

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبوغرارة  
كمال السنوسي

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم  
البياني للتدرج الموجه (HOG) مع مصنف آلة ناقل الدعم (SVM)

kamal Mohamed Elsanoussi Esmail aboghrara  
Faculty of Education – Traghen– University of Fezzan

### الملخص:

سرطان الرئة هو السبب الأكثر شيوعاً للوفاة بالسرطان بشكل عام في العالم، حيث يتسبب في وفاة 1.76 مليون شخص كل عام، وهو النوع الأكثر شيوعاً من السرطان بين الرجال والثالث الأكثر شيوعاً بين النساء وفي حين أنه غالباً ما يرتبط مباشرة بالتدخين، إلا أن 10-20% فقط من الأفراد الذين يصابون بالمرض لم يدخنوا من قبل وفي المراحل المبكرة من سرطان الرئة، قد تكون هناك علامات قليلة إن وجدت وهذا يدل علي وجود خطأ ما وهذا غالباً ما يؤدي إلى التشخيص المتأخر وقد أدى انتشاره في السنوات الأخيرة إلى زيادة معدلات الوفيات وهناك احتمال لزيادة عدد الإصابات في السنوات القادمة ونظراً للمشاكل التي يعاني منها التشخيص البصري لهذه الأمراض، بما في ذلك الحصول على نتائج غير صحيحة وغير دقيقة واستغراق الكثير من الوقت والجهد، فقد تم مؤخراً استخدام التقنيات الحديثة وتحديداً تقنيات التعلم الآلي Machine learning في تحليل وتصنيف هذه الأمراض من أجل مساعدة الأطباء في الكشف المبكر والدقيق عن الأمراض وكذلك التغلب على مشكلة نقص الأطباء الخبراء في المناطق الريفية حيث يتم استخدام صور الأشعة السينية للصدر للكشف عن هذا المرض وقد اتبع الباحث في هذه الدراسة المنهج التجريبي وتشكلت عينة الدراسة من 3886 صورة التي أخذت من مستودع GitHub مفتوح المصدر و تقترح هذه الدراسة طريقة الرؤية الآلية من أجل الكشف عن سرطان الرئة في صور الأشعة السينية للصدر و تم بناء نموذج التصنيف باستخدام السمات المأخوذة من صور الأشعة السينية باستخدام الرسم البياني للتدرج الموجه (HOG) وقد أظهر مصنف آلة الدعم

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

المتجه (SVM) المستخدم في هذا البحث أداءً رائعاً من حيث التصنيف وقد وفرت تقنية  
ميزة HOG المقترحة مدى دقة عالية تصل إلي (96.6%).  
الكلمات المفتاحية: سرطان الرئة، صور الأشعة السينية للصدر، صور الأشعة، التعلم  
الآلي.

#### **Abstract:**

Lung cancer is the most common cause of cancer death overall in the world, killing 1.76 million people each year. It is the most common type of cancer in men and the third most common in women. While it is often directly linked to smoking, 10–20% Only % of individuals who develop the disease have never smoked before. In the early stages of lung cancer, there may be few, if any, signs that indicate that something is wrong. This often leads to late diagnosis, and its spread in recent years has led to an increase in death rates. There are There is a possibility that the number of infections will increase in the coming years. Given the problems that visual diagnosis of these diseases suffers from, including obtaining incorrect and inaccurate results, and taking a lot of time and effort, modern technologies have recently been used, specifically machine learning techniques, in analyzing and classifying these diseases. Diseases in order to help doctors in early and accurate detection of diseases, as well as overcome the problem of the shortage of expert doctors in rural areas, where chest GitHub is open source. This study proposes a machine vision method in order to detect lung cancer in chest X-ray images. A classification model was built using the features taken

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

from The one used in this research has great performance in terms of classification, and the proposed HOG feature technology has provided a high accuracy range of up to (96.6%).

**Keywords:** lung cancer, chest X-ray images, radiology images, machine learning.

**المقدمة:** سرطان الرئة هو نوع من الأمراض الخبيثة التي تبدأ عندما تنمو الخلايا غير الطبيعية بشكل لا يمكن السيطرة عليه داخل الرئتين و سرطان الرئة هو مشكلة صحية خطيرة يمكن أن تسبب ضرراً شديداً أو حتى الموت و يمكن أن يكون لها تأثير كبير على القدرة على التنفس والصحة العامة ويمكن أن يشمل علاج سرطان الرئة الجراحة قليلة التوغل والعلاج الموجه والعلاج المناعي وتشمل الأعراض الشائعة صعوبة التنفس والسعال المزمن وألم الصدر وفقدان الوزن غير المبرر والضعف العام و يتم تصنيف سرطان الرئة وفقاً لنظام TNM والذي يحدد مدى انتشار الورم بأنه من الأمراض القاتلة [1]. وينقسم سرطان الرئة إلى أنواع مختلفة بناءً على الخلايا التي يبدأ فيها الورم حيث يعتبر سرطان الرئة من أكثر الأمراض انتشاراً في العالم وينشأ السرطان عندما تبدأ خلايا الجسم الغير طبيعية بالانفصال والتلامس مع الخلايا الطبيعية وجعلها خبيثة ويصيب السرطان الرئتين ويسبب تلف الأنسجة لدى المصابين قد لا يعاني بعض الأشخاص من أي أعراض في المراحل المبكرة، في حين أن غالبية الأشخاص الذين يعانون من الأعراض الأولية هي الحمى والسعال وآلام العضلات والصداع كلها أعراض ثانوية محتملة. ومع ذلك، تشير الأدلة حتى الآن إلى أن مجموعتين من الأشخاص أكثر عرضة للإصابة بالمرض هم: كبار السن (الأشخاص الذين تزيد أعمارهم عن 70 عاماً) والأشخاص المدخنون [2]. وينتشر المرض حالياً بسبب نقص طرق الكشف السريع عنه. ومن أجل تقليل عدد الوفيات الناجمة عن هذا الفيروس، يجب أن يكون العاملون في مجال الصحة والأطباء في جميع أنحاء العالم قادرين على اكتشافه بسرعة وبدقة جيدة ويعد التصوير الطبي واختبارات الدم والاختبارات الفيروسية من أكثر الإجراءات الطبية التشخيصية استخداماً [3] وفي التصوير

