

مترجم افتراضي قائم على التعلم الفيدرالي (الموزع) للتعرف على لغة الإشارة العربية في المدن الذكية

أحمد الزعبي¹، توفيق الحضرمي²، أمجد البشايرة³، لجين بني يونس¹

¹ قسم علوم الحاسوب، جامعة العلوم والتكنولوجيا الأردنية، إربد، الأردن

lhbaniyounis19@cit.just.edu.jo، agalzubi@just.edu.jo

² كلية العلوم والتكنولوجيا، جامعة نوتنجهام ترينت، نوتنجهام، المملكة المتحدة

tawfik.al-hadhrami@ntu.ac.uk

³ قسم علوم الحاسوب، الجامعة الأردنية، عمان، الأردن

amalbashayreh20@cit.just.edu.jo

الملخص:

تتميز لغة الإشارة العربية بقواعد نحوية وتركيبية منظمة مما يستلزم الالتزام بهذه القواعد في توليد لغة الإشارة الآلية للأشخاص ذوي الإعاقة. وبالتالي فإن التوليد الفعال للغة الإشارة يعتمد بشكل كبير على استخدام المؤشر الافتراضي ثلاثي الأبعاد. وتقدم هذه الورقة إطاراً رائداً يستخدم التعلم الفيدرالي لتطوير مترجم افتراضي للغة الإشارة العربية بهدف تعزيز إمكانية النفاذ والتواصل لمجتمع الصم في المدن الذكية. ويضع النموذج المقترح خط الأساس لتسهيل إنشاء تطبيقات مبتكرة تولد ترجمة دقيقة للغة الإشارة من حيث السياق والقواعد النحوية. ويحافظ هذا الإطار من خلال استخدام التعلم العميق الفيدرالي على خصوصية المستخدم مع السماح بالتحسين المستمر لأداء المترجم. ويهدف هذا النموذج إلى تعزيز الشمول والتكنولوجيا المساعدة من خلال دمج لغة الإشارة في الحلول التكنولوجية الحضرية.

الكلمات الرئيسية: لغة الإشارة العربية؛ التعلم العميق الفيدرالي؛ إمكانية النفاذ؛ الخصوصية الافتراضية؛ المدينة الذكية.

1. المقدمة

تعتمد المدن الذكية على أجهزة استشعار إنترنت الأشياء (IoT) لجمع البيانات وهي تدعم مجموعة من التطبيقات في مجالات تشمل الخدمات العامة وإدارة الموارد والاتصالات (Zheng et al., 2022). كما تقدم المدن الذكية حلولاً فعالة للتحديات الرئيسية مثل تطوير إنترنت الأشياء (Li et al., 2020) والرعاية الصحية (Ghazal et al., 2021) والنقل (Ushakov et al., 2022) وأنظمة الاتصالات (Guan et al., 2018). وتنتج أجهزة الاستشعار كميات هائلة من البيانات خلال عملية تبادل البيانات المكثفة هذه. وغالبًا ما يعتبر الصم وضعاف السمع (DHH) الذين يستخدمون لغة الإشارة بشكل أساسي أنفسهم جزءًا من أقلية لغوية وثقافية حيث يفتقرون إلى النفاذ الكامل إلى نفس الموارد اللغوية المتاحة للآخرين (Bramwell et al., 2000).

عادة ما تكون خدمات ترجمة لغة الإشارة التقليدية وتدابير إمكانية النفاذ غير قادرة على تلبية احتياجات الأشخاص ذوي الإعاقة البصرية وخاصة في المناطق غير الغربية (Othman et al., 2024). ويقدم المؤشرون "المترجمون" الافتراضيون العاملين بالذكاء الاصطناعي حلاً واعداً لسد فجوة إمكانية النفاذ هذه في المدن الذكية. ويمكن للمترجم الافتراضي أن يعزز دمج لغة الإشارة وظهورها ضمن المحتوى الرقمي والخدمات العامة مما يخلق بيئات حضرية أكثر قابلية للنفاذ. ومن خلال دمج المترجم الافتراضي للغة الإشارة في البنية التحتية للمدينة الذكية سيكون بإمكان الأشخاص ذوي الإعاقة البصرية التمتع بالنفاذ إلى الاتصالات في مجالات مثل النقل والرعاية الصحية والمعلومات العامة مما يعزز الهوية الثقافية واللغوية لمجتمعات الصم على مستوى العالم.

إن التعلم الفيدرالي هو نهج لامركزي للتعلم الآلي يمكن أجهزة أو كيانات متعددة من تدريب نموذج مشترك بشكل تعاوني مع الاحتفاظ بالبيانات مخزنة محلياً (Kairouz et al., 2021). كما أن استخدام الشخصيات الافتراضية في المدن الذكية التي تستفيد من التعلم الفيدرالي يمكنها من أن تصبح أكثر تكيفاً مع احتياجات المستخدم مع الحفاظ على خصوصيته. وبهذه الطريقة تعمل كل شخصية افتراضية محلياً على جهاز المستخدم أو داخل بيئة معينة مثل كشك عام أو مركز رعاية صحية وفق بياناتها الفريدة والخاصة. فعلى سبيل المثال يمكن للشخصيات الافتراضية التي تخدم الصم وضعاف السمع معالجة المعلومات المحلية للرد على استفسارات المستخدم أو ترجمة لغة الإشارة أو تقديم إرشادات في التنقل ويتم كل ذلك مع تخزين بيانات التفاعل الحساسة هذه محلياً.

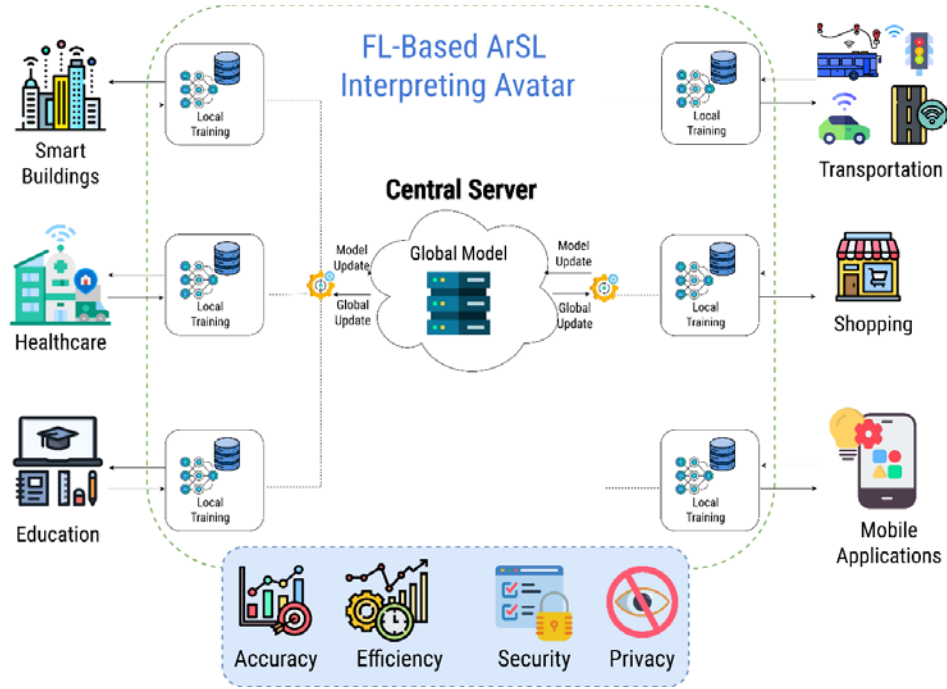
نقدم في هذه الورقة نموذجاً أولياً للشخصية الافتراضية العربية الواقعية التي تترجم لغة الإشارة العربية

