



الجمهورية العربية السورية

جامعة دمشق

كلية الهندسة الميكانيكية والكهربائية

قسم هندسة الحواسيب والأتمتة

ماجستير هندسة وبرمجة الروبوت

رسالة مقدمة لنيل درجة الماجستير في هندسة وبرمجة الروبوت في قسم هندسة الحواسيب والأتمتة بكلية
الهندسة الميكانيكية والكهربائية بعنوان:

تحسين تقانات التخاطب بين الإنسان والآلة باستخدام التعلم العميق

**Improving Human–Machine Communication Technologies Using
Deep Learning**

إعداد الطالبة: وفاء سلامة

بإشراف الدكتور: سمير كرمان

2024م

الملخص

نظام التعرف على الوجوه هو عبارة عن نظام تحقق يعمل بسرعة وفعالية عالية، فهو أسرع وأكثر ملاءمة من تقنيات المقاييس الحيوية الأخرى، مثل بصمات الأصابع أو مسح شبكية العين. ويعتمد نظام التعرف على الوجه في عمله على عدد أقل من نقاط الاتصال مقارنة بإدخال كلمات المرور أو أرقام التعريف الشخصية. وهو يتوافق مع المصادقة متعددة العوامل لإتاحة خطوة تحقق أمني إضافية. انطلاقاً من هذه الميزات والمخاطر التي تواجه أمن واستقرار الأنظمة الحاسوبية و غيرها في العالم، كان لا بد من العمل على ابتكار طرق آلية جديدة لتحليل الوجوه و الكشف عن ملامحها المختلفة باستخدام الروبوت.

يهدف البحث إلى تقديم خوارزمية و منهجية جديدة من أجل التعرف على الوجوه، وذلك من خلال استخدام مجموعة مختلفة من الخوارزميات التي يتم استخدامها في معالجة الصور، مثل خوارزميات التجزئة segmentation وطريقة زيادة البيانات Data augmentation، وتم استخدام الشبكات العصبية التلافيفية CNN لمعرفة النتائج النهائية الخاصة، وتم اختبار النظام على عدد كبير من الصور كي يتم التأكد من النتائج التي توصل إليها البحث. ثم تم تدريب هذه الخوارزمية واختبارها على 530 صورة وأظهرت النتائج دقة وصلت إلى أكثر من 97%.

الكلمات المفتاحية:

التعرف على الوجوه، معالجة الصورة الرقمية وتجزئتها وتصنيفها، خوارزمية زيادة البيانات، الشبكات العصبية التلافيفية

فهرس المحتويات:

الصفحة	المحتويات
2	الملخص
3	فهرس المحتويات
6	فهرس الأشكال
8	الفصل الأول: المقدمة Introduction
9	١_١ مقدمة:
10	١_٢ الدراسات والأبحاث المرجعية:
12	١_٣ مشكلة البحث:
13	١_٤ أهمية البحث:
13	١_٥ أهداف البحث:
13	١_٦ منهجية البحث:
13	الفصل الثاني: المبادئ النظرية للبحث Theoretical Principles of Research
14	٢_١ آلية عمل نظام التعرف على الوجوه باستخدام الحاسب:
15	٢_١_١ جمع البيانات من أجل إدخال الصور وتدريب الشبكة عليها:
15	٢_١_٢ عملية تقطيع الصورة:
16	٢_١_٣ إدخال الصور المقطعة إلى الشبكات العصبية التلافيفية:
17	٢_١_٤ بنية الشبكة العصبونية
18	٢_١_٤_١ الهيكل الهرمي للشبكات العصبية التلافيفية:
24	٢_١_٥ عملية زيادة البيانات:
25	٢_١_٦ عملية التصنيف للصور التي تم إدخالها للشبكة العصبونية:
26	٢_١_٧ اختبار النظام:
28	الفصل الثالث الخوارزمية المقترحة للنظام:

29	٣_١_خطوات خوارزمية التعرف على الوجوه المقترحة:
29	٣_١_١_تجميع وتحضير البيانات:
31	٣_١_٢_تحميل الشبكة العصبونية العميقة:
32	٣_١_٢_١_بنية الطبقات في الشبكة العصبونية المستخدمة:
33	٣_١_٢_٢_حجم وأبعاد الطبقات في الشبكة:
34	٣_١_٣_تحديد النموذج والتدريب:
34	٣_١_٣_١_استخدام نموذج شبكة المسار المرتفعة (highway Network):
34	٣_١_٣_٢_استخدام نموذج شبكة الانتباه (Attention Network):
35	٣_١_٤_٤_تقديم الصور للنموذج:
35	٣_١_٤_٥_تقييم أداء النموذج:
35	٣_١_٤_٦_تدريب واختبار الخوارزمية المقترحة:
37	٣_١_٤_٦_١_نتائج تطبيق عملية التجزئة على الصور المدخلة:
39	الفصل الرابع التقييم والمقارنة
40	٤_١_التقييم :
41	٤_١_١_حساب ال Precision:
41	٤_١_٢_حساب ال Recall:
42	٤_١_٣_حساب ال F1-score:
43	٤_١_٤_حساب ال Accuracy:
43	٤_١_٥_حساب ال Specifity:
49	٤_٢_المقارنة:
49	٤_٢_١_المقارنة بين الشبكة المدربة والغير مدربة:
51	٤_٣_النتائج والمناقشة:
52	٤_٤_الاستنتاجات:
53	المراجع:

فهرس الأشكال:

الصفحة	الأشكال
15	الشكل (1-2) مخطط آلية التعرف على الوجوه
18	الشكل (2-2) بنية الشبكة العصبية التلافيفية
19	الشكل (3-2) رسم تخطيطي للأثر المعتدل والتطبيع خلال المعالجة الأولية في طبقة الإدخال
20	الشكل (4-2) خريطة تأثير الارتباط والتبييض خلال المعالجة الأولية في طبقة الإدخال
21	الشكل (5-2) طبقة حساب الطي
22	الشكل (6_2) طبقة الحوافز
24	الشكل (7_2) عملية زيادة البيانات
26	الشكل (2 - 8) المخطط التدفقي لعملية التصنيف
27	الشكل (9-2) المخطط التدفقي لعملية الاختبار
29	الشكل (1-3) مخطط خوارزمية التعرف على الوجوه
30	الشكل (2-3) عينات مختلفة في شكل الوجه
30	الشكل (3-3) عينات مختلفة في لون الوجه
31	الشكل (4-3) عينات مختلفة في العمر
31	الشكل (5-3) عينات تحتوي أجسام غريبة
32	الشكل (6-3) بنية طبقات CNN
36	الشكل (7-3) الفرق بين بيانات الاختبار والتدريب
37	الشكل (8-3) الصور المجزأة التي ستدخل عملية التصنيف
38	الشكل (9-3) الصور المجزأة التي ستدخل في عملية التصنيف
42	الشكل (1-4) يوضح الفرق بين الاستدعاء والإحكام
44	الشكل (2-4) مخطط بياني لمعايير تقييم الأداء
45	الشكل (3-4) تابع الخسارة

47	الشكل (4-4) تابع الدقة
48	الشكل (5-4) يوضح confusion matrix
50	الشكل (6-4) مخطط بياني للمعايير لشبكة مدربة وغير مدربة و بدون transfer learning

جدول المصطلحات:

HOG	Histogram Of Oriented
SVM	Support Vector Machine
LBP	Local Binary Patterns
VGG	Visual Geometry Group
ResNet	Residual Network
ASM	Active Shape Model
IRRNet	Image Reconstruction and Recognition
Relu	Rectified Linear Unit
FC	Fully Connected Layer
TP	True Positive
NP	True Negative
FP	False Positive
FN	False Negative
CNN	Convolutional Neural Network

الفصل الأول

المقدمة introduction

١_١_ مقدمة:

نظام التعرف على الوجه Facial Recognition System هي تقنية قادرة وبشكل تلقائي على تحديد أو التحقق من شخص موجود في صورة رقمية أو إطار فيديو مأخوذ من مقطع فيديو.

هنالك العديد من الطرق التي تعمل بها أنظمة التعرف على الوجه إلا أنها تتطوي عموماً على سلسلة من الخطوات التي تعمل على التقاط ومعالجة وتحليل ومطابقة الوجه مع الوجوه المخزنة داخل قواعد البيانات، على الرغم من أن تكنولوجيا التعرف على الوجه لم تصل إلى ذروتها بعد، إلا أنها حققت ما يكفي من تقدم لإيجاد سلسلة من التطبيقات والبرامج، وتنوعت استخداماتها بشكل كبير لتصل إلى حد واسع مستخدم بكثرة حولنا. أدى الاستخدام المكثف لتقنيات مراقبة ومطابقة ملامح الوجوه إلى تعزيز البحث حول تقنية تعمل على تحديد الهوية الفردية في بيئة غير مقيدة، ومع ذلك يعد التعرف على الوجوه مهمة صعبة في بيئة غير مقيدة مع وجود العديد من المشاكل والتحديات التي تؤثر على هذه الأنظمة، لذلك كان لابد من العمل على تصميم برنامج قادر على التخلص منها إلى أقصى حد ممكن والوصول إلى أفضل النتائج.

١_٢_ الدراسات والأبحاث المرجعية:

١_٢_١_ الدراسة المرجعية الأولى[1]:

الكشف الدقيق عن تعابير الوجه عن طريق اختيار عدة معالم وسمات التحويل المحلية

يتألف النظام المقترح من أربع مراحل:

- 1- اكتشاف الوجه باستخدام تجزئة لون البشرة وملائمة الشكل البيضوي وإزالة منطقة العنق.
- 2- تعيين المعالم على الوجه من خلال رسم مربع محيطي على كل ملمح.
- 3 -استخراج السمات باستخدام المسافة الاقليدية و HOG و LBP
- 4-استخدام مصنف SVM شعاع دعم الآلة للتصنيف النهائي للشخص.

السلبيات :

لقد كان أداء النظام ذو دقة جيدة في التعرف على الوجه ولكن مع استخراج العديد من الملامح أدى ذلك إلى ارتفاع الوقت الحسابي مما يجعل النظام غير قادر على التصنيف في الزمن الحقيقي.

الخوارزمية المتبعة:

المسافة الاقليدية + SVM

١_٢_٢_ الدراسة المرجعية الثانية [2]:

التعرف على صور العائلات من خلال التشفير المدرب مسبقا

الطريقة المتبعة: يهدف البحث إلى تحديد ما اذا كانت صورتان مختلفتان للوجه تعودان لشخصين لهما قرابة أم لا وذلك للتحقق من قرابتهما. في البداية يتم استخراج السمات من

الصور باستخدام نموذج FaceNet بعد ذلك يتم حساب تابع الخسارة بين الصورتين التي تعطي المسافة بينهما ثم يتم تحويل المسافة إلى درجة احتمالية تعبر عن نسبة القرابة.

السلبيات: استخدم البحث سمات غير مرغوب بها والتي أدت إلى ظهور مشكلة

Overkill في الشبكات العصبية العميقة.

الخوارزمية المتبعة: VGG-Face-ResNet-50

١_٢_٣_ الدراسة المرجعية الثالثة [3]:

تحديد مواقع سمات الوجه التفاعلية

الطريقة المتبعة: يتم تمثيل الشكل كليا كمجموعة نقاط محددة مسبقا (المعالم) ثم يتم الربط بين هذه النقاط باستخدام كاشف الوجه والملاح البارزة ويدعى كاشف الوجه والملاح البارزة ويدعى ASM والذي يتألف من نموذجين هما:

1- نموذج الشكل: ويستخدم للتعبير عن معلم (أنف، فم، عينان، حاجبان) عن طريق مجموعة من الاحداثيات

2- نموذج الملف الشخصي: الذي يربط كل المكونات السابقة مع بعضها البعض وحساب Vetrbi المسافة بينها وبين جيرانها من خلال خوارزمية

التي تحسب احتمال انتماء كل بكسل معين لملمح الوجه، وفي النهاية الحصول على الشكل التقريبي للشخص

السلبيات: لوحظ أن أداء النظام ينخفض عندما احتوت بعض الصور المدخلة على أنواع مختلفة من الضجيج والتشويش، وبالتالي يجب تحسين الأداء مع مجموعات كبيرة من البيانات

المحتوية على نسبة من الضجيج وتجنب مشكلة عدم الملائمة للصور المختلفة وزيادة نسبة التدريب.

