تأثير عدد عينات التدريب وعدد نقاط الاختبار على دقة خرائط استخدام الأراضي والغطاء الأرضي والمشتقة من المرئيات الفضائية باستخدام خوارزميات التصنيف القائمة على تعلم الآلة: دراسة حالة على واحة الأحساء في المملكة العربية السعودية(*)

د. خديجة يحيى فقيه
أستاذ الخرائط المشارك
قسم الجغرافيا والاستدامة البيئية_كلية العلوم الإنسانية والاجتماعية
جامعة الأميرة نورة بنت عبد الرحمن

الملخص:

حققت هذه الدراسة في تأثير عدد عينات التدريب وعدد نقاط التحقق على دقة خرائط استخدام الأراضي والغطاء الأرضي (LULC) لواحة الإحساء في المملكة العربية السعودية، المشتقة من صور الأقمار الصناعية Landsat 8 باستخدام مصنفات آلة دعم المتجهات Support Vector Machines (SVM) والغابة العشوائية Random Forest (RF). تراوح عدد عينات التدريب بين 5 إلى 30 عينة لكل صنف. تم تقييم دقة التصنيف باستخدام الدقة الكلية (OA) ومعامل كابا، مع نقاط التحقق التي تراوحت من 5 إلى 30 لكل صنف، مقارنةً بالبيانات المرجعية من SAS Planet. بالإضافة إلى ذلك، تم تقييم دقة خرائط LULC من خلال مقارنة مساحات السطوح المائية الناتجة من أعداد عينات التدريب المختلفة مع تلك المحسوبة باستخدام مؤشر المياه المعياري (NDWI).

أشارت النتائج إلى أن SVM تفوقت على RF في معظم السيناريوهات، محققةً دقة عالية وثابتة حتى مع عدد قليل من عينات التدريب. أظهرت SVM قيم كابا تراوحت بين 0.93 إلى ١,٠٠ وOA بين 0.95 إلى ١,٠٠ مع 5 عينات، ومع زيادة عدد نقاط التحقق إلى 30، حافظت SVM على قيم كابا من 0.90 إلى ٠,٩٨ وOA من 0.92 إلى ٠,٩٩، مما يعكس استقرارها. بينما أظهرت RF أداءً جيداً ولكن مع تذبذب أكبر. كانت دقة RF أقل مع 5 عينات تدريب حيث تراوحت قيم كابا بين 0.80 إلى ٠,٩٠ وOA بين 0.85 إلى

(*)**مجلة كلية الآداب جامعة القاهرة المجلد (٨٤) العدد (٧) أكتوبر ٢٠٢٤.**

٠,٩٢. ومع 30 نقطة تحقق، تراوحت قيم كابا لـ RF بين ٠,٧٨ إلى 0.95 وOA بين ٠,٨٠ إلى 0.93، مما يشير إلى استقرار أقل.

يضاف إلى ذلك، أن تحليل مساحات السطوح المائية أظهر تراجع أداء RF بشكل ملحوظ مع عدد قليل من عينات التدريب ومن ثم تحسن هذا الأداء بشكل كبير مع زيادة عدد العينات، مع انخفاض الأخطاء من ١٥% إلى ٥% مع زيادة عدد العينات. على العكس من ذلك، حافظت SVM على أداء مستقر عبر جميع نطاقات عينات التدريب، مع أخطاء أقل من 5%. في الختام، كانت SVM عموماً أكثر دقة واستقراراً من RF، مما يجعلها المصنف المفضل لرسم خرائط LULC في معظم الحالات.

الكلمات المفتاحية: استخدام الأراضي والغطاء الأرضي (LULC)، خوارزمية آلة دعم المتجهات، خوارزمية الغابة العشوائية، عدد عينات التدريب، نقاط التحقق، الدقة الكلية، معامل كابا.

The impact of the number of training samples and the number of test points on the accuracy of land use and land cover maps derived from satellite imagery using machine learning-based classification algorithms: A case study on Al-Ahsa Oasis in Saudi Arabia

Abstract

This study investigated the effect of training sample size and check points on the accuracy of land use and land cover (LULC) maps for the Al-Ahsa oasis in Saudi Arabia, derived from Landsat 8 imagery using Support Vector Machines (SVM) and Random Forest (RF) classifiers. Training sample sizes ranged from 5 to 30 samples per class. The classification accuracy was evaluated using overall accuracy (OA) and the kappa coefficient, with check points varying from 5 to 30 per class, compared against reference data from SAS Planet. Additionally, the accuracy of LULC maps was assessed by comparing the water surface areas obtained from different training sample sizes with those calculated using the Normalized Difference Water Index (NDWI).

The results indicated that SVM generally outperformed RF in most scenarios, achieving high and stable accuracy even with a small number of training samples. SVM showed kappa values from 0.93 to

1.00 and OA from 0.95 to 1.00 with 5 samples. As the number of check points increased to 30, SVM maintained kappa values from 0.90 to 0.98 and OA from 0.92 to 0.99, reflecting its robustness. RF, while producing good results, exhibited greater variability in performance. With 5 training samples, RF's accuracy was lower, with kappa values from 0.80 to 0.90 and OA from 0.85 to 0.92. With 30 check points, RF's kappa values ranged from 0.78 to 0.95 and OA from 0.80 to 0.93, indicating less stability.

Furthermore, the analysis of water surface areas showed that RF performed significantly worse with fewer training samples but improved notably with more samples, with errors decreasing from 15% to 5% as samples increased. Conversely, SVM maintained consistent performance across all training sample ranges, with errors consistently below 5%. In conclusion, SVM was generally more accurate and stable than RF, making it the preferred classifier for LULC mapping in most cases.

