

مقاربات جديدة لتشخيص سرطان الثدي باستخدام خوارزميات الذكاء الصناعي

دعاء مهنا *

(تاريخ الإيداع 2022/5/9 . قُبِلَ للنشر في 2022/7/7)

□ ملخص □

يعتبر تحديد المنطقة التي تحتوي على الكتل السرطانية للثدي في الصورة الشعاعية مشكلة صعبة، وتكمن الحاجة في الحصول على هذه المنطقة لتوجيه الأخصائيين لاختيار العلاج الدقيق تبعاً لنوع الورم. الهدف من البحث هو العمل على تقنية لاستخراج منطقة الكتلة المشتبه بها من الصورة الطبية بشكل آلي دون الحاجة الى الاقتطاع اليدوي، وبكفاءة لإجراء مزيد من التشخيص الطبي، والعمل على تدريب مجموعة من خوارزميات الذكاء الصناعي لتصنيف بارامترات تحليل الأنسجة والتنبؤ بنوع الشذوذ. يقدم البحث منهجية علمية مقترحة لاستخراج منطقة الشذوذ بهدف إعطاء تشخيص أولي لها، وتصنيف بارامترات الأنسجة للتنبؤ بالنوع الدقيق للشذوذ وقد اعتمدت هذه المنهجية على:

- ✓ العمليات المورفولوجية لاستخراج منطقة الشذوذ.
- ✓ نظام التصنيف المرحلي TNM ومقياس الأشكال للشذوذ لإعطاء تشخيص مبدئي حول نوع وحجم الشذوذ بالإضافة لمرحلة المرض.
- ✓ استخدام مجموعة من خوارزميات الذكاء الصناعي لتصنيف بارامترات تحليل الأنسجة لإعطاء التشخيص النهائي للمرض.

الكلمات المفتاحية: الصور الطبية، سرطان الثدي، العمليات المورفولوجية، تشخيص أورام الثدي، الذكاء الصناعي، الشبكات العصبونية.

* ماجستير في قسم تكنولوجيا المعلومات-كلية هندسة تكنولوجيا المعلومات والاتصالات-جامعة طرطوس-طرطوس-سوريا

New approaches to breast cancer diagnosis using artificial intelligence algorithms

D'uaa mhnaa*

(Received 9/5/ 2022 . Accepted 7/7/2022)

□ ABSTRACT □

Determining the area containing the breast cancers masses on radiographs is a difficult problem, the need to have this area to guide the specialists, to choose the exact treatment according to the type of tumor.

The aim of the research is to work on an idea to extract the suspected mass area from the medical image automatically without the need for manual truncation and efficiently for further medical examination, working on training a set of artificial intelligence algorithms to classify parameters of tissue analysis and predict the type of abnormality.

The research presents a proposed approach to extract the anomaly area in order to give a preliminary diagnoses of the type of anomaly and classification of tissue parameters to predict the exact type of abnormality, this method relied on:

- ✓ Morphological operations to extract the anomaly area.
- ✓ TNM staging system and morphometric scale for anomalies to give an initial diagnosis about the type and size of anomalies in addition to the disease stage.
- ✓ Using a set of artificial intelligence algorithms to classify tissue analysis parameters to give the final diagnosis of the discuses.

Keywords: medical images, breast cancer, morphological processes, Diagnosis of breast tumors, artificial intelligence, neural networks.

*Postgraduate ٭ Department of Information Technology, Faculty of Information and Communication Technology, University of Tartous, Syria

1-المقدمة

يشكل سرطان الثدي السبب الثاني للوفيات بين النساء في العالم، ويعد الكشف المبكر عن سرطان الثدي ذو أهمية كبيرة، ويعتمد الكشف المبكر على الصور الطبية التي تلعب دوراً رئيسياً في تشخيص الأورام، إذ يعتبر التصوير الشعاعي أول إجراء يتم اتخاذه لتحديد احتمال الإصابة بالمرض، ويتبع هذا الإجراء تحليل الأنسجة للتأكد من الإصابة، ولتحديد نوع الشذوذ.

سابقاً تم الاعتماد على الأنظمة التقليدية لتشخيص سرطان الثدي والتي تتطلب من الأخصائي تحليل صور الأشعة وفي حال عدم القدرة على التشخيص يُنتدب أخصائي آخر كما أن التشخيص من قبل الأخصائيين يتطلب الاطلاع على تاريخ المرضى بالإضافة الى أن تأهيل طبيب حتى يصبح أخصائي أشعة قادراً على قراءة الماموغرام بكفاءة يستغرق أكثر من عشر سنين.

يتم استخدام الذكاء الصناعي كدعم تكنولوجي لصناع القرار في الحياة الطبية ولايقصد بالأنظمة المعتمدة على الذكاء الصناعي أن تحل مكان الأطباء ولكنها تعد مفيدة لسببين:

- تجنب الأشياء البسيطة والمعتادة التي تستغرق وقتاً طويلاً وتثقل كاهل الموارد الطبية بالكثير من الأنشطة التي يمكن للجهاز أن ينفذها بسهولة.
- الإشارة الى الحالات الطبية التي يصعب على البشر اكتشافها والتي تحدث فرقاً بين الحياة أو الموت.

يعتبر الحصول على المنطقة غير طبيعية ضمن الثدي أمراً مهماً لكن ذلك يتطلب البحث عن الميزات وفق

التالي:

الحجم والموقع والحدود.

تحليل المستويات الرمادية.

إيجاد الأنماط للمنطقة الهامة.

بالإضافة الى الحاجة الى طرق تصنيف للتقييم، من أجل تسجيل قيم الاحتمال للإيجابية الحقيقية

(true positive) والإيجابية السلبية (false positive) ومن هذه الطرق :

خوارزمية الجوار الأقرب ،مصنف بايزي ،الشبكات العصبونية،متجهات الدعم SVM،تحليل المكون الرئيسي

(Principal component analysis)

بالاعتماد على ماسبق يمكن تصميم نظام تشخيص مقترح للحصول على منطقة الشذوذ بشكل مؤتمت، بدلاً من

العمل على الاقتران اليدوي للمناطق غير الطبيعية من قبل أخصائي الأشعة.[1]

تكمن الفائدة من الحصول على المنطقة غير الطبيعية في التنبؤ بنوع الشذوذ، و توجد العديد من الطرق

المستخدمة منها تصميم نظام عصبوني ضبابي ARM CONTEX-M3 لتشخيص سرطان الثدي من حيث الكتل الحميدة والخبيثة.

إن الصور الطبية (الماموجرام) تتضمن درجة BIRADS (درجة تعبر عن نتائج التصوير الشعاعي للثدي

والتي تعد نظام لاكتشاف حالة الثدي، بحيث تشير الى نوع المنطقة غير الطبيعية والتي يتم التعبير عنها على شكل رموز رقمية معيارية.

تشير هذه الدراسة إلى أن مفهوم النظام الخبير يعتبر أداة لاتخاذ القرار بناءً على قاعدة معرفة ومحرك استدلال (يفسر الحقائق)، في حين أن مفهوم المنطق الضبابي هو عبارة عن طريقة للتعبير عن الاحتماليات، عن طريق الشبكات العصبونية التي تقوم بتقريب النتائج .