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

الطبي توجد فيه العديد من الاختبارات التشخيصية لسرطان الرئة التي يتم إجراؤها للتأكد ما إذا كان الشخص مصاب بورم حميد أو خبيث، مثل التصوير بالأشعة السينية والتصوير بالرنين المغناطيسي وتحليل صور الأنسجة وغيرها، وتعرف صور الأنسجة Histology Images بأنها الصور التي يتم التقاطها بالمجهر لعينة نسيج الورم التي يتم أخذها من جسم المريض، يؤدي التشخيص البصري لهذه الأورام السرطانية أحيانا إلى الحصول على نتائج خاطئة أو غير دقيقة وذلك نتيجة الأخطاء البشرية الناجمة عن القصور والمشاكل البصرية. تعتبر مهمة اكتشاف وتحديد الإصابة بالسرطان مهمة صعبة وتستغرق الكثير من الوقت ويؤديها اختصاصي الأشعة ذوي الخبرة والأطباء المختصين، حيث تعتمد دقة التشخيص على خبرة الطبيب فقط لذلك تم استخدام الأشعة السينية للصدر (CXR) مع التعلم الآلي كتقنية تشخيصية من الخط الأول [2]. تعتبر أشعة الصدر CXR واعدة للحالات الطارئة والعلاج نظرا لسرعة عملها وقلة تكلفتها وسهولة استخدامها بالنسبة لأخصائي الأشعة، كما يتعرض المريض لجرعات منخفضة من الإشعاعات المؤينة المتوفرة في العديد من البلدان. تُستخدم الأشعة السينية للصدر للتشخيص في معظم الإعدادات السريرية؛ فهي تستغرق وقتاً أقل لإعداد المرضى وتوفر نتائج سريعة ونتيجة لذلك، يمكن استخدام الأشعة السينية للصدر للتعرف على المرض وتخصيص الموارد الطبية والأشعة السينية رغم مميزاتها الكبيرة إلا أنها تحتاج إلى طبيب أشعة. يلعب أخصائيو الأشعة دوراً حيوياً نظراً لخبرتهم الواسعة في هذا المجال ومع ذلك، ونظراً للعدد المحدود من أخصائي الأشعة، فمن الصعب توفير أطباء خبراء لكل مستشفى. لذلك، قد تساعد نماذج الذكاء الاصطناعي الدقيقة والسريعة في التغلب على هذه المشكلة وتزويد المرضى بالمساعدة في الوقت المناسب. يمكن أن تساعد تقنيات الذكاء الاصطناعي في الأشعة في الحصول على تشخيص دقيق [5]. علاوة على ذلك، يمكن أن تساعد أساليب الذكاء الاصطناعي في إزالة العيوب مثل تكاليف الاختبار ووقت انتظار نتائج الاختبار لذلك أصبح استخدام تكنولوجيا المعلومات ضروري للتغلب على هذه القيود والتي تساعد الأطباء في اكتشاف وتشخيص هذا السرطان [5،6]. يمكن أن تساعد تقنيات الذكاء الاصطناعي المتقدمة (AI) جنباً إلى جنب مع التصوير الإشعاعي في

الكشف الدقيق عن هذا المرض. وكذلك التغلب على مشكلة نقص الأطباء الخبراء في المناطق الريفية وفي مجال التصوير الطبي، تم استخدام خوارزميات التعلم الآلي (ML) لتحسين أداء معالجة الصور بشكل كبير على العديد من التطبيقات منها [4]، [5] صور المجهر [6]، تصنيف أورام المخ [7]، صور الرنين المغناطيسي [8]، وصور الشبكية [9] وغيرها من التطبيقات.

**مشكلة الدراسة:** تعتبر مهمة اكتشاف وتحديد الإصابة بالسرطان مهمة صعبة وتستغرق الكثير من الوقت الجهد وتعتمد دقة التشخيص على خبرة الطبيب فقط المشكلة الرئيسية في الكشف عن سرطان الرئة تتلخص في الآتي:

1. قلة استخدام التقنيات الحديثة في تشخيص سرطان الرئة محليا.
2. تأثر نماذج الشبكات العصبية الالتفافية بمجموعة من العوامل مثل حجم البيانات المستخدمة وعدم توازن خصائصها وغيرها من العوامل.

