HNSJ Humanities & Natural Sciences Journal
ISSN: (e) 2709-0833
www.hnjournal.net

مجلة العلوم الإنسانية والطبيعية
مجلة علمية محكمة (التصنيف: NSP)
معامل التأثير العربي للعام 2024 = 0.44

عنوان البحث

تقييم دقة أدوات الذكاء الاصطناعي في تحليل صور الأشعة السينية (X-ray) للكشف عن الأمراض التنفسية: دراسة مقارنة للدراسات الحديثة

د.ميلاد عرير[1]، مسرة صالح محمد الشعافي[1]، أمل أحمد يوسف العماري[1]، انتصار عمار رحومه عيسى[1]

[1] كلية تقنية الحاسوب، وزارة التعليم التقني والفني، دولة ليبيا.

بريد الكتروني الباحث المراسل: meladmohammed@gmail.com

HNSJ, 2026, 7(8); https://doi.org/10.53796/hnsj78/7

المعرف العلمي العربي للأبحاث: https://arsri.org/10000/78/7

تاريخ الاستقبال: 2026/07/05م تاريخ القبول: 2026/07/15م تاريخ النشر: 2026/08/01م

المستخلص

شهدت السنوات الأخيرة تطوراً ملحوظاً في توظيف تقنيات الذكاء الاصطناعي، وبشكل خاص نماذج التعلم العميق، في تحليل صور الأشعة السينية للصدر بهدف دعم وتشخيص الأمراض التنفسية المختلفة. تهدف هذه الدراسة إلى إجراء تحليل مقارن لخمس دراسات حديثة استخدمت تقنيات التعلم العميق في تشخيص الأمراض التنفسية من خلال صور الأشعة السينية للصدر. وقد تم تحليل الدراسات وفق مجموعة من المعايير الموحدة، شملت نوع النموذج المستخدم، وحجم البيانات وجودتها، واستراتيجيات التدريب، بالإضافة إلى مقاييس تقييم الأداء مثل الدقة والحساسية والنوعية ومنحنى ROC. أظهرت النتائج أن نماذج التعلم العميق حققت أداءً مرتفعاً، حيث تراوحت الدقة بين 93% و98% تبعاً لنوع المرض وطبيعة البيانات المستخدمة. كما ساهم استخدام تقنيات نقل التعلم في تحسين الأداء بشكل ملحوظ، خاصة في الحالات التي تعاني من محدودية البيانات التدريبية. ومع ذلك، أظهرت الدراسة وجود عدد من التحديات المنهجية التي تؤثر في موثوقية النتائج، من أبرزها عدم توازن البيانات بين الفئات المختلفة، ونقص البيانات عالية الجودة، وغياب التحقق الخارجي في بعض الدراسات. كما بينت المقارنة أن النماذج المدربة على بيانات متعددة المصادر أو المأخوذة من أكثر من مركز طبي أظهرت أداءً أكثر استقراراً وقدرة على التعميم مقارنة بالنماذج المعتمدة على مصدر واحد، والتي كانت أكثر عرضة لمشكلة فرط التكيف. وتشير هذه النتائج إلى أهمية تنوع البيانات واعتماد أساليب تقييم أكثر شمولاً لتعزيز موثوقية أنظمة الذكاء الاصطناعي وزيادة إمكانية تطبيقها في البيئات السريرية الواقعية.

الكلمات المفتاحية: التعلم العميق، الذكاء الاصطناعي، صور الأشعة السينية للصدر، تشخيص الأمراض التنفسية، الشبكات العصبية الالتفافية (CNN).

RESEARCH TITLE

Evaluating the Accuracy of Artificial Intelligence Tools in Analyzing X-Ray Images for Detecting Respiratory Diseases: A Comparative Study of Recent Research

Abstract

Recent years have witnessed remarkable advances in the application of artificial intelligence technologies, particularly deep learning models, to chest X-ray analysis for supporting the diagnosis of various respiratory diseases. This study aims to conduct a comparative analysis of five recent studies that employed deep learning techniques to diagnose respiratory diseases using chest X-ray images. The studies were analyzed according to a set of standardized criteria, including the type of model used, dataset size and quality, training strategies, and performance evaluation metrics such as accuracy, sensitivity, specificity, and the receiver operating characteristic (ROC) curve. The findings showed that deep learning models achieved high performance, with accuracy rates ranging from 93% to 98%, depending on the type of disease and the nature of the data used. Transfer learning techniques also contributed significantly to improving performance, particularly in cases involving limited training data. However, the study identified several methodological challenges affecting the reliability of the findings, most notably class imbalance, a lack of high-quality data, and the absence of external validation in some studies. The comparison further revealed that models trained on multi-source data or data collected from multiple medical centers demonstrated more stable performance and greater generalizability than models relying on a single data source, which were more susceptible to overfitting. These findings highlight the importance of data diversity and the adoption of more comprehensive evaluation methods to enhance the reliability of artificial intelligence systems and increase their applicability in real-world clinical settings.