الخوارزمية المتبعة : ASM + Vetrbi

١_٢_٤ _ الدراسة المرجعية الرابعة [4]:

خوارزمية التعلم العميق من أجل التعرف على الوجوه غير الكاملة: شبكة استرداد ذات نسبة منخفضة
الطريقة المتبعة :

يستخدم النموذج LEENET والتي تحتوي على طريقة إكمال المصفوفة لاستعادة وحدات البكسل المفقودة أو التالفة كما تم استخدام شبكة CNN لتقليل الضجيج والتشويش في الصورة، ثم استخدام مصنف شعاع دعم الآلة svm لتحديد الأفراد في الصور التالفة.

السليبيات: إن استخدام منهجية استعادة الصورة يساعد النظام ليعمل بشكل أكثر دقة حتى مع الصور التالفة، لكنه لم يأخذ في الاعتبار اختلاف الوضع وحالة الإضاءة والظلال ووجود الشعر أو نظارات على الوجه.

الخوارزمية المتبعة :

LRRNET + CNN + SVM

١_٣ _ مشكلة البحث:

لا يزال التعرف على الوجوه يواجه العديد من التحديات عند التقاط الصور في ظروف الإضاءة السيئة، حيث تتأثر الصور الملتقطة بما يلي :

1_ بظروف الإضاءة 2_ وتغير الأوضاع 3_ واختلافات التعابير وعلامات الشيخوخة،

بالإضافة إلى 4-الاختلافات الفردية لكل شخص، وبالتالي تؤثر على توفير معلومات دقيقة عن الأفراد.

١_٤_ أهمية البحث:

- 1_ إن التعرف على الوجوه بالكامل يعتبر أدق وأفضل من الطرق القائمة على مكون واحد كالعين أو البصمة أو اليد لأنها تعتمد على عدة مكونات بدلا من مكون واحد.
- 2_ إمكانية التطبيق في العديد من المجالات كالتفاعل بين الإنسان والحاسوب، والتعليم الالكتروني، وتحديد هوية الإنسان، ونظام الأمن، وكشف حالات الغش والتزوير.

١_٥_ أهداف البحث:

الهدف الأساسي من هذا البحث هو تحسين دقة التعرف على الوجوه من خلال التغلب على المشاكل التي عانت منها الدراسات السابقة، بالإضافة إلى تحسين وقت المعالجة والوصول إلى النتائج خلال أقل زمن ممكن.

١_٦_ منهجية البحث:

ستتم هذه الدراسة وفقا لمنهج التحليل والتجريب باستخدام لغة بايثون للتأكد من فعالية وأداء النظام المقترح.

يتكون النظام المقترح من الشبكات العصبية التلافيفية CNN لاستخراج السمات عالية المستوى (منحنيات الوجه وتمييز علامات على الوجه والشكل المميز لمكونات الوجه).

حيث تتعلم تقنية التعلم العميق ملامح الوجه المميزة والدقيقة لكل فرد وتصنف الصور المعينة بناء على هذه السمات.

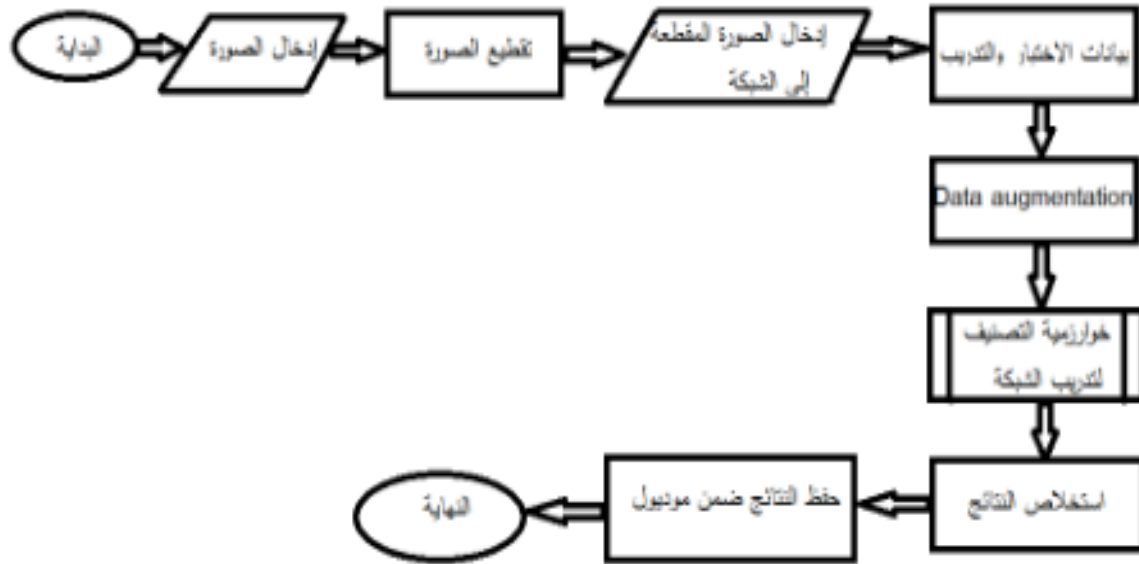
الفصل الثاني

المبادئ النظرية للبحث

Theoretical principles of research

٢_١_آلية عمل نظام التعرف على الوجوه باستخدام الحاسب:

يمثل المخطط الآتي العمليات التي يتم إجراؤها من أجل التعرف على الوجوه باستخدام الحاسب:



الشكل (2-1) مخطط آلية التعرف على الوجوه

٢-١-١_ جمع البيانات من أجل إدخال الصور وتدريب الشبكة على الصور:

تم تجميع Dataset من صور الأشخاص في المرحلة الأولى وذلك من أجل تقسيم الصور بالبداية إلى قسمين القسم الأول من أجل تدريب الشبكة على الصور بعد تجميع الصور يتم إدخالها إلى النظام المقترح من أجل تدريب الشبكات عليها.

٢-١-٢_ عملية تقطيع الصورة:

في هذه المرحلة يتم تقطيع الصور والتي هي مرحلة هامة من مراحل معالجة الصور الرقمية، وهي عملية تجزئة للصورة إلى مناطق مترابطة ومتجانسة وفقا لمعيار محدد كاللون، اتحاد هذه المناطق يجب أن ينتج عنه إعادة تكوين الصورة الأصلية.

التقطيع مرحلة مهمة تسمح باستخراج معلومات نوعية عن الصورة إذ توفر وصف عالي المستوى حيث إن كل منطقة مرتبطة بالمناطق المجاورة لها ضمن شبكة من العقد تمثل فيها كل عقدة منطقة في الصورة وتحمل هذه العقدة بطاقة تحوي معلومات نوعية عن المنطقة كحجمها ولونها وشكلها وتوجهها، أما الأقواس التي تربط العقد فيمكن أن توسم بمعلومات عن العلاقة بين المناطق المتجاورة كأن تكون مثلا منطقة ما محتواه في أخرى أو تكون تحتها أو فوقها إلى غير ذلك. مستوى التعقيد في تكوين الشبكة يختلف تبعا للتقنية المستخدمة في التقطيع [5].

٢-١-٣_ إدخال الصور المقطعة إلى الشبكة العصبونية :

تم بناء الشبكة العصبونية من أجل تدريب الشبكة العصبونية التي تم إنشاؤها وذلك من أجل إدخال الصور التي تم تقطيعها إلى الشبكة العصبونية ليتم بعدها تقسيم الصور إلى قسمين وهما قسم بيانات التدريب وقسم بيانات الاختبار.

١_ بيانات التدريب:

بيانات التدريب هي مجموعة بيانات من الأمثلة المستخدمة للتعلم، وهي تناسب المتغيرات (مثل القيم)، تميل معظم المقاربات التي تبحث من خلال بيانات التدريب عن العلاقات التجريبية إلى مناسبة البيانات، بمعنى أنه يمكنها تحديد العلاقات الواضحة في بيانات التدريب التي لا تتم بشكل عام .

٢_ بيانات التحقق:

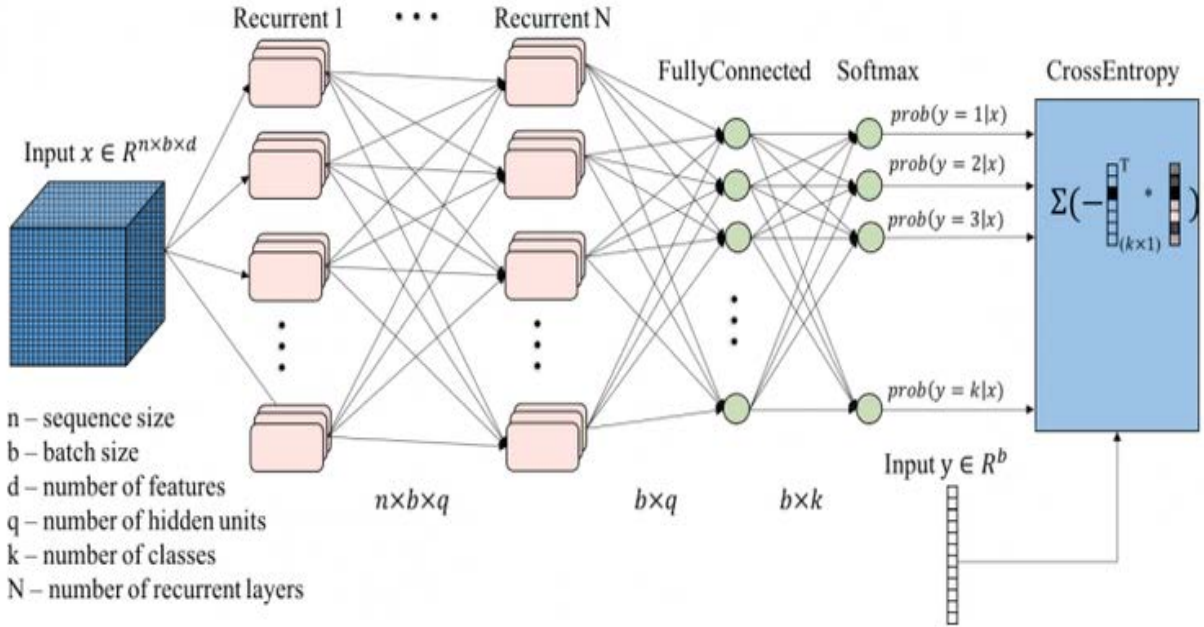
مجموعة بيانات التحقق عبارة عن مجموعة بيانات من الأمثلة المستخدمة لضبط بنية المتغيرات من المصنف. تسمى أحياناً مجموعة التطوير. في الشبكات العصبية الاصطناعية، تكون بنية المتغيرات على سبيل المثال ، عدد الوحدات المخفية. يجب أن تتبع مجموعة الاختبار (كما ذكر أعلاه) نفس التوزيع الاحتمالي مثل مجموعة بيانات التدريب.

٣_ بيانات الاختبار:

مجموعة بيانات الاختبار هي مجموعة بيانات مستقلة عن مجموعة بيانات التدريب، ولكنها تتبع نفس التوزيع الاحتمالي كمجموعة بيانات تدريبية. إذا كان النموذج المناسب لمجموعة بيانات التدريب مناسباً أيضاً لمجموعة بيانات الاختبار، فقد حدث الحد الأدنى من المطابقة الزائدة. عادةً ما يشير التركيب الأفضل لمجموعات البيانات التدريبية بخلاف مجموعة بيانات الاختبار إلى المطابقة الزائدة، وبالتالي فإن مجموعة الاختبار هي مجموعة من الأمثلة المستخدمة فقط لتقييم الأداء (أي التعميم) لمصنف محدد بالكامل [6].

