DETECTION OF PLANT DISEASES BY USING MACHINE LEARNING

الكشف عن أمراض النباتات عن طريق التعلم الآلي

Marwa Mohammed Obaid
MSc of Electrical Engineering, Control and Computers at the department of Computer Science, Dijlah University, Iraq

أ. مروه محمد عبيد
معيدة في قسم علوم الحاسوب
بجامعة دجلة الجامعة ، العراق

Majid Salal Naghmash
PhD, Ass. Prof. of Electrical and Electronic at the department of Computer Science, Dijlah University, Iraq

د. ماجد صلال نغمش
أستاذ مساعد في قسم علوم الحاسوب
بجامعة دجلة الجامعة ،العراق

Scientific Supervisor — المشرف العلمي

Abstract: *Every year, millions of dollars are lost due to plant diseases. these diseases can lead to reduced yields and sometimes complete crop failure, the agricultural landscape is facing serious challenges. Plant diseases are on the rise, threatening global food security. Traditional methods are often slow and costly, leaving farmers vulnerable to crop losses. therefor Farmers need efficient ways to identify and manage these issues before they escalate. However, advances in artificial intelligence (AI) provide a fresh approach to address these critical challenges.*

Keywords: *plant disease; machine learning, Artificial intelligence, Pattern recognition, convolutional neural networks (CNNs).*

الخلاصة : كل عام، يتم فقدان ملايين الدولارات بسبب أمراض النباتات. يمكن أن تؤدي هذه الأمراض إلى انخفاض الغلة وأحيانًا فشل المحاصيل تمامًا، ويواجه المشهد الزراعي تحديات خطيرة. تتزايد أمراض النباتات، مما يهدد الأمن الغذائي العالمي. غالبًا ما تكون الطرق التقليدية بطيئة ومكلفة، مما يجعل المزارعين عرضة لخسائر المحاصيل. لذلك يحتاج المزارعون إلى طرق فعالة لتحديد هذه المشكلات وإدارتها قبل تفاقمها. ومع ذلك، توفر التطورات في مجال الذكاء الاصطناعي نهجًا جديدًا لمعالجة هذه التحديات الحرجة.

الكلمات المفتاحية : أمراض النبات؛ التعلم الآلي، الذكاء الاصطناعي، التعرف على الأنماط، الشبكات العصبية التلافيفية (CNNs)

Introduction

In this report used basic photos of the leaves of both healthy and diseased plants to train and convolutional neural networks (CNNs) architectures in order to create an automated method for detecting and diagnosing plant diseases. The one that the supplied dataset included photos taken under actual field growing circumstances as well as experimental (laboratory) setups. Compared to shallow techniques, which learn from fewer data but are crop-specific, the suggested deep learning approach could discover more accuracy results [1, 2].

المقدمة

في هذا التقرير، تم استخدام صور أساسية لأوراق النباتات الصحية والمريضة للتدريب وهندسة الشبكات العصبية التلافيفية (CNNs) من أجل إنشاء طريقة آلية للكشف عن أمراض النباتات وتشخيصها. تلك التي تضمنتها مجموعة البيانات المقدمة، والتي تضمنت صورًا تم التقاطها في ظل ظروف زراعة الحقل الفعلية بالإضافة إلى الإعدادات التجريبية (المختبرية). وبالمقارنة بالتقنيات الضحلة، والتي تتعلم من بيانات أقل ولكنها خاصة بالمحاصيل، فإن نهج التعلم العميق المقترح يمكن أن يكتشف نتائج أكثر دقة [1، 2].

Results and discussion

Training of the models was performed with the use of an open database more than 80,000 images, containing 20 different plants in a set of 54 distinct classes of [plant, disease] combinations, including healthy plants. Several model architectures were trained, with the best performance reaching a 98.50% success rate in identifying the corresponding [plant, disease] combination (or healthy plant). The significantly high success rate makes the model a very useful advisory or early warning tool, and an approach that could be further expanded to support an integrated plant disease identification system to operate in real cultivation conditions, even though the developed system has a really high success rate, it is still far from being a general tool that may be applied in real-world situations for a number of reasons. The incorporation of 54 distinct to the best of our knowledge, the largest plant disease detection assignment currently being addressed with deep learning approaches consists of classes of [plant, illness] pairings of 20 distinct plant species. The next near future step, though, should be to expand the current database to include a greater range of plant species and illnesses. This can be a time- consuming and difficult procedure in a number of ways. Another crucial point that needs to be mentioned and to be addressed is that the testing dataset, which was utilized to evaluate the models, was a component of the training dataset. This is a standard procedure in machine learning models, but testing data that would come from real-world scenarios should demonstrate the true worth of the created system, particularly in terms of its applicability. distinct databases and/or sources. Some early experiments in that direction, using a small but sufficient amount of data, revealed a significant decrease in model performance, ranging from 25 – 30% percent, depending on the data source. This is comparable to comparable performances that have been reported, which showed a model accuracy of 30 % percent in such data for a problem involving 35 plant disease classes. A far greater range of training data from several sources covering various regions, cultivation circumstances, and image capture modes and sets to be gathered in order to make progress in that direction. A greater range of training data from several sources covering various regions, cultivation circumstances, and picture capture modes and sets to be gathered in order to make progress in that direction. Given the enormous potential of the suggested deep learning technique, it is a matter of quantity and quality of the data that is now accessible to enhance the system, make it more robust in actual cultivation circumstances, and expand its scope (in terms of plant species and illnesses that may be diagnosed).