Key Words: Deep learning; artificial intelligence; chest X-ray images; respiratory disease diagnosis; convolutional neural networks (CNNs).

1. مقدمة البحث :

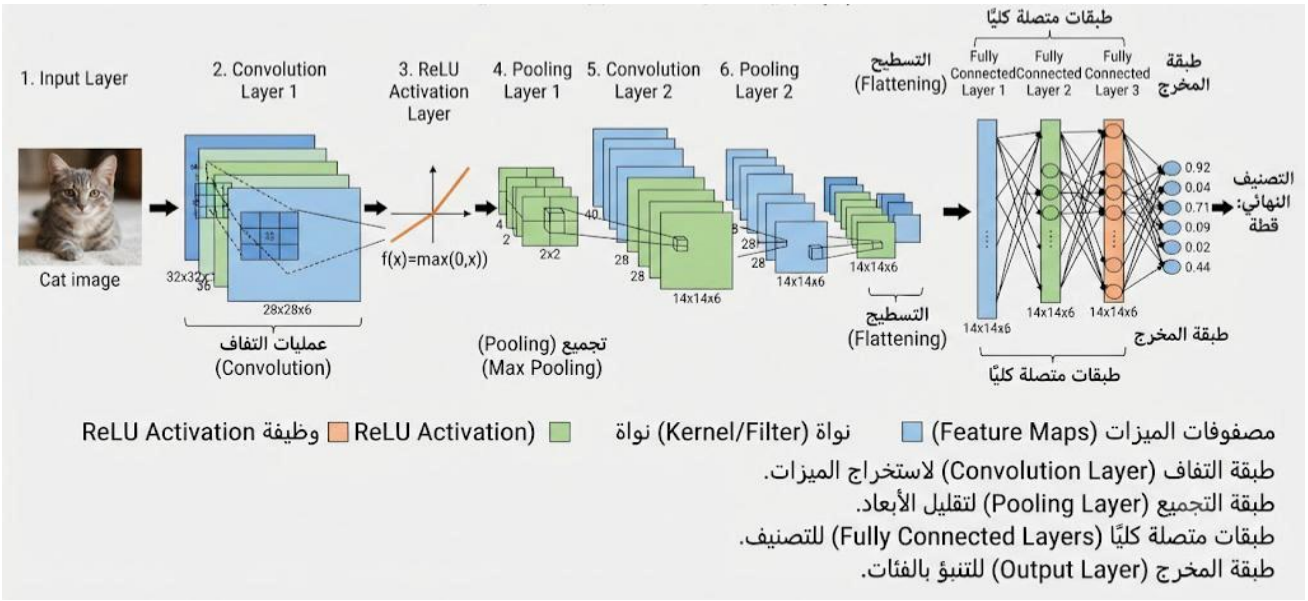
شهد القطاع الطبي خلال السنوات الأخيرة تطوراً ملحوظاً في استخدام تقنيات الذكاء الاصطناعي (Artificial Intelligence) بهدف تحسين عمليات التشخيص الطبي ودعم اتخاذ القرار السريري. ومن أهم المجالات التي استفادت من هذا التطور مجال تحليل الصور الطبية، وخاصة صور الأشعة السينية للصدر (Chest X-ray)، التي تُعد من أكثر أدوات التصوير استخداماً للكشف عن مجموعة واسعة من الأمراض التنفسية مثل الالتهاب الرئوي، السل، كوفيد-19، وسرطان الرئة (Khan et al., 2021; Rahman et al., 2020).

أسهمت تقنيات التعلم العميق (Deep Learning)، وعلى رأسها الشبكات العصبية الالتفافية (Convolutional Neural Networks – CNN)، في إحداث نقلة نوعية في تحليل الصور الطبية، حيث أثبتت قدرتها على تعلم السمات الهرمية المعقدة من بيانات الصور مباشرة، متجاوزة الحاجة إلى استخلاص الخصائص يدوياً (Siddiqi & Javaid, 2024). وقد طُبقت نماذج مثل ResNet وDenseNet وVGG بنجاح في مهام تصنيف وكشف الأمراض من صور الصدر الشعاعية، محققة أداءً ينافس، بل ويتفوق أحياناً، على أطباء الأشعة في ظروف بحثية محكومة (Guo et al., 2024; Showkat & Qureshi, 2022).

