

Evaluation of the Impact of Spatial Resolution of Satellite Imagery and Supervised Classification Algorithms on the Accuracy of Land Cover Maps Case study: Latakia Governorate

Kholood Khaled Aldabel*

(Received 1 / 11 / 2023. Accepted 14 / 12 / 2023)

□ ABSTRACT □

Remote sensing technology is considered one of the most effective methods for obtaining land cover maps, which are produced by applying satellite image classification techniques that are influenced by factors.

In this research, the impact of the type of classification algorithm and the spatial resolution of the satellite image used on the accuracy of the land cover map for a part of Latakia Governorate was studied. The supervised classification was applied with Maximum Likelihood, Mahalanobis distance, and minimum distance algorithms on Landsat 8 satellite image and Sentinel A image after resampling at resolutions of 15m and 10m, respectively. The results showed that the Maximum Likelihood algorithm is the most accurate in classifying the Sentinel 2A satellite imagery for land cover mapping, with an overall accuracy of 88.2% and a kappa coefficient of 0.81. As for Landsat 8 imagery, the Minimum Distance algorithm was the most accurate, with an overall accuracy of 83.3% and a kappa coefficient of 0.79. On the other hand, the results indicated an improvement in the overall accuracy values for all algorithms with increasing spatial resolution of the classified imagery. The overall accuracy increased from 80.1% to 88.2% for the Maximum Likelihood algorithm, from 78.2% to 80.9% for the Mahalanobis Distance algorithm, and from 83.3% to 85.3% for the Minimum Distance algorithm. We also observed an improvement in the kappa coefficient values from 0.75 to 0.81 for the Maximum Likelihood algorithm, from 0.71 to 0.75 for the Mahalanobis Distance algorithm, and from 0.79 to 0.80 for the Minimum Distance algorithm. Finally, the results indicated that with increased image spatial accuracy, there is less overlap between class ranges, resulting in a reduction in misclassification of certain pixels.

Key words: Supervised classification, Image resolution, Land cover, Overall accuracy, Kappa coefficient.

Copyright



:Tishreen University journal-Syria, The authors retain the copyright under a CC BY-NC-SA 04

* Academic Assistant, Department of Topographic Engineering, Faculty of Civil Engineering, Tishreen University, Lattakia, Syria. kholoodkhaledaldbl@ @tishreen.edu.sy

تقييم تأثير دقة التمييز الهندسية للمرييات الفضائية وخوارزميات التصنيف المراقب على دقة خرائط الغطاء الأرضي - حالة دراسة: محافظة اللاذقية

خلود خالد الدبل*

(تاريخ الإيداع 1 / 11 / 2023. قُبِلَ للنشر في 14 / 12 / 2023)

□ ملخص □

تعتبر تقنية الاستشعار عن بعد من أكثر الطرق فعالية في الحصول على خرائط الغطاء الأرضي والتي يتم إنتاجها بتطبيق تقنية تصنيف المرييات الفضائية وهي عملية تتأثر دقتها بالعديد من العوامل. تم في هذا البحث دراسة تأثير نوع خوارزمية التصنيف المراقب ودقة تمييز الصورة الفضائية المستخدمة على دقة خريطة الغطاء الأرضية لجزء من محافظة اللاذقية، حيث تم تطبيق التصنيف المراقب بخوارزميات الاحتمالية العظمى، مسافة ماهالانوبيس Mahalanobis والمسافة الدنيا على مرئية القمر الصناعي Landsat 8 وعلى مرئية Sentinel A بعد عملية إعادة الاعتيان عند الدقة 15 m و 10 m على الترتيب. بينت النتائج أن خوارزمية الاحتمالية العظمى هي الأدق في تصنيف المرئية الفضائية Sentinel 2A لأغراض تخريط الغطاء الأرضية، حيث بلغت قيمة الدقة الكلية للتصنيف 88.2% مع معامل كبا مساو لـ 0.81. أما بالنسبة للمرئية Landsat 8 فقد كانت خوارزمية المسافة الدنيا هي الأدق حيث بلغت الدقة الكلية للتصنيف القيمة 83.3% مع معامل كبا مساو لـ 0.79. من ناحية أخرى، بينت النتائج تحسن قيم الدقة الكلية للتصنيف لكل الخوارزميات مع زيادة دقة التمييز المكانية للمرئية المصنفة، حيث تحسنت قيمة الدقة الكلية للتصنيف من 80.1% إلى 88.2% للاحتمالية العظمى ومن 78.2% إلى 80.9% لخوارزمية مسافة Mahalanobis ومن 83.3% إلى 85.3% لخوارزمية المسافة الدنيا. كما لاحظنا تحسن قيم المعامل كبا من 0.75 إلى 0.81 للاحتمالية العظمى ومن 0.71 إلى 0.75 لخوارزمية مسافة Mahalanobis ومن 0.79 إلى 0.80 لخوارزمية المسافة الدنيا. وأخيراً، أوضحت النتائج أنه مع زيادة دقة تمييز المرييات، يوجد تداخل أقل بين مدى الفئات بين مما قلل من تصنيف بعض الخلايا تصنيفاً خاطئاً.

الكلمات المفتاحية: تصنيف مراقب، دقة تمييز مكانية، غطاء أرضي، دقة كلية، المعامل كبا.



حقوق النشر : مجلة جامعة تشرين - سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص

CC BY-NC-SA 04

*قائم بالأعمال . قسم الهندسة الطبوغرافية . كلية الهندسة المدنية . جامعة تشرين . اللاذقية . سورية.

بريد الكتروني: kholoodkhaledaldbl@tishreen.edu.sy

مقدمة:

تتمتع خرائط استعمال الأراضي وتغطيتها Land Use/Land Cover (LULC) maps بأهمية خاصة في تخطيط وإدارة الأراضي والموارد (Thanh Noi & Kappas, 2017)، وهي خرائط تظهر التغطية الحيوية للأرض، في حين تظهر خرائط استخدام الأراضي الأنشطة البشرية في نوع محدد من التغطية (Yang et al., 2017)، (Tien Bui et al., 2019). هذا، وتتوافر عدة طرق لإنتاج خرائط LULC ولكن استخدام مرئيات فضائية مستشعرة عن بعد يملك ميزات التطبيق على نطاق واسع باستخدام أجزاء مختلفة من الطيف الكهرطيسي Electromagnetic spectrum لتمثيل الظواهر Features إضافة إلى التكلفة المنخفضة وسرعة التحليل (خاصة للمناطق الممتدة) (Shahabi et al., 2020).

يتوافر العديد من الخوارزميات من أجل تصنيف المرئيات الفضائية (Halder et al., 2011) وهذه الخوارزميات تنتمي إلى التصنيف المراقب Supervised classification، غير المراقب Unsupervised classification أو إلى تحليل المرئيات القائم على العنصر Object-based image analysis ولكل منها نموذج رياضي إحصائي خاص به. تختلف نتائج التصنيف حسب المصنف المطبق فالمصنّفات غير المراقبة لا تتطلب معرفة أي شيء عن منطقة الدراسة وهي تستند في عملها على مواصفات الانعكاس Reflection فقط (Sathya & Malathi, 2011). إن الهدف من التصنيف غير المراقب هو فصل وحدات البكسل ألياً للمرئية الفضائية إلى مجموعات ذات طابع طيفي متمائل، وهي عملية تتم باستخدام واحد من عدة إجراءات إحصائية تسمى بشكل عام "التجميع" Clustering حيث يتم إنشاء فئات البكسل بناءً على البصمات الطيفية المشتركة Spectral signatures. وكأمثلة على خوارزميات التصنيف غير المراقب نذكر: K-means (Al-Doski et al., 2013) و ISODATA (Lillesand & Kiefer, 1999). أما التصنيف المراقب فهو يتطلب استخدام بيانات أو عينات تدريب Training samples والتي تعتبر ممثلة عن كل ظاهرة من الظواهر الأرضية ويتم جمع هذه العينات بناءً على معرفة المستخدم بالمشهد المصور. يملك التصنيف المراقب حسنة عدة مقارنة بغير المراقب، ففي التصنيف المراقب يتم أولاً تمييز أصناف المعلومات المفيدة ومن ثم يتم اختبار الفصل الطيفي بينها. نذكر من خوارزميات التصنيف المراقب: خوارزمية الاحتمالية العظمى Maximum Likelihood، خوارزمية المسافة الدنيا Minimum Distance، مسافة ماهالانوبيس Mahalanobis، الشبكة العصبونية الصناعية Artificial neural network (Haykin, 2009)، الغابة العشوائية Random forest وآلات شعاع الدعم Support vector machines (SVM) (Boser, 1992).

من ناحية أخرى، تلعب دقة التمييز المكانية دوراً مهماً في تحديد دقة التصنيف وذلك لأنها تحدد مستوى التفاصيل التي يستطيع المستخدم ملاحظتها على سطح الأرض. في الواقع، تظهر أهمية دقة التمييز المكانية عند استخدام المرئيات الفضائية عالية الدقة (Zhou & Robson, 2001).