النتائج والمناقشة

تم تدريب النماذج باستخدام قاعدة بيانات مفتوحة تحتوي على أكثر من 80000 صورة، تحتوي على 20 نباتًا مختلفًا في مجموعة من 54 فئة مميزة من مجموعات [النبات، المرض]، بما في ذلك النباتات الصحية. تم تدريب العديد من بنيات النموذج، حيث بلغ أفضل أداء معدل نجاح 98.50٪ في تحديد مجموعة [النبات، المرض] المقابلة (أو النبات الصحي). يجعل معدل النجاح المرتفع بشكل ملحوظ النموذج أداة استشارية أو تحذير مبكر مفيدة للغاية، ونهجًا يمكن توسيعه بشكل أكبر لدعم نظام متكامل لتحديد أمراض النبات للعمل في ظروف الزراعة الحقيقية، على الرغم من أن النظام المتطور يتمتع بمعدل نجاح مرتفع حقًا، إلا أنه لا يزال بعيدًا عن كونه أداة عامة يمكن تطبيقها في مواقف العالم الحقيقي لعدد من الأسباب. إن دمج 54 حالة مميزة على حد علمنا، أكبر مهمة للكشف عن أمراض النبات يتم التعامل معها حاليًا باستخدام أساليب التعلم العميق تتكون من فئات من أزواج [النبات، المرض] من 20 نوعًا نباتيًا مميزًا. ومع ذلك، يجب أن تكون الخطوة التالية في المستقبل القريب هي توسيع قاعدة البيانات الحالية لتشمل مجموعة أكبر من أنواع النباتات والأمراض. قد يكون هذا إجراءً يستغرق وقتًا طويلاً ويصعب القيام به بعدة طرق. هناك نقطة حاسمة أخرى يجب ذكرها ومعالجتها وهي أن مجموعة بيانات الاختبار، التي تم استخدامها لتقييم النماذج، كانت أحد مكونات مجموعة بيانات التدريب. هذا إجراء قياسي في نماذج التعلم الآلي، ولكن بيانات الاختبار التي تأتي من سيناريوهات العالم الحقيقي يجب أن توضح القيمة الحقيقية للنظام الذي تم إنشاؤه، خاصة من حيث قابليته للتطبيق. قواعد بيانات و/أو مصادر مميزة. كشفت بعض التجارب المبكرة في هذا الاتجاه، باستخدام كمية صغيرة ولكن كافية من البيانات، عن انخفاض كبير في أداء النموذج، يتراوح من 25 إلى 30٪، اعتمادًا على مصدر البيانات. هذا قابل للمقارنة بالأداء المماثل الذي تم الإبلاغ عنه، والذي أظهر دقة نموذج بنسبة 30٪ في مثل هذه البيانات لمشكلة تتضمن 35 فئة من أمراض النبات. مجموعة أكبر بكثير من بيانات التدريب من عدة مصادر تغطي مناطق مختلفة وظروف الزراعة وأوضاع التقاط الصور والمجموعات التي يجب جمعها من أجل إحراز تقدم في هذا الاتجاه. إن مجموعة أكبر من بيانات التدريب من مصادر متعددة تغطي مناطق مختلفة وظروف الزراعة وأنماط التقاط الصور ومجموعات الصور التي يجب جمعها من أجل إحراز تقدم في هذا الاتجاه. ونظراً للإمكانات الهائلة لتقنية التعلم العميق المقترحة، فإن الأمر يتعلق بكمية ونوعية البيانات التي يمكن الوصول إليها الآن لتعزيز النظام وجعله أكثر قوة في ظروف الزراعة الفعلية وتوسيع نطاقه (من حيث أنواع النباتات والأمراض التي يمكن تشخيصها).

Conclusion

This study specialized customized deep learning models based on certain convolutional neural network designs to identify plant illnesses from straightforward photos of healthy leaves. or ill plants [3]. a publicly accessible collection of 80,000 photos that were collected in actual agriculture fields as well as in lab settings was used to train the models. The information includes some healthy plants as well as 20 plant species in 54 different classes of [plant, illness] pairings.

الخاتمة

تخصصت هذه الدراسة في نماذج التعلم العميق المخصصة بناءً على تصميمات معينة للشبكات العصبية التلافيفية لتحديد أمراض النباتات من صور بسيطة لأوراق صحية أو نباتات مريضة [3]. تم استخدام مجموعة متاحة للجمهور من 80.000 صورة تم جمعها في حقول زراعية فعلية وكذلك في إعدادات المختبر لتدريب النماذج. تتضمن المعلومات بعض النباتات الصحية بالإضافة إلى 20 نوعًا من النباتات في 54 فئة مختلفة من أزواج [النبات، المرض].

Literature المراجع والمصادر

1. H. Orchi, M. Sadik, and M. Khaldoun, “On using artificial intelligence and the internet of things for crop disease detection: A contemporary survey,” Agric., vol. 12, no. 1, 2022, doi: 10.3390/agriculture12010009.
2. M. A. Nawaz et al., “Plant disease detection using internet of thing (IoT),” Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 1, pp. 505–509, 2020, doi: 10.14569/ijacsa.2020.0110162.
3. K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Comput. Electron. Agric., vol. 145, no. January, pp. 311–318, 2018, doi: 10.1016/j.compag.2018.01.009.