

## Detection of Vivax and Falciparum Malaria Using Deep Learning Techniques

Mohamad Hikmat Mohamad\*

(Received 7 / 12 / 2023. Accepted 22 / 2 / 2024)

### □ ABSTRACT □

Although malaria is uncommon in temperate climates, malaria is still common in tropical and subtropical countries. Every year, approximately 290 million people are infected with malaria, and more than 400,000 people die from this disease.

In this research, the dataset was downloaded from the Kaggle website, which consists of 27,558 images of infected people and healthy people, divided into: 13,779 images of healthy people, 6890 images of people infected with falciparum malaria, and 6889 images of people infected with vivax malaria. Initially, the images were pre-processed (fogification - noise removal - morphological operations), after which the convolutional neural network (CNN) was built, trained and tested, where the data was divided into two groups, 80% training and 20% testing, and then comparison was made between several pre-trained models. The proposed model achieved the best evaluation accuracy among the pre-trained models, and the results gave an accuracy rate of 96.5%, a sensitivity rate of 95%, and a specificity rate of 97.6%. Thus, the results demonstrate the effectiveness of the proposed algorithm as an assistive model for doctors in diagnosing malaria.

**Keywords:** convolutional neural networks , vivax malaria, falciparum malaria, sensitivity, specificity.

**Copyright**



:Tishreen University journal-Syria, The authors retain the copyright under a CC BY-NC-SA 04

---

\* Master - Industrial Automation Engineering – Tartus University – Tartus - Syria.  
mohamadhikmatmohamad@gmail.com

## الكشف عن الملاريا النشيطة والمنجلية باستخدام تقنيات التعلم العميق

محمد حكيمات محمد \*

(تاريخ الإيداع 7 / 12 / 2023. قُبِلَ للنشر في 22 / 2 / 2024)

### □ ملخص □

على الرغم من أن مرض الملاريا غير شائع في المناخات المعتدلة، فلا تزال الملاريا شائعة في البلدان الاستوائية وشبه الاستوائية. حيث يصاب كل عام ما يقرب من 290 مليون شخص بالملاريا، ويموت أكثر من 400 ألف شخص بسبب هذا المرض.

تم في هذا البحث تحميل مجموعة البيانات من موقع Kaggle وهي عبارة عن 27558 صور لأشخاص مصابين وأشخاص سليمين مقسمة إلى: 13779 صورة لأشخاص سليمين و 6890 صورة لأشخاص مصابين بالملاريا المنجلية و 6889 صور لأشخاص مصابين بالملاريا النشيطة. في البداية تمت المعالجة الأولية للصور (التعتيب - إزالة الضجيج - العمليات المورفولوجية)، بعد ذلك تم بناء الشبكة العصبونية الالتفافية CNN وتدريبها واختبارها، حيث تم تقسيم البيانات إلى مجموعتين 80% تدريب و 20% اختبار، ثم تمت المقارنة بين عدة نماذج مدربة مسبقاً. حقق النموذج المقترح أفضل دقة تقييم من بين النماذج المدربة مسبقاً وأعطت النتائج نسبة دقة 96.5% ونسبة الحساسية 95% ونسبة خصوصية 97.6% وبالتالي توضح النتائج فعالية الخوارزمية المقترحة كنموذج مساعد للأطباء في تشخيص مرض الملاريا.

الكلمات المفتاحية: الشبكات العصبونية الالتفافية - ملاريا نشيطة - ملاريا منجلية - الحساسية - الخصوصية.



حقوق النشر : مجلة جامعة تشرين - سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص

CC BY-NC-SA 04

\* ماجستير - قسم هندسة الأتمة الصناعية - كلية الهندسة التقنية - جامعة طرطوس - طرطوس - سورية.  
mohamadhikmatmohamad@gmail.com

**مقدمة:**

الملاريا مرض مميت وفقاً لتقرير الملاريا في العالم لعام 2019 الذي نشرته منظمة الصحة العالمية، الملاريا أو البُرْداء أو الأجمية هو مرض طفيلي معدي يسببه طفيلي (بلازموديوم)، ينتقل عن طريق لعاب أنثى بعوض الأنفوليس، والمعروفة باسم ناقلات الملاريا malaria vectors. لقد وجدت منذ 30 مليون سنة وتم تحديدها على أنها سبب رئيسي للوفاة في الحضارات القديمة في جميع أنحاء العالم اليوم، لا تزال الملاريا مرضاً خطيراً حيث يتعرض ما يقارب من نصف سكان العالم للخطر. لعلّ الرغم من أن منظمة الصحة العالمية قد حددت المنظمة الإفريقية على أنها تحمل نصيباً كبيراً بشكل غير متناسب مع عبء الملاريا العالمي. حيث تضم المنظمة 92% من سكان العالم، حالات الملاريا 93% من وفيات الملاريا في عام 2017.

كان هناك ما يقدر بنحو 405000 حالة وفاة، المرض قابل للشفاء عند الاكتشاف المبكر له تشمل الأساليب الحالية المستخدمة للكشف عن الملاريا الكشف المجهرى عن الخلايا المصابة في المختبر. هذه الطريقة باهظة الثمن يُقترح بديل أسرع ومنخفض التكلفة وموثوق به للكشف المجهرى عن الملاريا وذلك بتطوير خوارزميه للكشف عن خلايا الملاريا باستخدام تقنية معالجه الصور [1]، كما يعطينا كل الدوافع لبناء نموذج يمكنه اكتشاف الخلايا من صور خلايا متعددة في الدم على شرائح المجهر القياسية وتصنيفها على أنها سواء كانت مصابة أو غير مصابة بالاختبارات المبكرة والفعالة وإجراء التصنيف على صورة خلية مصابة باستخدام تقنيات تعلم الآلة [2].

مشكلة البحث:

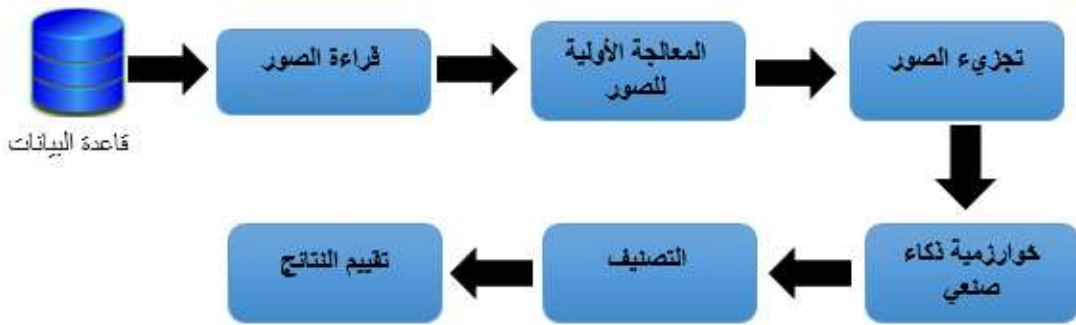
تركز مشكلة البحث الرئيسية في بناء نموذج للتعرف التلقائي مساعد للطبيب الذي يقوم بتشخيص الملاريا آلياً وذلك بتحليل صور لعينات دم مصابة ويعطي تشخيصاً دقيقاً عما إذا كان المريض مصاب بالملاريا أم لا باستخدام تقنيات معالجة الصورة وذلك يوفر الكثير من الوقت الطويل الذي يحتاجه الطبيب للتشخيص والمخبري من أجل التحليل كما يقلل الخطأ الذي يمكن أن يحصل (الخطأ البشري) فضلاً عن الدقة في التشخيص.

**أهمية البحث وأهدافه:**

إن استخدام الأتمتة الحاسوبية في هذا المجال أصبح أمراً في غاية الأهمية نظراً لأنه سيعطي نتائج أكثر دقة تساعد الأطباء في تحسين النتائج وتخفيض معدل الخطأ، والذي من الممكن أن يساعد في إنقاذ أرواح البشر. يهدف البحث إلى تطوير طريقة جديدة لاكتشاف مرض الملاريا والذي يمكن أن يحل محل طرق التشخيص التقليدية الحالية من قبل الطبيب ويساهم في تخفيض عدد عينات الدم والمسحات غير الضرورية وزيادة الموثوقية وذلك بالمساعدة على التشخيص الصحيح.