٢_١_٤_ الشبكات العصبية التلافيفية:

يمثل المخطط التالي بنية الشبكات العصبونية التلافيفية:



الشكل (2-2) بنية الشبكة العصبية التلافيفية

٢_١_٤_١_ الهيكل الهرمي للشبكات العصبية التلافيفية:

1. طبقة إدخال البيانات / طبقة الإدخال

2. طبقة تحويل

3. طبقة الحوافز / طبقة ReLU

4. طبقة التجميع

5. طبقة متصلة بالكامل / طبقة FC

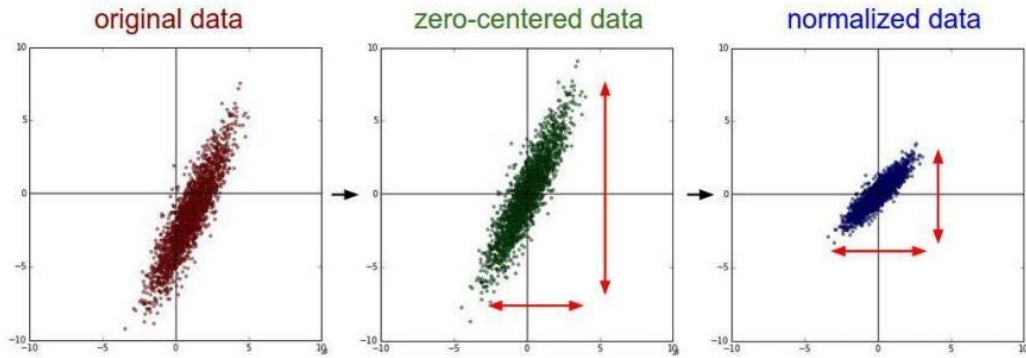
١_طبقة إدخال البيانات:

تتم المعالجة في هذه الطبقة بشكل أساسي للمعالجة المسبقة لبيانات الصورة الأصلية، بما في ذلك :

إزالة المتوسط: مركزة كل بُعد من بيانات الإدخال إلى 0 ، كما هو موضح في الشكل أدناه ، والهدف هو سحب مركز العينة إلى أصل نظام الإحداثيات.

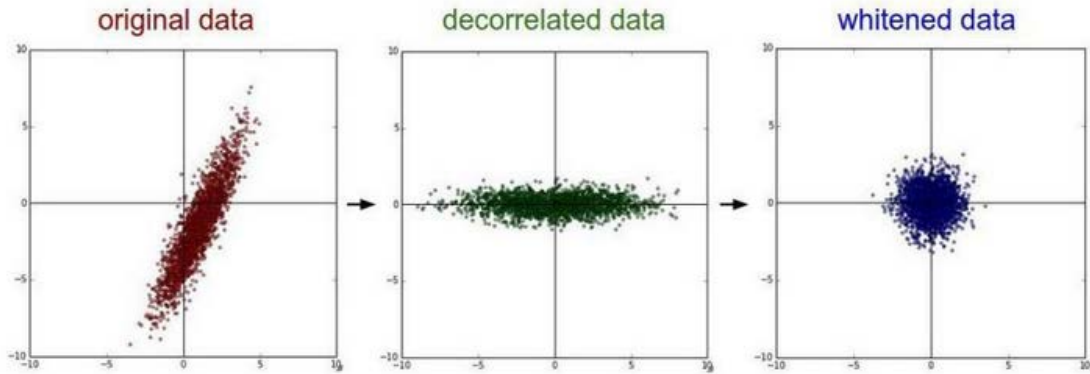
•التطبيع: يتم تطبيع السعة إلى نفس النطاق ، كما هو موضح أدناه ، مما يقلل من التداخل الناجم عن الاختلاف في نطاق قيمة البيانات في كل بُعد. على سبيل المثال، لدينا سمتين من سمات الأبعاد A و B. نطاق A هو من 0 إلى 10، ونطاق B من 0 إلى 10000. إذا كنت تستخدم هاتين الميزتين بشكل مباشر، فإن ذلك يمثل مشكلة. وتتمثل الممارسة الجيدة في التطبيع، أي يتم تغيير بيانات A و B. النطاق من 0 إلى 1.

التبويض: تقليل البعد مع PCA التبويض هو تطبيع السعة على كل محور مميز للبيانات . رسم تخطيطي للأثر المعتدل والتطبيع:



الشكل (2-3) رسم تخطيطي للأثر المعتدل والتطبيع خلال المعالجة الأولية في طبقة الإدخال

خريطة تأثير الارتباط والتبييض :



الشكل (2-4) خريطة تأثير الارتباط والتبييض خلال المعالجة الأولية في طبقة الإدخال

٢_طبقة حساب الانحناء :

هذه الطبقة هي أهم طبقة في الشبكة العصبية التلافيفية ومصدر اسم "الشبكة العصبية التلافيفية".

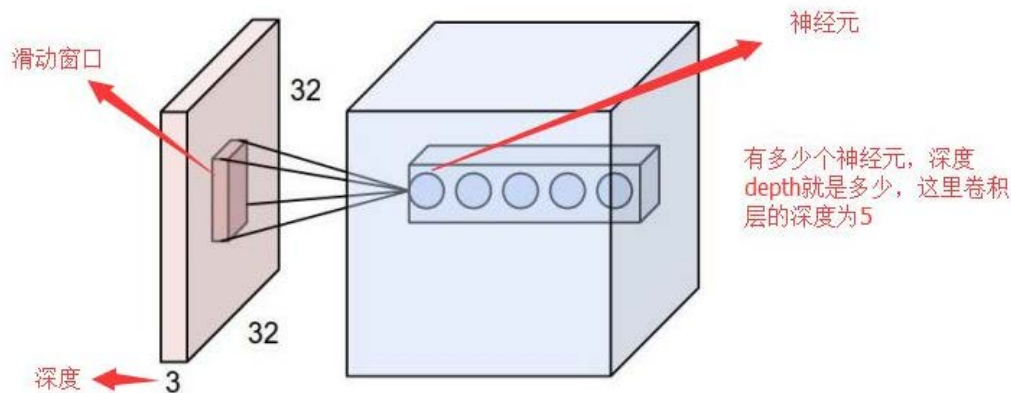
في هذه الطبقة التلافيفية ، هناك عمليتان رئيسيتان:

- ارتباط محلي. يتم التعامل مع كل الخلايا العصبية كمرشح.
- تتفتح النافذة (حقل الاستلام) ويحسب عامل التصفية للبيانات المحلية.
- لندخل أولاً بعض المصطلحات التي واجهتها الطبقة التلافيفية:

•العمق / العمق

•خطوة / خطوة (طول شريحة النافذة)

•قيمة المساحة / بدون مساحة



الشكل (5-2) طبقة حساب الطي

ما هي قيمة الحشو؟ الصورة السابقة هي مثال لتوضيح فكرة الحشو. على سبيل المثال، إذا كانت لدينا صورة بحجم 5×5 (شبكة واحدة تحتوي على بكسل واحد)، وأخذنا نافذة منزلة بحجم 2×2 مع خطوة بحجم 2، فنلاحظ أنه هناك بكسل واحد متبقي لا يمكن معالجته. في هذه الحالة، ماذا نفعل؟ نقوم بإضافة طبقة من الحشو إلى المصفوفة الأصلية لجعل حجمها 6×6 ، بحيث تصبح النافذة قادرة على اجتياز جميع وحدات البكسل. هذا هو دور قيمة الحشو. حساب الطي: في الصورة التالية نلاحظ أن هناك مربعاً رمادياً حول المصفوفة الزرقاء، وهذا يمثل الحشو كما ذكرنا سابقاً. المصفوفة الزرقاء هي صورة الإدخال، والمصفوفة الوردية تمثل الخلايا العصبية للطبقة التلافيفية. يتم تمثيل العصبونات هنا بواسطة

$W1, W0$

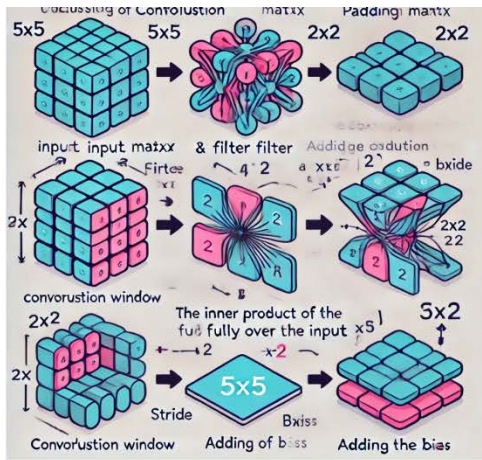
بينما المصفوفة الخضراء هي مصفوفة الإخراج بعد عملية الالتفاف، حيث يتم تعيين حجم الخطوة إلى 2

في عملية الطي، تقوم المصفوفة الزرقاء (صورة الإدخال) بحساب المنتج الداخلي للمصفوفة مع المصفوفة الوردية (المرشح)، ثم يتم إضافة نتائج العمليات الحسابية إلى قيمة الإزاحة كما هو موضح في الشكل أدناه يتم حساب القيمة كالتالي :

$$(2 + (2 - 1) + (2 - 2) + (1 - 2) + (-2) + 1) = 2 - 3 - 3 + 1 = -3$$

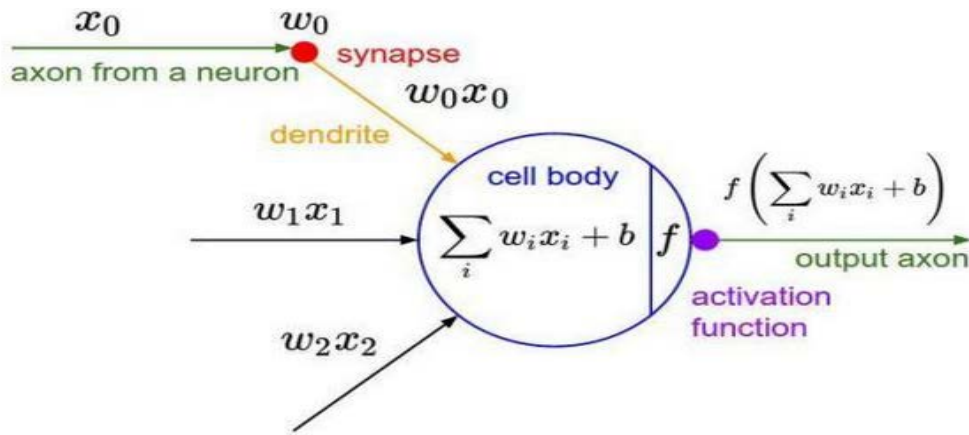
القيمة المحسوبة هنا هي عنصر من مصفوفة الإطار الأخضر

الطبقة التلافيفية في هذه الطبقة يتم تثبيت وزن كل نافذة بيانات اتصال بالخلايا العصبية كل خلية عصبية تركز على ميزة واحدة فقط من الصورة يتم استخدام الخلايا العصبية كمرشحات في معالجة الصور مثل فلتر sobel لاكتشاف الحواف كل مرشح في طبقة الـ 5x5 سيكون له ميزات خاصة بالصورة مثل الحواف الرأسية ،الحواف الأفقية ، الالوان ،والقوام إلخ.. جميع هذه الخلايا العصبية تساهم في مجموعة مستخرج الميزة للصورة بأكملها كما في الشكل



٣_ طبقة الحواف :

التعيين غير الخطي لمخرجات طبقة الالتفاف.