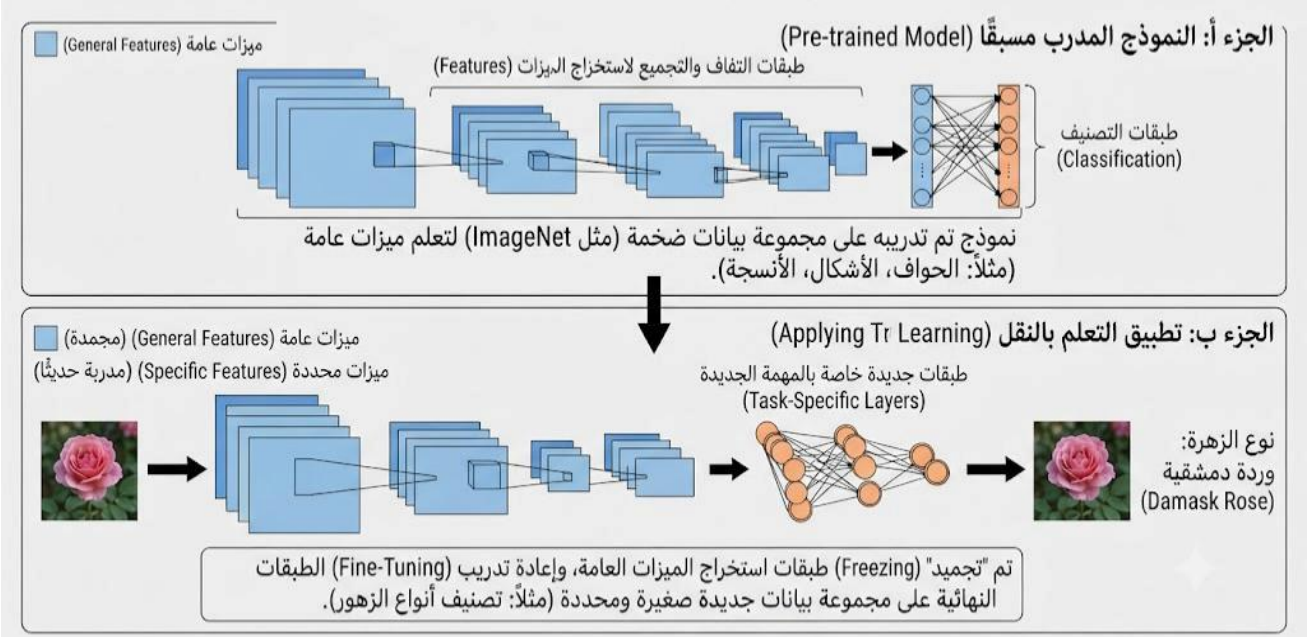
غير أن الانتقال من البيئة البحثية إلى التطبيق السريري الحقيقي يواجه تحديات كبيرة، أهمها التفاوت الملحوظ في دقة النماذج المُعلن عنها عبر الدراسات المختلفة. هذا التفاوت يُعزى إلى عوامل عدة، منها اختلاف أحجام قواعد البيانات وخصائصها الديموغرافية، تباين أساليب المعالجة المسبقة للصور، اختلاف معماريات النماذج واستراتيجيات التدريب، بالإضافة إلى تباين مقاييس التقييم المعتمدة (Harris et al., 2019; Medical Reports, 2024). لذلك، تبرز الحاجة إلى دراسة مقارنة تحلل هذه العوامل بشكل منهجي لتقديم فهم أعمق لمصادر التباين في أداء هذه النماذج.

يهدف هذا البحث إلى إجراء دراسة مقارنة لخمس أوراق علمية منشورة حديثاً، تمثل تطبيقات متنوعة للذكاء الاصطناعي في تحليل صور الأشعة السينية للكشف عن الأمراض التنفسية. تسعى الدراسة إلى تقييم دقة هذه الأدوات وفق معايير موحدة، وتحديد نقاط القوة والقيود المرتبطة بكل نهج، واقتراح اتجاهات مستقبلية لتطوير أنظمة أكثر موثوقية وقابلية للتعميم في السياق السريري.

