تقييم أداء التصنيف القائم على الكائن للمناطق المبنيّة في مرئيات القمر الصناعيّ Sentinel-2B مقارنة مع التصنيف القائم على البكسل

د. م. فادي عز الدين شعبان*

الملخص

يمثل التوسع السريع للمناطق الحضرية تحديا كبيرا أمام التنمية المستدامة وإدارة الموارد لذا توفر صور الأقمار الصناعية المجانية (Sentinel-2) ذات الدقة (10م) بيانات أساسية لإنتاج خرائط دقيقة للمناطق المبنية. إذ تعدّ هذه الخرائط عند دمجها في نظم المعلومات الجغرافية (GIS) أداة قوية لدعم التخطيط الحضري، والرصد البيئي واستجابة الكوارث.

يهدف هذا البحث إلى تقييم أداء خوارزميات التصنيف القائم على الكائن مقارنة بالتصنيف القائم على البكسل في استخراج المناطق المبنية من صور القمر الصناعي(Sentinel-2B). تم تطبيق مجموعة متنوعة من الخوارزميات على منطقة دراسة بمحافظة طرطوس في سورية، بما في ذلك شجرة القرار (DT)، والشجرة العشوائية (RT)، والجار الأقرب (KNN) وبايز (Bayes) للتصنيف القائم على الكائن بالإضافة إلى تطبيق أسلوب قواعد القرار. بينما تم استخدام التصنيف المراقب (خوارزمية الاحتمالية العظمى) للتصنيف القائم على البكسل.

أظهرت النتائج تفوق التصنيف القائم على الكائن بشكل عام على التصنيف القائم على البكسل، خاصة عند استخدام قواعد القرار وخوارزمية بايز. ويعزى ذلك إلى قدرة التصنيف القائم على الكائن على الاستفادة من المعلومات السياقية والخصائص الهيكلية للمناطق المبنية، مما يحسن من دقة التصنيف.

حقق أسلوب قواعد القرار أعلى دقة، حيث بلغت الدقة الكلية (94%) ومعامل كابا (0.85)، مقارنةً بـ (89%) و(0.74) للتصنيف القائم على البكسل. ومع ذلك، فإن اختيار أفضل خوارزمية يعتمد على التوازن المطلوب بين دقة المستخدم ودقة المنتج، وعلى طبيعة البيانات والمنطقة المدروسة. تعكس الفروقات في نسب المساحات المبنية وغير المبنية بين الطرق المختلفة تباينا في دقة التصنيفات، مما يستدعي دراسة خصائص كل طريقة لاختيار الأنسب وفقا للهدف والبيئة المدروسة.

كلمات المفتاحية: المناطق المبنية، Sentinel-2B، نظم المعلومات الجغرافية GIS، التصنيف القائم على الكائن، التصنيف القائم على البكسل.

* أستاذ مساعد في قسم الهندسة الطبوغرافية –كلية الهندسة المدنية –جامعة تشرين – اللاذقية – سورية.
✉ fadi.chaaban@tishreen.edu.sy

Evaluation of Object-Based Classification Performance for Built-Up Areas in Sentinel-2B Satellite Imagery Compared to Pixel-Based Classification

Dr.Eng Fadi CHAABAN*

Abstract

The rapid expansion of urban areas poses a significant challenge to sustainable development and resource management. Free Sentinel-2 satellite imagery, with its 10m resolution, provides essential data for producing accurate maps of built-up areas. When integrated into Geographic Information Systems (GIS), these maps serve as powerful tools for urban planning, environmental monitoring, and disaster response.

This research aims to evaluate the performance of object-based classification algorithms compared to pixel-based classification in extracting built-up areas from Sentinel-2B satellite images. Various algorithms were applied to a study area in Tartous Governorate, Syria, including Decision Tree (DT), Random Tree (RT), k-Nearest Neighbors (KNN), and Bayes for object-based classification, in addition to the application of decision rules. Meanwhile, supervised classification (Maximum Likelihood Algorithm) was used for pixel-based classification.

The results showed that object-based classification generally outperformed pixel-based classification, especially when using decision rules and the Bayes algorithm. This is attributed to the ability of object-based classification to utilize contextual information and structural characteristics of built-up areas, improving classification accuracy.

The decision rules approach achieved the highest accuracy, with an overall accuracy of 94% and a Kappa coefficient of 0.85, compared to 89% and 0.74 for pixel-based classification. However, the choice of the best algorithm depends on the required balance between user accuracy and

producer accuracy, as well as the nature of the data and the studied area. Differences in the proportions of built-up and non-built-up areas between methods reflect variations in classification accuracy, necessitating an examination of each method's characteristics to choose the most suitable one according to the objective and studied environment.

Keywords: Built-up areas, Sentinel-2, Geographic Information Systems(GIS), Object-based classification, Pixel-based classification.

* Associate Professor, Department of Topographic Engineering, Faculty of Civil Engineering, Tishreen University, Lattakia, Syria.
✉ fadi.chaaban@tishreen.edu.sy

1 مقدمة

تُصنف تقنيات استخراج المناطق المبنية من صور الأقمار الصناعية إلى نوعين رئيسيين: الأول هو التقنيات المعتمدة على التصنيف التقليدي للصور متعددة الأطياف، وتشمل أساليب متنوعة مثل: التصنيف المراقب، والتصنيف غير المراقب والتصنيف القائم على الكائن(كما يسمى التصنيف الهدفي)، والتصنيف بالتعلم العميق. الثاني هو تقنيات تستند إلى مؤشرات الفرق الطبيعي مثل مؤشر الفرق الطبيعي للمناطق المبنية (NDBI)، ومؤشر المناطق المبنية القائم على تحليل المكونات الرئيسية (PCABI)، ومؤشر المباني المُحسّن (EBI)، ومؤشر المناطق الحضرية (UI)، إلخ [1]. من جهة أخرى يمكن تصنيف خوارزميات تصنيف صور الأقمار الصناعية إلى تقنيات تعتمد على البيكسل (Pixel-base) كوحدة تحليل أساسيّة ، وتقنيات تعتمد على الكائن كوحدة تحليل (Object-based) ويتفرّع ضمن هذه التقنيات خوارزميات عديدة ومتنوّعة[2].

يعتمد التصنيف القائم على البكسل على تحليل كل بكسل على حدة بناءً على خواصه الطيفية .هذا النوع من التصنيف مناسب للصور ذات الطبيعة البسيطة ودقة مكانية متوسطة، وقد استُخدم على نطاق واسع في العديد من الأبحاث السابقة [3]. مع ذلك، ومع التطور التكنولوجي للمستشعرات وزيادة الدقة المكانية لصور الاستشعار عن بعد، أصبحت نتائج هذه الخوارزميات غير كافية .ولذلك، توجّه الباحثون إلى تطوير خوارزميات جديدة تعتمد على مبادئ الرؤية الحاسوبية، والتي تقوم بتقسيم الصورة إلى كائنات بناءً على مجموعة

من الخصائص المكانية والطيفية. وبالتالي، يمكن استخلاص عناصر وكائنات مختلفة من صور الاستشعار عن بعد بدقة أكبر [4].

يتم تصنيف الصور القائم على البكسل إما بطريقة مراقبة أو غير مراقبة. في التصنيف غير المراقب، يتم تقسيم الصورة إلى عدة فئات (Classes) بناءً على الخصائص الإحصائية الداخلية للصورة. أما في التصنيف المراقب، فيعتمد البرنامج على توجيهات المستخدم، حيث يتم تزويده مسبقا بمجموعة بيانات تدريبية تمثل فئات التصنيف المختلفة، وتُعرف هذه البيانات بـ "عينات التدريب" (Training data). من خلال هذه العينات، يصبح البرنامج قادرا على التعرف على الفئات المختلفة داخل الصورة.