الشكل (2_6) طبقة الحواف

دالة الإثارة التي تستخدمها CNN هي RELU (الوحدة الخطية المصححة) بشكل عام وتتميز بالتقارب السريع و حساب التدرج البسيط .

٤_ **طبقة التجميع (Pooling Layer)** : تم استخدام 2×2 في هذه الطبقة. طبقة التجميع محصورة بين طبقات تلافيفية متتالية بهدف ضغط كمية البيانات والمعلومات وتقليل التجهيز الزائد. باختصار، إذا كان الإدخال عبارة عن صورة، فإن الوظيفة الرئيسية لطبقة التجميع هي ضغط الصورة. عملية التجميع تساعد في تحقيق الثبات في الميزات، مما يعني الحفاظ على الحجم المنتاسق للميزات التي غالباً ما يتم ذكرها في معالجة الصور. تجميع الصورة يعني تغيير حجمها وتقليل التفاصيل غير المفيدة. فيما يتعلق بتقليل الميزة، من المعروف أن الصور تحتوي على العديد من المعلومات والميزات، ولكن بعض هذه المعلومات قد تكون غير مفيدة أو متكررة عند معالجة الصور. عملية التجميع تساهم في إزالة هذه المعلومات الزائدة واستخراج أهم الميزات. كما أن طبقة التجميع تلعب دوراً كبيراً في تقليل الإفراط في التجهيز (Overfitting)، مما يسهل تحسين أداء النموذج. تستخدم طبقة التجميع طريقتين رئيسيتين هما: الحد الأقصى للتجميع (Max Pooling) و التجميع المتوسط (Average Pooling). ولكن الطريقة الأكثر شيوعاً هي الحد الأقصى للتجميع (Max Pooling) في هذه الطريقة، لكل نافذة 2×2 ، يتم تحديد أكبر قيمة كعنصر مكافئ في مصفوفة الإخراج. على سبيل المثال، إذا كانت أكبر قيمة في نافذة 2×2 الأولى لمصفوفة الإدخال هي 6، فإن العنصر الأول في مصفوفة الإخراج سيكون 6، وهكذا .

٥_ **طبقة الاتصال الكاملة (Fully Connected Layer)** :في هذه الطبقة، جميع الخلايا

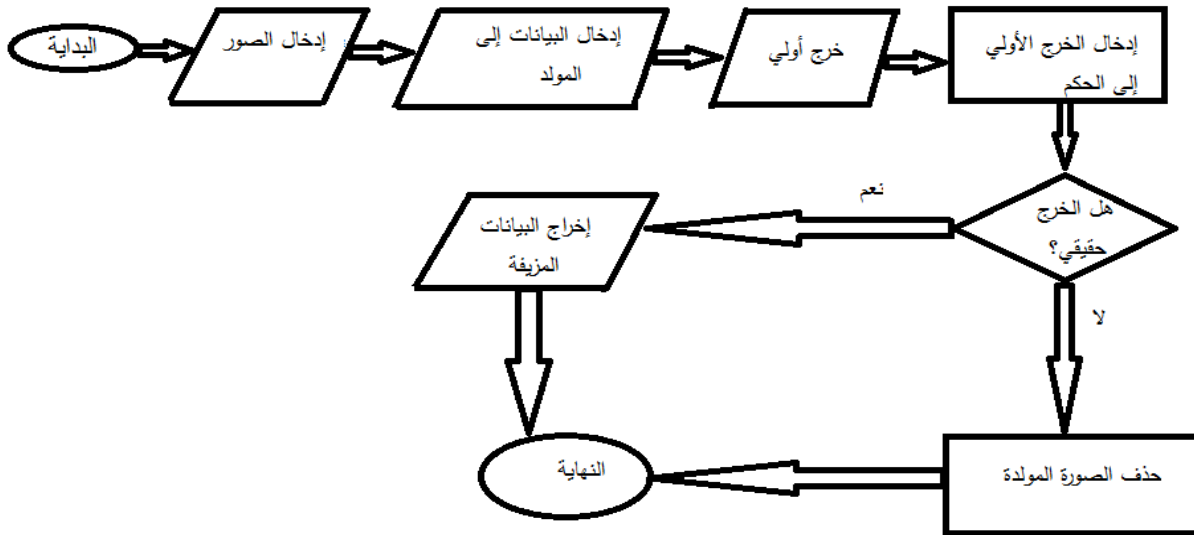
العصبية بين الطبقتين تتصل ببعضها البعض. عادةً ما تكون الطبقة المتصلة بالكامل في نهاية الشبكة العصبية التلافيفية، وتشبه في طريقة عملها طريقة الاتصال العصبي في الشبكات

العصبية التقليدية. هذه الطبقة تهدف إلى تحويل الميزات المكتشفة من الصور إلى تصنيفات نهائية [7]

٢_١_٥ عملية زيادة البيانات Data Augmentation:

هي العملية التي يتم استخدامها من أجل زيادة البيانات حيث يتم استخدام هذه العملية من أجل زيادة التنوع في الصور المدخلة إلى الشبكات العصبونية بهدف الحصول على أفضل دقة ممكنة دون الحاجة إلى جمع البيانات من جديد، تستخدم تقنية زيادة البيانات مثل تقنيات الاقتصاص والحشو والتقليب بشكل أفقي وذلك بشكل كبير من أجل زيادة البيانات تُساعد عملية تعزيز البيانات أيضاً على تلافي مشكلة فرط الملاءمة التي تحدث عند تدريب النماذج باستخدام مجموعة بيانات صغيرة الحجم وذات جودة منخفضة [8] .

ويمكن تمثيل عملية ال Data Augmentation حسب المخطط التالي:



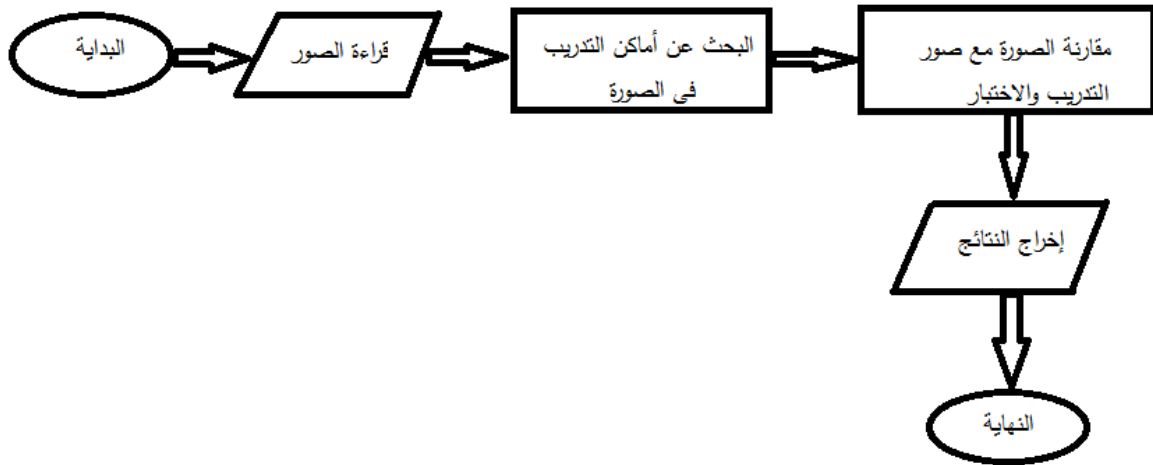
الشكل (7_2) عملية زيادة البيانات

وذلك عبر الخطوات التالية:

- 1- إدخال صور التدريب
 - 2- أولاً يتم الإدخال إلى الجزء الأول من عملية ال Augmentation: حيث يتم في هذا الجزء توليد صور جديدة من الصور التي يتم إدخالها عبر بيانات التدريب باتباع التقنيات التي تم ذكرها مسبقاً
 - 3- يتم إخراج مجموعة من الصور الأولية بعد العملية السابقة
 - 4- إدخال الخرج الأولي إلى شبكة جديدة تدعى الحكم: والتي تقوم بعملية مقارنة الخرج الأولي مع بيانات التدريب
 - 5- إذا كانت البيانات المخرجة حقيقية
 - 6- يتم حفظ الصورة الإضافية
 - 7- وإلا يتم حذف الصور
- وهذه العملية تتكرر على جميع صور التدريب التي تم إدخالها للشبكة بحيث يتم تكرار جميع البيانات المدخلة.

٢_١_٦_ عملية التصنيف للصور التي تم إدخالها للشبكة العصبونية:

في هذه المرحلة يتم تطبيق خوارزميات التصنيف التي يتم تطبيقها على الشبكات العصبونية من أجل تصنيف الصور المخرجة إلى نوعين، ويمكن تمثيل عملية التصنيف حسب المخطط التدفقي التالي:



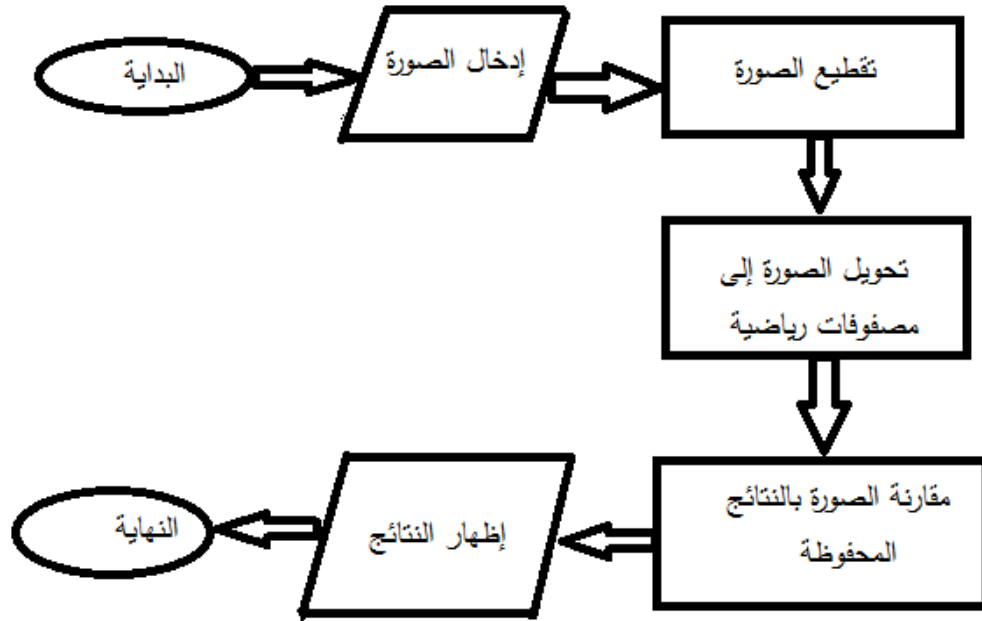
الشكل (2 - 8) المخطط التدفقي لعملية التصنيف

وحسب الخطوات التالية:

- 1- قراءة الصور التي سيتم إدخالها للخوارزمية
 - 2- البحث عن الأماكن المميزة في الصورة من أجل عملية تصنيف الصورة
 - 3- مقارنة الصورة مع صور التدريب والاختبار
 - 4- حفظ النتائج من عملية المقارنة
 - 5- إخراج النتائج
- يتم حفظ النتائج على شكل مصفوفات رياضية ضمن ملف منفصل وذلك من أجل الاحتفاظ بها من أجل عمليات الاختبار حتى لا نضطر لتدريب الشبكة في كل مرة نريد اختبار صورة جديدة وبهذه العملية نكتفي بمقارنة الصورة الجديدة بـ النتائج السابقة للحصول على النتيجة [10].