**طرائق البحث ومواده:**

تم في البداية جمع الصور لإنشاء قاعدة البيانات اللازمة حيث تم جمع 27558 صورة، ومن ثم تم معالجة الصور عن طريق تطبيق العمليات الشكلية (المورفولوجية) وبعض تقنيات معالجة الصورة ويوضح الشكل (1) المخطط الصندوقي لأهم مراحل العمل في هذا البحث.



الشكل (1): المخطط الصندوقي لخوارزمية العمل.

الدراسات المرجعية:

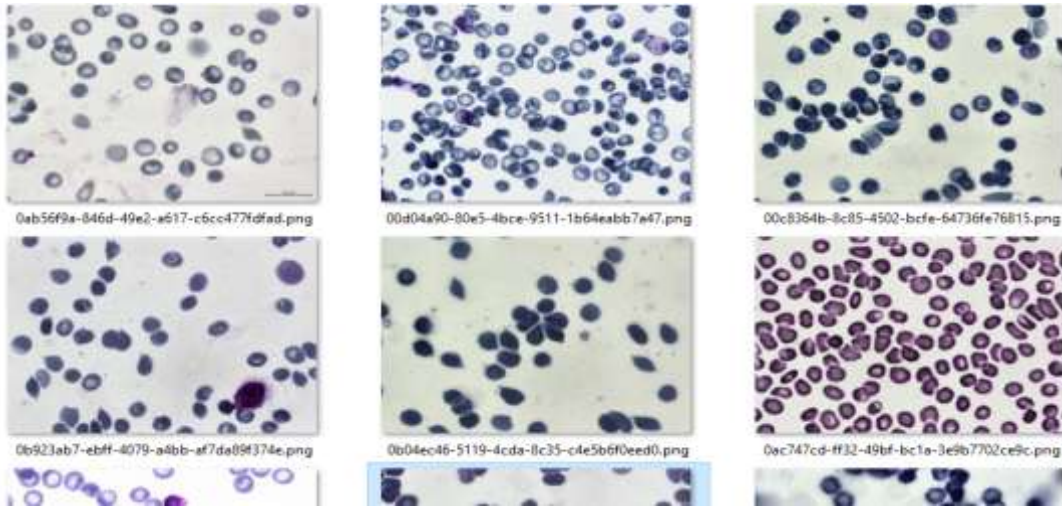
- في عام 2019 [1] قام الباحثون Mahesh Jangid و Sanidhya Kumar و Satabdi Nayak بتدريب نماذج مختلفة من التعلم العميق وراقبوا أياً من هذه النماذج يوفر دقة أفضل ودقة أسرع من نماذج التعلم العميق المستخدمة سابقاً. أظهرت النتائج أن نموذج Resnet 50 أعطى أعلى دقة بلغت 95%.
- في عام 2020 [2] قام الباحثون Muhammad Muinul Islam و Md. Mohidul Hasan Sifat باقتراح الاكتشاف التلقائي لطفيليات الملاريا بمراحلها من مسحة الدم في هذا العمل. تم جمع صور مسحة الدم لمرضى الملاريا من قاعدة البيانات على الإنترنت. بعد تطبيق بعض المعالجة المسبقة، تم استخدام U-Net لتقسيم RBC (خلايا الدم الحمراء) من صور مسحة الدم ، وتم استخدام شبكة CNN للتعرف على كرات الدم الحمراء المصابة بطفيليات الملاريا ، وأخيراً ، تم استخدام شبكة عصبية VGG16 للتعرف على الأنواع المختلفة ومراحل الملاريا. كانت دقة التجزئة وخصوصية نموذج U-Net 97.67% و 92.05% على التوالي. كانت دقة الكشف ونوعية كرات الدم الحمراء المصابة 100% و 95% باستخدام نموذج CNN ، متوسط دقة وخصوصية نموذج VGG16 للكشف عن أنواع الملاريا 95.55% و 94.75% على التوالي. كان متوسط دقة التدرج والنوعية لأنواع مختلفة من طفيليات الملاريا للمرحلة الحلقية 96.25% و 94.82% على التوالي بتطبيق VGG16.
- في عام 2020 [3] قام الباحثون Guang Z. Yang و Bruno M. G. Rosa بتقديم أداة جديدة للكشف السريع عن الملاريا تعتمد على محلل مقاومة الدم المحمول. يسمح الحقن المتزامن متعدد الألوان للكشف عن التيار والجهد بالجهاز بتقليل وقت الفحص (ترتيب الثواني) مع تعزيز الاختلافات في إشارة الممانعة المسجلة بين الترددات لزيادة مستوى التمييز في مستوى الطفيليات. ذهبوا إلى أبعد من ذلك لإثبات إمكانية تطبيق الجهاز مباشرة على عينات الدم التي تم جمعها من المتطوعين للتمييز بين العينات المصابة وغير المصابة. كان الحد الأدنى من مستوى التطفل في الدم الذي تم تمييزه 0.0078% (أو ما يعادل 390 طفيلياً / ميكرو لتر) ، وهو ما يقترب عن كثب من المعيار الدولي المحدد عند 200 طفيلي / ميكرو لتر. أخيراً ، نظهر أيضاً الاعتماد على التردد لعينات الدم المتزامنة وغير المتزامنة ، والتي يمكن أن تشكل الأساس لتطوير طريقة نقطة رعاية وخالية من المختبر للتمييز بين المراحل المختلفة لعدوى الملاريا في المجتمعات الراكدة اقتصادياً.
- في عام 2018 [4] قام الباحثون P. L. Paikrao (Assistant و Bhaiyya D. Ghanmode Professor) بالتحديد الدقيق لطفيل الملاريا حيث أن المعيار التقليدي قد يتغلغل فيه الخطأ البشري. كما أن الوقت

المطلوب لإجراء الاختبار أكثر. يمكن للخبير مراجعة عدد قليل من الشرائح المصابة. لذلك تم اقتراح هذه التقنية التي تتبع المعالجة المسبقة لتصنيف الصورة وتقسيم الصور وأخيراً الكشف عن طفيلي الملاريا.

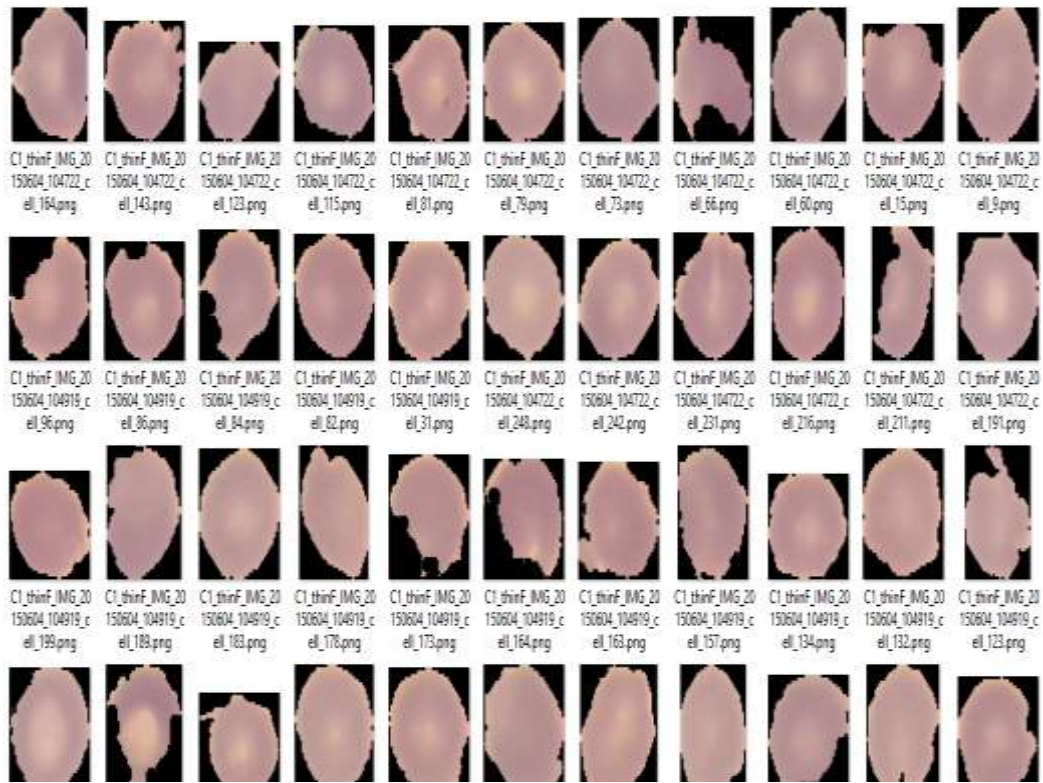
• في عام 2021 [5] قام الباحثون Abhishek Kanojiya و Abhishek Yadav و Kirti Motwani ببناء نموذج يمكنه اكتشاف الخلايا من صور خلايا متعددة في مسحة دم رقيقة على شرائح مجهرية قياسية وتصنيفها على أنها إما مصابة أو غير مصابة من خلال الاختبار المبكر والفعال باستخدام معالجة الصور. وأيضاً إجراء التصنيف على صورة الخلية المصابة باستخدام التعلم الآلي، حيث تم استخدام خوارزمية Cubic SVM وكانت دقة التصنيف عند استخدام هذه الخوارزمية 86.1%.