تشير النتائج أن النظام المقترح يعمل وفق القواعد التي تؤخذ من محرك الاستدلال للنظام الخبير، ثم إيجاد المجموعات الضبابية من خلال توظيف تقنيات التقريب في الشبكات العصبونية.
ويتكون النظام من ثلاث طبقات:

✓ طبقة الدخل (شكل الكتلة، حجم الكتلة، كثافة الكتلة، التكلس، توزيع التكلس).

✓ طبقة القواعد الضبابية بحيث يتم تحديد صيغة القاعدة وفق النموذج التالي:

(إذا كان شكل الكتلة.....وحجم الكتلة.....وكثافة الكتلة.....والتكلس.....وتوزيع

التكلس.....فان BIRADS هو....).

✓ طبقة الخرج (0-غير مكتمل، 1-سلبى، 2-حميدة، 3-حميدة على الأرجح، 4-متمل

خبيث،

5-خبيث، 6-خزعة) [2].

في هذه البحث تم العمل على:

❖ استخلاص الشذوذ من الصورة الطبية Mammograms باستخدام العمليات المورفولوجية

(الفتح والإغلاق opening&closing).

❖ حساب حجم المنطقة الشاذة.

❖ التنبؤ بنوع الشذوذ.

❖ استخدام الشبكة العصبونية لإعطاء التشخيص النهائي.

❖

2-أهمية البحث وأهدافه

يعتمد التشخيص الأولي لسرطان الثدي على إمكانية تحديد المنطقة غير الطبيعية في الصورة الشعاعية والتي توفر المعلومات اللازمة، قدر الإمكان، لاستخدامها في التصنيف المرحلي للمرض والذي بدوره يوجه الأخصائيين لاتخاذ الخطوة التالية، وإعطاء تشخيص نهائي مع تحديد العلاج الدقيق الذي يتناسب مع مرحلة المرض. إن الهدف من البحث هو الحصول على منطقة الشذوذ من الصورة الطبية وإبراز حوافها بشكل مؤتمت وإبراز مناطق أخرى ضمن الثدي لم تكن واضحة يُشتبه بإصابتها، وتوفير تشخيص أولي من خلاله قد يطلب تحليل للأنسجة، والاعتماد على البارامترات الناتجة عنه لإعطاء التشخيص النهائي.

3-طرائق البحث ومواده

أنجز هذا البحث للمساهمة في التشخيص النهائي لسرطان الثدي وفق الخطوات التالية:

❖ استخراج منطقة الشذوذ من الصورة الطبية وإبراز الحواف.

❖ إيجاد مركز وحجم منطقة الشذوذ.

❖ الاعتماد على مقياس الأشكال لمنطقة الشذوذ ونظام التصنيف المرحلي TNM لتحديد التشخيص الأولي (نوع ومرحلة الشذوذ).

❖ الاعتماد على التشخيص الأولي في تحديد إجراء تحليل للأنسجة وتوفير التشخيص النهائي. تم الاعتماد على الصور الطبية المتوفرة في قاعدة البيانات mini-MIAS بحجم 1024×1024 pixel للصورة الواحدة حيث تضمنت معلومات عن طبيعة أنسجة الثدي، الفئة التي ينتمي لها الشذوذ (المنطقة غير الطبيعية)، نوع الشذوذ (خبيث، حميد).

كما تم الاعتماد على قاعدة البيانات المتوفرة من Breast Cancer Wisconsin (Diagnostic) Data Set للحصول على بارامترات تحليل الأنسجة وتحتوي على 699 حالة لتشخيص سرطان الثدي حيث تم اعتماد الرقم 1 تشخيص لنوع الشذوذ الحميد والرقم 2 تشخيص لنوع الشذوذ الخبيث وفق البارامترات التالية:

- ✓ كثافة الكتلة (Clump thickness)
- ✓ حجم الكتلة المنتظم (Uniformity of cell size)
- ✓ شكل الكتلة المنتظم (Uniformity of cell shape)
- ✓ الاندماج (Marginal Adhesion)
- ✓ النوى العارية (Bare nuclei)
- ✓ حجم الخلية الظهارية (Single epithelial cell size)
- ✓ Bland chromatin
- ✓ النوى الطبيعية (Normal nucleoli)
- ✓ الانقسام النووي (Mitoses)

تم تنفيذ خطوات البحث من خلال المحاكاة الحاسوبية بالاعتماد على برنامج MATLAB الذي يحوي على العديد من المكتبات الداعمة.

3-1 تجزئة الصورة (Image segmentation)

تعد أهم مرحلة من مراحل معالجة الصور الطبية، وتهدف الى تقسيم الصورة إلى مناطق مترابطة ومتجانسة وفقاً لمعيار محدد بهدف إبراز المناطق الشاذة وهناك ثلاثة معايير لتقييم طريقة تجزئة معينة وهي:

- ✓ الدقة: الحصول على صورة محسنة واضحة المعالم مقارنة مع الصورة الأصلية.
- ✓ الجودة: عدم حذف أي معلومات أو تفاصيل في الصورة نتيجة عملية تحسين الدقة.
- ✓ الكفاءة: الوقت المطلوب لإنجاز عملية التجزئة.

تتم أهمية هذه المرحلة باستخراج معلومات نوعية من الصورة تقوم بتوفير تحليل عالي المستوى للصورة وقد طبقت العديد من الخوارزميات لتجزئة الصورة.[5]

3-2 استخلاص الشذوذ

العمليات المورفولوجية: تستخدم عمليات المورفولوجيا الرياضية في مجال الصورة الرقمية، بوصفها وسيلة لاستخراج عناصرها، تكمن الفائدة الكبيرة لها في تقييم منطقة ما ووصف شكلها، مثل انتزاع حدود الكائنات في الصور، والتعرف على الهياكل العظمية في المجالات الطبية، والتعرف على الأجسام المحدبة وغيرها؛ لذا تستخدم لغة المورفولوجيا الصورية مجموعة من العمليات والنظريات لإعادة تشكيل الأجسام في الصورة [3].

تستخدم بشكل واسع العمليات المورفولوجية لمعالجة الصور الرقمية، وذلك بهدف الحصول على المنطقة الهامة (Region of interest) في الصورة بالاعتماد على:

(a) **العنصر البنويوي (structuring element)**: وهو عبارة عن مصفوفة عشوائية من الواحدات والأصفار يتم إسقاطها على صورة الدخل بحيث يتم إزالة العنصر البنويوي فوق الصورة، وعند كل بكسل من الصورة تتم مقارنة عناصره بمجموعة البكسلات الأساسية بهدف اختزال المساحات البيضاء بعملية التآكل Erosion أو زيادتها بعملية التوسع Dilation.

(b) **التوسع (Dilation)**: تقوم هذه العملية على زيادة مساحة المناطق البيضاء الموجودة في الصورة المصدر بإزالة النقاط البيضاء الداخلية بحيث يتم تحديد قيمة البكسل في الصورة الناتجة عن عملية التوسع بالقيمة العظمى لجميع البكسلات المجاورة [4].

