



التعرف على تعبيرات الوجه باستخدام التعلم العميق

مقدمة من

م/فاطمة مازن على مازن

رسالة مقدمة إلى كلية الهندسة - جامعة الفيوم
كجزء من متطلبات الحصول على درجة دكتوراه الفلسفة في العلوم الهندسية

قسم الهندسة الكهربائية
تخصص هندسة الإلكترونيات و الاتصالات الكهربائية
كلية الهندسة - جامعة الفيوم

تحت إشراف

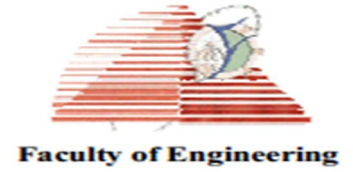
أ.د./ رانيا أحمد عبد العظيم ابو السعود

أستاذة الاشارات الكهربائية - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم

أ.م.د./ أحمد على نشأت إسماعيل

أستاذ مساعد - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم

كلية الهندسة، جامعة الفيوم
الفيوم- جمهورية مصر العربية



التعرف على تعبيرات الوجه باستخدام التعلم العميق مقدمة من

م/فاطمة مازن على مازن

رسالة مقدمة إلى كلية الهندسة - جامعة الفيوم
كجزء من متطلبات الحصول على درجة دكتوراه الفلسفة في العلوم الهندسية

قسم الهندسة الكهربائية
تخصص هندسة الإلكترونيات و الاتصالات الكهربائية
كلية الهندسة - جامعة الفيوم

يعتمد من لجنة الممتحنين:

- أ.د/ عمرو محمد رفعت - أستاذ الاشارات الكهربائية - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم
أ.د/ رانيا أحمد عبد العظيم ابو السعود- أستاذة الاشارات الكهربائية - قسم الهندسة الكهربائية - كلية الهندسة -
جامعة الفيوم
أ.م.د / أحمد على نشأت إسماعيل -أستاذ مساعد - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم
أ.د/ أحمد حسن مدين - أستاذ الإلكترونيات الدقيقة - هيئة الطاقة الذرية - مركز البحوث وتكنولوجيا الاشعاع

كلية الهندسة، جامعة الفيوم
الفيوم- جمهورية مصر العربية

التعرف على تعبيرات الوجه باستخدام التعلم العميق

مقدمة من

م/فاطمة مازن على مازن

رسالة مقدمة إلى كلية الهندسة - جامعة الفيوم
كجزء من متطلبات الحصول على درجة دكتوراه الفلسفة في العلوم الهندسية

قسم الهندسة الكهربائية
تخصص هندسة الإلكترونيات و الاتصالات الكهربائية
كلية الهندسة - جامعة الفيوم

يعتمد من لجنة الممتحنين:

أ.د/ عمرو محمد رفعت - أستاذ الاشارات الكهربائية - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم

التوقيع:

أ.د/ رانيا أحمد عبد العظيم ابو السعود - أستاذة الاشارات الكهربائية - قسم الهندسة الكهربائية - كلية الهندسة -
جامعة الفيوم

التوقيع:

أ.م.د / أحمد على نشأت إسماعيل-أستاذ مساعد - قسم الهندسة الكهربائية - كلية الهندسة - جامعة الفيوم

التوقيع:

أ.د/ أحمد حسن مدين- أستاذ الإلكترونيات الدقيقة - هيئة الطاقة الذرية - مركز البحوث وتكنولوجيا الاشعاع

التوقيع:

كلية الهندسة، جامعة الفيوم
الفيوم- جمهورية مصر العربية

2021

ملخص البحث

يعد التعرف على تعبيرات الوجه البشري أحد أكثر المهام تحديًا في التواصل الاجتماعي. إنها تلعب دورًا مهمًا في مجال الرؤية عن طريق الكمبيوتر والتفاعل بين الإنسان والآلة. إنه مجال بحث يحتوي على تطبيقات ضخمة في المجال الطبي ، والتحقيق في الجرائم ، والتسويق ، والتعلم عبر الإنترنت ، وألعاب الفيديو. يصف الجزء الأول من هذه الأطروحة نموذج مبنى على نموذج الشبكة العصبية العميقة للتعرف على الأنواع السبعة الرئيسية لتعبيرات الوجه ، والتي توجد في جميع الثقافات. هذه هي الغضب والاشمئزاز والخوف والسعادة والحزن والمفاجأة والحياد. تتضمن المنهجية المقترحة أربع مراحل: (أ) المعالجة التمهيدية المسبقة لقاعدة البيانات FER2013 من خلال إعادة توزيع الصور على الشعبة المناسبة لتجنب النتائج المضللة ، والتخلص من الصور التي لا تحتوي على وجوه وصور الوجوه غير الأمامية ؛ (ب) تصميم (CycleGAN) فعال وعلى درجة عالية من الاستقرار والكفاءة ، والذي يوفر ترجمة غير موجهة من تعبير إلى تعبير آخر. تم تصميم وتدريب CycleGAN باستخدام cycle consistency loss جديدة. (ج) إنتاج صور جديدة للتغلب على مشكلة class imbalance ، وخاصة بالنسبة لفئة الاشمئزاز ؛ وأخيرًا (د) بناء بنية الشبكة العصبية العميقة للتعرف على تعبيرات الوجه ، باستخدام نموذج VGG-Face المدربين مسبقًا مع أوزان vggface.

تم اختبار النموذج على النسخة الأصلية لقاعدة البيانات FER2013 والنسخة المتوازنة المعدلة. أيضًا ، تم استخدام النموذج المصمم لتحديد تعبيرات الوجه في الصور بعد تحديد الوجوه باستخدام الشبكات التلافيفية المتتالية متعددة المهام (MTCNN). تظهر النتائج أن النموذج قوي. وقت تشغيل النموذج المصمم للتعرف على تعبيرات الوجه هو 0.44 ثانية. بالإضافة إلى ذلك ، تمت زيادة متوسط دقة الاختبار من 64% لمجموعة بيانات FER2013 الأصلية إلى 91.76% للنسخة المتوازنة المعدلة باستخدام نفس النموذج.

يشتمل الجزء الثاني من الأطروحة على تصميم نظام التعرف على تعبيرات الوجه المعجل بوحدة معالجة الرسومات (GPU) للتعرف على التعبير في تسلسلات الفيديو في الوقت الفعلي.

يشتمل أي نظام للتعرف على تعبيرات الوجه على مرحلتين أساسيتين ، اكتشاف الوجه لتحديد موقع الوجه والتعرف على تعبيرات الوجه لتصنيف التعبير. لسوء الحظ ، تتطلب خوارزميات اكتشاف الوجه قوة حسابية مكثفة ، مما يجعلها خيارًا غير مناسب لأداء مهام اكتشاف الوجه في تسلسلات الفيديو في الوقت الفعلي. لالتقاط تدفقات الفيديو في الوقت الفعلي في بيثون ، تم استخدام مكتبة OpenCV وهي مكتبة مفتوحة المصدر تم بناؤها باستخدام لغة بيثون.

للتغلب على قيود المعالجة ، يجب دفع العمليات الحسابية إلى وحدة معالجة الرسومات (GPU) باستخدام NVIDIA Compute Unified Device Architecture (CUDA). لكن إصدارات OpenCV المتاحة ، للأسف ، لا تحتوي على دعم CUDA لتحقيق الأداء الأمثل على محطات العمل التي تدعم GPU. من أجل الاستخدام الأمثل لموارد الأجهزة ، يجب أن يكون هناك حل للتعامل مع هذه الفجوة بين موارد الأجهزة المتاحة ومكتبات Python التي تستخدم وحدة المعالجة المركزية كخلفية. وحدة OpenCV CUDA وهي عبارة عن مجموعة من الفئات والوظائف لاستخدام قدرات CUDA الحسابية هي المفتاح. إنها أداة مؤثرة للتنفيذ السريع لخوارزميات الرؤية عن طريق الكمبيوتر المتسارعة من CUDA. يشتمل