#### 1- تحضير البيانات:

تم الحصول على قاعدة بيانات خارجية مستخدمة حالياً، تم تحميلها من موقع Kaggle والتي تحتوي على 27558 صورة لخلايا الدم ومقسمة إلى قسمين الأول 13779 صورة لأشخاص غير مصابين والثاني 13779 صورة لأشخاص مصابين (6890 ملاريا منجلية - 6989 ملاريا نشيطة) ويبين الشكل (2) صور من قاعدة البيانات لأشخاص مصابين والشكل (3) صور من قاعدة البيانات لأشخاص غير مصابين.



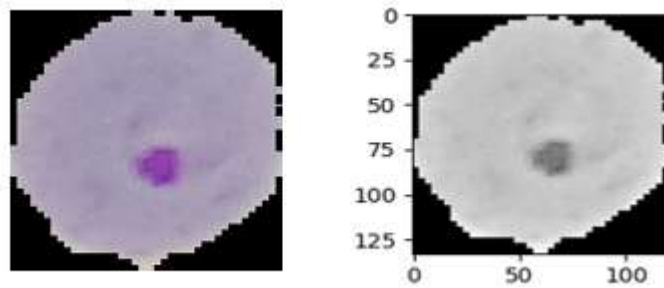
الشكل (2): صور من قاعدة البيانات لأشخاص مصابين.



الشكل (3): صور من قاعدة البيانات لأشخاص غير مصابين.

#### 1-1- المعالجة الأولية:

في البداية قمنا بأخذ صورة من قاعدة البيانات وطبقنا عليها مراحل المعالجة ومن ثم وبنفس الطريقة والأسلوب تم التطبيق على قاعدة البيانات ومن بعدها تم البدء بعملية المعالجة المبدئية. بما أن الصور 8 units أي مكعبة بـ 8 بتات سيتم تحويل الصور الملونة الى صور رمادية كما في الشكل (4) حيث يجب أن تكون جميع الصور الداخلة إلى الشبكة نفس النوع والحجم والتدرج اللوني [6].



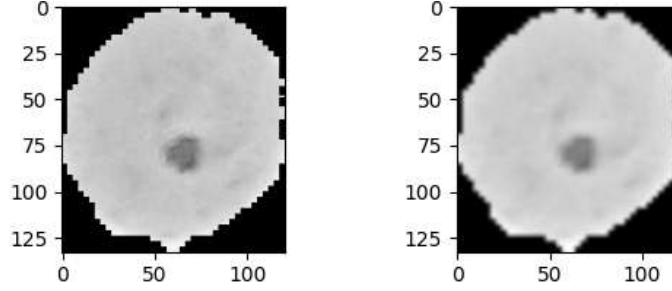
الشكل (4): الفرق بين الصورة الأصلية والصورة الرمادية

#### 1-2- التخلص من الضجيج:

➤ المرشحات الخطية:

هي المرشحات المتوسطة مثل مرشح المتوسط الحسابي، مرشح المتوسط الهندسي، مرشح المتوسط التوافقي وكل مرشح يناسب نوع ضجيج مختلف مثلاً مرشح المتوسط التوافقي يناسب الضجيج النبضي أما المرشحات الحسابية والهندسية

فهي مناسبة للضوضاء العشوائية مثل مرشح Gaussian، يتميز هذا المرشح أنه يقوم بتنعيم الصورة وفي نفس الوقت الحفاظ على أجزاء الصورة الأصلية وأبرزها الحواف ويبين الشكل (5) ناتج تطبيق المرشح الغاوسي على الصورة [6-7].

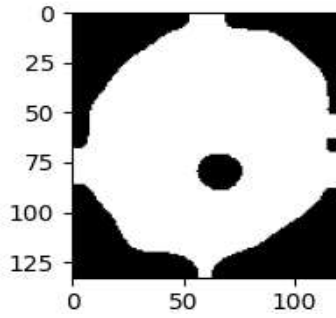


الشكل (5): الفرق بين الصورة الأصلية والصورة بعد تطبيق المرشح (مرشح Gaussian)

1-3- تجزئة الصورة:

a. تعتیب الصورة:

قمنا بعملية التعتیب لتحويل الصورة الرمادية (المرشحة) إلى ثنائية وتم ذلك بشكل آلي باستخدام تقنية أوتسو وذلك حسب كل صورة لكونها عملية أساسية وضرورية للحصول على المركبات المتصلة وتطبيق العمليات الشكلية عليها ويظهر الشكل (6) الصورة بعد عملية التعتیب.

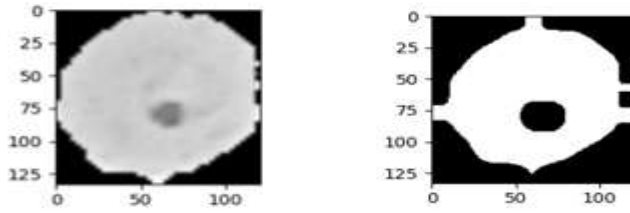


الشكل (6): الصورة بعد تطبيق عملية التعتیب.

b. العمليات المورفولوجية:

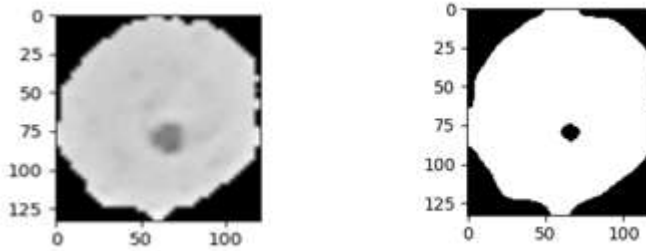
تتكون الصورة المستخدمة من منطقة بيضاء ومنطقة سوداء وتكون المنطقة السوداء (النقطة السوداء) التي تحتوي على المرض لذلك من خلال التحويلات الشكلية (لمورفولوجية) يمكن استخراج منطقة المرض حيث أن العمليات الشكلية هي بعض العمليات البسيطة التي تجري على شكل الصورة وتعد أداة لاستخراج مكونات الصورة المفيدة، تنفذ على الصور الثنائية يحتاج لتنفيذ هذه العملية إلى دخلين الأول هو الصورة الأصلية والدخل الثاني هو العنصر التركيبي structuring element أي المصفوفة المسؤولة عن تحديد القناع kernel الذي يحدد طبيعة العملية. يتم عادة اختيار العنصر التركيبي بشكل مشابه للكائن المراد معالجته في الصورة المدخلة لذلك تم اختيار العنصر البيضي ellipse، يوجد عمليتان أساسيتان وهما عملية التآكل Erosion وعملية التمدد Dilation [6-7].