الهدف من هذه العملية هي جعل الأشكال الموجودة في الصورة أكثر وضوحاً من خلال معالجة التشوهات الناتجة عن وجود مناطق غير مملوءة (ثقوب)

$$\text{Image}_{\text{new}} = \text{im} \oplus \text{e}_{\text{old}} \quad \text{structuring element} \quad (1)$$

(c) **التآكل (Erosion)**: تقوم هذه العملية على تقليص مساحة المناطق البيضاء الموجودة في الصورة بحيث يتم تحديد قيمة البكسل في الصورة الناتجة عن عملية التوسع بالقيمة الصغرى لجميع البكسلات المجاورة.

الهدف من هذه العملية المحافظة على الكائنات الأساسية في الصورة المصدر.

$$\text{Image}_{\text{new}} = \text{imeag}_{\text{old}} \ominus \text{structuring element} \quad (2)$$

حيث أن:

image_{new}: الصورة بعد تطبيق العملية المورفولوجية.

Image_{old}: الصورة قبل تطبيق العملية المورفولوجية.

\oplus رمز عملية التوسع

\ominus رمز عملية التآكل

ومن دمج هاتان العمليتين تولدت العمليات المشتقة الإغلاق (closing) والفتح (opening).

(d) **الفتح (opening)**: يعرف بأنه تآكل يليه تمدد يطبق عدة مرات باستخدام نفس العنصر البنويوي والغرض من هذه العملية هو الحفاظ على المناطق الأمامية (تحتوي عدد أكبر من البكسلات) في الصورة وفق المعادلة (3).

$$\text{Opening} = \text{IM} \ominus \text{SE} \oplus \text{SE} \quad (3)$$

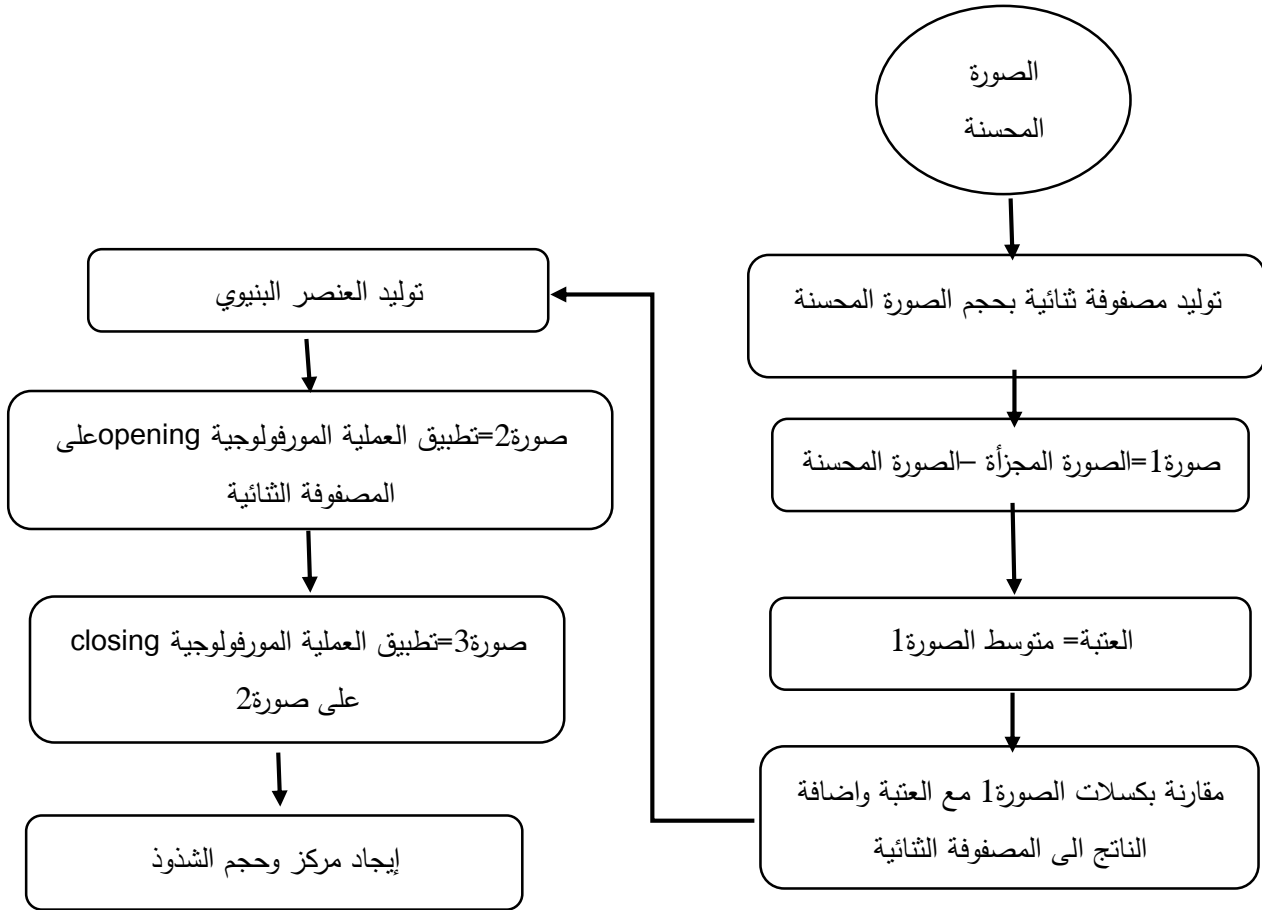
(e) **الإغلاق (closing)**: يعرف بأنه تمدد يليه تآكل يطبق عدة مرات باستخدام نفس العنصر البنويوي والغرض من هذه العملية هو الحفاظ على مناطق الخلفية (تحتوي عدد أقل من البكسلات) في الصورة وفق المعادلة (4).

$$\text{Closing} = \text{IM} \oplus \text{SE} \ominus \text{SE} \quad (4)$$

IM: صورة الدخل.

SE: العنصر البنويوي.

الخطوات المتبعة لاستخلاص الشذوذ من الصورة الطبية مبينة بالشكل (1):



الشكل (1) استخلاص منطقة الشذوذ

تم توليد صورة ثنائية بالاعتماد على العتبة الناتجة عن متوسط فرق الصورة المحسنة والمجزأة، ثم اختيار عنصر بنيوي مناسب لتطبيق العملية المورفولوجية opening ثم تطبيق العملية المورفولوجية closing للحصول على الصورة الخرج يليها إيجاد مساحة وحجم الشذوذ في الصورة.

3-3 نظام التصنيف المرحلي TNM:

تكمن الفائدة من التصنيف المرحلي في توجيه الأخصائيين لاختيار العلاج الدقيق الذي يحتاجه المريض يُستخدم نظام تصنيف الورم والعقد والانتشار (Tumour, Node, Metastasis)، لتحديد حجم الورم الأساسي، و انتشار الورم إلى العقد للمفاوية، أو انتشاره إلى أجزاء أخرى من الجسم.