٢_١_٧- اختبار النظام :

يمكن تمثيل عملية اختبار النظام البرمجي من خلال المخطط التدفقي التالي:



الشكل (2-9) المخطط التدفقي لعملية الاختبار

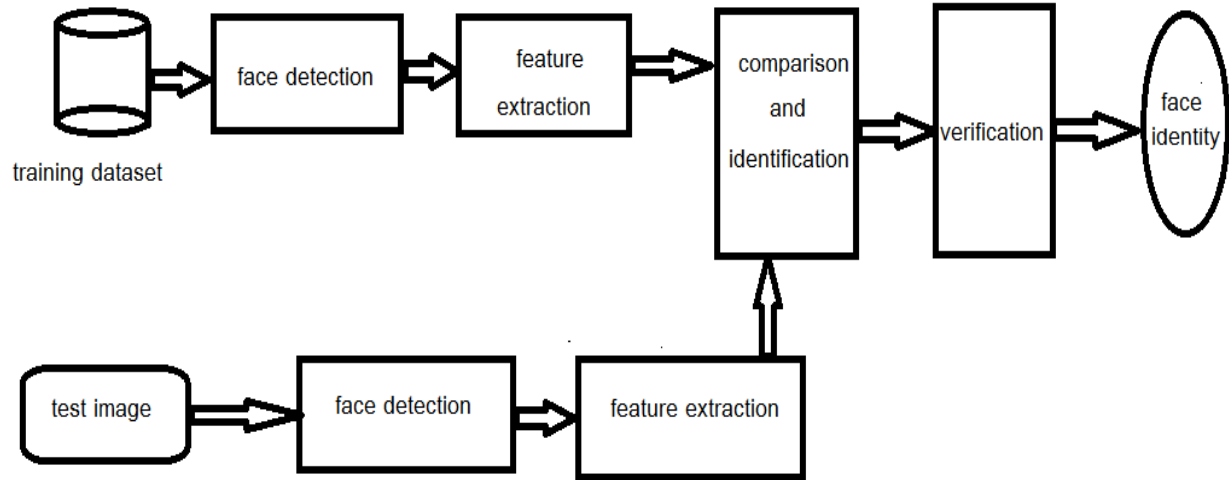
وذلك حسب الخطوات التالية:

- 1- إدخال الصورة المراد الكشف عنها
- 2- تمرير الصورة إلى خوارزميات التقطيع وتحويلها إلى اللون الرمادي
- 3- تحويل الصورة إلى شكل مصفوفات رمادية
- 4- تحويل المصفوفة الناتجة مع نتائج خوارزميات التصنيف
- 5- إظهار النتيجة النهائية

الفصل الثالث

الخوارزمية المقترحة للنظام

٣_١_خطوات خوارزمية التعرف على الوجوه المقترحة:



الشكل (3-1) مخطط خوارزمية التعرف على الوجوه

٣_١_١_تجميع وتحضير البيانات:

تم جمع مجموعة واسعة من الصور الأصلية والصور التي تم تعديلها بمختلف طرق التعديل مثل تعديل الإضاءة والتباين..، حيث تم تأمين قاعدة بيانات مكونة من عدة صور (أكثر من 2000 صورة)

لعدة أشخاص بمختلف الأعمار وظروف إضاءة مختلفة وأوضاع متعددة كل صورة تم حفظها كنوع Jpg.

١_ اختلاف العينات في أشكال الوجه:



الشكل (2-3) عينات مختلفة في شكل الوجه

٢_ اختلاف العينات في لون الوجه:



الشكل (3-3) عينات مختلفة في لون الوجه

٣_ اختلاف العينات في عمر الشخص:



الشكل (3-4) عينات مختلفة في العمر

٤_ وجود أجسام غريبة كالنظارات:

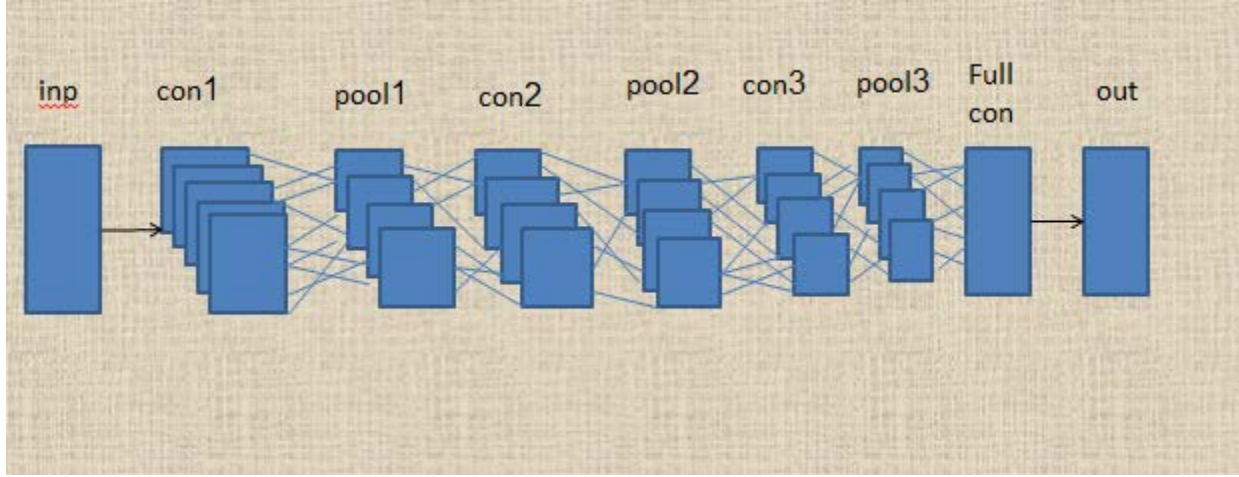


الشكل (3-5) عينات تحتوي أجسام غريبة

٣_١_٢_ تحميل الشبكة العصبونية العميقة:

تم استخدام CNN مع تعديلات محددة في طبقاتها لتمكينها من التعرف على تلك التعديلات بشكل دقيق.

٣_١_٢_١ بنية الطبقات في الشبكة العصبونية المستخدمة:



الشكل (3-6) بنية طبقات CNN

١_طبقة الدخل (Input Layer):

حجم الصورة الأصلية يحدد أبعاد هذه الطبقة.

٢_طبقات الفلترة Convolutional Layers :

تستخدم لاكتشاف الميزات البصرية في الصور.

يحدد حجم الفلاتر وعددها بناءً على تعقيد الوظيفة المطلوبة.

٣_طبقات الانحاء (Pooling Layers):

تستخدم لتقليل حجم الخرج من طبقات الفلترة.

غالباً ما تكون بحجم الفلتر والمستخدم هو 2

٤_طبقات التفعيل Activation Layers :

تستخدم وظيفة التنشيط ReLU لإضافة عنصر غير خطي إلى الشبكة.

٥_طبقة الاتصال الكامل (Fully Connected Layer):

تقوم بدمج الميزات المكتشفة من الطبقات السابقة.

٦_طبقة تكامل نموذج شبكة النفقة العالية ونموذج شبكة الانتباه :

حيث يتم من خلال هذه الطبقة تكامل عمل كل من النماذج السابقة من أجل كشف التعديلات على الصور وإعادتها إلى الشكل المحدد من أجل كشف الوجه حتى في الصور التي تم تعديلها مسبقاً

يتم استخدام الميزات التي تم استخراجها من طبقة الاتصال الكاملة من أجل تفعيل النماذج وتمثيل الميزات كبارامترات لهذه النماذج.

٧_طبقة الإخراج (Output Layer) :

عدد الوحدات في هذه الطبقة يعتمد على عدد الفئات المختلفة التي يجب التعرف عليها.

٣_٢_٢_١_حجم وأبعاد الطبقات في الشبكة:

١_حجم وأبعاد الفلاتر (Convolutional Layers): يتم استخدام 3×3 في هذه الطبقة
٢_حجم طبقات الانحناء (Pooling Layers): يتم استخدام 2×2 في هذه الطبقة اختيار الحجم جاء بناء على عدة اعتبارات عملية أولاً يعد هذا الحجم من الفلاتر الأكثر شيوعاً في الشبكات العصبية العميقة لأنه يسمح بتقليل الأبعاد بشكل فعال دون فقدان كبير للمعلومات الهامة. حيث يعمل 2×2 على تقليص الأبعاد بنسبة 50%، مما يساهم في تقليل عدد الحسابات المطلوبة، وبالتالي تسريع عملية التدريب وتحسين الكفاءة الحسابية للنموذج. كما أن هذا الحجم يعزز قدرة النموذج على التعميم بشكل جيد على بيانات جديدة دون التسبب في فقدان تفاصيل مهمة، مقارنةً باستخدام أحجام فلاتر أكبر مثل 5×5 التي قد تؤدي إلى فقدان التفاصيل الدقيقة. بالإضافة إلى ذلك، يعتبر 2×2 الحجم الأمثل الذي يوازن بين تقليل الأبعاد والحفاظ على الخصائص الرئيسية للصور، مما يجعله الخيار الأكثر استخداماً في الأدبيات العلمية والتطبيقات العملية .

٣_حجم طبقات التفعيل (Activation Layers): يعتمد على عدد الخرج من الطبقات

السابقة، وفي حالة النموذج المقترح فإن خرج طبقة الانحناء هو 3×3

٤_حجم طبقة الاتصال الكاملة (Fully Connected Layer): يتم استخدام هذه الطبقة من أجل تحويل الميزات المكتشفة من الصور إلى تصنيفات. تم حفظ عدد البيانات التي كانت

خرج الطبقة السابقة وإدخالها إلى هذه الطبقة من خلال الكود البرمجي وذلك من أجل المرونة وعدم تحديد الحجم بشكل مسبق .

٥_حجم طبقة الإخراج (Output Layer) : يتم تحديدها بناءً على عدد الفئات التي تحتاج إلى كشفها .

٣_١_٣ تحديد النموذج والتدريب : تم تحديد هيكل النموذج النهائي بما يسمح بالكشف عن التعديلات، ثم يتم تدريب النموذج باستخدام الصور الأصلية والصور المعدلة.

٣_١_٣_١ استخدام نموذج Highway Network :
الآلية الأساسية لها هي:

1. طبقة التوصيل المرتفع: تعمل كوسيط بين طبقة الإدخال والطبقات الخفية الأخرى، حيث تقوم بتحليل البيانات وتقديم توصيات للطبقات الخفية الأخرى حول كيفية تحسين التنبؤات.
2. الطبقات الخفية الأخرى: تستخدم توصيات الطبقة المرتفعة لتحسين التنبؤات التي تقدمها بناءً على البيانات التي تم تحليلها [12] .

٣_١_٣_٢ استخدام نموذج Attention Network :

١_ توجيه الانتباه: يتم تحديد مدى الانتباه لكل جزء من البيانات المدخلة. يتم تحديد هذا الانتباه عادة باستخدام دالة الانتباه تستخدم لحساب مؤشرات توجيه الانتباه التي تتوافق مع القيم الأخرى في البيانات.

٢_ تحديث الأوزان: بناءً على مؤشرات توجيه الانتباه، يتم تحديث الأوزان في الشبكة لتحسين التنبؤات التي تقدمها الشبكة. هذا يتم عادة عن طريق تحديث الأوزان بناءً على الأخطاء التي تتوقف عليها الشبكة في التنبؤات السابقة.

٣_ التكرار: يتم تكرار عملية التحديث وتوجيه الانتباه على مدى عدة دورات تدريب، حيث تتعلم الشبكة كيفية تحسين الأداء على المهام المحددة [13-14].

٣-١-٤- تقديم الصور للنموذج:

تم تقديم الصور الجديدة التي قد تم تعديلها إلى النموذج للكشف عن أي تعديلات قد تكون تمت عليها.

٣-١-٥- تقييم أداء النموذج:

تم قياس أداء النموذج بالنسبة لقدرته على كشف التعديلات المختلفة.

٣-١-٦- تدريب واختبار الخوارزمية المقترحة:

تم اختبار هذه الخوارزمية من أجل تعزيز دقة التعرف وفي نفس الوقت المحافظة على الاستقرار لكميات مختلفة من عينات الاختبار وبالتالي يمكن لهذه الشبكة التعرف على الوجوه بشكل دقيق بالنسبة لجميع الصور المختلفة من عينات الاختبار.