**أسئلة الدراسة:** بالنظر إلى مشكلة البحث تم وضع مجموعة التساؤلات التالية لحلها:

1. هل استخدام نموذج الرسم البياني للتدرج الموجه HOG سيؤدي إلى حصول على نتائج بنسبة دقة عالية؟
  2. هل مصنف ناقل الدعم SVM جيد في عملية التصنيف مقارنة مع نتائج الخبير.
- أهداف الدراسة:** من خلال الأسئلة السابقة يهدف البحث إلى تحقيق الآتي:
1. تصنيف صور الأنسجة لسرطان الرئة باستخدام نموذج الرسم البياني للتدرج الموجه.
  2. تحسين النموذج المستخدم من خلال دراسة العوامل التي تؤثر على نسبة دقته.
- أهمية الدراسة:** يحتاج الأطباء إلى طرق جديدة ومتطورة لمساعدتهم في الكشف المبكر عن الأمراض وتقليل احتمالية انتشار المرض وتدهور صحة المريض، إن استخدام نموذج الشبكات العصبية المقترح سيؤدي إلى الحصول على نتائج أكثر دقة في الكشف عن سرطان الرئة ومساعدة الأطباء في التشخيص المبكر وبالتالي تجنب أخطاء التشخيص الناجمة عن القصور والأخطاء البشرية.

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

سيتم تقديم النموذج المقترح وما تم التوصل إليه من نتائج إلى مركز الأورام بسببها لمساعدة الأطباء في الحصول على نتائج تشخيص دقيقة وبنسبة موثوقية عالية، ومن خلال العوامل التي تم دراستها والخطوات التحسينية التي تم تطبيقها في هذا البحث يمكن للباحثين في هذا المجال الاعتماد عليها في تجاربهم المستقبلية.

**حدود الدراسة:** سيتم التركيز في هذا البحث على استخدام أحد نماذج التعلم الآلي في تصنيف صور الأنسجة لسرطان الرئة، تم بالتحديد استخدام نموذج الرسم البياني للتدرج الموجه في التصنيف المتعدد الصور أنسجة سرطان الرئة المأخوذة من مجموعة بيانات GitHub.

**مصطلحات الدراسة:**

HOG: Histogram of Oriented Gradients

support vector machine SVM:

:Neural Networks NN

Artificial Intelligence AI:

chest x-rays CXR:

ML: Machine Learning

**منهجية البحث:** تعد منهجية البحث أحد أهم أجزاء البحث العلمي وهي عبارة عن مجموعة من الخطوات التي يقوم بها الباحث في دراسته و بمعنى آخره الخطوات التي يجب أن يسير عليها الباحث أثناء قيامه بالبحث العلمي والتي من خلالها سيجد الحل المناسب للمشكلة، يتبع هذا البحث الخطوات الأساسية المتبعة في الدراسات ذات الصلة والتي قامت باستخدام النماذج المختلفة للشبكات العصبية الاصطناعية في تصنيفا لأمر اضخم ،تتبع جميع هذه الدراسات نفس الخطوات الأساسية المتمثلة في تحميل مجموعة البيانات المعالجة المسبقة تقسيم مجموعة البيانات و تدريب النموذج، تقييم النموذج مع وجود بعض الاختلاف انتظر الاختلاف مشكلة و الشكل التالي يوضح الخطوات الأساسية للمنهجية المتبعة في الدراسات ذات الصلة:

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبوغرارة  
كمال السنوسي



الشكل (1.1) يوضح المنهجية المتبعة في الدراسات ذات الصلة

في هذا البحث سيتم إتباع جميع الخطوات الأساسية للمنهجية المتبعة في الدراسات السابقة، مع إضافة بعض الخطوات للمنهجية للتوافق مع الخطوات الفعلية التي تم اتباعها، وتتمثل الخطوات التي تمت إضافتها في خطوة تحسين النموذج.

#### الدراسات السابقة:

لقد أجريت الدراسة [3] Razzak et al للكشف عن COVID-19 باستخدام تقنية CNN وحققوا دقة 98.89% للكشف الثنائي و94.39% لاختبارات التصنيف المتعددة، على التوالي.

في دراسة لكوماروكوماري [4] حيث وجمعوا السمات COVID-19 وزودوا مصنف SVM بدقة نهائية بلغت 92.7 و92.9%، على التوالي، باستخدام VGG16 وVGG19. في دراسة [4] تم استخدام نموذج Shuffle Net لاستخراج السمات تلقائياً من صور الأشعة السينية، والتي تم إدخالها بعد ذلك في أربع مصنفات مميزة: SVM وSoftMax وRandom Forest وKNN. استخدم Sethy وBehera صورة الأشعة السينية لتصنيف السمات التي تم الحصول عليها من نماذج مختلفة للشبكة العصبية التلافيفية (CNN) مع مصنف آلة الدعم المتجه (SVM)، وحققوا دقة نهائية تبلغ 95.33 وفي هذه الدراسة تم استخدامنا مجموعة بيانات مرجعية لسرطان الرئة والالتهاب الرئوي الفيروسي والحالات الطبيعية لتدريب نموذج يعتمد على الرسم البياني للتدرجات الموجهة (HOG).