أهمية البحث وأهدافه:

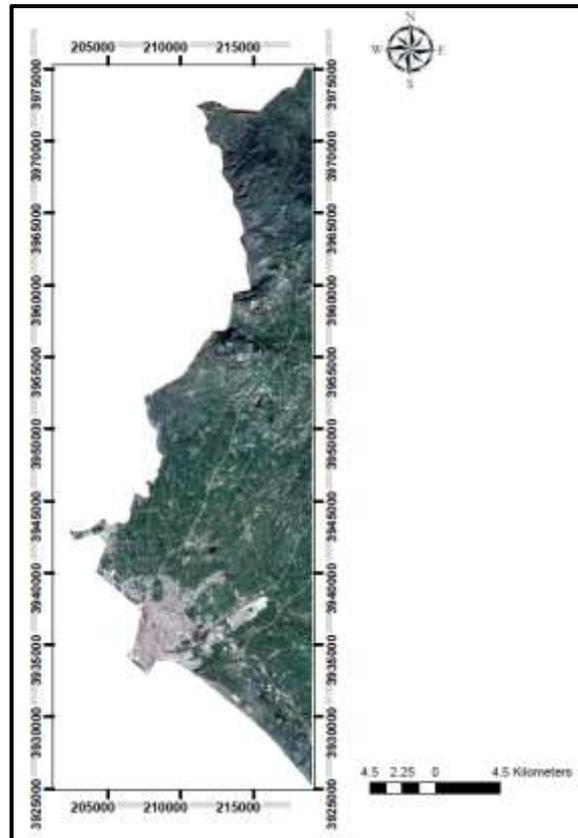
تأتي أهمية البحث من أنه يناقش بعض العوامل المؤثرة على دقة خارطة الغطاء الأرضي التي تعتبر مدخلاً هاماً في العديد من التطبيقات مثل التخطيط وإدارة الموارد ودراسة التوسع العمراني. أما أهداف البحث فهي:

- 1- تقييم تأثير خوارزمية التصنيف المراقب المطبقة على دقة تصنيف مرئيات الأقمار الصناعية والحصول منها على خارطة غطاء الأراضي.
- 2- تقييم تأثير دقة تمييز المرئية على دقة التصنيف المراقب بهدف الحصول على خارطة غطاء أراضي دقيقة. سنستفيد في البحث من إمكانيات المنصات المجانية التي تتيح الوصول إلى قواعد بيانات المرئيات الفضائية المجانية، مثل المنصة Earth Explorer والمنصة scihub.copernicus.eu في الحصول على مرئيات فضائية مجانية Landsat 8-OLI و Sentinel 2A لمنطقة الدراسة (محافظة اللاذقية) حيث سنقوم بتطبيق خوارزميات التصنيف المراقب وهي: الاحتمالية العظمى، مسافة Mahalanobis والمسافة الدنيا على هذه المرئيات وذلك لمعرفة تأثير هذين العاملين على دقة التصنيف وعلى خارطة غطاء الأراضي لمنطقة الدراسة.

طرائق البحث ومواده:

1- منطقة الدراسة

منطقة الدراسة تشكل جزءاً من محافظة اللاذقية الواقعة على الساحل السوري. تمتد هذه المنطقة (الشكل 1) بين خطي العرض $35^{\circ}15.608'$ جنوباً و $35^{\circ}55.127'$ شمالاً وبين خطي الطول $36^{\circ}18.196'$ شرقاً و $35^{\circ}42.290'$ غرباً وتبلغ مساحتها التقريبية 407 كيلومتر مربع تقريباً. المنطقة ذات تضاريس متنوعة وهي ذات نسيج عمراني مؤلف من المدينة ومن الضواحي والقرى المحيطة بها ومن مناطق خضراء وتكشفات صخرية مع وجود بعض المسطحات المائية.

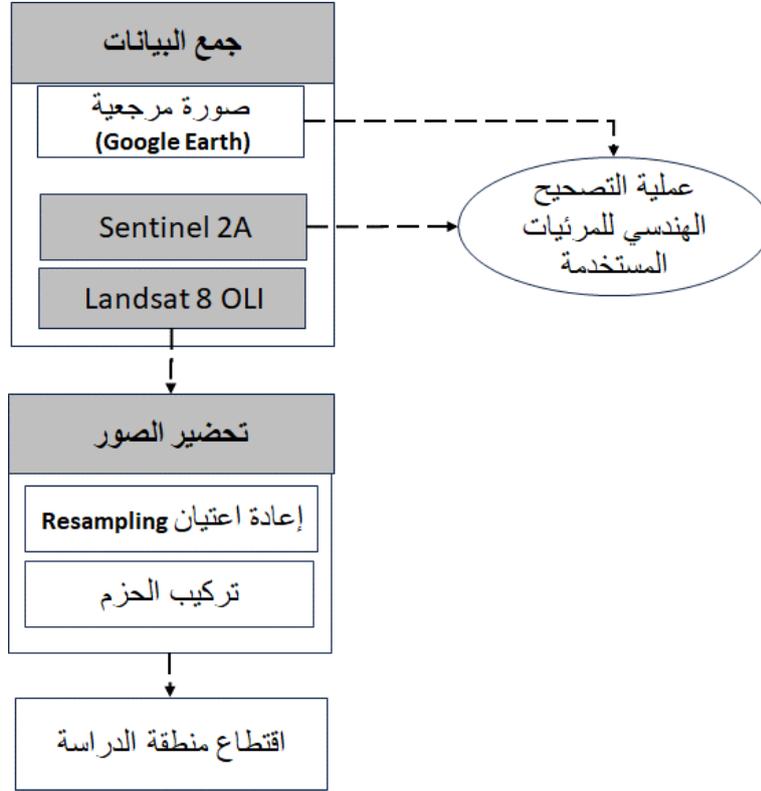


الشكل (1). حدود منطقة الدراسة.

2- منهجية البحث

مرت منهجية البحث بالخطوات التالية:

- المرحلة التحضيرية: نوضح في الشكل (2) الخطوات المطبقة في المرحلة التحضيرية للبيانات الخاصة بمنطقة الدراسة.



الشكل (2). خطوات المرحلة التحضيرية.

أولاً: جمع البيانات

البيانات عبارة عن ثلاث مرئيات فضائية تغطي منطقة الدراسة:

- 1- المرئية الأولى من النوع Sentinel 2A ملتقطة بتاريخ 10/09/2023 وتضم 13 حزمة طيفية ملتقطة باستخدام المستشعر MSI بدقة تمييز تساوي 10 m (أربع حزم هي الحزم المرئية وحزمة تحت الأحمر القريب)، 20 m (ست حزم هي حزمة الحافة الحمراء Red edge وحزم تحت الأحمر قصير الموجة) و60 m (ثلاث حزم لتصحيح تأثيرات الغلاف الجوي) (الجدول (1)). تم تنزيل هذه المرئية مجاناً من موقع وكالة الفضاء الأوروبية <https://scihub.copernicus.eu/dhus/#/home>. هذه المرئية معطاة ضمن نظام الإحداثيات العالمي WGS_1984_UTM_Zone_37N.

الجدول (1). أهم الحزم الطيفية للمرنيات Sentinel 2A.

دقة التمييز (متر)	طول الموجة (ميكرومتر)	الحزم Bands
60	0.443	Band 1 – Coastal aerosol السواحل
10	0.490	Band 2 – Blue الأزرق
10	0.560	Band 3 – Green الأخضر
10	0.665	Band 4 – Red الأحمر
20	0.705	Band 5 – Visible and Near Infrared (VNIR) تحت الحمراء المرئية والقريبة
20	0.740	Band 6 – Visible and Near Infrared (VNIR) تحت الحمراء المرئية والقريبة
20	0.783	Band 7 – Visible and Near Infrared (VNIR) تحت الحمراء المرئية والقريبة
10	0.842	Band 8 – Visible and Near Infrared (VNIR) تحت الحمراء المرئية والقريبة
20	0.865	Band 8a – Visible and Near Infrared (VNIR) تحت الحمراء المرئية والقريبة
60	0.940	Band 9 –SWIR تحت الحمراء المتوسطة
60	1.375	Band 9 –SWIR تحت الحمراء المتوسطة
60	1.610	Band 9 –SWIR تحت الحمراء المتوسطة
20	2.190	Band 9 –SWIR تحت الحمراء المتوسطة

2- المرئية الثانية من النوع Landsat 8-OLI ملتقطه بتاريخ 10/09/2023 وتضم 11 حزمة طيفية دقة تمييزها 30 m ماعدا الحزمة البانوكروماتية Panchromatic (الحزمة 8) التي تبلغ دقة تمييزها 15 m (الجدول (2)). تم تحميل هذه المرئية من موقع Earth Explorer (<https://earthexplorer.usgs.gov>) التابع لهيئة المسح الجيولوجي الأمريكية USGS. هذه المرئية معطاة ضمن نظام الإحداثيات العالمي WGS_1984_UTM_Zone_37N.

