنيات وخاصة تقنيات الذكاء الاصطناعي، من خلال تطوير النماذج التي لديها القدرة على التعرف على لغة الإشارة، وإمكانية تحويل لغة الإشارة إلى نص أو كلام، مما يُسهّل التواصل بين مجتمع الصم والمجتمع التقليدي.

تأتي تقنيات التعلم العميق وخوارزميات التعلم الآلي، كأحد ابرز هذه الحلول، حيث اكتسبت تقنيات الرؤية الحاسوبية والتعلم العميق اليوم شعبية كبيرة في هذا المجال، وقد جاءت العديد من الأبحاث مشتملة على نماذج تكفل القدرة على ترجمة لغة الاشارة غلى نص، معتمدة على تقنية الشبكات العصبية التلافيفية Convolution

Neural Networks (CNN) والتي تعد أشهر خوارزميات الشبكات العصبية في ذلك، وهي خوارزمية تُستخدم على نطاق واسع لمهام الصور والفيديو.



الشكل رقم (1) يوضح بنية الشبكات العصبية التلافيفية [Varshini, 2020] (CNN).

وفي ذلك جاءت هذه الدراسة معتمدة على تقنية الرؤية الحاسوبية للتعلم العميق للتعرف على إيماءات اليد من خلال بناء هياكل شبكات عصبية عميقة، حيث يكون لدى النموذج القدرة على التعرف على صور إيماءات اليد آنيًا كما هو في الشكل رقم 1.

ويجدر اللإشارة أيضا أن تصميم نموذج لترجمة لغة الإشارة بصورة دقيقة وكاملة يعد تحديًا كبيرًا، رغم الجهود البحثية العديدة المبذولة خلال السنوات القليلة الماضية، وذلك لأسباب عديدة، أهمها هو اختلاف مظهر الإشارات بشكل كبير باختلاف مستخدميها ووجهات نظرهم.

أهداف الدراسة:

يعد الهدف الرئيسي من هذه الدراسة هو المساهمة هو القدرة على الاكتشاف الآني والتحقق من لغة الإشارة وترجمتها إلى نص، وذلك من خلال تطوير نموذج ذكي لديه القدرة على التعرف على لغة الإشارة، ويمكن استخدامه دون اتصال بالإنترنت. وقد طُوّر هذا النموذج اعتمادا على الرصد للحركات من خلال الرؤية الحاسوبية للحصول على البيانات من المتحدث بلغة اللإشارة، ومن أهم مميزات هذه الدراسة قدرة النموذج على تحديد الكلمات المُدرجة في مجموعة من البيانات المحلية (مجموعة بيانات مُخصصة من قِبل الباحث) بهدف اختبار النموذج للتعرف عليها وترجمتها لنص.

في ذات الوقت تركز هذه الدراسه على التعرف على إشارات اليد في لغة الإشارة الديناميكية باستخدام الشبكات العصبية العميقة (DNN) للتعرف على أربع إشارات مشتركة بين لغة الإشارة العربية ولغة الإشارة الإنجليزية وهي: (إشارة الإبهام للأعلى – إشارة شكرًا لك – إشارة عش طويلا – إشارة الإبهام للأسفل)، وذلك من خلال إنشاء مصنف شبكات عصبية التفافية CNN يمكنها تصنيف الإشارات الأربعة، ثم البدء لتدريب النموذج بواسطة مكتبة برمجيات TensorFlow.

تعد أبرز المزايا المتعددة لاعتماد على مكتبة برمجيات Tenser–Flow في بناء نموذج للتعرف على لغة الإشارة وترجمتها لنص في الوقت الفعلي في قدرتها على ترجمة إيماءات اليد إلى نص بشكل يتسم بالدقة، وهو الامر الذي سيسهل أن يتم تبني هذا النموذج في أنظمة المكتبات ومراكز المعلومات للاستخدام فضلا عن كافة المؤسسات الأخرى والمطارات ومكاتب البريد أو المستشفيات، وذلك لتمكين التواصل بين الأشخاص العاديين والأشخاص ذوي الاحتياجات الخاصة،، ليس هذا فحسب بل يمكن استخدام هذا النموذج مستقبلا مع الأجهزة المضمنة / اولتطبيقات المستقلة / وتطبيقات الويب دون استنفاد للموارد بصورة كبيرة.

المراجعة العلمية:

يوضح كل من **[Cheok, 2019] [Rastgoo, 2021]** أن لغة الإشارة تعتمد على استخدام أجزاء مختلفة من الجسم، مثل الأصابع، واليد، والذراع، والرأس، فضلا عن تعابير الوجه ، هناك خمسة معايير رئيسية في لغة الإشارة، وهي: شكل اليد، واتجاه راحة اليد، والحركة، والموقع، والإشارات التعبيرية/غير اليدوية.

تستفيد العديد من التطبيقات من مزايا التعرف على لغة الإشارة مثل أنظمة الترجمة وخدمات الترجمة والترجمة البشرية عن بعد عبر الفيديو والتفاعل بين الإنسان والحاسوب **[Tang, 2015] [Yang, 2015] [Supancˇicˇ, 2018] [Zheng, 2017] [Deng, 2017] [Wang, 2016].**

ورغم ذلك يوضح **[Rastgoo. 2021]** أن أحد التحديات الرئيسية في تطوير نظم تقليدية للتعرف على لغة الإشارة لترجمة الكلمات أو الجمل الفردية إلى نص أو صوت لتسهيل التواصل بين مجتمع الصم والمجتمع التقليدي هو ما تعتمد عليه هذه النظم من الاعتماد على مجموعات بيانات تتضمن مقاطع فيديو لإشارة واحدة فقط أو صورًا لشخصية واحدة فقط للتعرف على لغة الإشارة، ومن ثم تعجز هذه الانظمة عن أن يتم تطبيقها

في ترجمة لغة الإشارة لمحادثة حقيقية تتم في الوقت الآني بين شخص أصم وآخر يسمع للقيام بذلك، ومن ثم يجب أن تتسم الأنظمة المقترحة بالفاعلية والذكاء بما يكفي لتقسيم مقاطع الفيديو المدخلة، بما في ذلك بعض الأحرف أو الكلمات أو الجمل، إلى أحرف أو كلمات أو جمل منفصلة، فضلا عن أبرز أحد التحديات الرئيسية في هذا المجال هو الاختلاف الكبير بين لغات الإشارة بين مختلف البلدان والدول.