إضافة إلى ذلك، فإن جودة الصور الشعاعية الأولية تؤدي دورًا حاسمًا في أداء النماذج، حيث تُسهم تقنيات تحسين الصور المعتمدة على الذكاء الاصطناعي في تقليل التشويش وتحسين التباين، مما قد ينعكس إيجابًا على دقة التشخيص الآلي (Song et al., 2025)



شكل (1): بنية الشبكة العصبية الالتفافية CNN



شكل (2): مفهوم Transfer Learning

2. مشكلة البحث :

على الرغم من الانتشار الواسع والنتائج المشجعة لاستخدام الذكاء الاصطناعي في تحليل صور الأشعة السينية، يظل هناك تباين واضح في نتائج الدراسات المنشورة فيما يتعلق بدقة التشخيص وكفاءة النماذج المُستخدمة (Almalki et al., 2021; Uppin & George, 2024). هذا التباين يثير تساؤلات جوهرية حول إمكانية الاعتماد على هذه التقنيات كأداة مساعدة في البيئات السريرية المتنوعة. تكمن مشكلة البحث المركزية في غياب تحليل مقارن موحد ومتعمق يوضح العلاقة بين العوامل المنهجية الرئيسية – مثل معمارية النموذج، حجم وجودة بيانات التدريب، طرق التحقق من النتائج – ومدى تأثير هذه العوامل مجتمعةً على دقة الكشف عن الأمراض التنفسية.

ومن هنا يطرح البحث تساؤله الرئيسي: ما مدى دقة أدوات الذكاء الاصطناعي المستخدمة في تحليل صور الأشعة السينية للكشف عن الأمراض التنفسية؟ وما العوامل المنهجية والتقنية التي تؤثر على كفاءتها وإمكانية تعميم نتائجها ؟

3. أهداف البحث :

يسعى البحث إلى تحقيق الأهداف التالية :

1. استعراض تطبيقات الذكاء الاصطناعي المعاصرة المُستخدمة في تحليل صور الأشعة السينية للصدر.
2. مراجعة وتحليل خمس دراسات علمية حديثة بشكل نقدي تغطي الكشف عن أمراض تنفسية مختلفة باستخدام الذكاء الاصطناعي.
3. مقارنة أداء النماذج المختلفة وفق معايير تقييم الأداء الموحدة.
4. تحديد وتحليل العوامل المؤثرة على تباين دقة النماذج.
5. توضيح أبرز التحديات المنهجية والعملية التي تواجه ترجمة هذه النماذج إلى الممارسة الطبية.
6. تقديم توصيات مبنية على الأدلة لتحسين تصميم وتقييم هذه التقنيات في الدراسات المستقبلية.

4. منهجية البحث :

لتحقيق أهداف الدراسة ، تم اعتماد منهج المراجعة التحليلية المقارنة (Comparative Analytical Review)، والذي يهدف إلى تحليل ومقارنة أداء نماذج الذكاء الاصطناعي المستخدمة في تفسير صور الأشعة السينية للصدر في الكشف عن الأمراض التنفسية.

تم اختيار خمس دراسات علمية أصلية منشورة حديثًا ضمن الفترة (2020–2024)، بحيث تغطي نطاقًا متنوعًا من الأمراض التنفسية، وتستخدم نماذج مختلفة من تقنيات التعلم العميق، مما يسمح بإجراء مقارنة منهجية عادلة بين الأساليب المختلفة.

وقد شملت الدراسات المختارة ما يلي :

Rahman et al. (2020): دراسة ركزت على الكشف عن الالتهاب الرئوي باستخدام تقنيات التعلم بالنقل (Transfer Learning).

Showkat & Qureshi (2022): دراسة تناولت الكشف عن كوفيد–19 باستخدام نماذج ResNet مدعومة بآلية الانتباه (Attention Mechanism).

Bharati et al. (2020): دراسة طورت نموذجًا هجينًا يجمع بين الشبكات العصبية الالتفافية (CNN) وLSTM لتصنيف أمراض رئوية متعددة.

Homayounieh et al. (2021): دراسة متعددة المراكز لتقييم نموذج ذكاء اصطناعي للكشف عن العقيدات الرئوية وتحليل تأثيره على أداء أطباء الأشعة.

Guo et al. (2024): دراسة مقارنة بين أداء أطباء الأشعة ونموذج تعلم عميق في تشخيص عدة أمراض صدرية باستخدام مجموعة بيانات كبيرة الحجم.

تم تحليل كل دراسة وفق إطار موحد يضمن الاتساق في المقارنة، حيث شمل التحليل مجموعة من المعايير المنهجية والتقنية، وهي: سنة النشر، هدف الدراسة، نوع المرض المستهدف، خصائص وحجم قاعدة البيانات، بنية نموذج الذكاء الاصطناعي المستخدم، استراتيجية التدريب والتحقق، إضافة إلى بروتوكولات التقييم المعتمدة ومقاييس الأداء، والتي تضمنت: الدقة (Accuracy)، الحساسية (Sensitivity)، النوعية (Specificity)، الدقة التنبؤية (Precision)، ودرجة F1.