في التدريب يجب أن تجزأ قاعدة البيانات إلى ثلاثة أقسام:

- قسم للتدريب
- قسم ثاني للاختبار
- قسم ثالث للتحقق،

حيث قسم الاختبار يمثل 30% من قاعدة البيانات،

وقسم التحقق يمثل 20% من قاعدة البيانات،

وفيما يتعلق بقسم التدريب فهو يمثل 50% من قاعدة البيانات.

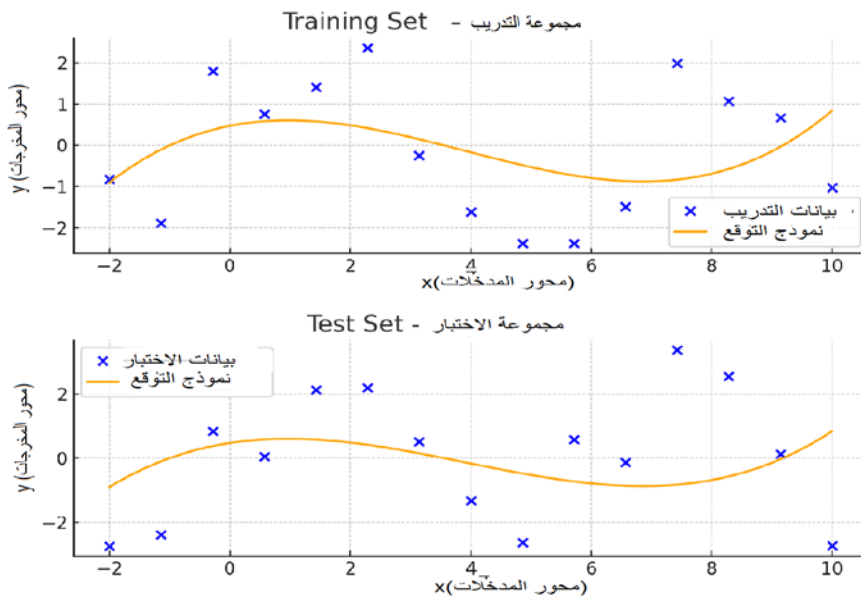
- إن القيام بعملية تقسيم قاعدة البيانات حسب النسب السابقة يضمن الحصول على نتائج

أدق، بالإضافة إلى تحسين عملية التدريب، لكنها تسبب في زيادة زمن التدريب [15].

تم إجراء 530 اختبار على 530 صورة وتظهر النتائج على هذه الاختبارات حسب الجدول التالي:

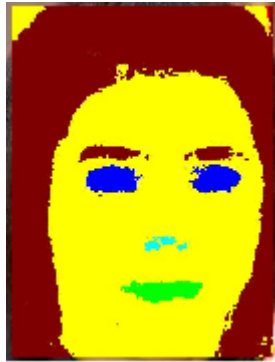
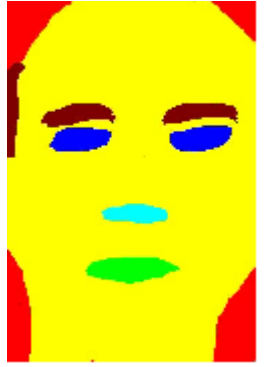
العملية	العدد	الرمز
صور تم توقع أنها لشخص وهي فعلا لهذا الشخص	272	TP
صور تم توقع أنها ليست لشخص وهي فعلا" ليست له	245	TN
صور تم توقع أنها لشخص وهي ليست لهذا الشخص	6	FP
صور تم توقع أنها ليست لشخص وهي لهذا الشخص	7	FN

يمثل الشكل الفرق بين بيانات الاختبار وبيانات التدريب:



الشكل (3-7) الفرق بين بيانات الاختبار والتدريب

٣-١-٦_١_ نتائج تطبيق عملية التجزئة على الصور المدخلة:



الشكل (8-3) الصور المجزأة التي ستتدخل عملية التصنيف



الشكل (9-3) الصور المجزأة التي ستدخل في عملية التصنيف

الفصل الرابع

التقييم والمقارنة

٤_ ١_ التقييم:

سيتم تقييم البحث على أساس الوثوقية التي يتم حسابها من أجل معرفة حالات الوثوقية لتطبيقات الذكاء الصناعي وهذه المعايير هي:

F1 Score -1

Precision -2

Recall -3

Accuracy -4

Specificity - 5

حيث يتم حساب كل عامل من هذه المعاملات من خلال قانون رياضي ولتطبيق هذه

القوانين يجب حساب كل من **TP , TN , FP, FN**

TP: True Positive وهي التي تم التوقع أنها صحيحة وهي فعلاً صحيحة في مثالنا (النتيجة المتوقعة true والنتيجة الحقيقية true).

TN: True Negative وهي التي تم التوقع أنها خاطئة وهي فعلاً خاطئة في مثالنا (النتيجة المتوقعة false والنتيجة الحقيقية false).

FP: False Positive وهي التي تم التوقع بأنه صحيح وهو في الحقيقة خاطئ في مثالنا (النتيجة المتوقعة true وفي الحقيقة false).

FN: False Negative وهي التي تم التوقع بأنها خطأ وهو في الحقيقة صحيح في مثالنا (النتيجة المتوقعة false ولكن في الحقيقة true).

وبعد حساب المعاملات الأربعة يمكن حساب دقة الخوارزميات التي تم تطبيقها [15- 16] .

٤_١_١_ حساب ال Precision :

ويتم حسابها من خلال المعادلة التالية

$$\text{Precision} = \frac{TP}{TP+FP}$$
$$= \frac{272}{272+6} = 0.98$$

وبهذا يكون معيار ال Precision يساوي 0.98 أي 98%

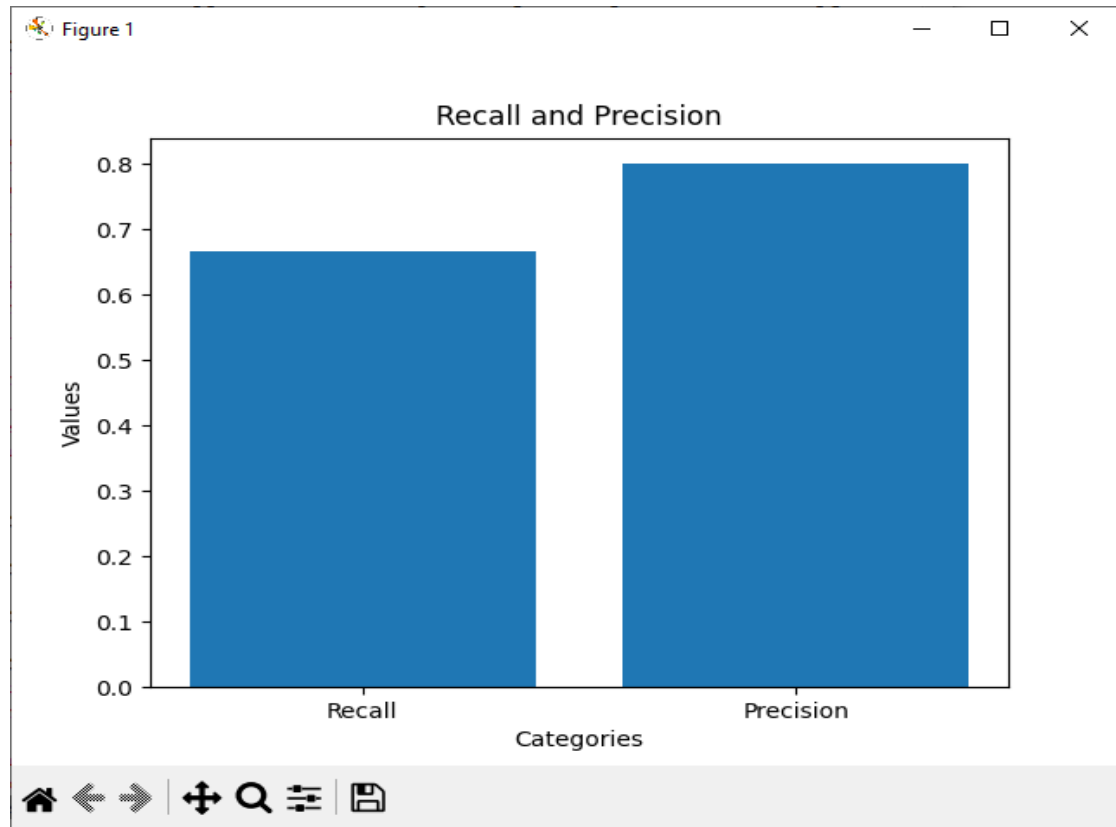
٤_١_٢_ حساب ال Recall :

ويتم حسابها من خلال المعادلة التالية:

$$\text{Recall} = \frac{TP}{TP+FN}$$
$$= \frac{272}{272+7} = 0.97$$

وبهذا يكون معيار ال Recall يساوي ال 0.975 أي 97.5%

يوضح الشكل التالي الاختلاف بين معايير الاستدعاء و الإحكام:



الشكل (1-4) يوضح الفرق بين الاستدعاء والإحكام

٤_١_٣ _ حساب F1 score :

ويتم حسابها من خلال المعادلة التالية:

$$F1score = \frac{2 * precision * Recall}{Precision + Recall}$$

$$= \frac{2*0.98*0.975}{0.98+0.975} = 0.98$$

وبهذا يكون معيار ال **F1 score** يساوي 0.98 أي 98%

٤_١_٤_ حساب الدقة **Accuracy** :

ويتم حسابه من خلال المعادلة التالية:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$= \frac{272+245}{272+245+6+7} = 0.975$$

وبهذا يكون معيار ال Accuracy يساوي 0.975 أي 97.5%

٤_١_٥_ حساب الحساسية **Specificity** :

ويتم حسابه من خلال المعادلة التالية:

$$\text{Specificity} = \frac{TN}{FP+TN}$$

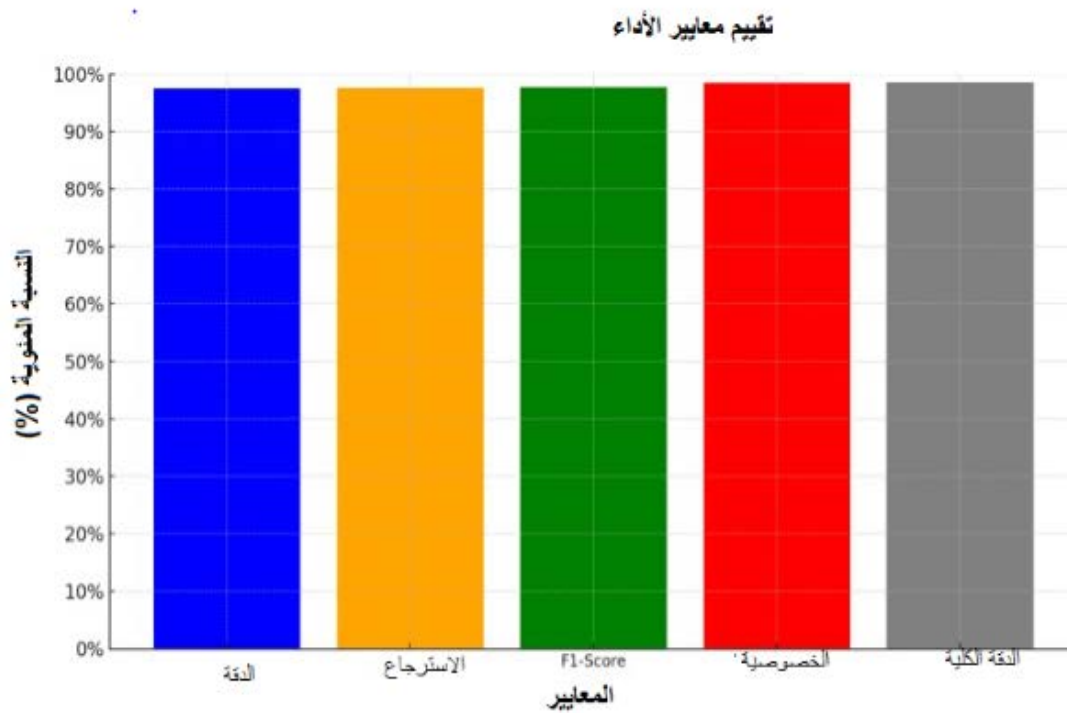
$$= \frac{241}{241+6} = 0.975$$

وبهذا يكون معيار ال Specificity يساوي 0.975 أي 97.5%

ويمكن تقييم المعايير السابقة حسب الجدول التالي:

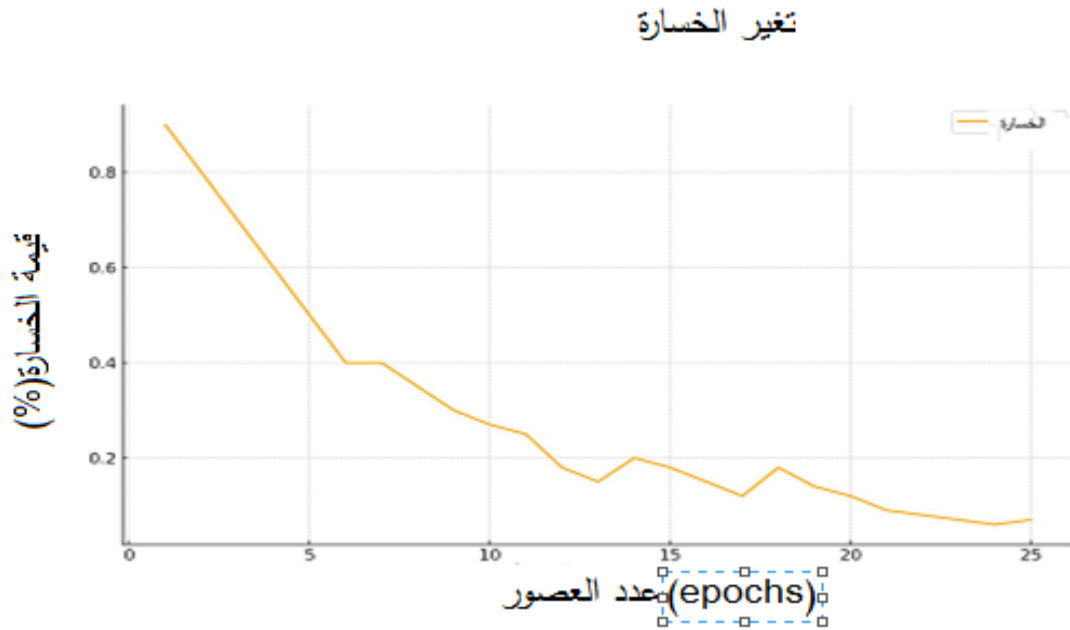
المعيار	Precision	Recall	F1 Score	Accuracy	Specificity
النتيجة	98%	97.5%	98%	97.5%	97.5%

ويمكن تمثيل الجدول السابق بالمخطط التالي:



الشكل (4-2) مخطط بياني لمعايير تقييم الأداء

مخطط يوضح تابع الخسارة:



الشكل (3-4) تابع الخسارة

العلاقة التي تمثل تابع الخسارة بشكل معادلة رياضية هي:

$$L = \frac{1}{N} \sum_{i=1}^N \mathcal{L}(y_i, \hat{y}_i)$$

حيث:

- L : قيمة الخسارة الكلية.
- N : عدد العينات (الأمثلة) المستخدمة في التدريب
- y_i : القيمة الحقيقية للتصنيف للعيينة i
- $\hat{y}_i$: القيمة المتوقعة (المخرجات) من النموذج للعيينة i

- دالة الخسارة المحددة (مثل الخسارة التربيعية أو Cross-Entropy Loss)

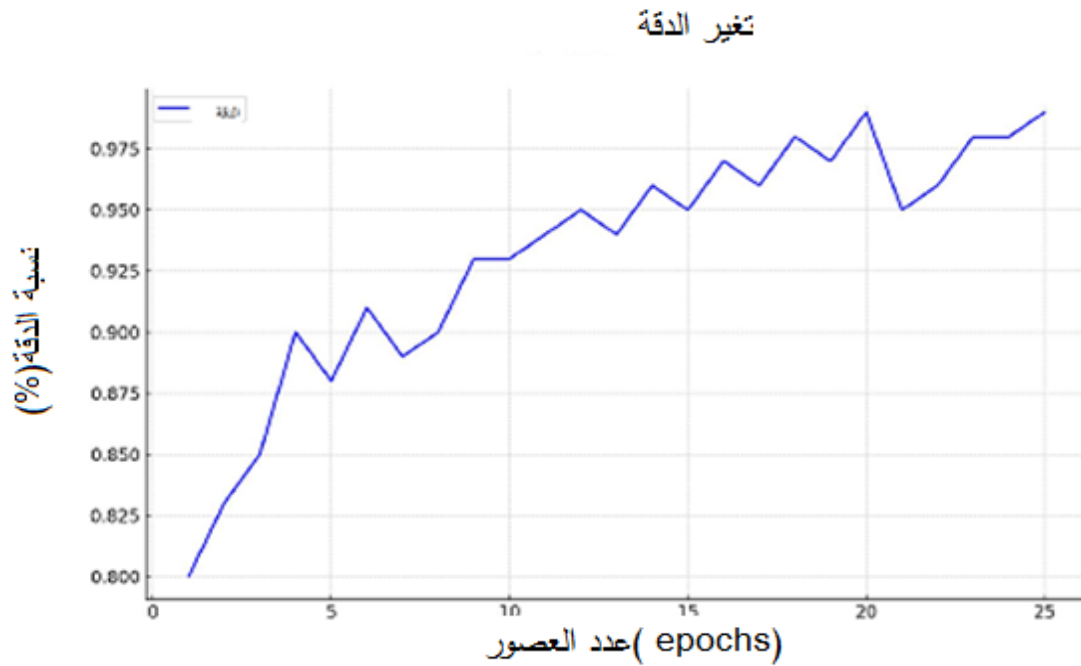
بالنسبة لخسارة Cross-Entropy (المستخدمة عادةً في تصنيف الصور)

$$L = - \frac{1}{N} \sum_{i=1}^N \left[y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i) \right]$$

التفسير:

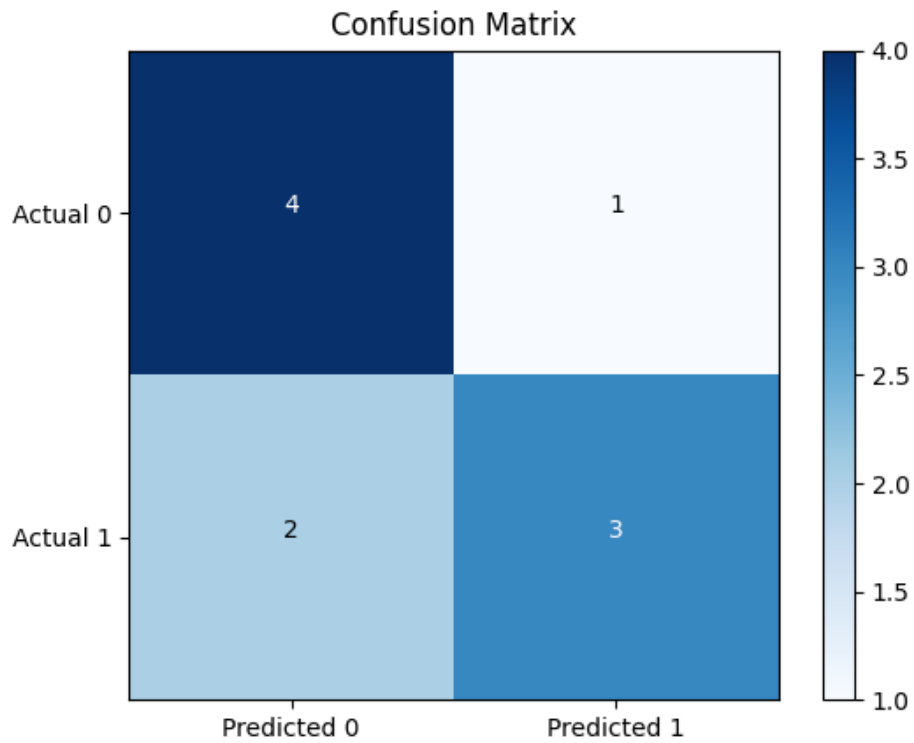
- القيمة L تتناقص مع زيادة عدد العصور (Epochs) بفضل تحسين الأوزان أثناء التدريب باستخدام خوارزمية الانحدار التدريجي (Gradient Descent)
- الهدف هو تقليل L قدر الإمكان حيث تعكس قيمة الخسارة مدى خطأ النموذج.

مخطط يوضح تابع الدقة:



الشكل (4-4) تابع الدقة

وتظهر المخططات التي يتم إنشاؤها من الشبكة بحد ذاتها على الشكل التالي:
معياري ال Confusion Matrix:



الشكل (4-5) يوضح confusion matrix

المعايير السابقة بدون استخدام full connection

المعيار	Precision	Recall	F1 Score	Accuracy	Specificity
النتيجة بدون استخدام الـ full connection	98%	97.5%	98%	97.5%	97.5%

٤_٢_ المقارنة :

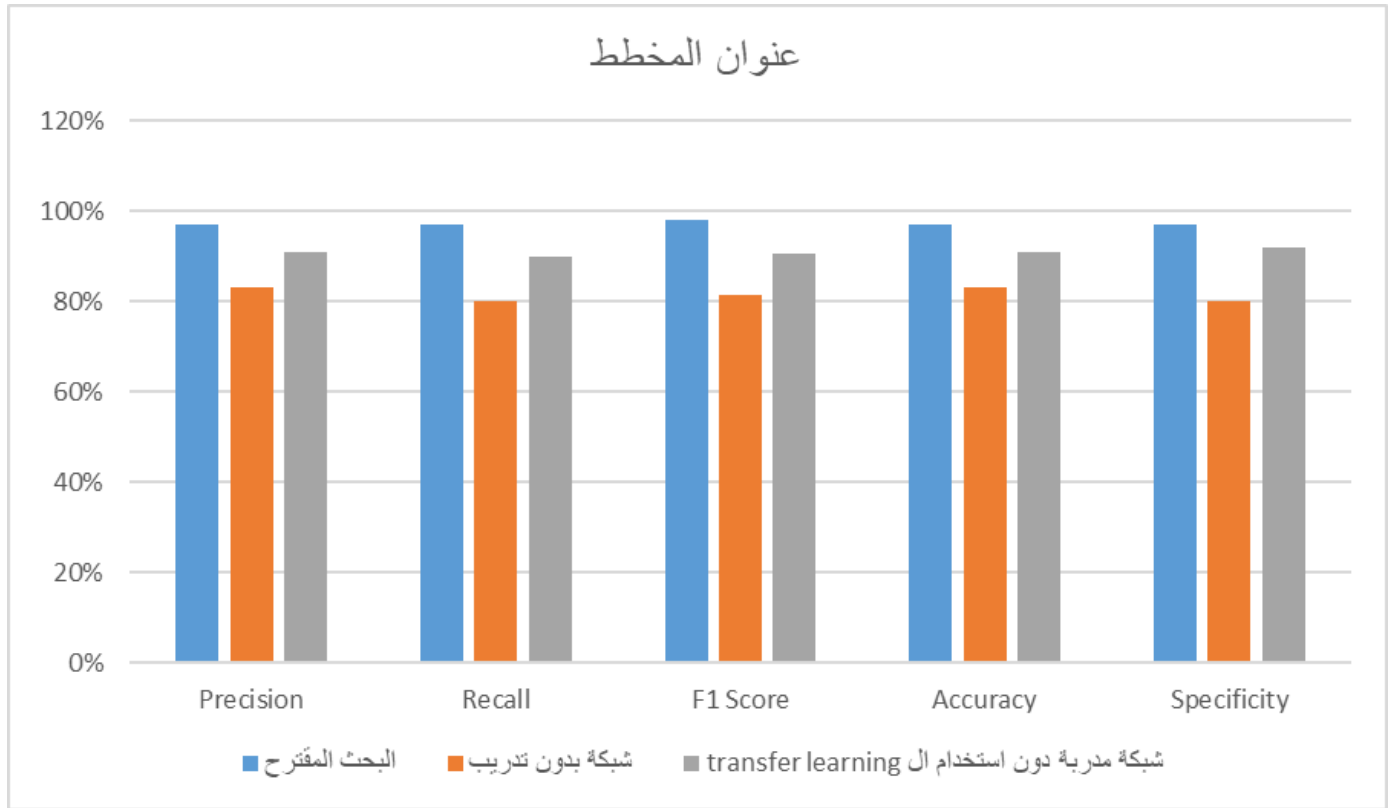
سيتم تقييم مقارنة البحث بعدة أشكال بحيث يوجد عدة أشكال لتقييم عمل الشبكات العصبونية وذلك من أجل تقييم عملها بشكل صحيح ودقيق.

