

الكاميرات الرقمية المدمجة والاستعداد للتعرف على الوجوه

(صوب والتقط الصورة)

حيدر طالب مهدي الأحمر

الجامعة العراقية - كلية الإدارة والاقتصاد

السماني عبد المطلب أحمد

جامعة النيلين - كلية علوم الحاسوب وتقانة المعلومات

Haeder.ahmar@gmail.com profsamani@gmail.com

الخلاصة

اجتمعت تكنولوجيا الكاميرا الرقمية "المدمجة المتمثلة بـ (صوب والتقط الصورة)" وغير المكلفة مع تكنولوجيا الشبكات التواصل الاجتماعية لإعطاء العامة الدافع لاستخدام تقنية التعرف على الوجوه. يتوقع المستخدمون منها الكثير؛ فهم يريدون التقاط الصور، وتصوير الفيديو، والتحميل. ويكون التعرف على أصدقائهم، والأسرة، والمعارف - قل عددهم أو أكثر - تلقائياً. وعلى الرغم من البساطة الواضحة للمشكلة، فإن التعرف على الوجه في هذا السياق أمراً صعباً. وبصفة عامة، فإن معدلات الفشل الشائعة في هذا النطاق هي من 4 إلى 8 من أصل 10. وعلى النقيض من ذلك، تنخفض معدلات الخطأ إلى ما يقرب من 1 في الـ 1000 من الصور التي يتم التحكم فيها بشكل جيد. ولتعزيز التقدم في التعرف على الوجه والشخص، تقدم هذه الورقة البحثية تحدي التعرف على الوجه (صوب والتقط الصورة) ويشمل التحدي 9369 صورة ثابتة لـ 292 شخصاً متساويين فيما يتعلق بالبعد عن الكاميرا، وأجهزة الاستشعار البديلة، العرض الأمامي مقابل العرض غير الأمامي، وتغير المواقع. وهناك أيضاً 2810 مقطع فيديو لـ 264 شخصاً؛ مجموعة فرعية من الـ 292 شخصاً. وتقدم نتائج التحقق لخوارزميات خط الأساس العامة وخوارزمية تجارية لثلاث حالات: مقارنة الصور الثابتة إلى الصور الثابتة، ومقاطع الفيديو لمقاطع الفيديو، والصور الثابتة لمقاطع الفيديو.

الكلمات الدالة:- صوب والتقط الصورة، معدل التحقق، بيسيرغ للتعرف على الأنماط، وقاعدة بيانات بانكا.

Compact Digital Cameras and GetReadyforFaceRecognition(SelectAnd Capture Photo)

HaederTalibMahdi Al-Samaniabdulmuttalib Ahmed

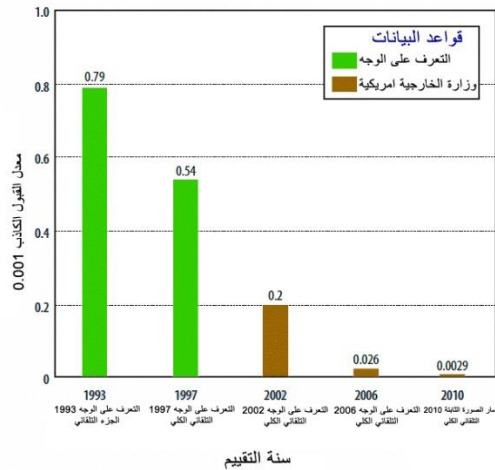
Abstract:

Integrated camera technology ("point-and-shoot") and inexpensive network technology combined social communication to give the public the impetus to use face recognition technology. Users expect a lot; they want to take pictures, shoot video, and download, and their friends, family, and knowledge - fewer or more - automatically be identified. Despite the obvious simplicity of the problem, face recognition in this context is difficult. In general, the common failure rates in this range are from 4 to 8 out of 10. By contrast, error rates fall to nearly 1 in 1000 of well-controlled images. To enhance progress in facial and personality recognition, this paper presents the challenge of face recognition ("point-and-shoot"). The challenge includes 9369 fixed images of 292 equal people with regard to distance from the camera, alternative sensors, front view vs. non-front view, and location change. There are also 2,810 videos for 264 people: a subset of the 292 people. The verification results provide general baseline algorithms and a commercial algorithm for three situations: comparing fixed images to still images, video clips to videos, and still images of videos.

Keywords:-point and shoot, Rate of verification, Pittsburgh PatternRecognition, BANCA Databases.

١ . المقدمة

صعوبة التعرف على الوجه تلقائياً تزداد بشكل كبير عندما لا يكون هناك قيود على ظروف التصوير . في ظل الظروف الأكثر تحكماً، والمعروفة بأنها صور الوجه الأمامية الملتقطة في ملف عرض الصور بالهاتف المحمول أو بيانات التصوير الجنائي التعريفي، ذكر التقييم البارومتري المتعدد ٢٠١٠ أن معدل التحقق لـ ٠,٩٩٧ بمعدل قبول كاذب من ١ في ١٠٠٠ (باتريك جي ،آخرون، ٢٠١٠)، كما هو مبين في الشكل رقم (١). عندما تكون ظروف الإضاءة غير مقيدة ويتم الحصول على صور الوجه الأمامي مع كاميرا ذات عدسة منعكسة رقمية مفردة تحت ظروف الإضاءة الطبيعية في الأماكن المغلقة والهواء الطلق، ينخفض معدل التحقق ذو الصلة إلى ٠,٨٠ (حسن صاحب زادة، ٢٠١١). هذا هو الأداء العام للثلاثة أقسام في مجموعة البيانات الجيدة والسيئة والقيحة. مشكلة تحدي الوجوه المعينة في البرية تحتوي على صور من المشاهير والنجوم تم جمعها من شبكة الإنترنت (تمارا بيرغ ،آخرون، ٢٠٠٧) . تم الحصول على هذه الصور في الأصل من قبل المصورين الصحفيين وتنظم قبل النشر على شبكة الإنترنت. حتى الآن أفضل أداء لمعدل التحقق لـ ٠,٥٤ بمعدل قبول كاذب = ٠,٠٠١ (Belhumeur, ٢٠١٢) .



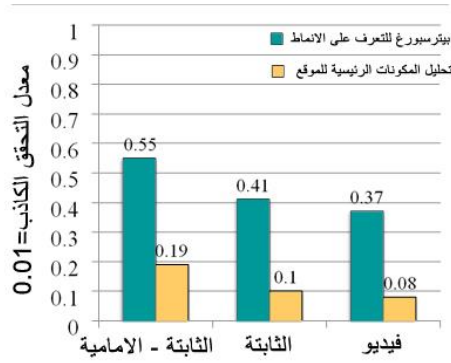
الشكل رقم (١) يبين نسبة معدل التحقق

ينخفض الأداء تدريجياً عندما يتحول من سيناريوهات متحكم بها إلى ظروف غير متحكم بها تصويب والتقاط الصورة.

واليوم يتم التقاط معظم صور الوجوه المكتسبة في جميع أنحاء العالم من قبل الهواة الذين يستخدمون كاميرات التصوير المدمجة والهواتف المحمولة. وتشمل الصور الناتجة تعقيدات نادرة في السيناريوهات السابقة، بما في ذلك ضعف التدرج، الطمس، زيادة وقلة التعرض، وأدوات الضغط. هذا على الرغم من عدد الميجابكسل تصل إلى من ١٢ إلى ٢٢ في بعض هذه الكاميرات.