• يدل رمز T على حجم الورم ومدى ارتشاحه في النسيج المجاورة، ويأخذ الأرقام 1 أو 2 أو 3

أو 4، حيث أن الرقم 1 للسرطانات الصغيرة والرقم 4 للسرطانات الأكبر حجماً.

- يدل رمز N على وجود نقائل الى العقد اللمفاوية وتأخذ الأرقام من 0 (لا تحتوي العقد على خلايا سرطانية) الى الرقم 3 (العديد من العقد اللمفاوية تحتوي على خلايا سرطانية).
- يدل رمز M على وجود نقائل بعيدة الى أجزاء أخرى من الجسم وتأخذ الرقم 0 (لا وجود للنقائل البعيدة) والرقم 1 (يوجد نقائل بعيدة).

□ يتم تعيين مرحلة الورم بالاعتماد على الحجم T

✓ 0: السرطان ضمن الثدي (1 سم) [6]

✓ 1A: الورم بحجم (2سم أو أصغر) T1N0M0

1B: الورم بحجم (2سم أو أصغر) T1N1M0

✓ 2A: الورم بحجم (2سم الى 5سم) T2N0M0

2B: الورم بحجم (2سم الى 5سم) T2N1M0

الورم بحجم (2سم الى 5سم) T2N2M0

الورم بحجم (أكثر من 5سم) T2N0M0

✓ 3A: الخلايا السرطانية تنتشر في 4~9 عقد لمفاوية

3B: الخلايا السرطانية انتشرت الى جدار الصدر

✓ 4: الخلايا السرطانية انتشرت في أجزاء الجسم كالكبد والرئة [6]

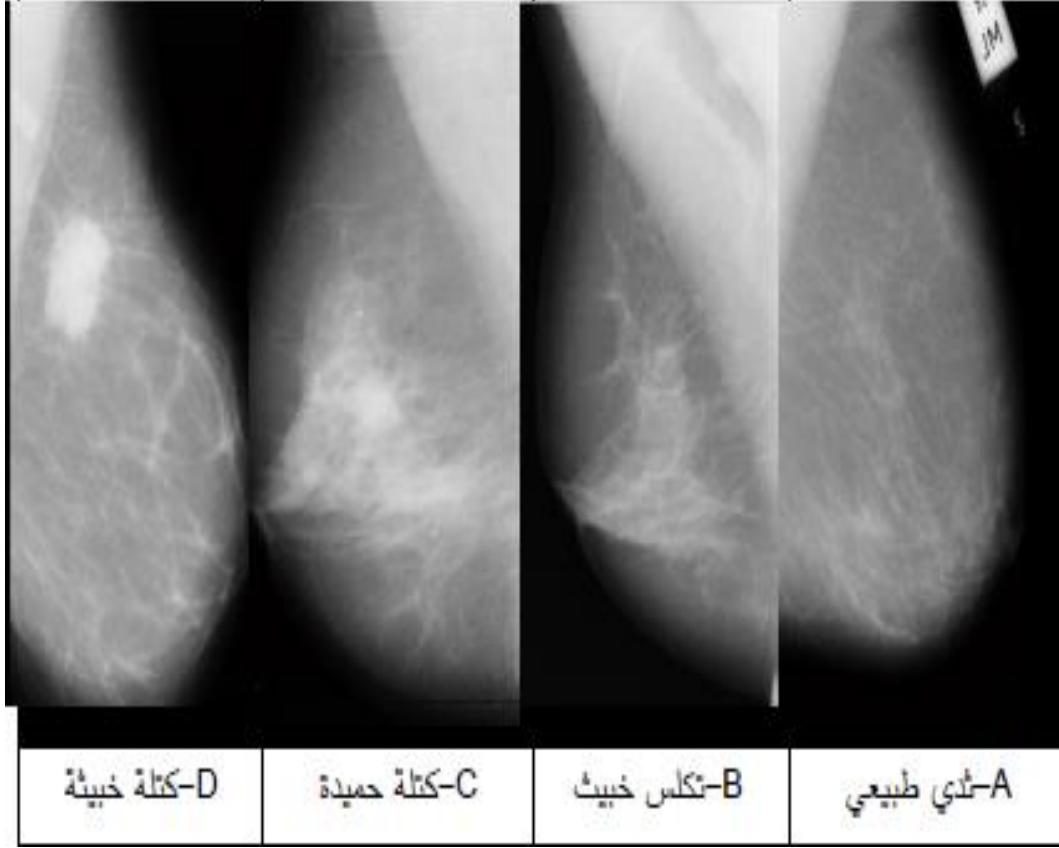
هذه الأنظمة لديها مقياس من 1 الى 4 وبشكل أقل شيوعاً من A الى D والتي تشير الى انتشار الورم الى العقد اللمفاوية وأعضاء الجسم الأخرى بحيث يستخدم الطبيب هذه الأحرف للحصول على مزيد من التقسيمات.

3-4 التنبؤ بالمرض:

وفقاً للمفهوم الطبي فإن الصورة الطبية تعطي تشخيص مبدئي للمرض، أي أنها تحدد احتمال اشتباه الإصابة بالمرض، إلا أن تأكيد الإصابة بالمرض يعتمد على البارامترات المستخرجة من تحليل الأنسجة. ضمن هذا المجال يتم الاعتماد على الذكاء الصناعي كدعم تقني لصناع القرار في المجال الطبي وذلك لان الانظمة التي تعتمد على الذكاء الصناعي تجمع بين ميزتين أساسيتين: دقة الرياضيات وقوة التقنيات الحالية. كما أن هذه الأنظمة تساعد على ملاحظة الأشياء غير المحسوسة (الأشياء التي من الصعب ملاحظتها أو التي تنتج عن عمليات حسابية معقدة أو التي تنتج بفعل مجموعة من العوامل) .

