"دمج تقدير عدم اليقين والتعلّم النشط لتعزيز موثوقية التشخيص الطبي المعتمد على التعلّم العميق"

"Integrating Uncertainty Estimation and Active Learning for Reliable Deep Learning-Based Medical Diagnosis"

الملخص

تهدف هذه المقالة إلى تعزيز موثوقية نماذج التعلّم العميق الطبية من خلال دمج منهجيات تقدير عدم اليقين والتعلّم النشط (Active Learning)، وتعالج مشكلتين محوريتين في تحليل الصور الشعاعية: ضعف القدرة على تفسير ثقة النموذج، وندرة البيانات المعلّمة. تستعرض المقالة أبرز أساليب تقدير عدم اليقين، موضّحة دورها في كشف الحالات الغامضة، وتقييم ثبات النموذج دون الحاجة إلى إعادة تدريبه. كما تناقش كيفية مساهمة التعلّم النشط في اختيار العينات الأكثر أهمية، مما يقلّل الحاجة إلى مجموعات بيانات كبيرة، ويعزّز تمثيل الحالات النادرة. تُظهر حالتا COVID-19 والالتهاب الرئوي اختلاف دور عدم اليقين بين المرضين؛ حيث ساعد تقدير عدم اليقين في COVID-19 على مواجهة الضبابية، بينما ركّز في حالة الالتهاب الرئوي على تمييز الأنماط المتداخلة.

تشير النتائج إلى أن الجمع بين عدم اليقين والتعلّم النشط لا يحسّن الدقة فحسب، بل يقدّم آلية أفضل لتقدير الثقة السريرية. وبذلك، تصبح النماذج الطبية أكثر قابلية للاعتماد العملي في بيئات التشخيص.

كلمات مفتاحية:

تقدير عدم اليقين، التعلّم النشط، عدم اليقين العشوائي، عدم اليقين المعرفي، التعلم العميق.

Abstract

This paper aims to enhance the reliability of deep learning models in medical imaging by integrating uncertainty estimation methods with active learning. It addresses two central challenges in radiological image analysis: the limited ability to interpret model confidence and the scarcity of labeled medical data. The paper reviews major uncertainty estimation approaches, highlighting their role in identifying ambiguous cases and evaluating model stability without the need for retraining. It also discusses how active learning contributes to selecting the most informative samples, thereby reducing the dependence on large datasets and improving the representation of rare cases.

Case studies on COVID-19 and pneumonia demonstrate the differing roles of uncertainty across the two diseases. In COVID-19, uncertainty estimation helps address visual ambiguity, whereas in pneumonia it focuses on distinguishing overlapping radiological patterns.

The results indicate that combining uncertainty estimation with active learning not only improves predictive accuracy, but also provides a more effective mechanism for assessing clinical confidence. Consequently, medical deep learning models become more suitable for practical deployment in diagnostic environments.

Keywords:

Uncertainty Quantification, Active Learning, Aleatoric Uncertainty, Epistemic Uncertainty, Deep Learning.

1. المقدمة

شهدت السنوات الأخيرة توسعًا ملحوظًا في اعتماد تقنيات التعلّم العميق (Deep Learning) داخل التطبيقات الطبية، ولا سيما ضمن تحليل صور الأشعّة السينية والرنين المغناطيسي والتصوير المقطعي، حيث أثبتت الشبكات العصبية الالتفافية (Convolutional Neural Networks, CNNs) قدرتها على اكتشاف الأورام والآفات بدقة عالية تنافس أداء الاختصاصيين في العديد من المهام التشخيصية [1]. وعلى الرغم من هذا التقدّم، ما تزال هذه النماذج تُعامل غالبًا كـ "صناديق سوداء" تُصدر أحكامًا نهائية بثقة رقمية مرتفعة دون أن تعبّر فعليًا عن درجة اليقين وراء القرار، إذ قد يظهر النموذج واثقًا بنسبة 99% رغم أن هذه الثقة قد تكون غير مبرّرة نتيجة ضجيج البيانات أو نقص العينات أو التمثيل غير المتوازن للفئات النادرة أثناء التدريب [2]. وقد دفع هذا التباين بين "الثقة العددية" و"اليقين الحقيقي" إلى تركيز الأبحاث الحديثة على مفهوم عدم اليقين بوصفه عنصرًا أساسيًا لضمان موثوقية النظم الطبية المعتمدة على التعلّم العميق [3].

كما برزت الحاجة إلى معالجة مشكلة جوهرية في التطبيقات الطبية تتمثل في ندرة البيانات المعلَّمة وارتفاع تكلفة الحصول عليها. وهنا ظهر التعلّم النشط (Active Learning) بوصفه منهجية تعتمد على قراءات عدم اليقين نفسها لتحديد العينات الأكثر فائدة للنموذج، الأمر الذي يُحسّن الأداء بكلفة أقل، ويجعل عدم اليقين عنصرًا مركزيًا لا في تفسير القرار وحسب، بل في توجيه مسار التعلّم ذاته [4].

