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

SRSI Regions Recognition Using Hybrid Machine Learning

**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 High Diploma in
Computer Science
By**

Rafal Nazar Younis Abdulrahman

Supervised by

Dr. Ruba Talal Ibrahim Ahmed

2025 A.D. **1446 A.H.**

ABSTRACT

Deserts pose environmental and economic challenges, and remote sensing satellite imagery (SRSI) is an important tool for studying them. Machine learning (ML) contributes to the analysis of these images to support sustainable development. This thesis aims to improve four ML approaches by combining them with pre-training methods to achieve multi-model learning. Two pre-training methods (Xception and DeneseNet201) were used to extract features, connect them, and feed them into machine learning algorithms (Light GBM, Decision Tree, Naive Bayes, and K-Nearest Neighbor). Furthermore, ensemble voting was used to improve the results of machine learning algorithms (DT, NB, and KNN) and overcome their shortcomings. The models were tested on data (SRSI) taken from Kaggle and data collected from several different sources, such as NASA and Nimbo. On both datasets, the proposed model was compared with previous and traditional studies. The results showed that the LGBM method outperformed previous, traditional methods by 99% in precision, recall, and F1-score, and by 100% in AUC-ROC. The ensemble voting method also achieved 96% in precision, recall, and F1-score on the first dataset but 91% on the combined dataset, as well as 99% AUC-ROC on both datasets. The previous study achieved an accuracy of 90-92%. Compared to standard methods, Xception, applied alone for feature extraction without combining it with DensNet201, achieved 97% accuracy on the first dataset but only 95% accuracy on the second. DensNet201, used alone for feature extraction, outperformed Xception, achieving 98% accuracy across both datasets, while Xception achieved 97% accuracy on the SRSI dataset and 95% accuracy on the combined dataset.

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

تمييز المناطق لصور الاستشعارعن بعد باستخدام التعلم الآلي الهجين

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة الدبلوم العالي في
علوم الحاسوب

من قبل
رفل نزار يونس عبد الرحمن

بإشراف
م.د ربى طلال إبراهيم احمد

١٤٤٦ هـ ٢٠٢٥م

الملخص

تشكل الصحارى تحديات بيئية واقتصادية، ويعد الاستشعار عن بعد عبر صور الأقمار الصناعية (SRSI) أداة مهمة لدراستها. يساهم التعلم الآلي (ML) في تحليل هذه الصور لدعم التنمية المستدامة. تهدف هذه الرسالة إلى تحسين أربعة مناهج MLعبر دمجها مع أساليب التدريب المسبق لتحقيق التعلم متعدد النماذج.وقد تم استخدام طريقتين للتدريب المسبق (Xception و DeneseNet201) لاستخراج الميزات، وربطها وإدخالها في خوارزميات التعلم الآلي (Light GBM وDecision Tree وNaive Bayes وK-Nearest Neighbor). بالإضافة إلى ذلك،استخدام التصويت الجماعي لتحسين نتائج خوارزميات التعلم الآلي (DT و NB و KNN) والتغلب على عيوبها.واختبار النماذج على البيانات الأولى (SRSI) مأخوذة من موقع Kaggle والثانية تم جمعها من عدة مواقع مختلفة مثل موقع NASA وNimbo. في كلتا مجموعتي البيانات، تم مقارنة النموذج المقترح مع الدراسات السابقة والتقليدية. أظهرت النتائج أن طريقة LGBM تفوقت على الطرق السابقة والتقليدية بنسبة ٩٩٪ في الدقة والتذكر ودرجة F1، وبنسبة ١٠٠٪ في AUC-ROC. كما حققت طريقة التصويت الجماعي ٩٦٪ في الدقة والتذكر ودرجة F1 في مجموعة البيانات الأولى ولكن ٩١٪ في مجموعة البيانات المجمعه، بالإضافة إلى ٩٩٪ AUC-ROC في كلتا مجموعتي البيانات. بينما حققت الدراسة السابقة دقة ٩٠-٩٢٪. بالمقارنة مع الطرق القياسية، تم تطبيق Xception بمفرده في استخراج الميزات دون دمجه مع DensNet201، وحققت دقة ٩٧٪ في مجموعة البيانات الأولى، ولكن دقة ٩٥٪ فقط في الثانية.وقد استخدام DensNet201 بمفرده لاستخراج الميزات وتفوق على Xception، محققًا دقة بنسبة ٩٨٪ عبر كلتا مجموعتي البيانات بينما حقق ال Xception في مجموعة البيانات SRSI نسبة ٩٧٪ وفي مجموعة البيانات اللتي تم جمعها حقق دقة بنسبة ٩٥٪.