١. لجذب الانتباه إلى هذه القضايا، نحن نقدم تحدي التعرف على الوجه (صوب والتقط الصورة). الشريط الأخير في الشكل (١) والذي يظهر معدل التحقق على تحدي التعرف على الوجه (صوب والتقط الصورة) الصور الثابتة باستخدام نسخة لمجموعة تطوير البرمجيات ٥,٢,٢ من خوارزمية وضعت من قبل بيتسبرغ للتعرف على الأنماط^١.

^١ تم استخدام مجموعة تطوير البرمجيات لبيتسبرغ للتعرف على الأنماط. في التجارب لأنها كانت متاحة بموجب ترخيص استخدام الحكومة الأمريكية.



الشكل رقم (٢) يبين معدلات التحقق من الصور الثابتة الامامية

الشكل رقم (٢) يبين معدلات التحقق من الصور الثابتة الامامية في تحدي التعرف على الوجه (صوب والنقط الصورة) (يسار)، والصور الامامية والجبهة الثابتة في تحدي التعرف على الوجه (صوب والنقط الصورة) (الوسط)، ومقارنات مقاطع الفيديو إلى الفيديو في تحدي التعرف على الوجه (صوب والنقط الصورة). يتم الإبلاغ عن الأداء ل مجموعة تطوير البرمجيات بيتسبرغ للتعرف على الأنماط. وخوارزمية تحليل المكون الرئيس للمنطقة المحلية (حسن صاحب زادة، ٢٠١١). المنطقة المحلية تحليل المكونات الرئيسية هو خط أساس مفتوح المصدر. لاحظ أنه نظرا لصعوبة المشكلة، يتم توفير معدلات التحقق بمعدل قبول كاذب لـ ٠,٠١، بدلا من معدل القبول الكاذبة التقليدي لـ ٠,٠٠١.

يتضمن تحدي التعرف على الوجه (صوب والنقط الصورة) بواسطة الصور الثابتة ومقاطع الفيديو. ويخلص الشكل (٢) أداء التحقق على الصور الثابتة الامامية، ومزيج من العرض الامامي وغير الامامي، ومقارنات مقاطع الفيديو إلى الفيديو. واعترافا بالصعوبة المتزايدة لتحدي التعرف على الوجه (صوب والنقط الصورة)، يتم الإبلاغ عن معدلات التحقق في الشكل رقم (٢) بمعدل قبول كاذب = ٠,٠١ بدلا من ٠,٠٠١. يوفر الإدراج الثابت والفيديو فرصة قوية لاستكشاف القيمة النسبية لكل منهما.

٢. نظرة عامة على تحدي التعرف على الوجه (صوب والنقط الصورة)

هناك ٩٣٧٦ صورة من ٢٩٣ شخصاً في القسم الثابت من تحدي التعرف على الوجه (صوب والنقط الصورة). تم تنفيذ مجموعة من الصور وفقا لتصميم تجريبي التي غيرت بشكل منهجي الاستشعار، والمكان، والوضعية والمسافة من الكاميرا. ويشمل تحدي التعرف على الوجه (صوب والنقط الصورة) أيضا ٢٨٠٢ مقطع فيديو من ٢٦٥ شخصاً عبر تنفيذ إجراءات بسيطة. هؤلاء الناس هم مجموعة فرعية من الـ ٢٩٣ شخصاً في قسم الصورة الثابتة من تحدي التعرف على الوجه (صوب والنقط الصورة). ويسهل هذا التصميم التحليل الإحصائي لكيفية تأثير العوامل التي تم تعدادها فقط على التعرف على الوجه في الصور الثابتة ومقاطع الفيديو.



الشكل رقم (٣) أمثلة من الصور

الشكل رقم (٣) عبارة عن أمثلة من الصور في تحدي التعرف على الوجه (صوب والتقط الصورة) التي اتخذت في أربع جلسات. لاحظ أن المواقع كانت متنوعة بين الجلسات، في حين أن الحساس، والمسافة إلى الكاميرا ووضع الوقوف قد تنوعت داخل الجلسات.



الشكل رقم (٤) صور الوجه المقطوعة

الشكل رقم (٤) صور الوجه المقطوعة المستخرجة من الصور الثابتة في تحدي التعرف على الوجه (صوب والتقط الصورة). هذه الصور تظهر بعض المضاعفات التي تنشأ في صور صوب والتقط، الإضاءة، ضبابية الحركة وضعف التركيز.

ولتوفير مزيد من المعلومات الأساسية عن عملية جمع البيانات، يعرض الشكل رقم (٣) صوراً من أربع جلسات مختلفة، بينما يبين الشكل رقم (٤) عينة من الوجوه المكتشفة والمقطعة. ويبين الشكل رقم (٣) التباين في الموقع، وأجهزة الاستشعار، وبعد المسافة إلى الكاميرا، والوضعية (أمامية مقابل غير أمامية). في حين يظهر الشكل رقم (٣) الصور المصغرة، والصور الفعلية تصل إلى ما يقرب من ٣٠٠٠ إلى ٤٠٠٠ بكسل: الحجم الدقيق يختلف عن طريق أجهزة الاستشعار. متوسط حجم الوجه يعب عنه عن طريق البيكسل بين العينين يساوي ٦٨؛ والحدود الرباعية هي ١٥ و ٥٤ و ٦٨ و ١١٦ و ٤٧٦ (بما في ذلك الحد الأدنى والحد الأقصى). توضح المقاربات في الشكل رقم (٤) بعض جوانب هذه البيانات التي تجعلها صعبة.

وتبين الأشكال رقم (٤) و(٥) إطارات من فيديو واحد مأخوذ بأعلى دقة لكاميرا فيديو محمولة باليد (١٢٨٠ × ٧٢٠). ويطلب من الأشخاص القيام بأعمال من أجل الحصول على مجموعة واسعة من العروض.

وهناك حافز آخر هو تثبيط السيناريوهات حيث ينظر الشخص مباشرة إلى كاميرا الفيديو لمدة طويلة، مما يقلل من مهمة التعرف على الصور الأمامية مع اللقطات المتعددة. تم كتابة تحركات مختلفة لمواقع وأيام مختلفة. في جميع الحالات، رغم ذلك، تظهر مقاطع الفيديو الشخص الذي يدخل المشهد بعيداً نسبياً عن الكاميرا، حيث يجري فعلاً ما، ثم يترك مجال الرؤية. عادة ما يكون الشخص قريباً نسبياً من الكاميرا عندما يغادر مجال

الرؤية، ولكنه لا ينظر مباشرة إلى الكاميرا.



الشكل رقم (٥) لقطات من فيديو

الشكل رقم (٥) أربع لقطات من فيديو واحد يظهر الشخص وهو يقوم بفعل ما، في هذه الحالة يقوم الشخص بالتمرين على لعبة التنس. لقطات الإطار تتخفف بمقدار النصف في هذا الشكل لتتناسب بشكل أفضل كشكل من أربع لوحات.

البيانات لهذه المشكلة الصعبة متوافرة عند الطلب من خلال الموقع الآتي (تم حذف الرابط عند مراجعة مخطط الكتابة). موقع على الإنترنت منفصل، (تم حذف الرابط)، يدعم هذا التحدي بالطرق الآتية:

أولاً، لديه صلاحية التنزيل حيث يمكن لأي شخصاً تنزيل البرامج والبيانات الوصفية المرتبطة بهذا التحدي. ويشمل هذا البرنامج خوارزميات خط الأساس الموصوفة في أدناه وكذلك رمز لتوليد مصفوفات التشابه ومنحنيات خصائص المستقبل التشغيلية المرتبطة بها. يتم تنظيم الموقع ولديه صلاحيات للباحثين للتسجيل كمساهمين نشطين وتحميل النتائج بطريقة تنسق تتبع التقدم المحرز في تحدي التعرف على الوجه (صوب والتقط الصورة).