Key words: Land use and land cover (LULC), Support Vector Machines (SVM), Random Forest (RF), Training sample size, Check points, Overall accuracy, Kappa coefficient.

١- مقدّمة:

تعدّ المرئيات الفضائية مدخلات أساسية في مجال إنتاج خرائط استخدامات الأراضي والغطاء الأرضي Land Use/Land Cover (LU/LC) maps (Taha, ٢٠١٦). فيما يخص خرائط استخدام الأراضي فهي توثق الأنشطة البشرية من بناء وزراعة، في حين تظهر خرائط الغطاء الأرضي الأصناف الفعلية للأرض من غابات ومياه ومناطق مبنية (Yang, et al., 2017) ولهذه الخرائط تطبيقات كثيرة مثل التخطيط الحضري وإدارة الموارد ومراقبة البيئة. هذا، ويعد التصنيف المراقب Supervised classification أحد أهم التقنيات المطبّقة للقيام بتصنيف المرئيات الفضائية واشتقاق خرائط استعمالات الأراضي والغطاء الأرضي. تحتاج هذه التقنية إلى عيّنات تدريب Training samples والتي تعتبر ممثلةً عن كل ظاهرة من الظاهرات الأرضية، ويتم جمع هذه العينات بناءً على معرفة المستخدم بالمشهد المصور وهي تستخدم لتدريب المصنّف Classifier للتعرف

على الظاهرات المتوافرة في المرئيات الفضائية (Richards & Richards, 2022) .

نذكر من خوارزميات التصنيف المراقب التقليدية خوارزمية الاحتمالية العظمى Maximum Likelihood (ML)، خوارزمية المسافة الدنيا Minimum Distance، مسافة ماهالانوبيس Mahlanobis. من ناحية أخرى، تتوافر خوارزميات تسمى بخوارزميات تعلم الآلة Machine learning مثل خوارزمية الشبكة العصبونية الصناعية Artificial Neural Network (ANN) Haykin, (2009)، الغابة العشوائية Random Forest (RF) وآلة دعم المتجهات Support vector machines (SVM) (Boser, 1992) والتي أصبحت من أكثر الأدوات فعاليةً في تصنيف المرئيات الفضائية حالياً (Rahman, et al., 2020; Ouchra, et al., 2022). تضم هذه الخوارزميات مجموعة كبيرة من الأساليب والتقنيات التي تمكن البرامج الحاسوبية من استخدام البيانات المستخرجة من مرئيات الأقمار الصناعية لتحديد وتصنيف مختلف الظاهرات الموجودة على سطح الأرض (Maulik & Chakraborty, 2017). تقوم خوارزمية RF على مجموعة من الأشجار القرارية Decision trees تعمل معاً لتحسين الدقة في التصنيف وتتميز بقدرتها على التعامل مع بيانات ذات أعداد كبيرة، أما خوارزمية SVM فتهدف إلى تعيين حدود تفصل بين أصناف البيانات (Laban, et al., 2019). بسبب حاجة هاتين الخوارزميتين، شأنهما شأن كل خوارزميات التصنيف المراقب، إلى عينات التدريب، فإن دقتهما تتبع لعدد هذه العينات (Figueroa, et al., 2012) بالإضافة إلى تأثر دقتهما بعدد نقاط الاختبار Check points المستخدمة في تقييم دقة التصنيف. ويهدف هذا البحث إلى دراسة تأثير عدد عينات التدريب وعدد نقاط الاختبار على دقة خرائط استخدام الأراضي المشتقة من تصنيف مرئية فضائية من النوع Landsat 8 باستخدام خوارزميتي SVM و RF.

٢- مشكلة الدراسة:

يمكن صياغة مشكلة الدراسة كما يأتي:

١. كيف يؤثر عدد عينات التدريب على دقة تصنيف المرئيات الفضائية باستخدام خوارزميتي تعلم الآلة SVM وRF؟

٢. ما هو عدد نقاط الاختبار المستخدمة في تقييم دقة التصنيف على دقة الخوارزميتين SVM وRF؟

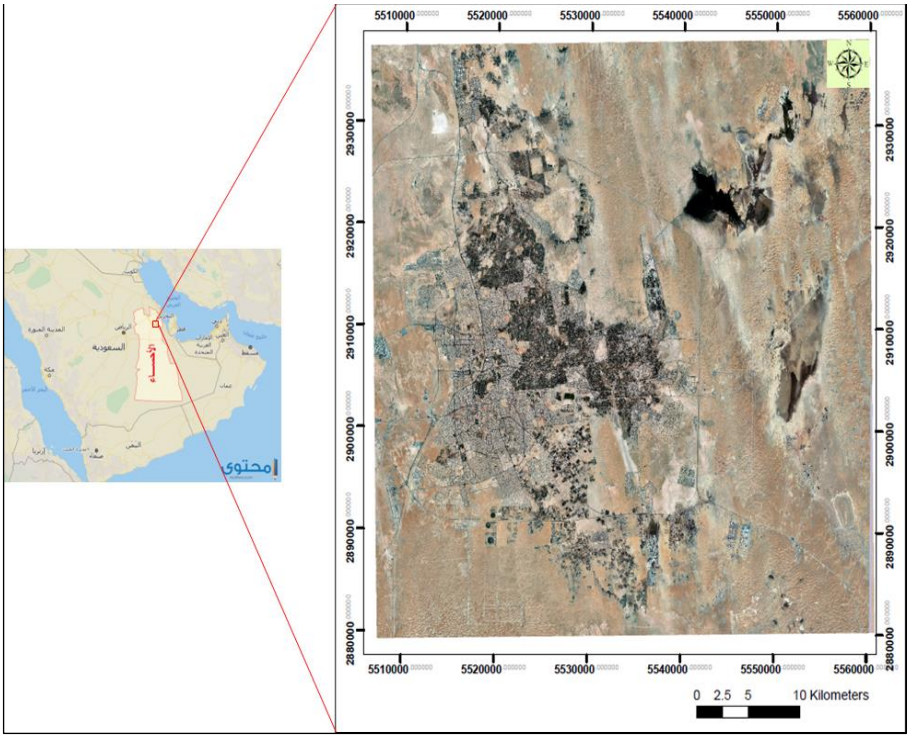
٣- أهداف الدراسة وأهميتها:

نسعى في هذه الدراسة إلى تحليل تأثير أعداد مختلفة من عينات التدريب على دقة تصنيف المرئيات الفضائية باستخدام خوارزمية SVM وخوارزمية RF وكذلك إلى دراسة تأثير عدد نقاط الاختبار المستخدمة في تقييم دقة التصنيف على أداء هاتين الخوارزميتين. أما أهمية البحث فهي تتجلى في مساعدة الباحثين على فهم تأثير عدد عينات التدريب وعدد نقاط الاختبار على دقة تصنيف المرئيات وعلى تحديد العدد الأمثل للعينات ونقاط الاختبار اللازمة لتحقيق دقة عالية.

٤- منطقة الدراسة

تقع واحة الأحساء في المنطقة الشرقية من المملكة العربية السعودية، حيث يقطنها ١٥% من إجمالي السكان السعوديين. تقع الأحساء على بعد ١٥٠ كم جنوب غرب الدمام و٣٣٠ كم شمال الرياض. ورغم كونها واحة صحراوية، إلا أن وجود العيون التي تتدفق منها المياه الطبيعية جعلها أكبر واحة زراعية في العالم، حيث تحتوي على أكثر من ٢,٥ مليون شجرة نخيل. تقع منطقة الدراسة في الإحساء وهي مدينة الإحساء والتي تُعَدُّ عاصمة محافظة الأحساء التابعة للمنطقة الشرقية. تأخذ المحافظة شكل زاوية قائمة، وتوجد وتتوزع الأراضي الزراعية في الشمال والشرق والجنوب وهي تضم مدن الهفوف، المبرز، العيون، العمران، مدينة الجفر، وجواثا. تقع هذه الواحة بين خطي العرض 15′ 25°

و40′ 25º وبين خطي الطول 30′ 49º و50′ 25º (وزارة الشؤون البلدية والقروية، ٢٠١٩) كما هو موضح في شكل (١).



المصدر: من إعداد الباحثة.

شكل (١). منطقة الدراسة.

٥- مصطلحات الدراسة:

٥-١- عيّنات التدريب:

نقصد بعينات التدريب مجموعات البيانات المستخدمة في تدريب خوارزميات تعلم الآلة وهي تساعد هذه الخوارزميات على فهم الأنماط والخصائص التي تملكها البيانات (المرئيات الفضائية) (Tuia, et al., 2011). بمعنى آخر أنه تتعلم خوارزميات التصنيف عبر عمليات التدريب على هذه العينات، حيث

تحاول فهم العلاقات والأنماط التي تسمح لها بإنجاز تصنيف دقيق. يؤثر عدد عينات التدريب على أداء خوارزمية تعلم الآلة بشكل كبير وعادة ما تؤدي زيادة هذا العدد إلى زيادة دقة التصنيف، وذلك لأنه يمكن للخوارزمية أن تتعلم بشكل أفضل ما هي الأنماط الدقيقة وتستطيع التأقلم مع الاختلافات في البيانات. ولكننا نشير هنا إلى أن زيادة عدد عينات التدريب بشكل غير مدروس يمكن أن يؤثر على تكلفة التحليل وموارد التخزين (Koggalage & Halgamuge, 2004)، خاصةً في حالة البيانات الكبيرة مثل المرئيات الفضائية، كما أنه قد يزيد من تعقيد الخوارزمية إذا كانت البيانات غير متوازنة أو غير ممثلة بشكل جيد.

٥-٢- نقاط الاختبار:

في سياق تصنيف المرئيات الفضائية، تمثل نقاط الاختبار الجزء الأهم من عملية تقييم أداء خوارزمية تعلّم الآلة (Sen, al., 2020). وهذه النقاط هي عينات حقيقية من البيانات المقاسة حقلياً أو على صور فضائية عالية الدقة أو على خرائط دقيقة لاستعمالات الأراضي والغطاء الأرضي وهي تعكس البيانات التي يجب أن تكون الخوارزمية قادرةً على تصنيفها بدقة. إذا كانت الخوارزمية قادرةً على تحديد نقاط الاختبار بدقة عالية فإن هذا يشير إلى أنها تستطيع التعامل مع أنماط البيانات العامة بشكل جيد. هذا ويجب أن تكون نقاط الاختبار متنوعة وتمثل بشكل جيد جميع الظاهرات والأنماط التي من المحتمل أن تواجهها الخوارزمية في البيانات الحقيقية لتجنب التحيز.