ومع ظهور شبكات التعلم العميق في السنوات الأخيرة، تم إجراء العديد من الجهود البحثية في مضمار التعرف والترجمة الآنية لغة الإشارة لنص في العديد من اللغات

[Asadi-Aghbolaghi, 2017] [Guo, 2017 [Ferreira, 2019] [Lim, 2019].

يعود تاريخ الجهود المبكرة في مجال التعرف على إشارات اليد والإيماءات إلى عام 1987، حيث اقترح **[Zimmerman, et al. 1987]** نظام يعتمد على قفاز متصل بواجهة تعامل لتقدير موضع اليد واتجاهها باستخدام مستشعرات التدفق المغناطيسي في القفاز، ومع ذلك، برزت العديد من التحديات المختلفة التي وجب معالجتها في التعرف على لغة الإشارة البصرية (منها على سبيل المثال، التباين بين الأشخاص مدخلاتهم، وظروف الإضاءة، ونقاط مختلفة).

يوضح **[Li, 2012]** أن نماذج التعلم العميق المستخدمة في مجال الرؤية الحاسوبية وخاصة الشبكات العصبية التلافيفية CNN قد برزت بدورها في معالجة لغات الإشارة وحملت في طياتها وعودا مامولة في دقة اكتشافها وترجمة إيماءات اليد.

أحد الأهداف الرئيسية لنماذج التعلم العميق هو تجنب الحاجة إلى إجراء عمليات تدريب مسبقة وبناء أو استخراج الخصائص والميزات لمجموعات البيانات، مما يتيح للنماذج الحاسوبية أن تتعلم وفقا لنمط متعدد للمعالجة والتمثيل بمستويات تجريد متعددة، ومن ثم تعمل شبكات التعلم العميق على محاكات آلية عمل الدماغ البشري في قدرته على التعلم الذاتي من خلال الاستنباط والاستدراك والتحليل والتنبؤ والتقاط الهياكل المعقدة للبيانات واسعة النطاق ضمنيًا **[Rastgoo. 2021].**

تنطوي شبكات التعلم العميق على مجموعة غنية من الأساليب، كالشبكات العصبية، والنماذج الاحتمالية الهرمية، خوارزميات التعلم الخاضعة للإشراف وغير الخاضعة للإشراف، ويعد أحد أهم العوامل الرئيسية التي ساهمت في دفع التوجه والاعتماد على التعلم العميق هو ظهور مجموعات البيانات Dataset ذات التصنيف

والجودة المرتفعة والمتاحة للجمهور على نطاق واسع، فضلا عن تعاظم قدرة الحوسبة المتوازية لوحدات المعالجة السريعة في الحاسب **[Rastgoo. 2021]**.

قدمت تقنيات الشبكات العصبية التلافيفية Convolutional Neural Networks (CNN) لأول مرة من جانب **[LeCun, 1988]** والتي قام باقتراحها كبنية خاصة للشبكات العصبية الاصطناعية (ANN)، تعتمد تقنية شبكات CNN على الخصائص البصرية التي توجد في المدخلات سواء كانت صور أو فيديو أو نص، ومن أكثر استخدامات هذه التقنية شيوعًا هو الاعتماد عليها في التعرف على محتوى الصور وتصنيفها، فعلى سبيل المثال، يقوم موقع Facebook باستخدام تقنية CNN في إنشاء خوارزميات تحديد التيجان في الصور بشكل تلقائي، ويقوم موقع Amazon باستخدام هذه التقنية في اكتشاف المنتجات بشكل بصري وتوليد التوصيات للمستخدمين كما يقوم موقع Google بالاعتماد ايضا على هذه التقنية للبحث اعتمادا على الصور كاستفسار بدلا من النص، حيث تمكن تقنية CNN الحاسب من أن يرى ويكتشف ويتعرف على الصور في صورة مجموعة من البكسلات، على سبيل المثال، إذا كان حجم الصورة 300 × 300 بكسيل، في هذه الحالة، سيكون حجم المصفوفة 300 × 300 × 3 حيث يمثل 300 العرض، و300 التالي الارتفاع، و3 قيم قناة RGB، يتم تعيين قيمة من 0 إلى 255 للحاسب لكل من هذه الأرقام، بحيث يصف تدرج هذه القيمة في هذا المدى كثافة البكسل في كل نقطة **[Wu, 2017]**.

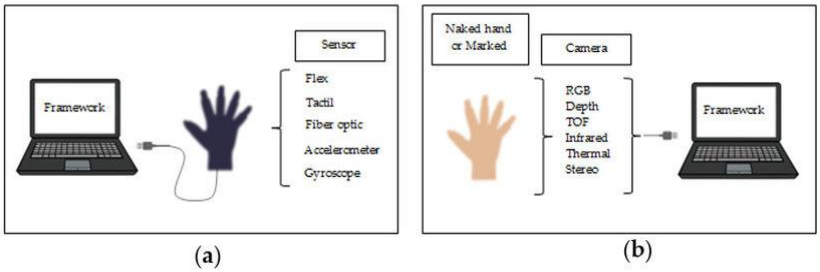
تعد أطار عمل TensorFlow واحدة من اقوى المكتبات البرمجية التي تم تطويرها ضمن شبكات CNN للتعرف على لغة الإشارة **[Rastgoo. 2021] [TensorFlow, 2023]**.

أطر عمل TensorFlow هي مكتبة برمجية مجانية ومفتوحة المصدر لتدفق البيانات والبرمجة التفاضلية عبر مجموعة من الدوال، وهي مكتبة تعتمد على الرياضيات رمزية، وتُستخدم أيضًا في تطبيقات التعلم الآلي مثل الشبكات العصبية، كما تستخدم هذه المكتبة البرمجية في مهام البحث والتطوير داخل شركة Google، اما عن لغات البرمجة التي يمكن ان يتم توظيف هذه المكتبة من خلالها هي لغة C++ ولغة Python ولغة CUDA **[TensorFlow, 2023]**.