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

**أوجه الاستفادة** من الدراسات السابقة: حيث تم تطوير نموذج آلة دعم التصنيف (SVM) لصور الأشعة السينية للصدر من أجل الحصول على أفضل أداء في تصنيف الالتهاب الرئوي الناتج عن السرطان الرئة والالتهاب الرئوي الفيروسي والمرضى العاديين. باستخدام المنهجية المقترحة، يمكننا تقييم المرضى المشتبه في إصابتهم بالالتهاب الرئوي الناتج عن المرض بسرعة وتوفير الرعاية الطبية الفورية. يمكن أن يساعد الأطباء ليس فقط في فحص السرطان بشكل أفضل، ولكن أيضًا في تحديد استراتيجية العلاج وتطوير خطة علاجية بناءً على مصدر العدوى.

**المواد والأساليب الدراسية:** لقد استخدمنا صور الأشعة السينية في هذه الدراسة للكشف عن الأفراد المصابين بمرض سرطان الرئة باستخدام ميزة HOG بالإضافة إلى تقنية التصنيف القائمة على آلة الدعم المتجه (SVM) تم استخدام SVM بدلاً من المصنف القائم على التعلم العميق للتصنيف. لأن الأخير يتطلب مجموعة بيانات كبيرة للتدريب والتحقق. يتم استخراج الميزات أو الخصائص من نموذج HOG وإعطائها لآلة الدعم المتجه (SVM) للتصنيف. تميز آلة الدعم المتجه (SVM) لتصنيف صور الأشعة السينية المتأثرة بسرطان الرئة من صور الأشعة السينية غير المتأثرة بسرطان الرئة، هذا النهج مفيد للمهنيين الطبيين للتمييز بين مرضى الأورام ومرضى الالتهاب الرئوي والأشخاص الأصحاء. باستخدام ميزات HOG، تم تقييم آلة الدعم المتجه (SVM) لتحديد المرض حيث قدمت آلة الدعم المتجه (SVM) أفضل النتائج للكشف عن المرض وتمت مناقشة مجموعة البيانات والمنهجية المقترحة (استخراج الميزات والتصنيف) بالتفصيل.

**1-عملية وصف مجموعة البيانات:** تم دمج مجموعة بيانات مختلفة لسرطان الرئة، والتي تم الحصول عليها من مستودع GitHub مفتوح المصدر <https://github.com/ieee8023/covid-chestxray-dataset/tree/master> [ /images ]. يبلغ إجمالي الصور في مستودع مجموعة البيانات 3886. مصنفة حسب الأشعة السينية للصدر منها 1200 صورة لأشخاص مصابين بالسرطان و1341 صورة لمرضى عاديين و1345 صورة للالتهاب الرئوي الفيروسي كما هو موضح من الجدول 1.

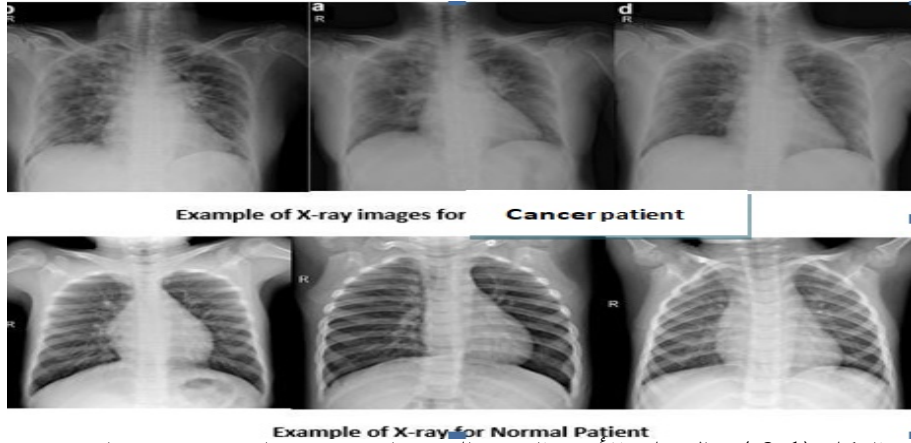


استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

كانت صور الأشعة السينية مختلفة في مستويات الرمادي وحجم الرقعة والأبعاد والخصائص التي تم الحصول عليها ويوضح الشكل 2.1 صورة أشعة سينية نموذجية لمريض مصاب بسرطان وشخص عادي تم استخدامها للبحث.

Table 1. The proposed COVID chest x-ray dataset.