٥-٣- تقييم دقة التصنيف:

يتم إجراء تقييم الدقة Accuracy assessment لتحديد مدى كفاءة أداء طريقة تصنيف معينة، وهذا التقييم يقارن تقييم الدقة صورة مصنفة بصورة مرجعية تتطابق مع الصورة المصنفة من حيث المقياس والتفاصيل والتاريخ والفئات والإسقاط. يتم إجراء هذا التقييم غالباً من خلال مقارنة بكسل- بكسل Pixel by pixel بين الصورة المصنفة وصورة يُفترض أنها تمثل سطح الأرض بشكل

صحيح (Stehman & Foody, 2009). على الرغم من أن المقارنة بكسل-بكسل تقيّم الاتفاق العام بين الصورتين، فإننا في سياق الاستشعار عن بعد نعتبر الصورة المرجعية معياراً للدقة ونبلغ عن الاختلافات بين الصورتين كأخطاء. عملياً، يتم استخدام مصفوفة الخطأ Error Matrix للتحقق من دقة تصنيف Classification accuracy استخدامات الأراضي والغطاء الأرضي LULC انطلاقاً من المرئيات الفضائية بطريقة كمية (Congalaton & Green, 2019). وتستخدم مصفوفة الخطأ لحساب مقاييس مختلفة لدقة التصنيف مثل الدقة الكلية OA (Overall Accuracy)، دقة المستخدم UA (User Accuracy) (أخطاء الحذف)، دقة المنتج PA (Procedure's accuracy) (أخطاء الإدراج)، المعامل كابا k (Kappa coefficient). فيما يخص الدقة الكلية Overall Accuracy) OA) فهي تمثل نسبة مجموع خلايا عينة الاختبار المصنفة بشكل صحيح إلى العدد الإجمالي للخلايا في مصفوفة الأخطاء بأكملها (Fitzgerald & Lees, 1994). أما دقة المستخدم فهي مقياس لموثوقية التصنيف؛ لأنها تخبر المستخدم بعدد المرات التي سيكون فيها الصنف على الخريطة موجوداً فعلياً على الأرض (Congalton, 1991) وخطأ الإدراج لأي صنف، يحدث عندما تقوم عملية التصنيف بتعيين وحدات بكسل لصنف معين لا تنتمي إليه في الواقع، وكلما زادت أخطاء الإدراج انخفضت دقة المستخدم (Congalton, 1991). بالنسبة لدقة المنتج فهي تشير إلى عدد المرات التي تظهر فيها البيانات الحقيقية على الأرض بشكل صحيح على الصورة المصنفة، وهي تساوي عدد المواقع المرجعية المصنفة بدقة مقسومة على العدد الإجمالي للمواقع المرجعية لذلك الصنف (Stehman, 1997). أما خطأ الاستبعاد فهو يشير إلى استبعاد المنطقة من الصنف الذي تنتمي إليه وكلما زادت أخطاء الاستبعاد انخفض مستوى دقة المنتج (Stehman, 1997). وأخيراً، يستخدم المعامل كابا لقياس العلاقة بين الاتفاق المتوقع لاحقاً (beyond chance agreement) وعدم الاتفاق (الاختلاف) المتوقع (Expected disagreement)،

وتستخدم في إيجاده كل عناصر مصفوفة الخطأ (Vanbelle, ٢٠١٦). يجب أن تكون قيمة ٠,٧٥K≥ لنحصل على نتائج جيدة جداً في التصنيف. وإذا كانت قيمة ٠,٤>K فذاك مؤشر عن دقة متواضعة (Taufik, 2016).

٥-٤- الخوارزمية SVM والخوارزمية RF

خوارزمية آلة دعم المتجهات SVM هي خوارزمية تعلم آلة تستخدم بشكل شائع في مجال تصنيف المرئيات الفضائية والهدف منها هو إيجاد أفضل حد فاصل بين الظاهرات المختلفة في هذه المرئيات. تقوم الخوارزمية SVM بإيجاد خط يفصل بين الفئات المختلفة من البيانات بأكبر مسافة ممكنة ويسمى بالحد الأقصى للفصل Maximum Margin ثم تستخدم تابع النواة Kernel لتحويل البيانات إلى حيّز ذي أبعاد أعلى حيث يمكن أن تكون قابلة للفصل بشكل خطي (Chugh, et al., ٢٠٢٠). في الأساس، طريقة تصنيف آلة دعم المتجهات SVM مشتقة من نظرية التعلم الإحصائي. الخطوة الأولى في هذه الطريقة هي تحويل متجهات البكسل الأصلية إلى مجموعة جديدة من الميزات، ثم يتم العثور على المستوي الفائق Hyperplane الأمثل الذي يزيد من الهامش بين الفئات في فضاء الميزات (Nemmour & Chibani, 2006). الغرض من تحويل المتجهات هو تحسين الفصل قدر الإمكان. دالة القرار لآلة دعم المتجهات هي:

$$f(x) = sign(\sum_{i=1}^{N} a_i^* y_i K(x, x_i) + b^*) \qquad (1)$$

حيث أن a_i^* هي مضاعفات لاغرانج، وN هو عدد المتجهات الداعمة التي تمثل بيانات التدريب والتي تتراوح قيمتها بين $0 \le a_i^* \le C$ و C هو معامل يحدده المستخدم للتحكم في التوازن بين عدد البكسلات غير القابلة للفصل وتعقيد الآلة. $K(x, x_i)$ هي دالة النواة المستخدمة لتحويل متجهات البكسلات. والانحياز b^* هو قيمة عددية محسوبة باستخدام أي متجه داعم.

