



وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

تشخيص أورام الأنسجة الرخوة باستخدام تقنيات ذكائية

رسالة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
البرمجيات

من قبل

نور يحيى غانم شاكر

بإشراف

أ.م.د. جمال صلاح الدين مجيد سيد حمو

المستخلص

إن النمو الهائل للذكاء الاصطناعي (Artificial Intelligent (AI)) على مدى السنوات العشر الماضية هو الرأي المساند لاتخاذ قرارات دقيقة لتشخيص نوع الأورام من خلال بناء برامج ذكية تعتمد على أحدث تقنيات الذكاء الاصطناعي كالتعلم الآلي (Machine Learning (ML)) والتعلم العميق (Deep Learning (DL)) وخوارزميات الأسراب (Swarm Algorithm) والتهجين فيما بينهم للوصول الى أعلى دقة في التشخيص، وفي هذه الرسالة تم اعتماد صور التشريح المرضي (Histopathology) للحالات المرضية وهي (اورام الثدي واورام الرئة واورام القولون) وسبب اختيار هذه الأنواع من الأورام كونها منتشرة انتشاراً واسعاً في جميع أنحاء العالم وهي أكثر الأمراض فتكاً. تم الحصول على البيانات من خلال جمع Datasets (الصور النسيجية) المعتمدة من شركة Kaggle والبالغ عددها 32783 صورة نسيجية بما في ذلك (7783 صورة لأورام الثدي الحميد والخبيث، و10.000 صورة لأورام القولون الحميد والخبيث، و15.000 صورة لأورام الرئة (الحميدة، الخبيثة والحرفشية))، على وفق مستويات تكبير مختلفة لصور التشريح المرضي للانسجة، وبعد دمج مجموعات البيانات بمجموعة بيانات واحدة تم تقسيمها إلى (80% من بيانات التدريب والتحقق من الصحة و 20% من بيانات الاختبار) وتحليلها بإجراء المعالجة السابقة عليها ثم تصميم وتنفيذ ال 13 نموذجاً ذكياً هجيناً باستخدام تقنيات التعلم العميق والتعلم الآلي وخوارزمية الأسراب لتشخيص الأورام، والوصول الى أعلى دقة في التشخيص الطبي ففي التعلم العميق تم بناء الشبكة العصبية التلافيفية (Convolution Neural Network (CNN)) واستخدام خوارزميات (ResNet50, ResNet101,) و (DenseNet21, DenseNet,169, DenseNet201) (ResNet152 المدربة مسبقاً (Pre-trained) بالاعتماد على تقنية نقل التعلم (Transfer Learning) التي تتيح تكييف النماذج لأداء المهام الجديدة، عن طريق بناء بنية جديدة تتناسب مع مجموعة البيانات المحددة في هذه الرسالة ، وتهجينها مع احدى تقنيات التعلم الآلي وهي خوارزمية آلة متجه الدعم ((Support Vector Machine (SVM)) وللوصول الى المعلمات المثلى لمتغيرات SVM تم تهجينها مع خوارزمية تحسين الحوت (Whale Optimization Algorithm) لتحقيق أعلى دقة لمصنف SVM بأقل وقت وجهد في التدريب ليتم تشخيص سبع حالات مرضية (التصنيف المتعدد)، وقد تم الوصول الى أعلى دقة في التشخيص والتي تبلغ (99.92%) للنموذج المهجن DenseNet201-SVM-Whale لحل مشكلة وقت التدريب للنماذج المهجنة التي تم

بناؤها في هذه الرسالة والذي تجاوز تدريبها 24 ساعة، ولاختزال وقت التدريب الى ساعة ونصف كحد اقصى لكل نموذج مهجن تم استخدام تقنية وحدة معالجة الرسومات (Graphics Processing Unit(GPU) – (GPU-A100)).

لقد أدى تطبيق نهج الكشف بمساعدة الكمبيوتر ((Computer Aided Diagnosis (CAD)) باستخدام تقنيات التعلم الآلي والتعلم العميق إلى نتائج واعدة إلى جانب تقنيات التصوير الطبي والتشخيص التشريحي المرضي للانسجة، لأنها تطورت بسرعة في السنوات الأخيرة لحل المشكلات المعقدة. وأخيرا تم اختبار النماذج المهجنة واعتماد النموذج الأعلى دقة في التصنيف، إن هذه المنهجية الهجينة تزيد من أداء التنبؤ مما يساعد في مختبرات الفحص النسيجي على دعم الطبيب في تشخيص نوع الأورام لدى المصابين.

**Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Software**



Diagnosing Soft Tissue Tumors Using Intelligent Technologies

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Master of Science
in
Software**

**By
Noor Yahya Ghanim Shaker**

**Supervised by
Assist.Prof. Dr. Jamal Salahaldeen Majeed Sayed Hamo**

2024 A.D.

1445 A.H.

Abstract

Over the past ten years, the growth of artificial intelligence (AI) is the basis for making accurate decisions to diagnose the type of tumors by building smart programs that rely on the latest artificial intelligence technologies such as Machine Learning (ML) and Deep Learning. Learning (DL) and swarm algorithms and hybridization between them to reach the highest accuracy in diagnosis. In this research, histopathology images were adopted for pathological cases (breast tumors, lung tumors, and colon tumors). The reason for choosing these types of tumors is that they are widespread throughout the world and is the most deadly disease, The data was obtained by collecting datasets (histological images) approved by Kaggle company, amounting to 32,783 histological images, including (7783 images of benign and malignant breast tumors, 10,000 images of benign and malignant colon tumors, and 15,000 images of lung tumors (benign, malignant and squamous) according to different magnification levels of the histopathological images of the tissues, and after merging the data sets with one data set, it was divided into (80% of the training and validation data and 20% of the test data) and analyzed by performing pre-processing on them, then design and implementation the 13 models. Smart, hybrid using deep learning, machine learning, and swarm algorithm techniques to diagnose tumors, and reaching the highest accuracy in medical diagnosis. In deep learning, a convolutional neural network (CNN) was built and the algorithms (201DenseNet, 169DenseNet, 21DenseNet) and (ResNet50) were used. , ResNet101, ResNet152) that are pre-trained based on the Transfer Learning technique, which allows models to be adapted to perform new tasks, by building a new structure that suits the data set specified in this thesis, and hybridizing it with one of the machine learning techniques. It is the Support Vector Machine (SVM) algorithm. To reach the optimal parameters for the SVM variables, it was hybridized with the Whale Optimization Algorithm to achieve the highest accuracy of the SVM classifier with the least training time and effort in order to diagnose seven disease cases (multiple classification). The highest diagnostic accuracy of (99.92%) was achieved for the hybrid model DenseNet201-SVM-Whale to solve the problem of the training time for the hybrid models that were built in this thesis, the training of which exceeded 24 hours, and to reduce the training time to a maximum of an hour and a half for each hybrid model. Graphics Processing Unit (GPU) technology - (GPU-A100) was used the

application of Computer Aided Diagnosis (CAD) approaches using machine learning and deep learning techniques has yielded promising results along with medical imaging and histopathological diagnosis techniques, as they have developed rapidly in recent years to solve complex problems. Finally, the hybrid models were tested and the highest accuracy model was adopted in classification. This hybrid methodology increases prediction performance, which helps in histological examination laboratories to support the doctor in diagnosing the type of tumors in patients.