



الجمهورية العربية السورية
وزارة التعليم العالي والبحث العلمي
جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
قسم هندسة الإلكترونيات والاتصالات

تقييم أداء خوارزميات ترميم الصور الرقمية

رسالة مقدمة لنيل درجة الماجستير في هندسة الإلكترونيات التطبيقية

إعداد

م. تقى الكيال

إشراف

د.م. طلال حمود

العام الدراسي

2023-2022

صفحة قرار الحكم

تصريح خطي

أنا الموقعة أدناه أصرح بعلمي الكامل وقبولي:

- أن كل ما ينتج عن البحث والرسالة هو ملكية فكرية ومالية بالتشارك مع جامعة دمشق، وأنني ألتزم بأخذ موافقة الجامعة في حال رغبتني في نشر البحث أو جزء منه نصاً أو مضموناً خارج إطار الجامعة (من دور نشر أو مكتبات أو مواقع إلكترونية وغيرها من وسائل النشر).
- أنه لا يوجد أي جزء من هذه الرسالة اقتبس من عمل آخر أو أنجز للحصول على شهادة أخرى في جامعة دمشق، أو أية جامعة، أو معهد تعليمي داخل الجمهورية العربية السورية أو خارجها.

الاسم والتوقيع

م. تقى الكيال

إهداء:

إلى من منحتني الحب والحنان وسهرت على راحتي وتعليمي منذ الصغر صاحبة القلب الأبيض
الواسع والرأي السديد أمي الحبيبة.

إلى من أفنى شبابه وعمره في سبيل تربيتي وتعليمي وأزال من دربي الأشواك ليمهد لي
طريق العلم والحياة أبي الحبيب.

إلى من منحتني أعظم مشاعر الحب والأمومة إلى الغائبة الحاضرة في قلبي إلى عصفورة الجنة
... ابنتي فرح.

إلى رفيق الدرب وشريك الحلو والمر إلى الصديق والسند زوجي الغالي.

إلى أصدقاء الحياة وبلسم الجرح والسند أخوتي الأعزاء.

إلى عائلتي الثابتة وأصحاب القلوب البيضاء المحبة أهل زوجي.

إلى من لم يبخل علي بعلمه وخبرته وجهوده ومتابعته وتحمل مشقة الطريق معي الدكتور
المهندس طلال حمود.

إلى أعضاء قسم هندسة الإلكترونيات والاتصالات في جامعة دمشق على ملاحظاتهم واقتراحاتهم
في جلسات المداورة.

وإلى جميع الأساتذة والزلاء في كلية الهندسة الميكانيكية والكهربائية في جامعة دمشق،
ولكل من أسهم في إنجاز هذا البحث حتى يرى النور.

فهرس المحتويات

2.....	صفحة قرار الحكم
3.....	تصريح خطي
4.....	إهداء
6.....	فهرس المحتويات
10.....	فهرس الأشكال
13.....	فهرس الجداول
14.....	قائمة المصطلحات والإختصارات
15.....	المستخلص
16.....	Abstract
17.....	الفصل الأول: الإطار العام للبحث General Framework of the Research
18.....	1-1 مقدمة
19.....	2-1 بعض الدراسات السابقة
19.....	3-1 مشكلة البحث
20.....	4-1 الهدف من البحث
20.....	5-1 أهمية البحث
20.....	6-1 مسوغات البحث
21.....	7-1 محددات البحث
21.....	8-1 حدود البحث
21.....	9-1 منهج البحث وأدواته
22.....	10-1 بنية الرسالة
23.....	الفصل الثاني: الأساسيات النظرية للبحث Theoretical Basis of the Research
24.....	1-2 مقدمة

26.....	2-2 أنواع الضجيج في الصور الرقمية
26.....	3-2 مرشحات التخلص من الضجيج في المجال الحيزي
27.....	4-2 ترابط الإشارات المتقطعة زمنياً Correlation of discrete-time signals
28.....	5-2 الشبكات العصبونية (NN) Neural Networks
29.....	1-5-2 الطبقات المكونة للشبكات العصبونية
30.....	2-5-2 بنية العصبون
31.....	3-5-2 تدريب الشبكات العصبونية
32.....	4-5-2 تابع التنشيط Activation function
38.....	6-2 الشبكة العصبونية الالتفافية (CNN) Convolutional Neural Network
39.....	1-6-2 طبقة الطي Convolutional layer
40.....	2-6-2 طبقة التجميع Pooling Layer
41.....	3-6-2 طبقة الترابط الكلي Fully Connected (FC) Layer
42.....	7-2 البنية U-Net
44.....	8-2 الشبكة المتعاكسة المولدة (GAN) Generative Adversarial Network
45.....	1-8-2 المميز (D) Discriminator
46.....	2-8-2 المولد (G) Generator
47.....	3-8-2 بعض تطبيقات البنية GAN

الفصل الثالث: ترميم الصور وتطبيقاته وبعض خوارزميات الترميم Image Inpainting

49.....	Applications and Some Image Inpainting Algorithms
50.....	1-3 ترميم الصور وتطبيقاته
51.....	2-3 الطرائق المتبعة في ترميم الصور الرقمية
51.....	1-2-3 الطرائق التقليدية Traditional Methods
51.....	2-2-3 الطرائق المعتمدة على التعلم Learning-based Methods
51.....	3-3 بعض خوارزميات ترميم الصور الرقمية

1-3-3 الخوارزمية الأولى: ترميم الصور بواسطة تعلّم الشبكة المتعاكسة المولّدة للحواف	
(EdgeConnect).....	52
2-3-3 الخوارزمية الثانية: ترميم الصورة بواسطة الخطوط العامة والشكل العام للصورة	
(StructureFlow).....	56
3-3-3 الخوارزمية الثالثة: ترميم الصور بواسطة شبكة الدمج العميق (DFNet).....	59
الفصل الرابع: تقدير جودة الصور الرقمية وبارامترات التقييم التقليدية والمُقترحة لتقييم جودة الصور	
المرمّمة Digital Image Quality Assessment and Traditional and Proposed	
63 Evaluation Parameters for Inpainted Image Quality Evaluation	
1-4 تقدير جودة الصور الرقمية.....	64
2-4 طرائق تقدير جودة الصور الرقمية.....	64
3-4 بارامترات التقييم التقليدية لتقدير جودة الصور المرممة.....	65
1-3-4 نسبة الإشارة العظمى إلى الضجيج(PSNR).....	65
2-3-4 معامل التشابه البنيوي (SSIM).....	65
4-4 البارامترات المُقترحة لتقدير جودة الصور المرممة.....	66
1-4-4 البارامتر المُقترح الأول: مربع متوسط الخطأ (MSE).....	66
2-4-4 البارامتر المُقترح الثاني: تقدير جودة الصور بالاعتماد على الحواف والبكسل	
(EPIQA).....	67
3-4-4 البارامتر المُقترح الثالث: الترابط المُقيّس(NCC).....	69
70 Implementation and Results الفصل الخامس: التنفيذ العملي والنتائج	
1-5 تمهيد.....	71
2-5 المرحلة الأولى: تطبيق خوارزميات الترميم على المجموعة نفسها من الصور المشوّهة.....	72
3-5 المرحلة الثانية: حساب بارامترات التقييم التقليدية لنتائج المرحلة الأولى.....	75
4-5 المرحلة الثالثة: حساب بارامترات التقييم المُقترحة لنتائج المرحلة الأولى.....	77
5-5 المرحلة الرابعة: تقييم فعالية البارامترات المُقترحة وتقييم أداء خوارزميات الترميم بناءً عليها	78

79.....	6-5 الخاتمة
80	الآفاق المستقبلية
81	References المراجع

فهرس الأشكال

- الشكل (1-2) مثال على المرشح المتوسط..... 25
- الشكل (2-2) مثال على المرشح الوسطي..... 26
- الشكل (3-2) التشابه بين عصبون الشبكة العصبونية وعصبون الدماغ البشري..... 27
- الشكل (4-2) بنية الشبكة العصبونية..... 29
- الشكل (5-2) بنية العصبون..... 29
- الشكل (6-2) مثال على المسار الأمامي..... 31
- الشكل (7-2) مثال على المسار الخلفي..... 32
- الشكل (8-2) تابع الخطوة الثنائي 33
- الشكل (9-2) تابع التنشيط Sigmoid..... 34
- الشكل (10-2) تابع التنشيط Tanh..... 34
- الشكل (11-2) تابع التنشيط Relu..... 35
- الشكل (12-2) تابع التنشيط LRelu..... 35
- الشكل (13-2) تابع التنشيط PRelu..... 36
- الشكل (14-2) تابع التنشيط Elu..... 36
- الشكل (15-2) مثال على حساب احتمال كل صنف بواسطة تابع التنشيط softmax 37
- الشكل (16-2) مثال على بنية CNN..... 38
- الشكل (17-2) مثال على عملية الطي في صورة رمادية..... 39
- الشكل (18-2) مثال على عملية التجميع الأعظمي 40
- الشكل (19-2) مثال على طبقة الترابط الكلي 41
- الشكل (20-2) مثال على مفهوم التقسيم على مستوى البكسل..... 42

43.....	الشكل (21-2) بنية الشبكة U-Net
44.....	الشكل (22-2) بنية GAN
45.....	الشكل (23-2) مسار تدريب المميز
46.....	الشكل (24-2) مسار تدريب المولّد
49.....	الشكل (1-3) أمثلة على تطبيقات ترميم الصور
52.....	الشكل (2-3) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم EdgeConnect
53.....	الشكل (3-3) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم EdgeConnect
53.....	الشكل (4-3) مثال على مراحل عمل خوارزمية الترميم EdgeConnect
53.....	الشكل (5-3) المخطط التفصيلي لخوارزمية الترميم EdgeConnect
54.....	الشكل (6-3) المخطط الوظيفي لخوارزمية الترميم EdgeConnect
55.....	الشكل (7-3) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم StructureFlow
56.....	الشكل (8-3) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم StructureFlow
56.....	الشكل (9-3) مثال على مراحل عمل خوارزمية الترميم StructreFlow
57.....	الشكل (10-3) المخطط التفصيلي لخوارزمية الترميم StructreFlow
57.....	الشكل (11-3) المخطط الوظيفي لخوارزمية الترميم StructreFlow
58.....	الشكل (12-3) بنية الدمج Fusion block
59.....	الشكل (13-3) المخطط التفصيلي لخوارزمية الترميم DFNet
59.....	الشكل (14-3) المخطط الوظيفي لخوارزمية الترميم DFNet
64.....	الشكل (1-4) مرشحات استخلاص الحواف
69.....	الشكل (1-5) الصور المشوّهة التي طُبقت على خوارزميات الترميم الثلاث
70.....	الشكل (2-5) نتائج ترميم الخوارزمية الأولى (EdgeConnect) للصور المشوّهة

- الشكل (3-5) نتائج ترميم الخوارزمية الثانية (StructreFlow) للصور المشوّهة 71
- الشكل (4-5) نتائج ترميم الخوارزمية الثالثة (DFNet) للصور المشوّهة 72

فهرس الجداول

الجدول (1-5) قيم البارامتر التقليدي PSNR لنتائج المرحلة الأولى	73
الجدول (2-5) قيم البارامتر التقليدي SSIM لنتائج المرحلة الأولى	73
الجدول (3-5) قيم البارامتر المقترح MSE لنتائج المرحلة الأولى	74
الجدول (4-5) قيم البارامتر المقترح EPIQA لنتائج المرحلة الأولى	74
الجدول (5-5) قيم البارامتر المقترح NCC لنتائج المرحلة الأولى	75

قائمة المصطلحات والإختصارات

CNN	Convolutional neural network	شبكة عصبونية التلافيفية
EC	EdgeConnect :Generative Image Inpainting with Adversarial Edge Learning	ترميم الصور بواسطة تعلّم الشبكة المتعاكسة المولدة للحواف
EPIQA	Edge and Pixel-Based Image Quality Assessment	تقدير جودة الصور بالاعتماد على الحواف والبكسل
GAN	Generative adversarial network	الشبكة المتعاكسة المولدة
MSE	Mean squared error	مربع متوسط الخطأ
NCC	Normalized Cross-Correlation	الترابط المُقيّس
PSNR	Peak signal to noise ratio	نسبة الإشارة العظمى إلى الضجيج
SSIM	Structural similarity index	معامل التشابه البنيوي
SF	StructureFlow :Image Inpainting via Structure-aware Appearance Flow	ترميم الصورة بواسطة الخطوط العامة والشكل العالم للصورة
DFNet	Deep Fusion Network for Image Completion	ترميم الصور بواسطة شبكة الدمج العميق

الملخص

تتعرض الصور بفعل التقادم والعمليات المطبقة عليها إلى تشوهات مختلفة، ومن ثم إلى نقص في المعلومات.

تعمل خوارزميات ترميم الصور على إعادة بناء الأجزاء الناقصة من الصور، ومن ثم إلى التخلص من التشوهات والضجيج واسترجاع المعلومات المفقودة. تعتمد كل خوارزمية من خوارزميات ترميم الصور - عامةً - في اختبار أدائها على إدخال عدد من الصور التالفة إلى الخوارزمية - كل خوارزمية تختار مجموعة صور تالفة مختلفة عن الأخرى - ثم التوصل إلى نتائج ترميم كل منها ثم حساب بارامترات التقييم التقليدية (PSNR, SSIM) وتقييم أداء خوارزمية الترميم اعتماداً على قيم تلك البارامترات.

وُجد في هذا البحث أن بارامترات التقييم التقليدية المستخدمة في تقييم الأداء غير كافية فهي تخلق حالة من الشك يصعب معها اتخاذ القرار أيّ خوارزمية من خوارزميات الترميم هي الأفضل.

يهدف هذا البحث إلى تقييم أداء خوارزميات الترميم من خلال توحيد الصور المشوّهة المدخلة واقتراح بارامترات تقييم جديدة تزيل حالة الشك التي وُجدت في البارامترات التقليدية وتدعم عملية اتخاذ القرار.

اعتمد في هذا البحث على الأدوات البرمجية (Python, MATLAB) في الحصول على نتائج ترميم ثلاث خوارزميات ترميم لعدد من الصور المشوّهة، وحساب قيم بارامترات التقييم التقليدية والمُقترحة.

من أجل تحقيق تقييم أداء صحيح لابدّ من اختبار عمل خوارزميات الترميم على مجموعة الصور المشوّهة ذاتها واختيار بارامترات تقييم مناسبة.

يُبين في هذا البحث امكانية استخدام ثلاث بارامترات إضافية هي:

NCC, MSE, EPIQA لتقييم أداء خوارزميات الترميم لإعادة بناء وتشكيل الأجزاء الناقصة من الصورة.

الكلمات المفتاحية:

ترميم الصور، تقييم الأداء، تقدير جودة الصور، بارامترات تقييم الأداء.

Abstract

Digital images are subject to a wide variety of distortions because of time and operations on them which means lose of information.

Image inpainting algorithms aim to reconstruct the missing areas in images, therefore get rid of distortions, noises and missing information recovery. In general to evaluate performance of image inpainting algorithms, every algorithm inpaints different group of distorted images, calculate the traditional evaluation parameters (PSNR, SSIM) and evaluate the performance based on them.

In this research was found that the traditional parameters are not suitable for performance evaluation because of confusion that make decision, which algorithm is better, so hard.

This research aims to evaluate the performance of image inpainting algorithm by inpaint same group of distorted images and propose new evaluation parameters to solve problems in traditional parameters.

In this research, based on Python and MATLAB, three distorted images were inpainted via three image inpainting algorithms and calculate both traditional and proposed evaluation parameter.

In order to make fair evaluation the performance evaluation on same images and choose good evaluation parameters is a must.

In this research was found that the parameters (NCC, MSE, EPIQA) are more suitable to evaluate image inpainting algorithms performance.

Key words:

Image inpainting, performance evaluation, image quality assessment, evaluation parameters.

الفصل الأول

الإطار العام للبحث

General Framework of the Research

1-1 المقدمة:

رغم أن الإهتمام بمجال ترميم الصور - أو عملية إصلاح الصور القديمة والمشوهة - موجود منذ سنوات عدة إلا أنه يحظى حالياً باهتمام أكبر من السابق، وذلك بسبب التطور الكبير الذي شهدته تقانات معالجة الصور. ومع تطور أدوات معالجة الصور أصبح لترميم الصور أوتوماتيكياً تطبيقات مهمة عديدة إضافة إلى كونه محط اهتمام كبير للباحثين في هذا المجال.

لم تقتصر تطبيقات ترميم الصور على إعادة إحياء الصور القديمة فقط، بل تعدتها إلى تطبيقات أخرى كالتخلص من أنواع التشوهات المختلفة التي قد تعاني منها الصورة إضافة إلى حذف جزء غير مرغوب من الصورة يحدده المستخدم.

وعلى الرغم من أن هناك عدداً كبيراً من الخوارزميات المهمة بمجال ترميم الصور حيث تهدف جميعها إلى التوصل لصورة مرممة أقرب ما تكون للصورة السليمة إضافة إلى خلوها من العيوب والتشوهات وخاصة عند الحواف الفاصلة بين الجزأين السليم والمرمم إلا أن لكل خوارزمية بنية خاصة بها وجوانب معينة تهتم بها تختلف عن غيرها من الخوارزميات.

تعتمد كل خوارزمية في تقييم أدائها على إدخال مجموعة ما من الصور المشوهة للحصول على الصورة المرممة لكل منها، ثم حساب بارامترات التقييم PSNR و SSIM، وتقييم الأداء اعتماداً على قيم تلك البارامترات.

من غير الممكن تحقيق مقارنة بين أداء عدد من الخوارزميات ما لم يتم تطبيقها على ذات المجموعة من الصور ثم حساب بارامترات التقييم التي بدورها يجب أن تكون قادرة على تحقيق الغاية المرجوة منها وهي اتخاذ القرار؛ أي الخوارزميات ذات الأداء الأفضل؟

يُسهّم هذا البحث في تقييم أداء خوارزميات ترميم الصور من خلال تطبيق عدد من هذه الخوارزميات على مجموعة الصور المشوهة نفسها، ثم حساب بارامترات التقييم التقليدية التي تبين أنها غير قادرة على تحقيق التقييم والمقارنة، لذلك تم اقتراح بارامترات تقييم جديدة وهي EPIQA و NCC و MSE بعد ما تبين أنها أعطت نتائج أكثر تميزاً من البارامترات التقليدية.

1-2 بعض الدراسات السابقة:

يعتمد هذا البحث على دراسة ثلاث خوارزميات ترميم صور تختلف في ما بينها بالبنية وأسلوب العمل؛

فمثلاً خوارزمية ترميم الصور بوساطة تعلّم الشبكة المتعاكسة المولّدة للحواف EdgeConnect [11] ذات البنية GAN - تعتمد في ترميم الصور المشوّهة على مرحلتين؛ أولاً ترميم الخطوط وثنائياً تقدير الكثافات اللونية للبكسلات المرممة.

ومن خوارزميات الترميم التي تعمل على مرحلتين أيضاً هي خوارزمية ترميم الصور بوساطة الخطوط العامة والشكل العام للصورة StrucureFlow [12] ذات البنية GAN - إلا أن أسلوب العمل هنا يختلف عن خوارزمية الترميم السابقة؛ حيث تهتم هذه الخوارزمية بتقسيم الخطوط المرممة إلى قسمين (خطوط عامة وخطوط تفصيلية).

أما خوارزمية الترميم بوساطة شبكة الدمج العميق DEFNet [13] ذات البنية CNN - فهي تهتم بمعالجة الحدود الفاصلة بين المنطقتين السليمة والمرممة لحل مشكلة الغباش في هذه المنطقة.