وانطلاقًا من هذه المعطيات، تكمن أهمية هذه المقالة في توضيح الدور المتكامل لمنهجيات تقدير عدم اليقين والتعلّم النشط في تعزيز موثوقية نماذج التشخيص الطبي، خصوصًا في ظل ندرة البيانات المعلّمة وارتفاع متطلبات الدقة والاعتمادية السريرية. وتهدف هذه المقالة إلى استعراض أبرز منهجيات تقدير عدم اليقين، وبيان آلية تكاملها مع استراتيجيات التعلّم النشط، وتحليل أثر هذا التكامل في تحسين موثوقية النماذج الطبية وتقليل الحاجة إلى البيانات الموسومة، وذلك بالاستناد إلى تطبيقات في تشخيص مرضى COVID-19 والالتهاب الرئوي.

2. مفهوم عدم اليقين وأنواعه

يشكّل عدم اليقين عنصرًا أساسيًا في تقييم موثوقية أنظمة التعلّم العميق الطبية، إذ يعبّر عن مدى وعي النموذج بحدود معرفته قبل إصدار القرار. يُقسّم هذا المفهوم عادةً إلى نوعين رئيسيين: عدم اليقين العشوائي المرتبط بخصائص البيانات نفسها، وعدم اليقين المعرفي الناتج عن محدودية النموذج أو نقص التمثيل في بيانات التدريب.

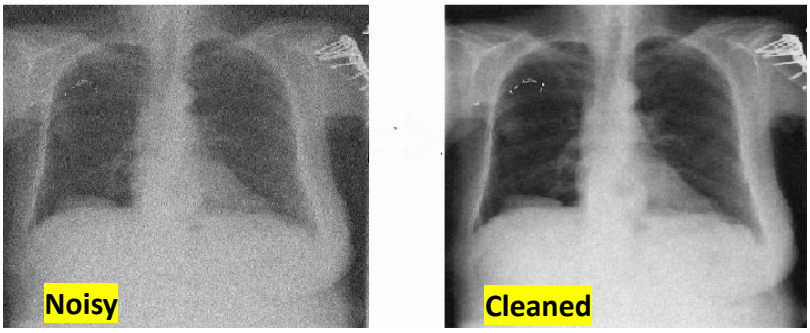
2.1. عدم اليقين العشوائي (Aleatoric Uncertainty)

يمثّل هذا النوع الغموض المرتبط بجودة البيانات وليس بإمكانات النموذج، إذ يظهر عندما تتأثر الصورة بتشويش بصري، أو ضجيج ناجم عن حركة المريض، أو ضعف في الإضاءة، أو انخفاض في التباين بين الأنسجة، إضافةً إلى الأخطاء البشرية في وضع التسميات. ويُعد هذا الغموض غير قابل للاختزال لأن مصدره فيزيائي وتجريبي، وبالتالي لا يمكن للنموذج تقليله حتى مع زيادة حجم البيانات أو تحسين قدراته [3].

تزداد احتمالية ظهور هذا النوع من الغموض في المهام التي تتسم بحدود غير واضحة أو بتفاصيل دقيقة يصعب تمييزها، مثل:

التباس حدود الأورام في صور الدماغ، انخفاض التباين في صور الصدر، وجود أدوات أو ظلال تشويشيه تعيق رؤية التفاصيل الحيوية.

يوضح الشكل (1) مثالًا على هذا النوع في صور الأشعّة الصدرية، حيث يؤدي الضجيج أو ضعف أجهزة التصوير إلى صور غير واضحة تقلّل دقة التنبؤ مهما كانت خبرة النموذج. وعلى الرغم من أن تقنيات تحسين الصور والمعايرة المسبقة قد تخفّف من هذا الأثر، فإنها لا تستطيع إزالته بالكامل [3].



الشكل (1): يوضح الشكل عدم اليقين العشوائي في صور الأشعّة الصدرية

2.2. عدم اليقين المعرفي (Epistemic Uncertainty)

يعكس هذا النوع حدود المعرفة المكتسبة لدى النموذج ذاته، أي الجهل بالعلاقات البنيوية داخل البيانات نتيجة محدودية التعلّم أو نقص التنوّع في العينات التدريبية [3].حيث يظهر عندما يواجه النموذج بيانات جديدة لم تُشاهد أثناء التدريب، أو ما يُعرف ببيانات خارج التوزيع (Out-of-Distribution – OOD)، حيث ترتفع درجة عدم اليقين رغم أن المخرجات قد تبدو واثقة رقميًا [3].

يمكن تقليل عدم اليقين المعرفي من خلال:

توسيع البيانات لتشمل عينات أكثر تنوعًا، خصوصًا الحالات النادرة [3]، وتحسين بنية النموذج لزيادة قدرته التمثيلية، بالإضافة الى استخدام نماذج التجميع (Ensembles) للتخفيف من انحياز نموذج واحد محدود الخبرة [5].

