

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



# **Semantic Segmentation of Natural Images Using Machine Learning Models**

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

**Submitted by  
Mohammed Saaduldin Jassim Mohammed**

**Supervised by  
Dr. Mohammed Chachan Younis  
Assistant Professor**

---

**2023 A.D.**

**1444 A.H.**

## **ABSTRACT**

The replication of human intellectual processes by machines, particularly computer systems, is known as artificial intelligence (AI). AI is an intelligent tool that is utilized across sectors to improve decision-making, increase productivity, and eliminate repetitive tasks. Machine learning (ML) is a key component of AI since it includes understanding and developing ways that can learn to conduct actions or make suggestions to improve performance on tasks.

This work has solved the image segmentation problem for the natural images to segment out water, land, and sky using machine learning approaches. Instead of applying image segmentation directly to the images, images are preprocessed, and statistical and textural features are then passed through a neural network for the pixel-wise semantic segmentation of the images. A 5X5 data chunk was chosen over the pixel-by-pixel technique since it requires less resources and time for training and testing. The error is calculated between model predictions and ground truth labels. The model weights are tuned using this calculated error.

The training images in this work are 53 type PNG format containing only one class (i.e., water, land, or sky with 27, 14, and 12 images, respectively). While 7 PNG images with their respective ground truth containing all three classes in an image have been used as testing images. A 64-bit Windows 10 OS, with an Intel(R) Core i7 processor (2.20GHz) with 16GB of RAM was used for the experimentation. The classification was achieved by using the tools available in the MatLab R2022b Machine Learning Toolbox.

The Neural architecture used in this work takes a batch size of 16. A Scaled Conjugate Gradient Algorithm is used to train the model on cross-entropy loss. After the learning of the artificial neural network, the trained model is used to predict the unseen testing data. The numerical results from confusion matrices and ROC plots for both training and testing stages show that the trained model achieved an accuracy of 97.5%, and the accuracy for unknown three tested images is 91.0%, 94.2%, and 96.8% which indicates that the trained model is performing well on unknown images.



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

## التقطيع الدلالي للصور الطبيعية باستخدام التعلم الآلي

رسالة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل  
كجزء من متطلبات نيل شهادة دبلوم عالي في  
علوم الحاسوب

من قبل

محمد سعد الدين جاسم محمد

بإشراف

أ.م.د. محمد جاجان يونس

## الخلاصة

يمكن تعريف الذكاء الاصطناعي على انه استخدام الآلات، وخاصة اجهزة الحاسوب، لتكرار العمليات الفكرية التي يتم تنفيذها من قبل الانسان. فالذكاء الاصطناعي هو أداة ذكية يتم استخدامها في مختلف القطاعات لتحسين عملية صنع القرار وزيادة الإنتاجية والقضاء على المهام المتكررة. ويعد التعلم الآلي المكون الرئيس للذكاء الاصطناعي، فهو يتضمن فهم وتطوير طرائق يمكنها التعلم لإتخاذ إجراءات معينة أو تقديم بعض الاقتراحات لتحسين أداء الكثير من المهام المختلفة.

تم في هذه الدراسة تصنيف الصور الطبيعية باستخدام تقنيات التعلم الآلي القائم على التجزئة لتقسيم الصور المختبرة الى ماء، أرض، وسماء. فبدلاً من تطبيق تجزئة الصور مباشرة يتم معالجة الصور مسبقاً ثم يتم تمرير الميزات الإحصائية والتركيبية عبر شبكة عصبية لتجزئة الصور حسب البكسل. تم اختيار نافذة بابعاد 5x5 بكسل لأنها تتطلب موارد ووقتاً أقل لتدريب واختبار الصور. كما تم حساب الخطأ بين التنبؤ العشوائي للبيانات والقيم الحقيقية، ثم تم ضبط أوزان الشبكة العصبية باستخدام الخطأ المحسوب.

تم تدريب 53 صورة من نوع PNG تحوي صنف واحد فقط (27 صورة فئة ماء، 14 صورة فئة أرض، و 12 صورة فئة سماء). وأما في مرحلة الاختبار فقد تم اختبار 7 صور من نوع PNG مرتبطة مع صور الحقيقة الارضية (Ground-truth) والتي تحتوي على الاصناف الثلاثة في كل صورة مختبرة. تم العمل على نظام التشغيل ويندوز 10 بعدد 64 بت، ومعالج Intel (R) Core i7 (2.20 جيجاهرتز) مع 16 جيجابايت من ذاكرة الوصول العشوائي (RAM). كما تم استخدام صندوق ادوات تدريب الآلة

(Machine Learning Toolbox) المتوفر في لغة ماتلاب الاصدار R2022b.

تم في هذه الدراسة استخدام خوارزمية التدرج المترافق/المتقارب المقياس ( Scaled Conjugate Gradient Algorithm ) لتدريب النموذج بالاعتماد على معدل فقدان كمية المعلومات (Cross-entropy)، وبحجم دفعة (Batch-size) يتكون من 16 مدخل بعدد الخصائص التي تم استخلاصها. وبمجرد الانتهاء من تدريب وتعلم الشبكة العصبية تم استخدام نموذج الشبكة هذا للتنبؤ ببيانات الاختبار غير المعروفة بالنسبة للشبكة. وقد أوضحت نتائج مصفوفة الارتباك (Confusion-matrix) وصندوق منحنى الخصائص (ROC Curve) لمرحلي التدريب والاختبار أن النموذج المدرب حقق دقة (Accuracy) بلغت 97%. واما في مرحلة الاختبار فقد بلغت دقة اختبار ثلاثة صور غير معروفة 91.0%، 94.2% و 96.8% مما يدل على أن النموذج المدرب يعمل بشكل جيد على البيانات الجديدة التي لم يتم تدريب الشبكة عليها سابقاً.