

تطوير أنظمة الذكاء الاصطناعي لترجمة لغة الإشارة في الفيديو الأستاذ المشارك أوزر سيليك، أحمد أفجي أوغلو

جامعة إسكي شهير عثمان غازي، ozer@ogu.edu.tr ، قسم الرياضيات وعلوم الكمبيوتر، كلية العلوم،

جامعة إسكي شهير عثمان غازي، إسكي شهير، تركيا

ahmet.avcioglu@engelsizceveri.com جامعة إسكي شهير عثمان غازي مبنى ETGB تكنوبارك رقم:

106/44 إسكي شهير/تركيا

الملخص:

غالبًا ما يواجه الأشخاص الصم أو ضعاف السمع تحديات في فهم الترجمة المكتوبة بالكامل في مقاطع الفيديو بسبب الاختلافات بين لغة الإشارة واللغة المنطوقة، حيث يتمتع كل منهما بقواعد نحوية وبنية فريدة. وفي حين توفر العديد من منصات الوسائط الاجتماعية ترجمة آلية إلا أنها غالبًا ما تكون غير كافية لتحقيق إمكانية النفاذ. لا بد من تضمين ترجمة لغة الإشارة لجعل محتوى الفيديو متاحًا بالكامل. وتتضمن إضافة لغة الإشارة إلى مقاطع الفيديو تقليديًا عملية تستغرق وقتًا طويلًا لتسجيل وإدخال مقاطع فيديو ترجمة منفصلة يجب إعادة تسجيلها عند إجراء أي تغييرات على الفيديو الأصلي. ويبسط البرنامج المساعد الإضافي (plugin) هذه العملية من خلال ترجمة التحديثات ديناميكيًا مباشرة من ملفات الترجمة مما يقلل بشكل كبير من الوقت والجهد والتكلفة. فهو يسمح بدعم لغة الإشارة دون الحاجة لتعديل المحتوى الأصلي. ويتكامل البرنامج المساعد الإضافي مع مشغلات الفيديو مما يوفر نافذة لغة إشارة قابلة للتخصيص يتحكم فيها المستخدم ويمكن تنشيطها أو نقلها أو تغيير حجمها أو إيقاف تشغيلها حسب الرغبة.

الكلمات الرئيسية: ترجمة لغة الإشارة، الذكاء الاصطناعي، البرمجة اللغوية العصبية، النفاذ الرقمي.

1. المقدمة

غالبًا ما تخطي الحواجز السمعية التي يواجهها الأفراد الصم أو ضعاف السمع التحديات السمعية، وهي تستبعدهم من أنشطة أساسية في الحياة اليومية وخاصة عندما يتعلق الأمر بفهم اللغة المكتوبة. وتختلف اللغات المنطوقة ولغة الإشارة اختلافًا كلياً ليس فقط من حيث الشكل ولكن أيضًا في هياكلها النحوية وقواعدها ودلالاتها. ويعني هذا الاختلاف الأساسي أن الاعتماد على الترجمة المكتوبة وحدها غير كافٍ لضمان النفاذ الكامل لمجتمعات الصم وضعاف السمع. وحتى مع الانتشار الواسع للترجمة الآلية عبر منصات الوسائط الاجتماعية وخدمات البث فإن العديد من الأفراد يواجهون صعوبة في تفسير هذه التسميات التوضيحية لأنها لا تعكس تمامًا قواعد اللغة أو الفروق الدقيقة في لغة الإشارة. وبالتالي فإن جهود إمكانية النفاذ الرقمي التي تعتمد فقط على الترجمة المكتوبة تتجاهل جانبًا أساسيًا من جوانب التواصل - المعلومات المنقولة من خلال لغة الإشارة.