يمتلك هذا النوع أهمية خاصة في البيئات السريرية الحساسة، لأنه يشير إلى الحالات التي يفتقر فيها النظام الآلي للخبرة الكافية، مما يستدعي تدخل الطبيب. فعلى سبيل المثال، قد تُنتج شبكة مدرّبة على أورام شائعة ناتجًا "واثقًا" تجاه حالة نادرة رغم غياب المعرفة الحقيقية التي تبرّر هذا اليقين الظاهري، وهو ما يرفع خطر التنبؤات الخاطئة إذا لم يُؤخذ عدم اليقين بالحسبان [3].

تقترح بعض الدراسات تقسيم عدم اليقين المعرفي إلى فروع أكثر تحديدًا [6]، مثل:

- عدم اليقين البنيوي للنموذج الناتج عن حدود التصميم أو المعايير.
- عدم اليقين خارج التوزيع (OOD) الناتج عن مواجهة بيانات تختلف جذريًا عن بيانات التدريب.

يمهّد فهم عدم اليقين للانتقال إلى منهجيات تقديره، والتي تُعدّ أساسًا لاختيار أفضل الأساليب في نماذج التعلّم العميق الطبية، كما سيُعرض في القسم التالي.

3. منهجيات تقدير عدم اليقين (Uncertainty Estimation Methods)

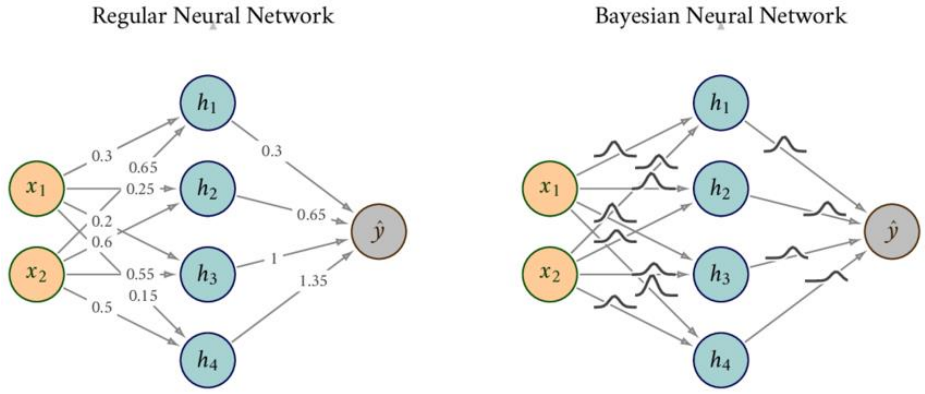
تهدف منهجيات تقدير عدم اليقين إلى تمكين نماذج التعلّم العميق من التعبير عن مستوى ثقتها في القرارات التشخيصية، وتتوزع هذه المنهجيات بين عدة أساليب، نستعرض أبرزها:

3.1. المنهجيات البايزية (Bayesian Methods)

تمثّل المنهجية البايزية أحد الأسس النظرية الأكثر رسوخًا في تقدير عدم اليقين، إذ تفترض أن أوزان الشبكة ليست قيمًا ثابتة وإنما توزيعات احتمالية تعبّر عن المعرفة المكتسبة وحدودها. غير أن تطبيق الشبكات العصبية البايزية الكاملة عبر خوارزميات مثل Variational Inference (VI) يُعدّ مرتفع الكلفة حسابيًا، مما يحدّ من قابليته للتطبيق في

البيئات السريرية ذات المتطلبات الزمنية الصارمة؛ لذلك شكّل أسلوب Monte Carlo Dropout (MC-Dropout) بديلًا عمليًا لهذه النماذج [6]، إذ يعتمد على إبقاء طبقات Dropout مفعّلة أثناء الاختبار وتمرير الصورة نفسها عبر النموذج عدة مرات، ثم حساب التباين بين المخرجات بوصفه مؤشرًا مباشرًا على مستوى عدم اليقين.

تمتاز هذه المنهجية بخفّتها وسهولة دمجها ضمن أي شبكة عصبية دون الحاجة إلى تعديل معماري، مما جعلها من أكثر الأساليب انتشارًا في تطبيقات التصوير الطبي، يوضح الشكل (2) الفرق بين الشبكات التقليدية التي تتعامل مع الأوزان كقيم ثابتة وبين الشبكات البايزية التي تمثل الأوزان كتوزيعات احتمالية قادرة على التعبير عن مستوى الثقة في المخرجات.