يُسهم هذا البحث في اقتراح ثلاثة بارامترات تقييم جديدة لتقييم أداء خوارزميات الترميم هي:

- مربع متوسط الخطأ MSE الذي اعتمد عليه في تقييم أداء دراسة تهتم بتحسين جودة الصور الملنقطة تحت أسطح البحار والمحيطات [16].
- تقدير جودة الصور بالاعتماد على الحواف والبكسل EPIQA الذي أثبتت الدراسة [23] تفوقه على بارامترتي التقييم PSNR, SSIM حيث يعمل هذا البارامتر على تقييم جودة الصورة المرممة من خلال مقارنتها مع الصورة السليمة من حيث الحواف والكثافات اللونية للبكسلات.
- الترابط المُقيّس NCC الذي يعمل على تقدير جودة الصورة المرممة من خلال حساب الترابط بين الصورتين الأصلية والمرممة [1,2].

1-3 مشكلة البحث:

تعتمد خوارزميات الترميم في تقييم أدائها على إدخال مجموعة ما من الصور المشوّهة، ثم حساب بارامترات التقييم التقليدية لنتائج الترميم التي تبين أنها غير قادرة على تحقيق المقارنة بين خوارزميات الترميم و تعزيز القرار؛ أي من الخوارزميات ذات الأداء الأفضل؟

من أجل تحقيق مقارنة صحيحة بين خوارزميات عدّة لابدّ من إدخال مجموعة الصور ذاتها، ثم حساب بارامترات التقييم لنتائج الخوارزميات، وبناء القرار اعتماداً على قيم تلك البارامترات.

تأتي مشكلة البحث من كون الخوارزميات لا تعتمد في تقييم نتائجها على مجموعة موحدة من الصور، وهذا ما يعطي مقارنة غير صحيحة إضافة إلى ضرورة إدخال بارامترات تقييم جديدة أكثر تميزاً من البارامترات التقليدية.

1-4 الهدف من البحث:

تقييم أداء خوارزميات ترميم الصور الرقمية من خلال اقتراح بارامترات تقييم أكثر تميزاً من البارامترات التقليدية المستخدمة في الدراسات السابقة.

1-5 أهمية البحث:

تأتي أهمية البحث من :

أولاً: التطبيقات الواسعة في مجال ترميم الصور وما لها من أهمية في إعادة إحياء الصور القديمة والتخلص من أنواع التشوهات المختلفة التي تعاني منها الصور.

ثانياً: الضرورة في كون بارامترات التقييم ذات تمييزية وقدرة على تحقيق الغاية المرجوة منها – اتخاذ قرار أيّ من خوارزميات الترميم ذات الأداء الأفضل؟ – وإلا فلا فائدة من هذه البارامترات.

1-6 مسوغات البحث:

العمل على إيجاد بارامترات جديدة لتقييم أداء خوارزميات ترميم الصور لتحل محل البارامترات المستخدمة حالياً، التي لم تعط تمييزية جيدة، ولم تكن مناسبة لاتخاذ القرار.

1-7 محددات البحث:

تناول هذا البحث دراسة بعض خوارزميات الترميم اعتماداً على الإمكانيات الفيزيائية المتاحة، كما جرى تدريب كل خوارزمية اعتماداً على بنية معلومات واحدة تدعى (places2).

1-8 حدود البحث:

- الحد المكاني: جرى تنفيذ هذا البحث في قسم هندسة الإلكترونيات والاتصالات في كلية الهندسة الميكانيكية والكهربائية في جامعة دمشق.
- الحد التقني: جرى تنفيذ خوارزميات الترميم وحساب قيم البارامترات التقليدية والمقترحة باستخدام حاسوب.
- الحد الزمني: استغرق تنفيذ البحث سنتين ونصف منذ التاريخ المحدد في قرار التسجيل.
- الحد العلمي: جرى اتباع المنهج التجريبي في اقتراح بارامترات التقييم، ومن ثم تقييم أداء خوارزميات ترميم الصور.

1-9 منهج البحث وأدواته:

اعتمد في هذا البحث المنهج العلمي البحثي المبني على أسس برمجية وجرى الاستعانة بالبرمجيات والأدوات البرمجية الآتية:

- البرمجية Python 3.7 مفتوحة المصدر.
- البرمجية MATLAB R2019a مفتوحة المصدر.

1-10 بنية الرسالة:

تتضمن بنية الرسالة الفصول الآتية:

- الفصل الأول الإطار العام للبحث.
- الفصل الثاني الأسس النظرية التي بُني عليها البحث.
- الفصل الثالث دراسة بعض خوارزميات الترميم وتبيان بنية عمل كل منها وآليته.
- الفصل الرابع دراسة بارامترات التقييم التقليدية والمقترحة.
- الفصل الخامس التنفيذ العملي والنتائج.

الفصل الثاني

الأساسيات النظرية للبحث

Theoretical Basis of the Research

2-1 تمهيد:

تُعد حاسة الرؤية من أكثر الحواس تطوراً عند الإنسان، وهذا ما جعل للصورة دوراً أساسياً في المعرفة والإدراك لدى البشر.

وعلى الرغم من أن الإبصار لدى البشر محدود بالمجال المرئي من الطيف الكهرومغناطيسي إلا أنه يمكن لآلات التصوير أن تلتقط الصور ضمن هذا المجال بكامله، بدءاً من أشعة غاما وانتهاءً بالأمواج الراديوية وهذا ما جعل لمعالجة الصور الرقمية مجالاً واسعاً من التطبيقات.

تُعرف الصورة أنها تابع ثنائي البعد $f(x,y)$ حيث (x,y) هي الإحداثيات الحيزية (مواقع البكسلات)، وقيمة التابع f عند كل نقطة (x,y) تسمى الكثافة أو المستوى الرمادي للصورة عند تلك النقطة.

عندما تكون قيم كل من (x,y) و قيم التابع f (الكثافة أو المستوى الرمادي) قيماً محدودة (قيم منقطعة) تسمى الصورة بالصورة الرقمية.

للصور الرقمية ثلاثة أنواع هي:

- الصور الثنائية Binary Images: هي صور باللونين الأبيض والأسود، يُعبر عن كل بكسل ب (1-bit) وبالتالي يأخذ كل بكسل إحدى القيمتين 0 (اللون الأسود) أو 1 (اللون الأبيض).

- الصور بالمستوى الرمادي Gray-scale Images: يُعبر عن كل بكسل ب (8-bit)؛ أي تتضمن هذه الصور مستويات رمادية مختلفة عددها 256، تتدرج بدءاً من الأسود (0) وحتى الأبيض (256).

- الصور الملونة Color Images: يُعبر عن كل بكسل من بكسلات الصورة الملونة (24-bit) تقسم إلى ثلاثة أقسام، كل منها (8-bit)، كل قسم يقابل أحد الألوان الأساسية الثلاثة (أحمر وأخضر وأزرق) أو ما يعرف ب RGB.

تسمى العملية التي يتم من خلالها معالجة الصور الرقمية بواسطة الحاسوب الرقمي بمعالجة الصور الرقمية.

تتضمن معالجة الصور الرقمية مجالات كثيرة، هي:

- التقاط الصور Image Acquisition: هي عملية معالجة أولية أساسية لأي عملية أخرى؛ فهي تهدف إلى الحصول على الصور بشكلها الرقمي.
 - تحسين الصور Image Enhancement: هي عملية تعديل الصورة بهدف زيادة جودتها، مثل ضبط النصوص والتباين وإزالة الضجيج.
 - إعادة بناء الصور Image Restoration: هي عملية إعادة بناء الأجزاء الناقصة من الصور المشوّهة أو استعادة المعلومات الناقصة من الصورة بسبب تشوه تعرضت له تلك الصورة.
 - معالجة الصور الملونة Color Image Processing .
 - الضغط Compression: تتضمن عمليات تقليل المساحة اللازمة لتخزين صورة أو تقليل عرض مجال الإرسال اللازم لإرسال الصورة.
 - التقسيم Segmentation: عملية تقسيم الصورة الرقمية إلى أجزاء أو عناصر تستخدم لتحديد مكونات الصورة وحدود تلك المكونات.
 - الوصف Description: هي العملية التي تلي عملية التقسيم حيث يتم استخلاص صفات كل مكون من مكونات الصورة وخصائصه.
 - التعرف على الأشكال Object Recognition: هي العملية التي تلي عملية الوصف، حيث يتم إسناد كل مكون من مكونات الصورة إلى الصنف class الخاص به.
- تُقسم عملية معالجة الصور الرقمية بشكل عام إلى ثلاثة مستويات وهي [9]:
- مستوى المعالجة المنخفض Low-level Process: يتضمن عمليات المعالجة الأولية للصور الرقمية لتحسين التباين وإزالة الضجيج. في هذا المستوى كل من الدخل والخرج عبارة عن صور.
 - مستوى المعالجة المتوسط Mid-Level Process: يتضمن هذا المستوى عمليات تقسيم الصور وصف الصور والتعرف على الأشكال. في هذا المستوى الدخل صورة، أما الخرج فهو وصف لهذه الصورة.
 - مستوى المعالجة العالي High-Level process: يتضمن عمليات فهم محتويات الصورة وإدراكها، وما يرتبط بالرؤية الحاسوبية.

2-2 أنواع الضجيج في الصور الرقمية:

قد تتعرض الصورة الرقمية عند الالتقاط أو في أثناء عمليات المعالجة أو الترميز أو النقل الى أنواع عدة من الضجيج الذي يؤدي إلى انخفاض جودة تلك الصورة. من أنواع الضجيج [24]:

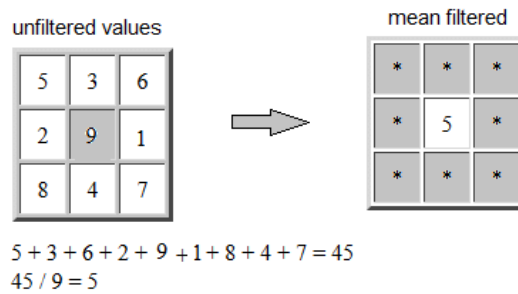
1-2-2 الضجيج الغاوسي Gaussian noise: يعاني كل بكسل من بكسلات الصورة من ضجيج ذو تابع توزيع احتمالي غاوسي.

2-2-2 الضجيج النبضي Impulse noise: هو ضجيج عشوائي يعاني منه عدد من بكسلات الصورة حيث يصبح البكسل الأبيض أسود اللون، والبكسل الأسود أبيض اللون.

3-2-2 الضجيج النقطي Speckle noise: ينتج هذا النوع من الضجيج عند ضرب قيم البكسلات بقيم عشوائية.

3-2 مرشحات التخلص من الضجيج في المجال الحيزي [24]:

1-3-2 المرشح المتوسط: يتم حساب متوسط مجموعة محددة، تُحدد بحجم المرشح، من قيم البكسلات في الصورة واستبدال البكسل بهذه القيمة، وتكرر هذه العملية على كامل الصورة، يبين الشكل (1-2) مثالاً على المرشح المتوسط .

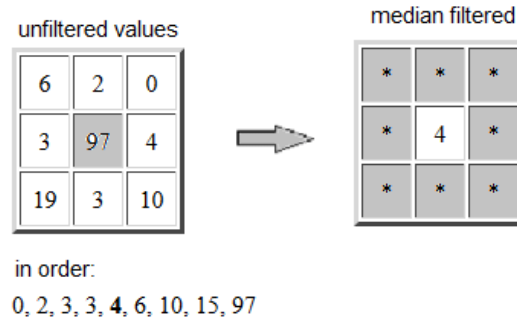


الشكل (1-2) مثال على المرشح المتوسط [20].

2-3-2 المرشح الأصغري: تُستبدل قيمة كل بكسل من بكسلات الصورة بالبكسل ذي القيمة الصغرى من بين مجموعة محددة، تُحدد بحجم المرشح من البكسلات.

3-3-2 المرشح الأعظمي: تُستبدل قيمة كل بكسل من بكسلات الصورة بالبكسل ذي القيمة العظمى من بين مجموعة محددة، تُحدد بحجم المرشح من البكسلات.

4-3-2 المرشح الوسطي: تُستبدل قيمة كل بكسل من بكسلات الصورة بالبكسل ذي القيمة الوسطى من بين مجموعة محددة، تُحدد بحجم المرشح، من البكسلات، يبين الشكل (2-2) مثالاً على المرشح الوسطي.



الشكل (2-2) مثال على المرشح الوسطي [20].

4-2 ترابط الإشارات المتقطعة زمنياً Correlation of discrete-time signals[1]

يحدد الترابط بين إشارتين مدى التشابه بينهما وله كثير من التطبيقات في كثير من المجالات، مثل الاتصالات الرقمية، الرادار، والسونار؛ فعلى سبيل المثال يُحدد وجود الهدف ضمن مجال مراقبة جهاز الرادار من خلال حساب الترابط بين الإشارتين المرسل والمستقبل، إذا كان ناتج حساب الترابط يدل على وجود ترابط بين الإشارتين، دلّ ذلك على أن الإشارة المستقبلية تتضمن نسخة من الإشارة المرسل مزاحة زمنياً ولذلك فإن الهدف موجود-ومقدار الإزاحة الزمنية يحدد المسافة بين الهدف والرادار - أما إذا كان ناتج حساب الترابط لا يدل على وجود ترابط بين الإشارتين المرسل والمستقبل دلّ ذلك على أن الإشارة المستقبلية هي عبارة عن إشارة ضجيج فقط ولا وجود للهدف في المنطقة المراقبة.

أنواع الترابط:

- الترابط المتبادل Crosscorrelation: يمثل مقدار التشابه بين إشارتين مختلفتين $x[n]$, $y[n]$ ويعطى بالعلاقة (1-2) [1]:

$$r_{x,y}(l) = \sum_{n=-\infty}^{n=+\infty} x[n] \times y[n-l] \quad (1-2)$$

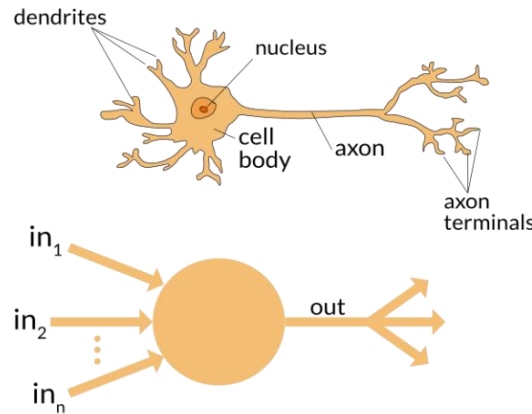
حيث $l = 0, \pm 1, \pm 2, \pm 3, \dots$ هو مقدار إزاحة الإشارة $y[n]$.

- الترابط الذاتي Autocorrelation: يمثل مقدار الترابط بين إشارتين متماثلتين إحداها متأخرة زمنياً عن الأخرى بمقدار l ويعطى بالعلاقة (2-2) [1]:

$$r_{x,x}(l) = \sum_{n=-\infty}^{n=+\infty} x[n] \times x[n-l] \quad (2-2)$$

5-2 الشبكات العصبونية (NN) Neural Networks [19]:

تعرف الشبكات العصبونية (NN) Neural Network، التي تسمى أيضاً الشبكة العصبونية الذكية (Artificial Neural Network (ANN)، أنها نظام تعلم حاسوبي يستخدم مجموعة من التتابع من أجل فهم المعلومات المدخلة وتحويلها من نمط ما إلى نمط آخر. تشابه العصبونات Neurons - التي هي عبارة عن الحجر الأساس في بناء الشبكات العصبونية- من حيث بنيتها ومبدأ عملها عصبونات الدماغ البشري التي تعمل معاً وفق آلية معينة من أجل فهم المعلومات الواردة إليها من الحواس المختلفة. يبين الشكل (2-3) التشابه بين عصبون الشبكة العصبونية وعصبون الدماغ البشري.



الشكل (2-3) التشابه بين عصبون الشبكة العصبونية وعصبون الدماغ البشري [8].

تعد الشبكات العصبونية إحدى الأدوات المستخدمة في خوارزميات تعليم الآلة؛ فهي تعمل على معالجة المعلومات المدخلة المعقدة بحيث يمكن للحاسوب التعامل معها. لا تعتمد خوارزميات تعليم الآلة المستخدمة للشبكات العصبونية على البرمجة وفق قواعد و خطوات يتحدد معها مسار العمل إنما تعتمد على التعلم من عدد من الأمثلة المدخلة حيث تعمل الشبكة العصبونية على استخلاص خصائص تلك الأمثلة المدخلة وصفاتها من أجل التوصل إلى الخرج الصحيح.

عندما يتم تزويد الشبكة العصبونية بالعدد الكافي من الأمثلة - حيث يؤدي هذا العدد إضافةً إلى تنوع الأمثلة دوراً أساسياً في دقة النتائج - وإنهاء عملية التدريب تصبح تلك الشبكة العصبونية جاهزة لمعالجة مدخلات أخرى جديدة لم تكن ضمن أمثلة التدريب والحصول على النتائج الصحيحة.

للشبكات العصبونية مجال واسع من التطبيقات كالتعرف على الوجوه والأشكال والإسهام في تنبؤات الطقس والكشف عن الأمراض والسرطانات.

2-5-1 الطبقات المكونة للشبكات العصبونية:

الشبكات العصبونية عموماً تتكون من ثلاثة أنواع من الطبقات:

- طبقة الدخل input layer: تعمل هذه الطبقة على استقبال المعلومات وتمثيلها إلى الطبقات التالية.
- الطبقة المخفية hidden layer: يمكن أن تتضمن الشبكات العصبونية طبقة واحدة مخفية أو أكثر، هذه الطبقة - أو الطبقات - هي المسؤولة عن إجراء العمليات على البيانات الواردة من طبقة الدخل أو الطبقات المخفية السابقة، كما أنها تحدد مدى تعقيد وكفاءة الشبكة العصبونية.

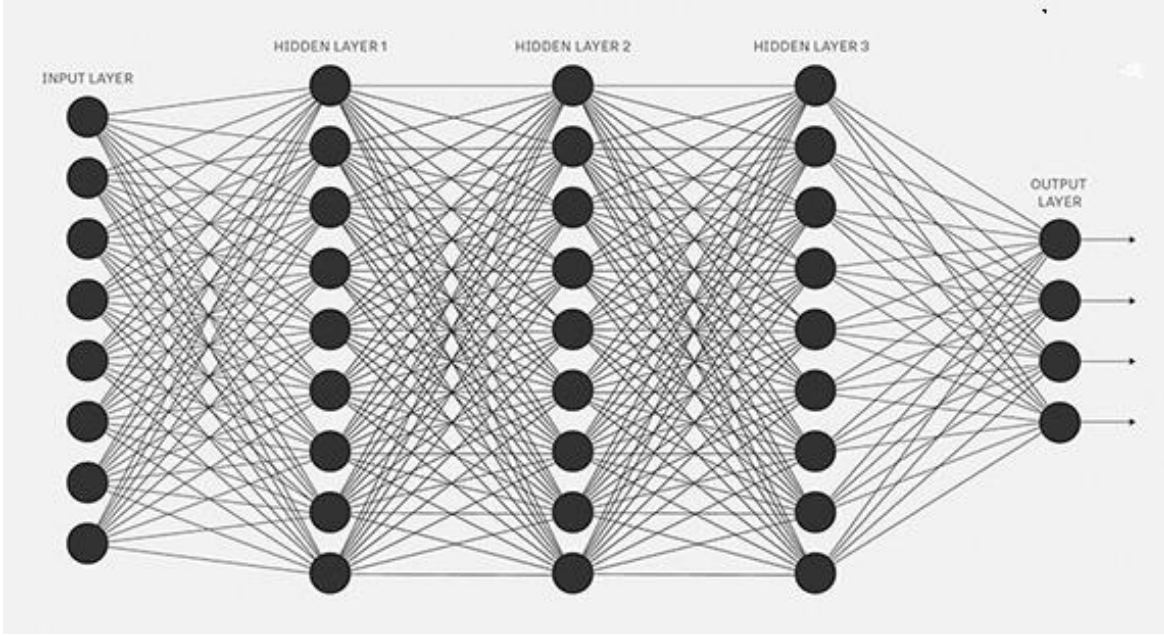
الشبكة العصبونية ذات الطبقات المخفية المتعددة (ثلاث طبقات فأكثر) تسمى الشبكة العصبونية عميقة التعلم deep learning neural network.