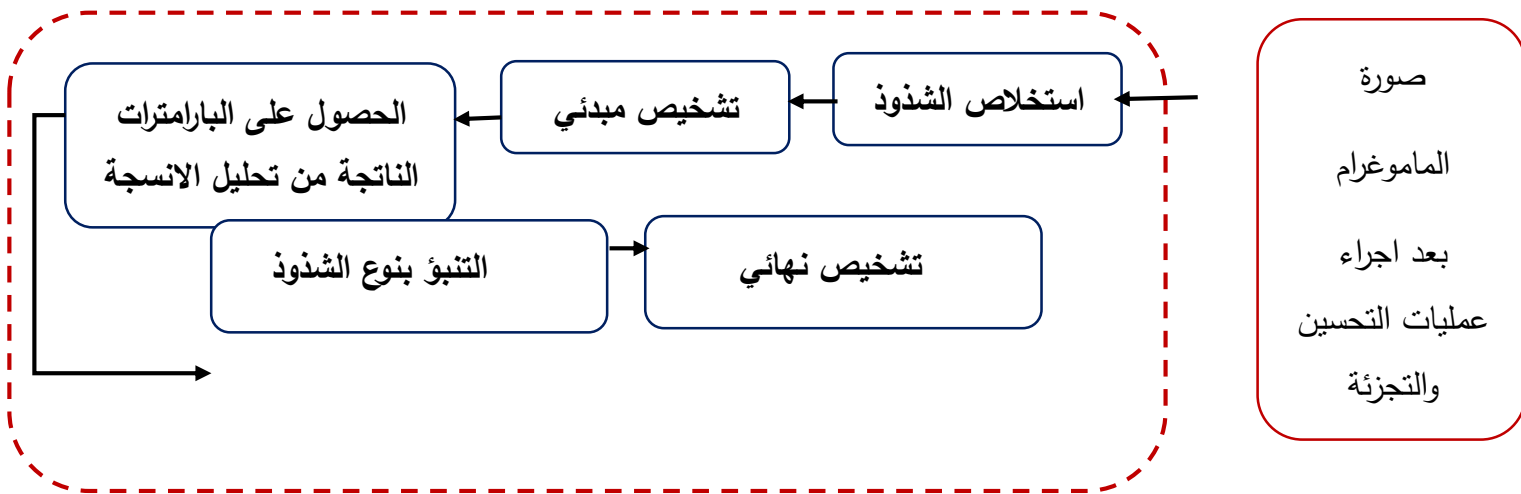
4- المناقشة

نُفذ البحث بتطبيق سلسلة من الخطوات على صور الماموغرام لإيجاد منطقة الشذوذ (في حال وجودها) واستخلاصها بشكل مؤتمت وإعطاء تشخيص دقيق بنوع الشذوذ على الحالات الموضحة بالشكل (2).



الشكل(2) الصور الطبية لحالات الثدي المدروسة قبل المعالجة

يظهر المخطط المبين بالشكل(3) الخطوات المتبعة للحصول على تشخيص نهائي للمرض.



يتم العمل على الصورة الطبية (الماموغرام) المحسنة والمجزئة لإجراء خطوات استخلاص الشذوذ للحصول على المنطقة غير طبيعية (في حال وجودها) بهدف اعطاء تشخيص مبدئي يتم الاعتماد عليه في تحديد ضرورة اجراء فحص الأنسجة ومن ثم إعطاء تشخيص نهائي. تكمن أهمية البحث من خلال القدرة على تبيان المنطقة غير الطبيعية بشكل مؤتمت وتحديد دقيق لنوع الشذوذ بزمن قصير بالإضافة الى الاشارة الى الحالات التي لايمكن ملاحظتها من قبل الطبيب.

4-1 استخلاص الشذوذ

يعتبر العمل على الصور الطبية للحصول على المنطقة المشتبه بإصابتها ذو أهمية كبيرة لتحديد التشخيص المبدئي لمرض السرطان (في حال وجوده) والذي يبين حجم الورم وإلى أي مدى قد نما. تكمن الأهمية من هذه المرحلة في الاشارة الى مناطق ضمن الثدي قد لايلحظ الطبيب إصابتها مما يمكن في بعض الحالات من المحافظة على الثدي.

تضمنت الخوارزمية المقترحة لاستخراج منطقة الاشتباه وفق الشكل(1):

1. توليد مصفوفة أرقام ثنائية بحجم 1024×1024 pixel
 2. اعتماد طريقة العتبة.
 3. الاعتماد على العمليات المورفولوجية الفتح والإغلاق (opening&closing)
 4. اختيار العنصر النيوي على شكل مصفوفة واحدة بحجم 8×8 pixel
 5. إيجاد مساحة ومركز الشذوذ مقدراً بـ cm
 6. حساب حجم الشذوذ من خلال:
 - a. عدد البكسلات لمنطقة الشذوذ= الجذر التربيعي لمساحة الشذوذ.
 - b. الحصول على قياس شاشة العرض مقدر بالانش (inch).
 - c. الحصول على حجم المنطقة بـ inch بتطبيق (عدد البكسلات/قياس شاشة العرض)
 - d. حجم المنطقة المدروسة بـ سم (cm) = حجم المنطقة بـ 2.54 .inch
- حيث أن inch=2.54cm

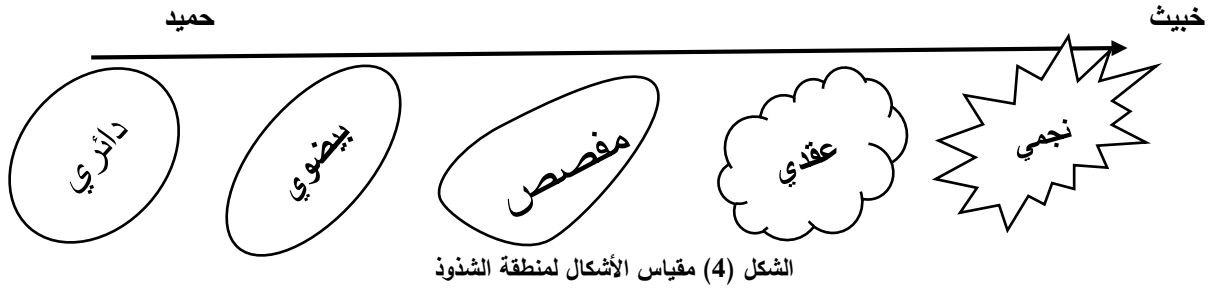
4-2 التشخيص المبدئي:

يوجه التشخيص المبدئي الأخصائيين لاختيار العلاج الدقيق الذي يحتاجه المريض ففي حال كان السرطان في مكان واحد (غير منتشر)، يكون العلاج موضعياً كالجراحة أو العلاج الإشعاعي والذي يعد كافيّاً للتخلص من السرطان بشكل كامل.

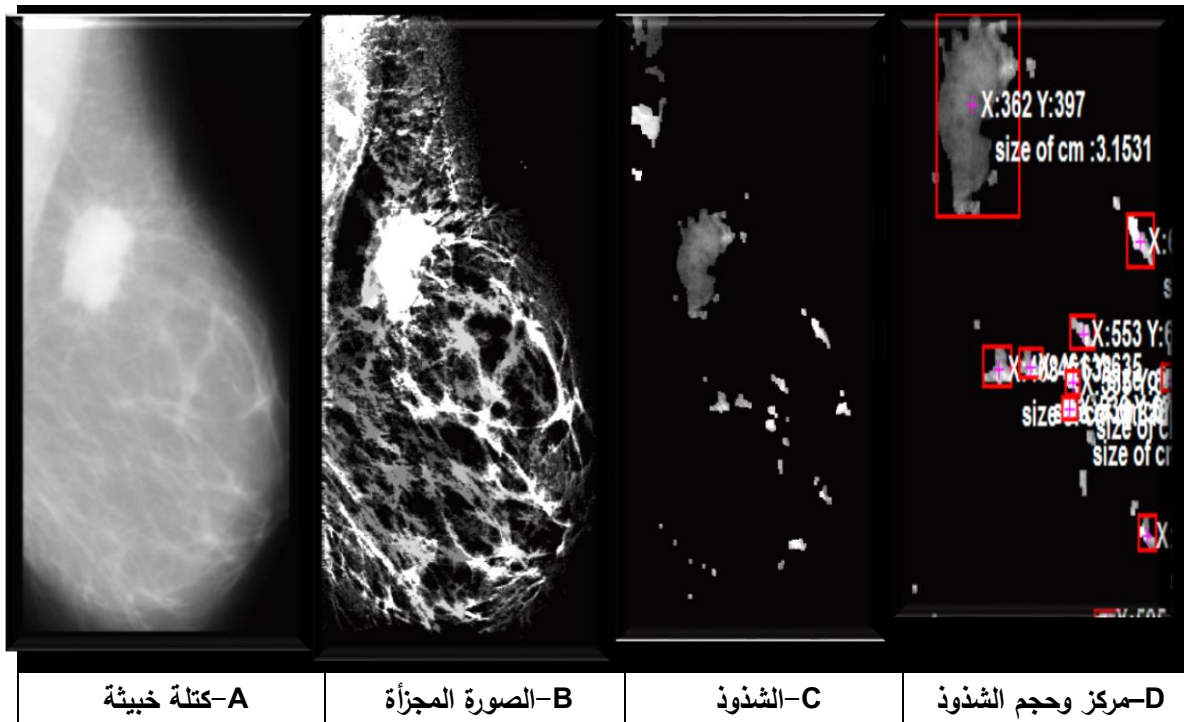
أما في حال انتشار السرطان، سنحتاج الى علاج جهازي (Systemic treatment) كالعلاج الكيميائي والهرموني الذي يدخل مجرى الدم وينقل الى جميع أجزاء الجسم.