كما تم استخراج النتائج الرئيسة لكل دراسة بشكل منظم، ثم تصنيفها ضمن جدول مقارن يهدف إلى تسهيل التحليل النقدي وتحديد أوجه التشابه والاختلاف بين النماذج من حيث الأداء والموثوقية وقابلية التعميم.

5. تحليل ومقارنة الدراسات السابقة :

يقدم هذا القسم تحليلاً للدراسات الخمس المختارة، يليه جدول مقارن يلخص أبرز جوانبها.

الدراسة الأولى: Rahman et al. (2020)

استهدفت هذه الدراسة الكشف عن الالتهاب الرئوي الفيروسي والبكتيري مقابل الحالات الطبيعية باستخدام صور أشعة سينية من قاعدة بيانات مفتوحة (Guangzhou Medical Center). طبقت الدراسة أربعة نماذج لنقل التعلم

(AlexNet, ResNet18, DenseNet201, SqueezeNet) مدربة مسبقاً على ImageNet. تم تقييم الأداء على مجموعة اختبار منفصلة، وحقق نموذج DenseNet201 أعلى أداء بدقة كلية 93.3%، وحساسية 93.1%، ونوعية 93.4%. أظهرت النتائج أن نقل التعلم يقدم حلاً فعالاً في حالات نقص البيانات الطبية، لكن النموذج لم يُختبر على بيانات خارجية من مراكز أخرى.

الدراسة الثانية: Showkat & Qureshi (2022)

ركزت هذه الدراسة على تصنيف صور الأشعة السينية لكوفيد-19 والالتهاب الرئوي الفيروسي والحالات الطبيعية. اعتمدت على نماذج ResNet50 وResNet101 مع دمج آلية الانتباه (Attention Mechanism). استخدمت مجموعة بيانات مدمجة من مصادر متعددة بلغت حوالي 7000 صورة. حقق نموذج ResNet50 مع الانتباه نتائج استثنائية بدقة 98.1%، وحساسية 98.3%، ونوعية 97.8% لمهمة كشف كوفيد-19. تميزت الدراسة بتقنيات متقدمة للتعامل مع عدم التوازن في البيانات، إلا أن تجميع البيانات من مصادر مختلفة قد يُدخل تحيزات تؤثر على القابلية للتعميم.

الدراسة الثالثة : Bharati et al. (2020)

قدمت نموذجاً هجيناً يجمع بين CNN التقليدية وطبقات الذاكرة طويلة-قصيرة المدى (LSTM) لتصنيف صور الصدر إلى فئات متعددة تشمل الالتهاب الرئوي، استرواح الصدر، وأمراض أخرى، بالإضافة إلى الحالات الطبيعية. استخدمت الدراسة قاعدتي بيانات عامتين (NIH ChestX-ray14 و CheXpert) بإجمالي يفوق 110,000 صورة. حقق النموذج الهجين دقة كلية 93.7% في تصنيف 14 مرضاً، متفوقاً على نماذج CNN المنفردة. أظهرت الدراسة إمكانية التقاط التبعيات المكانية والزمنية (عبر LSTM) في الصور، لكن تعقيد النموذج يجعله مرتفع التكلفة الحسابية ويحتاج كميات ضخمة من البيانات.

الدراسة الرابعة : Homayounieh et al. (2021)

في هذه الدراسة المتعددة المراكز، طُور نموذج ذكاء اصطناعي قائم على CNN للكشف عن العقيدات الرئوية (كمؤشر مبكر لسرطان الرئة) في صور الأشعة السينية من ثلاثة مراكز طبية في ألمانيا والولايات المتحدة. بلغ حجم مجموعة البيانات حوالي 6,000 صورة. كان التقييم فريداً حيث قارن أداء 17 طبيب أشعة بمساعدة النموذج وبدونه. أظهرت النتائج أن حساسية الأطباء ارتفعت من 48.1% إلى 56.3% عند استخدام النموذج كأداة مساعدة (بفارق إحصائي دال)، بينما انخفضت النوعية بشكل طفيف. تشير هذه النتائج إلى القيمة الحقيقية للذكاء الاصطناعي كأداة داعمة وليس بديلة، وتبرز أهمية التحقق السريري في بيئات متعددة.