هذا الجزء من الأطروحة على بناء مكتبة OpenCV من البداية مع دعم CUDA و CUDNN ، وهو حجر الزاوية لنظام التعرف على تعبيرات الوجه في الوقت الفعلي من تدفقات الفيديو. في مرحلة التعرف على تعبيرات الوجه ، تم استخدام النموذج المجمع على قاعدة البيانات FER2013 المتوازنة المعاد تسميتها الجديدة. تم استخدام المخطط المصمم في معالجة الفيديو في الوقت الفعلي لتصنيف الإطارات إلى واحدة من تعبيرات الوجه العالمية: الغضب والاشمئزاز والخوف والسعادة والحزن والمفاجأة والحياد.

بالنسبة لمرحلة تحديد الوجه ، تم استخدام Haar Cascaded والتعلم العميق واختبارهما باستخدام كل من وحدة المعالجة المركزية CPU ووحدة معالجة الرسومات GPU كخلفية وتمت مقارنة النتائج. فيما يتعلق بمقياس الإطار في الثانية (FPS) لبدء دفق الفيديو العام الذي يتضمن اكتشاف الوجه والتعرف على تعابير الوجه ، كان هناك تحسن كبير في FPS بعد استخدام GPU كخلفية في كل من Haar والتعلم العميق بفضل وحدة CUDA الخاصة بالنسخة مجمعة حديثًا من OpenCV. باستخدام وحدة OpenCV's Deep Neural Network (DNN) مع NVIDIA GPUs و CUDA و cuDNN ، تم تحسين FPS من 7.41 على وحدة المعالجة المركزية إلى 23.12 حقة استدلاليةً أسرع بنسبة 312.01٪ للنهج القائم على الميزات. تم تحسين سرعة الاستدلال أيضًا بنسبة تصل إلى 169.74٪ للنهج القائم على التعلم العميق حيث تمت زيادة FPS من 30.30 باستخدام وحدة المعالجة المركزية كخلفية إلى 51.43 باستخدام GPU كخلفية. يوصى باستخدام التعلم العميق في مرحلة تحديد الوجه حيث وجد أنه أكثر دقة وأسرع من Haar Cascaded.

قُسمت الرسالة إلى خمسة فصول وفيما يلي موجزاً عن محتويات كل فصل :

الفصل الأول: يتناول المقدمة والمشكلة التي تركز حولها هذه الدراسة والهدف منها .

الفصل الثاني : يستعرض الابحاث السابقة في مجال التعرف على تعبيرات الوجه عن طريق الآلة والنتائج التي توصلت لها هذه الابحاث ويعرض شرحا مبسطا للمفاهيم التي تم استخدامها في هذه الأطروحة.

الفصل الثالث : ويقوم الفصل الثالث بشرح نقطة البحث و الوسائل المستخدمة في التجارب .يتناول ايضا شرح تفصيلي لقاعدة البيانات المستخدمة وكيف تم تجهيزها من اجل الحصول على ادق النتائج.يسنعرض هذا الفصل ايضا كيفية اختيار قيم اهم المتغيرات الخاصة بالخوارزميات المستخدمة من اجل الوصول لافضل اداء وكفاءة.

الفصل الرابع : يستعرض الفصل الرابع كيفية تقييم اداء نموذج التعلم العميق الذي تم بناؤه عن طريق استخدام بعض المنحنيات و المقاييس الاحصائية .

الفصل الخامس : وهو الفصل الاخير الذي يضم ملخص لنتائج البحث وبعض المقترحات المستقبلية لتحسين كفاءة هذا النظام.

وقد ذُيِّلت الرسالة بالمراجع