أما خوارزمية تصنيف الغابة العشوائية (RF) فهي مصنف جماعي يستخدم مجموعة من أشجار القرار لإجراء التنبؤ بالتصنيف (Breiman, 2001). يتم إنشاء الأشجار عن طريق اختيار مجموعة فرعية من عينات التدريب مع الاستبدال. هذا يعني أن عينة واحدة قد يتم اختيارها عدة مرات، في حين أن عينات أخرى قد لا يتم اختيارها على الإطلاق (Belgiu & Drăguţ, 2016) وبالتالي يتم استخدام حوالي ثلثي عينات التدريب لتدريب الأشجار وتُستخدم العينات المتبقية لتقدير أداء الغابة العشوائية. يتم اتخاذ قرار تصنيف بكسلات الصورة بواسطة تصويت جميع الأشجار، وسيتم في النهاية اختيار الفئة التي تحصل على أكبر عدد من الأصوات.

٦- دراسات سابقة:

تمت دراسة تأثير عدد عينات التدريب على دقة التصنيف المراقب بتطبيق خوارزميات تعلم الآلة في العديد من الدراسات السابقة ففي (Myburgh & Van Niekerk, 2014) تمت مقارنة مجموعة من خوارزميات التصنيف وهي الخوارزمية SVM وخوارزمية الشبكة العصبونية Neural Network (NN) وخوارزمية الاحتمالية العظمي ML مع التركيز على عدد عينات التدريب المستخدمة. أشارت نتائج البحث إلى أن أداء جميع المصنفات تحسن بشكل ملحوظ مع زيادة عدد عينات التدريب مع توضيح أن أداء المصنف ML كان ضعيفاً عند استخدام عدد قليل من عينات التدريب (<١٠ لكل فئة)، وأداء المصنف NN كان ضعيفاً مقارنة بأداء المصنف SVM. أما في (Thanh Noi & Kappas, 2017) فقد كان الهدف هو تقييم أداء أكثر ثلاث مصنفات شيوعاً، وهي RF، والجار الأقرب Nearest Neighbor (NN)، والمصنف SVM عبر تطبيقها على صورة من القمر الصناعي Sentinel-٢ مع تقييم تأثيرات عدد عينات التدريب على دقة نتائج التصنيف للمصنفات السابقة. وفيما يخص عدد عيّنات التدريب فقد أعطت المصنفات الثلاثة دقة إجمالية (OA) متشابهة وعالية كان عدد عيّنات التدريب كبيراً بما فيه الكفاية. وفي (Fu, et al, 2023).

تمت دراسة تأثير عدد عينات التدريب على دقة تصنيف خوارزميات تعلم الآلة حيث أوضحت نتائج الدراسة التحسن المستمر لدقة التصنيف مع زيادة عدد عينات التدريب وتمثيلها للظاهرة المدروسة مع وجود فارق طفيف في الدقة الإجمالية ودقة المنتج، ودقة المستخدم عند استخدام ١٠٠٠ و١٠٠٠٠ عينة تدريب. أما في (Zhang, et, al, 2023) فقد تم العمل على ثلاث مرئيات هي GF،Sentinel 1-2 ، وLandsat 8، وتطبيق ثلاث طرق لاختيار عينات التدريب: الاختيار المجموعي، والاختيار المعتمد على الإنتروبيا، والاختيار المباشر، ثم تم استخدام العينات التدريبية المختارة لتدريب ثلاث مصنفات هي SVM،RF ، والجار الأقرب NN. ووفقًا لنتائج التجارب، أظهرت النماذج الثلاثة للتصنيف أداءً متماثلاً. وأظهرت طريقة الاختيار المجموعي أفضل أداء حيث زادت دقة تصنيف المرئيات مع زيادة عدد العينات ضمن نطاق معين. من ناحية أخرى، أظهرت نتائج (Maxwell, et al, 2018) أن استخدام عينات تدريبية كبيرة ومتنوعة يمكن أن يحسن دقة التصنيف بشكل كبير، خصوصاً عند استخدام خوارزميات مثل RF وSVM وهذا ما أوضحته أيضاً دراسات (Mountrakis, et al, 2011) حيث وجد الباحثون زيادة فعالية خوارزمية RF في تحسين دقة تصنيف الغطاء الأرضي من خلال استخدام أعداد عينات تدريبية كبيرة ومتنوعة. وفي (Wang, et al., 2014) وجد الباحثون أن استخدام عينات تدريبية كبيرة يمكن أن يعزز دقة تصنيف المرئيات الفضائية بشكل ملحوظ، كما تم في ((Belgiu & Drăguţ, 2016 استنتاج أن استخدام مرئيات Landsat–٨ مع عينات تدريبية كبيرة يتيح تحقيق دقة تصنيف أعلى، خصوصاً عند استخدام خوارزميات مثل SVM وRF. وأخيراً، أظهرت دراسة (Grabocka, et al., 2016) أن اختيار عدد كافٍ من نقاط الاختبار يمكن أن يحسن من دقة التقييم العام للمصنفات كما أظهرت دراسة (Foody, 2002) أهمية عدد عينات التدريب في تحسين دقة النماذج التصنيفية، مؤكدة أن زيادة عدد العينة التدريبية

يساهم بشكل كبير في تحسين دقة التصنيف للنماذج المستخدمة في تحليل المرئيات الفضائية.

مما سبق نجد أن الدراسات السابقة تناولت تأثير عدد عينات التدريب على دقة خريطة استخدام الأرضي والغطاء الأرضي بتطبيق مصنفات تعلّم الآلة، في حين أنها لم تناقش التأثير المحتمل لعدد نقاط الاختبار المستخدمة على هذه الدقة وهذه هي الثغرة البحثية التي حاولت دراستنا سدها إضافة إلى استنادها إلى آليات مقارنة الخوارزميتين SVM وRF المطبقة في الدراسات السابقة.