وقد أبرزت العديد من الدراسات أهمية لغة الإشارة في تعزيز إمكانية النفاذ للصم وضعاف السمع مؤكدة أن الترجمة على الرغم من كونها مفيدة ليست بديلًا عن الترجمة بلغة الإشارة [1]. وعلى سبيل المثال يسلط تقرير صادر عن الاتحاد العالمي للصم الضوء على أن أكثر من 70 مليون شخص حول العالم يستخدمون لغة الإشارة كوسيلة أساسية للتواصل مما يؤكد على

أهمية توفير وسائل الإعلام القابلة للنفاذ والتي تتضمن ترجمة لغة الإشارة [3]. ونتيجة لذلك فإنه من الأهمية بمكان تجاوز تدابير إمكانية النفاذ غير الكافية ودمج الترجمة بلغة الإشارة في محتوى الفيديو لتلبية احتياجات الأشخاص الصم وضعاف السمع.

يستكشف هذا البحث الأساليب التقليدية لدمج لغة الإشارة في محتوى الفيديو والقيود المفروضة على هذه الأساليب وكيف يمكن للتكنولوجيا الحديثة وخاصة الذكاء الاصطناعي إحداث ثورة في هذه العملية. ونقترح مجموعة من البرامج المساعدة الإضافية المبتكرة للويب والفيديو وملفات PDF التي تسمح بإنشاء ترجمة بلغة الإشارة في الوقت الفعلي متزامنة مع الترجمة. ويوفر هذا النظام حلًا مستدامًا وبأسعار معقولة ويمكنه التكيف مع التحديثات على الفيديو الأصلي مما يضمن للمستخدمين الصم وضعاف السمع النفاذ المستمر إلى المحتوى القابل للنفاذ بالكامل.

المنهجية

تتطلب عملية تطوير منصة محتوى فيديو قابل للنفاذ بالكامل للصم وضعاف السمع التغلب على العديد من التحديات سواء الفنية أو العملية. ويتضمن توفير الترجمة بلغة الإشارة في مقاطع الفيديو تقليديًا عملية متعددة الخطوات تبدأ بترجمة النص إلى لغة الإشارة بواسطة مترجم مدرب. وتتطلب هذه العملية إعداد استوديو احترافي حيث يتم تسجيل فيديو المترجم وتحريه ثم إدخاله على محتوى الفيديو الأصلي. وفي كل مرة يتم فيها تحديث الفيديو أو الترجمة الأصلية يجب أيضًا إعادة تسجيل فيديو لغة الإشارة وهو أمر يستغرق وقتًا طويلًا ويعتبر مكلفًا. إن هذه الطريقة تحد من إمكانية وقابلية التوسع في مجال إضافة دعم لغة الإشارة إلى مجموعة واسعة من محتوى الفيديو.

وغير هذا النهج المقترح هذه العملية من خلال الاستفادة من الذكاء الاصطناعي وتقنيات مزامنة الترجمة المتقدمة. حيث يستخرج النظام الذي طورناه المعلومات من ملف الترجمة في الفيديو ويستخدمها لتوليد ترجمة بلغة الإشارة في الوقت الفعلي لك إطار فيديو. ويلغي هذا الأمر الحاجة إلى إعادة التسجيل المكلفة أو التحرير بعد الإنتاج كلما تم إجراء تغييرات على محتوى الفيديو الأصلي أو الترجمة (الشكل 1).



الشكل 1. لقطة شاشة من فيديو مؤسسة إسطنبول للفنون والثقافة حيث يتم استخدام البرنامج الإضافي (SignForDeaf) لترجمة لغة الإشارة بالفيديو من موقعهم على الويب

(<https://www.iksv.org/tr/haberler/iksv-alt-kat-yepyeni-bir-cevrimici-seriye-basliyor-kimbu-cizgi-filmciler>).

ويكمن جوهر هذا الحل في قدرته على ترجمة التحديات من الترجمة النصية إلى لغة الإشارة بشكل مستدام مما يوفر ترجمة بلغة الإشارة تكون مستمرة ودقيقة ومتزامنة مع الفيديو. كما أن دمج هذا النظام مع منصات الفيديو الحالية مثل اليوتيوب يعد أمراً بسيطاً ولا يتطلب أي

تعديل على ملف الفيديو الأصلي. حيث يعمل البرنامج الإضافي بدلاً من ذلك كطبقة فوقية مما يسمح للمستخدمين بتنشيط أو نقل أو تغيير حجم أو تعطيل نافذة لغة الإشارة بناءً على تفضيلاتهم. وتضمن هذه الميزة سهولة الاستخدام إمكانية تخصيص النظام وفقاً للاحتياجات المحددة لكل مشاهد مما يعزز إمكانية النفاذ دون المساس بالسلامة البصرية للفيديو.