هناك العديد من الأساليب المستخدمة في التصنيف المراقب، من بينها الطرق الإحصائية التي تُعد الأكثر شيوعا، بالإضافة إلى الطرق غير الإحصائية والطرق الهندسية، حيث تعتمد الأخيرة على تقسيم الفئات التصنيفية باستخدام أسطح هندسية [2]

يُعد التصنيف بطريقة الاحتمالية العظمى (Maximum Likelihood Classification) من أكثر الأساليب شيوعاً واستخداماً في التصنيف الإحصائي. تعتمد خوارزمية هذا التصنيف على منحنيات التوزيع التكرارية (Histogram) للنطاقات التي تتبع توزيعاً طبيعياً. يستند هذا التصنيف إلى قاعدة قرار الاحتمال الأعظمي، التي تعتمد على مبدأ الاحتمالات، والمعروفة أيضاً باسم تصنيف بايز (Bayes Classification) [5]، والذي سيعتمد في دراستنا الحالية.

يقوم التصنيف القائم على الكائن باستخراج الفئات من الصور من خلال اتباع تسلسل هرمي محدد. يعتمد هذا النوع من التصنيف على تحليل عناصر الصورة بدقة، حيث يأخذ في الاعتبار خصائص اللون والشكل والنسيج والحجم، بالإضافة إلى الخصائص الطيفية وحالة كل عنصر وعلاقته بالعناصر المجاورة. الهدف من ذلك هو الحصول على نتائج دقيقة تحاكي قدرة الإنسان على التفريق بين الكائنات المرئية [6]، وذلك على عكس التصنيف القائم على البكسل الذي يعتمد فقط على قيمة النطاق في كل بكسل على حدة.

تُعد عملية التجزئة (Segmentation) الخطوة الأولى والأكثر أهمية في التصنيف القائم على الكائن. فهي تقوم بتقسيم الصورة إلى مناطق متجانسة (كائنات) بناءً على خصائصها

المشتركة. وبهذا، يتم تحديد الكائنات التي نرغب في تصنيفها بشكل دقيق، وذلك استنادا إلى خصائصها المميزة مثل القيم الطيفية والشكل والنسيج [7]، [8].

أظهرت العديد من الدراسات الحديثة، مثل دراسات[11] ,[10] ,[9] ,[1]. فعالية كبيرة لصور القمر الصناعي (Sentinel-2) في استخراج المناطق المبنية بدقة عالية. وقد استخدمت هذه الدراسات مجموعة متنوعة من الأساليب والخوارزميات لتحقيق ذلك، مما أسفر عن نتائج واعدة. حيث حققت هذه الدراسات دقة كلية تراوحت بين (79.95%) و (93.33%)، وقيم معامل كابا بين (0.785) و (0.922)، مما يشير إلى قدرة هذه التقنية على توفير معلومات دقيقة حول التوسع العمراني وتغيرات استخدام الأراضي على نطاق واسع.

استخدمت العديد من الدراسات السابقة أساليب تصنيف متنوعة لأغراض متعددة وباعتماد صور فضائية بدقات مختلفة [1]، [2]، [3]، [7]، [9]، [10]. ومع ذلك، تظل المقارنة بين التصنيف القائم على الكائن والتصنيف القائم على البكسل لاستخراج المناطق المبنية من صور الأقمار الصناعية (Sentinel-2B) محدودة نسبيا. لذا، يقترح هذا البحث إجراء مقارنة علمية بين أداء التصنيف القائم على الكائن وأداء التصنيف القائم على البكسل في استخراج المناطق المبنية باستخدام صور الأقمار الصناعية من نوع (Sentinel-2B) ، وذلك ضمن منطقة دراسة تقع في محافظة طرطوس، سورية. سيتم تطبيق مجموعة من خوارزميات التصنيف القائم على الكائن لتحليل الصور، بينما سيتم استخدام طريقة التصنيف المراقب في إطار التصنيف القائم على البكسل. سيتم تقييم النتائج باستخدام مقاييس دقة معتمدة وشائعة. تهدف هذه المقارنة إلى تحديد المنهجية الأكثر كفاءة ودقة في استخراج المناطق المبنية، بما يسهم في تحسين جودة ودقة الخرائط المستخلصة من بيانات الأقمار الصناعية.

2 هدف البحث وأهميته

يهدف البحث بشكل أساسي إلى تقييم دقة وفعالية التصنيف القائم على الكائن للمناطق المبنية باستخدام صور (Sentinel-2B) مقارنةً مع التصنيف القائم على البكسل، لتحديد

أي منهما يوفر تمثيلاً أكثر دقة للواقع .يمكن تحقيق هذا الهدف من خلال الأهداف الفرعية التالية:

- تطبيق مجموعة متنوعة من خوارزميات التصنيف القائم على الكائن، تشمل: شجرة القرار (Decision Tree - DT)، والشجرة العشوائية (Random Tree - RT)، والجار الأقرب (K-Nearest Neighbors - KNN)، وخوارزمية بايز (Bayes)، بالإضافة إلى تطبيق أسلوب قواعد القرار (Decision Rules) من خلال تطوير مجموعة من القواعد على منطقة دراسة محددة في محافظة طرطوس.
- تطبيق التصنيف القائم على البكسل باستخدام خوارزمية الاحتمالية العظمى لنفس المنطقة.
- تقييم دقة التصنيف لكل خوارزمية باستخدام مؤشرات تقييم الدقة مثل الدقة الكلية، معامل كابا، ودقة المنتج والمستخدم.
- تقييم أداء خوارزميات التصنيف المختلفة وتحديد الخوارزمية التي توفر أعلى دقة في استخلاص المناطق المبنية.

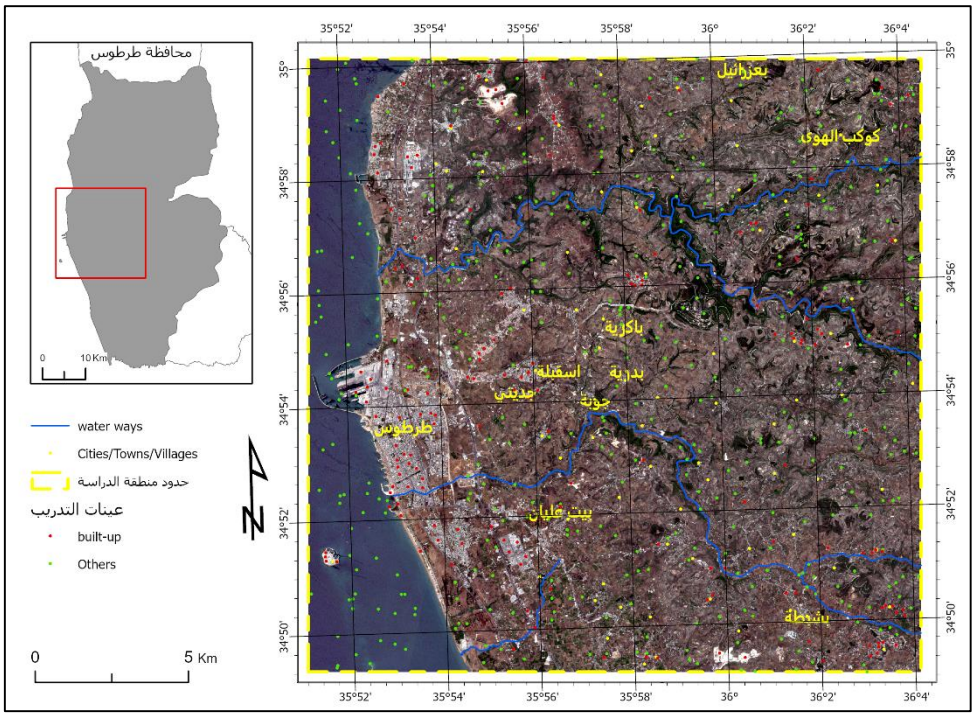
تأتي أهمية البحث من أنه يساهم في تطوير المعرفة العلمية في مجال الاستشعار عن بعد ونظم المعلومات الجغرافية وتصنيف الصور، وذلك من خلال تقييم أداء خوارزميات مختلفة في استخراج المناطق المبنية من صور الأقمار الصناعية المجانية (Sentinel-2B) ذات الدقة (10م). يمكن أن توفر النتائج المتوقعة إرشادات قيمة للباحثين والمختصين. كما ستوفر أداة قوية لدعم عملية صنع القرار في مجالات التخطيط العمراني وإدارة الموارد، من خلال تحسين دقة الخرائط المكانية واتخاذ قرارات أكثر استنارة في مجالات مثل تقييم المخاطر، وتخطيط البنية التحتية، ورصد التغيرات البيئية.