الجدول (2). أهم الحزم الطيفية للمرنيات Landsat 8.

دقة التمييز (متر)	طول الموجة (ميكرومتر)	الحزم Bands
30	0.45-0.43	Band 1 – Coastal aerosol السواحل
30	0.51-0.45	Band 2 – Blue الأزرق
30	0.59-0.53	Band 3 – Green الأخضر
30	0.67-0.64	Band 4 – Red الأحمر

30	0.88–0.85	Band 5 – Near Infrared (NIR) تحت الحمراء القريبة (الغطاء النباتي)
30	1.65–1.57	Band 6 –SWIR 1 تحت الحمراء المتوسطة
30	2.29–2.11	Band 7 –SWIR 2 تحت الحمراء المتوسطة
15	0.68–0.50	Band 8 – Panchromatic البانكروماتية
30	1.38–1.36	Band 9 – Cirrus (السحب)
100	11.19–10.60	Band 10 – Thermal Infrared (TIRS) 1 تحت الحمراء الحرارية
100	12.51–11.50	Band 11 – Thermal Infrared (TIRS) 2 تحت الحمراء الحرارية

3- الصورة الثالثة هي صورة بالألوان الحقيقية تم تصويبها بالاستعانة بالبرنامج المجاني Google earth دقة تمييزها المكانية 3 m. تم استخدام هذه الصورة في البحث كصورة مرجعية من أجل تقييم دقة التصنيف المراقب وذلك بعد تصحيحها هندسياً كي تكون ضمن نفس الإحداثيات الخاص بالصورتين السابقتين.

ثانياً: تحضير المرئيات

فيما يخص المرئية Sentinel 2A، تم إنجاز إعادة إعتيان الحزم B08A (المتوافقة مع الحزمة تحت الحمراء المرئية والقريبة VNIR) و B11 (المتوافقة مع الحزمة تحت الحمراء قصيرة الموجة SWIR) و B12 (المتوافقة مع الحزمة تحت الحمراء قصيرة الموجة SWIR) عند دقة التمييز 10 m ومن ثم دمجها مع الحزم B01 (المتوافقة مع الحزمة الزرقاء)، B03 (المتوافقة مع الحزمة الخضراء)، B04 (المتوافقة مع الحزمة الحمراء) و B08 (المتوافقة مع الحزمة تحت الحمراء المرئية والقريبة VNIR). أما في حالة المرئية Landsat 8، فقد أعدنا إعتيان الحزم B02 (المتوافقة مع الحزمة الزرقاء) و B3 (المتوافقة مع الحزمة الخضراء) و B4 (المتوافقة مع الحزمة الحمراء) و B5 (المتوافقة مع الحزمة تحت الحمراء القريبة NIR) و B6 (المتوافقة مع الحزمة تحت الحمراء قصيرة الموجة SWIR) و B7 (المتوافقة مع الحزمة تحت الحمراء قصيرة الموجة SWIR) عند دقة التمييز 15 m ومن ثم دمجها مع الحزمة البانوكروماتية B08. تم بعد ذلك تركيب الحزم Bands composite واقتطاع منطقة الدراسة. تمت هذه العمليات بالاستعانة بالبرنامج ArcGIS 10.8.

ثالثاً: الإرجاع الهندسي للصور

بهدف توحيد نظم إحداثيات الصور الثلاثة، تم تطبيق عملية التصحيح الهندسي Geometric correction للصورة Google Earth لنقلها إلى جملة الإحداثيات WGS_1984_UTM_Zone_37N وذلك باستخدام المرئية Sentinel 2A كمرئية مرجعية، وبمساعدة البرنامج ERDAS IMAGINE 14، تم استخدام 7 نقاط ضبط وتطبيق كثير الحدود من الدرجة الثانية للتصحيح. يوضح الجدول (3) نتائج عمليات الإرجاع لهذه الصورة.

الجدول (3). نتائج إرجاع الصورة Google Earth.

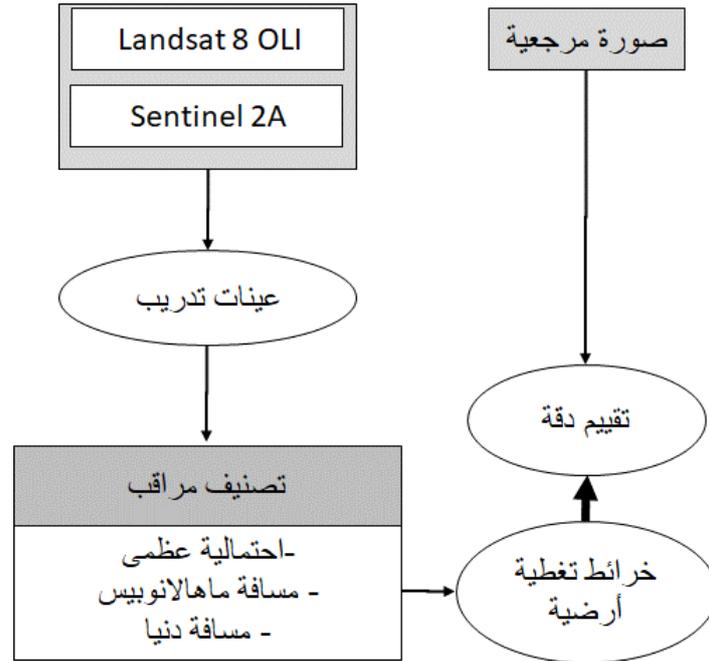
عدد نقاط الضبط	درجة كثير الحدود	الخطأ متوسط التربيع الأفقي RMS_T
7	الثانية	5.38 m

للحكم على دقة التصحيح، يُعطى حد التساهل للإرجاع الهندسي للصور الفضائية بالعلاقة (Alkhalil, 2016):

$$RMS_T \leq (4 \rightarrow 6).GR(\text{in meters}) \quad (1)$$

حيث GR هي دقة التمييز الأرضية للصورة. بتطبيق العلاقة نجد أن حدود التساهل لهذه الصورة تتراوح بين 12 m و 18 m، وبالتالي فالتصحيح مقبول.

• **مرحلة التصنيف:** تم استخدام البرنامج ERDAS IMAGINE 2014 لإنجاز هذه المرحلة، حيث تم تطبيق خوارزميات التصنيف المراقب: الاحتمالية العظمى، مسافة Mahalanobis وخوارزمية المسافة الدنيا. نوضح في الشكل (3) خطوات هذه المرحلة.



الشكل (3). خطوات مرحلة التصنيف.

فيما يخص خوارزمية الاحتمالية العظمى، فيتم في هذه الطريقة حساب قيم الاحتمالية لكل بكسل من بكسلات المرئية بناءً على نظرية الاحتمالات حيث يتم وضع كل خلية في الصنف الأكثر احتمالاً بالاعتماد على قيم انعكاس هذه الخلايا وعليه فإن تطبيق هذا الأسلوب يحتاج إلى عمليات حسابية طويلة نسبياً، خاصةً إذا كانت الحزم المستخدمة في التصنيف كثيرة (Yonezawa, 2007). هذا، وتبنى طريقة الاحتمالية العظمى على احتمال أن الإحصاءات لكل فئة في كل نطاق تتوزع توزيعاً طبيعياً، كما أن كل بكسل ينتمي إلى فئة معينة إلا إذا حددنا عتبة احتمال، لذا تعتبر هذه الطريقة من أكثر الطرق المستعملة في التصنيف المراقب كونها تعمل على الترجيح الأقصى على إنتاج خطوط متساوية الاحتمالية على شكل قطوع ناقصة لكل فئة. أما خوارزمية المسافة الدنيا عن الوسط الحسابي فتعتمد على تحديد المتوسطات الطيفية (الانعكاسية) لكل فئة من فئات التصنيف حيث يتم في منطقة الدراسة لكل فئة من فئات التصنيف حساب متوسطات القيم الرقمية لكل نطاق من النطاقات المستخدمة، وتحديد الوسط في كل فئة (Abburu & Golla, 2015). وأخيراً، تعتبر خوارزمية التصنيف القائمة على مسافة Mahalanobis قريبةً من طريقة التصنيف بأسلوب أقصر مسافة عن الوسط الحسابي، والاختلاف بينهما يتمثل في أن معادلة هذا الأسلوب تستخدم مصفوفة التباين المشترك، حيث يأخذ هذا الأسلوب تباين الطبقات بعين الاعتبار ويعتمد اعتماداً كبيراً على التوزيع الطبيعي من

البيانات في كل الأطياف المدخلة. وينتج عن استخدام مصفوفة التباين في المعادلة أن التجمعات شديدة الاختلاف فيما بينها ستقود الى أصناف شديدة الاختلاف بشكل مشابه والعكس بالعكس، وتستخدم الطريقة احصاءات عن كل فئة، ويتم تعيين البيكسل لأقرب فئة من فئات التدريب إلا إذا قمنا بتحديد عتبة للمسافة، في هذه الحالة يمكن أن تكون بعض البكسلات غير مصنفة إذا لم تلب الحد الأدنى (Jog & Dixit, 2016).