Dataset	Viral pneumonia	Normal	Cancer	Total images
Total	1345	1341	1200	3886



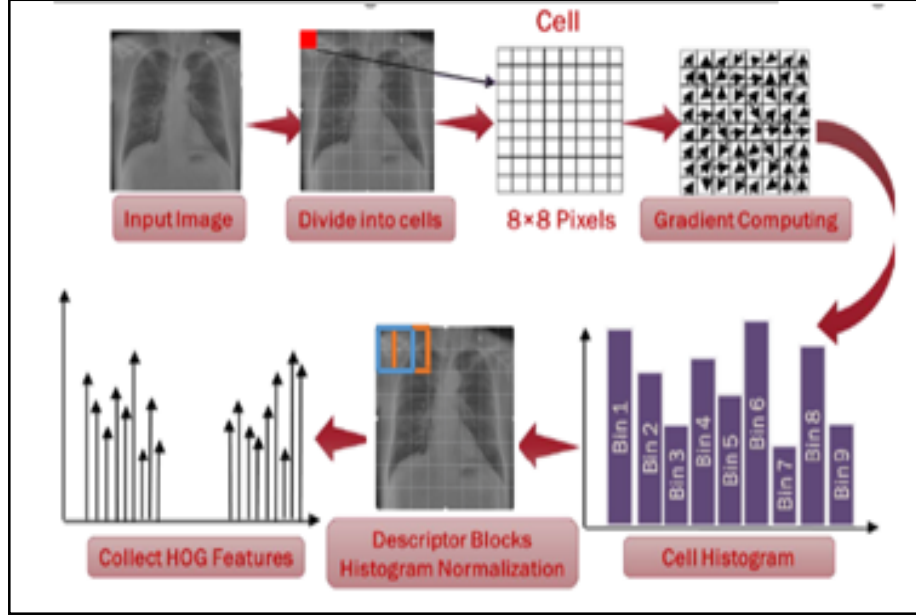
الشكل (2.1) مثال على الأشعة السينية للصدر لمريض مصابين ومرضى عاديين.

## 2- استخراج الميزات باستخدام HOG:

الرسم البياني للتدرج الموجه (HOG) هو تقنية وصفية لاستخراج الميزات من صور مرضى السرطان المصابين والصور العادية للأشعة السينية مع منطقة موجبة وسالبة. يتم تضمين مراحل مختلفة في الرسم البياني للتدرج الموجه (HOG). تُستخدم هذه الميزات لبناء نماذج التعلم الآلي.

في هذه الدراسة، سنستخدم صورة الأشعة السينية المدخلة بدرجات الرمادي المحددة. يجب أن تكون الصورة المدخلة قيمة حقيقية غير مبعثرة. يصف الشكل 2.2 طريقة الرسم البياني للتدرج الموجه (HOG) لاستخراج الميزات من الصورة [21].

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
 إسماعيل أبوغرارة  
 كمال السنوسي



الشكل (1.2) التدفق الأساسي لخوارزمية استخراج الميزات HOG (التدرج الموجه بالمدرج الموجه)

### 3 - تنفيذ خوارزمية HOG:

**الخطوة 1:** يتم توزيع صورة الإدخال إلى صور فرعية أصغر ذات صلة تسمى الخلايا، كل خلية بحجم 8×8.

**الخطوة 2:** حساب التدرج، وهي الخطوة الثانية في استخراج السمات HOG. قناع المشتق المنفصل ذو النقطة المركزية أحادية البعد في كل من الاتجاهين الأفقي (Dx) والرأسي (Dy) هو أفضل طريقة. باستخدام عملية الالتفاف، يتم الحصول على المشتقات الأفقية والرأسية (I<sub>x</sub>) و (I<sub>y</sub>)، انظر المعادلتين (2.1) و (2.2).

$$D_x = [-1,0,1]I_x = I * D_x \quad (2.1)$$

$$D_y = [-1,0,1]$$

$$I_y = I * D_y \quad (2.2)$$

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

$$I_x = I(x + 1, y) - I(x - 1, y) \quad \text{المعادلة (2.1) تعني أن}$$

$$I_y = I(x, y + 1) - I(x, y - 1). \quad \text{بينما المعادلة (2.2) تعني أن}$$

بالنسبة لكل بكسل  $(x, y)$ ، يتم تحديد مقدار التدرج كما هو موضح في المعادلة (2.3).

$$|G(x, y)| = \sqrt{I_x^2 + I_y^2} \quad (2.3)$$

يتم تحديد اتجاه واتجاه التدرج بواسطة المعادلة (2.4).

$$\theta(x, y) = \tan^{-1} \frac{I_y}{I_x} \quad (2.4)$$

**الخطوة 3:** تحديد كمية الرسم البياني/التجميع حسب الاتجاه

المرحلة الثالثة في استخراج ميزات HOG هي إنشاء رسم بياني خلوي ويتم تصنيف المنطقة ذات الاهتمام (ROI) إلى منطقة صغيرة تسمى الخلايا ويبلغ حجم الخلية  $8 \times 8$  بكسل ويساهم كل بكسل في الخلية في قناة الرسم البياني بناءً على الاتجاه بقيمة مرجحة. يمكن تمثيل مساهمة البيكسل بحجم التدرج أو دالة الحجم. يتم استخدام قيمة حجم التدرج كوزن للتصويت في [3] وقد تكون الخلايا مربعة (R-HOG) أو دائرية (C-HOG). (HOG). اعتماداً على ما إذا كان التدرج غير موقَّع أو موقَّع، تكون قنوات الهستوجرام متباعدة بالتساوي من 0 درجة إلى 180 درجة أو من 0 درجة إلى 360 درجة. يتم تباعد صناديق التوجيه في هذه الدراسة بالتساوي من 0 درجة إلى 180 درجة ويتم توفيرها بواسطة تسعة من 20 درجة، كما هو موضح في الجدول 2. على سبيل المثال تحتوي خلية واحدة على متجه سمات ذي 9 أبعاد. يتم المساهمة في كل صندوق عن طريق إضافة مقدار التدرج في الصندوق المقابل لاتجاه التدرج.