1 مواد وطرائق البحث

1.1 منطقة الدراسة

تقع منطقة الدراسة في محافظة طرطوس، سورية، وتمتد على مساحة 400 كيلومتر مربع (20 كم × 20 كم) وتشمل مركز مدينة طرطوس وعدد من القرى والبلدات المحيطة بها.

حيث تمتد بين خطي الطول °35.850026، °36.075144 شرقا، وخطي العرض °34.817516، ° 35.002960 شمالاً بنظام الإحداثيات الجيوديزي العالميWGS 84 كما هو موضح في الشكل1. يسود المنطقة المناخ المعتدل، مع متوسط هطول أمطار سنوي يبلغ حوالي840 ملم وتتميز بتنوع في الغطاء الأرضي، حيث تتراوح الارتفاعات بين 0 و622 متر فوق مستوى سطح البحر. يقطع المنطقة نهري الغمقة والحصين من الشرق إلى الغرب.



الشكل1 منطقة الدراسة

1.2 البيانات المستخدمة

استُخدم في هذه الدراسة بيانات القمر الصناعي (Sentinel-2B) التي تم الحصول عليها في 24 حزيران 2023 لاســـتخراج المنـاطق المبنيـة. تم تنزيـل المرئيـة من الموقع

(https://browser.dataspace.copernicus.eu/)، إذ يتم الحصـــول على بيانات Sentinel-2B ومعالجتها بواسطة وكالة الفضاء الأوروبية [1] .

تم تحميل المرئية بغطاء سحابي أقل من (0.01٪)، بأبعاد(110 كم × 110 كم). مستوى معالجة المرئية هو المسـتوى (A2)، والذي يتضـمن التصـحيح الإشـعاعي والهندسـي، والتصـــحيح العمودي (orthorectification)، والتســـجيل المكاني على نظام مرجعي عالمي، وهو (WGS84) وإسـقاط ((Universal Transverse Mercator (UTM)

..

تُعد أقمار (Sentinel-2) مهمة تصـــوير متعددة الأطياف عالية الدقة، حيث تحتوي على (13) نطاقا طيفيا تمتد من الجزء المرئي والقريب من الأشـــعة تحت الحمراء (VNIR)إلى جزء الأشـــعة تحت الحمراء قصـــيرة الموجة (SWIR) . من بين هذه النطاقات، تتمتع أربعة نطاقات بدقة مكانية تبلغ (10) أمتار، وســـتة نطاقات بدقة (20) متر، والنطاقات الثلاثة المتبقية بدقة (60) متر. عادةً ما تُفضـــل الصـــور ذات الدقة المكانية العالية للتطبيقات المتعلقة بالمناطق الحضـرية [12] ,[1]. يوضـح الجدول (1) منتجات البيانات المتاحة للمستخدمين .

الجدول 1 : خصائص نطاقات (Sentinel-2) [12]

Sentinel-2 bands	Central wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation red edge	0.705	20
Band 6 - Vegetation red edge	0.740	20
Band 7 - Vegetation red edge	0.783	20
Band 8 - NIR	0.842	10

Band 8A - Vegetation red edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

1.3 البرامج والأدوات المستخدمة

- **برنامج ArcGIS Pro**

برنامج (ArcGIS Pro) هو تطبيق من بيئة (GIS) مكتبي احترافي متكامل مطور من قبل شركة (ESRI) (https://www.esri.com). إذ يتيح استكشاف البيانات، عرضها، وتحليلها؛ وإنشاء خرائط ثنائية الأبعاد ومشاهد ثلاثية الأبعاد. يتميز البرنامج بدعمه الكامل للبيانات ثلاثية الأبعاد والقدرة على التكامل مع (ArcGIS Online)، بالإضافة إلى دعمه للغة البرمجة (Python 3).

يتيح برنامج (ArcGIS Pro) استخدام الأدوات التحليلية المكانية على البيانات ثنائية وثلاثية ورباعية الأبعاد لتحديد الأنماط، وإجراء التنبؤات والاستفسارات الوصفية والمكانية. كما يوفر البرنامج مجموعة من الأدوات القوية لإدارة وتحليل البيانات الشعاعية، البيانات المصفوفية وصور الطائرات بدون طيار ومرئيات الأقمار الصناعية و بيانات الليدار.

تم استخدام نسخة (ArcGIS Pro 3.0) في جمع البيانات ومعالجتها وإجراء التصنيف المراقب باستخدام طريقة الاحتمالية العظمى وإجراء تقييم الدقة، بالإضافة إلى الإخراج الكارتوغرافي للخرائط الغرضية.

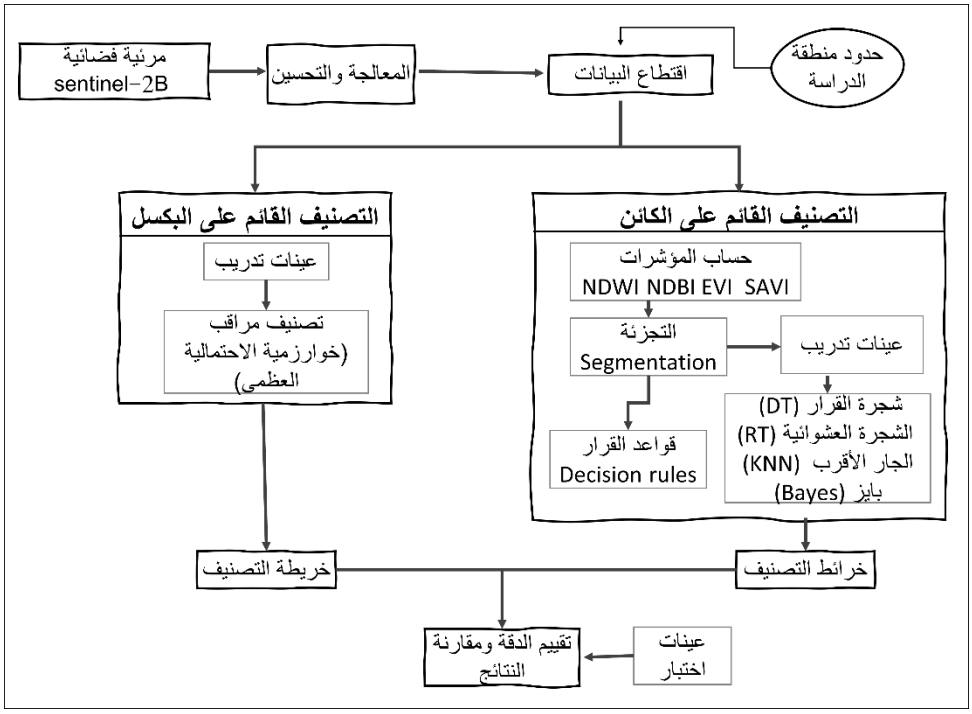
- **برنامج eCognition Developer**

يُعد (eCognition Developer) مجموعة من الأدوات والخوارزميات البرمجية لتحليل الصور وتصنيفها، طورتها شركة (Trimble Geospatial Inc). ويركز على التحليل القائم على الكائنات ودمج البيانات والتحليلات المتكاملة[6] . يتضمن البرنامج أحد عشر خوارزمية تجزئة مختلفة. الخوارزمية المستخدمة في هذا البحث هي خوارزمية التجزئة متعددة الدقة ((Multiresolution segmentation (MRS). تُعد خوارزمية التجزئة متعددة الدقة خوارزمية تُنمّي المناطق من خلال البدء بوحدة بكسل واحدة، تُعرف باسم وحدة

بكسل "البذرة"، ثم دمجها مع وحدات البكسل المجاورة حتى تصل إلى قيمة عتبة الحد الأقصى لعدم التجانس. تم اختيار هذه الخوارزمية لأنها تعمل بشكل جيد مع الصور متعددة النطاقات، مثل مرئيات الأقمار الصناعية التي ينتجها (Sentinel-2) [12]. تم استخدام النسخة (eCognition Developer 10.3) من هذا البرنامج في اشتقاق المؤشرات الضرورية وعملية التجزئة والتصنيف القائم على الكائن للمرئية.