كما تم تصميم البرنامج الإضافي مع أخذ القدرة على التكيف في المستقبل في الحسبان. فمع إضافة لغات إشارة جديدة يمكن تحديث النظام بسهولة لاستيعاب أقاليم ولغات مختلفة مما يضمن تطبيقه في سياقات لغوية متنوعة. ويسمح التصميم المعياري للبرنامج أيضاً بمزيد من التطورات في مجال الذكاء الاصطناعي مثل التحسينات في التعرف على اللغة في الوقت الفعلي والترجمة المتنوعة لبنى الجمل المعقدة ليتم دمجها دون الحاجة إلى إجراء تغييرات كبيرة في الإطار الحالي.

المزامنة التلقائية للترجمة

إن أحد أكثر الجوانب المبتكرة في البرنامج الإضافي الخاص بنا هو قدرته على الاستفادة من الذكاء الاصطناعي لمزامنة الترجمة مع الترجمة المناسبة بلغة الإشارة. حيث أنه يجب في الأنظمة التقليدية ضبط توقيت مقاطع فيديو لغة الإشارة بعناية ومزامنتها يدوياً مع محتوى الفيديو وهي عملية شاقة خاصة عند التعامل مع مقاطع الفيديو التي يتم تحديثها بشكل متكرر. ففي كل مرة تتم فيها إضافة ترجمة جديدة أو تعديلها يلزم إعادة تسجيل الترجمة بلغة الإشارة بالكامل متبوعة بإعادة دمجها في الفيديو الأمر الذي يستهلك الوقت والموارد.

ويتكيف حل (SignForDeaf) المدعوم بالذكاء الاصطناعي بشكل ديناميكي مع أي تغييرات يتم إجراؤها في ملفات الترجمة ويقوم تلقائياً بتحديث الترجمة بلغة الإشارة المقابلة دون

الحاجة إلى تدخل يدوي. وتتيح هذه الميزة مزامنة تستجيب للتحديثات المتكررة، ولا يضمن مستوى الأتمتة توفير الدقة فحسب بل يوفر أيضًا القدرة على التعامل مع كميات كبيرة من المحتوى مما يجعله قابلاً للتطوير للمؤسسات التي تنتج مخرجات فيديو متكررة ومتنوعة وهي ميزة تفتقر إليها الحلول السابقة.

نافذة لغة الإشارة القابلة للتخصيص

تتمثل إحدى الميزات الرئيسية التي تميز البرنامج الإضافي الخاص بنا في نافذة لغة الإشارة القابلة للتخصيص بالكامل. حيث إنه غالبًا ما يتم انتقاد طول إمكانية النفاذ لكونها غير مرنة ولكن هذا النظام يعطي الأولوية لتجربة المستخدم في تصميم نظام ترجمة بلغة الإشارة بالفيديو. ويتمتع المشاهدون بالتحكم الكامل في عرض نافذة لغة الإشارة مما يضمن إمكانية تكييفها مع تفضيلاتهم الفردية. ويتضمن هذا التخصيص خيارات لتحريك النافذة أو تغيير حجمها أو حتى تعطيلها حسب الحاجة مما يمنح المستخدمين المرونة لضبط العرض بناءً على بيئة المشاهدة والراحة الشخصية. فقد يفضل المستخدم الذي يشاهد مقطع فيديو على شاشة صغيرة مثل الهاتف المحمول على سبيل المثال تصغير حجم نافذة لغة الإشارة أو نقلها إلى زاوية الشاشة حيث لا تحجب العناصر المرئية المهمة. وعلى العكس من ذلك فقد يختار المستخدم الذي يشاهد على شاشة أكبر تكبير النافذة لمنحه مزيداً من مجال الرؤية.