كذلك يعد من المكتبات الأكثر استخدامًا لمهام اكتشاف وترجمة لغة الإشارة مكتبة OpenCV وهي مكتبة برمجية مفتوحة المصدر تشمل عدد من الدوال التي تستهدف بشكل رئيسي الرؤية الحاسوبية في الوقت الفعلي، وقد طُوّرت هذه المكتبة في الأصل بواسطة شركة Intel، ثم دعمتها معامل Willow Garage، من مزايا هذه المكتبة أنها عبر منصات متعددة، وهي مجانية الاستخدام بموجب ترخيص مفتوح المصدر **[OpenCV, 2023].**

كذلك تأتي مكتبة Keras كأحد أبرز المكتبات البرمجية التي تُستخدم لإنشاء أنظمة التعرف على لغات الإشارة في الوقت الفعلي، تعد Keras مكتبة شبكات عصبية متطورة وهي مفتوحة المصدر كتبت بلغة بايثون، كما يمكن تشغيلها باستخدام مكتبات أخرى كمكتبة TensorFlow، أو أداة Microsoft Cognitive Toolkit، أو لغة R، أو Theano، أو PlaidML، وقد صُممت هذه المكتبة لتمكين التجارب السريعة مع الشبكات العصبية العميقة، وتتميز بسهولة الاستخدام، وتعدد وحداتها، وقابليتها للتوسع **[Keras, 2023].**

يوضح **[Rastgoo. 2021]** أن النمطين الرئيسيتين اللذان يمكن الاعتماد عليهما في الوقت الراهن في بناء نماذج الذكاء الاصطناعي للتعرف على لغة الإشارة هما نمط الرؤية الحاسوبية من خلال رصد حركات اليد بكاميرا تقوم بالتقاط هذه الحركات وترجمتها لنص، أو نمط القفازات من خلال ارتداء المتحدث لهذا القفاز ورصد حركة اليد من خلال مستشعرات لترجمتها لنص، بينما تستخدم النماذج القائمة على القفازات مستشعرات ميكانيكية أو بصرية مُثبتة على القفاز لاستخدام الإشارات الكهربائية لتحديد وضعية اليد وترجمة الحركة، جدير بالذكر ان النماذج القائمة على الرؤية توفر أنظمةً أكثر واقعيةً وطبيعيةً، استنادًا إلى المعلومات التي يمكن للإنسان استشعارها من البيئة المحيطة، كما هو موضح في الشكل رقم (الشكل 2). وسوف تعتمد هذه الدراسة على نماذج الرؤية المعتمدة على الكاميرات.



الشكل ٢: يوضح نمط النماذج القائمة على الرؤية ونمط النماذج القائمة على القفازات [Oudah, 2020]

في الفترة الأخيرة برزت بعض أنماط الإدخال المرئي للبيانات في مجال التعرف على لغة الإشارة، حيث جاءت بيانات إدخال RGB والبيانات العميقة Depth Inputs لكيونا أكثر نمطين شائعين من أنماط الإدخال للبيانات داخل نماذج التعرف على لغة الإشارة، فصور RGB أو وفق ما يعرف بمقاطع الفيديو اتسمت بكونخا نمط إدخال لمحتويات تتسم بكونها مرتفعة الدقة، بينما جاءت تحتوي البيانات العميقة على معلومات دقيقة تتعلق بالمسافة بين مستوى الصورة والجسم المقابل لها، جدير بالذكر أن بعض النماذج اتسمت في القوت الراهن بمزايا القدرة على نمطي إدخال في آنٍ واحد **[Rastgoo. 2021]**.

أما عن أجهزة أدخال البيانات لنماذج التعرف على لغة الاشارة، فتعد جهاز الكاميرا هي الأكثر شيوعًا، حيث تدعم الكاميرات المختلفة تنسيقات وجودات مختلفة لبيانات الإدخال للبيانات **[Rastgoo. 2021]**.

أما عن انماط البيانات التي تعتمد عليها نماذج التعرف على لغة الأشارة، فهناك نوعان من بيانات الإدخال، البيانات الثابتة والبيانات الديناميكية وذلك لإدراك والتعرف على اللإيماءات ولاستخراج الخصائص اللازمة لترجمة الحركات الملتقطة، وقد تم تقديم العديد من النماذج القائمة على البيانات العميقة لاستخدام المدخلات الثابتة أو المتسلسلة في السنوات الأخيرة. أما فيما يتعلق بالبيانات الديناميكية فقد طورت العديد من النماذج التي تُفيد في تحسين دقة التعرف على لغة الإشارة **[Rastgoo. 2021]**.

في هذا السياق كان من إرهاصات نماذج التعرف على لغة الإشارة ما اقترحه **[Nandy et al, 2010]** كنهجًا لاكتشاف إيماءات لغة الإشارة الهندية والتعرف عليها من الصور ذات التدرج الرمادي، في هذا النهج، يتم تحويل الفيديو المسجل الذي يحتوي على إيماءات الإشارة إلى إطارات ذات تدرج رمادي، حيث يتم استخراج السمات من خلال رسم بياني Histogram وأخيرًا، يُستخدم يتم تطبيق خوارزمية التجميع Clustering لتصنيف الإشارات وفقًا لسماتها إلى إحدى الفئات المحددة مسبقًا، وقد استطاعت هذه الدراسة بتحقيق معدل تعرف على الإشارات بلغت دقته نسبة ١٠٠٪ ، وخلصوا إلى أن طريقة الرسم البياني Histogram ذات الـ ٣٦ خانة كانت أكثر دقة من طريقة الرسم البياني Histogram ذات الـ ١٨ خانة.

كما اقترح **[Mekala et al, 2011]** نموذج ذو بنية شبكة عصبية للتعرف على لغة الإشارة وتتبعها وتوليد نص فوري من بث الفيديو في الوقت الفعلي، حيث تتكون بنية النموذج من عدة مراحل كمرحلة إنشاء

الإطارات ثم مرحلة المعالجة المسبقة للصور ثم مرحلة استخراج الحركات والخصائص بناءً على حركة اليد وموضعها، يتم تمثيل سمات وحركات اليد باستخدام ما يعرف بنقطة اهتمام (POI) لليد، واعتمادا على هذه المنهجية، يتم استخراج 55 خاصية وحركة مختلفة، والتي سيتم استخدامها كمدخلات لبنية النموذج المقترحة، جدير بالذكر أن بنية النموذج تعتمد على شبكات CNN لتوفير القدرة على التنبؤ بالحركات، كذلك اشتملت هذه الدراسة على القيام بعملية تدريب واختبار للنموذج على الأبجدية الإنجليزية، من الألف إلى الياء، وقد قاموا بتحقيق معدل تعرف على الحركات بلغت دقتها نحو 100٪.