2 منهجية البحث

يوضح الشكل 2 مخططا انسيابيا يلخص المنهجية المقترحة والمتبعة في هذا البحث.



الشكل 2 المنهجية المقترحة للبحث

شملت المنهجية المراحل الرئيسية التالية:

2.1 الحصول على البيانات ومعالجتها

تمّ بدايةً تحميل بيانات القمر الصناعي (Sentinel-2B) لمنطقة الدراسة من الموقع المذكور سابقا. ثمّ، باستخدام برنامج (ArcGIS Pro)، تمّ تشكيل صورة ملونة مركّبة

(Composite Bands) بدقة مكانية 10 أمتار من ستة نطاقات طيفية للقمر الصناعي (Sentinel-2B)، هي:

- النطاق الأزرق (النطاق 2 عند 490 نانومتر)، دقة 10م.
- النطاق الأخضر (النطاق 3 عند 560 نانومتر) ، دقة 10م.
- النطاق الأحمر (النطاق 4 عند 665 نانومتر) ، دقة 10م.
- نطاق الأشعة تحت الحمراء القريبةNIR (النطاق 8 عند 842 نانومتر) ، دقة 10م.
- نطاق الأشعة تحت الحمراء القصيرةSWIR (النطاق 11 عند 1610 نانومتر) ، دقة 20م.
- نطاق الأشعة تحت الحمراء القصيرةSWIR (النطاق 12 عند 2190 نانومتر) ، دقة 20م.

أُضيف نطاقي الأشعة تحت الحمراء القصيرة (SWIR) للمساهمة في حساب المؤشرات الطيفية اللازمة لعملية التصنيف القائم على الكائن. بعد ذلك، تمّ اقتطاع منطقة الدراسة من الصورة المركّبة وفقا لحدود منطقة الدراسة (20 كم × 20 كم) للحصول على الصورة النهائية (الشكل 1) بستة نطاقات طيفية.

2.2 التصنيف القائم على البكسل

أجري تصنيف مراقب قائم على البكسل باستخدام طريقة الاحتمالية العظمى للصورة المقتطعة ضمن برنامج (ArcGIS pro)، حيث تم تصنيف منطقة الدراسة إلى فئتين رئيسيتين:

- مناطق مبنية :(Built-up) وتشمل جميع أنواع الأبنية والمنشآت.
- مناطق أخرى (غير مبنية) :((Others (non-Built-up) وتشمل جميع الغطاءات الأرضية الأخرى غير المبنية.

اعتمد التصنيف على مجموعة من عينات التدريب التي غطّت كلا فئتي التصنيف المذكورتين، حيث بلغ عدد نقاط العينة 1000 نقطة، منها 400 نقطة لصنف المناطق المبنية و600 نقطة لصنف المناطق غير المبنية، الشكل (1).

2.3 التصنيف القائم على الكائن

ضمن برنامج (eCognition)، تم استيراد المرئية المقتطعة وحساب مجموعة من المؤشرات شملت مؤشر المسطحات المائية (NDWI)، ومؤشر المناطق المبنية (NDBI)، ومؤشر الغطاء النباتي المحسن (EVI)، ومؤشر التربة (SAVI)، والتي تساعد في عملية التجزئة والتصنيف.

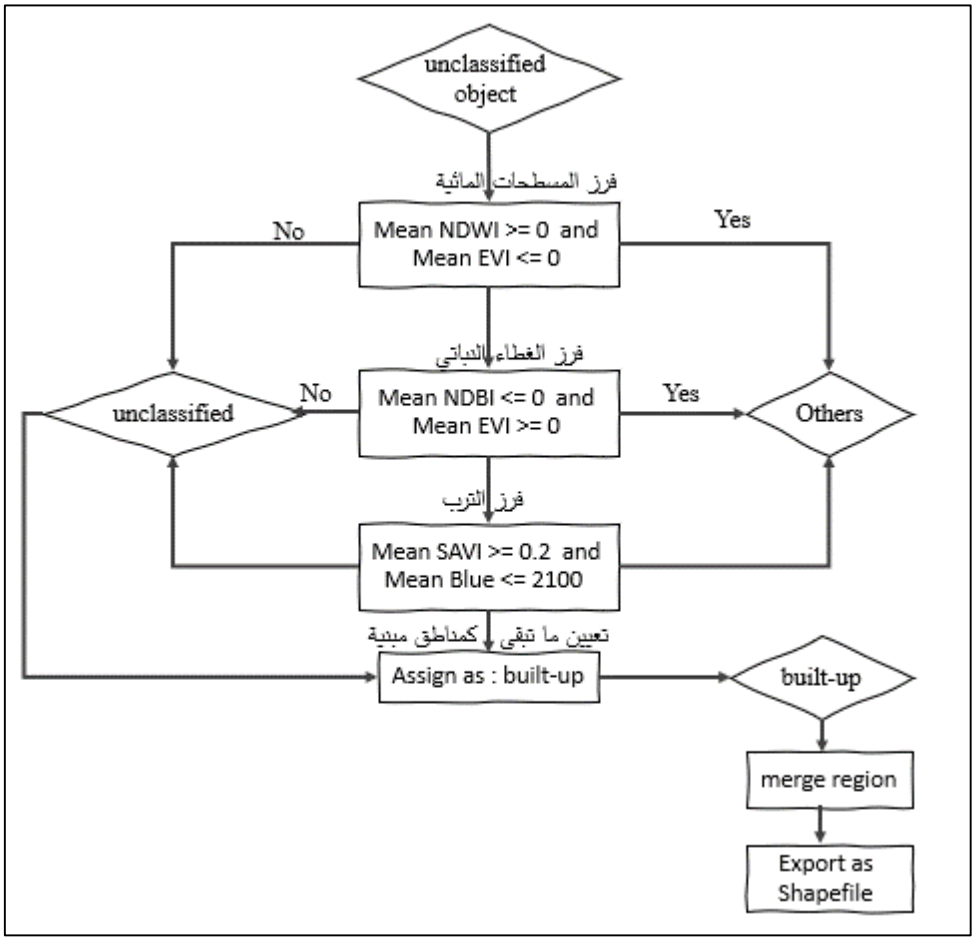
بعد ذلك، تمت عملية التجزئة باستخدام التجزئة متعددة الدقة مع تحديد معامل المقياس (Scale Parameter) بـ (25) للحصول على عناصر متوسطة الحجم. يتحكم هذا المعامل في مقدار التباين الطيفي داخل الكائنات، وبالتالي يؤثر على حجمها الناتج؛ فكلما كانت قيمته أكبر، كان حجم العناصر أكبر، والعكس صحيح. بمعنى آخر، يحدد هذا المعامل الحد الأقصى المسموح به لعدم تجانس الكائنات في الصورة الناتجة. وقد اعتمد العديد من الباحثين في العديد من الدراسات التطبيقية على نهج التجربة والخطأ لاختيار معامل المقياس المناسب لعملية التصنيف واستخراج كائنات الصورة [7]، [8].

تم اختيار قيمة معامل الشكل (Shape)، بالتجربة، مساوية لـ (00.2)، حيث يمثل هذا المعامل التوازن بين شكل الكائنات ولونها الطيفي. إذا كانت قيمته (0)، يتم أخذ اللون بالاعتبار فقط، بينما إذا كانت قيمته أكبر من (0)، يتم أخذ شكل الكائنات بعين الاعتبار مع اللون. كلما زادت القيمة، زادت أهمية أخذ الشكل بالاعتبار. وتم تعيين قيمة معامل الاكتناز (Compactness) بالتجربة بمقدار (0.5)، وهو معامل يوازن بين تمثيل انضغاط الأشياء المتكونة أثناء التجزئة. وتم إعطاء كافة الطبقات الداخلة في عملية التجزئة نفس الوزن.