الشكل (2): يوضح الشكل الفرق بين الشبكة العصبية التقليدية والشبكة العصبية البايزية

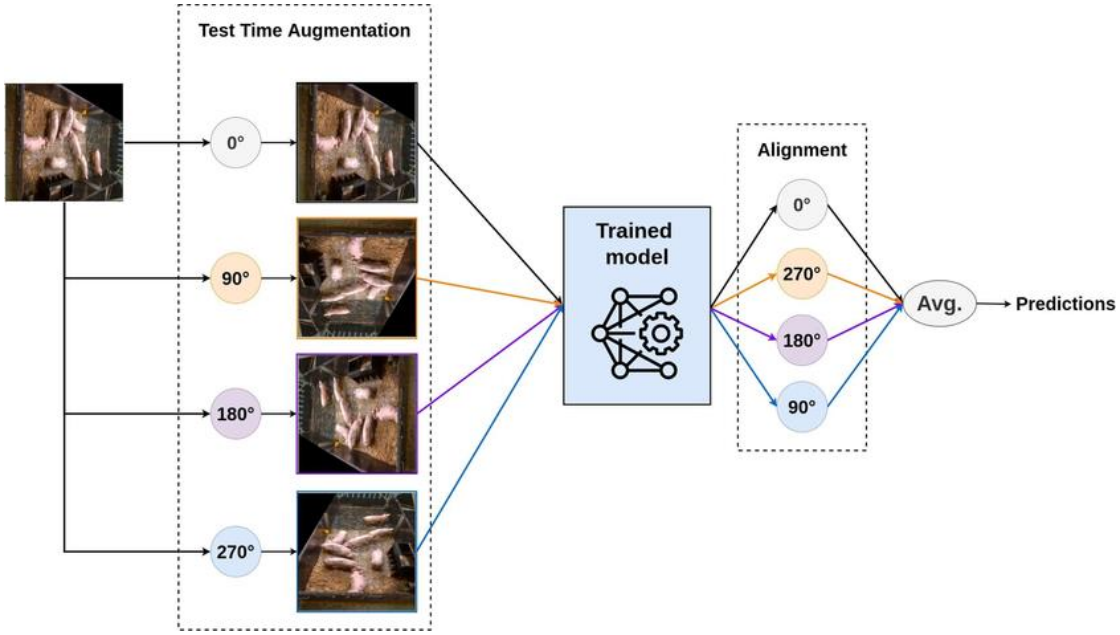
3.2. النماذج المجمّعة (Ensemble Models)

تعتمد النماذج المجمّعة على مبدأ إحصائي بسيط وهو أن اتفاق مجموعة من النماذج يعزّز الثقة بالقرار، بينما يشير اختلافها إلى درجة من الغموض في التنبؤ، ويقوم هذا الأسلوب على تدريب عدة نماذج متشابهة من حيث البنية لكنها تختلف في التهيئة الأولية أو في توزيع البيانات أثناء التدريب، ثم دمج مخرجاتها باستخدام متوسط التنبؤات أو التوزيع الاحتمالي لها، ويُعد تباين النتائج بين هذه النماذج مؤشرًا مباشرًا على مستوى عدم اليقين، إذ يعكس مدى اختلاف "وجهات النظر" داخل المجموعة. يُعد هذا الأسلوب من المنهجيات الشائعة الاستخدام في التطبيقات الطبية، إذ أثبت قدرته على تحسين الاستقرار وتقليل الأخطاء مقارنةً بالاعتماد على نموذج منفرد، لكن تواجه هذه المنهجية قيدًا رئيسيًا يتمثل في التكلفة الحسابية الكبيرة الناتجة عن تدريب وتشغيل عدة نماذج مستقلة، مما يحدّ من إمكانية استخدامه في الأنظمة الطبية ذات الموارد المحدودة أو في البيئات السريرية الحساسة التي تتطلب أزمنة استجابة قصيرة [5].

3.3. الاختبار بتعزيز البيانات (Test-Time Augmentation – TTA)

تعتمد منهجية الاختبار بتعزيز البيانات على قياس استقرار النموذج عند تطبيق تعديلات طفيفة على الصورة أثناء مرحلة الاختبار. بدلًا من الاعتماد على نسخة واحدة فقط، تُولّد عدة نسخ معدّلة من الصورة الأصلية – عبر تدوير بسيط أو تغيير في الإضاءة أو المقياس – ثم تُمرّر جميعها عبر النموذج، ويُقاس مدى التوافق بين المخرجات بوصفه مؤشرًا على مستوى الثقة. فإذا كانت القرارات متقاربة دلّ ذلك على ثبات النموذج، في حين يشير تباين المخرجات إلى ارتفاع عدم اليقين [7].