كما طور كل من **[Shivashankara & Srinath, 2018]** نموذجا ذكيا للتعرف على لغة الإشارة الأمريكية (ASL) وتحديد الإشارات/الإيماءات وترجمتها لنصوص، كما ما يعد قيمة مضافة لهذا النموذج ما اشتمل عليه هذا النموذج من خوارزمية عرفت باسم YCbCr، لتحليل مما حسّن أداء تجميع الحركات دون الالتفات لتبيان درجات ألوان البشرة لليد، كذلك استُخدم هذا لنموذج عدد من الصور في مرحلة المعالجة المسبقة للتعرف على الإيماءة، حيث اعتمدت عمليد تحديد الحركة على تحديد مركز ثقل اليد في الصورة المعالجة مسبقًا؛ ثم تم تحديد إزاحة المشتتات لتحديد الإيماءة بدقة، وقد بلغت الدقة الإجمالية لهذا النموذج في التعرف على حركة اليد نسبة ٩٣.٠٥٪.

منهجية الدراسة:

تنتمي هذه الدراسة إلى منهاج البحث التجريبية والتطبيقية، حيث تعد بمثابة مشروع تنفيذي لتصميم نموذج ذكي للتعرف على الجمل الشائعة استخامها والمشتركة في لغة الإشارة في مجتمع الصم المصري باللغتين العربية والأمريكية، ولإغراض إنشاء هذا النموذج، من الضروري المرور بالمراحل التالية:

1. مرحلة جمع البيانات.
2. مرحلة المعالجة المسبقة للبيانات.
3. مرحلة بناء النموذج.
4. مرحلة تدريب النموذج.
5. مرحلة اختبار النموذج.
6. مرحلة تقييم النموذج.

تـــم بنـــاء النمـــوذج المقتـــرح اعتمـــادا علـــى الشـــبكة العصـــبية الاصـــطناعية باســـتخدام تقنيـــة شـــبكات CNN ومكتبـــة TensorFlow ومكتبـــة OpenCV، وذلـــك بـــدءًا مـــن مراحـــل اســـتخدام كـــاميرا الويـــب لالتقـــاط صـــور لإيمـــاءات اليـــد، ثـــم القيـــام بتســـمية الصـــور الملتقطـــة، ثـــم تقسيمها إلى قسمين، أحدهما لتدريب النموذج، والآخر لاختبار النموذج.

البيانات والمدخلات:

تتوفر العديد من مجموعات البيانات التدريبية الضخمة موقع Kaggle والتي يمكن من خلالها تدريب النماذج لاكتشاف وترجمة لغة الإشارة على موقع Kaggle [Kaggle، 2025]، وتعرف هذه المجموعات من البيانات بشكل عام اسم MNIST وهي مجموعة بيانات عامة مجانية الاستخدام، تحتوي على بيانات لحوالي 1000 صورة، بحيث تعكس كل صورة حرف من الحروف لغة الإشارة الأمريكية الأربعة والعشرين الشائع استخدامها، لكن هذه المجموعات من البيانات مخصصة للغات الإشارة الثابتة التي تعتمد على استخدام الإيماءات لتمثيل الحروف والأرقام، وليست للغة الإشارة الديناميكية التي تُمثل الكلمات أو الجمل الشائعة؛ ولذلك تهدف هذه الدراسة إلى جمع الإشارات الديناميكية التي تعكس جمل وليس حروف عن طريق تسجيل إيماءات اليد بالكاميرا، مما يعني أن هذه الدراسة ستُعالج لغة الإشارة المُلتقطة آنيًا للتعرف على كلمات لغة الإشارة.

وفي هذه الدراسة قام الباحث بإنشاء مجموعة البيانات للتدريب والاختبار للنموذج المقترح، حيث جُمعت هذه البيانات من الباحث نفسه بالاعتماد على مكتبة OpenCV وكاميرا الويب، تُعبّر عن أربع جمل (إيماءات اليد) رئيسية مشتركة بين العربية والإنجليزية في لغة الإشارة وهي: الجملة الاولى: جيد (إبهام للأعلى) (الشكل 6)، الجملة الثانية: سيء (إبهام للأسفل) (الشكل 4)، الجملة الثالثة: شكرًا (الشكل 5)، والجملة الرابعة: يعيش طويلًا (الشكل 3).

شكل رقم (3) يوضح تسجيل رمز "عش طويلا" بلغة الإشارة للنموذج.

شكل رقم (4) يوضح تسجيل علامة سيئ بلغة الإشارة

شكل رقم (5) يوضح تسجيل علامة شكرا بلغة الإشارة.

شكل رقم (6) يوضح تسجيل علامة جيد بلغة الإشارة

قامت الدراسة بتجميع ٢٠ صورة لكل جملة، وجميعها تم التقاطها بجودة بدقة ١٩٢٠ × ١٠٨٠ بكسل، بمتوسط مدة ثانيتين، سُجِّلت الصور باستخدام كاميرا ويب واعتمادا على مكتبة

برمجيات OpenCV على جهاز كمبيوتر محمول وقد أُنشئت هذه المجموعة من البيانات في ظروف طبيعية – بدون أي سطوع إضافي، أو تحديد لاتجاهات اليد، أو مستشعرات للحركة، أو تعديلات في الخلفية، أو قفازات، أو التقاط للصور من زوايا مختلفة، مع حركة سريعة وبطيئة لليد.