• **مرحلة التحقق من دقة التصنيف:** للتحقق من دقة الخرائط الموضوعية لاستعمالات الأراضي، يتم استخدام مصفوفة الالتباس Confusion matrix أو مصفوفة الخطأ Error Matrix وهي مصفوفة مستخدمة كطريقة كمية لتقييم دقة تصنيف Accuracy assessment المرئيات الفضائية (Jin, 2013). هذه المصفوفة عبارة عن جدول يوضح التوافق بين نتائج التصنيف ومرئية مرجعية للأصناف (الشكل 4). يمثل قطر المصفوفة البيانات المتطابقة مع بعضها والبيانات غير القطرية بالصف هي عدد وحدات المرئية التي تم تصنيفها خطأ أما بالعمود فتمثل البيانات المحذوفة أو المهملة. ولتمثيل مصفوفة الخطأ رياضياً نفترض أن (n) هو عدد من العينات الموزعة على خلايا المصفوفة، بحيث تتسب كل عينة مصنفة إلى إحدى الفئات (k) في الصفوف، وبشكل مستقل تتسب مجموعة البيانات المرجعية إلى الفئات (k) نفسها في الأعمدة، ونفترض أن (j) تمثل الأعمدة في المصفوفة، بحيث $(j=1,2,3.. k)$ هي الأصناف أو الغطاءات الأرضية للبيانات المرجعية. ونفترض أن (i) تمثل الأعمدة في المصفوفة، بحيث $(i=1,2,3.. k)$ وهي الأصناف الغطاءات الأرضية في الخارطة المصنفة.

		العمدة - j			مجموع
		(البيانات المرجعية)			الصفوف
-i الصفوف		1	2	k	n_{i+}
(البيانات المصنفة)	1	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	k	n_{k1}	n_{k2}	n_{kk}	n_{k+}
مجموع الأعمدة	n_{+j}	n_{+1}	n_{+2}	n_{+k}	n

الشكل (4). التمثيل الرياضي لمصفوفة الخطأ

حيث $n_{i+} = \sum_{j=1}^k n_{ij}$ هو مجموع العينات المصنفة ضمن الصنف i ، و $n_{+j} = \sum_{i=1}^k n_{ij}$ هو مجموع العينات المصنفة ضمن الصنف j .

من هذه المصفوفة يمكننا أن نستنتج المعايير التالية للدقة:

• **الدقة الكلية (Overall Accuracy (OA))** والتي تعد الأكثر شيوعاً في تقييم الدقة إحصائياً، وهي مجموع الخلايا القطرية (أي خلايا عينة الاختبار المصنفة بشكل صحيح) مقسومة على العدد الإجمالي للخلايا في مصفوفة الأخطاء بأكملها. ويمكن حسابها على النحو التالي:

$$OA = \frac{\sum_{i=1}^k n_{ii}}{n} \quad (1)$$

• **دقة المستخدم (User Accuracy (UA))**: وهو مقياس لموثوقية التصنيف؛ لأنها تخبر المستخدم بعدد المرات التي سيكون فيها الصنف على الخارطة موجوداً فعلياً على الأرض ويمكن حسابها رياضياً على النحو التالي:

$$UA = \frac{n_{ii}}{n_{i+}} \quad (2)$$

- خطأ الإدراج (CE) Commission Error: بالنسبة لأي صنف، يحدث خطأ التكليف عندما تقوم عملية التصنيف بتعيين وحدات بكسل لصنف معين لا تنتمي إليه في الواقع، وكلما زادت أخطاء الإدراج انخفضت دقة المستخدم. ويمكن حسابه رياضياً على النحو التالي:

$$CE (\%) = 100\% - UA (\%) \quad (3)$$

- دقة المنتج (PA) Procedure's accuracy وهي تعبر عن دقة الخارطة من وجهة نظر صانعيها (المنتج)، حيث تشير إلى عدد المرات التي تظهر فيها البيانات الحقيقية على الأرض بشكل صحيح على الخارطة المصنفة، وهي تساوي عدد المواقع المرجعية المصنفة بدقة مقسومة على العدد الإجمالي للمواقع المرجعية لذلك الصنف. ونقيس دقة المنتج مدى دقة تصنيف منطقة معينة ويمكن حسابها كما يأتي:

$$PA = \frac{n_{jj}}{n_{+j}} \quad (4)$$

- خطأ الاستبعاد (OE) Omission Error والذي يشير إلى استبعاد المنطقة من الصنف الذي تنتمي إليه وكلما زادت أخطاء الاستبعاد انخفض مستوى دقة المنتج. ويمكن حسابه كما يأتي:

$$OE (\%) = 100\% - PA (\%) \quad (5)$$

- كما يمكن التعرف على دقة التصنيف باستخدام معامل كبا (k) kappa coefficient والذي يستخدم لقياس العلاقة بين الاتفاق المتوقع لاحقاً (beyond chance agreement) وعدم الاتفاق (الاختلاف) المتوقع (Expected disagreement)، وتستخدم في إيجاده كل عناصر مصفوفة دقة التصنيف وليس فقط العناصر القطرية في المصفوفة. وتم حسابه وفق المعادلة التالية:

$$K = \frac{n \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{n^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (6)$$

- حيث: عدد وحدات المرئية المستخدمة في تقييم الدقة تساوي n وعناصر المصفوفة هي x_{ii} . مجموع عناصر الأعمدة j في الصف i هي $\sum x_{ij} = X_i^+$ مجموع عناصر الصفوف في العمود j هي $\sum x_{ij} = X_{+j}$. يجب أن تكون قيمة $K \geq 0.75$ لنحصل على نتائج جيدة جداً في التصنيف، وإذا كانت قيمة $K < 0.4$ فذاك مؤشر عن دقة متواضعة

النتائج والمناقشة:

1- اقتطاع عينات التدريب

- بالاستعانة بالبرنامج ERDAS IMAGINE، تم أولاً قياس عينات تدريب تعبر عن الأصناف الخاصة بالغطاء الأرضية. تم اختيار العينات في مختلف أنحاء المرئية بحيث يكون عددها كافياً ومناسباً لعمليات التصنيف. أما الأصناف التي تم اعتمادها في دراستنا فهي موضحة في الجدول (2).

الجدول (2). أصناف الغطاء الأرضية المعتمدة في البحث.

وصف	نوع الغطاء الأرضية
المناطق السكنية وشبكة الطرق.	مناطق حضرية/مناطق مبنية (Built-up)
الأراضي الزراعية والأشجار.	مناطق خضراء (Vegetation)
مسطحات مائية/أنهار.	مياه (Water)

2- تصنيف المرئية Sentinel 2A

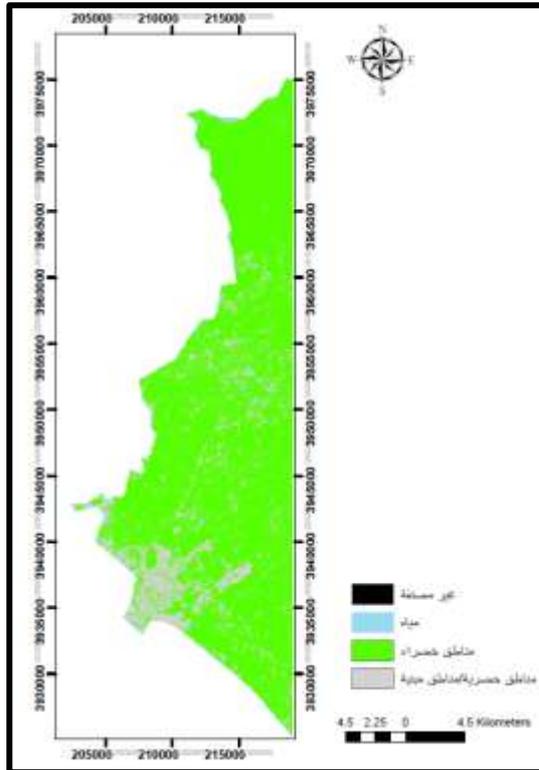
تم تنفيذ تصنيف المرئية الفضائية Sentinel 2A لمنطقة الدراسة باستخدام خوارزمية الاحتمالية العظمى، المسافة الدنيا وخوارزمية مسافة Mahalanobis باستخدام نفس ملف التوقع الطيفية للعينات التي تم قياسها على المرئية. تم قياس نفس العدد من العينات لكل صنف من الأصناف وبلغ عددها في كل صنف 10 عينات. وقبل عملية التصنيف كان لابد من اختبار جودة العينات المعتمدة وذلك عبر حساب مصفوفة التلاصق Contingency matrix بين هذه العينات. يمكننا أن نلاحظ من خلال الشكل (4) أن جميع الفئات التي تم تعيينها تقع بنسبة أكثر من 95% ضمن أصناف مختلفة وبأنه لا يوجد تداخل بين الفئات بنسبة أكثر من 4%، لذلك تعتبر العينات المختارة مقبولة. وبعد قبول هذه العينات من خلال المحلل نقوم بإكمال عملية التصنيف المراقب.