تتميز هذه المنهجية ببساطتها، ولكن تبقى الزيادة في زمن الاستجابة أحد أبرز قيودها نتيجة الحاجة إلى تمرير الصورة عدة مرات للحصول على توزيع كافٍ من المخرجات. كما موضح في الشكل (3)



الشكل (3): توضيح مبدأ اختبار تعزيز البيانات في تقدير عدم اليقين

3.4. التعلّم العميق الإثباتي (Evidential Deep Learning – EDL)

يمثل التعلّم العميق الإثباتي أحد الأساليب الحديثة لتقدير عدم اليقين، إذ يهدف إلى جعل النموذج يوفّر مع كل تنبؤ "دليلًا كمّيًا" يعبّر عن قوة الدعم الإحصائي للقرار بدل الاكتفاء بإعطاء قيمة احتمالية واحدة [8]. ويرتكز هذا النهج على دمج مبادئ نظرية دمبستر–شيفر (Dempster–Shafer Theory) للأدلة ضمن مخرجات الشبكة، بحيث تُنتج معاملات دليليه توضّح مقدار الثقة بكل فئة محتملة، يتميز EDL بكونه أحادي المرور، أي أنه يقدّم كلًا من التنبؤ ومقدار عدم اليقين في خطوة واحدة، مما يمنحه كفاءة حسابية مرتفعة وسهولة دمجه ضمن النماذج التقليدية. يحتاج هذا الأسلوب إلى معايرة دقيقة لدالة الخسارة لضمان التوازن بين قوة الدليل ودقة التنبؤ، إذ قد تؤدي المعايرة الضعيفة إلى ثقة مبالغ فيها في الحالات الغامضة أو في بيانات خارج التوزيع (OOD) [9]، مما يجعله خيارًا مناسبًا للأنظمة السريرية التي تسعى إلى تحسين الثقة بالمخرجات في الزمن الحقيقي.

3.5. أساليب المعايرة (Calibration Methods)

تهدف أساليب المعايرة إلى جعل احتمالات النموذج تعبّر بدقة عن صحتها الواقعية، بحيث تتوافق "الثقة الرقمية" مع الأداء الفعلي؛ فالنموذج الذي يصرّح بثقة 80% يجب أن يكون صحيحًا في ثمانية من كل عشرة حالات. يُعد التحجيم الحراري (Temperature Scaling) من أكثر أساليب المعايرة شيوعًا، إذ يضيف معامل حرارة إلى مخرجات دالة Softmax لضبط اتساق الثقة مع الدقّة دون الحاجة إلى إعادة تدريب النموذج. لتقييم جودة المعايرة، تُستخدم مؤشرات مثل Expected Calibration Error (ECE) وBrier Score، وتعمل هذه المقاييس على تحديد الفارق بين الثقة المعلَنة والدقّة الحقيقية، مما يوفر معيارًا موحّدًا للحكم على موثوقية النماذج وخاصة الطبية قبل دمجها في مهام التعلّم النشط التي تعتمد بشكل مباشر على قياسات الثقة في اختيار العينات [10].

4. التعلّم النشط:

يمثّل التعلّم النشط أحد الأساليب المحورية لمواجهة ندرة البيانات الطبية وارتفاع تكلفة عملية وضع التسميات السريرية للصورة (Labeling)، وهي مهمة يتولّاها الخبير الطبي لإنتاج الحقيقة الأرضية التي يعتمد عليها النموذج أثناء التعلّم. تعتمد منهجية التعلم النشط على وظائف اختيار تستند إلى مقاييس عدم اليقين بهدف تحديد العينات التي يجهلها النموذج أو يتردد في تصنيفها، مما يسمح بعرض أكثر الحالات أهمية على الخبير لإضافة التسميات اللازمة [4]، تستخدم هذه الوظائف مؤشرات متعددة مثل التباين التنبؤي (Predictive Variance) أو الإنتروبيا التنبؤية

(Predictive Entropy). تعتبر هذه المنهجية مناسبة للتصوير الطبي الذي يتسم بتفاوت كبير في جودة الصور ودرجة غموضها. ويكتسب أهمية خاصة عند دمجه مع تقدير عدم اليقين، إذ يتيح للنموذج التركيز على المناطق الغامضة في فضاء البيانات، سواء كان هذا الغموض عشوائيًا ناتجًا عن ضوضاء الصورة أو معرفيًا مرتبطًا بندرة الحالات. ويسهم ذلك في بناء نماذج أكثر قدرة على التعميم باستخدام بيانات أقل حجمًا ولكن أعلى قيمة من الناحية السريرية.

سنعرض في القسم التالي تطبيقات عملية توضّح كيفية تفاعل عدم اليقين والتعلّم النشط في مرضين صدريين يختلفان جذريًا من حيث طبيعة الصور الشعاعية: COVID-19 والالتهاب الرئوي.