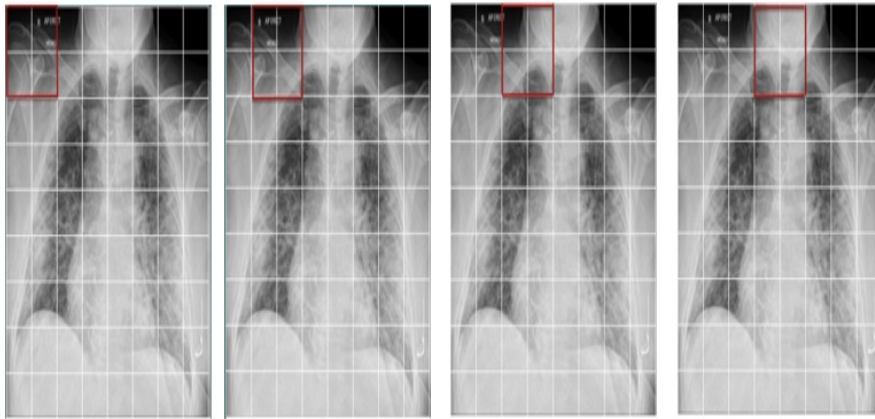
**Table 2: The orientation bins**

Bin number	Direction
1	0°-20°
2	20°-40°

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

3	40°-60°
4	60°-80°
5	80°-100°
6	100°-120°
7	120°-140°
8	140°-160°
9	160°-180°

**الخطوة 4:** تختلف قوى التدرج بشكل كبير بسبب الاختلافات المحلية في الإضاءة والتباين والخطوة الرابعة في الحصول على ميزات HOG هي تجنب التغييرات في الضوء والتباين. ويتم تحقيق ذلك من خلال إجراء تطبيع محلي للتدرج، والذي يستلزم دمج الخلايا في كتل أكبر متصلة مكانياً وإن موصوف HOG هو متجه يحتوي على مكونات جميع الهستوغرامات الخلوية الطبيعية في مناطق الكتل. الكتلة بحجم  $2 \times 2$  خلية. تتداخل هذه الكتل بشكل متكرر، مما يعني أن كل خلية تقدم مساهمات عديدة للميزة النهائية. وقد وجدت العديد من الدراسات أن كل أربع خلايا تشكل كتلة واحدة، كما هو موضح في الشكل 2.3.



الشكل (1.3) تقسيم الصورة إلى كتلة بحجم  $2 \times 2$

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

نقوم بتقدير خاصية HOG لكل خلية، ثم نقوم بدمج المصفوفات في مصفوفة واحدة مكونة من 36 مكونًا وتطبيعيها كما هو موضح في المعادلة (1.5).

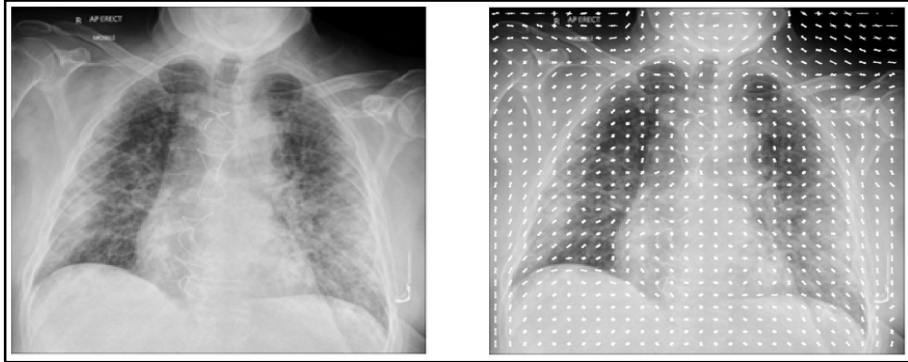
$$V = [a_1, a_2, a_3, \dots, a_{36}] \quad (1.5)$$

لتطبيع المتجه  $V$ ، يتم تقسيم جميع القيم في هذا المتجه على  $k$  كما هو محدد في المعادلة، وستكون النتيجة متجهًا موحدًا بحجم  $1 \times 36$  (1.6).

$$|V| = \left( \frac{a_1}{k}, \frac{a_2}{k}, \frac{a_3}{k}, \dots, \frac{a_{36}}{k} \right) \quad (1.6)$$

$$\text{where } K = \sqrt{(a_1)^2 + (a_2)^2 + (a_3)^2 + \dots + (a_{36})^2}$$

الخطوة الأخيرة لحساب ميزات HOG للصورة ولقد أنشأنا ميزات لكتلة واحدة فقط من الصورة. تم حساب الميزة الطبيعية لكل كتلة في الصورة المجزأة باستخدام المعادلة (1.6) وأخيرًا تم دمج كل هذه الميزات للحصول على ميزات الصورة الكاملة كما هو موضح في الشكل 3.4.



الشكل (2.3) (يسار) صورة إدخال. (يمين) ميزات HOG لصورة الإدخال.

#### 4: التصنيف Classification

سيتم استخدام خوارزمية التعلم الآلي (Support Vector Machine (SVM) بعد خطوة استخراج السمات ومع ذلك فهذه هي المرحلة النهائية في عملية التعرف على المرض، حيث يتم التعرف على المرض بناءً على استخراج مجموعة السمات من HOG.