❖ التآكل: هو عملية تقلص الكائنات في الصورة، تؤدي هذه العملية إلى تنعيم حدود السطح الأمامي (الذي يكون باللون الأبيض) حيث تقوم هذه العملية بإنقاص البكسلات ذو القيمة 1 وفق القناع المحدد وتؤدي بذلك إلى تقليل المساحات البيضاء، وعزل العناصر عن بعضها كما في الشكل (7).



الشكل (7): الصورة بعد عملية التآكل

❖ التمدد: هو عملية تكبير الكائنات في الصورة وهي عكس عملية التآكل وبالتالي يزداد حجم المنطقة البيضاء كما في الشكل (8).



الشكل (8): الصورة بعد عملية التمدد

#### 1-4- استعادة الصورة:

يمكن من خلال العمليات الشكلية الحصول على موقع الورم وحجمه ولكن نفس الوقت تؤدي هذه العمليات إلى إزالة الحواف من الصورة حيث الحواف هي خيط رفيع يمر بين المناطق غير المشتركة في الصورة أو المجالات اللونية غير المتجانسة والتي تظهر بشكل مفاجئ وبالتالي من أجل ترميم الصورة يجب استرجاع الحواف.

A. مرشحات كشف الحواف في المجال الفراغي:

يقصد بالحافة في الصورة الرقمية التغير المفاجئ في قيم السويات الرمادية وقد يكون هذا التغير متدرجا أو حادا. لذلك يتم التعامل مع مرشحات المشتق الأول والثاني:

إن المشتق الأول يظهر مكان الحواف في الصورة أما المشتق الثاني فيظهر حواف أكثر لأنه سينتج حافتين الأولى ذات قيمة عالية والثانية ذات قيمة منخفضة المشتق هنا هو (القيمة التالية-القيمة السابقة) يأخذ قيمة 1 لدى الانتقال من العمود ذو اللون الأسود إلى العمود ذو اللون الأبيض مشيرا إلى وجود حافة في هذه المنطقة من أهمها مرشح سويل الذي يعتمد على كشف حواف الصورة في اتجاه واحد فقط (عمودي - أفقي-شاقولي -قطري )، ومرشح روبيرت الذي يتميز بقناع صغير يكشف الزوايا المائلة إما 45 أو -45.

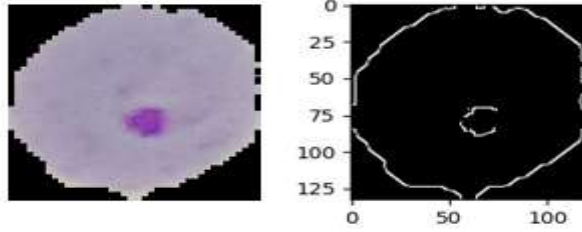
المشق الثاني يعطي قيمة (1) ثم قيمة (-1) أي قيمة عليا وقيمة دنيا مشيرا إلى أننا عبرنا بالصفير ورافعا قيمة الفرق بين السويات ومحددا بالضبط مكان الحافة وبالتالي المشتق الثاني أدق. من أهم هذه المرشحات مرشح اللابلاسيان الذي يقوم بكشف الحواف الأفقية والعمودية أو الأفقية والعمودية والمائلة معاً ويعتبر من أقوى المرشحات، ومرشح كاني Canny يقوم باستخلاص الحواف اعتمادا على مرحلتين أساسيتين [7]:

- تتعيم الصورة لتخفيف أثر الحواف.
- انتقاء الحواف بالاعتماد على عتبتين قوية وضعيفة فالعتبة القوية تنتج حوفا أما العتبة الضعيفة فتعمل على أخذ الحواف الضعيفة المرتبطة بالحواف القوية وإهمال باقي الحواف لذلك نحص على حواف أكثر دقة. حيث اذا كانت



الحواف ذات تدرج أكبر من العتبة القوية يتم أخذها أما إذا كانت ذات تدرج أقل ترفض مباشرة، وإن كانت بين العتبتين يتم قبولها إذا كانت على صلة بالحواف القوية إن كانت مجاورتها في الجوار الثماني قوية، أما إن كانت ضعيفة تحذف لأنها غير مرتبطة بحواف قوية .

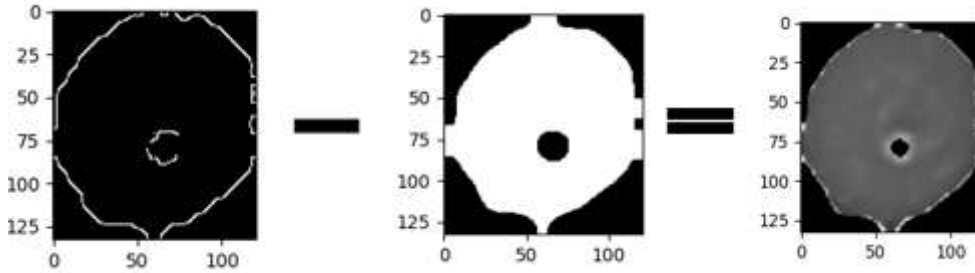
تم الحصول على حواف الصورة الأصلية من خلال مرشح كاني Canny وتم اجراء مسح للصورة لحذف عناصر الصورة التي لا تشكل الحافة، ويمتاز هذا الكاشف في أنه يقوم بتأدية عدة خطوات خلال عملية الكشف ينتج منها كشف دقيق للحواف. يظهر الشكل (9) الصورة المعتبة وبعد كشف الحواف



الشكل (9) الصورة المعتبة وبعد كشف الحواف

B. دمج الصورتين:

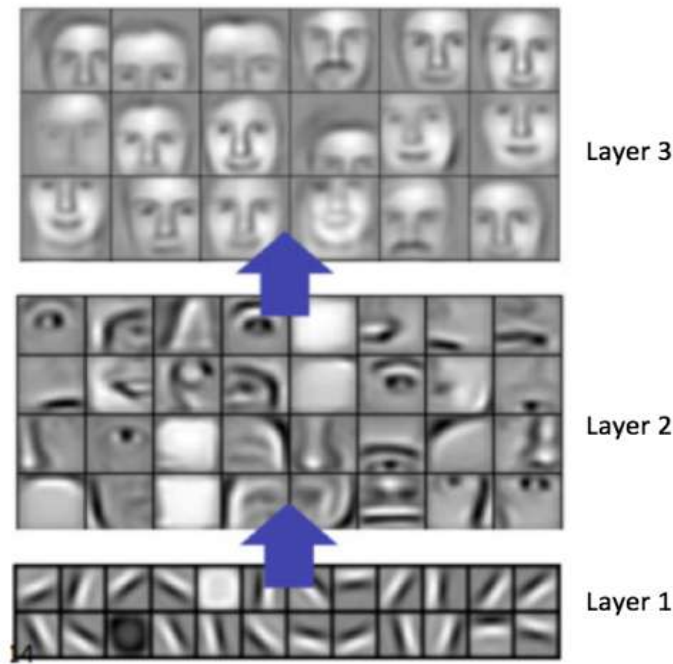
يتم بعد كشف الحواف إعادة بناء الصورة الشكل (10) من خلال دمج الصورة التي تم الحصول عليها بعد تطبيق العمليات الشكلية مع حواف الصورة الأصلية، ويظهر الشكل عملية الدمج والصورة الناتجة.



الشكل (10): نتيجة دمج الصورتين.

## 2- بنية الشبكات العصبونية العميقة:

تتكون الشبكة العصبونية العميقة من عدة طبقات: طبقة الالتفاف، طبقة الدمج، والطبقة المتصلة بالكامل. في المستوى الأول من التعامل مع صورة الإدخال، عادة ما تتعلم الشبكة على العصبية ميزات بصرية بسيطة مثل الحواف أو اللون، ثم في المستوى الثاني يجمع بين خصائص المستوى السابق. تؤدي إضافة المزيد من المستويات إلى العثور على ميزات ذات مستوى أعلى والشكل (11) يوضح ميزات كل مستوى [8].