5. تطبيقات عدم اليقين والتعلم النشط في المجال الطبي:

5.1. تشخيص COVID-19 باستخدام عدم اليقين والتعلّم النشط

يتّسم مرض COVID-19 بتنوّع واسع في الأنماط الشعاعية ضمن صور الأشعة السينية (X-ray) والتصوير المقطعي المحوسب، حيث تظهر الآفات عادةً على شكل مناطق ضبابية منخفضة التباين تتداخل مع أعراض أمراض صدرية أخرى مثل الالتهاب الرئوي والوذمات الرئوية. ويسهم هذا التداخل في إضعاف قدرة النماذج التقليدية على استخراج خصائص تمييزية واضحة، مما يرفع مستوى عدم اليقين العشوائي الناتج عن غموض حدود الإصابة أو ضعف جودة الصورة. كما أن معظم قواعد البيانات التي استُخدمت في تطوير نماذج COVID-19 خلال المراحل الأولى من الجائحة كانت صغيرة وغير متوازنة، مما زاد من أثر عدم اليقين المعرفي الناتج عن محدودية الخبرة المكتسبة أثناء التدريب [11,3]. وبالتالي، يُعد COVID-19 حالة نموذجية يظهر فيها النوعان معًا: غموض بصري بالإضافة الى نقص التمثيل.

5.1.1. دور منهجيات عدم اليقين في تحسين ثقة النموذج:

أظهر MC-Dropout قدرة واضحة على كشف الحالات المشكوك بها من خلال تحليل تباين التنبؤات المتكررة، خصوصًا في الآفات الدقيقة أو الصور منخفضة الجودة [6]. كما توفر نماذج التجميع العميق Deep Ensembles رؤية أشمل لمجال عدم اليقين من خلال دمج قراءات متعددة لنماذج ذات تهيئات مختلفة، مما يسمح برصد الحالات التي تختلف فيها النماذج بشكل جوهري، وهو مؤشر مهم لوجود بيانات خارج التوزيع أو حالات يصعب تصنيفها حتى على الخبراء [5]، كما يساعد الاختبار بتعزيز البيانات TTA في اختبار ثبات النموذج تحت تغيّرات طفيفة في زاوية التصوير أو شدة الإضاءة، وهي عوامل شائعة في البيئات السريرية اليومية [7]. تتيح هذه القياسات فهم مدى حساسية النموذج لمثل هذه التغيّرات وتحديد الحالات التي لا يمتلك فيها النظام ثقة كافية ليُعتمد عليه بشكل مستقل.

5.1.2. التعلّم النشط في تحسين النموذج:

أظهر التعلّم النشط نتائج لافتة في تشخيص COVID-19، إذ يسمح للنموذج بانتقاء العينات الأكثر غموضًا وطلب تسميتها من الطبيب، حيث تستهدف هذه الاستراتيجية مباشرةً مناطق عدم اليقين المعرفي، مما يتيح توسيع خبرة النموذج باستخدام أقل كمية ممكنة من البيانات المعلّمة [3,4].

وبالتالي عند دمج تقنيات عدم اليقين لتحديد أين يخطئ النموذج وتقنية التعلم النشط لتحديد من أين سيتعلم النموذج، أظهرت دراسات مختلفة انها أدت إلى تقليل حجم البيانات المطلوبة لتحسين الأداء بنسبة تتراوح بين 20% و40%، مع تحسين واضح في قدرة النموذج على التمييز بين العدوى الفيروسية والحالات المشابهة [12].

5.2. تشخيص الالتهاب الرئوي باستخدام عدم اليقين والتعلّم النشط

يمتلك الالتهاب الرئوي خصائص شعاعية أكثر وضوحًا مقارنةً بـ COVID-19 ضمن صور الأشعّة السينية (X-ray)، إذ تظهر الآفات غالبًا على شكل تكثّفات أو مناطق معتمة ذات حدود واضحة نسبيًا. ويسهم ذلك في توفير

تباين بصـــري أعلى يســـهل على النماذج اســـتخراج الأنماط التمييزية، ويؤدي بالتالي إلى مســـتوى أقل من عدم اليقين العشوائي شريطة أن تكون جودة الصورة مناسبة [3]. على الرغم من وضوح المعالم الشعاعية، يبقى التحدّي الأساسي في تمييز نوع الالتهاب (فيروســـي مقابل بكتيري)، وهو ما يرفع مســـتوى عدم اليقين المعرفي المرتبط بتمثيل العينات وتوازن الفئات [13].

5.2.1. دور منهجيات عدم اليقين (UQ) في تحسين ثقة النموذج:

أظهرت منهجيات تقدير عدم اليقين أداءً ملحوظًا في مهام تشـــخيص الالتهاب الرئوي، لكن وظيفتها هنا تختلف عن دورها في COVID-19. ففي حين يُستخدم عدم اليقين في COVID-19 لمواجهة الضـبابية البصـرية، يُستخدم في الالتهاب الرئوي للتحقق من موثوقية النموذج عند الحالات المتداخلة أو غير النموذجية. قدّم MC-Dropout تقديرات مستقرة لمستوى التشتت في التنبؤات، خصوصًا في الصور منخفضة الجودة أو في الأنواع الفرعية النادرة من الالتهاب [6]. كما أثبت Deep Ensembles فاعلية عالية في تقليل الأخطاء الناتجة عن الالتباس بين الالتهاب الفيروســـي والبكتيري عبر دمج وجهات نظر متعددة ورصـد الحالات التي تختلف فيها النماذج بوضـوح [5]، أما TTA فكان مفيدًا في تقييم ثبات النموذج عند تغيّر عوامل الإضاءة أو اختلاف أجهزة التصوير، مما يساعد على قياس حساسية النموذج لهذه التغيرات وتعزيز موثوقية النتائج [7].