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

SVM هي واحدة من خوارزميات التعلم الخاضع للإشراف والتي يمكن استخدامها في مشاكل التصنيف والانحدار، وعادة ما تستخدم في مشاكل التصنيف لفعاليتها والحصول على دقة ممتازة في معظم بيانات المستخدم وهي من نوعين: التصنيف الثنائي، حيث تصنف مجموعتين فقط، والتصنيف متعدد الفئات، حيث تصنف أكثر من مجموعة. الفكرة الرئيسية لخوارزمية SVM هي إيجاد أفضل مستوى فصل لفصل الفئات في هذه الدراسة تم استخدام مصنف SVM للتصنيف ويعتمد هذا المصنف على نظرية التعلم الإحصائي وبيانات التدريب المعطاة:  $\omega^T \cdot x = 0$

حيث يمثل كل من  $\omega$  و  $x$  متجهين ويمثلان  $\omega^T \cdot x$  حاصل الضرب النقطي لهذين المتجهين، وعادةً ما يُطلق على المتجه  $\omega$  اسم متجه الأوزان ويمكن تمثيل بيانات التدريب رياضياً على النحو التالي:

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \in R^n$$

وهذا يعني أن بياناتنا يتم تمثيلها بواسطة زوج من البيانات  $(x_n, y_n)$  حيث يمثل متجه  $x_n$  المعالم لمتجه المعالم وله أبعاد  $n$ ، ويمثل  $y_n$  تسمية  $x_n$ . عندما  $y_i = 1$ ، فهذا يعني أن البيانات تنتمي إلى المجموعة 1، وإذا  $y_i = -1$ ، فإنها تنتمي إلى المجموعة -1.

### النتائج والمناقشة Results and Discussion

للتحقق من صحة الإطار الذي تم إنشاؤه استخدمت هذه الدراسة ما مجموعه مكونة من 3886 صورة أشعة سينية لسرطانا لرئة للتدريب والاختبار وتحديد التحقق، كما هو موضح في الجدول رقم 1 لمعالجة مشكلة عدم اتساق البيانات، تم تغليف توزيع البيانات في هذه الدراسة تم التقاط مجموعة الاختبار قبل التدريب، ومع ذلك، تم جمع صور التحقق من صحة من مجموعة التدريب لأغراض التدريب، استخدمت هذه الدراسة 657 صورة طبيعية و588 صورة إيجابية لسرطان الرئة و659 حالة التهاب رئوي فيروسي. تم استخدام 402 صورة طبيعية و360 صورة إيجابية لسرطان الرئة و404 حالات التهاب رئوي فيروسي في الاختبارات تتضمن مجموعة بيانات التدريب صور الاختبار هذه كما هو موضح في الجدول

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

2، يحتوي هذا النظام أيضًا على 816 صورة تحقق لفئات الصور الطبيعية وسرطان الرئة والالتهاب الرئوي الفيروسي.

عدد الصور المستخدمة في التدريب والتحقق والاختبار في الفئات الطبيعية والسرطانية

Table3.

Dataset	Viral pneumonia	Normal	Cancer	Total images
Train	659	657	588	1904
Validation	282	282	252	816
Test	404	402	360	1166
Total	1345	1341	1200	3886

Table4. Confusion matrix

	Cancer	normal	Viral pneumonia
Cancer	346	4	10
Normal	1	382	19
Viral pneumonia	3	19	381

يظهر الجدول رقم 4 مصفوفة الارتباك التي تم الحصول عليها من خصائص الأداء المقاسة بالكامل للأنظمة أثناء خطوة التصنيف و أثناء التقييم، يتطلب النهج المقترح بيانات اختبار مسماة للتحقق من صحة البيانات الناتجة وتصور مصفوفة الارتباك الأداء العام للنظام ويكتشف النظام بشكل صحيح 346 صورة إيجابية من أصل 360 صورة تم اختبارها ويفشل في اكتشاف 14 صورة إيجابية لمرض السرطان، بينما يصنف النظام بشكل غير صحيح 20 صورة طبيعية للمرض ويجد بشكل صحيح 382 صورة طبيعية من أصل 402 صورة تم اختبارها ويفشل النظام في اكتشاف 22 صورة للالتهاب الرئوي الفيروسي ولكنه يكتشف بشكل صحيح 381 صورة للالتهاب الرئوي الفيروسي من أصل 404 صورة تم

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

اختباره وأظهرت النتائج أن دقة الطريقة المقترحة هي 96.6% مما يضمن أن تكون النتائج الحقيقية، سواء كانت إيجابية حقيقية أو سلبية حقيقية، عالية إلى حد ما الدقة العالية ضرورية في التشخيص وهذا يجعل الطريقة المقترحة موثوقة للتطبيقات السريرية حيث RAM8Gigabit وذاكرة وصول عشوائي Core i7 بلغ وقت المعالجة حوالي 0.56 ثانية على جهاز مزود بمعالج.