(ArSL) في أنظمة المدن الذكية مع الحفاظ على الكفاءة والدقة والخصوصية والأمان. وتقوم الشخصية الافتراضية القائمة على التعلم الفيدرالي بإلغاء الحاجة إلى مشاركة البيانات الخام مع خادم مركزي. وبدلاً من ذلك، فهي تقوم بمعالجة البيانات المحلية للتعلم والتكيف وتحديث النماذج لتنتمشي مع تفضيلات واحتياجات كل مستخدم. وترسل كل شخصية افتراضية بشكل دوري تحديثات النموذج مثل الأوزان أو الميزات المحدثة إلى مجمع مركزي. ويقوم هذا المجمع بالجمع بين هذه التحديثات لتطوير نموذج عالمي يتم توزيعه بعد ذلك مرة أخرى على الشخصيات الافتراضية. ويتيح هذا الأمر دمج المعرفة الجماعية لجميع المستخدمين مع ضمان عدم تبادل المعلومات الشخصية أبداً. كما تسمح هذه التقنية التعاونية لكل شخصية افتراضية بتحسين الدقة والوعي السياقي بمرور الوقت. فمن الممكن على سبيل المثال للشخصية الافتراضية التي تخدم الصم وضعاف السمع في المستشفى أن تتعلم كيفية تفسير التعليمات الطبية في حين يمكن لشخصية افتراضية تعمل في مركز النقل أن تعزز قدرتها على تقديم إرشادات التنقل في الوقت الفعلي. ويمكن للشخصيات الافتراضية كنتيجة لعملية التعلم الفيدرالية اكتساب المعرفة في مجموعة متنوعة من الظروف والاستفادة من الخبرات في بيئات متعددة.

تم تقسيم الجزء المتبقي من هذه الورقة على النحو التالي: يسلط القسم 2 الضوء على الفوائد الرئيسية لاستخدام التعلم الفيدرالي في لغة الإشارة العربية. ويستعرض القسم 3 الأعمال ذات الصلة القائمة على الشخصية الافتراضية. ويقدم القسم 4 هيكل الشخصية الافتراضية المقترحة للغة الإشارة العربية في المدن الذكية. ويوضح القسم 5 إطار اتصال الشخصية الافتراضية ويناقش القسم 6 نتائج المحاكاة التجريبية ليختتم القسم 7 أخيراً هذا العمل.

2. فوائد التعلم الفيدرالي للغة الإشارة العربية في المدن الذكية

يوفر تبني نموذج التعلم الفيدرالي للغة الإشارة العربية في المدن الذكية العديد من المزايا بما في ذلك الأداء المحسن وإمكانية النفاذ والتكيف في الوقت الفعلي وتعزيز الخصوصية وقابلية التوسع والتعلم التعاوني. ويوضح الشكل 1 مزايا استخدام الشخصية الافتراضية القائمة على لغة الإشارة العربية في المدن الذكية والتي سيتم مناقشتها بمزيد من التفصيل في الأقسام الفرعية التالية.



الشكل 1. نموذج عام للتعلم الفيدرالي في المدن الذكية.

2.1. دقة أفضل للترجمة

ويعني هذا قدرة الشخصية الافتراضية على تقديم استجابات ذات صلة ودقيقة ومناسبة سياقيًا لاحتياجات كل مستخدم (Garcia et al., 2023). حيث يمكن للشخصية الافتراضية الدقيقة فهم نية المستخدم سواء كانت تترجم لغة الإشارة أو تقدم المساعدة في حالات الطوارئ أو تساعد في التنقل في المدينة. ويسمح التكيف السياقي للشخصية الافتراضية بتخصيص استجاباتها بناءً على موقع المستخدم أو تفاعلاته الأخيرة أو احتياجاته الفريدة. ويحسن التعلم الفيدرالي هذه الدقة من خلال تمكين الشخصية الافتراضية من التعلم من التفاعلات المحلية عبر أجهزة مختلفة بدلاً من مركزية البيانات (Pandya et al., 2023). ورغم ذلك فإنه يمكن للعديد من القضايا أن تؤثر بشكل كبير على فعالية التطبيقات المختلفة في هذا المجال في المدن الذكية مثل إدارة المرور والسلامة العامة. وتساهم في هذه التحديات عوامل متعددة مثل عدم تجانس البيانات ونُدرة البيانات ومشاركة العملاء المتغيرة وقيود على الموارد الواردة من أجهزة إنترنت الأشياء. فغالبًا ما تؤدي البيانات غير المتسقة الواردة من مصادر متنوعة إلى جانب البيانات الديناميكية التي يمكن أن تتغير بسرعة إلى توزيعات بيانات غير مستقلة وغير موزعة بشكل متطابق (IID) وفجوات في بيانات التدريب مما يؤدي إلى الحصول على نماذج لا تعكس

ظروف المدينة بدقة.

يهدف النموذج المقترح القائم على التعلم الفيدرالي إلى معالجة هذه التحديات من خلال تعزيز تكامل البيانات من خلال توحيد المدخلات من مصادر مختلفة وتنفيذ تقنية التعلم اللامركزي التكيفي التي تضبط النماذج بناءً على الظروف المحلية. كما أنه يسهل التجميع الهرمي للعملاء بناءً على المناطق الجغرافية أو الخدمات مما يسمح بتدريب نموذج موجه بشكل أفضل. وعلاوة على ذلك يستخدم هذا النموذج تقنيات التجميع القوية لتعزيز مرونة النموذج العالمي مما يضمن دقته على الرغم من الطبيعة المتنوعة والديناميكية لمصادر البيانات في المدن الذكية. كما تعزز قدرة النموذج على التعرف على اللهجة العربية وتفسيرها بدقة من دقته الإجمالية من خلال تحسين استجابته للاحتياجات الفريدة لمجتمع الصم الناطقين بالعربية. وتعمل عملية التحسين المستمر هذه على تعزيز فهم النموذج مما يؤدي إلى استجابات أكثر دقة. وتمكن التحديثات في الوقت الفعلي النموذج من الاستجابة للمواقف المتغيرة مثل توفير تحديثات الطقس أو حركة المرور في الوقت المناسب بناءً على الظروف الفعلية للمدينة وبالتالي تعزيز دقة الدليل.