بعد الحصول على الصورة المجزأة إلى كائنات، تم التصنيف باستخدام أسلوبين: أسلوب بناء قواعد القرار، وأساليب استخدام خوارزميات التصنيف المتوفرة ضمن برنامج (eCognition)، وهي شجرة القرار (DT)، والشجرة العشوائية (RT)، والجار الأقرب (KNN)، وبايز (Bayes).

2.3.1 التصنيف باستخدام قواعد القرار Decision Rules

يوضح الشكل 3 شجرة قواعد القرار التي تم اتباعها، حيث تم تطوير وتطبيق مجموعة من قواعد القرار (decision rules) ضمن برنامج (eCognition) بهدف تصنيف واستخراج المناطق المبنية وغير المبنية. تم تعريف الكائنات المستهدفة من خلال إنشاء قواعد تستند إلى خصائصها، باستخدام القيم المتوسطة للمؤشرات.



الشكل 3 شجرة قواعد القرار المتبعة في التصنيف

تم استخدام مجموعة من المؤشرات بدلاً من الاعتماد على مؤشر المناطق المبنية فقط لتحسين دقة التصنيف وتقليل التداخل بين الفئات، خاصة في المناطق التي تتشابه فيها

القيم الطيفية بين المناطق المبنية وغير المبنية. كما أن دمج مؤشرات مثل مؤشر التربة ومؤشر الغطاء النباتي المحسن يتيح تمييزا أفضل للتباين الطيفي والمكاني.

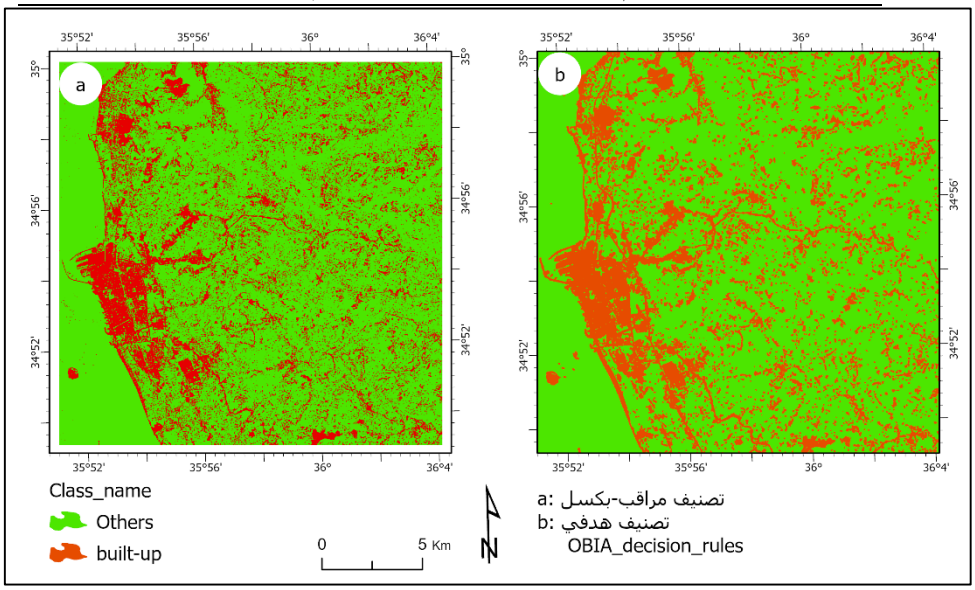
كما تم اعتماد هذا النهج التدريجي لأن المسطحات المائية والتربة والغطاء النباتي لها خصائص طيفية مميزة يسهل فصلها بدقة. بعد استبعاد هذه الفئات، يصبح ما تبقى أكثر وضوحا لتصنيفه كمناطق مبنية، مما يقلل التداخل الطيفي ويزيد من دقة التصنيف. هذا الترتيب يساعد على التعامل مع التباين الكبير في المناطق المبنية، خاصة تلك التي قد تتشابه طيفيا مع التربة أو الغطاء النباتي.

2.3.2 التصنيف باستخدام خوارزميات التصنيف

في هذه المرحلة، تم تصنيف الصورة المجزأة باستخدام أربعة خوارزميات مختلفة هي: شجرة القرار (DT)، والشجرة العشوائية (RT)، والجار الأقرب (KNN)، وبايز (Bayes). وقد تم الاعتماد على متوسط القيم الرقمية للبكسلات في جميع النطاقات الطيفية للصورة كمدخل لهذه الخوارزميات. قبل عملية التصنيف، تم تدريب هذه الخوارزميات باستخدام نفس عينات التدريب (1000 نقطة، منها 400 نقطة لصنف المناطق المبنية و600 نقطة لصنف المناطق غير المبنية، الشكل (1) التي استخدمت في التصنيف القائم على البكسل

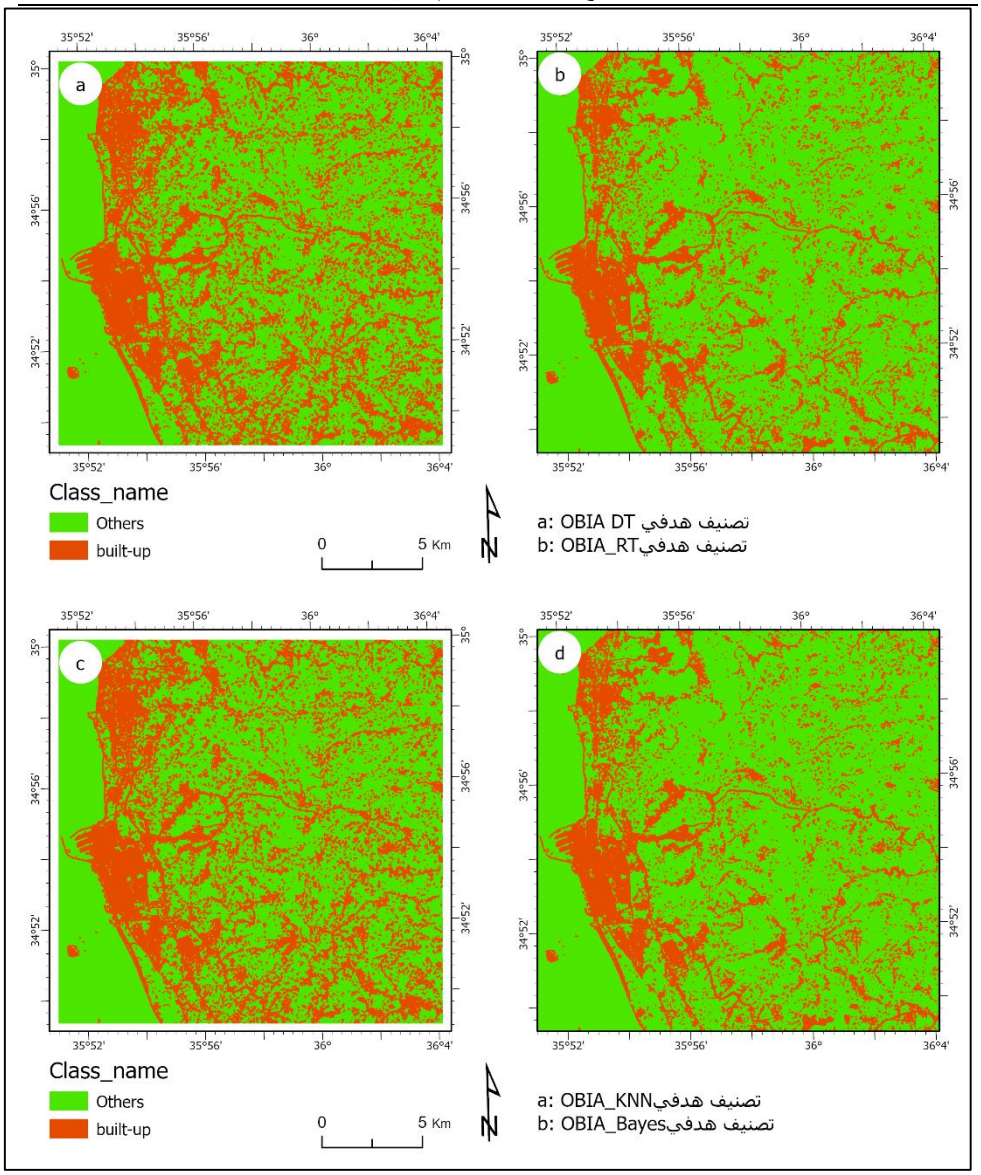
3 النتائج والمناقشة

قمنا بتصدير النتائج النهائية من برنامج (eCognition) إلى برنامج (ArcGIS Pro) لعرضها وتقييم دقة النتائج. يوضح الشكل 4 (a) نتيجة التصنيف اعتماداً على البكسل باستخدام خوارزمية الاحتمالية العظمى، بينما يوضح الشكل 4 (b) النتيجة اعتماداً على التصنيف القائم على الكائن باستخدام قواعد القرار (Decision Rules).