#### الخلاصة Conclusion:

تم استخدام الرسم البياني للتدرج الموجه (HOG) لصور الأشعة السينية لمرض السرطان الرئة وتصنيفها بطريقة دقيقة وبسيطة وتم اعتماد صور الأشعة السينية باعتبارها البنية الأكثر شيوعاً المستخدمة في التصوير الطبي الحيوي و أظهر النظام المطبق مقاييس أداء رائعة مع مصنف SVM من حيث الدقة وعلاوة على ذلك يمكن تعديل الخوارزمية بسهولة لإضافة صور إضافية (العادية والغير العادية) لتحسين الأداء و تفوق نظام المقترح (96.60 بالمائة) على تقنيات استخلاص الميزات الفردية الأخرى من حيث دقة التصنيف وأيضا تفوقت SVM على تقنيات التصنيف الأخرى مثل SoftMax من حيث دقة التصنيف يوفر النظام المقترح فائدة كبيرة لجميع الدول من حيث فحص وتشخيص المرض كبدل أسرع للطريقة الحالية المعتمدة على التفاعل المتسلسل.

#### التوصيات Recommendations:

- في ضوء النتائج التي توصلت إليها الدراسة، فقد توصلت الدراسة الي التوصيات الآتية:
1. تدريب النموذج آلة ناقل من خلال الدراسة على عينة صور أكبر للحصول على أفضل نتائج.
  2. تحسين جودة النموذج المستخدم من خلال دراسة العوامل التي تؤثر على نسبة دقته.
  3. استخدام نموذج الشبكات العصبية المقترح سيؤدي إلى الحصول على نتائج أكثر دقة في الكشف عن سرطان الرئة ومساعدة الأطباء في التشخيص المبكر وتجنب أخطاء التشخيص الناجمة عن القصور والأخطاء البشرية.



استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

4. تقديم النموذج المقترح وما تم التوصل إليه من نتائج إلى مركز الأورام بسببها لمساعدة الأطباء في الحصول على نتائج تشخيص دقيقة وبنسبة موثوقية عالية.

5. العوامل التي تم دراستها والخطوات التحسينية لنموذج آلة ناقل الدعم (SVM) التي تم تطبيقها في هذا البحث يمكن للباحثين في هذا المجال الاعتماد عليها في تجاربهم المستقبلية.

#### المراجع Reference:

- [1] Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., & Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine*, 54(24), 1451–1462.
- [2] Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., ... & Tian, J. H. (2020). Pei 465 YY et al: A new coronavirus associated with human respiratory disease in 466. China. *Nature*, 579(7798), 265–269.
- [3] Yang, X., Yu, Y., Xu, J., Shu, H., Liu, H., Wu, Y., ... & Shang, Y. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, 8(5), 475–481.
- [4] Wu, Z., & McGoogan, J. M. (2020). Outbreak in China: summary of a report of 72314 cases from the Chinese center for disease control and prevention. *JAMA*, 323(10.10), 1001.
- [5] Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., ... & Pillai, S. K. (2020). First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine*.

- [6] Kong, W., & Agarwal, P. P. (2020). Chest imaging appearance of COVID-19 infection. *Radiol Cardiothoracic Imaging* 2 (1): e200028.
- [7] Hu, Z., Ge, Q., Li, S., Jin, L., & Xiong, M. (2020). Artificial intelligence forecasting of covid-19 in china. *arXiv preprint arXiv:2002.07112*.
- [8] Zu, Z. Y., Jiang, M. D., Xu, P. P., Chen, W., Ni, Q. Q., Lu, G. M., & Zhang, L. J. (2020). Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology*, 296(2), E15-E25.
- [9] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4).
- [10] Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920-1930.
- [11] Sarker, I. H., Kayes, A. S. M., & Watters, P. (2019). Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. *Journal of Big Data*, 6(1), 1-28.
- [12] Mahdy, L. N., Ezzat, K. A., Elmousalami, H. H., Ella, H. A., & Hassanien, A. E. (2020). Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine. *Med Rxiv*.
- [13] Vickers, N. J. (2017). Animal communication: when i'm calling you, will you answer too? *Current biology*, 27(14), R713-R715.
- [14] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

learning in medical image analysis. Medical image analysis, 42, 60–88.

- [15] Altaf, F., Islam, S. M., Akhtar, N., & Janjua, N. K. (2019). Going deep in medical image analysis: concepts, methods, challenges, and future directions. IEEE Access, 7, 99540–99572.
- [16] Xing, F., Xie, Y., Su, H., Liu, F., & Yang, L. (2017). Deep learning in microscopy image analysis: A survey. IEEE transactions on neural networks and learning systems, 29(10), 4550–4568.
- [17] Muhammad, K., Khan, S., Del Ser, J., & de Albuquerque, V. H. C. (2020). Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey. IEEE Transactions on Neural Networks and Learning Systems, 32(2), 507–522.
- [18] Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., & Wang, J. (2018). Applications of deep learning to MRI images: a survey. Big Data Min Anal 1 (1): 1–18.
- [19] Seeböck, P., Orlando, J. I., Schlegl, T., Waldstein, S. M., Bogunovič, H., Klimescha, S., ... & Schmidt-Erfurth, U. (2019). Exploiting epistemic uncertainty of anatomy segmentation for anomaly detection in retinal OCT. IEEE transactions on medical imaging, 39(1), 87–98.
- [20] Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., & Ghassemi, M. (2020). Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988.

استخدام تقنيات الذكاء الاصطناعي للكشف عن حالات سرطان الرئة عن طريق الرسم البياني...  
إسماعيل أبو غرارة  
كمال السنوسي

[21] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886–893). Ieee.