ERROR MATRIX				
Classified Data	Reference Data			Row Total
	Building	Vegetation	Water	
Building	100.00	1.39	0.00	2272
Vegetation	0.00	98.61	1.50	4609
Water	0.00	0.00	98.50	853
Column Total	2207	4661	866	7734

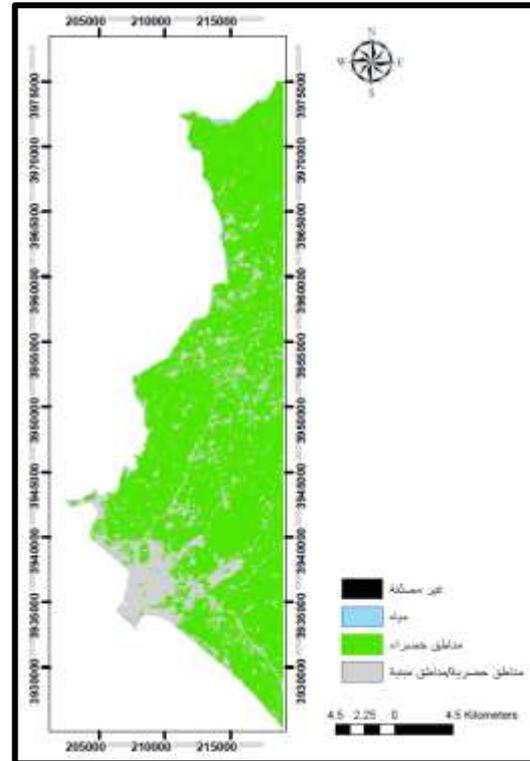
----- End of Error Matrix -----

الشكل (4). تقرير تقييم دقة عينات التدريب للمرئية Sentinel 2A.

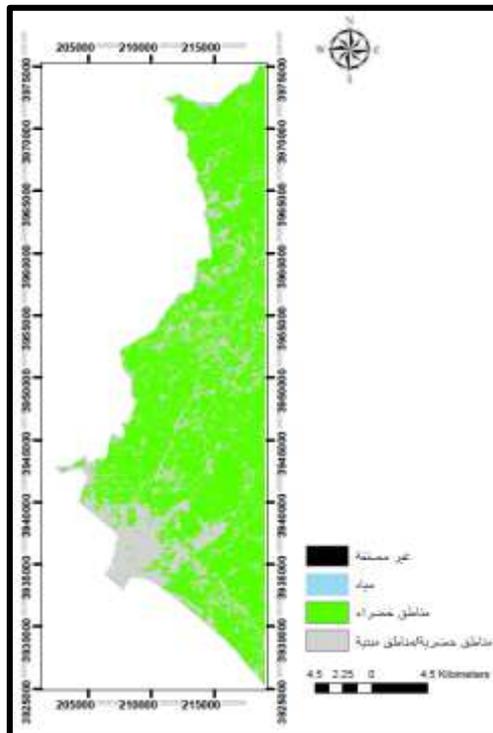
بعد التأكد من دقة عينات التدريب، قمنا بتصنيف المرئية Sentinel 2A بالطرق المختبرة في هذا البحث. توضح الأشكال التالية نتائج هذا التصنيف.



الشكل (6). نتائج المرئية Sentinel 2A باستخدام خوارزمية المسافة الدنيا.



الشكل (5). نتائج تصنيف المرئية Sentinel 2A باستخدام خوارزمية الاحتمالية العظمى.



الشكل (7). نتائج المرئية Sentinel 2A باستخدام خوارزمية مسافة Mahalanobis.

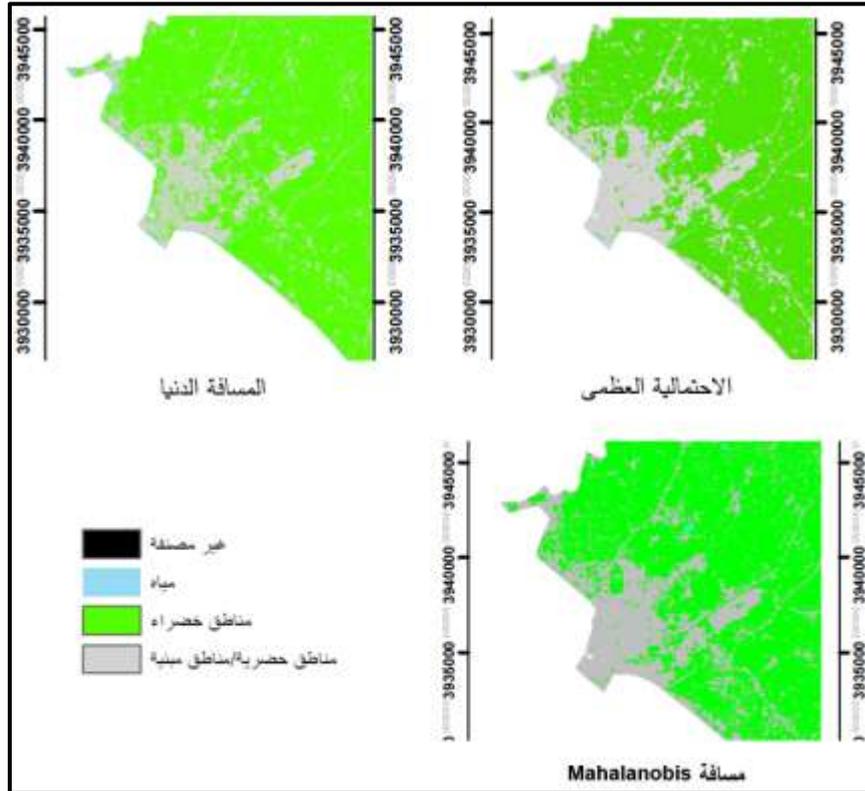
لتقييم دقة التصنيف بالخوارزميات الثلاثة السابقة، استخدمنا المرئية المرجعية التي تغطي منطقة الدراسة لتعريف الأصناف الواقعية ومقارنتها لاحقاً مع الأصناف التي تنتأب بها خوارزميات التصنيف. ولتوليد مصفوفة الالتباس أو

الخطأ، تم استخدام 100 نقطة اختبار عشوائية Random points قيست على الخرائط الموضوعية (المريثات المصنفة) ومن ثم تم فتح الملف ومقارنة هذه النقاط المصنفة بالنقاط الحقيقية من المريثة المرجعية. يبين الجدول (3) نتائج هذا التقييم.

جدول (3): خوارزميات التصنيف المراقب المستخدمة في تصنيف المريثة Sentinel2A مرتبة حسب دقة التصنيف.

المعامل كبا	الدقة الكلية	الخوارزمية
0.81	%88.2	الاحتمالية العظمى
0.80	%85.3	المسافة الدنيا
0.74	%80.9	مسافة Mahalanobis

نلاحظ من الجدول (3) أن أدق الخوارزميات هي خوارزمية الاحتمالية العظمى تليها خوارزمية المسافة الدنيا ثم خوارزمية مسافة Mahalanobis. ومن الناحية البصرية، نلاحظ أن المريثات المصنفة باستخدام خوارزمية الاحتمالية العظمى والمسافة الدنيا متوافقة أكثر مع الواقع مقارنةً بالمريثة الناتجة عن خوارزمية مسافة Mahalanobis والتي صنفت جزء من المناطق الخضراء على أنها مناطق حضرية (الشكل (8)).



الشكل (8). التقارب بين نتائج خوارزميتي الاحتمالية العظمى والمسافة الدنيا وتصنيف خوارزمية مسافة Mahalanobis جزء من المناطق الخضراء على أنها مناطق مبنية.