٣. العمل ذات الصلة

وكان تقييم فيريه (سيد رزيقي، ٢٠٠٠) لأول جهد كبير في التعرف على الوجه لتوزيع البيانات المشتركة الذي وضع جنباً إلى جنب مع البروتوكول القياسي المعمول بها. ومنذ ذلك الحين ساهمت مجموعة متنوعة من مجموعات البيانات والمسابقات والتقييمات ومشاكل التحدي في مجال التعرف على الوجه. هنا نسلط الضوء على عدد قليل منها.

وقد جمعت قاعدة بيانات تجربة تحسين عملية التعرف على الوجه بجامعة كارنيجي ميلون (أر كروس، ٢٠٠٤) بطريقة تدعم الاستكشافات التجريبية الممتازة للتفاعلات المتحكم بها بين الإضاءة والوضعية.

قاعدة بيانات تجربة العمليات المتعددة لتحسين التعرف على الوجه (غروس، ٢٠١٠) التي تم إصدارها في عام ٢٠١٠ تشمل ٣٣٧ شخصاً، مثل سابقتها، ويدعم التحليل التجريبي بعينات كثيفة والاختلافات المتحكم بها في الإضاءة والوضعية. التحدي الكبير للتعرف على الوجه (P. J. Phillips, ٢٠٠٥) كانت التجربة الرابعة بمطابقة صور الأماكن المغلقة المتحكم بها مع صور الأماكن الداخلية والخارجية غير الخاضعة للتحكم بالإضاءة، وشكلت تحدياً جديداً كبيراً.

تم إصدار مشروع قاعدة بيانات الوجه متعدد الوسائط (XM٢VTS) ومشروع قاعدة بيانات (BANCA) (al, ٢٠٠٣) مع بروتوكولات تقييم ذات صلة، ونظمت مسابقات في كل منهما (م. حاموز، ٢٠٠٠) (كاي ميسير، ٢٠٠٤) يمثل مشروع بيو سيكور الأوروبي جهداً منسقاً كبيراً للنهوض بالقياسات الحيوية متعددة الوسائط، بما في ذلك الوجه (جوليت، ٢٠٠٩) وأكدت حملة تقييم بيو سيكور DS٢ (التحكم بالدخول) المرتبطة بها (N. Poh and et, ٢٠٠٩) الاندماج جنباً إلى جنب مع استخدام تدابير الجودة الحيوية.

وأعرب جيداً عن الرغبة في الابتعاد عن سيناريوهات الحصول على الصور المتحكم بها بشكل جيد في مجموعة بيانات الوجوه المحددة في البرية (تمارا بيرغ، آخرون، ٢٠٠٧). اثنين من الجوانب البارزة من الوجوه المحددة في البرية هي التحول إلى صور الفرصة، للوجوه المحددة في البرية على شبكة الإنترنت، جنباً إلى جنب مع موقع الويب المنسق جيداً والمحدث لإبراز نتائج الأداء الحالي. ويزداد كشف الوجه أيضاً بصعوبة أكبر في السيناريوهات الأقل تحكماً، وأن مسابقة اكتشاف الوجه في مسابقة مجموعات البيانات الصعبة (احمد البرقوقي، ٢٠١١) جمعت العديد من المجموعات في جهد مشترك.

لمزيد من المعلومات عن نهج التعرف على الوجوه بمقاطع الفيديو، هناك استطلاع حديث ممتاز (مينغ دو، وآخرون، ٢٠١١). تم إظهار مجموعات بيانات مفتوحة، على سبيل المثال، مجموعة بيانات وجوه يوتيوب التي تتكون من ٣٤٢٥ مقطع فيديو من ١٥٩٥ شخصاً مختلفاً (L. Wolf, ٢٠١١) جمعت بروح "الوجوه

المحددة" في البرية، وتبني إطارا مشتركا من عشرة أضعاف، وليس بالقياسات الحيوية. وهناك مشكلة أخرى سابقة بتحدي الفيديو هو قسم تحدي الفيديو من التحدي الحيوي الكبير المتعدد وفي مثل هذا التحدي، يشارك الناس الذين يقومون بأنشطة.

٤. بروتوكول

لإنشاء أساس للمقارنة، وتمشيا مع البروتوكول المستخدم في مشاكل التحدي السابقة (باتريك جي، آخرون، ٢٠١٠) (حسن صاحب زادة، ٢٠١١)، يجب على خوارزميات التعرف على الوجه الثابتة حساب درجة التشابه لجميع أزواج من الصور التي تم الحصول عليها عن طريق مطابقة الصور في مجموعة الهدف إلى الصور في مجموعة الاستعلام. ويتم تنظيم قسم الفيديو من التحدي بالطريقة نفسها خوارزميات مقارنة باثنين من مقاطع الفيديو. تؤدي المقارنة بين مقطعي فيديو إلى نتيجة تشابه واحدة. وفي جميع الحالات، تصبح مصفوفة التشابه الناتجة أساسا للتحليل اللاحق، مما يمكن من التعبير عن الأداء من حيث منحى خصائص تشغيل المستقبل.

وهناك تقييد آخر في البروتوكول وهو أمر مهم لتجنب النتائج المفرطة التفاؤل والمضللة. يجب أن تتغير درجة التشابه $S(q,t)$ التي تم إرجاعها إلى زوج الاستعلام (q,t) استجابة للتغيرات في مجموعة الاستعلام Q أو مجموعة الهدف T التي استخلصت من q و t . هناك على الأقل طريقتان شائعتان لانتهاك هذا الحظر التي تعزز الأداء في التجارب:

• التدريب على الصور، أو الناس، في (T) أو (Q) هو انتهاك بروتوكول تحدي التعرف على الوجه (صوب والتقط الصورة).

• ضبط درجات التشابه على أساس المقارنات مع الصور الأخرى في أي من T أو Q هو انتهاك بروتوكول تحدي التعرف على الوجه (صوب والتقط الصورة).

وسيوذي أي من هاتين الممارستين أو كليهما إلى تحسين معدلات التحقق. ومع ذلك، فإن القيام بذلك يأتي على حساب العمومية ويصل إلى معالجة مسألة أكثر محدودية: "ما مدى أداء الخوارزمية عند ضبطها مسبقا لمجموعة مغلقة من الأشخاص المعروفين؟"

٥. الهدف من جمع البيانات والتصميم

هدفان يحفزان التعرف على الوجه (صوب والتقط الصورة). الأول هو إنشاء مجموعة قواعد بيانات وتحدي المشكلة التي من شأنها تشجيع المجتمع التعرف على الوجه لتطوير خوارزميات أفضل والتغلب على العديد من التحديات المرتبطة ببيانات صوب والتقط. تحقيقا لهذه الغاية تم الاحتفاظ بمجموعة بيانات صغيرة بما فيه الكفاية لأنظمة البحوث مع التركيز على تطوير خوارزمية أنواع رئيسية من الاختلاف: الموقع: تم التقاط الصور في تسعة مواقع، داخل المباني وخارجها. تم التقاط مقاطع فيديو في ستة مواقع (داخل المباني وخارجها).

الوضعية: كل من ينظر إلى الكاميرا وخارجها إلى الجانب.