٧- منهج الدراسة والبيانات المتوافرة:

المنهج العلمي المستخدم في الدراسة هو مزيج من المنهج التجريبي والاستقرائي. (أ) فيما يخص المنهج التجريبي فقد تم تطبيقه في جمع البيانات ومعالجتها حيث تم استخدام صور القمر الصناعي Landsat-8 وإجراء تصحيحات مسبقة، وهذا يتضمن إعداد تجارب مختلفة لمعالجة وتحليل البيانات. كما أن تطبيق المصنفات المختلفة ومقارنة نتائجها يشير إلى استخدام المنهج التجريبي لاختبار أداء هذه المصنفات. (ب) أما الجزء الاستقرائي فيخص تحليل النتائج واستنتاج الأنماط. فبعد تطبيق المصنفات، تم تحليل النتائج لاستنتاج الأنماط والتوجهات في دقة التصنيف بالنسبة لعدد عينات التدريب وعدد نقاط الاختبار والمصنفات المختلفة.

فيما يخص جمع البيانات فقد تم استخدام مرئية فضائية من النوع Landsat 8 OLI في هذه الدراسة ملتقطة في الشهر الخامس من العام ٢٠٢٤. تم الحصول على هذه المرئية مجاناً من المنصة Earth Explorer التابعة للمؤسسة الأمريكية للمسح الجيولوجي USGS (https://earthexplorer. usgs.gov/) هذه المرئية معطاة ضمن نظام الإحداثيات WGS_1984_UTM_Zone_39N وهي تضم ١١ حزمة طيفية (جدول (١)) دقة تمييزها 30 m مع حزمة البانوكروماتية Panchromatic (الحزمة ٨) التي تبلغ دقة تمييزها 15 m.

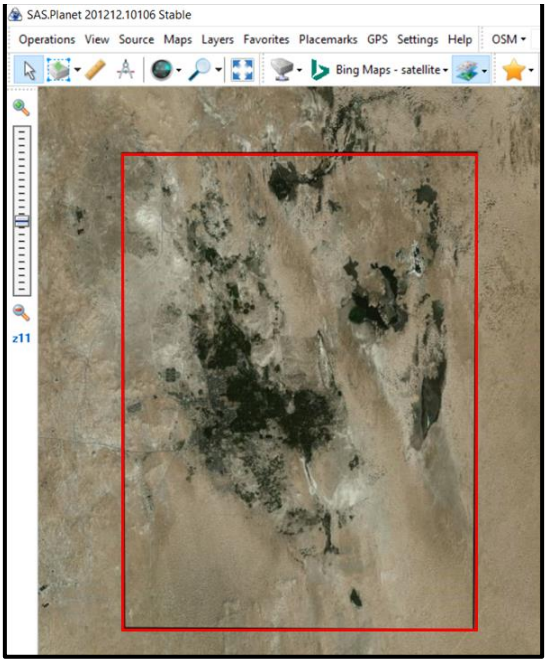
جدول (١). أهم الحزم الطيفية للمرئيات Landsat-8.

الحزم **Bands**	طول الموجة (ميكرومتر)	دقة التمييز (متر)
Band 1 – Coastal aerosol السواحل	0.45-0.43	30
Band 2 - Blue الأزرق	0.51-0.45	30
Band 3 - Green الأخضر	0.59-0.53	30
Band 4 - Red الأحمر	0.67-0.64	30
Band 5 – Near Infrared (NIR) تحت الحمراء القريبة (الغطاء النباتي)	0.88-0.85	30
Band 6 -SWIR تحت الحمراء المتوسطة ١	1.65-1.57	30
Band 7 -SWIR تحت الحمراء المتوسطة ٢	2.29-2.11	30
Band 8 - Panchromatic البانكروماتية	0.68-0.50	15
Band 9 - Cirrus السمحاق (السحب)	1.38-1.36	30
Band 10 – Thermal Infrared (TIRS) 1 تحت الحمراء الحرارية	11.19-10.60	100
Band 11 – Thermal Infrared (TIRS) 2 تحت الحمراء الحرارية	12.51-11.50	100

تم تطبيق التصحيحات الجوية والإشعاعية على هذه المرئية وذلك للحصول على بيانات أكثر دقة وأكثر موثوقية يمكن استعمالها بفعالية في عملية التصنيف والتحليل اللاحقة (Santamaria-Artigas et al., 2021). تم تحويل قيم الأعداد الرقمية DN الخاصة بكل حزمة من حزم المرئية Landsat إلى قيم إشعاع (Radiance) ومن ثم إلى قيم انعكاس Reflectance وتصحيح

تأثير زاوية ارتفاع الشمس Sun elevation وذلك بالاستعانة بالمعلومات المتوافرة في ملف الوصفية للمرئية (Metadata).

من ناحية أخرى، قمنا باقتطاع صورة ملونة بالاستعانة بالبرنامج SAS Planet[١] بدقة تمييز مكانية تساوي ١,٢٠ متر تغطي منطقة الدراسة (شكل (٢)) في فترة الدراسة لتلعب درو الصورة المرجعية اللازمة لتقييم دقة التصنيف المراقب للمرئية المستخدمة في هذا البحث.