هناك نوع آخر من العلاج هو العلاج المساعد (Adjuvant treatment) يستخدم عند عدم التأكد من انتشار السرطان من العقد اللمفاوية الى أجزاء الجسم الأخرى فيعطى للمريض الى جانب العلاج الأساسي (العلاج الكيميائي بعد الخضوع للجراحة) [6].

ضمن البحث تم الاعتماد على مقياس الأشكال لمنطقة الشذوذ جنباً إلى جنب مع خوارزمية استخلاص الشذوذ لإعطاء التشخيص المبدئي لطبيعة الشذوذ.



يظهر الشكل(5) نتائج تنفيذ خوارزمية استخلاص الشذوذ على إحدى حالات الدراسة:



الشكل(5) استخلاص منطقة الشذوذ من الصورة الطبية

يمكن إعطاء تشخيص مبدئي من خلال نتائج استخلاص الشذوذ للحالة السابقة وبالمطابقة مع مقياس الأشكال (4) وبالاعتماد على مفهوم نظام التصنيف المرحلي TNM أن نوع الشذوذ هو كتلة خبيثة من المرحلة الثالثة.

3-4 التنبؤ بنوع الشذوذ

ضمن البحث قمنا بتدريب مجموعة من خوارزميات الذكاء الصناعي لتصنيف بارامترات تحليل الأنسجة والتنبؤ بنوع الشذوذ.

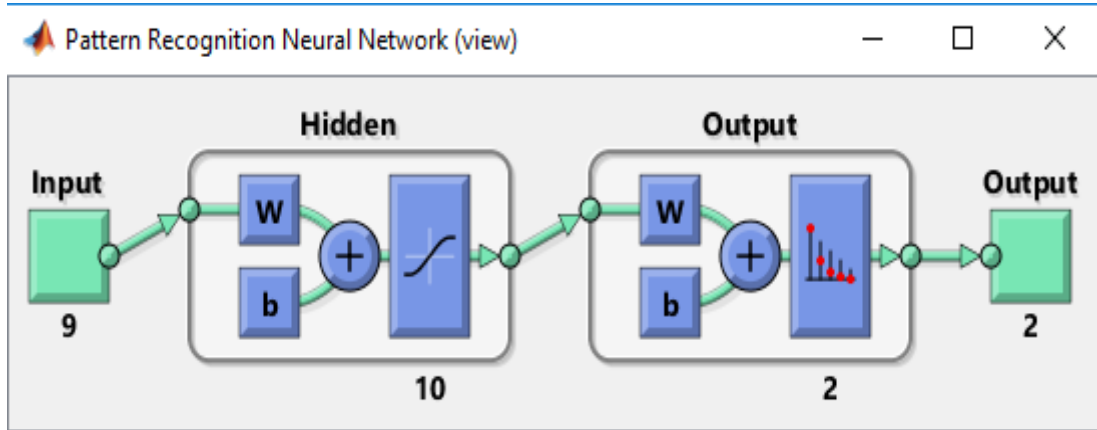
تحتوي قاعدة البيانات التي استخدمت لتدريب خوارزميات الذكاء الصناعي على 699 حالة لتشخيص سرطان الثدي (خبيث(2)، حميد(1)).

4-3-1 الشبكات العصبونية:

تم استخدام شبكة Perceptrons متعددة الطبقات بطبقة مخفية واحدة تتضمن 10 عصبونات (تجريبياً) وبالاعتماد على تابع التدريب (Levenbreg Marquardt) مع خرج للشبكة (2) وهو عبارة عن النوع (خبيث، حميد) المراد التصنيف وفقه.

تم تقسيم البيانات الى 3 مجموعات وفق التالي:

- ✓ مجموعة التدريب 70% بحيث تتضمن (489) عينة.
- ✓ مجموعة التحقق من صحة الاختبار 15% تتضمن (105) عينة.
- ✓ مجموعة الاختبار 15% تتضمن (105) عينة.



الشكل (6) الشبكة العصبونية المستخدمة للتصنيف

تم الاعتماد على الطريقة الاستمرارية في تقسيم البيانات بحيث تتضمن هذه الطريقة الافتراضات الأساسية حول مجموعة البيانات ، بما في ذلك اختيار مجموعة البيانات بشكل عشوائي وتوزيعها بشكل مستقل ومماثل بالإضافة إلى أن التوزيع مستقر ولا يتغير بمرور الوقت

يشير الشكل (6) الى الشبكة العصبونية المستخدمة في التصنيف حيث يمثل W (الوزن) و b (الانحياز) وتطلب تدريب الشبكة 8 دورات ضمن برنامج الماتلاب وأعطت نتائج تصنيف الشبكة العصبونية التالي:

• مصفوفة الارتباك

تعد مقياس الأداء لمشكلة التصنيف بحيث تعتبر:

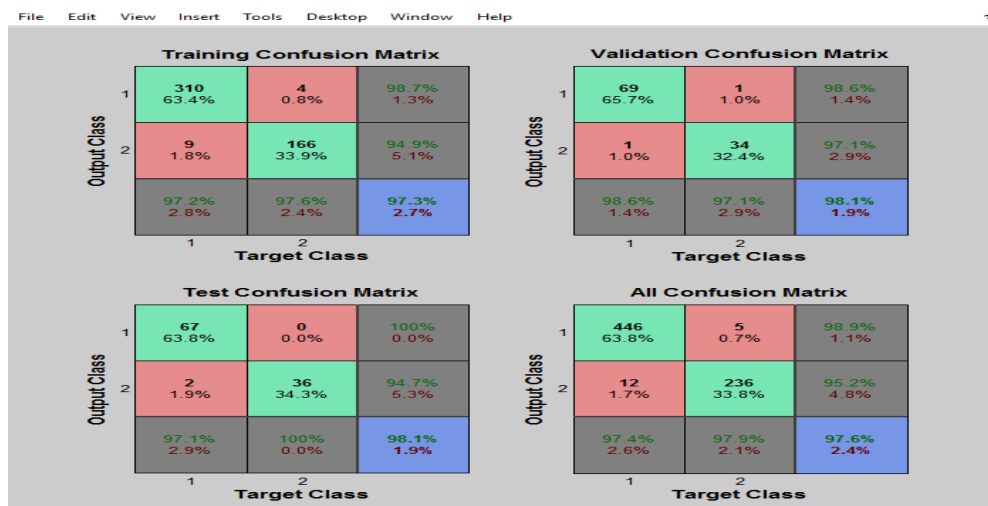
✓ التصنيفات الصحيحة هي عبارة عن الإيجابيات الحقيقية TP (التنبؤ بالمرض وهو موجود

فعلاً) أو السلبيات الحقيقية TN (التنبؤ بعدم وجود المرض)

✓ التصنيفات الخاطئة هي عبارة عن الإيجابيات الكاذبة FP (التنبؤ بوجود المرض وذلك غير

صحيح) أو السلبيات الخاطئة FN (التنبؤ بعدم وجود المرض وهو غير صحيح)

وبالاعتماد على هذه القيم يتم استخلاص الحساسية والنوعية والدقة.



الشكل (7) مصفوفة الارتباك

يمثل الشكل (7) مصفوفة الارتباك وفق التالي:

- مصفوفة الارتباك للتدريب أعطت تصنيف خاطئ لـ 4 عينات ضمن الفئة الأولى (خبيث) و9 عينات ضمن الفئة الثانية (حميد).
- مصفوفة الارتباك للتحقق أعطت تصنيف خاطئ لـ 1 من العينات ضمن الفئتين.
- مصفوفة الارتباك للاختبار أعطت تصنيف خاطئ لـ 2 من العينات ضمن الفئة الثانية (حميد).
- مصفوفة الارتباك الكلية أعطت تصنيف خاطئ لـ 5 عينات ضمن الفئة الأولى (خبيث) و12 عينات ضمن الفئة الثانية (حميد) بدقة 97.6%.

بالاستفادة من الشكل (7) تم الحصول على قيم التصنيفات من المصفوفة الكلية:

Target class

	TP الإيجابي الحقيقي (63.8)	FP الإيجابي الكاذب (0.7)
Output class	FN السلبي الكاذب (1.7)	TN السلبي الحقيقي (33.8)

الشكل (8) قيم مصفوفة الارتباك

الحساسية (Sensitivity): هي القدرة على كشف حالات السرطان عندما يكون فعلاً موجود أو بمعنى آخر هي نسبة المرضى اللواتي تم تشخيصهن بسرطان الثدي خلال سنة من إجرائهن الفحص بالماموغرام وكانت نتيجة الفحص لديهن بالماموغرام مشتبّه [7] وتحسب الحساسية وفق العلاقة:

$$\text{Sensitivity} = \frac{TP}{TP+FN} * 100 = \frac{63.8}{63.8+1.7} * 100 = 97.4$$

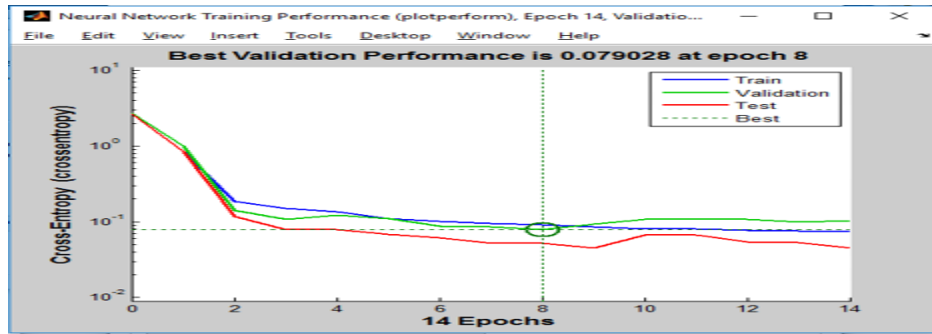
النوعية (Specificity): هي القدرة على تحديد المرضى الذين ليس لديهم فعلاً مرض السرطان أو نسبة المرضى اللواتي لم يشخص لهن سرطان ثدي خلال سنة من إجراء فحص الماموغرام وكانت نتيجة التحري عندهن سلبية أو طبيعية [7] وتحسب النوعية وفق العلاقة:

$$\text{Specificity} = \frac{TN}{TN+FP} * 100 = \frac{33.8}{33.8+0.7} * 100 = 97.9$$

تحسب الدقة [7] وفق العلاقة:

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} * 100 = \frac{63.8+33.8}{63.8+33.8+0.7+1.7} * 100 = 97.6$$

التدريب



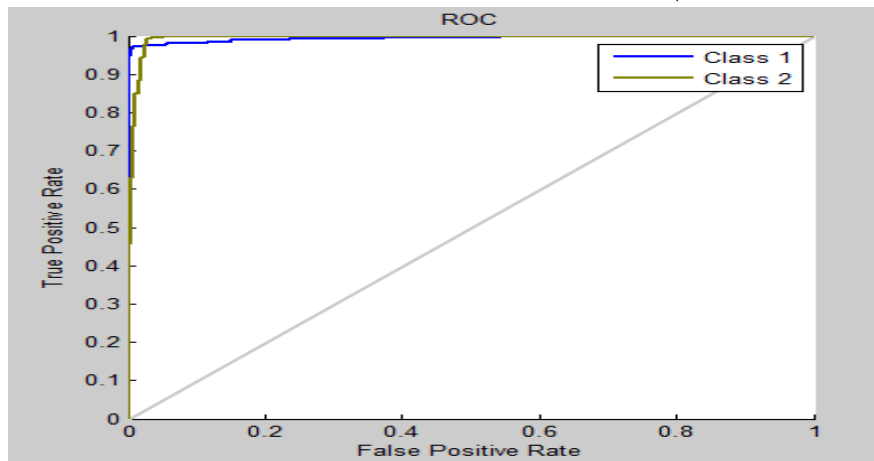
الشكل (9) مخطط الأداء باستخدام برنامج الماتلاب

يمكن الملاحظة من الشكل (9) أن الشبكة قد تعلمت عند التكرار 8 مع أفضل قيمة لأداء الشبكة هي 0.0737 وأفضل قيمة للتحقق هي 0.079028 هي قريبة جداً من 0 مما يدل على دقة تصنيف الشبكة.

Receiver operating characteristic curve ✓

يظهر مخطط Roc القدرة التشخيصية لمصنف نظام التشخيص بحيث يقوم بتطبيق قيم العتبة [0,1] على كل المخرجات.

يعطى مخطط Roc وفق معدل الإيجابي الحقيقي (احتمال التشخيص الصحيح) مقابل معدل الإيجابي الخاطئ (احتمال التشخيص الخاطئ).