المسافة: الناس على حد سواء القريب والبعيد عن الكاميرا.

المستشعر: خمس كاميرات تصوير ثابتة، وخمس كاميرات فيديو محمولة، وكاميرا فيديو تحكم واحدة.

الفيديو: العديد من إطارات الفيديو على النقيض من واحد ثابت

والهدف الثاني هو دعم التحليل الإحصائي القوي. وإعادة النظر في أسئلة مثل: "هل المسافة من الكاميرا إلى الشخص مهمة بالمقارنة مع اختيار الكاميرا؟" أو إجابات ذات دلالة إحصائية تعتمد على خطة جمع البيانات التي تضمن أخذ العينات المناسبة عبر العوامل وكذلك التوازن بين العوامل. ولدعم هذا الهدف، قبل جمع

الصور، قامت المؤسسات الرئيسية المساهمة في تحدي التعرف على الوجه (صوب والتقط الصورة) بوضع خطة لجمع البيانات. رسمياً، وهي عبارة عن طريقة جمع تشبه تصميم قطعة أرض متعددة الانقسام ممتدة على طول الوقت (دانيل، 1999) وجمعت الصور لتحقيق عينة جيدة من العوامل المذكورة أعلاه (في بعض الحالات لم يتم الشخص بزيارة مواقع كافية لتحقيق التوازن بهذه الطريقة وفي هذه الحالة تم اختيار 8 أو 12 صورة من موقع معين).

5,1. استهداف الصور الثابتة ومجموعات طلبات البحث

يقارن بروتوكول التقييم لقسم الصورة الثابتة من تحدي التعرف على الوجه (صوب والتقط الصورة) جميع الصور في مجموعة الاستعلام إلى الصور كافة في مجموعة الهدف. لذلك، يتم ربط تخصيص الصور إلى مجموعات الهدف والاستعلام بإحكام مع خطة جمع البيانات. يشارك جميع الأشخاص / المواد في عدد متساو من الصور بحيث لا يوجد شخصاً واحد له تأثير أكبر على النتائج من أي شخصاً آخر. على وجه التحديد، كل شخصاً يشارك في 4 صور من كل من 8 مواقع تم اختيارها عشوائياً إلى مجموعات الهدف والاستعلام. وعلاوة على ذلك، يتم تنظيم مجموعات الهدف والاستعلام بطريقة تحول دون المقارنات في اليوم نفسه المقارنات في اليوم نفسه معروفة جيداً لتكون سهلة ويؤدي وجودها في مجموعة بيانات إلى تقويض قيمتها.

الموقع يؤثر على الأداء لأن المواقع مختلفة توفر ظروف التصوير المختلفة. اختيار الصورة أيضاً يتوازن فيما يتعلق بالموقع، مع ما يقرب من العدد نفسه من الصور المكتسبة في كل موقع. ومع ذلك، المقارنات بين الصور التي تم جمعها في الموقع نفسه ولكن في أيام مختلفة نادرة، لذلك يفضل أزواج الموقع نفسه في إنشاء الهدف ومجموعات الاستعلام.

يمكن تفسير بناء مجموعة الهدف ومجموعة الاستعلام بناء من حيث الكتل. كل كتلة تحتوي عادة على 4 صور جمعت من شخصاً واحد في مكان واحد. ويوضح الشكل 3 هذا الهيكل الكتل لشخصاً واحد، وأربعة مواقع مختلفة، القرب مقابل المسافة البعيدة، واثنين من أجهزة الاستشعار. داخل كل كتلة، ويتم اختيار أربع صور عشوائية من تلك المتاحة مثل تلك المسافة والوضعية المتوازنة: صورة واحدة من كل أربعة أحوال قريب / أمامي، قريب / غير أمامي، بعيد / أمامي، وبعيد / غير أمامي. إذا كانت هناك صور غير كافية للشخص والمكان لتحقيق التوازن في هذه الأزياء ثم يتم إسقاط جميع الصور لهذا الشخص والموقع. وقد تم اختيار عشوائية الاستشعار لصالح التمثيل المتساوي من أجهزة الاستشعار عبر العوامل الثلاثة الأخرى.

الجدول رقم (1) يلخص بيانات الصورة الثابتة لتحدي التعرف على الوجه (صوب والتقط الصورة) حجم صغير بما فيه الكفاية أن معظم مطوري الخوارزمية يمكن بسهولة تجربته مع خوارزميات معقدة، ولكن الثابتة توافر عدد كافٍ من نقاط المطابقة لدعم التحليل الإحصائي القوي.

جدول رقم (1) ملخص بيانات الصور الثابتة لتحدي التعرف على الوجه (صوب والتقط الصورة).

عدد من الموضوعات	293
إجمالي الصور	9376
الصور لكل موضوع	32
مطابقة النتائج لكل موضوع	256
مجموع النتائج	21,977,344 (4688 × 4688)
إجمالي نتائج المطابقة	75,008 (265 لكل موضوع)
عدد المواقع	9
مطابقة النتائج لنفس الموقع	8096 (10,8%)

٥,٢. استهداف مقاطع الفيديو ومجموعات طلبات البحث

تم جمع عدد أقل من مقاطع الفيديو، ثم أجريت بعض التغييرات لتحقيق أفضل استفادة من مقاطع الفيديو المتاحة. على سبيل المثال، بدلا من وجود مجموعات استهداف واستعلام مفصولة، للفيديو هناك مجموعة واحدة تعمل ككل من الاستعلام ومجموعة الهدف. وبطبيعة الحال، فإن استخدام المجموعة نفسها يتضمن التأهل الذي لا يقارن أي فيديو من أي وقت مضى بالمقارنة مع نفسه.

وثمة تمييز آخر هو أن كل عمل تم تصويره من قبل كامرتين واحدة ذات جودة عالية، ١٩٢٠ × ١٠٨٠ بكسل، وأخرى باناسونيك على حامل ثلاثي واحد من خمسة كاميرات الفيديو المحمولة البديلة. تعمل بيانات باناسونيك المعتمدة على حامل ثلاثي بمثابة السيطرة. وتتراوح دقة الكاميرات المحمولة من ٦٤٠ × ٤٨٠ حتى ١٢٨٠ × ٧٢٠. وتتوفر صور التحكم للمقارنة مع صور الكاميرا الثابتة، ولكن لا ينبغي مقارنتها عموما بالفيديو المحمول، حيث يمكن بعد ذلك مقارنة أزواج من مقاطع الفيديو التي اتخذت في الوقت نفسه.

يلخص الجدول رقم (٢) بيانات فيديو تحدي التعرف على الوجه (صوب والنقط الصورة). يتم تقسيم المعلومات إلى حد كبير بين مقاطع الفيديو المحمولة والمتحكم بها، مما يعكس النقطة التحذيرية المذكورة أعلاه أن لكل فيديو محمول هناك فيديو تحكم مصاحب اتخذ من الشخص نفسه في الوقت نفسه وهو يفعل الشيء نفسه. وينبغي عموما تشغيل التجارب إما على الفيديو المحمول أو المتحكم به: وليس اعلى فيديو يجمع بينهم.

جدول رقم (٢) ملخص بيانات تحدي مقاطع الفيديو التعرف على الوجه (صوب والنقط الصورة).