2.2. تحسين التكيف وإمكانية النفاذ في الوقت الفعلي

يتمثل الهدف هنا في تقديم مساعدة سريعة وتستفيد من الموارد المتاحة وأكثر استجابة للمستخدمين. ويتم تحقيق ذلك من خلال دمج الحوسبة الطرفية مع التعلم الفيدرالي مما يتيح للشخصيات الافتراضية العمل بزم من أقل واستهلاك أقل للنطاق الترددي واستخدام محسن للموارد. ومن خلال معالجة البيانات على الأجهزة الطرفية المحلية مثل الخوادم القريبة أو أجهزة المستخدم يمكن للشخصية الافتراضية القائمة على التعلم الفيدرالي تجنب التأخيرات الناجمة عن معالجة السحابة المركزية وتقديم استجابات أسرع وردود فعل في الوقت الفعلي. ويمكن الوصول إلى البيانات المتعلقة بالظروف المحلية مثل حركة المرور أو جداول النقل العام على الفور مما يسمح للشخصية الافتراضية بتوفير أسرع طريق للعودة إلى المنزل إذا لزم الأمر. ومع ذلك فقد يحدث إفراط في استخدام الوقت والبيانات للاتصال في مثل هذه السيناريوهات بسبب الحجم الكبير المحتمل لتحديثات النموذج المرسلة إلى الخادم المركزي وخاصة عند التعامل مع النماذج المعقدة. كما يمكن أن يتسبب العدد الكبير من العملاء في تكاليف إضافية حيث أن كل عميل يرسل تحديثاته قد يسبب زيادة حركة المرور على الشبكة بشكل كبير. وقد يساهم عدم تجانس البيانات أيضاً في هذه المشكلة حيث يمكن أن تؤدي الاختلافات في البيانات عبر العملاء إلى اختلافات في أحجام التحديثات وتردداتها مما يؤدي إلى تعقيد عملية التجميع.

يعالج النموذج المقترح هذه التحديات ويحقق كفاءة عالية من خلال تقليل حجم تحديثات النموذج (على سبيل المثال، الأوزان والتدرجات) المرسلة إلى الخادم المركزي وبالتالي تقليل متطلبات النطاق الترددي بشكل كبير. كما يسمح استخدام التعلم الفيدرالي غير المتزامن (Xu et al., 2023) بمعالجة التحديثات من العملاء بشكل فردي مع تقليل وقت الانتظار ومنع زيادة حركة المرور على الشبكة. ويضمن هذا النهج لمشاركة البيانات الانتقائية أن يظل النموذج خفيفاً وفعالاً مع الاستفادة من تجارب المستخدمين المشتركة في جميع أنحاء المدينة. وبالتالي فإن التعلم الفيدرالي يجعل من الممكن تحسين نماذج لغة الإشارة العربية باستمرار من خلال التكيف المحلي وردود الفعل الموزعة للمستخدمين. ويخلق هذا الأمر لمستخدمي لغة الإشارة بيئات متنوعة في المدينة الذكية تكون أكثر شمولاً وتعمل على تحسين التفاعل والتواصل. كما تسمح الطبيعة اللامركزية للتعلم الفيدرالي للنموذج بالتكيف مع اللهجات الإقليمية وتفضيلات المستخدم والاحتياجات المحددة في الوقت الفعلي مما يجعله أكثر فعالية في البيئات الحضرية المتنوعة.

2.3. الحفاظ على الخصوصية والأمان

تعد قدرة التعلم الفيدرالي على الحفاظ على الخصوصية إحدى فوائده الرئيسية. ويتحقق ذلك من خلال استخدام التعلم الفيدرالي والحوسبة الطرفية من قبل الشخصيات الافتراضية المنتشرة في المدن الذكية. وتسمح هذه الطريقة للشخصيات الافتراضية بمعالجة البيانات المحلية والتعلم منها في مناطقها المخصصة مما يضمن بقاء المعلومات الحساسة داخل البيئة المحلية بدلاً من نقلها إلى خادم مركزي. ومع ذلك فإن تسرب المعلومات يمثل مصدر قلق كبير في المدن الذكية التي تعتمد على أنظمة التعلم الفيدرالي حيث قد يستخدم المتسللون استراتيجيات مثل تسميم النموذج أو تسميم البيانات للاستفادة من نقاط الضعف. كما يمكن لخوارزميات مثل الانحدار التدريجي تسريب معلومات المستخدم الحساسة مما يعرض نظام التعلم الفيدرالي لمزيد من الهجمات المحتملة (Pandya et al., 2023).

يتيح نموذج الاتصال المقترح القائم على التعلم الفيدرالي توزيع الشخصيات الافتراضية باستخدام تقنيات الحفاظ على الخصوصية التي تشارك فقط تحديثات النموذج المجهولة بدلاً من البيانات الخام وبالتالي تقليل انتشار البيانات. كما تحمي قنوات الاتصال المشفرة بين الشخصيات الافتراضية جميع التفاعلات بينما تسمح إدارة البيانات اللامركزية لكل شخصية افتراضية بالتحكم في البيانات محلياً مما يعزز الخصوصية. وهكذا فإن هذه التقنية تحترم قواعد الخصوصية وتحافظ أيضاً على ثقة المستخدم من خلال ضمان التعامل مع البيانات الحساسة بأمان على المستوى المحلي. علاوة على ذلك يحافظ إطار التعلم الفيدرالي المقترح على الأمان من خلال تقييد

تأثير الهجمات المحتملة على مناطق محددة بدلاً من تأثيرها على النظام بأكمله حيث يمكن إعطاء التحديثات المنتظمة وتصحيحات الأمان مباشرة لكل شخصية افتراضية مما يضمن حمايتها المستمرة دون الاعتماد على نظام مركزي لمهام الأمان الحرجة.