الشكل (10) مخطط Roc

يمكن ملاحظة أن المنحني يصل الى الزاوية اليسرى العلوية من خط عدم التمييز (كلما زادت المساحة الموجودة تحت المنحني، كلما كان الاختبار أكثر دقة).

4-3-2 خوارزميات التصنيف الأخرى:

✓ تم استخدام نفس قاعدة البيانات للتصنيف بالاعتماد على خوارزمية الجوار الأقرب والتي تخزن جميع الحالات المتاحة لديها من بيانات التدريب وتقوم بتصنيف الحالات الجديدة حسب أغلبية جيرانها، وذلك عن طريق وظيفة حسابية لقياس المسافة بينهم وضمن البحث قمنا بتحديد ($k=5$) (تم اختياره تجريبياً) وهو عبارة عن عدد عناصر الجوار. ✓ خوارزمية متجهات الدعم:

هي خوارزمية تعلم الآلة خاضعة للإشراف والتي تستخدم في مشاكل التصنيف، ضمن هذه الخوارزمية عنصر البيانات هو عبارة عن نقطة في مساحة الأبعاد n حيث (n هو عدد البارامترات) وتكون قيمة كل ميزة هي قيمة الإحداثيات المعينة. ثم يتم إجراء التصنيف من خلال إيجاد حد القرار الذي يفصل بين الفئتين.

✓ شجرة القرار Decision Tree:

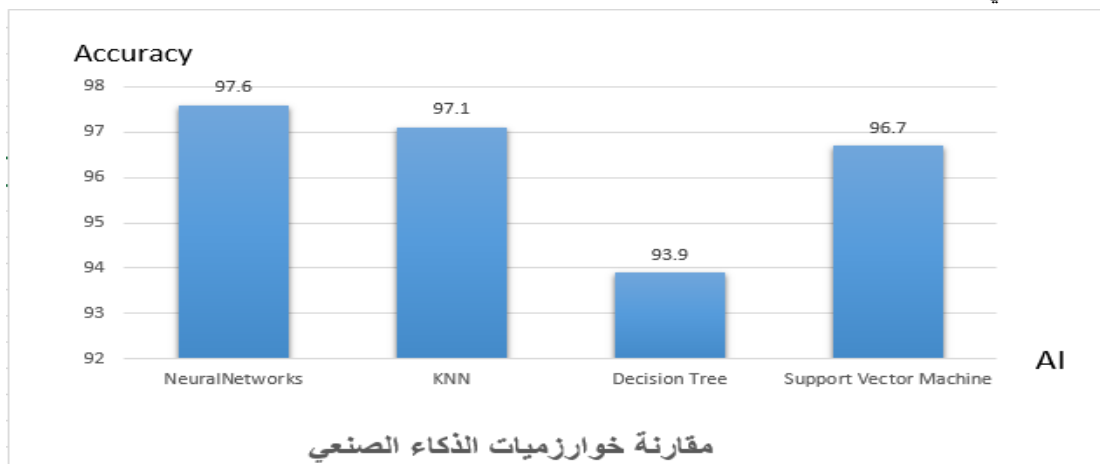
شجرة القرارات هي بنية تشبه المخطط الانسيابي، حيث تمثل كل عقدة داخلية اختباراً على سمة، ويمثل كل فرع نتيجة الاختبار، وتمثل كل عقدة ورقة الفئة (القرار المتخذ بعد حساب جميع السمات). تمثل المسارات من الجذر إلى الورقة قواعد التصنيف.

ويمكن تلخيص نتائج استخدام هذه الخوارزميات بالجدول (1):

جدول (1) ملخص نتائج تطبيق خوارزميات الذكاء الصناعي

AI	Sensitivity	Specificity	Accuracy
Neural Network	97.4	97.9	97.6
KNN ($k=5$)	97.5	96.2	97.1
svm	97.1	95.8	96.7
Decision Tree	94.9	92.1	93.9

أعطت الشبكة العصبونية أفضل دقة مقارنة بالخوارزميات التي تم ذكرها وبالاعتماد على نفس قاعدة البيانات المخطط التالي يظهر نتيجة المقارنة:



مقارنة خوارزميات الذكاء الصناعي

الشكل (11) مقارنة دقة خوارزميات الذكاء الصناعي

5- الاستنتاجات والتوصيات

- استخلاص الشذوذ أدى الى ظهور سمات لم تكن واضحة في الصورة الطبية مما أدى الى الحصول على معلومات أفضل للتشخيص من قبل الأطباء وبالتالي رفع سوية التنبؤ بوجود المرض
- الشبكات العصبونية أعطت أفضل دقة للتنبؤ بنوع الشذوذ.
- تتمتع الطريقة المقترحة بالسرعة والفعالية والتي تؤدي الى انخفاض تكلفة العلاج.
- استخدام الطريقة المقترحة ضمن المنظومة الطبية يوفر صورة كاملة المعلومات قدر الإمكان قابلة للتشخيص والتحليل الصحيح بزمن صغير بالإضافة الى تحديد نوع الشذوذ وبالتالي تشكيل رديف لرأي الطبيب.
- يمكن المتابعة في البحث من خلال:
- العمل على رفع دقة تنبؤ الشبكة العصبونية.
- العمل على تدريب شبكة عصبونية عميقة (CNN) على صورة الماموغرام الناتجة عن التجزئة لاستخلاص منطقة الشذوذ.
- العمل على الصور الطبية ثلاثية الأبعاد.

المراجع

- [1] Halalli, B. and Makandar, A., 2017. Computer Aided Diagnosis-Medical Image Analysis Techniques. Breast Imaging
- [2] Nagarajasri, B. and Padmavathamma, M., 2013. Threshold Neuro Fuzzy Expert System for Diagnosis of Breast Cancer. International Journal of Computer Applications, 66(8), pp.6-10
- [3] ساعي؛ ميا. اقتطاع منطقة الأذن من الصور الجانبية للوجه باستخدام كاشف الجلد المعتمد على الأرجحية والعمليات المورفولوجية، *Tishreen University Journal-Engineering Sciences Series* 2015, 35.1.
- [4] Said, K.A.M. and Jambek, A.B., 2016, August. A study on image processing using mathematical morphological. In 2016 3rd International Conference on Electronic Design (ICED) (pp. 507-512). IEEE.
- يعرب ديوب، جعفر سلمان، دعاء مهنا. حل لتحسين جودة الصورة الرقمية و ابراز معالمها لزيادة دقة تشخيص أمراض الثدي، مجلة جامعة طرطوس، المجلد(الثالث)، العدد(الثامن)، 2019.
- رياض، ربا، فادي، الوليد(26/6/2014). سلسلة ما هو السرطان-مراحل السرطان. تم الاطلاع عليه في [6] <https://www.syr-res.com/article/2360.html> تتم الاسترجاع من الرابط
- [7]] Li, H., Giger, M.L., Huo, Z., Olopade, O.I., Lan, L., Weber, B.L. and Bonta, I., 2004. Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: effect of ROI size and location. Medical Physics, 31(3), pp.549-555