كذلك يجدر الإشارة إلى أن الدراسة لم تعتمد على أية أدوات مساعدة من مستشعرات أو قفازات ذكية للكشف عن حركات اليد، كما هو موضح في الشكل 3، وقد بلغت المسافة بين الباحث وكاميرا التجميع للقطات حوالي 30 سم، مع إمكانية تجاوز هذه المسافة بمقدار 32 إلى 35 سم، وفيما يلي بعض التقنيات التي تم توظيفها لبناء النموذج:

- واجهة التعامل: تم الاعتماد على برنامج Jupyter لكتابة الاكواد البرمجية لبناء النموذج، وقد تم الاعتماد على لغة بايثون لما تتسم به من سهولة ودقة في بناء النماذج الذكية وقدرتها على استيعاب مجموعات البيانات.
- **بيئة نظام التشغيل:** حيث تم الاعتماد على نظام تشغيل Windows 10، 16 – RAM جيجابايت، بطاقة الرسومات – 6 جيجابايت، 1060 ROM تيرابايت.
- **البرمجيات:** لغة Python (3.7.4)، منصة Anaconda (2019–0.7)، برنامج IDE (Jupyter)، مكتبة برمجيات NumPy (الإصدار 1.16.5)، مكتبة برمجيات OpenCV (الإصدار 3.4.2)، مكتبة برمجيات TensorFlow (الإصدار 2.11.0)، مستودع Github، منصة Virtual Studio (2022)، تطبيق CUDA (10.1) وCuDNN (7.6)، وحدة معالجة الرسومات NIVIA لبناء نموذج تدريب بصورة سريعة.
- **مكتبة برمجيات TensorFlow:** وهي حزمة أكواد برمجية لتطبيقات الذكاء الاصطناعي، وهي مكتبة مفتوحة المصدر يتم الاعتماد عليها بشكل رئيسي في بناء نماذج ذكية، وتكفل هذه المكتبة البرمجية للمطورين من بناء شبكات عصبية واسعة النطاق متعددة الطبقات،

وعاد ما تستخدم مكتبة TensorFlow بشكل رئيسي في عمليات التصنيف، والإدراك، والفهم، والاكتشاف، والتنبؤ بالبيانات والكيانات.

- **واجهة برمجة تطبيقات** Tensor-Flow: وهي واجهة مفتوحة المصدر لتحديد الكيانات في الصورة والتعرف عليها.
- **مكتبة برمجيات** OpenCV: وهي مكتبة بايثون مفتوحة المصدر تم تصميمها لمعالجة مشاكل الرؤية الحاسوبية، وتُركز هذه المكتبة البرمجية بشكل أساسي على المعالجة في الوقت الفعلي، كما تتسم بالكفاءة في المعالجة الصورية لإدارة كميات هائلة من البيانات، يعد أبرز استخداماتها في معالجة الصور والأفلام للتعرف على العناصر والأشخاص، وحتى إدراك حروف الكتابة اليدوية.
- **برنامج** LabelImg: وهو أداة توضيحية رسومية للصور، عاد ما تستخدم في إنشاء علامات تبويب وتيجات Tags or Labels للكيانات داخل الصورة.

المعالجة المسبقة للبيانات والمدخلات Pre-processing Data:

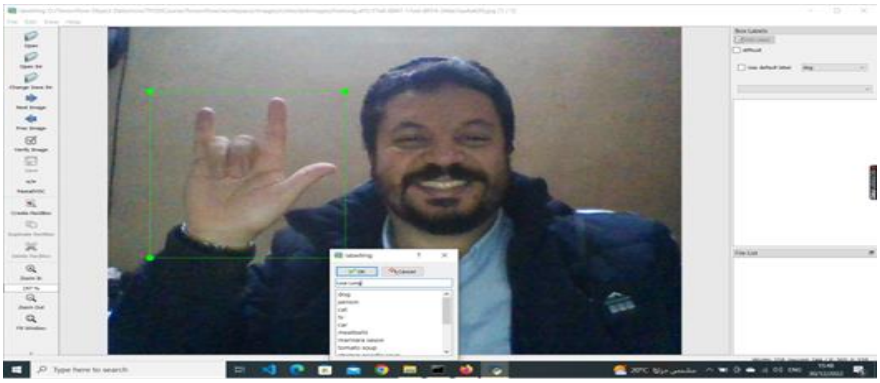
يعد الهدف الرئيسي من مرحلة المعالجة المسبقة هو تحسين بيانات الصورة عن طريق تقليل الانحراف غير المرغوب فيه أو تعزيز ميزات الصورة لمزيد من الدقة في المعالجة، ولضبط النتائج، كما تعمل مرحلة المعالجة المسبقة أيضًا على محاولة الالتقاط للعوامل والانماط الهامة التي تميز الصور وما تشمله من حركات اليد دون ضوضاء أو بيانات غير مرغوب فيها، والتي تتضمن عمليات المعالجة من استخراج وتغيير للحجم وتحديد التدرج الرمادي للصور، تنطوي مرحلة المعالجة المسبقة على عدد من العمليات الفرعية والتي تتمثل فيما يلي:

- **التقطيع والقص للصور:**

حيث تنطوي هذه المهمة على إزالة الأجزاء غير المرغوب فيها من الصورة لتحسين إنشاء الإطارات واللقطات أو إبراز الدلالات التعبيرية لحركة اليد أو تغيير نسبة العرض والارتفاع للصورة.

- **عنونة الصور:**

اعتمدت الدراسة في إنشاء عناوين وتيجان للقطات داخل الصور على برنامج LabelImg وكذلك تم الاعتماد على نفس البرنامج في إنشاء لوصف لكل صورة، حيث تقوم الدراسة برسم مستطيل على الإيماءة لغة الإشارة، مما سيسمح بفصل كل إيماءة إلى فئات مختلفة، وبمجرد الانتهاء من إضافة التيجان والعنونة للصور، تقوم الدراسة بحفظ الصور في مجلدات التدريب والاختبار الخاصة بالنموذج (الشكل 7).



كل رقم (7) يوضح عملية التحديد لحركة اليد والعنونة لها.

علاوة على ذلك قامت الدراسة بإنشاء ملفات XML (لغة الترميز القابلة للتوسعة) لكل صورة، حيث يتكون الملف من إحداثيات تحديد الإشارة داخل الصورة، وتحديد الفئة التي تمتني إليه حركة اليد، واسم الملف، وكافة البيانات التي تميز كل لقطة عن الأخرى.

بناء النموذج للدراسة Model:

يجدر الإشارة قبيل بناء النماذج الذكية اكتشاف وترجمة لغة الإشارة إلى وجود فئتان من النماذج وفقا لعملية التدريب والترجمة: الفئة الأولى نماذج لغة الإشارة المتقطعة والفئة الثانية هي نماذج لغة الإشارة المتصلة، ففي الفئة الأولى المستقلة المعزولة يتم تدريب النموذج على تمييز إيماءة واحدة بشكل مستقل، حيث يتم التعريف في كل صورة بتمثيل حرف أبجدي أو رقم أو إيماءة خاصة، بينما تشير الفئة الثانية من النماذج، إلى قدرة النماذج على تمييز الجمل

كاملة وترجمتها بشكل كامل بدلاً من التعرف على إيماءة واحدة وترجمتها بشكل فردي ومستقل، وينتمي نموذج الدراسة إلى الفئة الثانية.