3. الأعمال ذو الصلة

تم تطوير أنظمة الشخصيات الافتراضية المختلفة لتعزيز إمكانية النفاذ في مجالات التعليم والتواصل والتفاعلات في الوقت الفعلي للأشخاص ذوي الإعاقة وخاصة في المجتمع الناطق باللغة العربية. وتختلف هذه الشخصيات الافتراضية في الوظائف والتركيز والتكنولوجيات الأساسية. وقد تم تصميم عدد قليل فقط من هذه الشخصيات الافتراضية بشكل خاص لمعالجة مشكلات الاتصال مع الأشخاص الصم أو ضعاف السمع. وتقدم (Mindrockets, 2024) شخصيات افتراضية لترجمة النص إلى لغة الإشارة العربية ولغات الإشارة الأخرى. وتعد هذه الشخصيات الافتراضية خفيفة ويمكن دمجها بسهولة في مواقع الويب مع التركيز على إمكانية النفاذ للصم وغيرهم من الأشخاص ذوي الإعاقة مثل أولئك الذين يعانون من اضطراب نقص الانتباه وفرط النشاط (ADHD) أو عسر القراءة. وتتمثل المزايا الأساسية لشخصيات (Mind Rockets) الافتراضية في تغطيتها الواسعة للإعاقات وسهولة تكاملها وتوافقها مع مواقع الويب وتطبيقات الهاتف المحمول والمنصات الرقمية الأخرى. ومع ذلك فقدرتها محصورة على ترجمة النصوص إلى إشارات.

ويعد نظام الشخصية الافتراضية "إشارة" (2024، إشارة) تقدماً كبيراً آخر في مجال ترجمة لغة الإشارة العربية. ويترجم هذا النظام لغة الإشارة العربية إلى نص مكتوب أو منطوق وبالعكس عبر لهجات متعددة. ويركز "إشارة" على الاتصال ثنائي الاتجاه مما يجعله قادراً على ترجمة إيماءات لغة الإشارة إلى نص وترجمة النص إلى لغة الإشارة من خلال شخصية افتراضية متحركة. ويتوفر الآن تطبيق ويب فقط لهذا النظام الذي لا يزال جديداً نسبياً.

إن الشخصية الافتراضية "بوحمدة" (Othman & El Ghoul, 2022) هي ابتكار قطري يترجم النص إلى لغة الإشارة القطرية. وقد لاقت هذه الشخصية الافتراضية قبولا واسعاً من قبل مجتمع الصم المحلي مع ملابسها ذات الأهمية الثقافية (الغتره والثوب) ورسومها المتحركة الواقعية للغاية. وهي شخصية مدعومة بهندسة سحابية متطورة تمكنها من الاستجابة الفعالة في الوقت الفعلي حتى في بيئات النطاق الترددي المنخفض. وهي تعطي

الأولوية للأهمية الثقافية وإمكانية النفاذ وتعمل كجسر بين مجتمعي السمع والصم في قطر. وتتضمن بنيتها مكونات لترجمة النص العربي إلى إيماءات رسومية بلغة الإشارة وقاعدة بيانات من الإشارات والعبارات التوضيحية. ويزيد من روعة هذه الشخصية كل من تصميمها الواقعي وكفاءتها السحابية وتكاملها الثقافي. إن أحد القيود الأساسية على هذه الشخصية هي أنها متاحة باللهجة القطرية فقط مع مشكلات توسع محتملة في المناطق الأخرى الناطقة باللغة العربية.

يوفر نموذجنا المقترح القائم على التعلم الفيدرالي تطورات ملحوظة مقارنة بالأنظمة الحالية بما في ذلك تعزيز الخصوصية وتنوع في اللهجات بشكل أكبر وتحسين إمكانية النفاذ والقدرة على التكيف بشكل أفضل عبر مختلف القطاعات. ويتم تدريب نموذجنا على البيانات اللامركزية من خلال التعلم الفيدرالي على عكس الأنظمة المركزية التقليدية التي تعالج بيانات المستخدم على منصات سحابية وتكشف عن معلومات حساسة. ويضمن هذا الأمر الحفاظ على الخصوصية بشكل كبير مع تحسين قدرات الترجمة باستمرار. كما تم تصميمه للتفاعل بسلسلة عبر منصات وقطاعات مختلفة في المدن الذكية مما يضع معياراً جديداً للتواصل الشامل وإمكانية النفاذ لمجتمع الصم الناطقين باللغة العربية. وبالمقارنة تقدم أنظمة "إشارة" و"بوحمند" و(Mind Rockets) حلولاً محلية قوية ولكنها قد تفتقر إلى جوانب التكيف والخصوصية الواسعة التي يوفرها نموذجنا. فقد تم تصميمه للتفاعلات التكيفية والواعية للسياق حيث يقوم بتخصيص مظهره بناءً على القطاع المحدد (مثل الطب وخدمة العملاء والنقل). ويضمن هذا أن نموذجنا ليس صحيحاً لغوياً فحسب بل إنه مناسب للسياق أيضاً. وتعتمد الأنظمة الحالية على تصميمات ثابتة تفتقر إلى القدرة على التكيف مع القطاعات المختلفة مما يقلل من فعاليتها في تطبيقات المدن الذكية المتنوعة.

الجدول 1. ملخص للشخصيات الافتراضية المستندة إلى التعلم الفيدرالي مقارنة بالشخصيات الافتراضية ذات الصلة في لغة الإشارة العربية.

الشخصية القائمة على التعلم الفيدرالي	إشارة	مايند روكيت (Mind (Rockets	بو حمد	
مجتمع الصم العرب	مجتمع الصم العرب	مجتمع الصم العام والعربي	مجتمع الصم القطري	الفئة المستهدفة
اللهجات والتعبيرات العربية المتنوعة	دعم محدود للهجات	لهجات ولغات متعددة	مقتصر على لغة الإشارة القطرية	دعم اللهجة
اللهجات والتعبيرات العربية المتنوعة	دعم محدود للهجات	لهجات ولغات متعددة	مقتصر على لغة الإشارة القطرية	الأهمية الثقافية والإقليمية
التصميم التكميلي والزي الرسمي الخاص بالمنطقة	أهمية ثقافية محدودة	تخصيص ثقافي محدود	أهمية ثقافية قوية	الخصوصية والأمان
شخصيات افتراضية تكيفية بناءً على القطاع	شخصيات افتراضية قابلة للتخصيص	شخصيات افتراضية ثابتة قابلة للتكيف مع منصات مختلفة	تصميم ثابت لشخصية افتراضية للسياق القطري	تعبيرات الوجه والإيماءات
ملامح الوجه المعبرة والإيماءات الدقيقة	إيماءات لغة الإشارة القياسية	إيماءات لغة الإشارة القياسية	إيماءات ثلاثية الأبعاد واقعية	توافق المنصة