٢٦٥	عدد من الموضوعات
٢٨٠٢	إجمالي مقاطع الفيديو
١٤٠١	إجمالي فيديو التحكم
١٤٠١	إجمالي مقاطع الفيديو المحمولة
٤ إلى ٧	التحكم في مقاطع الفيديو حسب الموضوع
٤ إلى ٧	مقاطع الفيديو المحمولة حسب الموضوع
٦	عدد المواقع

٥,٣. الهدف الثابت مقابل فيديو الاستعلام

وبما أن تحدي التعرف على الوجه (صوب والنقط الصورة) يتضمن الصور الثابتة والفيديو من الناس انفسهم، بل هو مجموعة بيانات ممتازة للبدء في استكشاف أداء الاعتراف بين الطرائق: الثابت إلى الفيديو. على سبيل المثال، شخصاً ما هو معروف مسبقاً من قبل صورة ثابتة واحدة في وقت لاحق ينظر، أو يزعم هو ينظر، في مقطع فيديو

للحصول على مثل هذه التجربة، مجموعة الاستعلام يتكون من ١٤٠١ من مقاطع الفيديو المحمولة (أو المتحكم فيها بدلا من ذلك) ومجموعة الهدف تتكون من مجموعة الهدف للصورة الثابتة. وبموجب هذا الاختبار، يجب أن تكون خوارزمية الاعتراف قادرة على إرجاع درجة تشابه واحدة عند تقديم صورة ثابتة واحدة ومقطع فيديو.

٦. دعم البيانات والبرمجيات

تصف الأقسام السابقة من هذه الورقة الأهداف والبيانات والبروتوكولات الخاصة بتحدي التعرف على الوجه (صوب والنقط الصورة). يصف هذا القسم بيانات وبرامج إضافية اختيارية ولكنها تهدف إلى مساعدة الباحثين. ويأتي الدعم الأكثر أهمية في شكل خطوط أساس للمقارنات. يتم توفير خوارزميات خط الأساس المفتوح المصدر، تحليل الفوج الخطي التمييزية وتحليل المكون الرئيس للمنطقة المحلية، لجميع الإصدارات الثلاثة من

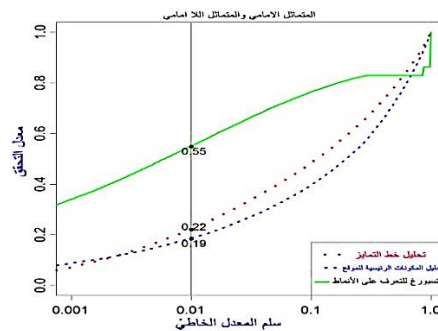
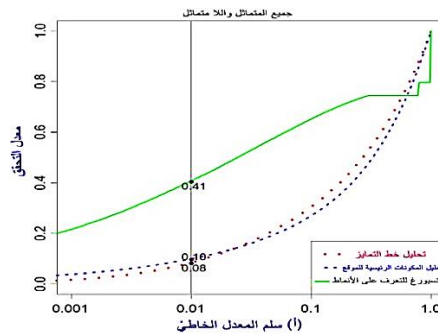
التحدي: الثابتة إلى الثابتة، والفيديو إلى الفيديو، والثابتة للفيديو. يتم تقديم نتائج الكشف عن الوجه المقدمة لكل من الصور الثابتة والفيديو عبر بيتسبرغ للتعرف على الأنماط.. بالنسبة للفيديو، يتم توفير نتائج لكل إطار. كما هو موضح في أدناه، هناك أخطاء وجوه لم يتم الكشف عنها في كل إطار. إتاحة نتائج الكشف عن الوجه يسمح للباحثين الذين ليس لديهم إمكانية الوصول إلى خوارزميات الكشف عن الوجه لتطوير خوارزميات التعرف على الوجوه على بيانات الفيديو.

٦,١ أساسيات مطابقة الصور الثابتة

تُعرض نتائج التحقق لثلاثة خوارزميات. اثنان من خوارزميات خط الأساس مفتوحة المصدر، تحليل الفوج الخطي التمييزية (جوناثون فليب، ٢٠١٢) وتحليل المكونات الرئيسية للمنطقة المحلية (حسن صاحب زادة، ٢٠١١)، وتتوفر من خلال شبكة الإنترنت. تم تطوير الخوارزمية الثالثة من قبل بيتسبرغ للتعرف على الأنماط. تم الحصول على النتائج المعروضة باستخدام مجموعة تطوير البرمجيات ٥,٢,٢.

٦,١,١ اكتشاف الوجه وتحديده

يجب أن تتضمن النتائج المقدمة على تحدي التعرف على الوجه (صوب والتقط الصورة) اكتشاف الوجه وتحديده يتم اعتباره عنصر مهم من الخوارزمية التي تم وضعها للتعرف على الوجه. وبعبارة أخرى، النتائج حيث توجد وجوه، وبشكل أكثر تحديدا العينين، إيجادها يدويا ليس ذات فائدة عملية كبيرة. ومع ذلك، فإن إحداثيات العين الناتجة عن خوارزمية التعرف على نمط بيتسبرغ متوفرة مع مجموعة البيانات. ويمكن أن تتضمن الدراسات الكاملة التي تعرض النتائج على تحدي التعرف على الوجه (صوب والتقط الصورة) نتائج تكميلية حيث يتم الكشف عن الجانب الآلي باستخدام الجانب الإحداثيات المقدمة عند استخدام هذا الإحداثيات المقدمة بوضوح.



الشكل رقم (٦) أداء التحقق من الصورة الثابتة لخطين

الشكل رقم (٦) أداء التحقق من الصورة الثابتة لخطين أساسيين وخوارزمية تجارية على تحدي التعرف على الوجه (صوب والتقط الصورة):

(أ) مشكلة تحدي كاملة.

(ب) صورة الوجه الأمامي جزء فقط من تحدي التعرف على الوجه (صوب والتقط الصورة).

لمعايرة صعوبة الكشف عن الوجه وتحديده على البيانات من تحدي التعرف على الوجه (صوب والتقط الصورة)، قمنا بمقارنة كاشف الوجه المقدمة كجزء من خوارزمية التعرف على نمط بيتسبيرغ مع معيار الكمبيوتر مفتوح الرؤية لكاشف الوجه على أساس العمل من فيولا وجونز (Jones, ٢٠٠٤). كاشف بيتسبيرغ التعرف على الأنماط فشل في تحديد موقع الوجوه في ٧,٠٪ من صور تحدي التعرف على الوجه (صوب والتقط الصورة). في المقابل، فشل كاشف الرؤية الكمبيوتر المفتوحة للكشف عن الوجه على ٢٧,٦٪ من الصور. قمنا بتوسيع خوارزمية رؤية الكمبيوتر المفتوحة لتشمل معلومات عن الكتفين وكذلك الوجوه. ناتج الخوارزمية خفض معدل الخطأ إلى ٩,٠٪، والتي تعد ليست نتيجة جيدة كما الخوارزمية التجارية ولكن أقرب بكثير. يتم تضمين شفرة خوارزمية الكشف عن الوجه الموسعة مع البرامج الموزعة كجزء من تحدي التعرف على الوجه (صوب والتقط الصورة).

٦,١,٢ مقدمة عن خوارزمية PCA

تعتبر خوارزمية تحليل المكونات الأساسية PCA Principal component analysis أحد أهم التقنيات الناجحة التي تم استخدامها في مجال التعرف على الصور وفي مجال ضغط الصور.

تصنف PCA على أنها أحد الطرق الاحصائية في هذا المجال.