الشكل 4 نتائج التصنيف: (a) القائم على البكسل باستخدام خوارزمية الاحتمالية العظمى. (b) التصنيف القائم على الكائن باستخدام قواعد القرار

كما يوضح الشكل 5 (a) نتيجة التصنيف اعتماداً على التصنيف القائم على الكائن باستخدام شجرة القرار (DT)، والشكل 5 (b) النتيجة باستخدام الشجرة العشوائية (RT)، والشكل 5 (c) النتيجة باستخدام الجار الأقرب (KNN)، والشكل 5 (d) النتيجة باستخدام خوارزمية بايز (Bayes).

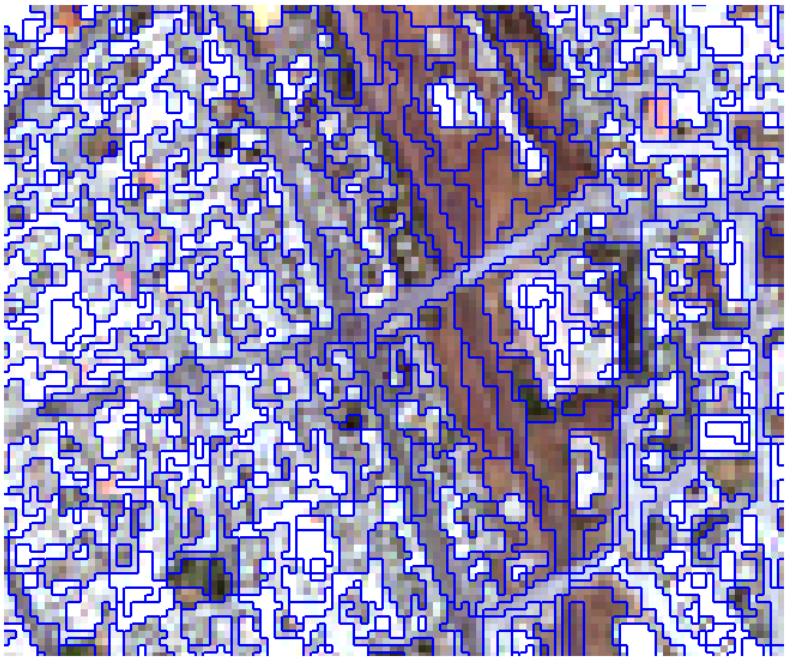


الشكل 5 نتائج التصنيف القائم على الكائن : (a) باستخدام شجرة القرار (DT). (b) باستخدام الشجرة العشوائية (RT). (c) باستخدام الجار الأقرب (KNN). (d) باستخدام خوارزمية بايز (Bayes).

3.1 مقارنة وتقييم طرق التصنيف

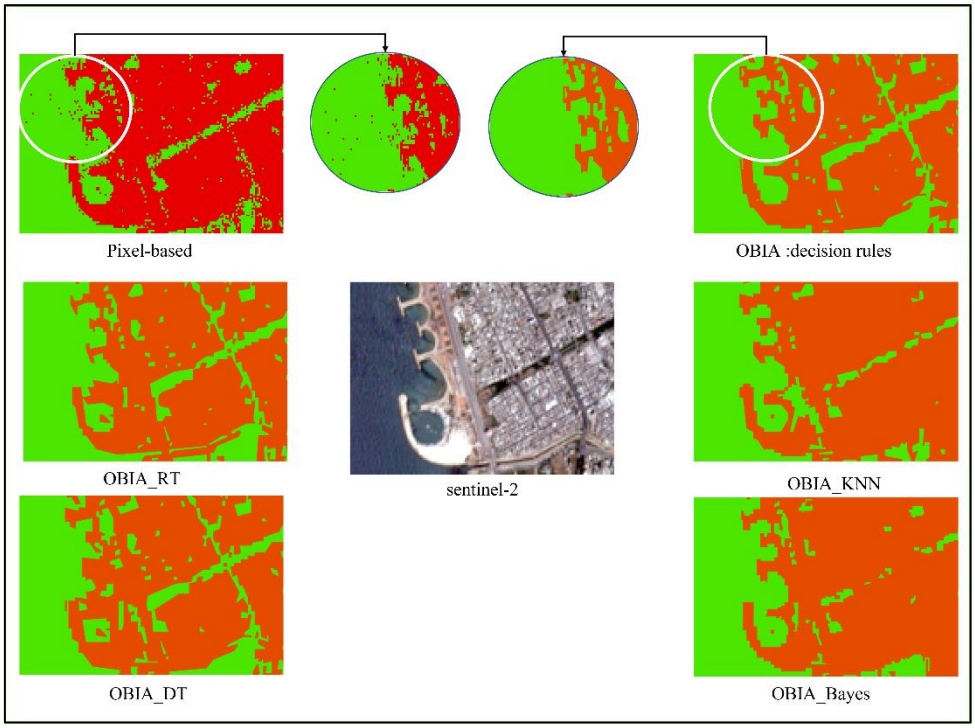
3.1.1 التقييم البصري

نجحت عملية التجزئة متعددة الدقة التي تم تنفيذها بواسطة برنامج (eCognition) في تحديد ملامح المعالم والمناظر الطبيعية والمبنية من خلال تقسيمها إلى كائنات، بحيث يتميز كل كائن بصفاته الطيفية والمكانية والنسيجية، مثل الكتل السكنية، الطرق، الأراضي البور، وأحزمة الأشجار في جزر الطرق والحدائق. تُعد التجزئة خطوة مهمة في التصنيف القائم على الكائن لأنها تتيح تحليل العلاقات المكانية بين الكائنات، وهو ما لا يتوفر في التصنيف القائم على البكسل. في التصنيف القائم على البكسل، يتم تصنيف كل بكسل بشكل فردي بناءً على خصائصه الطيفية دون النظر إلى السياق المكاني أو العلاقات مع البكسلات المجاورة. يُظهر الشكل 6 جزءا من تجزئة (eCognition) المعروضة على صور (Sentinel-2) الأصلية، مما يوضح أن التجزئة توفر تمثيلا تفصيليا مقبولا للمعالم والمناظر الطبيعية والمبنية.



الشكل 6 مقطع من تجزئة eCognition متعددة المستويات على صور Sentinel-2. عرض المشهد حوالي 1.5 كم.

من أجل المقارنة والتقييم البصري لنتائج التصنيف بالطرق المختلفة، تم مقارنة مشهد في جزء من منطقة الدراسة كما هو موضح في الشكل (7). حيث يمكن ملاحظة أن معظم طرق التصنيف، سواء القائم على البكسل أو الهدفي، قدمت ملامح وتفاصيل شبه مطابقة للواقع. ومع ذلك، فإن طريقة شجرة القرار لم تكن فعالة، حيث شوهت المنظر ولم تقدم تفاصيل واضحة بشكل جيد. كما تبين أن طريقة قواعد القرار أعطت أفضل تمثيل بصري وتطابقا معقولًا مع الواقع. يمكن أيضا ملاحظة أن التصنيف القائم على البكسل قدم مشهدا قريبا من الواقع، لكنه يعاني من مشكلة وجود بكسلات متفرقة لأنه يعتمد على تحليل كل بكسل بصورة فردية بناءً على خصائصه الطيفية. هذه الطريقة يمكن أن تقدم نتائج جيدة في بعض الحالات، لكنها غالبا ما تعاني من مشكلة الضجيج (salt-and-pepper). هذا التأثير يتمثل في ظهور بكسلات متفرقة تؤدي إلى تشويه الصورة والتقليل من دقة الحواف والتفاصيل، مما يجعل المشهد أقل واقعية.