- طبقة الخرج output layer: هي الطبقة المسؤولة عن إظهار النتائج.

يوضح الشكل (2-4) مثلاً على بنية شبكة عصبونية مكونة من طبقة دخل وطبقة خرج وثلاث طبقات مخفية.

2-5-2 بنية العصبون:

تتكون كل طبقة من طبقات الشبكة العصبونية من عدد من العصبونات التي تسمى أيضاً بالعقد.

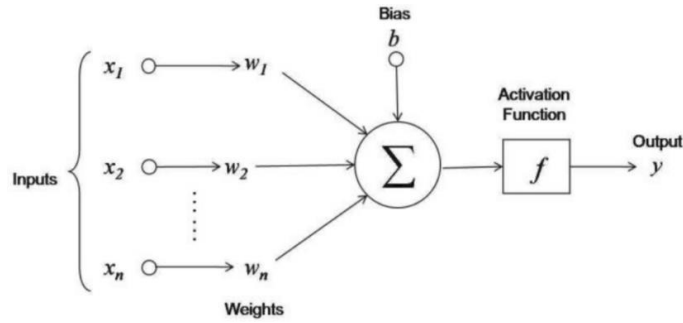


الشكل (2-4) بنية الشبكة العصبونية [19].

لكل عصبون neuron ثلاث مهام هي:

- استقبال الإشارات الواردة من عصبونات الطبقة السابقة.
- إجراء عمليات حسابية على الإشارات المستقبلة.
- إرسال الإشارة إلى العصبونات في الطبقة التالية.

يبين الشكل (2-5) بنية العصبون.



الشكل (2-5) بنية العصبون [10].

يعمل كل عصبون على استقبال الإشارات الواردة من العصبونات السابقة، ثم تُضرب كل إشارة بقيمة تسمى الوزن (W) Weight تحدد هذه القيمة مدى أهمية الإشارة الواردة ثم تُجمع قيم الإشارات الواردة كلها بعد ضربها بالأوزان الخاصة بكل منها وإزاحتها بقيمة محددة تسمى قيمة الانزياح (b) Bias ثم يعمل تابع التنشيط Transfer function , Activation function على تحديد قيمة خرج هذا العصبون ليمرر إلى عصبونات الطبقة التالية.

يعبر عن ذلك رياضياً بالعلاقة (3-2) [10]:

$$y = f\left(\sum_{i=1}^{i=n} x_i * w_i\right) + b \quad (3-2)$$

حيث: b قيمة الانزياح، i هو عدد إشارات الدخل للعصبون، x_i إشارة الدخل، w_i الوزن الخاص بإشارة الدخل x_i ، f هو تابع التنشيط، y إشارة خرج العصبون التي ستكون إشارة دخل لعصبونات الطبقة التالية.

يتم اختيار القيم المناسبة لكل من الأوزان والانزياح أثناء عملية تدريب الشبكة العصبونية بحيث تعطي النتائج المطلوبة وبخطأ يقترب قدر الإمكان من الصفر.

3-5-2 تدريب الشبكات العصبونية:

تتبع عملية تدريب الشبكات العصبونية في كل دورة تدريب epoch مسارين هما:

- المسار الأمامي Forward pass.

- المسار الخلفي Backward pass.

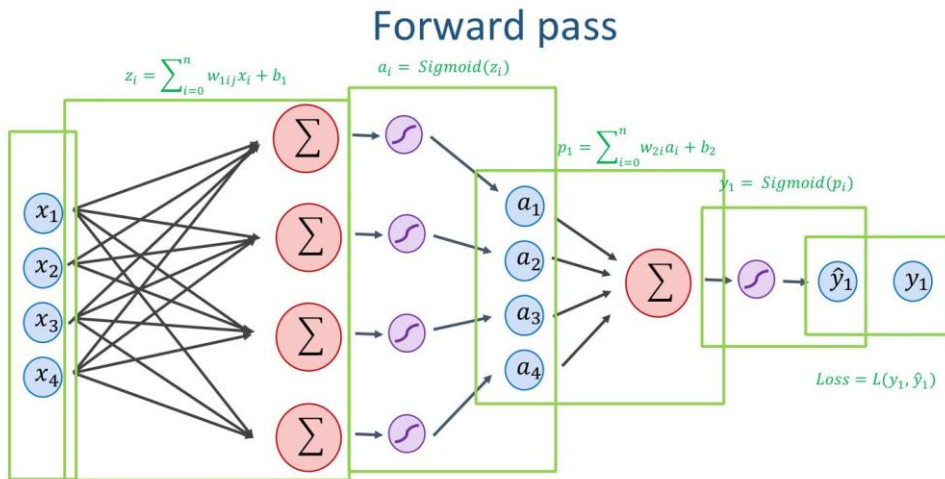
في المسار الأمامي يتم حساب قيمة خرج كل عصبون من عصبونات الطبقة الأولى اعتماداً على قيم أوزان الإشارات الواردة إلى كل عصبون (تسند إلى الأوزان عند أول دورة تدريب قيم ابتدائية يتم اختيارها عشوائياً بحيث تكون قيمة صغيرة جداً) وتابع التنشيط الخاص به، حيث يتم حساب تابع التنشيط لمجموع الأوزان المضروبة مع إشارات الدخل الخاصة بكل وزن، تُمرر الإشارات الناتجة عن كل عصبون من عصبونات الطبقة الأولى إلى عصبونات الطبقة الثانية ثم تُحسب قيم إشارات خرج كل عصبون من عصبونات الطبقة الثانية بالطريقة نفسها، ثم تُدخل إلى عصبونات الطبقة الثالثة، وهكذا حتى الوصول إلى طبقة الخرج والحصول على الناتج المطلوب. يبين الشكل (2-6) مثلاً يوضح المسار الأمامي.

في المسار الخلفي يتم تقييم ناتج الشبكة العصبونية من خلال تابع الخطأ الذي يحدد مدى اختلاف ناتج الشبكة العصبونية عن الناتج الصحيح، ثم تعديل قيم الأوزان ضمن الشبكة العصبونية بحيث تحقق اقتراب نتائج الشبكة من النتائج الصحيحة المطلوبة حيث يتم ذلك اعتماداً على قيم المشتق الأول الذي يحسب لكل طبقة من طبقات الشبكة، يبين الشكل (2-7) مثلاً يوضح المسار الخلفي.

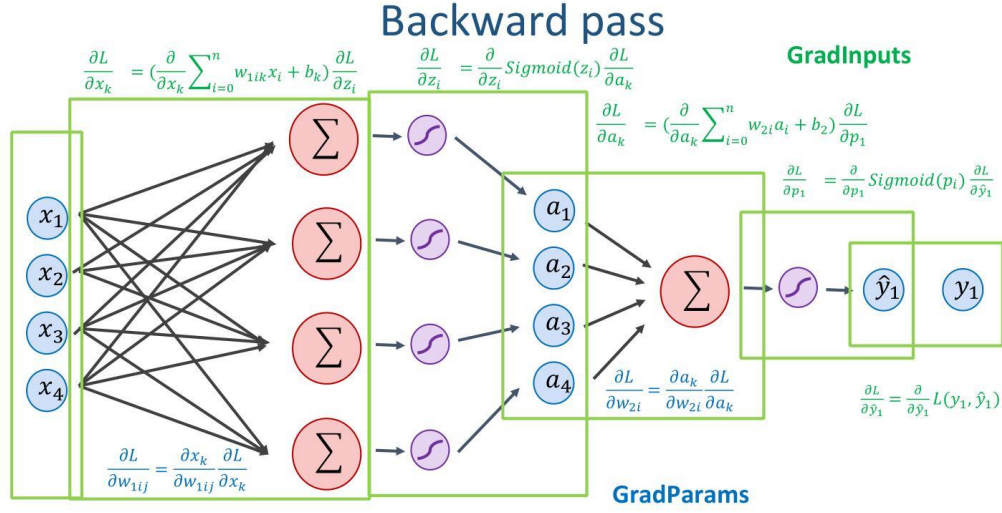
تتكرر عملية تدريب الشبكة العصبونية دورات عدة epochs لتحقيق النتائج المطلوبة.

2-5-4 تابع التنشيط Activation function:

هو تابع يتم اختياره عند تصميم الشبكة العصبونية لكل من طبقتي الخرج والطبقة المخفية بهدف جعل الشبكة العصبونية شبكة لاخطية .



الشكل (2-6) مثال على المسار الأمامي [8].



الشكل (2-7) مثال على المسار الخلفي [8].

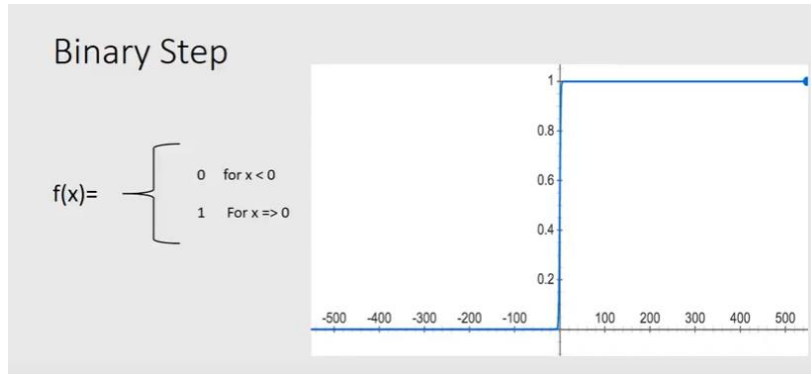
تأتي أهمية تابع التنشيط مما يأتي:

- يُدخل إلى كل عصبون ناتج مجموع الأوزان المضروبة مع إشارات الدخل كلها، ولذلك يمكن أن تكون هذه القيمة ضمن المجال $[-\infty, +\infty]$ ، يعمل تابع التنشيط على تحديد قيمة خرج العصبون؛ أي وبمعنى آخر أي العصبونات ستكون مفعلة، وأي العصبونات ستكون غير مفعلة.
- يعمل تابع التنشيط على حل مشكلة التعقيد الحسابي في الشبكات العصبونية العميقة الناتج عن تراكم مجموع الإنزياح مع الأوزان المضروبة بإشارات الدخل من طبقة إلى أخرى حيث يعمل تابع التنشيط على تحديد قيمة خرج كل عصبون بقيم محددة؛ فمثلاً يعمل تابع التنشيط softmax على تحديد قيم خرج كل عصبون ب 0 أو 1.
- تجعل اللاحظية المضافة إلى الشبكة العصبونية، بفضل تابع التنشيط، الشبكة قادرة على فهم والتعامل مع المعلومات المعقدة؛ فمثلاً الصور هي معلومات لاختية لا يمكن فهمها والتعامل معها بواسطة شبكة عصبونية خطية.
- تُسهم لاختية الشبكة العصبونية في عملية تدريب الشبكة في أثناء المسار الخلفي الذي يتم فيه تعديل قيم الأوزان بناءً على مشتق تابع التنشيط في كل مرحلة، بحال كانت الشبكة خطية سيكون المشتق عدداً ثابتاً، ولذلك لن يتم تعديل قيم الأوزان.

بعض أنواع توابع التنشيط شائعة الاستخدام [22]:

- تابع الخطوة الثنائي Binary step function:

يُفعل العصبون إذا كانت قيمة دخل تابع التنشيط أكبر من قيمة العتبة ولا يُفعل فيما عدا ذلك. يبيّن الشكل (2-8) تابع الخطوة.

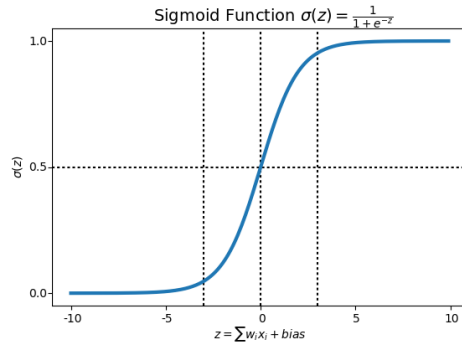


الشكل (2-8) تابع الخطوة الثنائي [17,22].

- تابع التنشيط السيني Sigmoid function:

هو تابع لا خطي- يشبه الحرف S- يأخذ قيماً بين 0 و 1 ؛ ولذلك هو مناسب لعصبونات طبقة الخرج في الشبكات العصبونية التي تعمل كمصنف ثنائي.

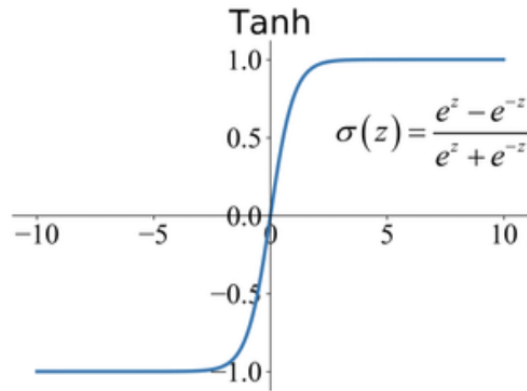
نلاحظ من الشكل (2-9) أن تابع التنشيط Sigmoid function يتحسس بصورة جيدة من أجل قيم الدخل بين 2 و -2 في حين تصبح أقل من أجل قيم الدخل الأخرى، ومن ثم من أجل قيم دخل كبيرة تصبح قيمة التابع قيمة ثابتة، ومن ثم مشتق التابع عند تلك القيم مساوٍ للصفر، ما يعرف باختفاء الأوزان، وعليه يصبح التدريب غير ممكن من أجل قيم دخل كبيرة. يعاني تابع التحويل Sigmoid من مشكلة أخرى وهي أنه غير متناظر بالنسبة إلى المبدأ ومن ثم تكون كل العصبونات موجبة فقط.



الشكل (9-2) تابع التنشيط Sigmoid [17,22].

- تابع الظل القطعي Tanh:

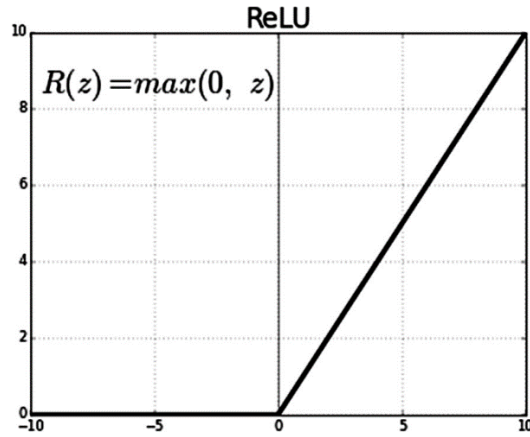
يبين الشكل (10-2) تابع التنشيط Tanh وهو يشابه تابع التنشيط Sigmoid إلا أنه متناظر بالنسبة إلى للمبدأ ومن ثم تأخذ العصبونات قيماً موجبة وسالبة.



الشكل (10-2) تابع التنشيط Tanh [17,22].

- تابع الوحدة الخطية المُقَوِّم Rectified Linear Unit (Relu):

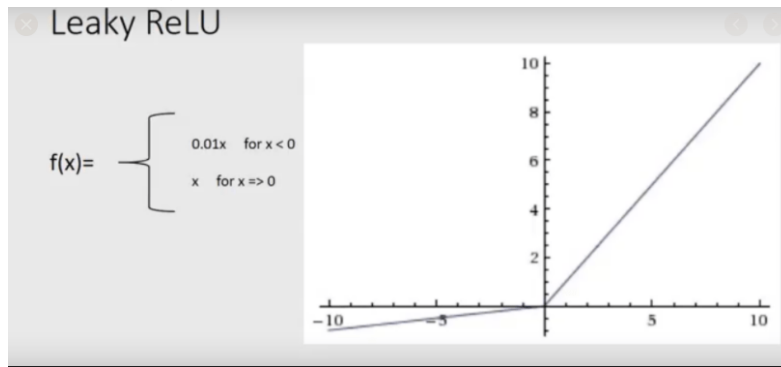
يعمل تابع التنشيط Relu على تفعيل بعض العصبونات - العصبونات ذات القيم الموجبة فقط - وهذا ما يقلل من التعقيد الحسابي في الشبكة العصبونية. يبين الشكل (11-2) تابع التنشيط Relu. تعاني الشبكة التي تستخدم تابع التنشيط Relu في أثناء المسار الخلفي - مرحلة تعديل قيم الأوزان - من مشكلة تعديل الأوزان في العصبونات ذات القيم الموجبة فقط، أما العصبونات الأخرى فلن تُعدل قيم أوزانها؛ وذلك لأن مشتق تابع التحويل مساوٍ للصفر، ولذلك ستبقى العصبونات ذات القيم السالبة ستبقى غير مفعلة.



الشكل (11-2) تابع التنشيط Relu [17,22] .

- تابع الوحدة الخطية المنفرج (Leaky Relu (LRelu:

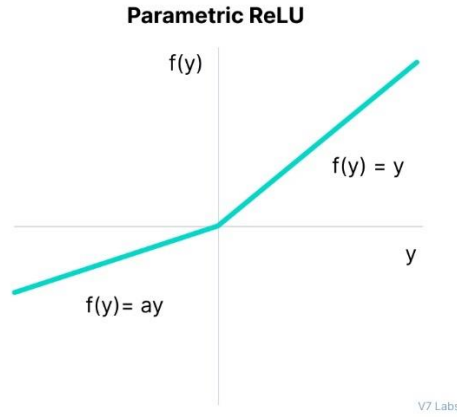
يبين الشكل (12-2) تابع التنشيط LRelu. وهو تابع تنشيط مشابه لتابع التنشيط RELU استخدم لجعل مشتق تابع تنشيط العصبونات السالبة غير مساو للصفر.



الشكل (12-2) تابع التنشيط LRelu [17,22].

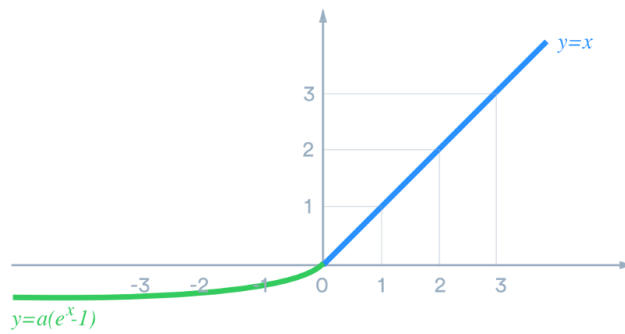
- تابع الوحدة الخطية البارامتري (Parameteric Relu(PRelu):

يبين الشكل (13-2) التابع PRelu. هو تابع Relu معدّل من أجل حل مشكلة المشتقات الصفرية في الجزء السالب من التابع، يضيف هذا التابع بارامتراً جديداً وهو ميل التابع a في الجزء السالب منه تحدد قيمة الميل في أثناء عملية التدريب.