الهدف الاساسي في خوارزمية PCA يكمن في تقليص الأبعاد الكبيرة في فضاء المعطيات إلى فضاءات بأبعاد أصغر. وعادة ما تكون الفضاءات الجديدة عبارة عن فضاءات الملامح (اي تحوي الملامح الاساسية والمهمة للمعطيات في فضاءاتها الاصلية)، وبالتالي وعبر تقليص الأبعاد هذا نكون قد وصفنا المعطيات بشكل اقتصادي أكثر مما يساعدنا فيما بعد.

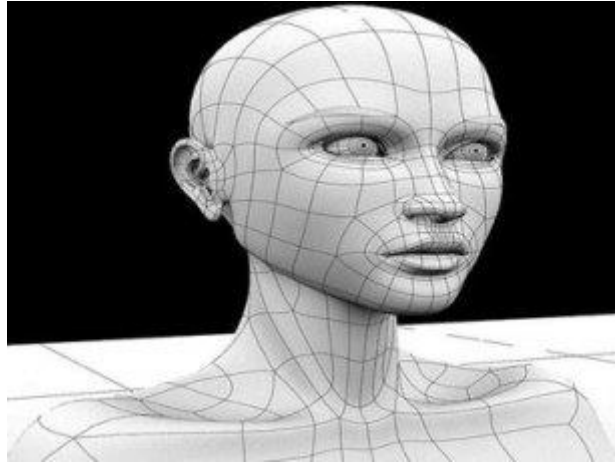
٦,١,٣ نظم التعرف ومشاكل الأبعاد الكبيرة

عادة ما تظهر المشاكل في نظم التعرف على الوجوه عندما تتعامل النظم مع فضاءات كبيرة الأبعاد (على سبيل المثال التعامل مع الصور).

يمكن اجراء العديد من التحسينات وذلك عبر مطابقة ونقل المعطيات الموجودة إلى معطيات بفضاء اقل ابعادا. وبالتالي نكون قد قمنا بإنقاص الأبعاد dimensionality reduction من الفضاء الأصلي ذو الأبعاد الكبيرة إلى الفضاء الجديد ذو الأبعاد الأصغر.

وليكن مثلا لدينا الشعاع التالي:

وذلك ضمن فضاء مؤلف من N بعد , نقوم عبر تقليص الأبعاد بالانتقال إلى شعاع آخر $\mathbf{Y} = [y_1, y_2, \dots, y_K]^T$ اي إلى فضاء مؤلف من K بعد بحيث $K < N$.



الشكل رقم (٧) فضاء أبعاد الوجه

من حيث المبدأ ، فإن إنقاص الأبعاد يؤدي بدوره إلى فقدان وضياح بالمعلومات، ولكن الهدف من خوارزمية PCA هو إنقاص أبعاد المعطيات مع الاحتفاظ بأكبر قدر ممكن ومهم من المعلومات الموجودة في المعطيات الأصلية.

هذه العملية مكافئة للاحتفاظ بأكبر قدر ممكن من التنوعات والتغيرات الموجودة ضمن المعطيات الاصلية.

ضمن هذا السياق ، تقوم PCA بحساب التحويل الخطي T linear transformation الذي بدوره يقوم بمقابلة المعطيات الموجودة ضمن فضاء الأبعاد الأعلى إلى المعلومات الموافقة لها ضمن الفضاء الجزئي ذو الأبعاد الأقل ، كما هو موضع ادناها:

$$\begin{cases} y_1 = t_{11}x_1 + t_{12}x_2 + \dots + t_{1N}x_N \\ y_2 = t_{21}x_1 + t_{22}x_2 + \dots + t_{2N}x_N \\ \dots \\ y_K = t_{K1}x_1 + t_{K2}x_2 + \dots + t_{KN}x_N \end{cases}$$

أو بعبارة أخرى

$$y = Tx$$

حيث ان

$$T = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1N} \\ t_{21} & t_{22} & \dots & t_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ t_{K1} & t_{K2} & \dots & t_{KN} \end{bmatrix}$$

إن التحويل الأمثل T هو عبارة عن التحويل الذي تكون فيه القيمة $\|x - y\|$ اصغرية وتبعاً لنظرية PCA فإنه يمكن تعريف الفضاء ذو الأبعاد الأقل الأمثل عبر الاستقادة من أفضل الأشعة الذاتية eigenvectors الخاصة بمصفوفة التباين التابعة للمعطيات covariance matrix of the data. (نعني بذلك : فإن الاشعة الذاتية الموافقة للقيم الذاتية الأكبر largest eigenvalues لمصفوفة التباين ، ويشار إليها ايضا بالمكونات الأساسية "principal components".

لنفترض بأن I_1, I_2, \dots, I_M عبارة عن مجموعة من M شعاع، وكل شعاع فيها له الأبعاد التالية $N \times 1$ (Schwarz tiger's Weblog، ٢٠١٣).

٦,١,٤ أداء التحقق

ويبين الشكل رقم (٦) منحنيات مميزة لتشغيل المستقبل للتحكم التعرف على الوجه (صوب والتقط الصورة). يتم حساب خوارزميات خط الأساس باستخدام صور التدريب المعينة التي تلتزم بالبروتوكول أعلاه (جوناثون فليب، ٢٠١٢) كل من الصور والناس (حسن صاحب زادة، ٢٠١١) هي منفصلة عن الهدف ومجموعات الاستعلام. يتم توفير مجموعة الصور التدريبية كجزء من تحدي التعرف على الوجه (صوب والتقط الصورة). وجاءت خوارزمية التعرف على نمط بيتسبرغ لنا من قبل المدربين.

لقد حاولنا تجنب التقسيم على الوجه (صوب والتقط الصورة) إلى مشاكل فرعية؛ يؤدي ذلك إلى إضعاف قيمة النقطة المرجعية الواضحة. ومع ذلك، التعرف على الوجه الأمامي هو تخصص متطور جدا من المشكلة وشعرنا أنه من المفيد أن تظهر في الشكل B٦ النتائج للصور الأمامية فقط. نلاحظ أيضا في الشكل ٦ أن معدلات التحقق لاحظت في معدل قبول كاذب = ٠,٠١. وهذا يمثل عدم التقيد بالمعايير النسبية لمعدل قبول كاذب أكثر صرامة = ٠,٠٠١ القياسي المستخدم في الدراسات السابقة (حسن صاحب زادة، ٢٠١١). في رأينا هذا التحول مناسب نظرا لصعوبة تحدي التعرف على الوجه (صوب والتقط الصورة).

الأدوات واضحة في منحنيات التعرف على نمط بيتسبرغ بدءا من اليمين معدل قبول كاذب = ٠,٠١. ويعود سبب ذلك إلى وضع خوارزمية درجات التشابه لقيم ثابتة خاصة، مما يفترض أن يكون هناك بعض القرارات الداخلية التي تتخذها الخوارزمية. في حين أن الأدوات بارزة وتستدعي التفسير، فهي ليست ذات أهمية عملية. ومن الناحية العملية، فإن معدلات التحقق بمعدلات قبول كاذبة أكبر من ٠,١ لها معنى ضئيل حيث أنه من الصعب تصور أن نظم ميدانية واسعة النطاق ذات قيمة تعمل بمعدلات قبول كاذبة عالية.

(أ) مقاطع الفيديو للكاميرا المحمولة.

(ب) مقاطع الفيديو للكاميرا ذات تحكم عالية الدقة (١٩٢٠ × ١٠٨٠).