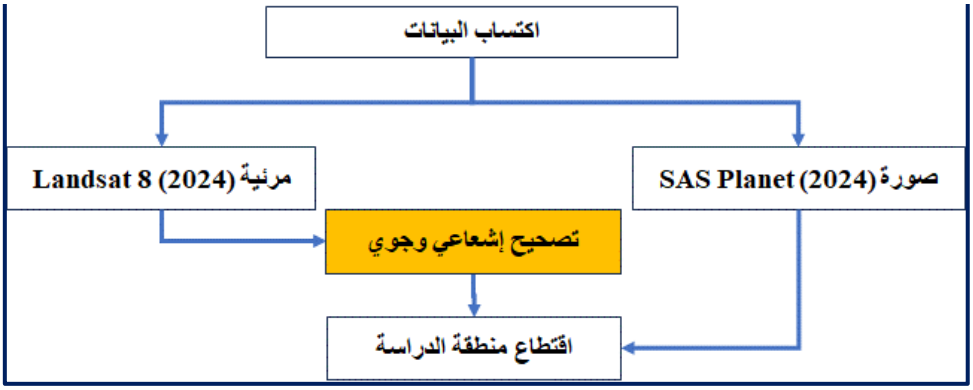
المصدر: من إعداد الباحثة.

شكل (٢). اكتساب الصور المرجعية باستخدام البرنامج **SAS Planet.**

تم بعد ذلك تركيب الحزم Bands composite للمرئية Landsat 8 OLI واقتطاع منطقة الدراسة (شكل (٣)).

---

[١] برنامج SAS Planet هو أداة مجانية تستخدم لاستعراض وتنزيل الخرائط من مجموعة متنوعة من مصادر الخرائط العالمية. يتيح البرنامج للمستخدمين الوصول إلى صور الأقمار الصناعية، والخرائط الطبوغرافية، وخرائط الشوارع، وغيرها من أنواع الخرائط.



المصدر: من إعداد الباحثة.

شكل (٣). اكتساب البيانات وتحضيرها.

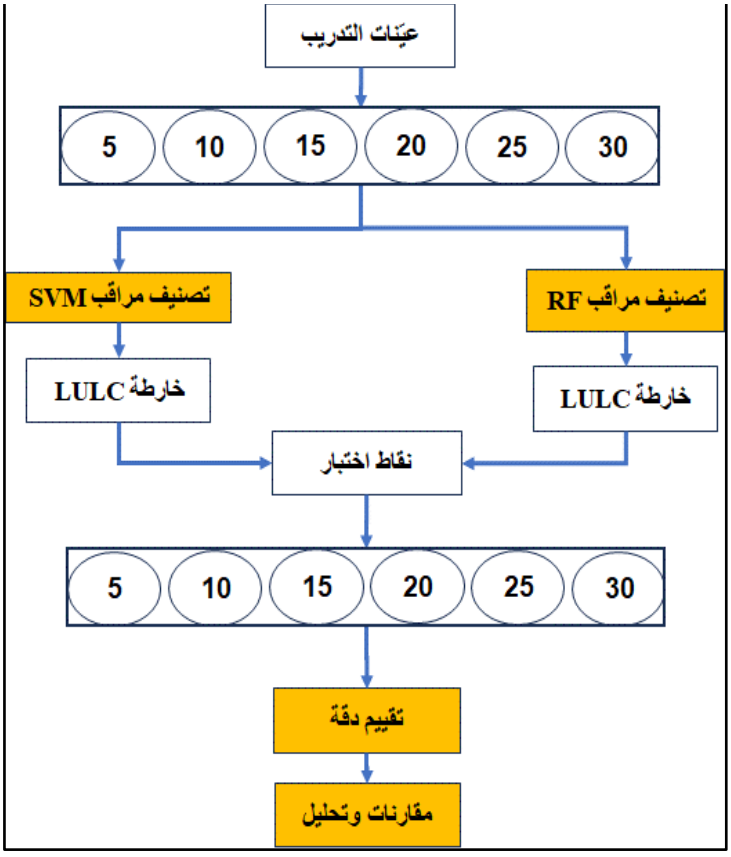
فيما يخص الخطوات العملية المطبقة في المعالجة فيمكن تلخيصها فيما يلي (شكل (٤)):

١- تغيير عدد عينات التدريب بدءاً من خمس عينات وصولاً إلى ثلاثين عيّنة تم قياسها في كل صنف من الأصناف المتوافرة في منطقة الدراسة. الزيادة التي تم اعتمادها هي خمس عينات في كل مرة.

٢- القيام بتصنيف المرئية الفضائية عند كل عدد من عينات التدريب وذلك باستخدام الخوارزميتين SVM وRF والحصول على الخرائط LULC الموافقة لكل عدد من أعداد عينات التدريب.

٣- تقييم دقة الخرائط LULC الناتجة باعتماد عدد متغير من نقاط الاختبار يبدأ بخمس نقاط وينتهي بثلاثين نقطة تم قياسها ضمن كل صنف في المرئيات المصنّفة ومقارنتها مع مقابلاتها على المرئية SAS Planet.

٤- القيام بتقييم الدقة والمقارنات وتحليل النتائج.



المصدر: من إعداد الباحثة.

شكل (٤). خطوات المرحلة الأولى من طريقة المعالجة.

٨- النتائج والمناقشة:

٨-١- سيناريوهات اقتطاع عينات التدريب:

الأصناف التي تم اعتمادها في الدراسة موضحة في جدول (٢).