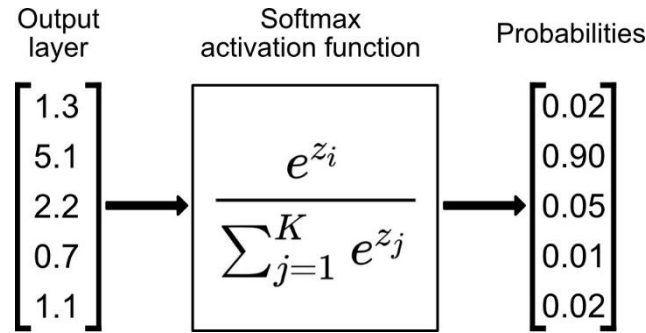
الشكل (2-13) تابع التنشيط PRelu [17,22].

- تابع الوحدة الخطية الأسية (Exponential Linear Unit (ELU):
يبين الشكل (2-14) التابع ELU . هو تابع Relu معدل يجعل الجزء السالب منه عبارة عن تابع أسي للتخلص من المشتق الصفري في الجزء السالب.



الشكل (2-14) تابع التنشيط ELU [17,22].

- تابع التنشيط الاحتمالي Softmax function:
هو عبارة عن مجموع توابع عدة Sigmoid يستخدم في طبقة الخرج للشبكات العصبونية التي تعمل عمل مصنف متعدد الأصناف، إذ يعطي التابع Softmax قيمة الاحتمال الخاصة بكل صنف.
في الشبكات العصبونية التي تعمل عمل مصنف متعدد تتضمن طبقة الخرج عدداً من العصبونات مساوٍ لعدد الأصناف classes، يعمل التابع Softmax على حساب قيمة الاحتمال لكل صنف. يبين الشكل (2-15) تطبيق التابع Softmax لحساب احتمال كل صنف في شبكة عصبونية تتضمن خمسة أصناف في طبقة الخرج.



الشكل (2-15) مثال على حساب احتمال كل صنف بواسطة تابع التنشيط softmax [21].

6-2 الشبكة العصبونية الالتفافية (CNN) Convolutional Neural Network

: [17-18]

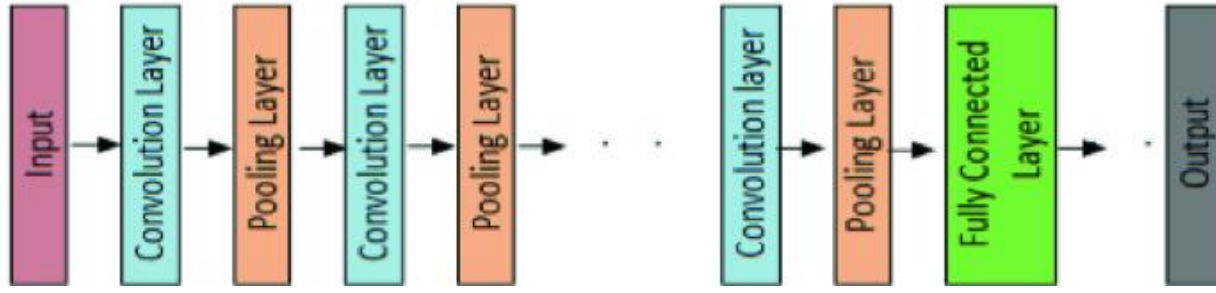
هي نوع خاص من أنواع الشبكات العصبونية متعددة الطبقات، أو بنية تعلم عميق مُلهمة من نظام الرؤية عند المخلوقات الحية مناسبة لتطبيقات الرؤية الحاسوبية تتميز بالقدرة على استخلاص الصفات والخصائص الخاصة بإشارة الدخل كالصور مثلاً بكفاءة عالية من دون الحاجة لعمليات معالجة أولية.

تتكون من عدد محدد من الطبقات تعمل كل منها على استخلاص خصائص إشارة الدخل؛ حيث تعمل الطبقات الأولى على استخلاص الخصائص عالية المستوى High Level Features في حين تعمل الطبقات الأكثر عمقاً على استخلاص الخصائص منخفضة المستوى Low Level Features.

تتضمن بنية الشبكة العصبونية الالتفافية عموماً خمسة أنواع من الطبقات هي:

- طبقة الدخل Input layer.
- طبقة الطي Convolutional layer.
- طبقة التجميع Pooling layer.
- طبقة الترابط الكلي A Fully connected layer.
- طبقة الخرج Output layer.

يبيّن الشكل (2-16) مثلاً على بنية الشبكة CNN.



الشكل (2-16) مثال على بنية CNN [17].

2-6-1 طبقة الطي Convolutional layer:

تُعد هذه الطبقة من المكونات الأكثر أهمية في شبكة CNN، تتضمن مجموعة من مرشحات الطي (convolutional kernels (or filters) التي تُجمع مع الصورة المدخلة للحصول على خريطة الخصائص feature map.

المرشح Filter or Kernel هو عبارة عن مصفوفة تتضمن أعداداً يسمى كل منها وزن Weight، عند البدء بعملية تدريب الشبكة CNN تُعطى لهذه الأوزان قيم عشوائية، تُعدل هذه القيم عند كل دورة تدريب للشبكة CNN لتصبح أوزان المرشح قادرة على استخلاص خصائص ذات معنى، حيث إن كل مرشح يمكن أن يستخلص خصائص مختلفة عن مرشح آخر.

عملية الطي: هي عملية تجميع المرشح مع الصورة المدخلة بهدف استخلاص الخصائص منها، تتم هذه العملية عن طريق ضرب كل قيمة من قيم المرشح مع القيمة المقابلة لها من الصورة المدخلة، ثم جمع نتائج الضرب للحصول على ناتج عملية الطي الأولى، وتكرر هذه العملية على كامل الصورة من خلال إزاحة المرشح وحساب قيمة ناتج الطي، وهكذا على كامل الصورة، وبهذا يتم الحصول على الخصائص المستخلصة من الصورة. يبين الشكل (2-17) مثلاً على عملية الطي مطبقة على صورة رمادية أبعادها 4×4 ، ومرشح أبعاده 3×3 .

تسمى المنطقة من الصورة التي يطبق عليها المرشح receptive field، يسمى مقدار انزياح المرشح stride وهو في المثال السابق مقداره بكسل واحد؛ أي يُزاح المرشح بمقدار بكسل واحد وكلما زادت قيمة الانزياح كلما كانت خريطة الخصائص الناتجة ذات أبعاد أقل.

ملاحظة: بحال كانت الصورة المدخلة صورة ملونة يُطبق المرشح على كل من القيم R,G,B، ثم يجمع النتائج للحصول على خريطة الخصائص الخاصة بتلك الصورة الملونة.

Input Matrix	Kernel	Result																													
<table><tr><td>45</td><td>12</td><td>5</td><td>17</td></tr><tr><td>22</td><td>10</td><td>35</td><td>6</td></tr><tr><td>88</td><td>26</td><td>51</td><td>19</td></tr><tr><td>9</td><td>77</td><td>42</td><td>3</td></tr></table>	45	12	5	17	22	10	35	6	88	26	51	19	9	77	42	3	<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>-45</td><td>12</td></tr><tr><td>22</td><td>10</td></tr></table>	-45	12	22	10
45	12	5	17																												
22	10	35	6																												
88	26	51	19																												
9	77	42	3																												
0	-1	0																													
-1	5	-1																													
0	-1	0																													
-45	12																														
22	10																														
<table><tr><td>45</td><td>12</td><td>5</td><td>17</td></tr><tr><td>22</td><td>10</td><td>35</td><td>6</td></tr><tr><td>88</td><td>26</td><td>51</td><td>19</td></tr><tr><td>9</td><td>77</td><td>42</td><td>3</td></tr></table>	45	12	5	17	22	10	35	6	88	26	51	19	9	77	42	3	<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>-45</td><td>103</td></tr><tr><td>22</td><td>10</td></tr></table>	-45	103	22	10
45	12	5	17																												
22	10	35	6																												
88	26	51	19																												
9	77	42	3																												
0	-1	0																													
-1	5	-1																													
0	-1	0																													
-45	103																														
22	10																														
<table><tr><td>45</td><td>12</td><td>5</td><td>17</td></tr><tr><td>22</td><td>10</td><td>35</td><td>6</td></tr><tr><td>88</td><td>26</td><td>51</td><td>19</td></tr><tr><td>9</td><td>77</td><td>42</td><td>3</td></tr></table>	45	12	5	17	22	10	35	6	88	26	51	19	9	77	42	3	<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>-45</td><td>103</td></tr><tr><td>-96</td><td>10</td></tr></table>	-45	103	-96	10
45	12	5	17																												
22	10	35	6																												
88	26	51	19																												
9	77	42	3																												
0	-1	0																													
-1	5	-1																													
0	-1	0																													
-45	103																														
-96	10																														
<table><tr><td>45</td><td>12</td><td>5</td><td>17</td></tr><tr><td>22</td><td>10</td><td>35</td><td>6</td></tr><tr><td>88</td><td>26</td><td>51</td><td>19</td></tr><tr><td>9</td><td>77</td><td>42</td><td>3</td></tr></table>	45	12	5	17	22	10	35	6	88	26	51	19	9	77	42	3	<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>-45</td><td>103</td></tr><tr><td>-176</td><td>133</td></tr></table>	-45	103	-176	133
45	12	5	17																												
22	10	35	6																												
88	26	51	19																												
9	77	42	3																												
0	-1	0																													
-1	5	-1																													
0	-1	0																													
-45	103																														
-176	133																														

الشكل (2-17) مثال على عملية الطي على صورة رمادية [17].

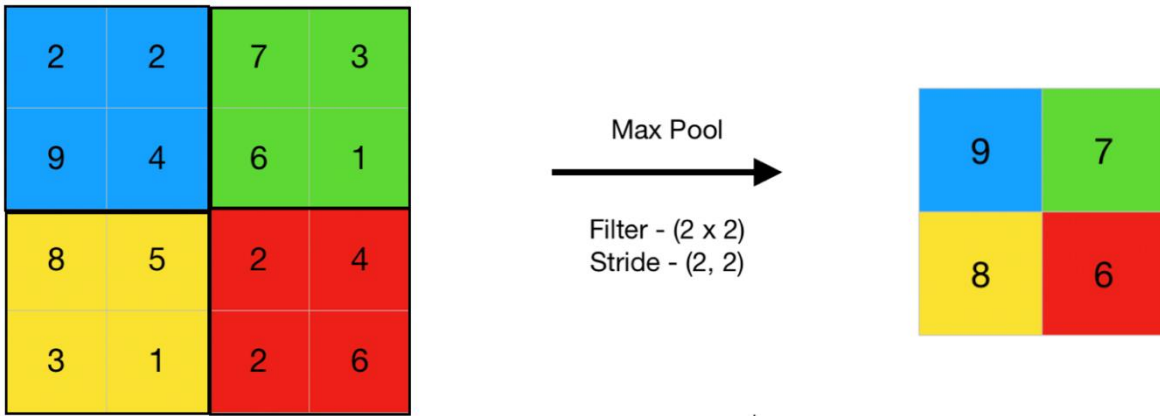
2-6-2 طبقة التجميع Pooling Layer:

هي طبقة تلي دوماً طبقة الطي تعمل على تقليص حجم خريطة الخصائص (الناتجة عن عملية الطي) بهدف تقليل البارامترات اللازمة للمعالجة، وتقليل عمليات المعالجة المطبقة، وتقليل سعة التخزين. تعمل هذه الطبقة على تقليل حجم خريطة الخصائص من خلال حذف جزء منها والإبقاء على الخصائص الأكثر تأثيراً، ومن ثم التركيز عليها، وذلك بهدف التعرف على خصائص العناصر المكونة للصورة بصرف النظر عن مكان وجودها بالصورة ومن ثم إمكانية التعرف على العناصر مع تغير مواقعها في الصورة أو حتى كون الصورة مقلوبة.

تُطبق هذه العملية بطريقة مشابهة لطبقة الطي؛ حيث يطبق مرشح على الصورة، ثم يزاح بمقدار Stride ليشمل كامل الصورة.

هناك كثير من أنواع تقانات التجميع منها: التجميع الأعظمي، والتجميع الأصغري، والتجميع المتوسط.

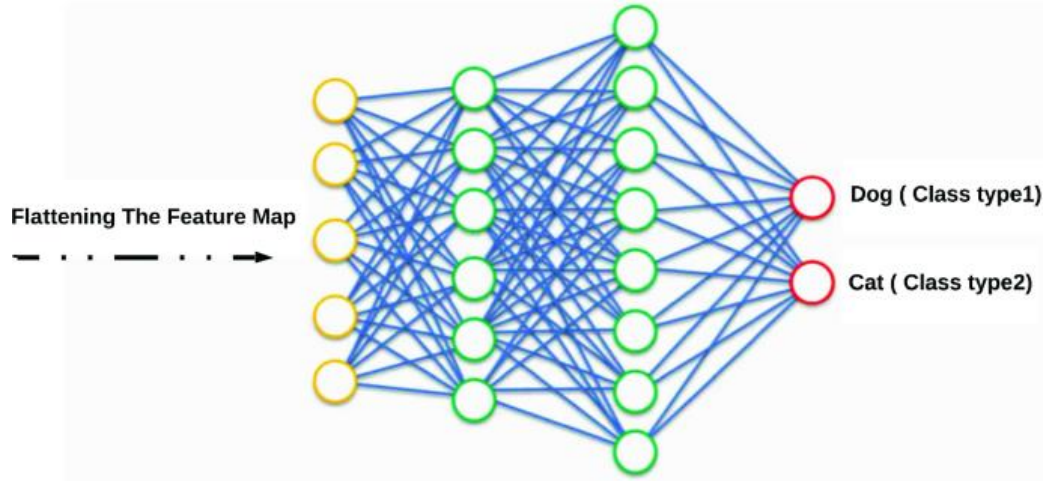
تُعد تقانة التجميع الأعظمي من أكثر أنواع التقانات المستخدمة في شبكات CNN حيث يتم الإبقاء على البكسل ذي القيمة الأعلى وحذف ما دون ذلك يبين الشكل (2-18) مثلاً يوضح عملية التجميع الأعظمي مطبقة على صورة رمادية بمرشح أبعاده 2×2 وإزاحة بمقدار بكسلين.



الشكل (2-18) مثال على عملية التجميع الأعظمي [18].

3-6-2 طبقة الترابط الكلي Fully Connected (FC) Layer:

تتكون هذه الطبقة من شبكة عصبونية NN واحدة أو أكثر، أي إن كل عصبون مربوط مع كامل عصبونات الطبقة التالية، تعمل هذه الطبقة عمل مصنف Classifier دخلها هو مصفوفة الخصائص الناتجة عن آخر طبقة طي أو آخر عملية تجميع بعد تحويلها إلى مصفوفة أبعادها $1 \times n$ (vector) تُربط الطبقة FC مع طبقة الخرج المسؤولة عن إظهار نتيجة التصنيف (إما نوع الصف الذي تنتمي إليه الصورة المدخلة أو احتمال إنتماء الصورة المدخلة إلى صف ما) يبين الشكل (2-19) مثلاً عن الطبقة الترابط الكلي .



الشكل (2-19) مثال على طبقة الترابط الكلي [17].

تتميز شبكة العصبونية الالتفافية CNN عن الشبكة العصبونية ANN أنه في الشبكة العصبونية يُربط كل عصبون من كل طبقة مع جميع العصبونات في الطبقة التالية بخلاف الشبكة العصبونية الالتفافية التي يُربط فيها كل عصبون مع عدد محدود من عصبونات الطبقة التالية مما يقلل عدد البارامترات اللازمة للتدريب.

في الشبكة العصبونية ANN وفي أثناء عملية التدريب تحتاج كل وصلة بين عصبونين إلى تحديد قيمة الوزن الخاص بها أما في شبكة العصبونية الالتفافية أثناء عملية التدريب يتم ضبط قيم الأوزان نفسها لكل عصبونات الدخل (التي هي قيم مرشح الطي) وهذا ما يقلل الزمن اللازم للتدريب وسعة التخزين اللازمة لتخزين الأوزان.

7-2 البنية U-Net:

هي بنية تستخدم الشبكة العصبونية الالتفافية تعمل على تقسيم الصورة على مستوى البكسل semantic segmentation؛ أي تعمل هذه البنية على إسناد كل بكسل من بكسلات الصورة إلى الصنف الخاص به استُخدمت أولاً في معالجة الصور الطبية [7] ثم أصبح لها كثير من التطبيقات مثل المركبات ذاتية القيادة التي يتوجب عليها مراقبة المسار الذي ستسير عليه، والتعرف على الأجسام التي ستصادفها والتمييز بينها وتحدد ما إذا كان سيتوجب عليها التوقف أو الإنعطاف أو متابعة السير، يوضح الشكل (2-20) مفهوم التقسيم على مستوى البكسل. تتكون بنية U-Net من فرعين متناظرين، كل منهما هو عبارة عن شبكة عصبونية التفافية يشكّلان معاً بنية تشابه الحرف U، يوضح الشكل (2-21) بنية الشبكة U-Net.

يسمى الفرع الأول مرمز encoder or contraction path يعمل على استخلاص الخصائص من الصورة ثم فهم محتوياتها، في هذا الفرع ستقل أبعاد الصورة تدريجياً، أما الفرع الثاني يسمى مفكك الترميز decoder or expanding path وهو معاكس للفرع الأول حيث يعمل على تحديد مواقع الأجسام التي تم التعرف عليها ضمن الفرع الأول، في هذا الفرع ستزداد أبعاد الصورة تدريجياً لتصبح مساوية لأبعاد الصورة المدخلة إلى الفرع الأول.



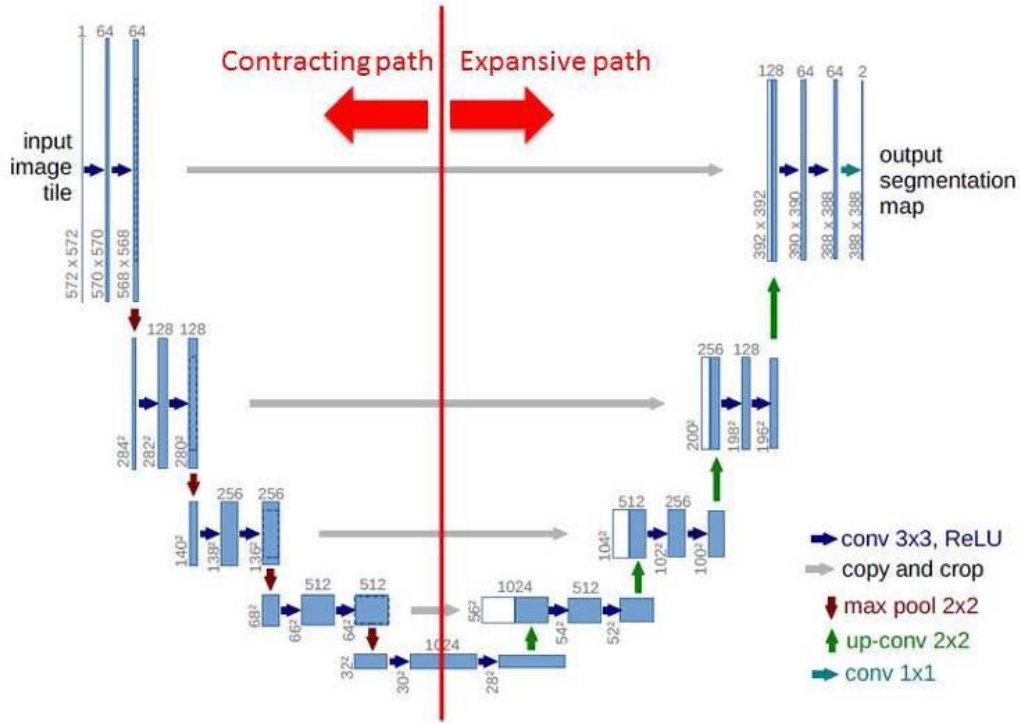
الشكل (2-20) مثال على مفهوم التقسيم على مستوى البكسل [14].