الدراسة الخامسة: Guo et al. (2024)

صممت هذه الدراسة الفريدة مسابقة تنافسية بين تسعة من أطباء الأشعة ونموذج تعلم عميق (DenseNet-121) للكشف عن أربعة أمراض صدرية (الالتهاب الرئوي، استرواح الصدر، الانصباب الجنبي، وكسور الأضلاع) في مجموعة بيانات صينية كبيرة (أكثر من 200,000 صورة). تم تدريب النموذج واختباره داخلياً، ثم أجري اختبار خارجي على بيانات مستشفى آخر. أظهر النموذج أداءً أعلى معنوياً من الأطباء غير المدعومين بالذكاء الاصطناعي (دقة 95.2% مقابل 89.7%). ولكن عند استخدام الأطباء لمخرجات النموذج، تحسن أداؤهم ليقترب من أداء النموذج (93.5%). هذه الدراسة تؤكد إمكانية تفوق النماذج في الظروف المضبوطة، ولكنها تشدد على أن "الذكاء الاصطناعي المعزز" (AI-augmented radiologist) يمثل السيناريو الأكثر أماناً وفاعلية.

جدول (1): مقارنة شاملة بين الدراسات الخمس

المعيار	Bharati et al. (2020)	Showkat & Qureshi (2022)	Rahman et al. (2020)	Homayouni eh et al. (2021)	Guo et al .(2024)
المرض المستهدف	الالتهاب الرئوي (فيروسي/بكتيري)	كوفيد-19 والالتهاب الفيروسي	أمراض رئوية متعددة (14 فئة)	العقيدات الرئوية	4 أمراض صدرية متنوعة
قاعدة البيانات (حجم)	~5,856 صورة	~7,000 صورة	صورة+ 110,000	صورة ~6,000	+ 200,000 صورة
معمارية النموذج	نقل تعلم(DenseNet201)	ResNet50+Attention	CNN + LSTM هجينة	((Retina Net CNN خاصة	DenseNet-121
الدقة Accuracy	%93.3	%98.1 (لكوفيد-19)	%93.7 (14 فئة)	غير متاحة بشكل كلي (قيم الحساسية)	%95.2
الحساسية Sensitivity	%93.1	%98.3	غير مفصلة لجميع الفئات	%56.3 (باستخدام AI)	%93.8
ميزة التقييم الرئيسية	اختبار على مجموعة منفصلة داخلية	استخدام آلية الانتباه	التصنيف المتعدد الفئات	دراسة متعددة المراكز مع أطباء	مسابقة ضد أطباء، واختبار خارجي
نقاط القوة البارزة	إثبات فعالية نقل التعلم	أداء ممتاز، آلية الانتباه	قدرة على تعلم أمراض متعددة	تقييم واقعي في بيئة سريرية متعددة	تصميم قوي، حجم بيانات كبير، اختبار خارجي
القيود الرئيسية	قاعدة بيانات واحدة، لا اختبار خارجي	خطر تحيز البيانات المدمجة	تعقيد النموذج، يحتاج بيانات ضخمة	حساسية أساسية منخفضة للعقيدات	يتطلب موارد حسابية هائلة

6. الفجوة البحثية :

باستعراض الأدبيات، يتضح أن الغالبية العظمى من الدراسات المنشورة (مثل Rahman et al., 2020; Showkat & Qureshi, 2022; Bharati et al., 2020) تتبنى نهج "تطوير واختبار" نموذج واحد على مجموعة بيانات محدودة، وغالباً ما يكون الاختبار داخلياً. هذه الدراسات تقدم أدلة أولية على جدوى التقنيات، لكنها لا تخوض في تحليل مقارن منهجي بين معماريات أو منهجيات مختلفة لاستخلاص أنماط الأداء العامة. بينما توجد مراجعات منهجية (مثل Khan et al., 2021; Harris et al., 2019) تقارن النتائج عبر أوراق متعددة، إلا أنها غالباً ما تتوقف عند سرد الفروقات الرقمية في الدقة دون تحليل عميق للعوامل المنهجية التي تقف وراء هذا التباين، مثل تأثير جودة البيانات وطرق تقسيمها، واستراتيجيات ما قبل المعالجة، وبروتوكولات التحقق. وبالتالي، تظهر فجوة بحثية واضحة تتمثل في الحاجة إلى تحليل مقارن نقدي يتجاوز مجرد عرض النتائج، ليربطها بشكل سببي بالمنهجيات المتبعة، وهو ما تسعى هذه الدراسة إلى سده.

7. الإضافة العلمية للبحث :

تتمثل الإضافة العلمية لهذا البحث، مقارنة بالمراجعات السابقة ، في النقاط التالية:

1. التحليل المقارن الموحد والمتعمق : لا يقتصر البحث على عرض نتائج الدراسات المختارة، بل يقدم تحليلاً نقدياً موازياً لها وفق مجموعة معايير تقييم محددة ومتناسقة، مما يمكن من إجراء مقارنة عادلة وذات مغزى.