3- تصنيف المريثة Landsat 8

تم تنفيذ تصنيف المريثة الفضائية Landsat 8 لمنطقة الدراسة باستخدام خوارزمية الاحتمالية العظمى، المسافة الدنيا وخوارزمية مسافة Mahalanobis وذلك باستخدام نفس ملف التوافق الطيفية للعينات التي تم قياسها على المريثة. تم قياس نفس العدد من العينات لكل صنف من الأصناف وبلغ عددها في كل صنف 10 عينات. كما في

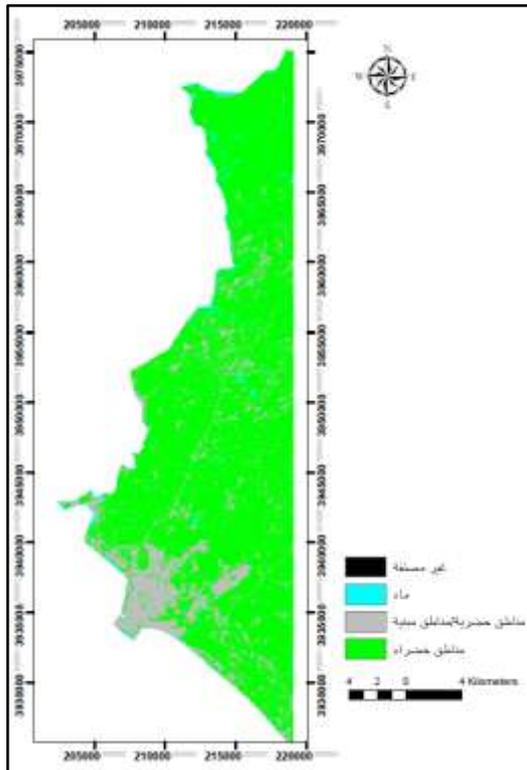
حالة المرئية Sentinel 2A قمنا بحساب مصفوفة التلاصق لتقييم جودة العينات المعتمدة. يمكننا أن نلاحظ من خلال الشكل (9) أن جميع الفئات التي تمّ تعيينها تقع بنسبة أكثر من 95% ضمن أصناف مختلفة وبأنه لا يوجد تداخل بين الفئات بنسبة أكثر من 4%، لذلك تعتبر العينات المختارة مقبولة. وبعد قبول هذه العينات من خلال المُحلل نقوم بإكمال عملية التّصنيف المُراقب.

ERROR MATRIX				
Classified Data	Reference Data			
	Building	Water	Vegetation	Row Total
Building	98.21	0.42	2.80	2052
Water	0.00	99.58	0.00	237
Vegetation	1.79	0.00	97.20	2503
Column Total	2016	238	2538	4792

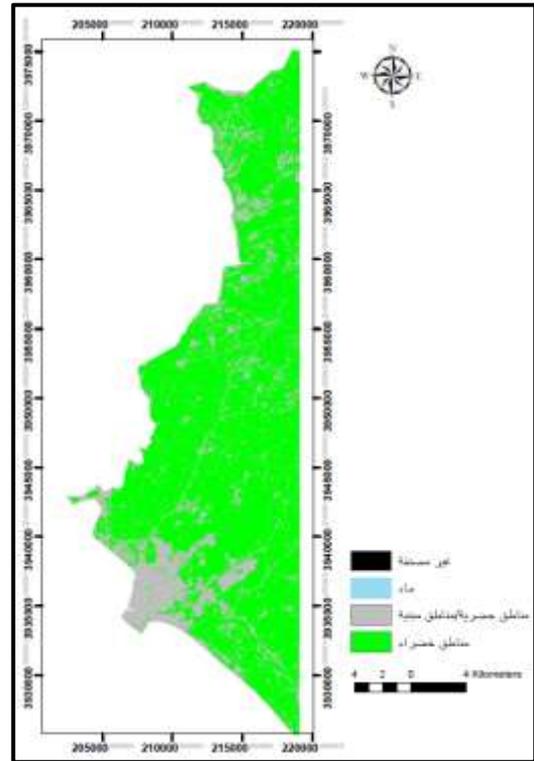
----- End of Error Matrix -----

الشكل (9). تقرير تقييم دقة عينات التدريب للمرئية Landsat 8.

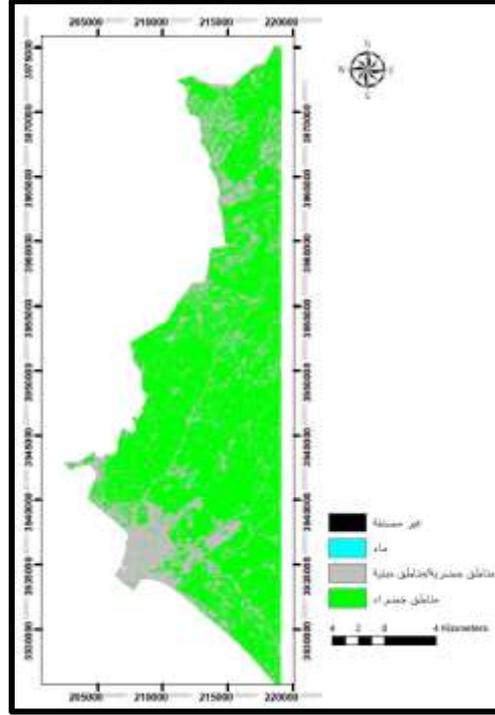
بعد التأكد من دقة عينات التدريب، قمنا بتصنيف المرئية Landsat 8 بالطرق المختبرة في هذا البحث. توضح الأشكال التالية نتائج هذا التصنيف.



الشكل (11). نتائج تصنيف المرئية Landsat 8 باستخدام خوارزمية المسافة الدنيا.



الشكل (10). نتائج تصنيف المرئية Landsat 8 باستخدام خوارزمية الاحتمالية العظمى.



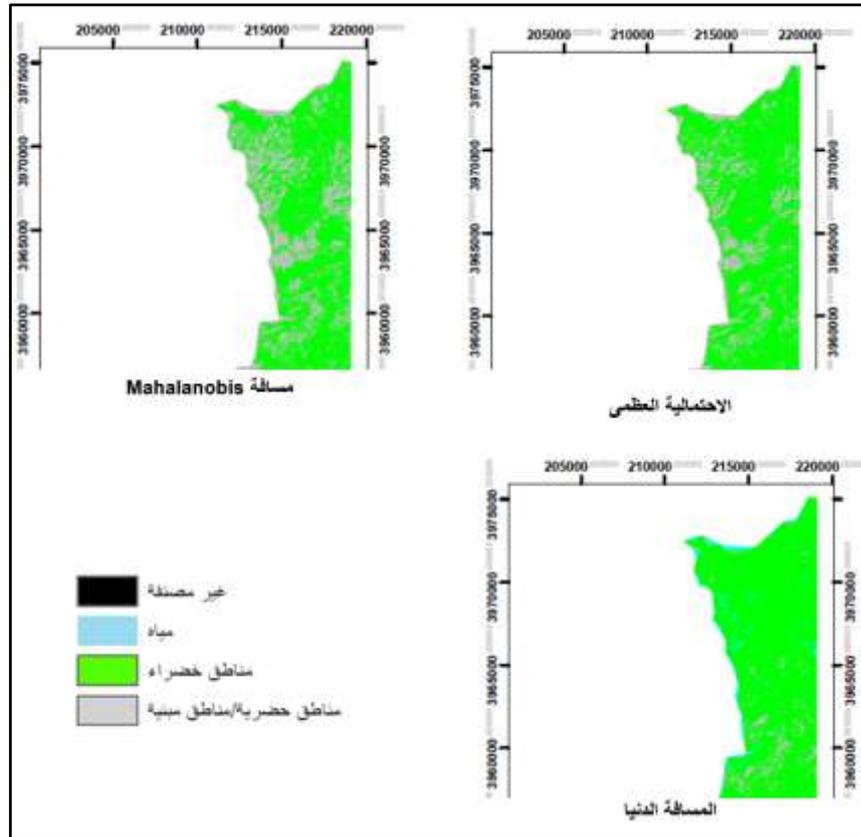
الشكل (12). نتائج تصنيف المرئية Landsat 8 باستخدام خوارزمية مسافة Mahalanobis.

بنفس الطريقة السابقة، تم تقييم دقة التصنيف بالخوارزميات الثلاثة وذلك باستخدام المرئية المرجعية التي تغطي منطقة الدراسة حيث تم توليد مصفوفة الالتباس أو الخطأ باستخدام 100 نقطة اختبار قيست على الخرائط الموضوعية (المرئيات المصنفة) بمعدل 33 نقطة لكل صنف. ومن ثم تم فتح الملف ومقارنة هذه النقاط المصنفة بالنقاط الحقيقية من المرئية المرجعية. يبين الجدول (4) نتائج هذا التقييم.

جدول (4): خوارزميات التصنيف المراقب المستخدمة في تصنيف المرئية Landsat 8 مرتبة حسب دقة التصنيف.

المعامل كايا	الدقة الكلية	الخوارزمية
0.79	%83.3	المسافة الدنيا
0.75	%80.1	الاحتمالية العظمى
0.71	%78.2	مسافة Mahalanobis

ونلاحظ أن خوارزمية المسافة الدنيا قد أعطت أدق النتائج تلتها خوارزمية الاحتمالية العظمى ومن ثم خوارزمية مسافة Mahalanobis. ومن الناحية البصرية، نلاحظ أن المرئيات المصنفة باستخدام خوارزمية الاحتمالية العظمى ومسافة Mahalanobis متوافقة أكثر مع بعضها ولكنها صنفت بعض الأراضي الزراعية على أنها مناطق حضرية. أما نتائج المسافة الدنيا فهي متوافقة أكثر الواقع مقارنة بالخوارزميتين الباقيتين (الشكل (13)).