ويضمن هذا النهج الذي يركز على المستخدم ألا تتداخل الترجمة بلغة الإشارة مع محتوى الفيديو الرئيسي مع الحفاظ على إمكانية النفاذ إليه بسهولة حسب الحاجة. كما يمكن للمستخدمين تمكين أو تعطيل ميزة لغة الإشارة في أي وقت مما يضمن مشاهدة الفيديو

دون تشتيت لأولئك الذين يختارون تعطيلها. ويمكن أن تعمل هذه المرونة على تحسين إمكانية النفاذ إلى محتوى الفيديو وشموله بشكل كبير وخاصة بالنسبة لمجتمعات الصم وضعاف السمع.

تقليل الوقت والتكلفة

عادةً ما تكون الطرق التقليدية لإضافة الترجمة بلغة الإشارة مكلفة وبطيئة. حيث تتضمن العملية عادةً توظيف مترجم لغة إشارة محترف وتسجيل ترجمته في الاستوديو ثم دمج فيديو لغة الإشارة في المحتوى الرئيسي من خلال التحرير بعد الإنتاج. وتتطلب هذه الطريقة موارد بشرية ومالية كبيرة مما قد يجعلها غير مجدية لمنشئي المحتوى الصغار أو المنظمات ذات الميزانيات المحدودة. وعلاوة على ذلك فإن أي تغييرات على الفيديو الأصلي أو نص الترجمة ستتطلب تكرار هذه العملية بالكامل مما يؤدي إلى خسائر في التكلفة والوقت.

ويقضي البرنامج الإضافي الخاص بنا على هذه التحديات من خلال إزالة الحاجة إلى إعادة التصوير المستمر والتكامل اليدوي. فبمجرد تنصيب البرنامج الإضافي يتم إنشاء الترجمة بلغة الإشارة تلقائيًا بناءً على الترجمة النصية مما يعني أن أي تغييرات على الترجمة ستنعكس على الفور في الترجمة بلغة الإشارة. ويقلل هذا بشكل كبير من خسائر الوقت والتكلفة المطلوبة للحفاظ على محتوى الفيديو القابل للنفاذ مما يتيح للمزيد من المبدعين والمعلمين والمنظمات تقديم الترجمة بلغة الإشارة دون العبء المالي الإضافي. ولنأخذ مؤسسة تعليمية كبيرة أو منصة إعلامية على سبيل المثال حيث يتم إنتاج مئات مقاطع الفيديو كل شهر وهنا يمكن أن يكون توفير الوقت كبيراً مما يسمح بتخصيص الموارد لإنشاء محتوى جديد أو تحسين ميزات إمكانية النفاذ الأخرى بدلا من ذلك.

بالإضافة إلى ذلك فإن انخفاض تكلفة التنفيذ يعني أن المنظمات الأصغر أو المبدعين المستقلين الذين ربما لم يتمكنوا في السابق من تحمل تكاليف دعم لغة الإشارة سيتمكنهم الآن تقديم مقاطع فيديو قابلة للنفاذ بالكامل. ويوسع هذا الأمر من نطاق المحتوى القابل للنفاذ عبر منصات وصناعات مختلفة من مقاطع الفيديو التعليمية والدورات التدريبية عبر الإنترنت إلى المواد الترفيهية والتدريبية للشركات. وفي نهاية المطاف لن تنحصر هذه الفائدة في مجتمع الصم وضعاف السمع من خلال توفير محتوى أكثر شمولاً، بل ستشجع أيضاً على تبني ممارسات إمكانية النفاذ على نطاق أوسع عبر المشهد الإعلامي.