2. ربط الأداء بالمنهجية : يركز البحث بشكل خاص على تحليل أسباب اختلاف نتائج الدراسات، رابطاً بين اختيارات التصميم (مثل حجم البيانات، مصدرها، معمارية النموذج) ومقاييس الأداء المُبلغ عنها. هذا يتجاوز وصف "ماذا" كانت النتيجة إلى تفسير "لماذا" حدثت.

3. التقييم متعدد الأبعاد للقابلية السريرية: لا يكتفي البحث بتقييم الدقة النظرية، بل يناقش بشكل نقدي مدى قابلية كل نموذج للاعتماد عليه كأداة مساعدة للأطباء في بيئة طبية حقيقية، مستفيداً من إدراج دراسات مثل Homayounieh et al. (2021) و Guo et al. (2024) التي اختبرت التفاعل بين الإنسان والآلة.

4. توصيات قابلة للتنفيذ: بناءً على الفجوات التي تم تحديدها، يقدم البحث توصيات مستقبلية محددة لتحسين تصميم وتطبيق وتقييم نماذج الذكاء الاصطناعي في هذا المجال الحيوي، مما يوجه جهود الباحثين والمطورين.

8. النتائج والتوصيات :

أكدت هذه الدراسة المقارنة أن تقنيات الذكاء الاصطناعي، وخصوصاً الشبكات العصبية الالتفافية العميقة ونماذج نقل التعلم، تمتلك قدرة عالية وكامنة على تحليل صور الأشعة السينية للصدر والكشف عن طيف من الأمراض التنفسية بدقة تنافسية، بلغت في أفضل الحالات 98% (Showkat & Qureshi, 2022). ومع ذلك، يتضح بشكل جلي أن كفاءة هذه النماذج ليست مطلقة، بل تعتمد بشكل حاسم على عدة عوامل منهجية محورية:

1. جودة وحجم البيانات: النماذج المدربة على قواعد بيانات ضخمة ومتنوعة مثل Guo et al. (2024) أظهرت قابلية أفضل للتعميم. على النقيض، النماذج المعتمدة على بيانات محدودة عانت من خطر الإفراط في التلاؤم (Overfitting) وانخفاض الأداء عند الاختبار الخارجي (Rahman et al., 2020).

2. معمارية النموذج واستراتيجية التدريب: نماذج نقل التعلم (Transfer Learning) تقدم حلاً قوياً لنقص البيانات، بينما النماذج المطورة من الصفر تحتاج موارد بيانات ضخمة (Bharati et al., 2020). كما أن إضافة آليات الانتباه

(Showkat & Qureshi, 2022) حسّنت الأداء.

3. طريقة التحقق من النتائج: الفجوة بين الأداء في الاختبار الداخلي والأداء في العالم الحقيقي تظل كبيرة. الدراسات الوحيدة التي تضمنت تحققاً خارجياً أو تفاعلاً مع أطباء في بيئات متعددة المراكز (Homayounieh et al., 2021; Guo et al., 2024) قدمت الصورة الأكثر واقعية حول الفائدة السريرية، والتي تمثلت في "دعم" القرار الطبي بدلاً من الاستبدال الكامل.

التوصيات المستقبلية :

بناءً على ما سبق ، يوصي البحث بما يلي لتحسين موثوقية وتطبيق هذه التقنيات :

1. بناء قواعد بيانات موحدة ومتعددة المراكز: يجب أن تتعاون المؤسسات الطبية لبناء قواعد بيانات ضخمة، متنوعة ديموغرافياً ومرضياً، ومُعنونة بدقة من قبل خبراء، لتكون معياراً ذهبياً لتدريب واختبار النماذج.
2. اعتماد بروتوكولات تقييم صارمة: يجب أن تتضمن الدراسات المستقبلية حتماً اختباراً خارجياً على بيانات من مراكز مستقلة، واعتماد مقاييس تقييم شاملة تتجاوز الدقة الكلية (مثل منحنى AUC، الحساسية، والنوعية لكل فئة).
3. تطوير نماذج قابلة للتفسير: لدعم ثقة الأطباء، يجب تصميم نماذج توفر خرائط تموضع (مثل Grad-CAM) لتوضيح المناطق التي استند إليها النموذج في قراره، مما يسهل عملية التحقق البشري.
4. التركيز على التفاعل بين الإنسان والآلة: يجب أن يتحول هدف الأبحاث من تطوير نموذج "مستقل" إلى بناء نظام "ذكاء اصطناعي معزز" يثبت فعاليته في تحسين أداء الطبيب وتقليل زمن التشخيص في تجارب سريرية مضبوطة.
5. الاستثمار في تحسين جودة الصور: ينبغي دمج تقنيات تحسين الصور القائمة على الذكاء الاصطناعي (مثل تقليل الضوضاء وزيادة حدة الحواف) كخطوة معالجة مسبقة معيارية، حيث أظهرت الدراسات الحديثة أن تحسين جودة صور الأشعة السينية يمكن أن يرفع من أداء نماذج التشخيص اللاحقة (Song et al., 2025).