يلاحظ من الشكل (2-21) ما يلي:

- الفرع الأيسر (المرمّز) هو عبارة عن شبكة CNN تتضمن طبقات طي عدّة تستخدم كل عملية طي مرشحاً بأبعاده 3×3 بهدف استخلاص الخصائص من الصورة المدخلة. يلي طبقة الطي طبقة التجميع الأعظمي التي تستخدم مرشحاً بأبعاده 2×2 بهدف تقليص أبعاد خريطة الخصائص الناتجة عن طبقة الطي السابقة. في كل عملية طي يتم مضاعفة عدد المرشحات المستخدمة بهدف التوصل إلى تنوع أكبر في الخصائص المستخلصة.

- الفرع الأيمن (مفكك الترميز) معاكس تماماً للفرع الأيسر من حيث البنية والعمل فهو يتكون من عدد من طبقات معاكسة للطّي تهدف إلى تحديد مواضع الأجسام التي تم التعرف إليها في الفرع الأول، تلي كل منها طبقة تعمل على زيادة أبعاد خريطة الخصائص حتى التوصل إلى صورة أبعادها مساوية تماماً لأبعاد صورة الدخل وتتضمن ناتج تقسيم الصورة المدخلة على مستوى البكسل.

Network Architecture



الشكل (21-2) بنية الشبكة U-Net [7].

8-2 الشبكة المتعاكسة المولدة (GAN) Generative Adversarial Network (GAN) [6]:

هي بنية مكونة من شبكتين عصبونيتين تعمل إحداهما عمل معاكس للأخرى تهدف إلى توليد معلومات جديدة سواء كانت هذه المعلومات صوراً أو فيديو أو صوتاً، يوضح الشكل (22-2) البنية المتعاكسة المولدة.

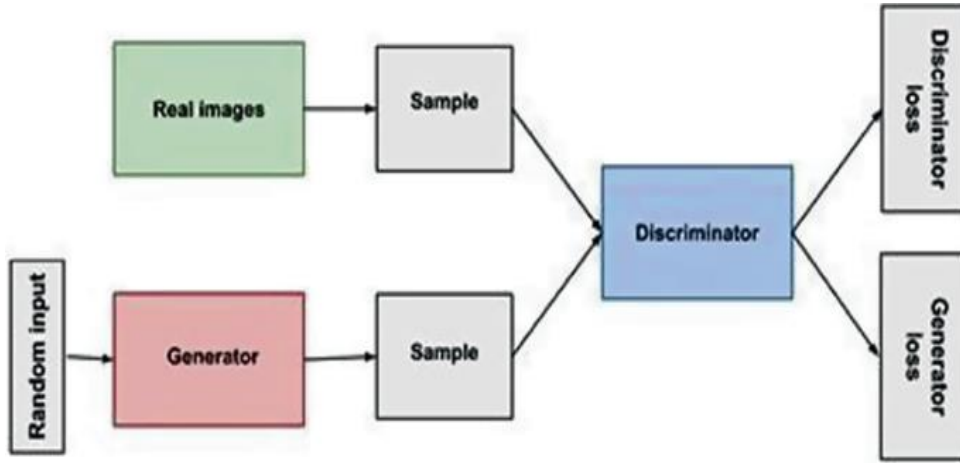
تتكون البنية GAN من قسمين:

- المولد (G) Generator: هو شبكة عصبونية تُدرب على توليد معلومات جديدة صحيحة؛ أي مشابهة بأكبر قدر ممكن لمثيلاتها من المعلومات الواقعية.

- المميز (D) Discriminator : هو شبكة عصبونية مهمتها التمييز بين المعلومات المولدة من المولد (تسمى بالمعلومات المزيفة) والمعلومات الحقيقية بهدف دفع المولد إلى توليد معلومات صحيحة تطابق المعلومات الحقيقية.

عند بدء عملية التدريب يعمل المولد على توليد معلومات عشوائية مزيفة يمكن للمميز وبكل سهولة التمييز بينها وبين المعلومات الحقيقية.

مع استمرار عملية التدريب واقترب المولد من توليد معلومات مشابهة للمعلومات الحقيقية تصبح عملية التمييز بينها وبين المعلومات الحقيقية أكثر تعقيداً، وهكذا حتى يصبح المميز غير قادر على التمييز بين المعلومات المولدة من المولد (المعلومات المزيفة) والمعلومات الحقيقية، عندها تنتهي عملية التدريب، وتصبح البنية GAN قادرة على توليد المعلومات الحقيقية المطلوبة.



الشكل (2-22) البنية GAN [26].

1-8-2 المميز (D) Discriminator :

مهمته التمييز بين المعلومات المزيفة (المولدة من المولد) والمعلومات الحقيقية، وهو بذلك يعمل عمل مصنف؛ أي يحول الخصائص الواردة إليه إلى أحد الصنفين، إما مزيف أو حقيقي.

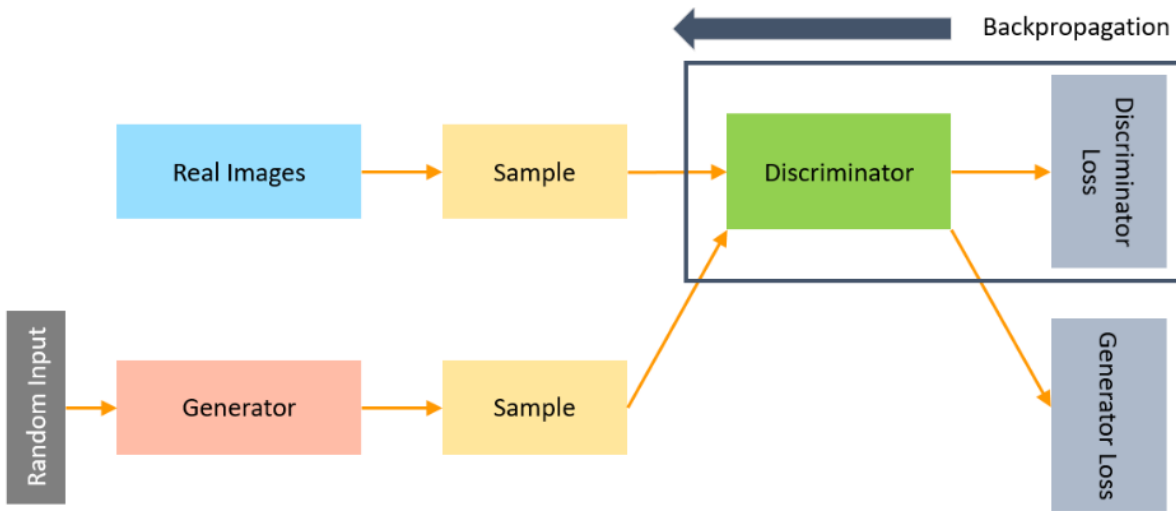
يستخدم المميز في أثناء عملية التدريب نوعين من المعلومات هما:

- المعلومات المزيفة المولدة من المولد.
- المعلومات الحقيقية مثل صور أشخاص بحال كان المطلوب من البنية GAN توليد صور لأشخاص.

عند تدريب المميز :

- يُعتمد عند التدريب على فقد المميز Discriminator loss فقط (فقد المولّد يستخدم عند تدريب المولّد فقط).
- لا يُدرّب المولّد في أثناء عملية تدريب المميز ؛ أي تبقى أوزان المولّد ثابتة.
- يُصنّف المميز كل من المعلومات الحقيقية والمعلومات المزيفة الواردة من المولّد.
- يعمل فقد المميز على حماية المميز من سوء التصنيف؛ أي منعه من تصنيف المعلومات الحقيقية على أنها مزيفة والمعلومات المزيفة على أنها حقيقية.
- تُعدّل أوزان المميز خلال المسار الخلفية (Backpropagation) بحيث يصبح فقد المميز أقل ما يمكن.

يوضح الشكل (23-2) مسار تدريب المميز.



الشكل (23-2) مسار تدريب المميز [26].

2-8-2 المولّد (G) Generator :

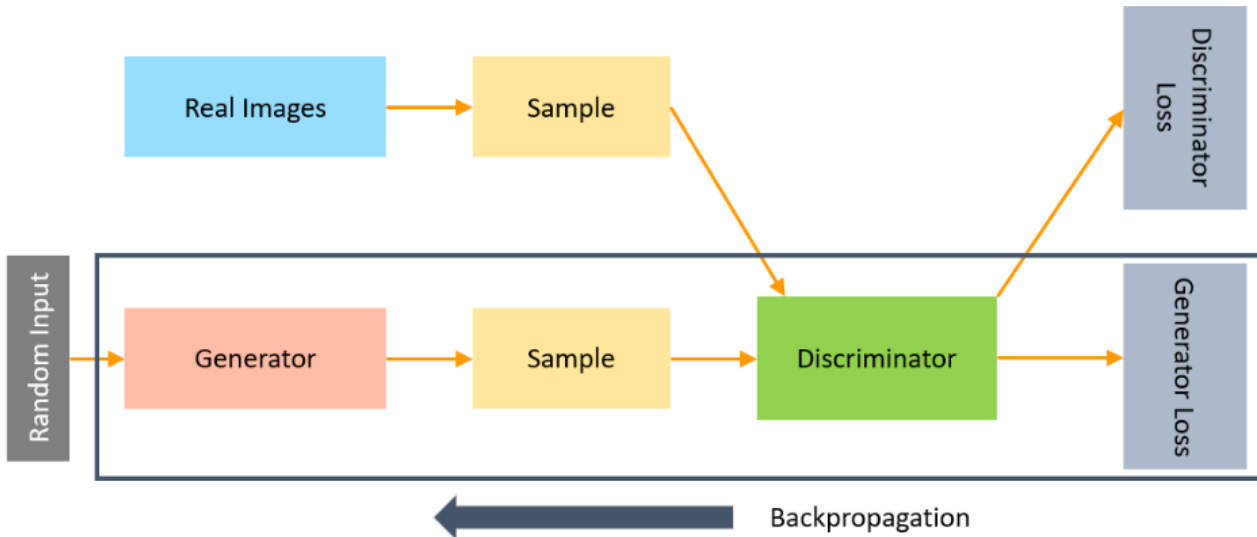
يعمل على توليد المعلومات المزيفة ويُدرّب حتى يجعل المميز يصنف المعلومات المزيفة على أنها حقيقية.

عند تدريب المولّد:

- المحافظة على أوزان المميز ثابتة.
- يعمل المولّد على توليد المعلومات.

- تُدخل المعلومات المولدة إلى المميز الذي سيعمل على تصنيفها.
- يُعتمد في أثناء تدريب المولّد على فقد المولّد (Generator loss) فقط.
- يعمل فقد المولّد على دفع المولّد إلى توليد معلومات لا يمكن للمميز تمييزها عن الحقيقية.
- في المسار الخلفي تُعدل أوزان المولّد بحيث يصبح فقد المولّد أقل ما يمكن.

يبين الشكل (2-24) مسار تدريب المولّد.



الشكل (2-24) مسار تدريب المولّد [26].

عند تدريب البنية GAN ككل يُدرّب أولاً المميز وتُثبت أوزانه ثم يدرّب المولّد وتُثبت أوزانه حتى يُدرّب المميز وهكذا حتى يصبح المميز غير قادر على التمييز بين المعلومات المزيفة والحقيقية؛ أي يكون احتمال كل صنف منهما هو 0.5 عندها يكون التدريب قد انتهى، يُحذف المميز وتكون البنية جاهزة لتوليد المعلومات المطلوبة.

2-8-3 بعض تطبيقات البنية GAN:

- توليد تفاصيل تُضاف إلى الصور منخفضة الدقة بهدف الحصول على الصورة نفسها بدقة أعلى.
- توليد معلومات انطلاقاً من صنف محدد سابقاً، مثل أن يقوم المستخدم بتحديد نوع أو مكونات صورة ما تعمل البنية على توليدها.
- التحويل من صورة إلى صورة بخصائص مختلفة.
- التحويل من الصورة الثنائية إلى صورة ملونة.

- توليد أشكال ثلاثية الأبعاد بدءاً من صور ثنائية البعد.
- توليد أطر جديدة وإضافتها إلى فيديو.
- الحصول على صور الوجوه من زوايا مختلفة بدءاً من صورة واحدة.
- التحويل من نص إلى صورة، حيث يتم إدخال نص إلى البنية وهي ستعمل على توليد صورة تصف النص المدخل.
- تعاني بعض الصور من النقص في المعلومات بسبب تشوه تعرضت له، تعمل البنية GAN على توليد تلك المعلومات الناقصة بهدف ترميم الصورة المشوهة.
- التحويل من نص مكتوب إلى خطاب.

الفصل الثالث

ترميم الصور وتطبيقاته وبعض خوارزميات الترميم

Image Inpainting, Applications and Some Image Inpainting Algorithms

3-1 ترميم الصور وتطبيقاته:

يُعرف ترميم الصور image inpainting أنه العملية التي تهدف إلى إصلاح الصور المشوهة من خلال إعادة بناء الأجزاء المفقودة وتوليد الخطوط العامة والتفصيلية ضمن تلك الأجزاء بما يتناسب مع الأجزاء السليمة المجاورة للأجزاء المشوهة بهدف الحصول على الصورة السليمة.

التحدي الأساسي في هذا المجال هو إعادة توليد خطوط عامة صحيحة وتفاصيل تتناسب مع الجزء السليم المجاور للجزء المفقود من دون ضجيج أو غباش في الحدود الفاصلة بين المنطقتين السليمة والمفقودة إضافةً إلى التناسق اللوني بينهما.

استحوذ مجال ترميم الصور الرقمية على اهتمام عدد كبير من الباحثين بعد التطور الكبير في تكنولوجيا المعالجة وأدواتها إضافةً إلى تطبيقات الترميم الممثلة بما يأتي:

- اعتمد علماء الآثار والتاريخ في جمع معلوماتهم على النقوش والصور المنحوتة؛ فقد كانت الصور هي وسيلة التواصل ونقل معلومات وأخبار الحضارات عبر العصور. ولكن بفعل الزمن والتقدم تعرضت هذه الصور لتشوهات وعيوب أصبح معها من الصعب فهم محتواها، وهنا يأتي دور ترميم الصور لإعادة إحياء تلك الصور والحصول على الصورة الأصلية السليمة.

- التخلص من أنواع الضجيج المختلفة التي قد تعاني منها الصور. من الممكن أن يسبب الضجيج نقصاً في معلومات الصور؛ لذلك تقوم عملية الترميم ببناء الأجزاء الناقصة ومن ثم التخلص من الضجيج.

- حذف جزء غير مرغوب من الصورة يحدده المستخدم، ومن ثم ترميم مكان الجزء المحذوف بما يتناسب مع الأجزاء المحيطة به.

تعتمد القدرة على ترميم الصور عامة على مدى التعقيد في مكونات الصورة بالإضافة إلى مقدار المعلومات المفقودة. يبيّن الشكل (3-1) أمثلة على تطبيقات ترميم الصور.



الشكل (3-1) أمثلة على تطبيقات ترميم الصور [15].

3-2 الطرائق المتبعة في ترميم الصور الرقمية:

3-2-1 الطرائق التقليدية Traditional Methods [15]:

- الطرائق المعتمدة على العينات Patch-based methods: تعتمد هذه الطرائق على ترميم المناطق المفقودة عينة عينة patch by patch من خلال البحث عن العينة الأكثر ملاءمة (الأكثر ترابطاً) للمنطقة المفقودة، ويتم ذلك إما في الجزء السليم من الصورة المراد ترميمها أو من صور أخرى مصدرية source images، ثم نسخه ولصقه في المنطقة المفقودة.
- الطرائق المعتمدة على النشر diffusion-based methods: تعمل هذه الطرائق على تمديد جوار المناطق المفقودة إلى داخلها.

يمكن للطرائق التقليدية أن تعطي نتائج مقبولة نوعاً ما إذا كان النقص في الصور بمساحة صغيرة جداً، أو كان النقص في خلفية الصورة، أما إذا تعدى النقص مكونات الصورة فلا تعطي نتائج جيدة.

3-2-2 الطرائق المعتمدة على التعلّم Learning-based Methods:

تتبع خوارزمية الترميم إحدى البنيتين الآتيتين [15]:

- الشبكة العصبونية الالتفافية (CNN).
- الشبكة المولدة المتعاكسة (GAN).

الميزة الأساسية في هذه الطرائق أنها قادرة على تعلّم تفاصيل الصورة وفهمها وهذه ميزة مهمة جداً في الصور ذات التفاصيل والمحتويات المعقدة مثل الوجوه والأماكن.

3-3 بعض خوارزميات ترميم الصور الرقمية:

على الرغم من أن جميع خوارزميات ترميم الصور تهدف عموماً إلى الحصول على أفضل نتيجة ترميم ممكنة؛ أي التوصل إلى صورة مرممة مطابقة ما أمكن للصورة الأصلية، فإن لكل خوارزمية بنية خاصة بها وجوانب معينة تهتم بها.

1-3-3 الخوارزمية الأولى: ترميم الصور بواسطة تعلّم الشبكة المتعكسة المولّدة للحواف (EdgeConnect) [11].

الفكرة الأساسية التي تعتمد عليها خوارزمية الترميم هذه هي: الخطوط أولاً، والألوان ثانياً. أي إنه عند توافر حواف الأجزاء المكونة للصورة أصبح العمل محصوراً في تقدير الكثافة اللونية للبكسلات.

يُقسم العمل إلى مرحلتين:

المرحلة الأولى: توليد الحواف (Edge generation (Image-to-Edges).

المرحلة الثانية: إتمام الصورة (Image completion (Edges-to-Images).

تتبع كلا المرحلتين بنية الشبكة المولّدة المتعكسة.

1-1-3-3 المرحلة الأولى: توليد الحواف Image-to-Edges:

يتم تقدير حواف الأجزاء المفقودة (المشوّهة) من الصورة فقط.

يبيّن الشكل (2-3) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم EdgeConnect.

دخل المرحلة الأولى:

Iingray المستوى الرمادي للصورة المراد ترميمها.

Cin صورة تتضمن حواف مكونات الصورة المراد ترميمها.

M صورة ثنائية توضح أماكن النقص (التشوه) في الصورة المراد ترميمها، الأبيض

يُعبّر عن الجزء المفقود أما الأسود يُعبّر عن الجزء السليم من الصورة المشوّهة.

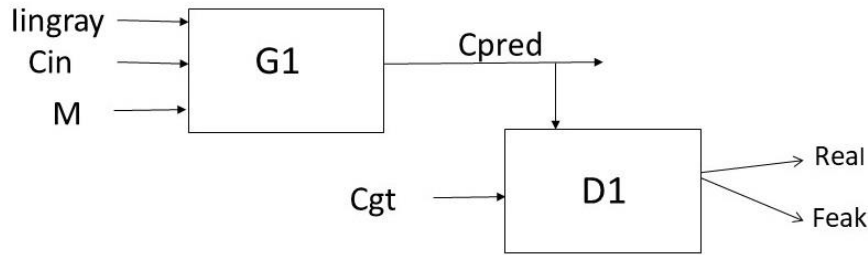
خرج المرحلة الأولى:

Cpred= G1(Iingray,Cin,M) هو عبارة عن الحواف المتوقعة لمكونات الأجزاء

المفقودة فقط.