٤_٢_١_ المقارنة بين الشبكة المدربة والشبكة الغير المدربة:

سيتم ضمن هذه الفقرة تقييم أداء عمل الشبكة العصبونية عند تدريبها وحفظ نتائج التدريب وعملها دون عملية التدريب وإظهار النتائج وتمت هذه العملية بتمرير نفس عدد الصور للشبكة العصبونية المدربة والتي نتائجها محفوظة ضمن مودل والشبكة العصبونية الغير المدربة وعلى نفس الصور وتم إجراء الاختبار على 180 صورة مختلفة وظهرت النتائج كالتالي:

المعيار	البحث المقترح	شبكة بدون تدريب	شبكة بدون تدريب وبدون استخدام ال full connection	شبكة مدربة دون استخدام ال transfer learning و دون استخدام ال full connection	شبكة مدربة دون استخدام ال transfer learning
Precision	%98	%83	%77	%85	%91
Recall	%97.5	%80	%73.5	%84	%90
F1 Score	%98	%81.5	%75	%83	%90.5
Accuracy	%97.5	%83	%77	%84	%91
Specificity	%97.5	%80	%74	%80	%92

ويمكن تمثيل الجدول السابق حسب المخطط التالي:



الشكل (4-6) مخطط بياني للمعايير لشبكة مدربة وغير مدربة و بدون transfer learning

الجدول يقارن بين نماذج مختلفة من الشبكات العصبية وأدائها تحت شروط متعددة. كل نموذج تم تقييمه عبر عدة طرق تدريب، سواء باستخدام أو بدون استخدام التدريب باستخدام النقل التعليمي (Transfer Learning) والتدريب بدون استخدام الاتصال الكامل (Full Connection) التقارب في النتائج يظهر لأن جميع النماذج أدت بشكل مشابه تحت نفس الظروف هذا قد يكون بسبب :

1. الشبكات متماثلة: النماذج المستخدمة قد تكون متماثلة في بنية الشبكة أو الطريقة التي يتم بها تدريبها
2. المهام المشابهة: إذا كانت المهام التي يتم تقييم النماذج عليها مشابهة جداً، فمن المتوقع أن تكون النتائج متقاربة.

3. نقل المعرفة: استخدام النقل التعليمي قد يساعد في تقليل الفوارق بين النماذج المختلفة لأن المعرفة المكتسبة من نموذج معين تُستخدم في نماذج أخرى.

٤_٣_ النتائج والمناقشة:

تم اقتراح نظام جديد للتعرف والكشف عن ملامح الوجوه وتصنيفها، وقد جرى تطبيق النظام واختباره على مجموعة واسعة من الصور الرقمية، إذ تم في البداية استخدام خوارزميات التجزئة Segmentation من أجل تقطيع الصور وتحليلها إلى الأجزاء الرئيسية الخاصة بها، وإظهار الأماكن المميزة ضمن هذه الصورة، ومن أجل إعطاء جميع الصور نفس الأبعاد كي لا يتم تغيير النتائج التي يتم تقديمها إلى الخوارزمية، بحيث يتم ضمان أن جميع الصور لها نفس الأبعاد.

وجرى بعد ذلك استخدام خوارزمية زيادة البيانات Data Augmentation من أجل زيادة عدد الصور التي سيتم الاختبار عليها ضمن الشبكات العصبونية، إذ أن خوارزمية زيادة البيانات تقوم بإعطاء نفس الصورة ولكن من عدة منظورات مختلفة وبأشكال مختلفة، وهذا يضمن التأكد من أن الشبكة العصبونية سيتم تدريبها بشكل جيد كونه من المعروف أن الشبكات العصبونية التي تعمل على الصور تعطي نتائج أدق كلما زدنا عدد الصور التي يتم إدخالها إلى الشبكة.

في النهاية تم إجراء عدد كبير من التجارب والتي تحقق أعلى دقة ممكنة وبأقل زمن معالجة ممكن مقارنة بالدراسات السابقة.

٤_٤_ الاستنتاجات:

١_ أظهر النظام أهمية دمج نظامي التجزئة والتصنيف، مقارنة بالأنظمة التي تعتمد على التجزئة فقط.

٢_ أظهر النظام كفاءة في التخلص من الصعوبات والمشاكل التي واجهت الدراسات السابقة، بسبب احتواء بعض الصور على مشاكل أثناء عملية أخذ الصور والتي تؤثر على جودة التصنيف.

٣_ كما أظهر النظام كفاءة عالية في التعرف على الصور التي تحتوي على نسبة ضجيج عالية.

٤_ استخراج ملامح الوجه وحوافه بشكل مفصل بعد أن تمت عملية التجزئة بشكل صحيح، وعلى الرغم من اختلاف السويات الرمادية والأبعاد للصور الرقمية الناتجة عن الاختلاف في عمر الشخص وشكل الوجه ولونه.

٥_ عدم استخلاص أي ملامح إضافية أو خاطئة في الصور غير الواضحة ومبهمة الملامح والتي تحتوي أشياء غير مهمة (كالنظارات والأقراط).

٦_ قام النظام بتحقيق عملية التشخيص بفاعلية عالية، من خلال إنجاز عملية التصنيف بشكل آلي، ومن دون تدخل المستخدم في أي مرحلة من المراحل.

المراجع :

- [1] L. Tian. Investigating Attention Maps for Partial Face Recognition. Research Internship. Technical University of Munich, 2022.
- [2] I. Belgacem. Vaccinate Your Images Against Deepfakes: Leveraging Adversarial Attacks to Fight Image-Based Abuse. Master's Thesis. Technical University of Munich, 2022.
- [3] M. S. Saliya. Multi-Camera Landmark-based 3D Ear Localization. Master's Thesis. Technical University of Munich, 2021.
- [4] W. Zhao. Deep Aggregation Network for Set-Based Face Recognition. Master's Thesis. Technical University of Munich, 2021.
- [5] T. Kong. Face Morphing with Generative Adversarial Network . Research Internship. Technical University of Munich, 2021.
- [6] Recognizing Families through Images with Pre-trained Encoder, Tuan-Duy H. Nguyen, Huu-Nghia H. Nguyen, and Hieu Dao, 2020.
- [7] Bhowmich. Conditional Face Image Generation and Its Application in 3d Pedestrian Modelling. Master's Thesis. Technical University of Munich, 2020.

- [8] Z. Wang. Automotive Powertrain Health Prediction based on Diagnostic Trouble Code. Master's Thesis. Technical University of Munich, 2020.
- [9] Z. Zhang. Attention-based Partial Face Recognition. Master's Thesis. Technical University of Munich, 2020.
- [10] Z. Cao. An Image Aggregation Network for Video Face Recognition. Master's Thesis. Technical University of Munich, 2020.
- [11] song, B., Xu, J., & Xu, L. (2019). Multilevel inverter fault diagnosis based on WPT-PCA-NN algorithm. *Journal of Shandong University of Science and Technology (Natural Science)*, 38(1), 111-120.
- [12] J. Graefe. Multicolumn Networks for Face Recognition. Senior Seminar. Technical University of Munich, 2019.
- [13] A. Moiz. Attention Based Fusion Using Feature Aggregation for Audio-Visual Person Verification. Master's Thesis. Technical University of Munich, 2019.
- [14] N. Kardileev. Enhanced Driver Gaze Classification using Video Data together with Car Bus Signals. Master's Thesis. Technical University of Munich, 2019.
- [15] Interactive Facial Feature Localization, Vuong Le, Jonathan Brandt, Zhe Lin, and Thomas S.Huang, 2019.

- [16] S. Hillebrand. Disentangled Representation Learning GAN for Pose-Invariant Face Recognition. Senior Seminar. Technical University of Munich, 2018.
- [17] L. N. G'artner. Exploring Disentangled Feature Representation Beyond Face Identification. Senior Seminar. Technical University of Munich, 2018.
- [18] El Khiyari, H. and Wechsler, H. (2016) Face Verification Subject to Varying (Age Ethnicity, and Gender) Demographics Using Deep Learning. Journal of Biometrics and Biostatistics , 7, 323. <https://doi.org/10.4172/2155-6180.1000323>.
- [19]Guo Wei, Cai Ning. Convolutional Network Coding [J]. Journal of China Academy of Electronic Science and Technology, 2016.
- [20] Parkhi, O.M., Vedaldi, A. and Zisserman, A. (2015) Deep Face Recognition. Proceedings of the British Machine Vision Conference (BMVC). <https://doi.org/10.5244/c.29.41>
- [21] Z. Liu, L. Meng, and Z. Zhu, "A Systematic Approach to Face Recognition Using Convolutional Neural Network," IEEE Access, 2024.
- [22] J. Martinez, H. Park, and R. Balasubramanian, "Second Edition FRCSyn Challenge at CVPR 2024: Face Recognition Challenge in the Era of Synthetic Data," Proceedings of CVPR 2024, 2024.

Abstract

Facial recognition is a verification system that works very quickly and effectively. It is faster and more convenient than other biometric technologies, such as fingerprints or retinal scanning. The facial recognition system relies on fewer touchpoints than entering passwords or personal identification numbers. It is compatible with multi-factor authentication to allow an additional security verification step. Based on these advantages and risks facing the security and stability of computer systems and others in the world, it was necessary to work on devising new automated methods for analyzing faces and detecting their various features using a robot.

The research aims to present a new algorithm and methodology for face recognition, by using a different set of algorithms that are used in image processing, such as segmentation algorithms and data augmentation method, and convolutional neural networks (CNN) were used to find out the final results. The system was tested on a large number of images to confirm the results of the research.

This algorithm was then trained and tested on 530 images, and the results showed an accuracy of more than 97%.

key words:

Face recognition, digital image processing, segmentation and classification, data augmentation algorithm, convolutional neural networks.

**Syrian Arab Republic
Damascus University
Faculty of Mechanical and Electrical Engineering
Department of Computers and Automation
Engineering
Master in Robot Engineering and Programming**



Improving Human–Machine Communication Technologies Using Deep Learning

A dissertation submitted To Damascus University in partial fulfillment of the requirements for the degree of Master in Robot Engineering and programming at Faculty of Mechanical and Electrical Engineering

**By:
Eng. Wafaa wafik salamh**

**Main Supervisor:
Prof.Eng.Dr. Samir karaman**

2024