٦,٢ نتائج خط الأساس للفيديو

وتنشأ عدة مستويات إضافية من التعقيد في معالجة قسم تحدي الفيديو. على سبيل المثال، كشف الوجه وتحديد في مقاطع الفيديو للكاميرا المحمولة هو أكثر صعوبة من قسم الصورة الثابتة من تحدي التعرف على الوجه (صوب والتقط الصورة). وهناك تعقيد آخر يتعلق بكيفية تقليل المقارنة بين مقطعي فيديو، كل منهما يستمر ما يقرب من ٥ إلى ١٠ ثوان في الطول، إلى درجة تشابه واحدة.

٦,٢,١ اكتشاف الوجه وتحديد في الفيديو

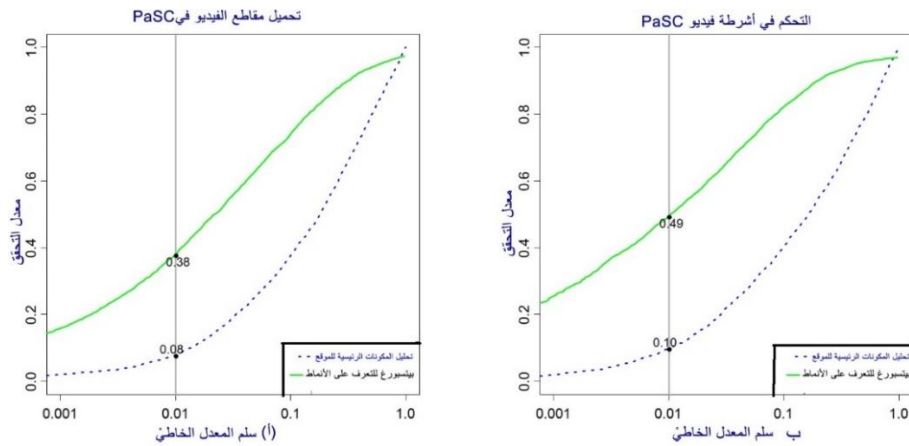
كشف الوجه وتحديد في مقاطع الفيديو مهمة صعبة. في المجموع، هناك ٣٣٤٨٧٩ و ٣٢٨٩٦٧ لقطة من الفيديو في ١٤٠١ فيديو كاميرا التحكم و ١٤٠١ فيديو الكاميرا المحمولة على التوالي. ويعود الفرق جزئيا إلى التباين المتواضع في طول الفيديو. عند تشغيل هذه الأطر، فإن جهاز كشف الوجه المفتوح في جهاز الكمبيوتر الذي يعتمد على فيولا وجونز (Jones, ٢٠٠٤) كثيرا ما يفشل في العثور على الوجوه في أقل بكثير من نصف إجمالي إطارات الفيديو. كشف الوجه وتحديد قدراته من برنامج التعرف على نمط بيتسبرغ يفعل أفضل بكثير.

في ٦٠ فقط من ١٤٠١ من كاميرا التحكم و ٣٤ من أصل ١٤٠١ فيديو الكاميرا المحمولة، خوارزمية بيتسبرغ للتعرف على الأنماط تفشل في العثور على وجه في أي إطارات. وبخلاف ذلك، يتم الإبلاغ عن مواقع الوجه

والتقديرات التقديرية التقريبية للإطار الواحد على الأقل، وعادة ما تكون العديد من الإطارات. لتلخيص الأداء، يتم الكشف عن وجوه في ١٢١،٠١٦ و ١٢٧،٦٢١ إطارات من مقاطع فيديو كاميرا التحكم والكاميرا المحمولة على التوالي. ويمثل هذا الكشف في ثلث جميع الإطارات تقريبا.

لمساعدة المجموعات البحثية التي تركز على جوانب الاعتراف من التحدي، سيتم توفير آلة الكشف عن الوجه التي تم إنشاؤها في ملف قيم مفصولة بفواصل التي تحتوي على موقع الوجه،

وحجم والوضعية التقريبية. تصدر الخوارزميات نتائج على قسم الفيديو من تحدي التعرف على الوجه (صوب) والتقط الصورة) باستخدام الكشف عن الوجه الخاصة بهم وسيتم الإبلاغ عن التحديد بشكل مختلف عن تلك التي تستخدم النتائج كشف الوجه المقدمة.



الشكل رقم (٨) أداء التحقق الفيديو على تحدي التعرف على الوجه (صوب والتقط الصورة):

٦,٢,٢ مقارنات الفيديو من إطار إلى إطار

وبين الشكل رقم (٨) نتائج مقارنة مقاطع الفيديو لمقاطع الفيديو باستخدام امتداد تحليل المكون الرئيس للمنطقة المحلية وخوارزميات بيتسبرغ للتعرف على الأنماط على وجه التحديد، يتضمن التمديد مقارنة جميع الإطارات مع الوجوه المكتشفة في فيديو واحد لجميع الإطارات مع الوجوه المكتشفة في فيديو آخر. ثم يتم فرز مجموعة كبيرة من نقاط التشابه الناتجة ويتم تحديد النقطة الواحدة على أساس الترتيب في قائمة فرزها.

الاكتشاف المبكر في الأداء هذا للمقارنة العديد من الإطارات يليها الاختيار على أساس رتبة من النتيجة هو أن الدرجة القصوى ليست دائما الأفضل. على وجه الخصوص، قمنا بفحص منحنيات تشغيل مستقبلات مميزة لخيارات مختلفة، بما في ذلك الدرجة القصوى، والنتيجة ١٠٪ في قائمة عشرات فرزها، ومتوسط النتيجة، وما إلى ذلك في حالة تحليل المكون الرئيس للمنطقة المحلية واختيار الدرجة القصوى أسفر عن أفضل خصائص تشغيل المتلقي: هذا هو ما هو مبين في الشكل ٧. بالنسبة لخوارزمية بيتسبرغ التعرف على الأنماط المستخدمة بهذه الطريقة، فقد أسفرت النتيجة القصوى عن خصائص تشغيل المتلقي أكثر من اختيار النسبة المئوية ٩٠: حيث تعرض خاصية تشغيل المستقبل.

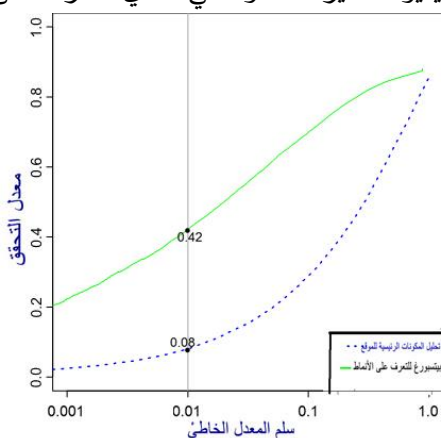
وأحد الأسباب التي حذفت من نتائج تحليل الفوارق الخطية التمييزية من الشكل رقم (٨) هو أن خصائص تشغيل المستقبل المرتبطة بخيارات الرتب المختلفة في درجات التشابه كانت غير متوقعة وشعرنا بأن هناك حاجة إلى مزيد من العمل لفهم السبب بشكل أفضل؟ بشكل عام، ما هو هذا الاعتماد على رتبة الترتيب بالتشابه؟ الذي

يخبرنا هو أنه لا يزال هناك المزيد من العمل الذي يتعين القيام به في تحديد كيفية إجراء مثل هذه الاختيارات وكيف تنتشر عواقب الانتقاء في المطابقة وعلى القدر نفسه من الأهمية توزيعات عدم التطابق.