ملاحظة: Cgt هي عينات تدريب تستخدم فقط أثناء عملية التدريب.

G1 هو المولد المسؤول عن توليد الحواف في الأجزاء المفقودة، D1 هو المميز المسؤول عن تدريب المولد G1.



الشكل (2-3) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم EdgeConnect [الباحث].

2-1-3-3 المرحلة الثانية: إتمام الصورة Edges-to-Image:

يتم في هذه المرحلة دمج الحواف المولدة من المرحلة السابقة (حواف الأجزاء المفقودة) مع حواف الأجزاء السليمة المحيطة بالأجزاء المفقودة، ثم تقدير الكثافة اللونية لكل بكسل من بكسلات الأجزاء المفقودة للحصول على الصورة المرممة.

يبين الشكل (3-3) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم EdgeConnect.

دخل المرحلة الثانية:

هي صورة تجمع بين الحواف المولدة في المرحلة السابقة والحواف في الأجزاء السليمة.

$$C_{comp} = C_{in} + C_{pred}$$

Iin الصورة المراد ترميمها.

خرج المرحلة الثانية:

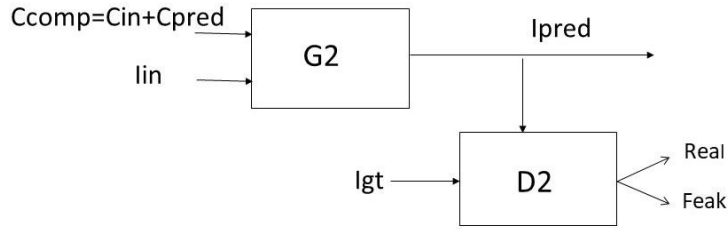
Ipred= G2(Ccomp,Iin) الصورة المرممة النهائية.

G1 هو المولد المسؤول عن توليد الكثافات اللونية للبكسلات في الأجزاء المفقودة، D1 هو المميز المسؤول عن تدريب المولد G1.

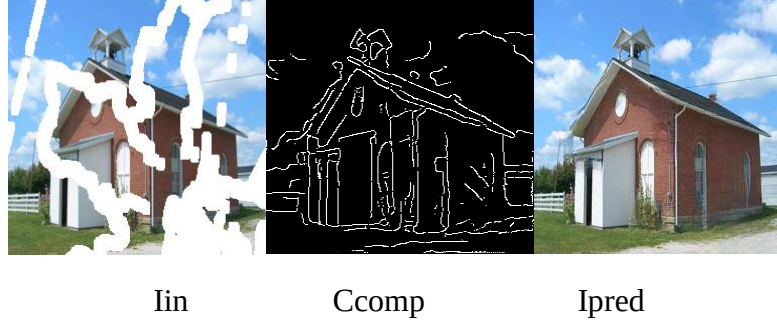
يبين الشكل (4-3) مثلاً يوضح مراحل عمل خوارزمية الترميم EdgeConnect.

يبين الشكل (5-3) المخطط التفصيلي لخوارزمية الترميم EdgeConnect.

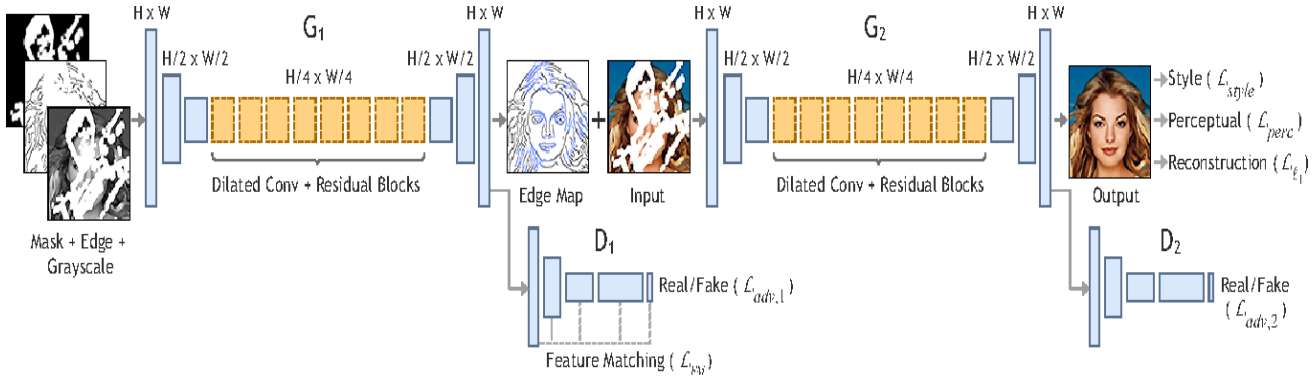
يبين الشكل (6-3) المخطط الوظيفي لخوارزمية الترميم EdgeConnect.



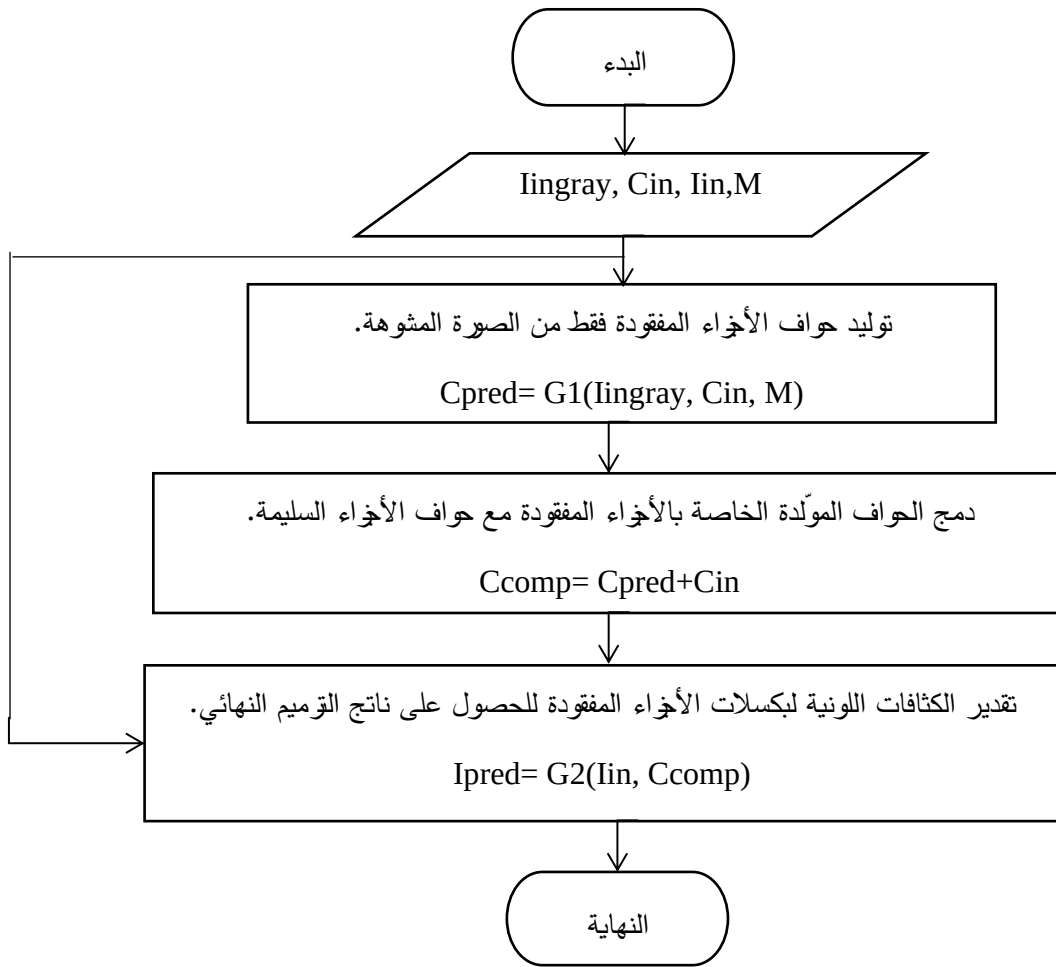
الشكل (3-3) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم EdgeConnect [الباحث].



الشكل (4-3) مثال على مراحل عمل خوارزمية الترميم EdgeConnect [11].



الشكل (5-3) المخطط التفصيلي لخوارزمية الترميم EdgeConnect [11].



الشكل (3-6) المخطط الوظيفي لخوارزمية الترميم EdgeConnect [الباحث].

معاملات تدريب الخوارزمية:

معدل التدريب: 0.0001

معاملات المحسن adam: $\beta_1=0$ و $\beta_2=0.9$

معدل تدريب المولد/ المميز: 0.1

عدد عينات التدريب في كل دورة تدريب: 8

العدد الأعظمي لدورات التدريب: $2e^6$

وزن تابع الخطأ L1: 1

وزن تابع الخطأ Lfm: 10

وزن تابع الخطأ Lstyle: 250

وزن تابع الخطأ Lprec: 0.1

وزن تابع الخطأ Ladv: 0.1

3-3-2 الخوارزمية الثانية: ترميم الصورة بوساطة الخطوط العامة والشكل العالم للصورة :[12] (StructureFlow)

الفكرة الأساسية التي تعتمد عليها خوارزمية الترميم هذه هي تقسيم الخطوط المولدة إلى قسمين هما: خطوط عامة، وخطوط دقيقة (تفاصيل).

يُقسم العمل إلى مرحلتين:

المرحلة الأولى: إعادة بناء الخطوط العامة Structure reconstruction.

المرحلة الثانية: توليد التفاصيل Texture generation.

تتبع كلا المرحلتين بنية الشبكة المولدة المتعكسة.

3-3-2-1 المرحلة الأولى: إعادة بناء الخطوط العامة Structure reconstruction.

في هذه المرحلة يكون التركيز فقط على استعادة الخطوط العامة.

يبيّن الشكل (3-7) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم StructureFlow.

دخل المرحلة الأولى:

Iin الصورة المراد ترميمها.

Sin صورة تتضمن الخطوط العامة فقط للصورة المراد ترميمها يمكن الحصول عليها من

حذف high-frequency details والإبقاء على low-frequency details.

M صورة ثنائية توضح أماكن النقص (التشوه) في الصورة المراد ترميمها، الأبيض يُعبّر عن

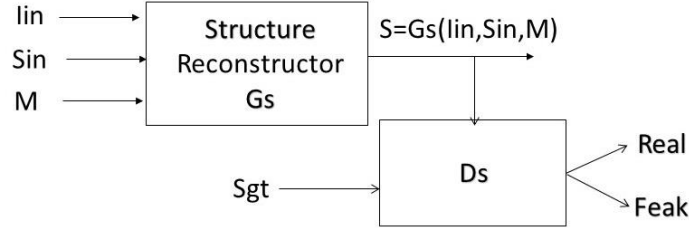
الجزء المفقود، أما الأسود فيُعبّر عن الجزء السليم من الصورة المشوهة.

خرج المرحلة الأولى:

S=Gs(Iin,Sin,M) هي صورة تتضمن الخطوط العامة المقدرة في الأجزاء المفقودة والخطوط

العامة في الأجزاء السليمة.

Sgt هي عينات التدريب تتضمن الخطوط العامة فقط من الصورة، G_s هو المولد المسؤول عن توليد الحواف فقط في الأجزاء المفقودة، D_s هو المميز المسؤول عن تدريب المولد G_s .



الشكل (3-7) المخطط الصندوقي للمرحلة الأولى لخوارزمية الترميم StructureFlow [الباحث].

3-2-2-3 المرحلة الثانية: توليد التفاصيل Texture generation:

في هذه المرحلة تُولّد الخطوط التفصيلية اعتماداً على الخطوط العامة المولّدة في المرحلة السابقة إضافةً إلى الأجزاء السليمة من الصورة. يتم ذلك بواسطة ما يسمى Appearance flow التي تعمل على حساب الترابط بين الخطوط العامة المولّدة في المرحلة السابقة و الخطوط العامة في الأجزاء السليمة من الصورة. تُخزن نتائج حساب الترابط في مصفوفة ثنائية الأبعاد appearance-flow field كل عنصر منها يشير إلى الجزء في المنطقة السليمة الأكثر تشابهاً مع الجزء في المنطقة المُرَمّة، وعليه فإن التفاصيل في الجزء المُرَمم ستكون مشابهة للتفاصيل في الجزء السليم الأكثر تشابهاً معه. يعمل المولّد G_t على توليد التفاصيل اعتماداً على المصفوفة appearance-flow field.

يبين الشكل (3-8) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم StructureFlow.

دخل المرحلة الثانية:

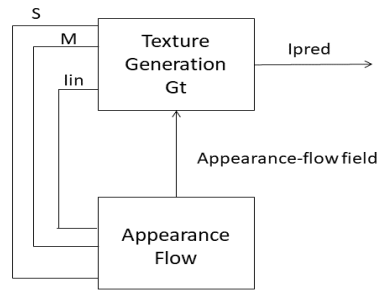
S صورة تتضمن الخطوط العامة الناتجة عن المرحلة السابقة.

M صورة ثنائية توضح أماكن النقص (التشوه) في الصورة المراد ترميمها، الأبيض يُعبّر عن الجزء المفقود أما الأسود فيُعبّر عن الجزء السليم من الصورة المشوهة.

I_{in} الصورة المراد ترميمها.

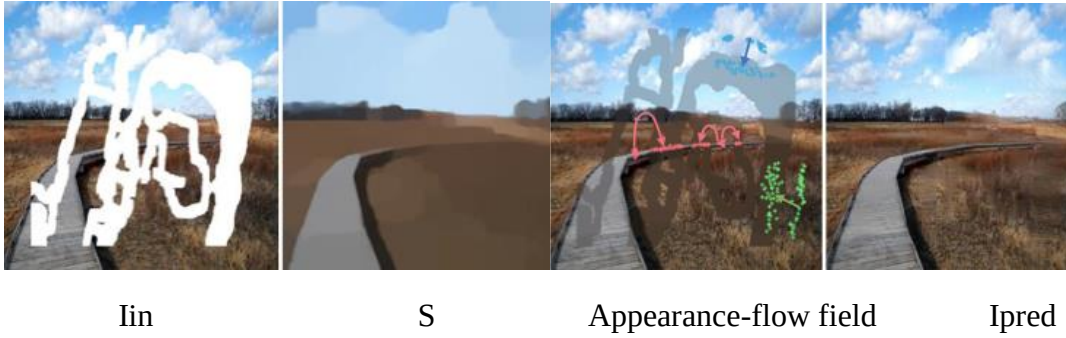
خرج المرحلة الثانية:

$I_{pred} = G_t(S, M, I_{in})$ الصورة المُرَمّة النهائية.



الشكل (8-3) المخطط الصندوقي للمرحلة الثانية لخوارزمية الترميم StructureFlow [الباحث].

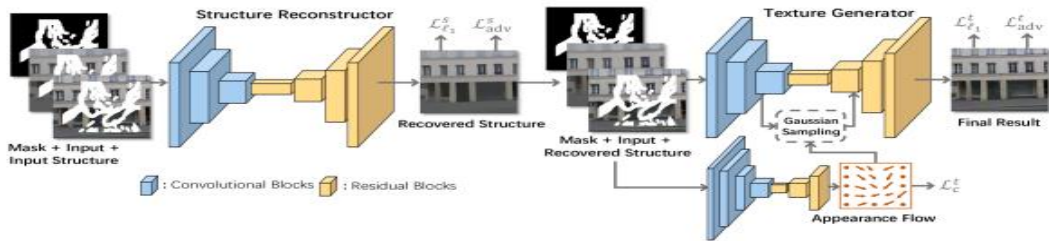
يبين الشكل (9-3) مثالاً على مراحل عمل خوارزمية الترميم StructureFlow.



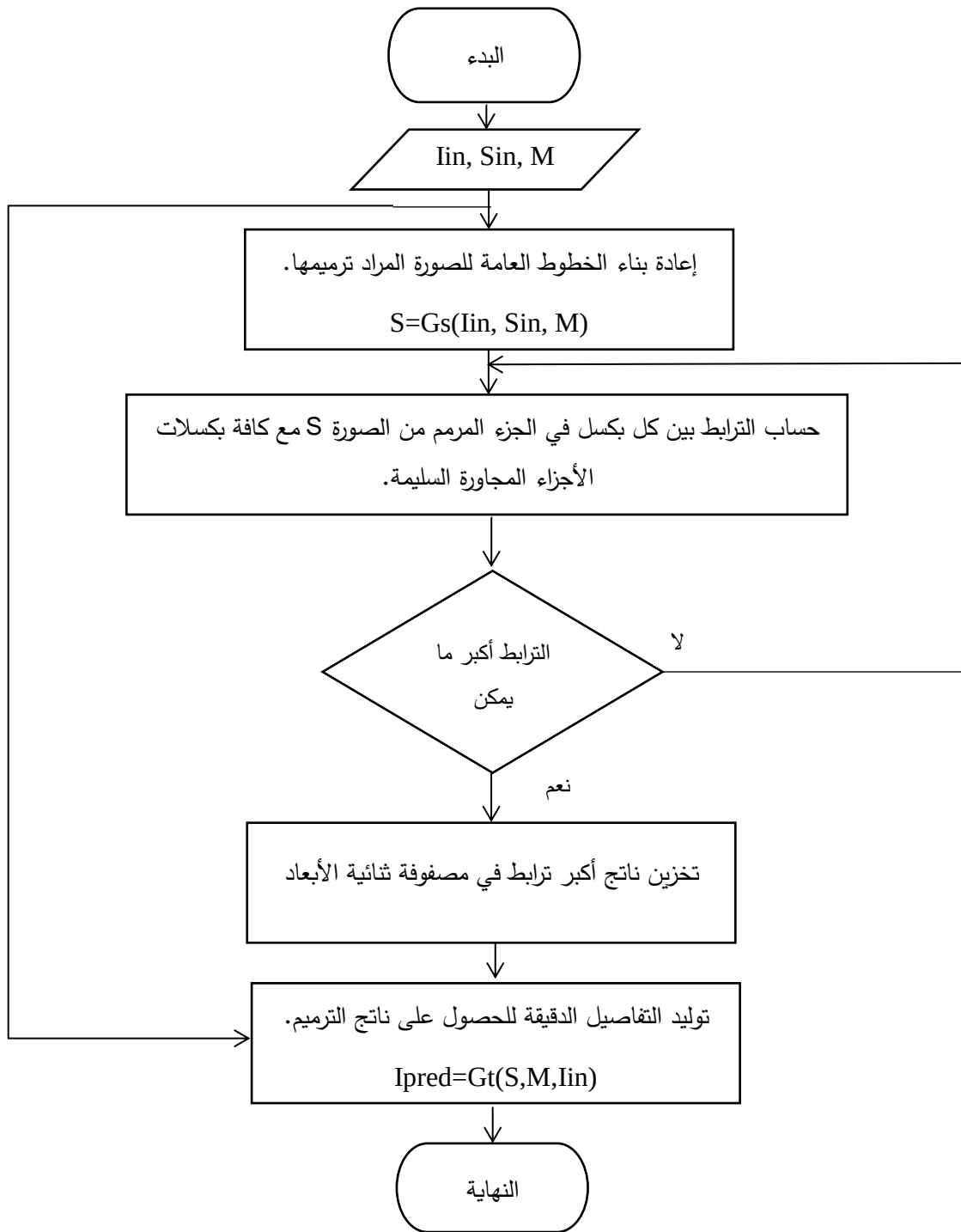
الشكل (9-3) مثال على مراحل عمل خوارزمية الترميم StructureFlow [12].