الشكل (11): الميزات التي نحصل عليها في كل مستوى.

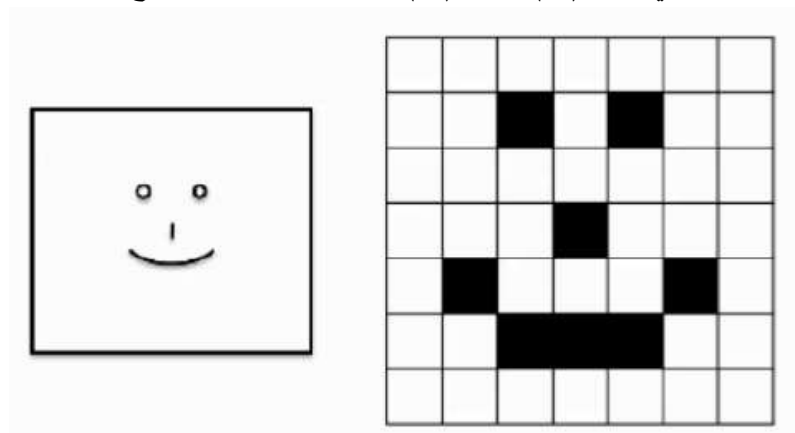
a. طبقة الالتفاف:

تعد طبقة الالتفاف أهم جزء في CNN وتستخدم دائماً كطبقة أولى. تتمثل المهمة الرئيسية لهذه الطبقة في تحديد السمات الموجودة في المناطق المحلية لصورة الإدخال، والتي تكون مشتركة لمجموعة البيانات بأكملها. يؤدي التعرف على الميزة إلى إنتاج خريطة السمات عن طريق تطبيق عوامل الفلترة، حيث تطبق طبقة الالتفاف فلتراً محلياً على صورة الإدخال، حيث ينتج أفضل ترابط بين وحدات البكسل الخاصة بالصورة المدخلة.

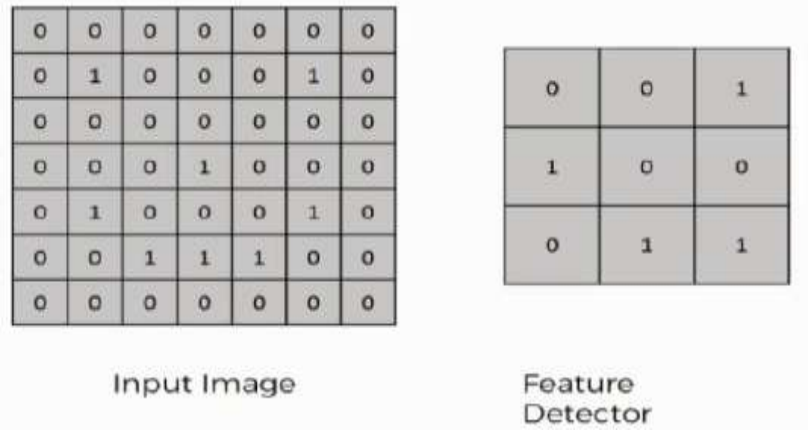
هناك ثلاث مكونات مهمة في طبقة الالتفاف [8-9]:

- 1- صورة الدخل.
- 2- كشف الميزات.
- 3- خريطة السمات.

لنفرض لدينا الصورة الأصلية المبينة في الشكل (12) والشكل (13) الصورة بعد تحويلها إلى نموذج إدخال.

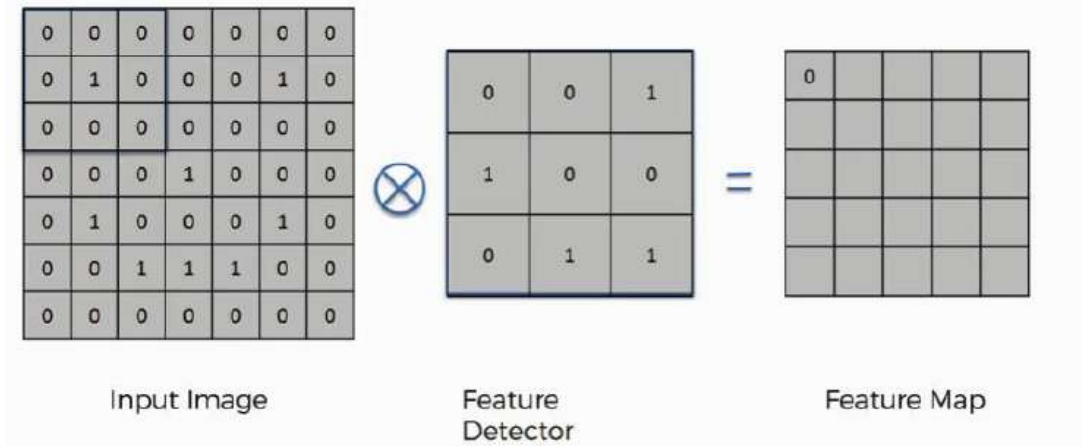


الشكل (12): صورة الدخل.

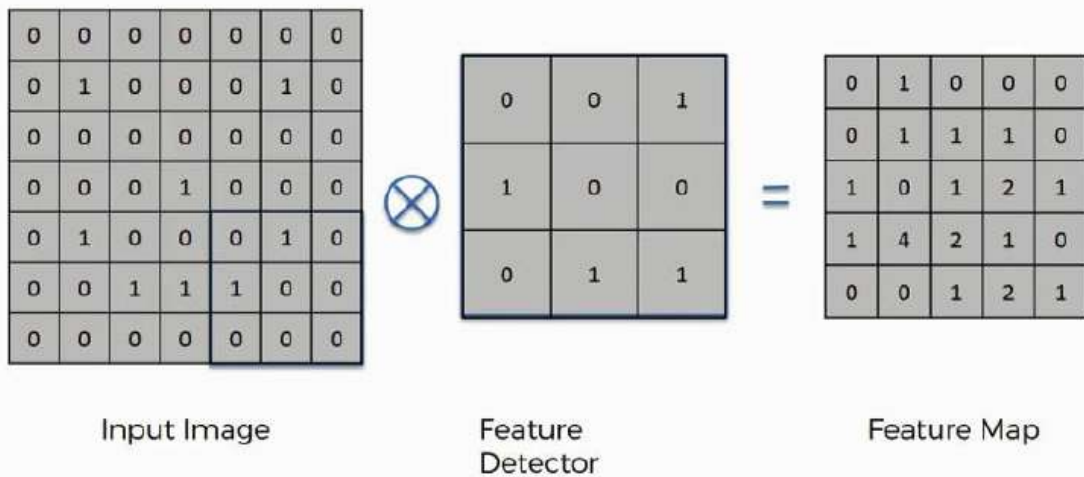


الشكل (13): الصورة المدخلة بعد تحويلها.

يستخدم الشكل (13) كاشف الميزة، ويسمى أيضاً بالفلتر بقياس  $3 \times 3$  يطبق هذا الفلتر على صورة الإدخال كما هو موضح بالشكل (14) وينطبق عامل الالتفاف على القيمة الأولى لخريطة السمات الجدولية، ثم تتكرر العملية ذاتها لإكمال الصف الأول من خريطة السمات، بعد الإنتهاء من الصف الأول ينتقل إلى الصف التالي ويتم تكرار هذه العملية حتى تكتمل خريطة السمات بالكامل والشكل (15) يوضح ذلك.