5.2.2. التعلّم النشط في تحسين النموذج:

في حالة الالتهاب الرئوي، يركّز التعلّم النشـط على العينات التي تُظهر تداخلًا بصـريًا أو عدم وضـوح في الحدود بين الأنواع المختلفة للالتهاب. وقد أثبت أســـلوب Uncertainty Sampling قدرته على انتقاء مثل هذه الحالات تلقائيًا وإرســـالها إلى الطبيب لتســـميتها، مما يعزز خبرة النموذج في الأنماط النادرة أو قليلة الظهور، وســـاهم هذا النهج في تحسين الأداء تدريجيًا دون الحاجة إلى توسيم آلاف الصور دفعة واحدة، إذ ركّز النموذج على العينات الأكثر قيمة بدلًا من الاعتماد على الحجم الكلّي للبيانات. وتبرز أهمية هذه الميزة في المســتشــفيات ذات الموارد المحدودة أو في قواعد البيانات الصغيرة [4].

يتضــح مما ســبق أن طبيعة عدم اليقين وآليات الاســتفادة من التعلم النشــط تختلف بين تشــخيص الالتهاب الرئوي و COVID-19 تبعاً لخصـائص البيانات الشـعاعية. ففي حالة COVID-19 يتركز دور تقدير عدم اليقين في التعامل مع الضــبابية البصــرية وعدم توازن البيانات الشــعاعية، بينما يبرز في الالتهاب الرئوي في دعم التمييز بين الحالات المتداخلة أو غير النموذجية، كما تشـير الدراسـات المسـتعرضـة إلى أن التعلم النشـط يسـهم في تحسـين كفاءة اختيار العينات في كلا التطبيقين ، مع اختلاف حجم تأثيره وفق طبيعة البيانات المتاحة . ويلخّص الجدول (1) أبرز الفروقات بين الحالتين من حيث طبيعة البيانات، وأنواع عدم اليقين، ومنهجيات التقدير، ودقة التصــنيف، إضــافةً إلى أثر التعلّم النشط في تقليل الحاجة إلى البيانات الموسومة.

معيار المقارنة	COVID-19	الالتهاب الرئوي
طبيعة البيانات الشعاعية	آفات ضبابية ومتفاوتة مع عدم توازن في البيانات	تكثفات رئوية مع توازن أفضل للبيانات
نوع عدم اليقين السائد	عدم يقين عشوائي ومعرفي	عدم يقين معرفي بصورة أساسية، مع وجود عدم يقين عشوائي بدرجة أقل
السبب الرئيس لارتفاع عدم اليقين	ضبابية الآفات، تداخل المظاهر الشعاعية، وعدم توازن البيانات	تشابه الأنواع الفرعية للالتهاب، وصعوبة التمييز بين الالتهاب الفيروسي والبكتيري

منهجيات تقدير عدم اليقين المستخدمة	MC-Dropout, Deep Ensembles, TTA	MC-Dropout, Deep Ensembles, TTA
دقة النموذج الأساسية	93.4% [11]	91.7% [11]
دقة النموذج المعتمد على تقدير عدم اليقين	92.6% [11]	83.5% [11]
الدور الرئيس لتقدير عدم اليقين	التعامل مع الضبابية البصرية وتحديد الحالات منخفضة الثقة	تحسين موثوقية التمييز بين الحالات المتداخلة وغير النموذجية
أثر التعلم النشط	خفض الحاجة إلى البيانات الموسومة مع تحسين كفاءة التعلم من خلال اختيار العينات الأكثر إفادة [4]	تحسن اختيار العينات الأكثر إفادة وتعزيز تعلم الأنماط النادرة، دون وجود نسبة رقمية موحدة في الدراسات [13]

الجدول (1): مقارنة بين تطبيقات تقدير عدم اليقين (UQ) والتعلّم النشط (Active Learning) في تشخيص COVID-19 والالتهاب الرئوي.