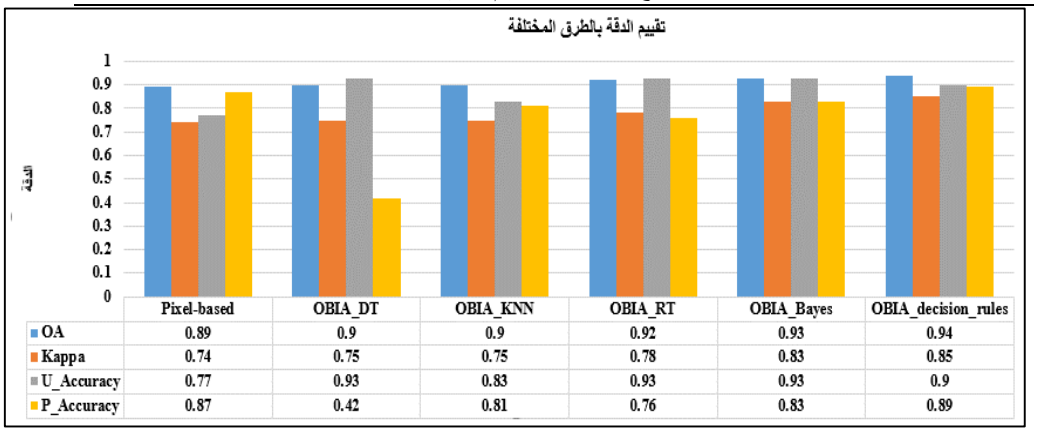
الشكل7 التقييم البصري لنتائج التصنيف في جزء من منطقة الدراسة (1.5*1كم)

3.2 تقييم الدقة

لمقارنة نتائج طرق التصنيف، قمنا بتقييم دقة النتائج عن طريق حساب الدقة الكلية(OA)، دقة المستخدم (U_Accuracy)، دقة المنتج (P_Accuracy)، ومعامل كابا (Kappa) لكل طريقة. قمنا بإنشاء طبقة من النقاط العشوائية باستخدام طريقة الاعتيان العشوائي الطبقي، مستعينين بأداة متاحة في ArcGIS Pro، بلغ عدد نقاط العينة (600) نقطة، منها (172) نقطة لصنف المناطق المبنية و(428) لصنف المناطق غير المبنية. تم اختيار هذه المنهجية لضمان تمثيل الفئات المختلفة بما يتناسب مع توزيعها الحقيقي في منطقة الدراسة. نظرا لأن المناطق غير المبنية تشغل حيزا أكبر من المساحة الكلية، فقد تم تخصيص لها عددا أكبر من النقاط لتعزيز دقة التقييم وضمان التوزيع العادل بين الفئتين. تم الحصول على البيانات المرجعية (الأرضية) بصريا من صور الأقمار الصناعية عالية الدقة في (Google Earth) لنفس فترة المرئية الفضائية، يعتبر هذا الخيار مثاليا نظرا لدقته المكانية العالية وقدرته على تغطية مساحات واسعة، مما يجعله أداة شائعة الاستخدام في الدراسات من هذا النوع. بعد ذلك، تم بناء مصفوفة الخطأ (Confusion Matrix) لكل طريقة. يظهر الجدول(2) والشكل (8) النتائج وقيم الدقة للطرق المختلفة.

الجدول 4: تقييم الدقة بالطرق المختلفة

	OA	Kappa	U_Accuracy	P_Accuracy
Pixel-based	0.89	0.74	0.77	0.87
OBIA_DT	0.9	0.75	0.93	0.42
OBIA_KNN	0.9	0.75	0.83	0.81
OBIA_RT	0.92	0.78	0.93	0.76
OBIA_Bayes	0.93	0.83	0.93	0.83
OBIA_decision_rules	0.94	0.85	0.9	0.89



	Pixel-based	OBIA_DT	OBIA_KNN	OBIA_RT	OBIA_Bayes	OBIA_decision_rules
OA	0.89	0.9	0.9	0.92	0.93	0.94
Kappa	0.74	0.75	0.75	0.78	0.83	0.85
U_Accuracy	0.77	0.93	0.83	0.93	0.93	0.9
P_Accuracy	0.87	0.42	0.81	0.76	0.83	0.89

الشكل 8 الدقة الكلية، دقة المستخدم، دقة المنتج، ومعامل كابا لكل طريقة

من الجدول 2 والشكل 8 يمكن ملاحظة ما يلي:

أظهر التصنيف القائم على البكسل بخوارزمية الاحتمالية العظمى أداءً جيدا نسبيا، مع دقة كلية بلغت (89%) ومعامل كابا (0.74)، مما يشير إلى فعاليته في تحديد المناطق المبنية. ومع ذلك، فإن دقة المستخدم كانت (77%)، وهي قليلة نوعا ما، مما يدل على وجود بعض الأخطاء في التصنيف.

أما بالنسبة للتصنيف القائم على الكائن باستخدام شجرة القرار، فقد تحسنت الدقة الكلية إلى (90%) ومعامل كابا إلى (0.75) مقارنةً بالتصنيف القائم على البكسل، لكن دقة المنتج كانت منخفضة بشكل كبير عند (42%)، مما يشير إلى أن التصنيف قد يحدد بشكل خاطئ بعض المناطق غير المبنية كمبنية.

وفي حالة استخدام خوارزمية الجار الأقرب (KNN)، كانت الدقة الكلية (90%) ومعامل كابا (0.75)، مع دقة المستخدم (83%) ودقة المنتج (81%). يعرض (KNN) توازنا جيدا بين دقة المستخدم ودقة المنتج، مما يجعله خيارا مناسبا للتصنيف القائم على الكائن مع تحسين الأداء العام للتصنيف.

حققت الشجرة العشوائية أداءً أفضل من شجرة القرار والجار الأقرب، حيث بلغت الدقة الكلية (92%) ومعامل كابا (0.78)، مما يشير إلى قدرتها العالية على التمييز بين المناطق المبنية وغير المبنية، مع دقة المستخدم (93%) ودقة المنتج (76%).

أما تصنيف بايز فأظهر أداءً قويا مع دقة كلية (93%) ومعامل كابا (0.83)، بالإضافة إلى توازن جيد بين دقة المستخدم (93%) ودقة المنتج (83%).

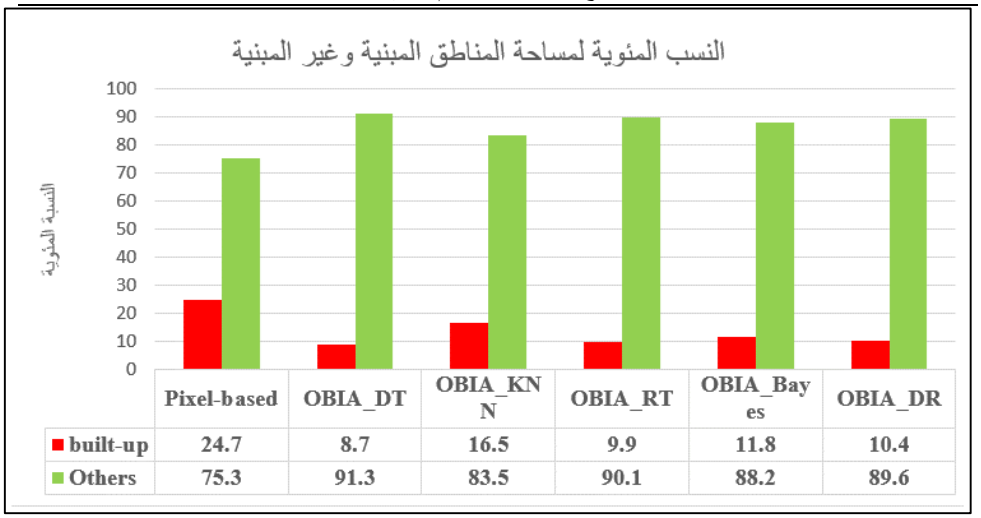
أخيرا، حققت قواعد القرار (Decision Rules) أعلى أداء بين جميع الطرق، مع دقة كلية (94%) ومعامل كابا (0.85)، مما يشير إلى فعاليتها في تحديد وتصنيف المناطق المبنية بدقة عالية، حيث بلغت دقة المستخدم (90%) ودقة المنتج (89%).