4. هيكل الشخصية الافتراضية

يتضمن الإطار المقترح القائم على لغة الإشارة شخصية افتراضية عربية واقعية تم تطويرها لتسهيل ترجمة لغة الإشارة العربية في أنظمة المدن الذكية مع التركيز على الكفاءة والدقة والخصوصية. وتتطوي ترجمة لغة الإشارة العربية على قضايا فريدة بسبب بنيتها اللغوية المعقدة والموارد المحدودة المتاحة لعلم لغويات لغة الإشارة العربية خاصة عند التعامل مع اللهجات العربية المتنوعة والتعبيرات الإقليمية. وغالبًا ما تعتمد الأنظمة القائمة على الشخصية الافتراضية الحالية على تقنية بدائية لترجمة كلمة بكلمة والتي تفشل في نقل الفروق الدقيقة للتعبير اللغوي الفعلي المطلوب لغة الإشارة العربية وتفتقر إلى الدعم المناسب للهجة (Othman & El Ghoul, 2022). ولمعالجة هذه التحديات تستخدم الشخصية الافتراضية في هذا العمل نهج التعلم الفيدرالي مما يمكنها من تعزيز قدراتها في الترجمة من خلال التدريب على مصادر بيانات متنوعة مع الحفاظ على الخصوصية. ويهدف هذا النهج إلى تحسين القدرة على التعرف على اللهجات العربية المختلفة وترجمتها بدقة وبالتالي زيادة الاستجابة لمطالب مجتمع الصم الناطقين بالعربية.

تمتد إمكانيات الشخصية الافتراضية القائمة على التعلم الفيدرالي إلى مجموعة متنوعة من قطاعات المدن الذكية بما في ذلك النقل العام والرعاية الصحية وخدمة العملاء والتعليم والمزيد. حيث يمكنها تسهيل التفاعلات الأكثر شمولاً وسهولة في النفاذ للأشخاص الصم. وسوف يتكيف مظهر الشخصية الافتراضية في هذه القطاعات ليناسب السياق المعني مما يعزز الموثوقية والواقعية في التواصل. ففي بيئات الرعاية الصحية على سبيل المثال سوف ترتدي الشخصية الافتراضية القائمة على التعلم الفيدرالي زيًا طبيًا وفي بيئات خدمة العملاء سوف ترتدي ملابس رسمية وفي وسائل النقل سيتم استخدام زي رسمي خاص. ويهدف هذا التصميم التكيفي إلى تقديم تجربة غامرة تحترم المستخدمين عبر تطبيقات مختلفة.

يجب التغلب على العديد من المشاكل الفنية لضمان تمكن الشخصية الافتراضية من تلبية احتياجات ترجمة لغة الإشارة في الوقت الفعلي وضمان الواقعية. ويجب أن تؤدي الشخصية الافتراضية رسومًا متحركة واقعية بما في ذلك ملامح الوجه المعبرة والإيماءات الطبيعية والتي تعد حيوية لنقل التفاصيل الدقيقة والعواطف كما هو موضح في الشكل 2. كما تهدف الشخصية الافتراضية القائمة على التعلم الفيدرالي إلى العمل بسلاسة عبر مجموعة متنوعة من المنصات بما في ذلك الهواتف المحمولة ومتصفحات الويب والساعات الذكية مع الحفاظ على الأداء على اتصالات الإنترنت منخفضة السرعة والقدرات الرسومية المحدودة. ومن المطلوب هنا تحقيق التوافق بين الأنظمة الأساسية وقدرات الاستجابة في الوقت الفعلي للتنفيذ الناجح في المدينة الذكية.



الشكل 2. عرض لملامح الجسم التعبيرية للشخصية الافتراضية.

يمثل النموذج المقترح للمترجم الافتراضي القائم على التعلم الفيدرالي تحسناً واعدًا في ترجمة لغة الإشارة القائمة على الشخصية الافتراضية من خلال الجمع بين النمذجة اللغوية المتقدمة والتعلم الفيدرالي ومعالجة أهداف الخصوصية وصحة اللغة والتشغيل الفعّال الذي تتطلبه تطبيقات المدن الذكية. ويهدف نموذج لغة

الإشارة القائمة على التعلم الفيدرالي إلى وضع معيار جديد للتواصل القابل للنفاز من خلال تعزيز بيئة حضرية أكثر شمولاً لمجتمع الصم الناطقين باللغة العربية. وبغرض تحريك الشخصية الافتراضية ثلاثية الأبعاد فإننا نستخدم مكونين مختلفين يحسنان تصويرها المرئي على النحو التالي:

الرسوم المتحركة القائمة على الهيكل العظمي: نقدم للمترجم الافتراضي للغة الإشارة القائم على التعلم الفيدرالي هيكلًا عظميًا ما يسمح بالتحكم الدقيق في حركات الجسم. وقد قمنا بتغليف هيكل الشخصية الافتراضية بدقة بالجلد ليتماشى مع التحولات الصلبة لكل عظمة في الهيكل العظمي. وتضمن هذه التقنية أنه عند تعديل حركة الهيكل سيتم تعديل الشبكة المطابقة بسلاسة مما يسمح بتنفيذ إيماءات الجسم واليد الضرورية للتواصل الفعّال في أنظمة لغة الإشارة بشكل طبيعي.

الرسوم المتحركة المتحركة عبر تشكيل النقاط الهندسية (vertex morphing): بالنسبة للرسوم المتحركة التفصيلية للوجه نستخدم تحويل رؤوس النقاط. ويوضح الشكل 3(أ) كيفية إنتاج تعبيرات وجه معبرة عن طريق تغيير مواضع رؤوس نقاط معينة في هندسة الشخصية الافتراضية. ويمكن استخدام وظائف رياضية محددة لحفظ أو تعديل الموضع الجديد لكل رأس نقطة. وبسبب تطلبها العالي للموارد فإن هذه التقنية تُستخدم فقط لإنشاء حركات الوجه المعقدة. ويعرض الشكل 3(ب) الكائن ثلاثي الأبعاد الذي يمثل بناء وجه الشخصية الافتراضية.

يتطلب خط الأساس المقترح مجموعة من الأدوات والخدمات المتطورة للمساعدة في تحريك الشخصية الافتراضية. وسيمكن هذا التصميم المطورين من دمج الشخصية الافتراضية القائمة على التعلم الفيدرالي بسرعة في تطبيقات مختلفة مع توفير القدرة على تحويل النص المكتوب إلى لغة إشارة عربية.



الشكل 3. تعابير الوجه الأمامي للشخصية الافتراضية (يسار) والوجه ثلاثي الأبعاد (يمين).

5. إطار الاتصالات الفيدرالي

يعمل نموذج المترجم الافتراضي للغة الإشارة القائم على التعلم الفيدرالي ضمن بنية إطار اتصال لامركزية حيث يستخدم كل قطاع نمودجه الخاص. ويتم تزويد كل نمودج بمجموعة بيانات محلية ونمودج محلي مما يسمح له بالتفاعل بشكل فعال مع مجتمع الصم مع تحسين أدائه بشكل مستمر من خلال استراتيجية التعلم الفيدرالية.