في هذا النموذج المقترح للدراسة، تم استخدام واجهة برمجة تطبيقات التعرف على الكيانات TensorFlow لإجراء علمية التعرف والاكتشاف لجمل لغة الإشارة في الوقت الفعلي.

يعتمد النموذج المقترح على سلسلة معالجة تبدأ بعملية الإدخال والالتقاط لإيماءات اليد للمتحدث من خلال كاميرا الويب، ثم القيام باستخراج إطارات مختلفة من الصور الملتقطة في الوقت الآني، مما يكفل توليد القدرة على تحديد الإشارة او الجملة او الكلمة التي تشير لها كل إيماءة.

تتمثل الخطوات الرئيسية لإنشاء نموذج دراسة فيما يلي:

1. جمع الصور باستخدام كاميرا الويب
2. استخدام مكتبة OpenCV لبناء شبكة CNN.
3. تسمية الصور باستخدام برنامج LabelImg.
4. إعداد مسار TensorFlow للبدء في بناء النموذج الذكي لاكتشاف وترجمة الكيانات.
5. استخدام نمط التعلم الانتقالي لتدريب نموذج التعلم العميق.
6. اكتشاف لغة الإشارة آنيًا من خلال النموذج اعتمادا على الكاميرا أيضا.

اعتمدت علمية البناء للنموذج على تقسيم مجموعة البيانات إلى ملفين، الملف الأول لتدريب النموذج المقترح على 70٪ من الصور الملتقطة، بينما اشتمل الملف الثاني للنموذج على الصور التي سيتم اختبار النموذج في ضوئها وقد بلغت 30٪ من الصور الملتقطة.

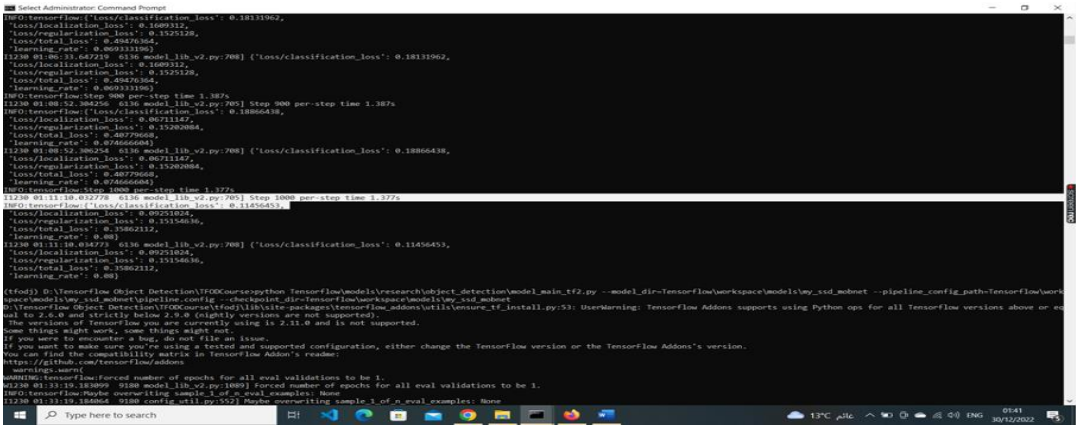
ثم تأتي المرحلة التالية في عملية البناء لنموذج الدراسة وهي إنشاء تسجيلات TF (TensorFlow)، والتي تعد خطوة هامة نظرا لكون تقبل واجهة تطبيقات TensorFlow لا تقبل البيانات إلا بتنسيقات ملفات TF-Record، هذه الملفات هي عبارة عن سلسلة من سلاسل البيانات الثنائية، قد تم تحويل الصور لها لإجراء عملية البناء والتدريب والاختبار

للنموذج، تكفل هذه التسجيلات جودة وسرعة مراحل علمية البناء للنموذج ودقة عملية التدريب للنموذج (الشكل 8).

شكل رقم (8) يوضح عمليات بناء النموذج للدراسة.

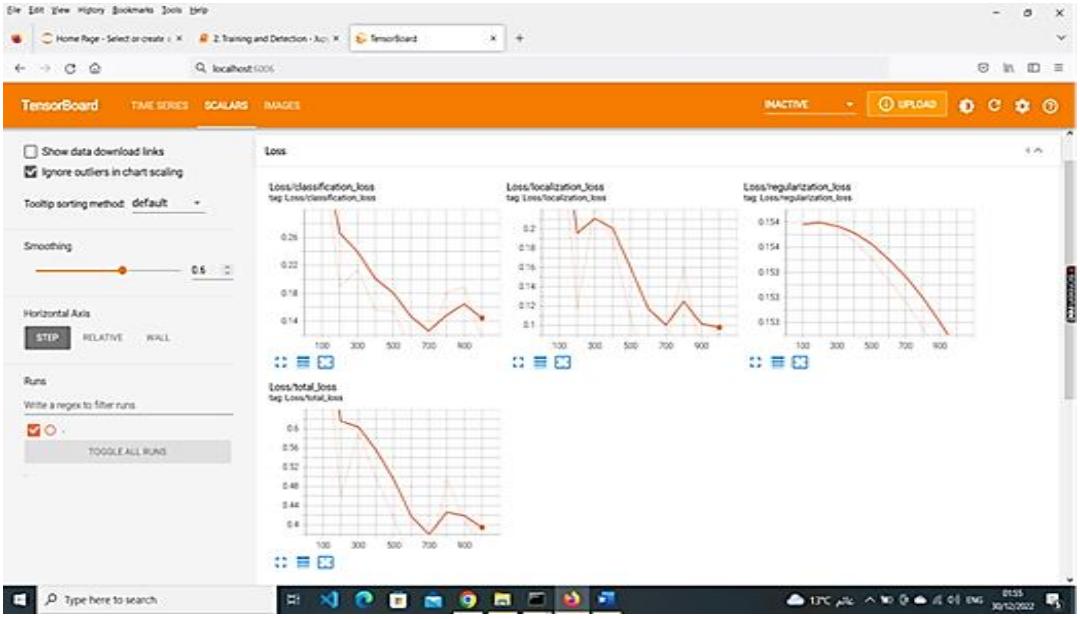
عقب إنشاء مجموعة تسجيلات TFR، قامت الدراسة بتوظيف أداة SSD-mobilenet-v2-pnlite-320x320-coco17 كاداة رئيسية لبناء النموذج المقترحوتهيئتها وفق متطلبات الدراسة وتعديلها وفقا للخصائص المطلوبة، مثل وضبط المعاملات ودوال التحقق من دقة النتائج بالأهداف وغيرها، حتى تتمكن الدراسة من تدريب النموذج على اكتشاف الكيانات.