الخاتمة

يمثل تطبيق أنظمة الترجمة الآلية بلغة الإشارة تقدماً مهماً في جعل المحتوى الرقمي أكثر نفاذاً بالنسبة لمجتمعات الصم وضعاف السمع. ومن خلال معالجته لقيود النفاذ التقليدية القائمة على الترجمة فإن هذا الحل يقدم خياراً أكثر شمولاً وديناميكية لتوفير ترجمة متزامنة بلغة الإشارة. ولا يخفف هذا النهج الأعباء الزمنية والمالية المرتبطة عادةً بإنشاء مقاطع فيديو بلغة الإشارة فحسب بل يضمن أيضاً إمكانية تحديث ميزات إمكانية النفاذ بسهولة مع تطور المحتوى. وهذا أمر مهم في السياقات التعليمية والترفيهية والمهنية حيث يتم تحديث المحتوى بشكل

متكرر. [2]

يفتح هذا النظام إمكانيات جديدة لمستقبل إمكانية النفاذ في وسائل الإعلام. فمع استمرار الذكاء الاصطناعي في التقدم في مجالات مثل معالجة اللغة الطبيعية والتعرف على الإيماءات يمكننا أن نتوقع ترجمات أكثر تطورًا ودقة من حيث السياق مما قد يؤدي إلى سد الفجوات بين لغات الإشارة المختلفة واللغات المنطوقة في جميع أنحاء العالم.[5] ويعتبر هذا الأمر هاماً ليس فقط لمجتمع الصم وضعاف السمع ولكن للمجتمع ككل لأنه يعزز الشمول ويكسر حواجز الاتصال عبر الخطوط اللغوية والثقافية.[4] فمن خلال دمج الترجمة بلغة الإشارة المدعومة بالذكاء الاصطناعي في منصات الفيديو السائدة ستتاح الفرصة لمنشئي المحتوى والمنظمات لتحسين تجربة المشاهدة بشكل كبير لجمهورهم. وتسمح الطبيعة القابلة للتخصيص وسهولة الاستخدام للبرنامج الإضافي للترجمة بلغة الإشارة بالفيديو للأفراد بتخصيص تجربة إمكانية النفاذ الخاصة بهم مما يجعله أداة متعددة الاستخدامات يمكنها استيعاب مجموعة واسعة من التفضيلات والاحتياجات الشخصية. كما يمكننا مع تزايد عدد المنظمات التي تتبنى هذه التقنيات أن نتوقع تحولاً ثقافياً أوسع نحو تطبيع إمكانية النفاذ إلى المحتوى الرقمي مما يعود بالنفع ليس فقط على مجتمع الصم وضعاف السمع بل وأيضاً على فهم المجتمع للشمول.[6]

في الختام نرى أنه للاستخدام الواسع النطاق لأدوات الترجمة بلغة الإشارة المدعومة بالذكاء الاصطناعي القدرة على إعادة تشكيل مشهد الفيديو على شبكة الإنترنت مما يجعل المحتوى الرقمي متاحاً للجميع. إنه حل لا يعالج تحديات إمكانية النفاذ الحالية فحسب بل يمهّد الطريق أيضاً للتقدم المستقبلي مما يضمن عدم تخلف أي شخص عن الركب مع تقدم التكنولوجيا.

الشكر

نود أن نشكر زملائنا وشركاء البحث على مساهماتهم القيمة في هذا المشروع.

- Wilson, M., & Moffat, P. (2018). The impact of subtitles and sign language on video accessibility. *Journal of Deaf Studies and Deaf Education*, 23(2), 204-215.
doi:10.1093/deafed/eny012
2. Arfé, B., Rossi, C., & Sicoli, S. (2014). The role of sign language in reading comprehension for deaf individuals. *Frontiers in Psychology*, 5, 1174.
doi:10.3389/fpsyg.2014.01174
3. World Federation of the Deaf (2021). Global accessibility report on sign language use in media. WFD Publications, pp. 12-36.
4. Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press, pp. 45-89. ISBN: 9780521016505
5. Berke, J. (2020). AI and Accessibility: Bridging the communication gap for the deaf community. *AI Journal of Linguistics*, 15(3), 235-250.
doi:10.1111/ail.153235
6. Napier, J., Leigh, G., & Goswell, D. (2016). *Sign Language Interpreting: Theory and Practice in Australia and New Zealand*. Federation Press, pp. 102-145.
ISBN: 9781760021162