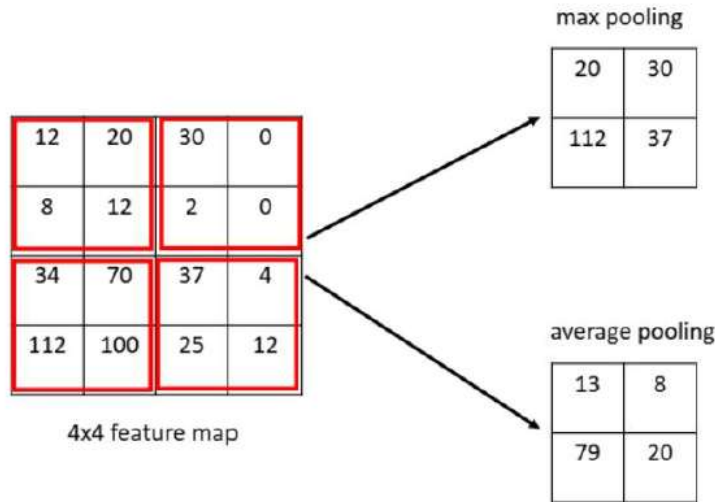
الشكل (14): كيفية الحصول على خريطة الميزات.



الشكل (15): خريطة الميزات.

b. طبقة الدمج:

عادة ما يتم استخدام هذه الطبقة بشكل دوري بين طبقتين متتاليتين من الالتفاف وتتمثل مهمتها في تقليل حجم خرائط السمات بالإضافة إلى استخراج الميزات المهمة في خريطة السمات. هناك طريقتين مهمتين من طبقات الدمج هما (max- average) والشكل (16) يبين الفرق بين الطريقتين [9].



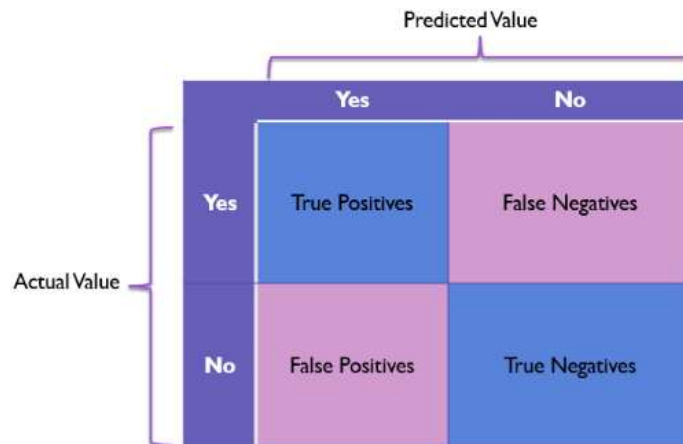
الشكل (16): الفرق بين max-average.

c. الطبقة المتصلة بالكامل:

هي طبقة لتصنيف البيانات وتعتبر طبقة تقليدية متعددة الطبقات تستخدم تابع التنشيط (SoftMax) مع إمكانية استخدام دوال أخرى، ويشير مصطلح (كاملة الاتصال) إلى أن كل عصبون في الطبقة السابقة متصلة بكل عصبون في الطبقة التالية. يمثل الخرج من طبقتي الالتفاف والدمج خصائص عالية المستوى للصورة المدخلة وتعمل الطبقة كاملة الاتصال على استخدام هذه الخصائص لعملية تصنيف الصورة المدخلة إلى عدة فئات بناءً على تدريب البيانات [10].

3- مصفوفة الارتباك Confusion Matrix :

تعد مصفوفة الارتباك بأنها جدول يُستخدم لتقييم أداء نموذج التصنيف حيث توضح عدد التنبؤات الصحيحة وغير الصحيحة التي قدمها النموذج مقارنة بالنتائج الفعلية. الشكل (17) يبين تمثيل مصفوفة الارتباك في بأربعة مكونات [11].



الشكل (17): مصفوفة الارتباك .

حيث:

1. إيجابية صحيحة (TP): هذه القيمة تعبر عن عدد الحالات الإيجابية الفعلية التي توقعها النموذج بشكل صحيح على أنها إيجابية.
2. سلبية صحيحة (TN): أما في ما يتعلق بهذه القيمة فهي تعبر عن عدد الحالات السلبية الفعلية التي توقعها النموذج بشكل صحيح على أنها سلبية.
3. إيجابية خاطئة (FP): قيمة FP تعبر عن عدد الحالات السلبية الفعلية التي توقعها النموذج بشكل غير صحيح على أنها إيجابية.
4. سلبية خاطئة (FN): تعبر عن عدد الحالات الإيجابية الفعلية التي توقعها النموذج بشكل غير صحيح على أنها سلبية.

-4 بعض المؤشرات المستمدة من مصفوفة الارتباك:

1. الصحة (accuracy): وهي عدد العينات التي تم توقعها بشكل صحيح على العدد الإجمالي للعينات وتعطى بالعلاقة [11]:

$$Acc = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \quad (1)$$

2. الحساسية (recall or sensitivity): وهي عدد العينات الصحيحة التي تم كشفها على عدد العينات الصحيحة الإجمالي، وتعطى بالعلاقة [11]:

$$SN = \frac{Tp}{Tp + Fn} \quad (2)$$

3. الخصوصية (Specificity): عدد العينات الخاطئة التي تم كشفها على العدد الإجمالي للعينات الخاطئة، وتعطى بالعلاقة [12]:

$$SP = \frac{Tn}{Tn + Fp} \quad (3)$$

4. مقياس القيمة المتوقعة الخاطئة (Negative Predicated Value): عدد العينات الخاطئة التي تم توقعها على العدد الإجمالي للعينات التي تم توقعها خاطئة، وتعطى بالعلاقة [12]:

$$NPV = \frac{Tn}{Tn + Fn} \quad (4)$$

5. المقياس (F1-score): ويعبر عن العلاقة التوافقية بين المقياسين (precision) و (recall)، ويعطى بالعلاقة [12]:

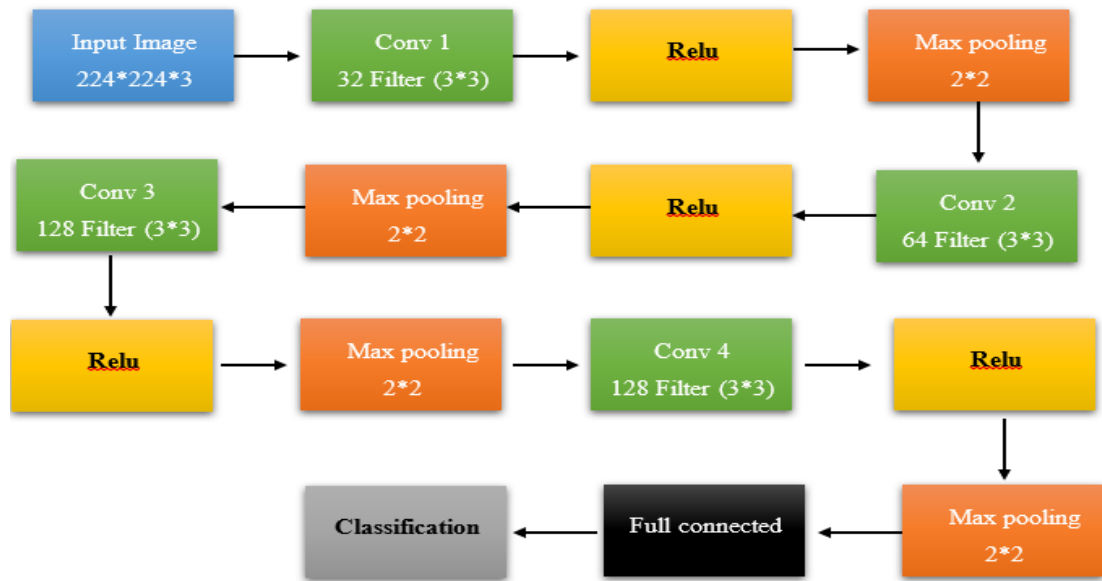
$$F1 - score = \frac{2 * Tp}{2 * Tp + Fn + Fp} \quad (5)$$

6. بناء نموذج الشبكة:

يوضح الشكل (18) نموذج الشبكة المستخدم للكشف عن الأشخاص المصابين بالملاريا حيث تتكون هذه الشبكة من:

1. مرحلة طبقة النفاق أولى تتكون من 32 نواه مختلفة بخطوة اختزال 3×3 تعطينا طبقة واحدة، ثم يلي ذلك تطبيق تابع التفعيل Relu ، يليها مباشرة طبقة تجميع أقصى مع حجم خطوة 2.
2. مرحلة طبقة النفاق ثانية تتكون من 64 نواه مختلفة بخطوة اختزال 3×3 تعطينا طبقة واحدة، ثم يلي ذلك تطبيق تابع التفعيل Relu ، يليها مباشرة طبقة تجميع أقصى مع حجم خطوة 2.
3. مرحلة طبقة النفاق ثالثة تتكون من 128 نواه مختلفة بخطوة اختزال 3\*3 تعطينا طبقة واحدة، ثم يلي ذلك تطبيق تابع التفعيل Relu ، يليها مباشرة طبقة تجميع أقصى مع حجم خطوة 2.
4. مرحلة طبقة النفاق رابعة تتكون من 128 نواه مختلفة بخطوة اختزال 3\*3 تعطينا طبقة واحدة، ثم يلي ذلك تطبيق تابع التفعيل Relu ، يليها مباشرة طبقة تجميع أقصى مع حجم خطوة 2.
5. مرحلة خامسة طبقة كاملة الاتصال تقوم بمعالجة السمات المستخرجة وتستخدم تابع التفعيل Relu.

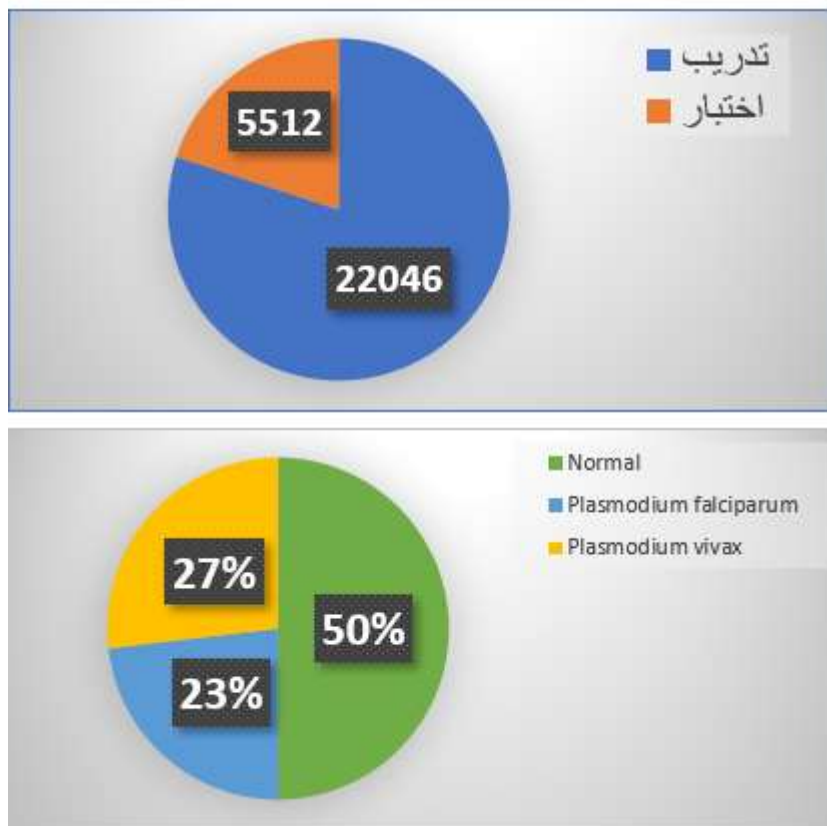
6. طبقة الخرج (التصنيف) وهي طبقة كاملة الاتصال وتستخدم تابع التفعيل Softmax.



الشكل (18): بنية الشبكة المصممة.

a. تدريب الشبكة:

قمنا بتقسيم البيانات إلى مجموعتين (80% تدريب - 20% اختبار) كما هو موضح بالشكل (19) وبيين الجدول (1) البارامترات التي تم ضبط فيها الشبكة المصممة.

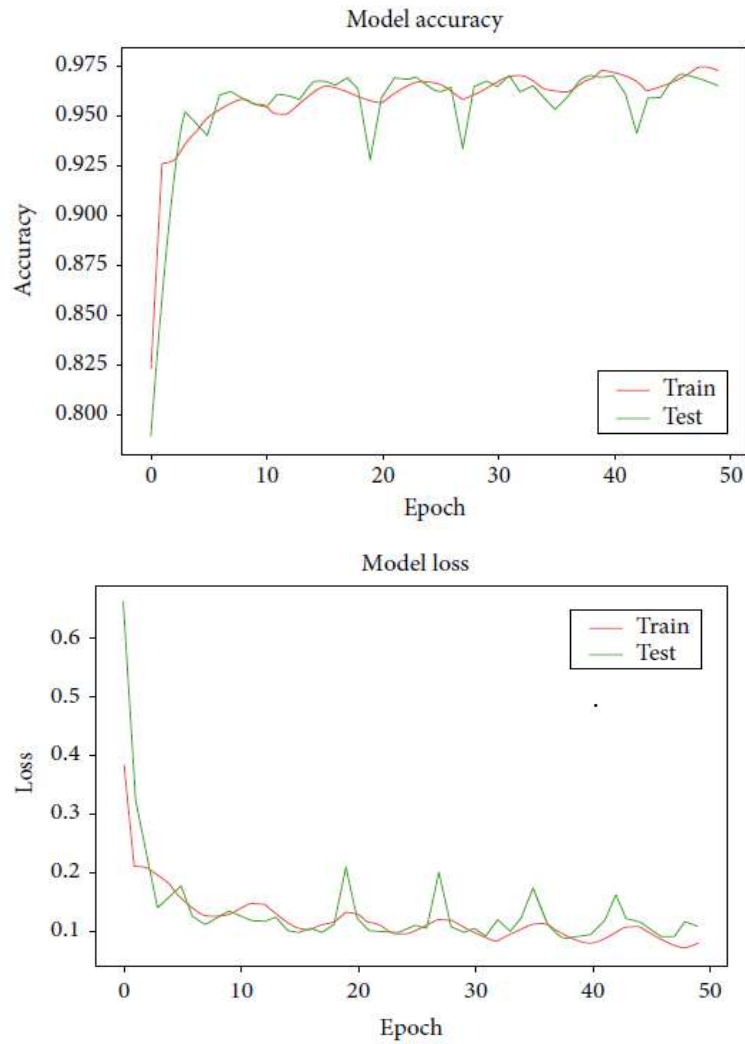


الشكل (19): تقسيم البيانات.

الجدول (1): بارامترات الشبكة

Parameter	Value
Epochs	50
Optimizer	Adam
Batch size	32
Input shape	224 × 224
Pooling	Max 2 × 2 (convolutional layers, flatten layer)
Activation	ReLU (convolutional layers), Softmax (final dense layer)

بعد الانتهاء من عملية التدريب والاختبار قمنا برسم منحنيات الدقة والخطأ لكل من معطيات التدريب والاختبار والشكل (20) يبين ذلك.

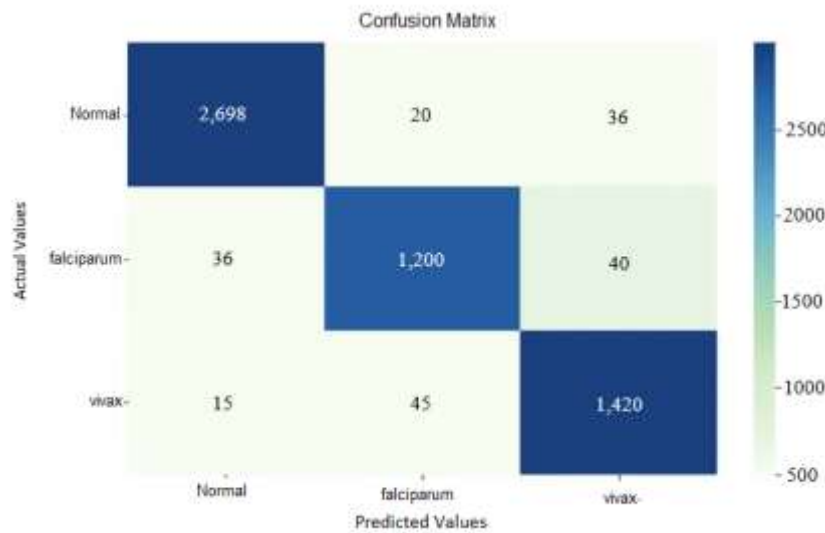


الشكل (20): منحنيات الدقة والخطأ لكل من معطيات التدريب والاختبار.