ملاحظة: تستند القيم الرقمية الواردة في هذا الجدول مستمدة من دراسات مختلفة وقد أدرجت لأغراض المقارنة المرجعية والتوضيح، ولا تمثل مقارنة مباشرة ضمن تجربة تجريبية موحدة

6. الخاتمة

يُظهر هذا البحث أن دمج منهجيات تقدير عدم اليقين مع التعلّم النشط يشكّل خطوة محورية نحو بناء نماذج طبية أكثر موثوقية وقدرة على التعميم. فمن خلال فهم مصادر عدم اليقين العشوائي والمعرفي، وتوظيف أساليب مثل MC-Dropout و Deep Ensembles و TTA و EDL، يصبح من الممكن قياس الثقة في القرارات التشخيصية بصورة أدق، وتحديد الحالات التي تحتاج تدخّلًا بشريًا أو إعادة فحص. وفي الوقت ذاته، يتيح التعلّم النشط تحسين جودة البيانات المعلّمة عبر التركيز على العينات الأكثر قيمة، مما يقلل التكلفة السريرية ويرفع كفاءة النماذج حتى في البيئات ذات الموارد المحدودة. وتُبيّن دراستا COVID-19 والالتهاب الرئوي أن تأثير هذه المنهجيات يعتمد بقوة على طبيعة البيانات وجودتها، حيث يبرز دور عدم اليقين بشكل خاص في الأمراض الغامضة أو ذات قواعد بيانات صغيرة، بينما يتقدّم التعلّم النشط في الحالات التي تتطلب تعزيز تمثيل الفئات أو الأنماط النادرة. ويمهّد هذا التكامل بين المفهومين لمرحلة جديدة من النظم الذكية القادرة على تقديم قرارات أكثر وعيًا وثقة، مما يجعلها أقرب للاستخدام السريري الفعلي ويعزز دور الذكاء الاصطناعي كشريك داعم للطبيب، لا بديلًا عنه.

وعلى الرغم من المزايا التي توفرها منهجيات مثل Deep Ensembles و TTA في تحسين موثوقية التنبؤات وتعزيز ثقة النماذج الطبية، فإن ارتفاع كلفتها الزمنية والحسابية قد يحدّ من استخدامها في البيئات السريرية ذات الموارد المحدودة. ويمكن الحد من هذه القيود من خلال استخدام نماذج أكثر كفاءة، أو تطبيق هذه المنهجيات على الحالات التي تُظهر مستويات مرتفعة من عدم اليقين فقط، أو الاستفادة من البنى الحاسوبية المتقدمة عند توفرها، بما يحقق توازنًا بين موثوقية التشخيص وكفاءة التنفيذ العملي. وتشير الدراسات المرجعية إلى أن التكامل بين تقدير عدم اليقين والتعلم النشط يمثل اتجاهاً بحثياً واعداً نحو تطوير أنظمة ذكاء اصطناعي طبي أكثر موثوقية، كما يفتح المجال أمام أبحاث مستقبلية تركز على ابتكار أساليب أكثر كفاءة وأقل تكلفة بما يسهم في تسريع تبني هذه التقنيات ضمن الممارسات الطبية الواقعية.

7. المراجع:

[1] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, "Medical Image Analysis Using Deep Learning Algorithms", Frontiers in Public Health, vol. 11, no. 1273253, 2023.

[2] K. Borys, Y. A. Schmitt, M. Nauta, C. Seifert, N. Krämer, C. M. Friedrich, and F. Nensa", Explainable AI in Medical Imaging: An Overview for Clinical Practitioners–Saliency-Based XAI Approaches," European Journal of Radiology, vol. 162, no. 110787, 2023.

[3] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi," A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges" ,Information Fusion, vol. 76, pp. 243–297, 2021.

[4] B. Settles, "Active Learning Literature Survey", Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2010.

[5] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles", in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.

[6] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning", in Proceedings of the 33rd International Conference on Machine Learning (ICML), pp. 1050–1059, 2016.

[7] A. Chen, Y. Li, W. Qian, K. Morse, C. Miao, and M. Huai, "Modeling and Understanding Uncertainty in Medical Image Classification", in Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Computer Science, vol. 15010, pp. 557–567, 2024.

[8] M. Sensoy, L. Kaplan, and M. Kandemir, "Evidential Deep Learning to Quantify Classification Uncertainty", Advances in Neural Information Processing Systems (NeurIPS), vol. 31, 2018.

[9] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?", Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 5574–5584, 2017.

[10] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On Calibration of Modern Neural Networks", in Proceedings of the 34th International Conference on Machine Learning (ICML), pp. 1321–1330, 2017.

[11] A. Whata, K. Dibeco, K. Madzima, and I. Obagbuwa, "Uncertainty Quantification in Multi-Class Image Classification Using Chest X-Ray Images of COVID-19 and Pneumonia", Frontiers in Artificial Intelligence, vol. 7, Art. no. 1410841, 2024.

[12] R. Mehta, C. Shui, and T. Arbel, "Evaluating the Fairness of Deep Learning Uncertainty Estimates in Medical Image Analysis", Proceedings of Machine Learning Research (PMLR), vol. 227, pp. 1453–1492, 2023.

[13] B. Lambert, F. Forbes, A. Tucholka, S. Doyle, H. Dehaene, and M. Dojat, "Trustworthy Clinical AI Solutions: A Unified Review of Uncertainty Quantification in Deep Learning Models for Medical Image Analysis", Medical Image Analysis, vol. 89, no. 102921, 2024.