5.1. عملية التعلم الفيدرالية

يوضح الشكل 4 عملية التعلم الفيدرالية. حيث يتمتع نمودج كل قطاع (i) بإمكانية النفاذ إلى مجموعة بيانات محلية (D_i) تحتوي على بيانات التفاعل مع المستخدمين الصم المحليين. ويستخدم النمودج هذه البيانات لتدريب نمودجه المحلي (M_i). ويمكن هيكلة عملية التدريب المحلية على النحو التالي:

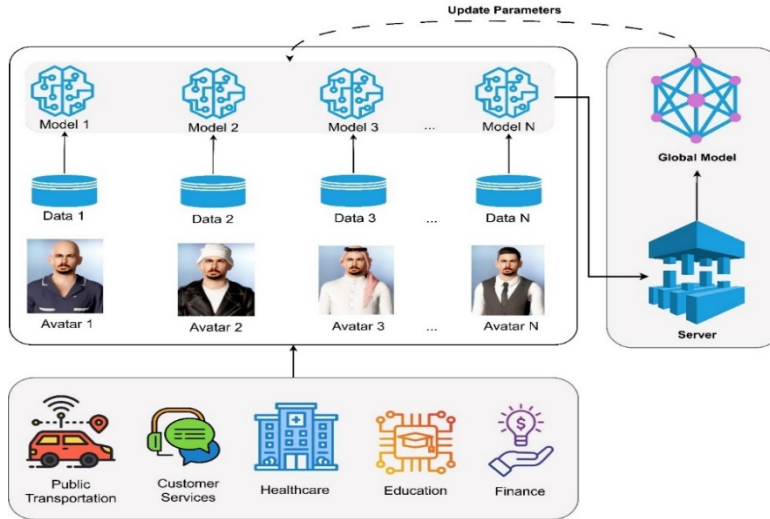
$$M_i^{(t+1)} = M_i^{(t)} - \eta \nabla L(M_i^{(t)}, D_i) \quad (1)$$

حيث

- $M_i^{(t+1)} = M_i^{(t)} - \eta \nabla L(M_i^{(t)}, D_i)$

- η هو معدل التعلم.

- L يمثل دالة الخسارة المحسوبة على مجموعة البيانات المحلية (D_i).



الشكل 4. تواصل الشخصية الافتراضية في التعلم الفيدرالي العميق.

بعد إكمال التدريب المحلي تقوم الشخصية الافتراضية بإنشاء تحديث نموذج (∇Mi) الذي يلتقط التغيرات التي تم إجراؤها على معلماته:

$$\nabla Mi = M_i^{(t+1)} - M_i^{(t)} \quad (2)$$

وفي نهاية فترة محددة مسبقاً (يوميًا على سبيل المثال) يتم إرسال جميع تحديثات النموذج المحلية من الشخصيات الافتراضية في القطاعات المختلفة إلى خادم مركزي للتجميع. ويمكن تمثيل عملية التجميع على النحو التالي:

$$M^{(t+1)} = M^{(t)} + \frac{1}{N} \sum_{i=1}^N \nabla Mi \quad (3)$$

حيث

- $(M^{(t)})$ هي معلمات النموذج العالمي عند التكرار (t) .
- (N) هو العدد الإجمالي من الشخصيات الافتراضية (النماذج المحلية) المشاركة في التحديث.

ثم يتم استخدام التحديث الذي تم جمعه لتحسين النموذج العالمي (M) على الخادم:

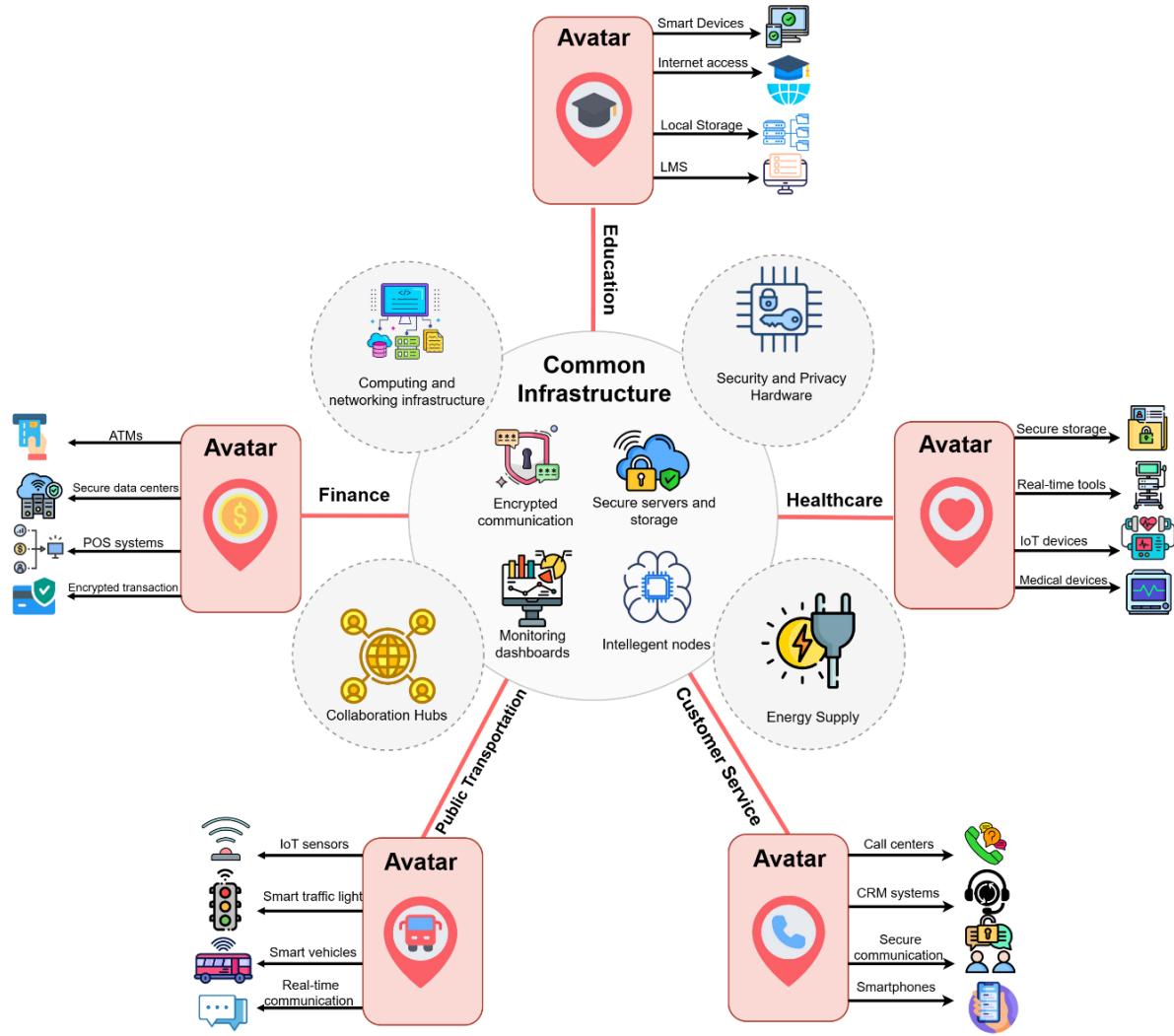
$$M^{(t+1)} = M^{(t)} + \alpha(M^{(t+1)} - M^{(t)}) \quad (4)$$