من الشكل (20) نلاحظ عدم وجود Overfitting وذلك لأن المنحني الذي يمثل دقة التدريب في مجموعة البيانات ارتفاح دقة النموذج بشكل ملحوظ بعد كل مرحلة تدريب اما بالنسبة للمنحني الذي يوضح دقة الاختبار فإنه يبين ارتفاح نسبة دقة الاختبار مع زيادة الدقة الخاصة بالتدريب، وكذلك الامر بالنسبة للمنحنيات التي تشير إلى معدل الخطأ الذي ينتجه النموذج يتضح أن معدل الخطأ لمجموعة التدريب ينخفض بعد كل (epoch) ونلاحظ أن معدل الخطأ لمجموعة الاختبار ينخفض مع انخفاض معدل الخطأ لمجموعة التدريب وبالتالي فإن النموذج قادر على التعميم لأنه تدرّب بشكل جيد أثناء مرحلة التدريب وعند اختباره على بيانات جديدة لم يتدرّب عليها أعطى نتائج جيدة.

#### b. تقييم النموذج:

لتوضيح نتائج نموذجنا قمنا برسم مصفوفة الارتباك حيث يتبين من خلالها نسبة الحالات التي توقع فيها النموذج بشكل صحيح وكذلك تبيّن الحالات التي فشل النموذج فيها بتوقع الخرج الصحيح، والشكل (21) مصفوفة الارتباك للشبكة المقترحة، ويبين الجدول (2) نتائج أداء النموذج بعد التدريب.



الشكل (21): مصفوفة الارتباك .

نلاحظ أنه من بين 1256 مريض مصاب بالملاريا (*Plasmodium falciparum*)، استطاع النموذج التنبؤ بشكل صحيح لـ 1200 شخص مريض وأخطأ النموذج بـ 40 شخص مصاب بالـ *falciparum* صنفهم على أساس *Plasmodium vivax* و 36 شخص مصاب صنفهم على أنهم Normal، وقام النموذج بتصنيف 1420 شخص مصاب بالملاريا (*Plasmodium vivax*) بشكل صحيح وأخطأ بتصنيف 45 شخص على أنهم مصابين بالـ *Plasmodium falciparum* و 15 شخص صنفهم على أساس أنهم Normal وهم مصابين، كما صنف 2698 شخصاً غير مصاب بشكل صحيح، وأخطأ النموذج بتصنيف 56 غير مصابين بالملاريا على أنهم مصابين.

الجدول (2): نتائج أداء النموذج بعد التدريب .

ACC	SN	SP	F1-Sorce
%96.5	%95	%97.6	%96.5

7. مقارنة مع الدراسات المرجعية:



قمنا بمقارنة النتائج التي حصلنا عليها بالدراسات المرجعية السابقة والجدول (3) يوضح ذلك

الجدول (3): مقارنة النتائج مع الدراسات المرجعية السابقة.

الدراسة	النموذج المستخدم	الدقة
Satabdi Nayak [1]	Resnet 50	95%
Md. Mohidul Hasan Sifat [2]	VGG16	94.8%
Abhishek Kanojiya [5]	Cubic SVM	86.1%
الدراسة الحالية	CNN	96.5%

نلاحظ من الجدول السابق أن نموذجنا حقق أفضل دقة في تشخيص الأشخاص المصابين بالملاريا 96.5% بالإضافة أنه يمكنه تحديد نوع الملاريا المصاب بها المريض.

### الاستنتاجات والتوصيات:

#### الاستنتاجات:

في هذه الدراسة، تم بناء نموذج التعلم العميق متعدد التصنيفات وتقييمه للكشف عن الملاريا. ويعد هذا النموذج هاماً لكونه يقوم بتصنيف أنواع أمراض الملاريا في نموذج واحد. ومن المهم تشخيص هذه الأمراض بشكل صحيح لتحديد العلاج المناسب.

من خلال النتائج التي تم الحصول عليها تفوق نموذج CNN على النماذج الأخرى المقترحة في الدراسات المرجعية، حيث حقق نموذج CNN دقة 96.5 %، وتمكن النموذج من التدريب على التصنيف المتعدد، كما أنه نلاحظ وجود تقارب كبير بين منحنى التدريب والاختبار وهو مؤشر جيد لغياب Over-fitting.

#### التوصيات:

- 1- نوصي بدمج أكثر من نوع من الشبكات العصبونية العميقة للحصول على أداء أفضل.
- 2- نوصي باستخدام تقنيات أخرى من الذكاء الصناعي مثل (المنطق الضبابي والخوارزمية الجينية و...) ومقارنة أدائها مع الشبكات العصبونية العميقة.
- 3- تدريب النموذج المستخدم على تصنيف صور لأمراض مختلفة.

### References:

- [1] S. Nayak, S. Kumar, and M. Jangid, "Malaria detection using multiple deep learning approaches," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), 2019; P:292-297. doi:10.1109/icct46177.2019.8969046
- [2] Kanojiya , Abhishek Yadav, Kirti Motwani " Malaria Detection using Image Processing and Machine Learning " IRJET.(2020);7(9):P:3433-3438.
- [3] Md. M. Hasan Sifat and M. M. Islam, "A fully automated system to detect malaria parasites and their stages from the blood smear," 2020 IEEE Region 10 Symposium (TENSYP), 2020;P:1351-1354. doi:10.1109/tensymp50017.2020.9230761

- [4] B. M. Rosa and G. Z. Yang, "Portable impedance analyzer as a rapid screening tool for malaria: An experimental study with culture and blood infected samples by early forms of plasmodium falciparum," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 12, pp. 3531–3541, 2020. doi:10.1109/tbme.2020.2990595
- [5] B. D. Ghanmode and P. L. Paikrao, "Malaria parasite concentration determination using digital image processing," 2018 4th International Conference for Convergence in Technology (I2CT), 2018;p:1-4. doi:10.1109/i2ct42659.2018.9058284.
- [6] L. Qi et al., "Photo acoustic tomography image restoration with measured spatially variant point spread functions," *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2318–2328, 2021; p:1-12.. doi:10.1109/tmi.2021.3077022
- [7] Ke Gu, Guangtao Zhai, Weisi Lin, Xiaokang Yang, and Wenjun Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3218–3231, 2015. doi:10.1109/tip.2015.2439035
- [8] D. Demirovic, E. Skejic, and A. Serifovic-Trbalic, "Performance of some image processing algorithms in tensorflow," 2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP), 2018; p:978-981. doi:10.1109/iwssip.2018.8439714
- [9] H. R. Roth et al., "Improving computer-aided detection using convolutional neural networks and random view aggregation," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1170–1181, 2016; p:114-125. doi:10.1109/tmi.2015.2482920
- [10] J. Nappi, W. Cai, and H. Yoshida, "Computer-aided detection of colorectal lesions for cathartic-free CT Colonography," 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009; p:911-914. doi:10.1109/isbi.2009.5193201
- [11] Aous M., Ghada S. Designing a Multi classification Convolutional Neural Networks Model for the Diagnosis of Lung Cancer and Covid-19. *Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series*. 2022; 44(6):185-20..
- [12] Dina M. I, Nada M. E., Amany M. S. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. *Elsevier*. 2021; 15(3):1-13.