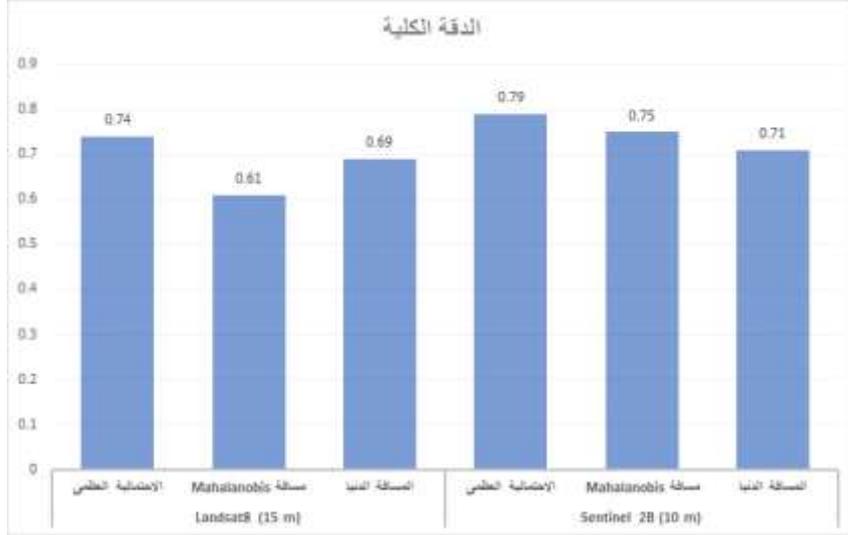


الشكل (13). التقارب بين نتائج خوارزميتي الاحتمالية العظمى ومسافة Mahalanobis مع تصنيفهما لبعض المناطق الخضراء على أنها مناطق مبنية. نتائج خوارزمية المسافة الدنيا هي الأقرب إلى الواقع.

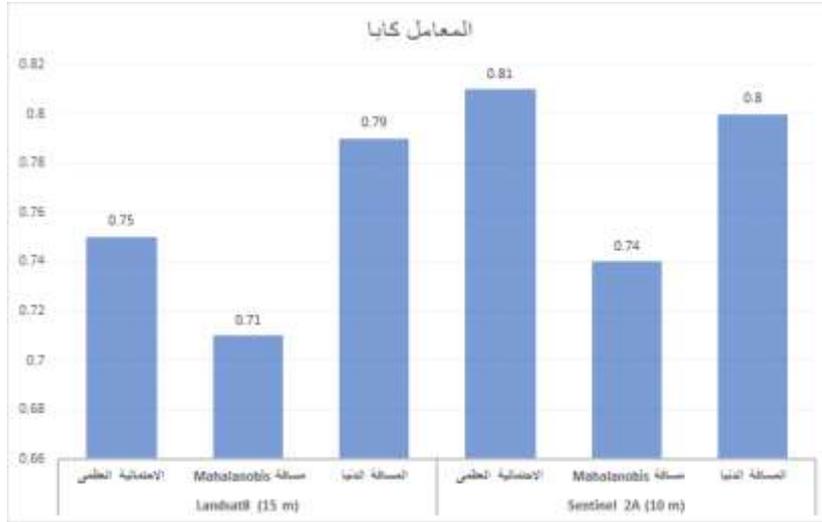
4- مقارنة النتائج

الدقة الكلية والمعامل كبا

يوضح الشكل (14) مقارنة قيم الدقة الكلية لكل خوارزميات التصنيف المراقب المختبرة وذلك كتابع للمرئية المصنفة، حيث نلاحظ تحسّن قيم هذه الدقة لكل الخوارزميات مع زيادة دقة التمييز المكانية للمرئية من 15 m إلى 10 m. من ناحية أخرى، يوضح الشكل (15) مقارنة قيم المعامل كبا الخاص بخوارزميات التصنيف المراقب المختبرة وذلك كتابع للمرئية المصنفة، حيث نلاحظ تحسّن قيم هذا المعامل لكل الخوارزميات مع زيادة دقة التمييز المكانية للمرئية. من ناحية أخرى نلاحظ أن خوارزمية التصنيف بالاحتمالية العظمى هي الخوارزمية الأدق في حالة المرئية Sentinel 2A وأن هنالك تحسناً في قيمة المعامل لحالة خوارزمية المسافة الدنيا عند نقصان دقة تمييز المرئية المصنفة.



الشكل (14). قيم الدقة الكلية للتصنيف المراقب كتابع للخوارزمية ولدقة تمييز المرئية.



الشكل (15). قيم المعامل كاتا للتصنيف المراقب كتابع للخوارزمية ولدقة تمييز المرئية.

يمكن تفسير الفروق في قيم الدقة في حالة المرئية 8 Landsat بالنسبة للمرئية 2A Sentinel بوجود تداخلات أكبر بين مدى الفئات بين الطبقات بسبب التقارب بين القيم الانعكاسية الطيفية للأهداف وبسبب دقة التمييز المكانية المنخفضة للمرئية المصنفة (15 m)، ويؤدي هذا التداخل إلى تصنيف بعض الخلايا تصنيفاً خاطئاً (قارن قيم مصفوفتي الالتصاق في الشكلين (4) و(9)).

مساحة أصناف الغطاء الأرضية

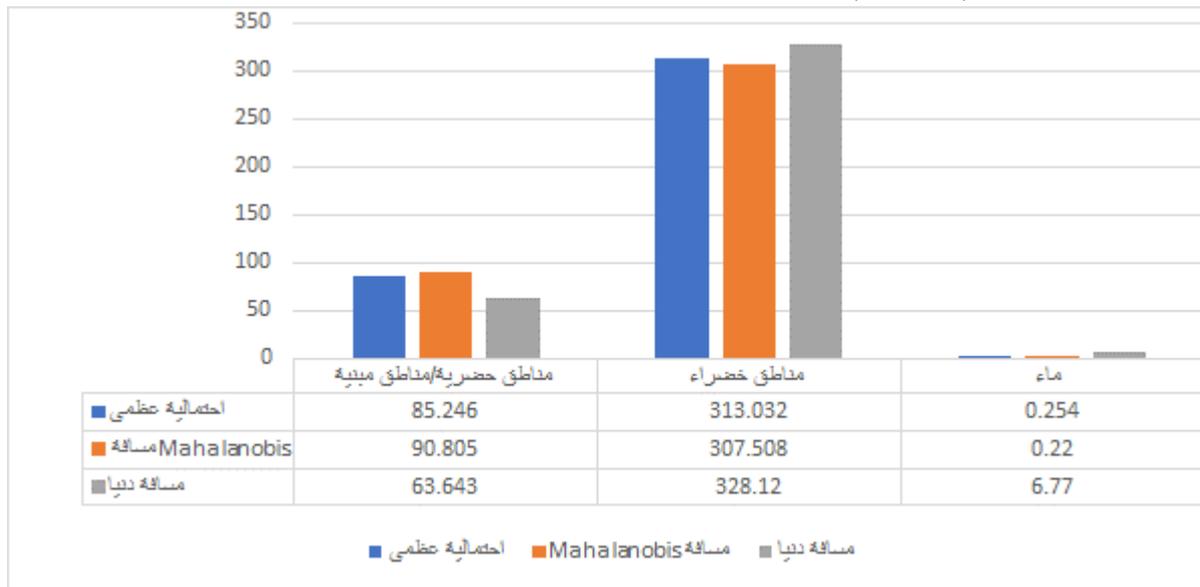
نوضح في الجدول (4) مساحات الأصناف المحددة (بالكيلومتر المربع) بواسطة كل من خوارزميات التصنيف المراقب المختبرة المطبقة على المرئية 8 Landsat (دقة التمييز 15 m) وعلى المرئية 2A Sentinel (دقة التمييز 10 m).

الجدول (4). مساحات الأصناف الناتجة كتابع لخوارزمية التصنيف المراقب ولدقة تمييز المرئية (بالكيلومتر المربع).