ويكون (α) هو معامل فائق يحدد مساهمة التحديث الذي تم جمعه في النموذج العالمي. وبعد تحديث النموذج العالمي يتم توزيع المعلومات الجديدة مرة أخرى على كل شخصية افتراضية مما يسمح لهم بإغناء نماذجهم المحلية بالمعرفة العالمية المحسنة. وتتم هذه العملية بشكل متكرر حيث تتعلم كل شخصية افتراضية بشكل مستمر من تفاعلاتها مع المستخدمين الصم المحليين وتقوم بتحديث نموذجها المحلي والمساهمة في المجموع العالمي. ويؤدي هذا النهج إلى نظام شخصية افتراضية متطورة بشكل مطرد قادر على فهم وترجمة لغة الإشارة العربية بشكل أفضل. كما يسمح نهج التعلم الفيدرالي للشخصيات الافتراضية القائمة على التعلم الفيدرالي في المدن الذكية بالحفاظ على خصوصية المستخدمين مع تعزيز قدرات الاتصال. ومن خلال استغلال البيانات المحلية للتعلم الشخصي وجمع تحديثات النموذج بانتظام فإنه سيكون بإمكان الشخصية الافتراضية أن تتكيف بسهولة مع متطلبات مجتمع الصم في مجموعة متنوعة من الصناعات مما يؤدي إلى تحسين الاتصال والإدماج في البيئات الحضرية الذكية.

5.2. البنية الأساسية للتعليم الفيدرالي في سيناريوهات العالم الحقيقي

يتطلب تنفيذ نموذج التعلم الفيدرالي في التطبيقات الواقعية بنية أساسية قوية وقابلة للتطوير ومصممة خصيصًا لتلبية الاحتياجات المحددة لكل قطاع مع الالتزام بالمبادئ المشتركة. ويعتمد التعلم الفيدرالي على الحوسبة الطرفية مع وجود عقد طرفية ذكية مجهزة للتدريب والاستدلال المحلي كما هو موضح في الشكل 5. وتستخدم هذه العقد أجهزة حديثة مثل وحدات المعالجة المركزية ووحدات معالجة الرسومات والمعالجات المتخصصة مثل وحدات معالجة تينسور (Tensor) للقيام بمهام التعلم الآلي المعقدة مما يختصر الوقت ويزيد من الاستجابة (Duan et al., 2023). وعلى سبيل المثال تسمح المعدات الطبية التي تدعم إنترنت الأشياء والخوادم المحلية في الرعاية الصحية بتدريب بيانات المرضى الحساسة بينما توفر وحدات الكمبيوتر الموجودة على متن السيارات معالجة في الوقت الفعلي لبيانات المرور وأجهزة الاستشعار في وسائل النقل العام. وعلى نحو مماثل توفر الأجهزة اللوحية وأجهزة الكمبيوتر المحمولة عند دمجها مع الخوادم المؤسسية التعلم الشخصي في مجال التعليم في حين تعمل خوادم الفروع وأجهزة الصراف الآلي وأجهزة المستهلك على تمكين نماذج التعلم الفيدرالي المحلية في مجال خدمة العملاء والتمويل.

تعتبر أنظمة التخزين عالية السعة ضرورية لإدارة مجموعات البيانات الضخمة والمتفرقة في مختلف الصناعات ولكن خصوصية البيانات تبقى محمية من خلال الحفاظ عليها موضعية في العقد الطرفية (Salh et al., 2023). كما أن حلول التخزين الآمنة في الموقع ذاته تعد ضرورية بما في ذلك محركات الأقراص الصلبة المشفرة (SSDs) للسجلات الطبية والتخزين الموضعي لبيانات المرور والركاب ومجموعات البيانات الأكاديمية في أنظمة إدارة التعلم وسجلات تفاعل العملاء في نماذج الخدمات والبيانات المالية المشفرة للكشف عن عمليات الاحتيال. كما أن هناك حاجة إلى شبكات اتصال فعالة وآمنة مثل خدمة الجيل الخامس أو الألياف الضوئية أو الانترنت اللاسلكي لتحديثات النماذج المستمرة بين العقد الطرفية ومراكز التجميع المركزية (Yang et al., 2023). وتسمح هذه الشبكات للعقد الطرفية بالتواصل بأمان مع معلمات النموذج مع الحفاظ على خصوصية البيانات.



الشكل 5. البنية التحتية في الوقت الحقيقي للمترجم الافتراضي للغة الإشارة العربية في المدن الذكية.

تدعم طبقة التنفيذ في البنية الأساسية للتعليم الفيدرالي مجموعة واسعة من أجهزة الحوسبة بناءً على حالات استخدام متنوعة بما في ذلك وحدات معالجة الرسوميات عالية الأداء لتدريب نماذج التشخيص المعقدة ووحدات المعالجة المركزية القوية للمواقف الديناميكية والأجهزة الطرفية خفيفة الوزن للنماذج الأصغر كما هو موضح في الشكل 5. وتتطلب هذه الأجهزة مصدر طاقة يمكن الاعتماد عليه مع أنظمة احتياطية مثل (UPS) أو المولدات التي تضمن التشغيل المستمر. وتضمن الأنظمة الفعالة استمرار الخوادم في العمل بشكل فعال بينما تيسر الأجهزة القوية إجراء العمليات في المواقف الصعبة مثل البيئات التعليمية البعيدة أو أنظمة النقل.

ويعد التنسيق المركزي أمراً بالغ الأهمية لتحقيق قابلية التشغيل المتبادل وقابلية التوسع والامتثال للمعايير في أنظمة التعلم الفيدرالي. حيث ينفذ الموزع المركزي البروتوكولات المعمول بها لتكامل النماذج مما يوفر تعاوناً سلساً بين العقد والقطاعات الطرفية. كما يعد الامتثال التنظيمي عاملاً حاسماً بما في ذلك أطر العمل مثل (GDPR) للخدمات المالية ومعايير (HL7) للرعاية الصحية. وتوفر لوحات المعلومات ميزة المراقبة في الوقت الفعلي لأداء النظام ومشاركة العقد وقابلية التوسع مما يتيح إدارة أكثر فعالية لهذا النظام.