يبين الشكل (10-3) المخطط التفصيلي لخوارزمية الترميم StructureFlow.

يبين الشكل (11-3) المخطط الوظيفي لخوارزمية الترميم StructureFlow.



الشكل (10-3) المخطط التفصيلي لخوارزمية الترميم StructureFlow [12]



الشكل (3-11) المخطط الوظيفي لخوارزمية الترميم StructureFlow [الباحث].

معاملات تدريب الخوارزمية:

معدل التدريب: 0.0001

معاملات المحسن adam: $\beta_1=0$ و $\beta_2=0.999$

عدد عينات التدريب في كل دورة تدريب: 8

العدد الأعظمي لدورات التدريب: $8e^6$

وزن تابع الخطأ L1 في المرحلة الأولى: 4

وزن تابع الخطأ Ladv في المرحلة الأولى: 1

وزن تابع الخطأ L1 في المرحلة الثانية: 5

وزن تابع الخطأ Ladv في المرحلة الثانية: 1

وزن تابع الخطأ Lc: 0.25

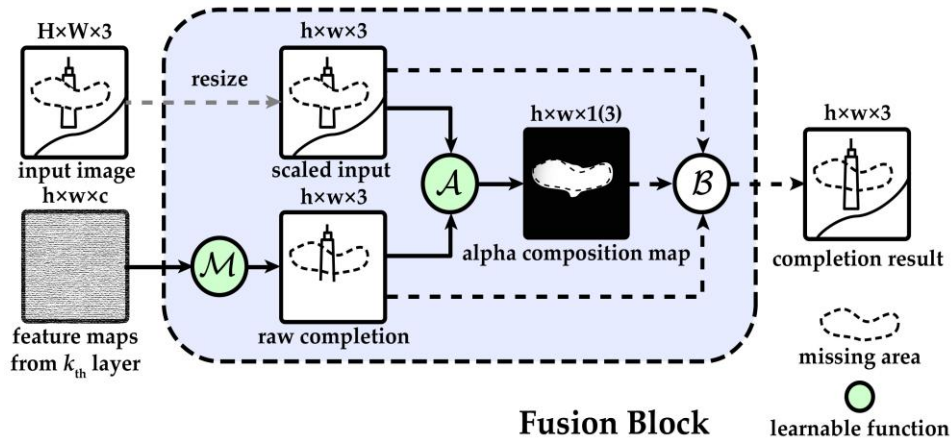
3-3-3 الخوارزمية الثالثة: ترميم الصور بواسطة شبكة الدمج العميق (DFNet)

[13]:

تعمل خوارزمية الترميم هذه على حل مشكلة تعاني منها معظم خوارزميات الترميم وهي الضجيج والغباش في الحدود الفاصلة بين الجزأين السليم والمرمم. تحقق ذلك باستخدام بنية دمج fusion block، الشكل (3-12)، مهمتها الربط بين الجزأين المرمم والسليم وتحقيق اندماج على مستوى البكسل في الحدود الفاصلة بينهما.

تُربط بنية الدمج مع الطبقة الأخيرة لمفكك الترميز من البنية U-Net للحصول على ناتج الترميم بدقةات مختلفة، يوضح ذلك في الشكل (3-13).

تتبع هذه الخوارزمية بنية CNN.



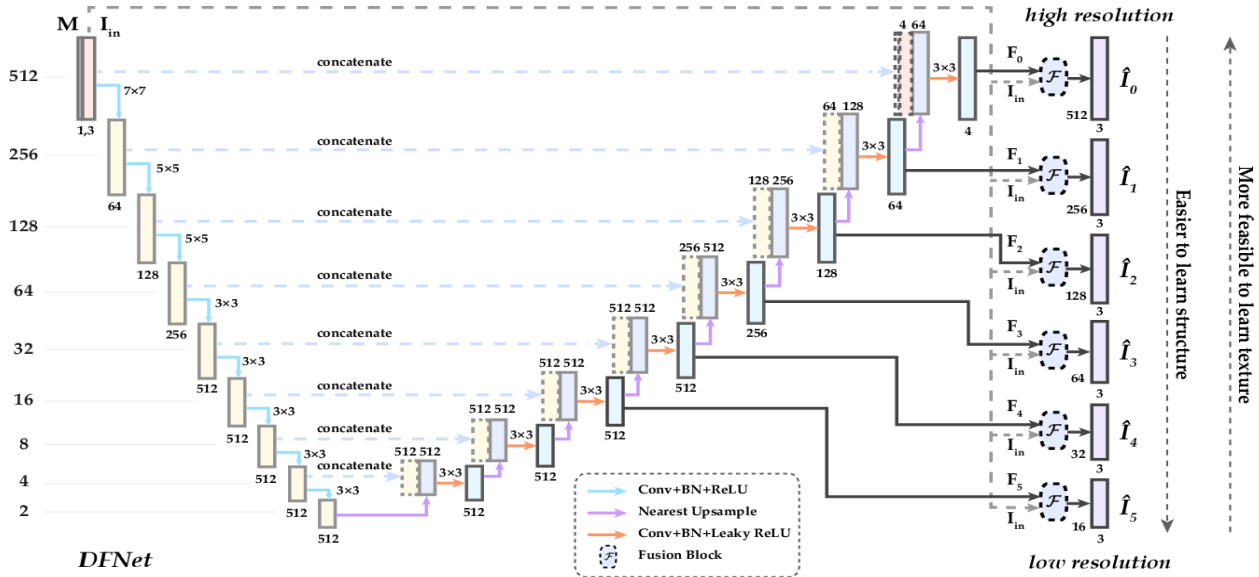
الشكل (12-3) بنية الدمج Fusion block [13].

M هي شبكة عصبونية تعمل على توليد ناتج الترميم الأولي للصورة المشوّهة.

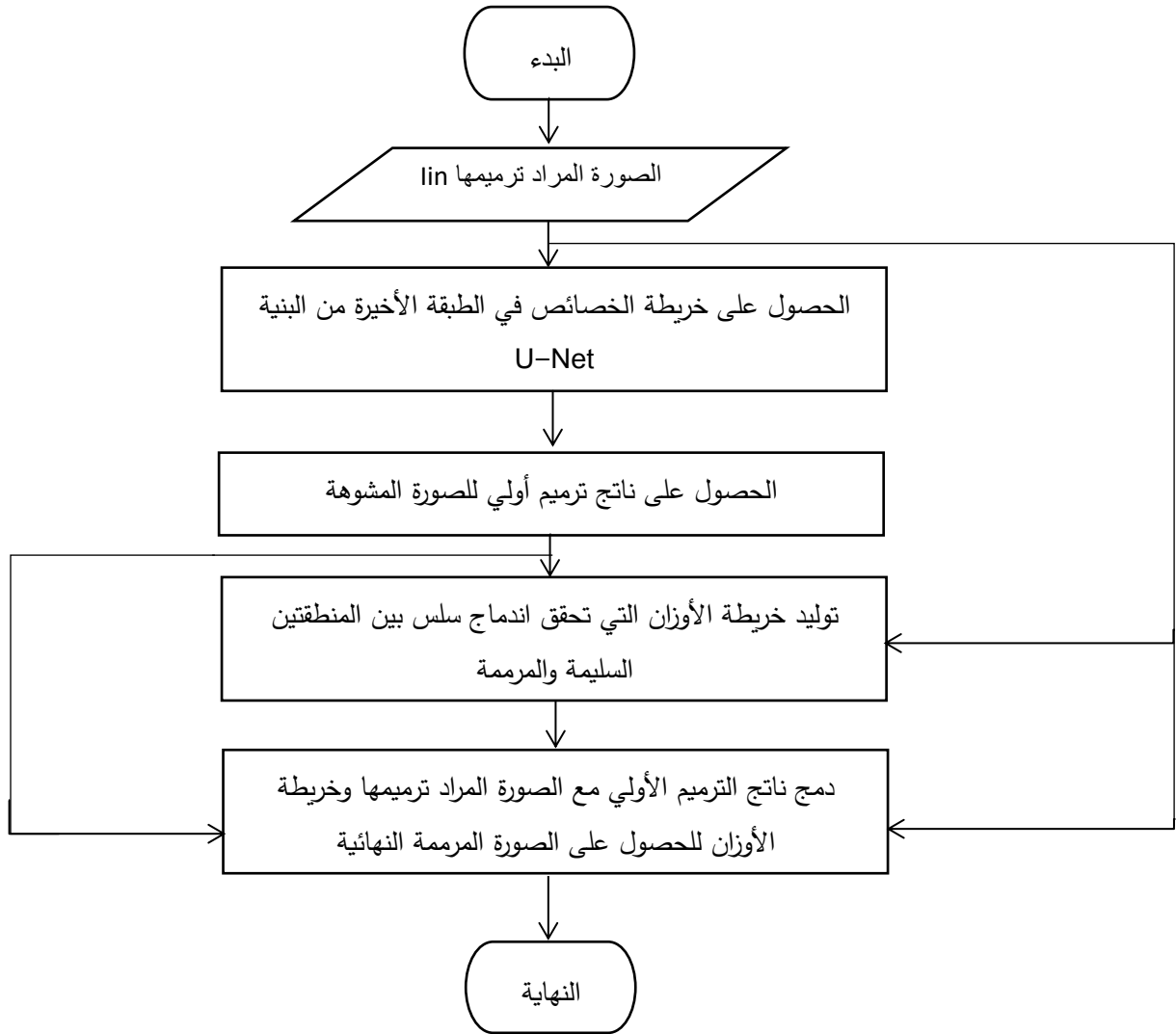
A هي شبكة عصبونية تعمل على توليد خريطة الأوزان التي تحقق اندماج سلس بين المنطقتين السليمة والمرممة بهدف التخلص من الغباش والضجيج في تلك المنطقة.

B هي شبكة عصبونية تدمج ناتج الترميم الأولي مع كل من الصورة المراد ترميمها وخريطة الأوزان للحصول على ناتج الترميم النهائي.

يبين الشكل (14-3) المخطط الوظيفي لخوارزمية الترميم DFNet.



الشكل (13-3) المخطط التفصيلي لخوارزمية الترميم DFNet [13].



الشكل (3-14) المخطط الوظيفي لخوارزمية الترميم DFNet [الباحث].

معاملات تدريب الخوارزمية:

معدل التدريب: 0.0002

معاملات المحسن adam: $\beta_1=0$ و $\beta_2=0.099$

عدد عينات التدريب في كل دورة تدريب: 6

العدد الأعظمي لدورات التدريب: $6e^6$

وزن تابع الخطأ L1: 6

وزن تابع الخطأ Lperc: 0.1

وزن تابع الخطأ Lstyle: 240

الفصل الرابع

تقدير جودة الصور الرقمية وبارامترات التقييم التقليدية والمُقترحة
لتقييم جودة الصور المرممة

**Digital Image Quality Assessment and
Traditional and Proposed Evaluation Parameters
for Inpainted Image Quality Evaluation**

4-1 تقدير جودة الصور الرقمية:

يهدف تقدير جودة الصور الرقمية إلى قياس مدى التشوه الذي تعرضت له الصور الرقمية نتيجة العمليات المطبقة عليها مثل المعالجة، والتخزين، والضغط، والإرسال.

يؤدي تقدير جودة الصور الرقمية دوراً هاماً جداً في تطبيقات معالجة الصور المختلفة، مثل مقارنة أداء عدة خوارزميات ضغط الصور، وإزالة الضجيج، وتقييم أداء الأنظمة، وترميم الصور، وتحسين الصور.

4-2 طرائق تقدير جودة الصور الرقمية[5]:

- Subjective measurement:

تُعرض الصور المطلوب تقييم جودتها على مجموعة من الأشخاص، يقوم كل منهم بوضع تقدير score لجودة الصورة. تُعد هذه الطريقة من الطرائق الأساسية في تقدير جودة الصورة؛ لأن الإهتمام بقياس الصور وتحسينها يهدف إلى جعلها مقبولة بالنسبة إلى الشخص المشاهد، إلا أن هذه الطريقة تعاني من مساوئ عدة حالت دون استخدامها فهي مكلفة، وتستغرق وقتاً طويلاً، وتعتمد على القدرات البصرية للأشخاص، ولا يمكن أن تكون بالزمن الحقيقي.

- Objective measurement:

تُقدر جودة الصور أوتوماتيكياً من خلال علاقات رياضية من دون الحاجة لتدخل البشر. تمتاز هذه الطريقة بالسرعة، والعمل بالزمن الحقيقي، وقلة كلفتها.

تُصنف هذه الطريقة تبعاً لوجود الصورة الأصلية (غير المشوهة) إلى :

- No-reference (NR): الصورة الأصلية غير متوفرة، تستخدم طريقة التقدير هذه في بعض التطبيقات التي لا تتوفر فيها الصورة الأصلية مثل تحسين دقة الصورة.

- Full-reference (FR): تتطلب وجود الصورة الأصلية كاملة، وتُعد من أكثر طرائق التقدير دقة وشيوعاً، تصنف إلى:

- تقدير جودة الصور بالاعتماد على قياس مدى الاختلاف بين بكسلات كل من الصورتين الأصلية والمشوّهة.

- تقدير جودة الصور بالاعتماد على قياس الترابط بين الصورتين الأصلية والمشوّهة.

- تقدير جودة الصور بالاعتماد على قياس مدى إختلاف الحواف بين الصورتين الأصلية والمشوهة.

- تقدير جودة الصور بالاعتماد على نظام الرؤية البشري (HVS).

- Reduced-reference (RR): طريقة تقدير بين FR و NR حيث يتوافر بعض المعلومات فقط عن الصورة الأصلية.

4-3 بارامترات التقييم التقليدية لتقدير جودة الصور المرممة:

اعتمد في تقييم جودة الصورة المرممة وكذلك تقييم أداء خوارزمية الترميم على بارامترين تقليديين هما:

- نسبة الإشارة العظمى إلى الضجيج (PSNR) Peak signal to noise ratio.

- معامل التشابه البنيوي (SSIM) Structural similarity index.

4-3-1 نسبة الإشارة العظمى إلى الضجيج (PSNR):

هو بارامتر تقييم full-reference يعمل على تقدير جودة الصور من خلال قياس الاختلاف بين بكسلات الصورتين الأصلية والمشوهة، يُقدر بالديسيبل، ويعطى بالعلاقة (4-1) [4]:

$$PSNR = 10 * \log \frac{MAX^2}{MSE(f,g)} \quad (1 - 4)$$

حيث: MAX هي أعظم قيمة للبكسل والمتواجدة إما في الصورة الأصلية أو المرممة .

f الصورة الأصلية.

g الصورة المرممة.

يأخذ هذا البارامتر قيمة بين الصفر واللانهاية، القيم الأكبر تعني جودة أفضل للصورة .

4-3-2 معامل التشابه البنيوي (SSIM):

هو بارامتر تقييم full-reference يعتمد على نظام الرؤية البشري. يجمع هذا البارامتر ثلاثة عوامل هي:

- فقد الترابط (s) loss of correlation.

- تشوه النصوص (l) luminance distortion.

- تشوه التباين (c) contrast distortion.

يعطى بالعلاقات (2-4) [3]، (3-4) [4]:

$$SSIM = s(f,g) \cdot l(f,g) \cdot c(f,g) \quad (2-4)$$

$$SSIM = \frac{(2\mu_f \mu_g + c1)(2\sigma_f \sigma_g + c2)}{(\mu_f^2 + \mu_g^2 + c1)(\sigma_f^2 + \sigma_g^2 + c2)} \quad (3-4)$$

f الصورة الأصلية.

g الصورة المرممة

μ_f, μ_g : متوسط النصوص.

σ_f, σ_g : الانحراف المعياري.

σ_{fg} : التباين بين الصورتين F و G.

c1, c2 : ثوابت.

يأخذ البارامتر SSIM قيمةً ضمن المجال [0,1]، القيمة 1 تعني أن الصورتين الأصلية والمرممة متطابقتان أما القيمة 0 فتعني أن الصورتين مختلفتان تماماً.

4-4 البارامترات المقترحة لتقدير جودة الصور المرممة:

اقترح ثلاثة بارامترات تقييم جديدة لتقييم جودة الصور المرممة، هي:

1-4-4 البارامتر المقترح الأول: مربع متوسط الخطأ (MSE):

يعمل البحث [16] على تحسين جودة الصور الملتقطة تحت أسطح البحار والمحيطات، بهدف مراقبة خطوط النقل والكبلات وصيانتها أو لأهداف البحث عن مصادر الطاقة، أو بهدف البحث العلمي، بهدف الاستخلاص الدقيق لحواف مكوناتها كعملية معالجة أولية لتكون جاهزة لإدخالها إلى خوارزميات التعرف، لأن تلك الصور تعاني من انخفاض كبير في جودتها بسبب خصائص الوسط الملتقطة به.

يعتمد هذا البحث في تقييم نتائجه ومقارنتها مع نتائج أبحاث أخرى على بارامتر تقييم يُعطى بالعلاقة (4-4) [16]:

$$MSE = \frac{1}{m * n} \sum_{i=1}^m \sum_{j=1}^n (f(i,j) - g(i,j))^2 \quad (4-4)$$

$f(i,j)$: الكثافة اللونية للبكسل ذو الموقع i,j من الصورة الأصلية f .

$g(i,j)$: الكثافة اللونية للبكسل ذو الموقع i,j من الصورة المرممة g .

i,j : إحداثيات البكسل من الصورة ذات الأبعاد $m*n$.

تدل القيم الأصل لمربع متوسط الخطأ على جودة أفضل للصورة.

ملاحظة: تقترح الدراسة [16] طريقة لتحسين جودة الصور الملتقطة في الأوساط المائية بالتالي هي دراسة لا ترتبط بمجال ترميم الصور، اعتمد هذا البحث على الدراسة [16] فقط في اقتراح بارامتر تقييم جديد لتقدير جودة الصور المرممة.

2-4-4 البارامتر المقترح الثاني: تقدير جودة الصور بالاعتماد على الحواف والبكسل (EPIQA):

يقترح البحث [23] بارامتر تقييم EPIQA يعمل على تقدير جودة الصور بأسلوب مشابه للبشر حيث عند عرض صورتين إحداهما تعرضت لتشوه ما على الإنسان فإنه سيعمل على المقارنة بينهما من خلال ملاحظة الفروقات في الحواف إضافة إلى اختلافات الكثافات اللونية للبكسلات.

انطلق هذا البحث في اقتراحه لبارامتر التقييم EPIQA من أهمية استخلاص الحواف التي هي عبارة عن الحدود الفاصلة بين العناصر و الأجزاء المكونة للصورة، و كون هذا البارامتر مطابقاً لنظام الرؤية البشري إضافة إلى أنه أثبت فعالية البارامتر EPIQA مقارنة مع البارامترات التقليدية PSNR و SSIM. مراحل حساب البارامتر EPIQA:

- حساب نسبة الإشارة العظمى إلى الضجيج لكل من الصورتين الأصلية والمرممة.
- استخلاص حواف كل من الصورتين الأصلية و المرممة باستخدام مرشح Sobel ذو القيم المبينة في الشكل (4-1) كون هذا المرشح يتميز بالدقة:

-1.2	-0.8	-0.6	1.2	0.8	0.6
0	0	0	0	0	0
1.2	0.8	0.6	-1.2	-0.8	-0.6

1.2	0	-1.2	-1.2	0	1.2
0.8	0	-0.8	-0.8	0	0.8
0.6	0	-0.6	-0.6	0	0.6

الشكل (4-1) مرشحات استخلاص الحواف [23].