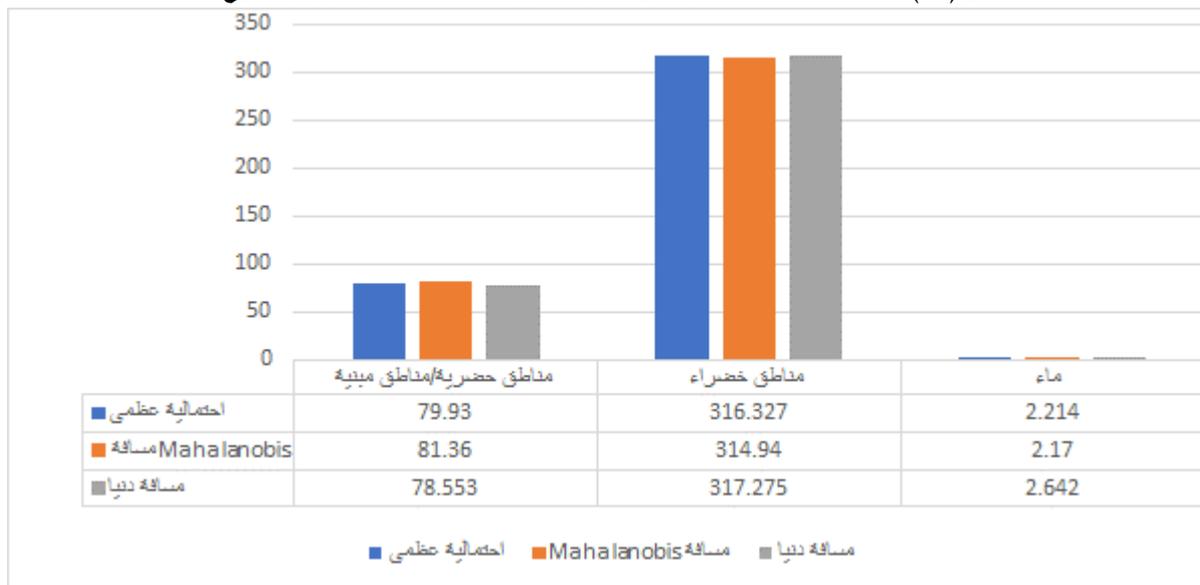
المرئية	الصف	احتمالية عظمى	مسافة Mahalanobis	مسافة دنيا
Sentinel 2A (10 m)	مناطق حضرية/مناطق مبنية	79.93	81.36	78.553

	مناطق خضراء	316.327	314.94	317.275
	مياه	2.214	2.17	2.642
	مناطق حضرية/مناطق مبنية	85.246	90.805	63.643
Landsat 8 (15 m)	مناطق خضراء	313.032	307.508	328.12
	مياه	0.254	0.22	6.77

بالنسبة للمرئية Sentinel 2A، فنلاحظ تقارباً بين نتائج جميع الخوارزميات وهذا عائد إلى زيادة دقة تمييز المرئية. أما فيما يخص المرئية Landsat 8، فنلاحظ التقارب بين نتائج خوارزمية الاحتمالية العظمى وخوارزمية مسافة Mahalanobis الدنيا (الشكل 16).



الشكل (16). مساحات الأصناف الناتجة عن تصنيف المرئية Landsat 8 بالكيلومتر المربع.



الشكل (17). مساحات الأصناف الناتجة عن عن تصنيف المرئية Sentinel 2A بالكيلومتر المربع.

إن النتائج التي توصل إليها البحث تتوافق مع نتائج معظم الأبحاث السابقة، ولكننا حاولنا في هذا البحث زيادة عدد خوارزميات التصنيف المراقب المختبرة بحيث نستطيع الحصول على نتائج أكثر عمومية. في الواقع، يتوفر عدد قليل من الدراسات التي تناولت خوارزميات التصنيف المراقب مثل مسافة ماهالانوبيس، المسافة الدنيا، من أجل مرئيات Sentinel، وهذا ما حاولنا إضافته في بحثنا. من ناحية أخرى، تتوافق نتائج بحثنا مع نتائج الأبحاث الأخرى من ناحية تأثير دقة التمييز الخاصة بالمرئيات المصنفة على دقة هذا التصنيف ولكن هذه الأبحاث استندت في عملية تقييم دقة التصنيف المراقب على مصادر مدفوعة ومكلفة عموماً في حين حاول بحثنا الاستفادة من صور ملونة بالألوان الحقيقية وذات دقة تمييز مناسبة متاحة مجاناً على شبكة الإنترنت في عملة التقييم.

الاستنتاجات والتوصيات:

وجدنا في هذا البحث ما يأتي:

- 1- يمكننا تطبيق طرائق الاستشعار عن بعد على المرئيات الفضائية (التصنيف) من أجل إنتاج خرائط الغطاء الأرضية بدقة تتبع لخوارزمية التصنيف المراقب المطبقة ودقة تمييز هذه المرئيات.
 - 2- بالنسبة لخوارزميات التصنيف المراقب، أثبتت خوارزمية الاحتمالية العظمى أنها الأدق في تصنيف المرئية الفضائية Sentinel 2A لأغراض تخريط الغطاء الأرضية حيث بلغت قيمة الدقة الكلية للتصنيف % 88.2 مع معامل كابا مساو لـ 0.81. أما بالنسبة للمرئية Landsat 8 فقد كانت خوارزمية المسافة الدنيا هي الأدق حيث بلغت الدقة الكلية للتصنيف القيمة % 83.3 مع معامل كابا مساو لـ 0.79.
 - 3- لاحظنا تحسّن قيم الدقة الكلية للتصنيف لكل الخوارزميات مع زيادة دقة التمييز المكانية للمرئية المصنفة. في الواقع، عند زيادة دقة التمييز من 15 m إلى 10 m، تحسّنت قيمة الدقة الكلية للتصنيف من 80.1% إلى % 88.2 للاحتتمالية العظمى ومن % 78.2 إلى % 80.9 لخوارزمية مسافة Mahalanobis ومن % 83.3 إلى % 85.3 لخوارزمية المسافة الدنيا.
 - 4- لاحظنا تحسّن قيم المعامل كابا مع زيادة دقة التمييز المكانية للمرئية المصنفة. في الواقع، عند زيادة دقة التمييز من 15 m إلى 10 m، تحسّنت قيمة هذا المعامل من 0.75 إلى 0.81 للاحتتمالية العظمى ومن 0.71 إلى 0.75 لخوارزمية مسافة Mahalanobis ومن 0.79 إلى 0.80 لخوارزمية المسافة الدنيا.
 - 5- مع زيادة دقة تمييز المرئيات، وجدنا أن هنالك تداخل أقل بين مدى الفئات بين مما قلل من تصنيف بعض الخلايا تصنيفاً خاطئاً.
- نوصي بتوسيع الدراسة بحيث ندخل خوارزميات تصنيف مراقب أخرى مثل خوارزمية SVM، خوارزمية متوازي السطوح Parallelogram algorithm، والتصنيف وفق مخطط الزاوية الطيفية Spectral angle وغيرها مع تغيير دقة تمييز المرئية بحيث نغطي طيفاً أوسع من دقات التمييز المتوافرة حالياً في المرئيات المستشعرة عن بعد، وذلك للوصول إلى نتائج أكثر شمولية.

References:

1. ABBURU, S., GOLLA, S.- *Satellite image classification methods and techniques: A review*. International journal of computer applications, 2015, 119.8.
2. AL KHALIL, O. -*Standards to evaluate the precision of geometric correction of satellite images*. Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series, 2016, 38, 2, 29-44.
3. AL-DOSKI, J., MANSORL, S., SHAFRI, H. -*Image classification in remote sensing*. Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia, 2013, 3.10.
4. BOSER, B., GUYON, I., VAPNIK, V.- *A training algorithm for optimal margin classifiers*. In: Proceedings of the fifth annual workshop on Computational learning theory. 1992, 144-152.
5. HALDER, A., GHOSH, A., GHOSH, S. -*Supervised and unsupervised landuse map generation from remotely sensed images using ant-based systems*. Applied Soft Computing, 2011, 11.8, 5770-5781.
6. HAYKIN, S. -*Neural networks and learning machines, 3/E*. Pearson Education India, 2009.
7. JIN, S. -*Accuracy Assessment for Classification and Modeling*. Remote Sensing of Natural Resources, 2013, 45.
8. JOG, S., DIXIT, M.-*Supervised classification of satellite images*. In: 2016 Conference on Advances in Signal Processing (CASP). IEEE, 2016, 93-98.
9. LILLESAND, T., KIEFER, R., CHIPMAN, J. -*Remote sensing and image interpretation*. John Wiley & Sons, 2015.
10. SATHYA, P., MALATHI, L. -*Classification and segmentation in satellite imagery using back propagation algorithm of ann and k-means algorithm*. International Journal of Machine Learning and Computing, 2011, 1.4, 422.
11. SHAHABI, H., et al. -*Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier*. Remote Sensing, 2020, 12.2, 266.
12. THANH NOI, P., KAPPAS, M.- *Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery*. Sensors, 2017, 18.1,18.
13. TIEN BUI, D., et al. -*Land cover change mapping using a combination of Sentinel-1 data and multispectral satellite imagery: A case study of Sanandaj county, Kurdistan, Iran*. Applied Ecology & Environmental Research, 2019, 17.3.
14. YANG, Di, et al.- *Open land-use map: a regional land-use mapping strategy for incorporating OpenStreetMap with earth observations*. Geo-spatial information science, 2017, 20.3, 269-281.
15. YONEZAWA, C.- *Maximum likelihood classification combined with spectral angle mapper algorithm for high resolution satellite imagery*. International Journal of Remote Sensing, 2007, 28.16, 3729-3737.
16. ZHOU, Q., ROBSON, M.- *Contextual information is ultimately necessary if one is to obtain accurate image classifications*. International Journal of Remote Sensing, 2001, 22, 3457-3470.