أخيرًا، قامت الدراسة بتدريب النموذج باستخدام التعلم الانتقالي Transfer Learning وقد اشتمل تدريب النموذج على 2000 خطوة تدريبية Threshold لتحقيق نتائج أفضل، حُفظ النموذج عند عدة نقاط تفتيش، وقد بلغ إجمالي نقاط الفحص المُنشأة 3 نقاط، وعقب إجراء عملية التدريب لبضع ساعات، تم إخفاض درجة الفارق بين النتائج والأهداف Loss Function إلى 0.242 الآن، ليتم تحميل النموذج لإجراء عملية الاختبار والتحقق على بيانات جديدة والكشف عن دقة مخرجاته كما هو موضح في الشكل رقم 9.



شكل رقم 9 يوضح إجراء عملية التدريب للنموذج

بعد إكمال عملية تدريب للنموذج المقترح، قامت الدراسة بتقييم فاعلية ودقة النموذج من خلال مقياس التحقق والاستدعاء وقياس دقة مخرجاته باستخدام مصفوفة Confusion Matrix، وقد بلغ معدل التحقق من صحة النموذج في دالة الاسترجاع نحو 50%، وبلغ معدل دقة النموذج حوالي 50%، وقد استُخدمت أداة Tensor Board لإنشاء المصفوفة وتحديد نسبة الإدراك (الشكل 10).



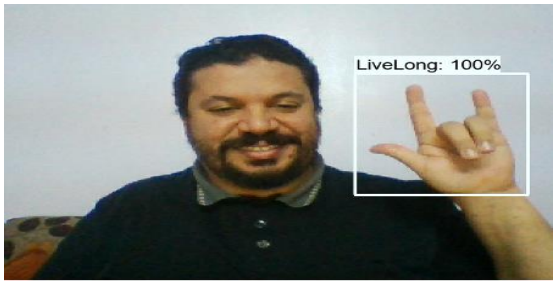
شكل رقم 10 يوضح عملية تقييم النموذج من خلال أداة Tensor Board

النتائج والتوصيات:

كان الهدف الرئيسي من هذه الورقة البحثية هو بناء نموذج للكشف عن لغة الإشارة آنيًا باستخدام كل من الشبكات العميقة الممثلة في مكتبة TensorFlow ومكتبة OpenCV، وقد هدف هذا النموذج إلى تقليص فجوة عدم القدرة على التواصل بين المحادثة بين فئة المستفيدين من مجتمع الصم والمستفيدين والعاملين في المكتبات العامة المصرية، ومن خلال هذا النموذج استطاعت الدراسة اكتشاف والتعرف وترجمة أربع إيماءات شائعة بين لغتي الإشارة العربية والأمريكية، وهي "إبهام لأعلى"، "إبهام لأسفل"، "شكرًا"، و"يحيا طويلًا"، وقد تراوحت دقة النموذج المقترح للكشف عن كل إيماءة من إيماءات لغة الإشارة والتعرف عليها بين 98% و100%، مع توفير قدرة النموذج على استيعاب إيماءات أكثر مع ارتفاع معدلات الدقة في ترجمتها.

تم اقتراح نموذج في هذه الورقة البحثية للكشف عن الإشارات وإيماءات اليد في الوقت الفعلي للجمل الشائعة بين لغة الإشارة الأمريكية والعربية وذلك بهدف تلبية احتياجات المكتبات المصرية للتعامل مع فئة المستفيدين من ذوي إعاقة الصمم وضعاف السمع والاحتياجات الخاصة ، تم بناء النموذج المقترح لاكتشاف وترجمة 4 إيماءات رئيسية شائعة الاستخدام ومشتركة بين لغة الإشارة العربية والانجليزية الامريكية، وقد التقطت الدراسة لكل إشارة 20 صورة منفصلة، ثم تم تقسيمها إلى فئتين من الملفات ملف التدريب وملف الاختبار، وقد التقطت الدراسة هذه الصور من عدة زوايا للحصول على نتيجة أفضل في عملية التدريب لتحقيق دقة عالية، وقد أظهرت النتائج التجريبية قدرة النموذج على اكتشاف الإشارة في الصورة والتعرف عليها بمعدل دقة تراوح ما بين 98٪ و 100٪، وقد تم توفير واجهة برمجة تطبيقات TensorFlow للكشف عن الكيانات وتحديد رمز الإشارة بدقة وأداءً أفضل والتعرف عليها في الوقت الفعلي وهو ما يتضح في كل من الشكل رقم 11 والشكل رقم 12.

شكل رقم 11 يوضح قدرة النموذج على تحديد وترجمة رمز إشارة "شكرا" بدقة بلغة 100%



شكل رقم 12 يوضح دقة النموذج في اكتشاف وترجمة رمز الإشارة "عيش طويلا" بدقة بلغة 100%

المراجع والمصادر:

1. عبد الفتاح، محمود أ. (2005). لغة الإشـــارة العربية: منظور. مجلة دراســـات الصـــم وتعليم الصـــم، 10(2): 212-221. متوفر على: DOI:10.1093/deafed/eni007 .

2-Asadi-Aghbolaghi, Maryam, Albert Clape´s, Marco Bellantonio, Hugo Jair Escalante, V´ıctor Ponce-Lo´pez, Xavier Baro´, Isabelle Guyon, Shohreh Kasaei, and Sergio Escalera. (2017). Deep learning for action and gesture recognition in image sequences: A survey. In Gesture Recognition, pages 539–578. Springer, 2017.

3-Boesch, Gaudenz. (2023). "LabelImg for Image Annotation". Available at: https://viso.ai/computer-vision/labelimg-for-image-annotation /

4- Bolt, Nancy & Wyber, Stephen. (2017) "Library service to people with disabilities". IFLA. Available at: https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/topics/librariesdevelopment/documents/report_to_medd_on_library_service_to_people_with_disabilities_final_2_14_18.pdf

5-Cheok, Ming. Jin, Zaid Omar, and Jaward Mohamed Hisham. (2019). A review of hand gesture and sign language recognition techniques. International Journal of Machine Learning and Cybernetics, 10(1):131-153.