9. الخاتمة :

يمثل الذكاء الاصطناعي نقلة نوعية في تحليل صور الأشعة السينية للصدر، مبشراً بتشخيص أسرع وأدق للأمراض التنفسية. لقد أوضحت هذه الدراسة المقارنة، من خلال تحليل خمس أوراق بحثية متنوعة، أن مستوى الدقة العالية الذي تم تحقيقه في البيئات البحثية (93-98%) مشروط بعوامل منهجية صارمة. يظل التحدي الأكبر ليس في تطوير خوارزميات أكثر تعقيداً، بل في ضمان قابليتها للتعميم وأمانها وثقة الأطباء بها عبر التحقق الخارجي الصارم، وبناء قواعد بيانات ضخمة عالية الجودة، وتصميم أنظمة تركز على تعزيز القدرات البشرية لا استبدالها. بهذه الخطوات، يمكن للأداة أن تنتقل بثبات من المختبر إلى العيادة، محققة أثرها الحقيقي في رعاية المرضى.

References

Almalki, Y. E., Qayyum, A., Irfan, M., Haider, N., Glowacz, A., Alshehri, F. M., Alduraibi, S. K., Alshamrani, K., Alkhalik Basha, M. A., Saeed, M. K., & Rahman, S. (2021). A novel method for COVID-19 diagnosis using artificial intelligence in chest X-ray images. Healthcare, 9(5), 522.

.2Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. Informatics in Medicine Unlocked, 20, 100391.

.3Guo, L., Zhou, C., Xu, J., Huang, C., Yu, Y., & Lu, G. (2024). Deep learning for chest X-ray diagnosis: Competition between radiologists with or without artificial intelligence assistance. Journal of Imaging Informatics in Medicine, 37, 922–934.

.4Harris, M., Qi, A., Jeagal, L., Torabi, N., Menzies, D., Korobitsyn, A., Pai, M., Nathavitharana, R. R., & Khan, F. A. (2019). A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. PLOS ONE, 14(9), e0221339.

.5Homayounieh, F., Digumarthy, S., Ebrahimian, S., Rueckel, J., Hoppe, B. F., Sabel, B. O., Conjeti, S., Ridder, K., Sistermanns, M., Preuhs, A., Ghesu, F., Mansoor, A., Moghbel, M., Botwin, A., Cartmell, S., Patti, J., Huemmer, C., Fieselmann, A., Joerger, C., . . . Kalra, M. (2021). An artificial intelligence-based chest X-ray model on human nodule detection accuracy from a multicenter study. JAMA Network Open, 4(12), e2141096.

.6Khan, W., Zaki, N., & Ali, L. (2021). Intelligent pneumonia identification from chest X-rays: A systematic literature review. IEEE Access, 9, 51747–51762.

.7Medical Reports. (2024). Recent advancement of deep learning techniques for pneumonia prediction from chest X-ray image. Medical Reports, 7, 100106.

.8Rahman, T., Chowdhury, M. E. H., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., & Kashem, S. (2020). Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Applied Sciences, 10(9), 3233.

.9Showkat, S., & Qureshi, S. (2022). Efficacy of transfer learning-based ResNet models in Chest X-ray image classification for detecting COVID-19 Pneumonia. Chemometrics and Intelligent Laboratory Systems, 224, 104534.

.10Siddiqi, R., & Javaid, S. (2024). Deep learning for pneumonia detection in chest X-ray images: A comprehensive survey. Journal of Imaging, 10(7), 176.

.11Song, L., Sun, H., Xiao, H., Lam, S. K., Zhan, Y., Ren, G., & Cai, J. (2025). Artificial intelligence for chest X-ray image enhancement. Radiation Medicine and Protection, 6(1), 61–68.

.12Uppin, C., & George, G. (2024). Assessing the efficacy of transfer learning in chest X-ray image classification for respiratory disease diagnosis: Focus on COVID-19, lung opacity, and viral pneumonia. Journal of Image Processing and Artificial Intelligence, 10(1), 11–20.