وقد تثبتت قدرة البنية الأساسية للتعلم الفيدرالي على التكيف في مجموعة متنوعة من المجالات. ففي مجال الرعاية الصحية يسمح التعلم الفيدرالي بالتدريب التعاوني لنماذج التشخيص عبر المؤسسات مع حماية معلومات المرضى الحساسة. كما يعمل التعلم الفيدرالي على تعزيز النقل العام من خلال السماح بتحسين أنماط المرور في الوقت الفعلي وجدولة المسار بناءً على بيانات المركبات. وتستخدم صناعة التعليم التعلم الفيدرالي لبناء نماذج تعليمية مخصصة بناءً على مجموعات بيانات متفرقة من العديد من الجامعات. كما يعمل التعلم الفيدرالي على تحسين تجربة العملاء من خلال إنشاء نماذج قابلة للتكيف تتضمن ملاحظات من مراكز الاتصال البعيدة. كما يساعد التعلم الفيدرالي في مجال التمويل في كشف الاحتيال من خلال السماح للمؤسسات بالتعاون في بيانات المعاملات وتعزيز الدقة مع الحفاظ على السرية. وتظهر الشخصية الافتراضية القائمة على التعلم الفيدرالي قوة التعلم الآلي اللامركزي الذي يحافظ على الخصوصية من خلال معالجة المخاوف الخاصة بمختلف القطاعات وتعزيز الابتكار على مستوى الصناعة مع ضمان الامتثال للمعايير والأمان والكفاءة.

6. النتائج التجريبية

6.1. إعداد المحاكاة

تم استخدام نماذج التعلم العميق (VGG16-VGG19) (Simonyan, 2014) في هذه الورقة كعمود فقري لإطار التعلم الفيدرالي العميق. وقد تم إجراء العديد من التجارب لاختيار أفضل المعلمات الفائقة لهذا النموذج. يتم تقييم الأداء على مجموعة بيانات لغة الإشارة العربية 2018 (Latif et al., 2019) من حيث متوسط وقت التدريب والدقة والضبط والتذكر ودرجة f1. وتستخدم كل تجربة 64 عينة بيانات مع تدريب خمسة عملاء (يمثلون 5 شخصيات افتراضية) على مدى عشر فترات تدريب كاملة. وتقوم طريقة المتوسط الفيدرالي بحساب التدرجات من جانب الخادم بينما يتم استخدام الإنترنت المتقاطعة التصنيفية كدالة خسارة. كما يتم استخدام (Softmax) كدالة تنشيط في حين يتم استخدام الانحدار التدرجي العشوائي كمحسن أساسي بمعدل تعلم 0.01.

6.2. الأداء في مجال التعرف على لغة الإشارة العربية

يلخص الجدول 2 نتائج أداء جميع النماذج القائمة على التعلم الفيدرالي. ويوفر هذا التقييم فهماً تفصيلياً لأداء النموذج عبر خصائص تصنيف الشخصية الافتراضية مما يضمن قدرتها على التعامل مع السيناريوهات المختلفة باستخدام عدة تقسيمات للبيانات. ومع ذلك فإن تقييم أداء النموذج يتم باستخدام مقاييس الدقة والمتوسط الكلي وذلك نظراً للطبيعة غير المتوازنة لمجموعة بيانات لغة الإشارة العربية. ويكون المتوسط الكلي في الظروف غير المتوازنة مفيداً للغاية لأنه يعامل جميع الفئات بالتساوي دون التأثير بالفئة الغالبة. ويتمتع نموذج (FL-VGG19) بأداء مرتفع وفق الاختبار حيث حقق دقة بنسبة 98.8% وضبط بنسبة 98.79% وتذكر بنسبة 98.78% ودرجة F1 بنسبة 98.78%.

الجدول 2. أداء (VGG16) و (VGG19) وفق مجموعة بيانات محاكاة الشخصية الافتراضية.

النموذج	دقة التحقق مقابل دقة	دقة الاختبار	الضبط	التذكر	درجة F1	وقت التدريب لكل عميل (ثانية)	وقت التدريب لكل جولة (دقيقة)
VGG16	97.30	98.70	98.72	98.71	98.71	60	5
VGG19	97.10	98.80	98.79	98.78	98.78	65	5.4

ومع ذلك فقد حقق كل من النموذجين (FL-VGG16) و (FL-VGG19) أداءً مماثلاً عبر جميع المقاييس. حيث يحقق (FL-VGG16) دقة تحقق أعلى قليلاً مع اختلاف طفيف بنسبة 0.2% مما يسلط الضوء على التناقضات الطفيفة بين أداء النموذجين. ويستغرق (VGG16) 60 ثانية لكل شخصية افتراضية لتدريب البيانات المحلية بينما يستغرق (VGG19) 65 ثانية. وتزيد الطبقات الإضافية في (VGG19) من متطلباته الحسابية قليلاً. ومن ناحية أخرى يستغرق (VGG16) 5 دقائق في كل جولة تدريب بينما يستغرق (VGG19) 5.4 دقيقة مما يعكس التعقيد الحسابي الإضافي لـ (VGG19). وبشكل عام فإن نموذج (VGG19) يوفر مقاييس أداء أفضل قليلاً ولكن على حساب زيادة وقت التدريب. ومع ذلك فإن الاختلافات بين النموذجين ضئيلة مما يشير إلى أن أيًا منهما قد يكون مناسباً اعتماداً على المتطلبات المحددة للتطبيق.

7. الخاتمة

تقدم هذه الورقة إطاراً أساسياً لتطوير مترجم افتراضي يعمل بالذكاء الاصطناعي يمكنه إنشاء وترجمة لغة الإشارة باستخدام نهج التعلم الفيدرالي. ويؤدي هذا المترجم الافتراضي المبتكر القائم على التعلم الفيدرالي حركات يدين واقعية وتعبيرات وجه حقيقية ويعطي الأولوية للخصوصية وأمان البيانات من خلال الاستفادة من نموذج التعلم الفيدرالي العميق. ويمكن استخدام هذا المترجم الافتراضي بشكل فعال في بيئة المدينة الذكية مما يتيح لتقنيات المدينة الذكية أن تكون أكثر شمولاً ونفاذاً. ويضع هذا الإطار المقترح الأساس لتعزيز دمج ترجمة لغة الإشارة في تطبيقات مختلفة مما يعزز في النهاية التواصل والتفاعل المحسنين لسكان الصم وضعاف السمع في المناطق الحضرية.

المراجع