6-Day, John Michael. (2004) "Guidelines for library services to deaf people". IFLA. Available at : https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/professional-report/62.pdf

7-Deng, Xiaoming, Shuo Yang, Yinda Zhang, Ping Tan, Liang Chang, and Hongan Wang. (2017). Hand3d : Hand pose estimation using 3d neural network. arXiv preprint arXiv:1704.02224.

8-Ferreira, Pedro M, Jaime S Cardoso, and Ana Rebelo. (2019). On the role of multimodal learning in the recognition of sign language. Multimedia Tools and Applications, 78(8):10035-10056 .

9-Guo, Hengkai, Guijin Wang, Xinghao Chen, and Cairong Zhang. (2017). Towards good practices for deep 3d hand pose estimation. arXiv preprint arXiv:1707.07248.

10-Kaggle. (2023). https://www.kaggle.com /

11-Keras. (2023). https://keras.io /

12-Kodandaram. Satwik Ram, N. Kumar, and Sunil Gl. (2021). "Sign language recognition". Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12 :994-1009, 08.

13-L, i Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. (2021). A survey of convolutional neural networks : analysis, applications, and prospects. IEEE transactions on neural networks and learning systems.

14-LeCun, Y., Touresky, D., Hinton, G., & Sejnowski, T. (1988, June). A theoretical framework for back-propagation. In Proceedings of the 1988 connectionist models summer school (Vol. 1, pp. 21-28).

15-Libguides: Deaf statistics: Africa amp; the Americas. Africa amp; the Americas - Deaf Statistics – Lib Guides at Gallaudet University Library.

16-LibGuides.(2023) .Deaf statisticsAfricatheAmericas. https://libguides.gallaudet.edu/c.php?g=773916p=5553200.

17-Lim, Kian Ming, Alan Wee Chiat Tan, Chin Poo Lee, and Shing Chi- ang Tan. (2019). Isolated sign language recognition using convolutional neural network hand modelling and hand energy image. Multimedia Tools and Applications, 78(14):19917-19944.

18-McInnes. John M and Treffry, Jacquelyn A. (1993). "Deaf-blind infants and children: A developmental guide". Available at: https://www.jstor.org/stable/10.3138/j.ctt13x1pr2 .

19-Mekala. Priyanka, Ying Gao, Jeffrey Fan, and Asad Davari. (2011). Real-time Sign language recognition based on neural network architecture. In 2011 IEEE 43rd Southeastern symposium on system theory, pages 195-199. IEEE, 2011.

20- MNIST. (2023). https://www.kaggle.com/datasets/hojjatk/mnist-dataset

21- مصطفى، ولاء م. (٢٠١٧). "خدمات المعلومات في مكتبات الصم في مصر". مجلات Cybrarians: http://journal.cybrarians.info/index.php?option=com_content&view=article&id=777%3Awmegahed&catid=303%3Areviews&Itemid=80

22-Nandy. Anup, Jay Shankar Prasad, Soumik Mondal, Pavan Chakraborty, and Gora Chand Nandi. (2010). Recognition of isolated Indian sign language gesture in real time. In International conference on business administration and information processing, pages 102-107. Springer.

23-OpenCV. (2023) https://opencv.org/

24-Oudah, Munir & Al-Naji, Ali Abdulelah & Chahl, Javaan. (2020). Hand Gesture Recognition Based on Computer Vision : A Review of Techniques. Journal of Imaging. 6. 73. 10.3390/jimaging6080073.

25-Rastgoo, Razieh, Kiani, Kourosh, and Escalera, Sergio. (2021). “Sign language recognition : A deep survey”. Expert Systems with Applications, 164 :113794, 2021. Available at : https://doi.org/10.1016/j.eswa.2020.113794

26-Sense (2023). “Sign language”. Available at: https://www.sense.org.uk/information-and-advice/communication/sign-language /

27-Shivashankar, a Ss and S Srinath. (2018). American sign language recognition system : an optimal approach. International Journal of Image, Graphics and Signal Processing, 11(8):18.

28-Supancˇicˇ, James Steven, Gregory Rogez, Yi Yang, Jamie Shotton, and Deva Ramanan. (2018). Depth-based hand pose estimation: methods, data, and challenges. International Journal of Computer Vision, 126(11):1180-1198.

29-Suthar, Jay. Parikh. Devansh, Sharma. Tanya, Patel. Avi (2021). "Sign Language Recognition for Static and Dynamic Gestures". Global Journal of Computer Science and Technology Volume XXI, Issue II, Version 1, Year 202. Available at: https://globaljournals.org/GJCST_Volume21/1-Sign-Language-Recognition-for-Static.pdf

30-Tang, Ao, Ke Lu, Yufei Wang, Jie Huang, and Houqiang Li. (2015). A real-time hand posture recognition system using deep neural networks. ACM Transactions on Intelligent Systems and Technology (TIST), 6(2):1-23.

31-TensorFlow. (2023). https://www.tensorflow.org/learn

32-Varshini, Chava & Hruday, Gurram & Chandu, Guguloth & Sharif, Shaik. (2020). Sign Language Recognition. International Journal of Engineering Research and. V9. 10.17577/IJERTV9IS050781.

33-Voulodimos, Athanasios, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. (2018). Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018.

34-Wang, Xingyou, Weijie Jiang, and Zhiyong Luo. (2016). Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers, pages 2428-2437.

35-Wu, Jianxin. (2017). to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5(23):495.

36-Yang, Yezhou, Yi Li, Cornelia Fermuller, and Yiannis Aloimonos. (2015). Robot learning manipulation action plans by" watching" unconstrained videos from the

world wide web. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 29.

37-Zheng, Lihong, Bin Liang, and Ailian Jiang. (2017). Recent advances of deep learning for sign language recognition. In 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pages 1–7. IEEE.

38-Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign language recognition: A deep survey. Expert Systems with Applications, 164, 113794. https://doi.org/10.1016/j.eswa.2020.113794

39-Zimmerman, Thomas G, Jaron Lanier, Chuck Blanchard, Steve Bryson, and Young Harvill. (1986). A hand gesture interface device. ACM Sigchi Bulletin, 18(4):189–192 .