- تقسيم كل صورة إلى كتل أبعاد كل كتلة 8*8 حيث وفقاً للدراسة الكتل ذات الأبعاد الأكبر تؤدي إلى فقد في معلومات الصورة والكتل بأبعاد أصغر تحتاج إلى زمن معالجة طويل.
- حساب قيم المصفوفات $F_{i,j}$ و $G_{i,j}$ التي تعبر عن خصائص الحواف في كل كتلة على حدة للصورتين الأصلية F والمرممة G :

$$F_{i,j} = (EOIf, ALEf, PLEf, NEPf)$$

$$G_{i,j} = (EOIg, ALEg, PLEg, NEPg)$$
حيث : EOI عدد الحواف في كل كتلة.

ALE متوسط أطوال الحواف ضمن كل كتلة.

PLE العدد الأعظمي للبكسلات في الكتلة التي تملك السوية الرمادية نفسها.

NEP عدد البكسلات المكونة للحواف.

j, i موقع الكتل.

- حساب المسافة الإقليدية العلاقة، (4-5) [23]، بين كل كتلة من كتل الصورة الأصلية مع الكتلة المقابلة لها من الصورة المرممة:

$$d_{i,j} = \sqrt{(F_{i,j} - G_{i,j})^2} \quad (5 - 4)$$

- حساب بارامتر التقييم $EPIQA$ وفق العلاقة (4-6) [23]:

$$EPIQA = PSNR + \frac{1}{m * n} \sum_{i=1}^m \sum_{j=1}^n d_{i,j} \quad (6 - 4)$$

تدل القيمة $EPIQA$ الأكبر على جودة أفضل للصورة.

3-4-4 البارامتر المقترح الثالث: الترابط المُقيّس (NCC) :

يُستخدم الترابط لحساب مدى التشابه بين إشارتين [1] وبالتالي يمكن الإعتماد عليه في حساب مدى التشابه بين الصورتين الأصلية والمرممة ومن ثم تقدير جودة الصورة المرممة.

يعطى البارامتر المقترح NCC بالعلاقة (7-4) [2]:

$$NCC = \frac{\sum_{i=1}^m \sum_{j=1}^n [(f_{i,j} - \bar{f})(g_{i,j} - \bar{g})]}{[(\sum_{i=1}^m \sum_{j=1}^n (f_{i,j} - \bar{f})^2)(\sum_{i=1}^m \sum_{j=1}^n (g_{i,j} - \bar{g})^2)]^{1/2}} \quad (7-4)$$

حيث: $f_{i,j}$ هو الكثافة اللونية للبكسل ذو الموقع i^*j في الصورة الأصلية.

$g_{i,j}$ هو الكثافة اللونية للبكسل ذو الموقع i^*j في الصورة المرممة.

$\bar{f}$ هو متوسط قيم بكسلات الصورة الأصلية.

$\bar{g}$ هو متوسط قيم بكسلات الصورة المرممة.

يأخذ هذا البارامتر قيمة بين 0 و 1.

$NCC=+1$ أي إن الصورة الأصلية تطابق تماماً الصورة المرممة.

$NCC=0$ أي إن الصورة الأصلية مختلفة تماماً عن الصورة المرممة.

ولذلك كلما اقتربت قيمة البارامتر من 1 كانت ذات جودة أفضل.

الفصل الخامس

التنفيذ العملي والنتائج

Implementation and Results

5-1 تمهيد:

من أجل تقييم أداء عدد من خوارزميات معالجة الصور والمقارنة بينها لا بدّ من اختبار عملها من أجل الدخل نفسه، ثم حساب بارامترات تقييم الخرج. يتم اختيار بارامترات تقييم الخرج لتكون قادرة على تعزيز قرار أيّ من نتائج هذه الخوارزميات هو أفضل؟ ومن ثم تعزيز قرار أيّ من هذه الخوارزميات له بنية أنسب للتطبيق.

يعمل هذا البحث على تقييم أداء خوارزميات ترميم الصور الرقمية من خلال إدخال عدد من الصور المشوّهة على ثلاث من خوارزميات الترميم، ثم اقتراح بارامترات تقييم جديدة، وتقييم أداء خوارزميات الترميم اعتماداً عليها، وذلك لأن بارامترات التقييم التقليدية لم تكن قادرة على تعزيز القرار.

يعتمد هذا البحث في جزئه العملي على البرمجيتين الآتيتين:

MATLAB R2019a: استُخدم لتقييم الأداء وحساب بارامترات التقييم التقليدية والمُقترحة.

PYTHON 3.7: استُخدم لتطبيق خوارزميات الترميم والحصول على نتائج ترميم عدّة صور مشوّهة.

يُقسم العمل إلى أربع مراحل:

المرحلة الأولى: تطبيق خوارزميات الترميم على المجموعة نفسها من الصور المشوّهة.

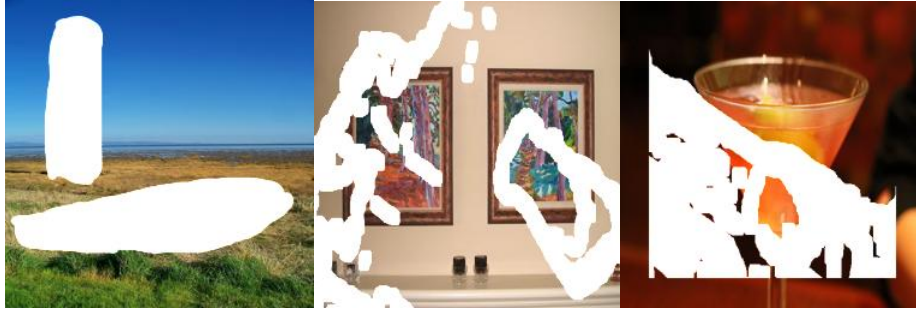
المرحلة الثانية: حساب بارامترات التقييم التقليدية لنتائج المرحلة الأولى.

المرحلة الثالثة: حساب بارامترات التقييم المُقترحة لنتائج المرحلة الأولى.

المرحلة الرابعة: تقييم فعالية البارامترات المُقترحة، وتقييم أداء خوارزميات الترميم بناءً عليها.

2-5 المرحلة الأولى: تطبيق خوارزميات الترميم على المجموعة نفسها من الصور المشوّهة.

يوضّح الشكل (1-5) الصور المشوّهة التي طُبِّقت عليها خوارزميات الترميم الثلاث وهي صور مُختارة بشكل عشوائي من ضمن مجموعة الاختبار في بنية المعلومات places2.



1

2

3



4

5

6



7

الشكل (1-5) الصور المشوّهة التي طُبِّقت عليها خوارزميات الترميم الثلاث [الباحث].

نتائج الترميم على النحو التالي:

الشكل (2-5) نتائج ترميم الخوارزمية الأولى (EdgeConnect) للصور المشوّهة.

الشكل (3-5) نتائج ترميم الخوارزمية الثانية (StructreFlow) للصور المشوّهة.

الشكل (4-5) نتائج ترميم الخوارزمية الثالثة (DFNet) للصور المشوّهة.



1

2

3



4

5

6

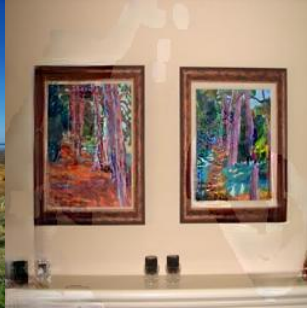


7

الشكل (2-5) نتائج ترميم الخوارزمية الأولى (EdgeConnect) للصور المشوّهة [الباحث].



1



2



3



4



5

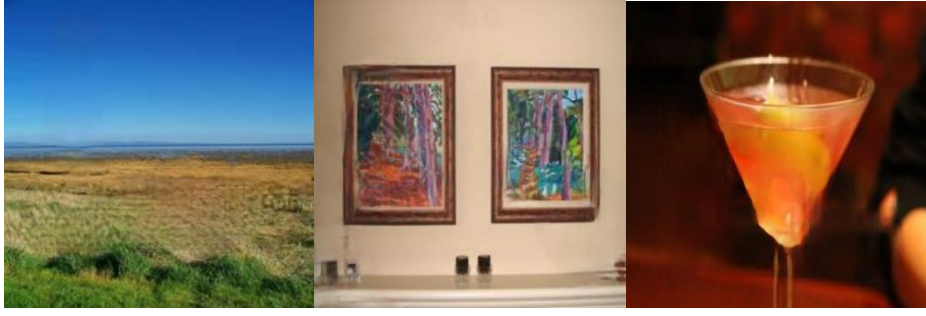


6



7

الشكل (3-5) نتائج ترميم الخوارزمية الثانية (StrutureFlow) للصور المشوهة [الباحث].



1

2

3



4

5

6



7

الشكل (4-5) نتائج ترميم الخوارزمية الثالثة (DFNet) للصور المشوّهة [الباحث].

3-5 المرحلة الثانية: حساب بارامترات التقييم التقليدية لنتائج المرحلة الأولى.

يبين الجدول (1-5) قيم البارامتر التقليدي PSNR لنتائج المرحلة الأولى.

يبين الجدول (2-5) قيم البارامتر التقليدي SSIM لنتائج المرحلة الأولى.

الجدول (1-5) قيم البارامتر التقليدي PSNR لنتائج المرحلة الأولى [الباحث].

DFNet	StructureFlow	EdgeConnect	Image
18.3	17.8	18.3	1
26.4	24.9	25.3	2
28	26.8	26	3
21.2	20.2	20.2	4
22	22.1	21.1	5
27.7	26	26.6	6
30.7	27.9	30.4	7

الجدول (2-5) قيم البارامتر التقليدي SSIM لنتائج المرحلة الأولى [الباحث].

DFNet	StructureFlow	EdgeConnect	Image
0.86	0.86	0.86	1
0.94	0.93	0.93	2
0.96	0.94	0.96	3
0.86	0.84	0.84	4
0.86	0.86	0.84	5
0.96	0.94	0.94	6
0.93	0.94	0.93	7

مناقشة النتائج: من الملاحظ من الجدولين السابقين أن قيم البارامترات التقليدية تكون متقاربة جداً أو متطابقة وهذا ما يخلق حالة من الشك يصبح معها من الصعب تقييم أداء الخوارزميات واتخاذ القرار؛ وهذا يخالف الغاية المرجوة من بارامترات تقييم الأداء عموماً؛ لذلك لابد من اقتراح بارامترات تقييم جديدة لإزالة حالة الشك والقدرة على اتخاذ القرار بناءً عليها.

4-5 المرحلة الثالثة: حساب بارامترات التقييم المقترحة لنتائج المرحلة الأولى.

البارامترات المقترحة لحل المشكلة التي ظهرت في البارامترات التقليدية هي: MSE و EPIQA و NCC.

يبين الجدول (3-5) قيم البارامتر المقترح MSE لنتائج المرحلة الأولى.

يبين الجدول (4-5) قيم البارامتر المقترح EPIQA لنتائج المرحلة الأولى.

يبين الجدول (5-5) قيم البارامتر المقترح NCC لنتائج المرحلة الأولى.

الجدول (3-5) قيم البارامتر المقترح MSE لنتائج المرحلة الأولى [الباحث].

DFNet	StructureFlow	EdgeConnect	Image
726.2	773.2	729.8	1
138.7	197.7	178.2	2
88	120.1	130.3	3
435.6	542.5	688.1	4
354.4	355.9	358.6	5
102.6	140.2	165.2	6
54.4	104.4	65.3	7

الجدول (4-5) قيم البارامتر المقترح EPIQA لنتائج المرحلة الأولى [الباحث].

DFNet	StructureFlow	EdgeConnect	Image
20.1	19	18.3	1
26.9	23.5	25.8	2
28.2	26.9	24.3	3
22	21	20.2	4
24.8	22.8	21.9	5
28.1	27.2	26.1	6
30.9	26.3	28.7	7

الجدول (5-5) قيم البارامتر المقترح NCC لنتائج المرحلة الأولى [الباحث].

DFNet	StructureFlow	EdgeConnect	Image
0.7418	0.7292	0.7416	1
0.9733	0.9707	0.9657	2
0.9819	0.9778	0.9734	3
0.9139	0.8832	0.8497	4
0.9646	0.9644	0.9554	5
0.9839	0.9794	0.9745	6
0.9954	0.9948	0.9902	7

5-5 المرحلة الرابعة: تقييم فعالية البارامترات المقترحة وتقييم أداء خوارزميات الترميم بناءً عليها.

نلاحظ من الجداول (3-5) و (4-5) و (5-5) أن قيم البارامترات المقترحة أعطت تمييزية جيدة، وأزالت حالة الشك التي كانت في البارامترات التقليدية، وأصبح من الممكن اعتماداً على البارامترات المقترحة اتخاذ القرار، وهو أن الخوارزمية الثالثة (DFNet) هي الخوارزمية الأفضل من بين خوارزميات الترميم المدروسة؛ وذلك لأن تلك الخوارزمية اهتمت بحل مشكلة الغباش في الحدود الفاصلة بين الجزأين السليم والمرمم التي تعاني منها معظم خوارزميات ترميم الصور.

5-6 الخاتمة:

قدّم هذا البحث الإسهام في تقييم أداء خوارزميات ترميم الصور الرقمية من خلال تطبيق مجموعة صور مشوّهة موحدة على عدد من خوارزميات الترميم لتحقيق مقارنة وتقييم أداء صحيحين، ثم حساب بارامترات تقييم جديدة (MSE و NCC و EPIQA) بعد ما تبين أنها ذات تمييزية أفضل من البارامترات التقليدية (PSNR و SSIM) وتقييم أداء الخوارزميات اعتماداً على البارامترات المقترحة.

الآفاق المستقبلية

دراسة مدى فعالية البارامترات المقترحة لتقييم جودة الصور المرممة في تقييم جودة ترميم الفيديو.

:References المراجع

- 1- John G.Proakis ,Dimitris G.Manolakis , (1995) “Digital Signal Processing: Principles, Algorithms, and Applications: Discrete-Time Signals and Systems” 3rd edition, Prentice Hall,pp.118-131.
- 2-Hepi Hapsari Handayani,(2010), “DIGITAL IMAGE MATCHING METHOD USING NORMALIZED CROSS-CORRELATION (NCC) “Geoid Journal of Geodesy and Geomatics Vol. 06, No. 01
- 3- Alain Horé, Djemel Ziou , (2010) “Image quality metrics: PSNR vs. SSIM “2010 International Conference on Pattern Recognition.
- 4- C.Sasi varnan, A.Jagan, Jaspreet Kaur, Divya Jyoti, Dr.D.S.Rao, (2011) “Image Quality Assessment Techniques pn Spatial Domain” IJCST Vol. 2, Issue 3.
- 5- Yusra A. Y. Al-Najjar, Dr. Der Chen Soong , (2012) “Comparison of Image Quality Assessment: PSNR, HVS, SSIM, UIQI” International Journal of Scientific & Engineering Research, Volume 3
- 6- Ian J. Goodfellow, Jean Pouget-Abadie , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair , Aaron Courville, Yoshua Bengio (2014) “Generative Adversarial Nets” arXiv:1406.2661v1.
- 7- Olaf Ronneberger, Philipp Fischer, and Thomas Brox ,(2015) “U-Net: Convolutional Networks for Biomedical Image Segmentation” arXiv:1505.04597v1.
- 8- Barnes Connely (2017, January, Tu Th) “Introduction to Computer Vision, Neural Networks + Convolutional Neural Networks” lecture, University of Virginia
- 9- Rafael C. Gonzalez , Richard E. Woods, (2018) “Digital Image Processing: introduction “ 4th ed, Pearson Education Limited , pp . 18-24.
- 10- Yu Li, Tao Zhang , Shuyu Sun, Xin Gao,(2018) “Accelerating Flash Calculation through Deep Learning Methods” arXiv:1809.07311v2
- 11- Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Qureshi, and Mehran Ebrahimi, (2019) “Edgeconnect: Generative image inpainting with adversarial edge learning”. arXiv preprint arXiv:1901.00212. 1, 2, 3, 5, 6, 7 .
- 12- Yurui Ren, Xiaoming Yu, Ruonan Zhang, Thomas H. Li, Shan Liu, Ge Li , (2019) “StructureFlow: Image Inpainting via Structure-Aware Appearance Flow” The IEEE International Conference on Computer Vision (ICCV), pp. 181-190.
- 13- Xin Hong , Pengfei wiong , Renhe Ji , Haoqiang Fan, (2019) “ Deep Fusion Network for image completion” th 27th ACM International Conference on Multimedia .
- 14- Bruno Artacho , Andreas Savakis,(2019), “Waterfall Atrous Spatial Pooling Architecture for Efficient Semantic Segmentation” Sensors 2019, 19, 5361; doi:10.3390/s19245361

- 15- Omar Elharrouss, Noor Almaadeed, Somaya Al-Maadeed, Younes Akbari, (2019) "Image Inpainting: A Review", arXiv:1909.06399
- 16- M Sudhakara, M Janaki Meena, (2020) "An edge detection mechanism using L*A*B Color-based contrast enhancement for underwater images" Indonesian Journal of Electrical Engineering and Computer Science Vol. 18, No. 1, pp. 41~48.
- 17- Anirudha Ghosh , Abu Sufian , Farhana Sultana , Amlan Chakrabarti , Debashis De, (2020) "Fundamental Concepts of Convolutional Neural Network" Intelligent Systems Reference Library book series (ISRL, volume 172).
- 18- Laith Alzubaidi , Jinglan Zhang , Amjad J. Humaidi , Ayad Al-Dujaili³, Ye Duan , Omran Al-Shamma , J. Santamaría , Mohammed A. Fadhel , Muthana Al-Amidie and Laith Farhan, (2021) "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions" Journal of Big Data volume 8, Article number: 53.
- 19- Roza Dastres, Mohsen Soori , (2021) "Artificial Neural Network Systems" International Journal of Imaging and Robotics, Volume 21; Issue No. 2.
- 20- Aristeidis Chandrinos, (2021), "The Challenge of Predicting OAG Progression from the Initial Visual Field Test" , ResearchGate DOI: 10.13140/RG.2.2.20967.27049
- 21- Nabi Nabiyevev , Saber Malekzadeh, (2021), "Anomalous sound localization estimation" ResearchGate DOI: 10.13140/RG.2.2.25949.95201
- 22- Shiv Ram Dubey , Satish Kumar Singh , Bidyut Baran Chaudhuri, (2022) "Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark", arXiv:2109.14545v3.
- 23- Seyed Muhammad Hossein Mousavi, S. Muhammad Hassan Mosavi, (2022) "A New Edge and Pixel-Based Image Quality Assessment Metric for Colour and Depth Images " 9th Iranian joint congress on Fuzzy and Intelligent Systems (CFIS).
- 24- Prashant Kumar Dubey, Pratima Dubey , Mayank Sharma , Soni Changlani, (2022) "Comparative performance analysis of spatial domain filtering techniques in digital image processing for removing different types of noise" ISSN: 2008-6822 (electronic).
- 25- <https://neptune.ai/blog/generative-adversarial-networks-gan-applications>

Syrian Arab Republic

Ministry of High Education and Scientific Research

Damascus University

Faculty of Mechanical and Electrical Engineering

Department of Electronics and Communication Engineering



Performance Evaluation of Digital Image Inpainting Algorithms

A dissertation submitted in partial fulfillment of the requirements for the degree

Master in Applied Electronic Engineering

Presented by

Eng. Touka Al kayyal

Supervised by

Dr. Eng. Talal Hammod

Academic year

2022-2023