يظهر من النتائج أن التصنيف القائم على الكائن يتفوق بشكل عام على التصنيف القائم على البكسل في معظم الحالات، خاصة مع استخدام قواعد القرار وخوارزمية بايز. تعكس هذه النتائج قدرة التصنيف القائم على الكائن على الاستفادة من المعلومات السياقية والخصائص الهيكلية للمناطق المبنية، مما يحسن من دقة التصنيف.

من ناحية أخرى، تظهر بعض الطرق مثل الشجرة العشوائية والجار الأقرب أداءً متوازناً، مما يجعلها بدائل جيدة عند الحاجة لتوازن بين دقة المستخدم ودقة المنتج.

3.3 مقارنة النسب المئوية لمساحات نتائج التصنيف

الهدف من هذه المقارنة هو تقييم التباين النسبي بين طرق التصنيف المختلفة باستخدام نموذج مرجعي يعتمد على أعلى دقة تم تحقيقها، من خلال مقارنة النسب المئوية لمساحات المناطق المبنية وغير المبنية الناتجة عن كل طريقة. توفر هذه المقارنة فهما أعمق لكيفية اختلاف التصنيفات وتوزيع المساحات النسبية، مما قد يكون مفيدا في تحديد الميل نحو المبالغة أو التقليل في تصنيف المناطق المبنية، حتى في غياب بيانات مرجعية ميدانية. نظرا لكون طريقة قواعد القرار هي الأفضل دقةً، اعتمدنا عليها كمرجع لنقارن بها نسب مساحة المناطق المبنية وغير المبنية في منطقة الدراسة في كل طريقة من طرق التصنيف. تم حساب مساحة هذه المناطق لكل طريقة، ثم حساب نسبها من المساحة الكلية لمنطقة الدراسة. يُظهر الشكل (9) النسب المئوية لمساحة المناطق المبنية وغير المبنية لكل طريقة.



	Pixel-based	OBIA_DT	OBIA_KNN	OBIA_RT	OBIA_Bayes	OBIA_DR
built-up	24.7	8.7	16.5	9.9	11.8	10.4
Others	75.3	91.3	83.5	90.1	88.2	89.6

الشكل 9 النسب المئوية لمساحة المناطق المبنية وغير المبنية لكل طريقة

يتضح من الشكل 9 أن نسبة المناطق المبنية في طريقة قواعد القرار والتي تُعتبر المرجع الأكثر دقة بلغت 10.4% مقارنة بالمناطق غير المبنية التي بلغت 89.6%. الطريقة القائمة على البكسل أظهرت أعلى نسبة للمناطق المبنية (24.7%) مقارنة بالطرق الأخرى، مما يشير إلى احتمالية زيادة تقدير المناطق المبنية. في المقابل، تُظهر طريقة OBIA_DT أدنى نسبة للمناطق المبنية (8.7%)، مما قد يدل على تقليل تقدير هذه المناطق. تتراوح نسب المناطق المبنية في الطرق الأخرى مثل OBIA_KNN (16.5%) وOBIA_RT (9.9%) وOBIA_Bayes (11.8%) مما يعكس تباينا في دقة التصنيفات. يشير هذا التباين إلى اختلافات في كيفية تعريف ومعالجة المعالم المبنية في كل طريقة، مما يتطلب دراسة خصائص كل طريقة لاختيار الأنسب وفقا للهدف والبيئة المدروسة.

4 الاستنتاجات والتوصيات

أظهرت الدراسة تفوق التصنيف القائم على الكائن في استخراج المناطق المبنية باستخدام صور القمر الصناعي (Sentinel-2B)، حيث حقق التصنيف باستخدام قواعد القرار أعلى دقة كلية بنسبة (94%) ومعامل كابا(0.85). يعكس هذا الأداء قدرة التصنيف القائم على الكائن على الاستفادة من المعلومات السياقية والخصائص الهيكلية للمناطق المبنية، مما

يزيد من دقة وموثوقية النتائج. بالمقارنة، سجل التصنيف القائم على البكسل دقة كلية بلغت (89%) ومعامل كابا (0.74)، مما يشير إلى بعض القيود في التعامل مع التعقيدات المكانية والطيفية.

كشفت النتائج عن اختلافات في التقديرات المساحية للمناطق المبنية، حيث أظهرت طريقة التصنيف القائم على البكسل نسبة بلغت (24.7%)، في حين قدمت قواعد القرار نسبة أقرب للواقع عند (10.4%). يبرز هذا التباين أهمية اختيار الطريقة المناسبة بناءً على الهدف من الدراسة وخصائص المنطقة المدروسة.

توصي الدراسة بتبني التصنيف القائم على الكائن في التطبيقات التي تتطلب دقة عالية وتفاصيل سياقية للمناطق المبنية، خاصة في البيئات المعقدة والمتنوعة. يمكن أن يتم دمج تقنيات التعلم العميق لتحسين الأداء والنتائج. ومن أجل زيادة دقة وفاعلية التصنيف، توصي الدراسة بإجراء أبحاث إضافية تستكشف تأثير العوامل البيئية والجغرافية على دقة التصنيف في مناطق مختلفة.

5 المراجع

[1] S. Vigneshwaran and S. Vasantha Kumar, “Extraction of built-up area using high resolution Sentinel-2A and Google satellite imagery,” in International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 2018. doi: 10.5194/isprs-archives-XLII-4-W9-165-2018.

[2] H. Darwishe, A. Mohammad, and F. Chaaban, “ Developing a Model of Deep Learning by ANNs for Urban Areas Extraction from Remote Sensing Images- Study Area: Homs-Tartous,” Al-Baath University Journal, vol. 43, no. 7. (Published in Arabic), 2021.

[3] L. Bruzzone and D. F. Prieto, “Unsupervised retraining of a maximum likelihood classifier for the analysis of multitemporal remote sensing images,” IEEE Transactions on Geoscience and Remote Sensing, vol. 39, no. 2, 2001, doi: 10.1109/36.905255.

[4] N. Gupta and H. S. Bhadauria, “Object based Information

Extraction from High Resolution Satellite Imagery using eCognition," IJCSI International Journal of Computer Science Issues, vol. 11, no. 3, 2014.

[5] J. A. Richards, Remote sensing digital image analysis. 2022. doi: 10.1007/978-3-030-82327-6.

[6] Trimble Germany GmbH, "Trimble Documentation eCognition Developer 9.5 Reference Book," Trimble Germany GmbH: Munich, Germany, 2019.

[7] A. M. El-naggar, "Determination of optimum segmentation parameter values for extracting building from remote sensing images," 2018. doi: 10.1016/j.aej.2018.10.001.

[8] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A review of supervised object-based land-cover image classification," 2017. doi: 10.1016/j.isprsjprs.2017.06.001.

[9] M. C. Maloney, S. J. Becker, A. W. H. Griffin, S. L. Lyon, and K. Lasko, "Automated Built-Up Infrastructure Land Cover Extraction Using Index Ensembles with Machine Learning, Automated Training Data, and Red Band Texture Layers," Remote Sens (Basel), vol. 16, no. 5, 2024, doi: 10.3390/rs16050868.

[10] R. Saini et al., "Improving Built-up Extraction Using Spectral Indices and Machine Learning on Sentinel-2 Satellite Data in Mumbai Suburban District, India.," Traitement du Signal, vol. 41, no. 3, 2024.

[11] W. Zhu, C. Yuan, Y. Tian, Y. Wang, L. Li, and C. Hu, "A New High-Resolution Rural Built-Up Land Extraction Method Based on Artificial Surface Index with Short-Wave Infrared Downscaling," Remote Sens (Basel), vol. 16, no. 7, p. 1126, 2024.

[12] K. Sideris, Review of image segmentation algorithms for analysing Sentinel-2 data over large geographical areas. JNCC, 2020.