٦,٣. الهدف الثابت مقابل نتائج طلب بحث الفيديو

ويعرض الشكل رقم (٩) النتائج حيث تتكون مجموعة الاستعلام من مقاطع الفيديو للكاميرا المحمولة، ومجموعة الهدف هو من الصورة الأمامية الثابتة إلى اختبار الصورة الأمامية الثابتة المذكورة أعلاه. بطريقة مشابهة لخوارزمية خط الأساس للفيديو، تتم مقارنة الصورة المستهدفة الواحدة مع كل إطارات الفيديو التي تم اكتشاف وجه لها. يتم فرز النتائج الناتجة وتنشأ المتغيرات خوارزمية مختلفة على أساس اتخاذ درجة "الحد الأقصى" و '١٠٪' في درجة الترتيب، وما إلى ذلك ومن المثير للاهتمام أن نلاحظ أن معدل التحقق لهذه التجربة أعلى إلى حد ما من تجربة مقاطع الفيديو إلى الفيديو.

الهدف الثابت مقابل اطلب بحث فيديو الكاميرا المحمولة في تحدي التعرف على الوجه (صوب والتقط الصورة).



الشكل رقم (٩) نتائج مجموعة الاستعلام

٧. الخاتمة

تتوافر صور الوجه وبيانات مقاطع الفيديو والبيانات الوصفية المرتبطة بها من أجل تحدي التعرف على الوجه (صوب والتقط الصورة) عند الطلب. برنامج الدعم يشمل خوارزميات تحليل الفوج الخطي التمييزي تحليل المكونات الأساسية للمنطقة المحلية ورمز التهديد سيتم تحميلها من خلال شبكة الإنترنت. ونحن سوف نبكر ونحافظ على الموقع المنسق حيث مجموعات العمل على تحدي التعرف على الوجه (صوب والتقط الصورة) قد تقدم النتائج. يمكن للمشاركين تقديم نتائج الأداء والجودة. ويظهر موقع الوجوه المبوبة في الموقع البري بوضوح قيمة نقطة اتصال مشتركة حيث تتوافر معلومات محدثة.

تم دعم مشاكل التحدي التي شيدت بشكل جيد وتعزيز البحوث على مدى سنوات عديدة، ويحدونا الأمل في أن تحدي التعرف على الوجه (صوب والتقط الصورة) سيكون حافزا للتحسينات الموضوعية والضرورية في التعرف من صور صوب والتقط. صور صوب والتقط تعرض واجهات جديدة لمواجهة الاعتراف الذي تتاولناه في هذا التحدي. من وجهة نظر متعاضدة، فإن خطة جمع البيانات لتحليل التعرف على الوجه (صوب والتقط الصورة) هي أن الدراسات التي أجريت على تحدي التعرف على الوجه (صوب والتقط الصورة) يمكن أن تعالج مباشرة الأهمية النسبية للعوامل مثل أجهزة الاستشعار، والموقع، الوضعية، المسافة إلى الكاميرا، وما إلى ذلك من أن إدراج الفيديو يفتح مجالاً حاسماً آخر للبحوث، بما في ذلك الأسئلة الأساسية فيما يتعلق فقط بحجم الفائدة الإضافية، إن وجدت، ويرتبط مع امتلاك الفيديو المقصود بصلبة لصورة واحدة ثابتة جيدة .

المصادر العربية:

- احمد البرقوقي واخرون. (٢٠١١). كشف الوجه والعين على مجموعات البيانات الصلبة. المؤتمر الدولي المشترك بشأن القياسات الحيوية . (ص ١-١١). واشنطن: معهد مهندسي الكهرباء والإلكترونيات.
- أر كروس واخرون. (٢٠٠٤). التعرف على الوجه عبر الشكل والإضاءة. تأليف S. Baker R. Gross، Design and Analysis of Experiments (ص ١٩٣ - ٢١٦). نيويورك: دار سبرينغر.
- أنجيلا م.، فوس ،دانيال. (١٩٩٩). تصميم وتحليل التجارب. تأليف A. M. Dean and D. Voss، Design and Analysis of Experiments (ص ١٠١-١١٥). لندن: دار سبرينغر للنشر العلمي.
- باتريك جي ،آخرون. (٢٠١٠). تقييم خوارزميات التعرف علي الوجه في ظل صورة ذات ابعاد ثنائية. غايثرسبيرغ، الولايات المتحدة: المعهد الوطني للمعايير والتكنولوجيا الشعار.
- تمارا بيرغ ،آخرون. (٢٠٠٧). التعرف على الوجه في البيئات غير المقيدة. امهيرست: جامعه ماساتشوستس.
- جوناثون فليب واخرون. (٢٠١٢). الدراسات الاولية علي الجيد والسيء والقيبح. Preliminary Studies on the Good, the Bad, and the Ugly Face Recognition Challenge Problem (ص ٩-١٦). غايثرسبيرغ، الولايات المتحدة الأمريكية: المعهد الوطني للمعايير والتكنولوجيا.
- حسن صاحب زادة واخرون. (٢٠١١). بداية للأفضل، والرديء ، والمشكلة التحدي التعرف على الوجه المشوه. تقنيات التعرف على الوجوه (ص ٣٤٦ - ٣٥٣). سانتا بريارة: المعهد الوطني للمعايير والتكنولوجيا.
- رالف غروس. (٢٠١٠). الصورة والرؤية عبر الكمبيوتر. تأليف Ralph Gross، Image and Vision Computing (ص ٨٠٧-٨١٣). بوتوروث: بوتوروث-هاينمان نيوتن، الولايات المتحدة الأمريكية.
- سيد رزيقي واخرون. (٢٠٠٠). منهجية التقييم لخوارزميات التعرف على الوجه. تحليل الأنماط وذكاء آلات (ص ١٠٩٠ - ١١٠٤). واشنطن: معهد مهندسي الكهرباء والإلكترونيات.
- كاي ميسير واخرون. (٢٠٠٤). اختبار مصادقة الوجه على قاعدة بيانات banka. المؤتمر الدولي للتعرف علي الأنماط (ص ١-١٠). الولايات المتحدة الامريكية: معهد مهندسي الكهرباء والإلكترونيات.
- م. حاموز واخرون. (٢٠٠٠). مقارنة نتائج التحقق من الوجه على قاعدة بيانات XM٢VTFS. المؤتمر الدولي المعني بالتعرف علي الأنماط (ص ٢٠٣-٢٢٠). برشلونة: معهد مهندسي الكهرباء والإلكترونيات.
- مينغ دو، وآخرون. (٢٠١١). التتبع والتعرف على الوجه في الفيديو. تأليف مينغ دو، وآخرون، Face Tracking and Recognition in Video (ص ٣٢٣ - ٣٥١). لندن.

المصادر الأجنبية:

- al, E. B.-B. (2003). The BANCA database and evaluation protocol. *Audio- and Video-based Biometric Person Authentication* (pp. 625–638). Guildford: Guildford UK.
- Belhumeur, T. B. (2012). Tom-vs-pete classifiers and identity-preserving alignment for face verification. *In Proceedings of the British Machine Vision Conference*, (pp. 129.1–129.11.). LONDON.
- Jones, P. V. (2004, May 5-6). Robust real-time face detection. *International Journal of Computer Vision*, pp. 57(2):137–154.
- L. Wolf, T. H. (2011). Face recognition in unconstrained videos with matched background similarity. *In Computer Vision and Pattern Recognition (CVPR)* (pp. 529–534). WA.USA: IeeeExplore.
- P. J. Phillips, P. J. (2005). Worek. Overview of the face recognition grand challenge. *In IEEE Conference on Computer Vision and Pattern Recognition*. WA.USA: IEEE Explore.