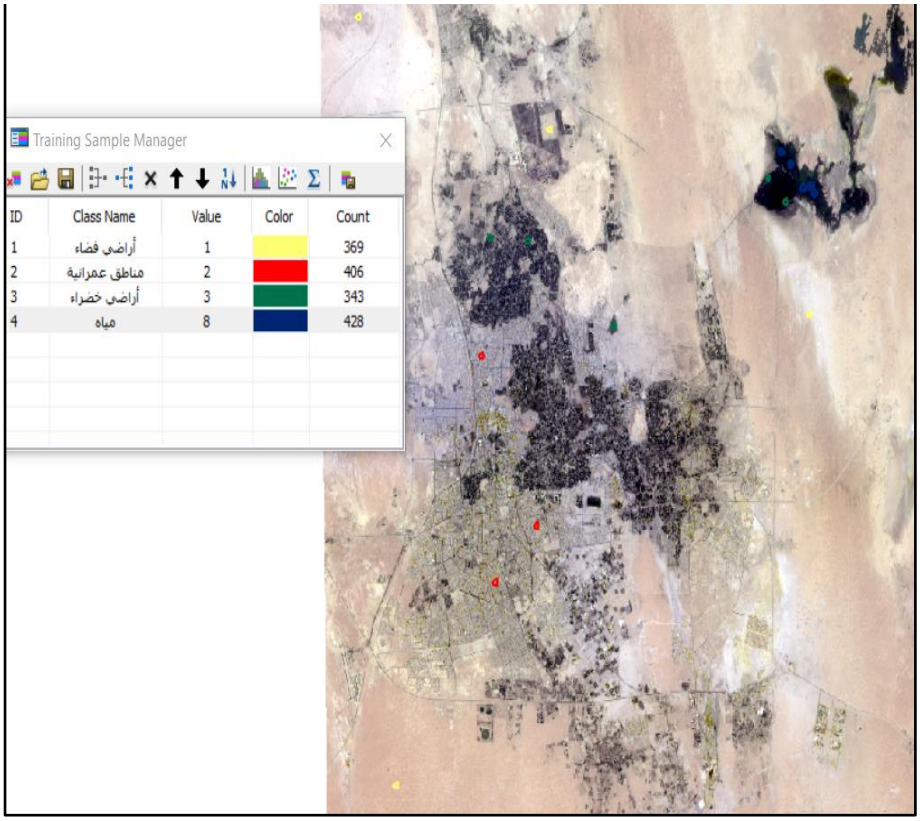
جدول (٢). أصناف استخدام الأرضي المعتمدة في البحث.

نوع الاستخدام الأرضي	وصف
مناطق عمرانية (Built-up Areas)	المناطق السكنية والمختلطة والتجارية والقرى وشبكة الطرق.
مناطق الأراضي الفضاء (Baren land)	المناطق السكنية والمختلطة والتجارية والقرى وشبكة الطرق.
مناطق خضراء (Vegetation)	الأراضي الزراعية والمراعي والنباتات الطبيعية والأشجار والحدائق والمتنزهات والملاعب.
مياه(Water)	البحيرات والسبخات

تم اقتطاع عينات التدريب التي تعبر عن الأصناف الخاصة باستخدام الأرضي بالاستعانة بالبرنامج ArcGIS ١٠,٨ حيث تم العمل وفق السيناريوهات التالية:

- اقتطاع ٢٠ عينة تدريب موزعة بمعدل ٥ عينات لكل صنف.
- اقتطاع ٤٠ عينة تدريب موزعة بمعدل ١٠ عينات لكل صنف.
- اقتطاع ٦٠ عينة تدريب موزعة بمعدل ١٥ عينة لكل صنف.
- اقتطاع ٨٠ عينة تدريب موزعة بمعدل ٢٠ عينة لكل صنف.
- اقتطاع ١٠٠ عينة تدريب موزعة بمعدل ٢٥ عينة لكل صنف.
- اقتطاع ١٢٠عينة تدريب موزعة بمعدل ٣٠ عينة لكل صنف.

تم توزيع العيّنات بشكل موحد داخل الصنف الواحد بحيث نستطيع تجاوز مشكلة التحيّز. يوضح شكل (٥) توزع هذه العينات وفق السيناريو الأول.

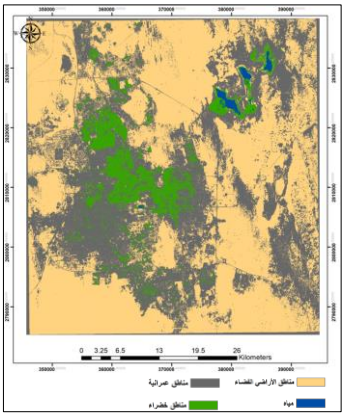
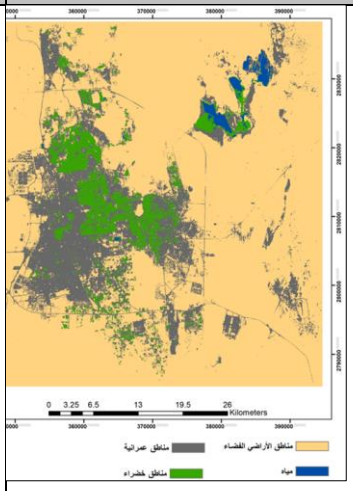
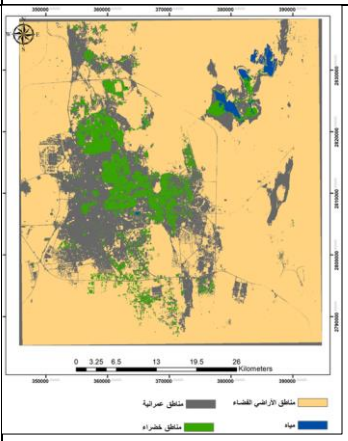
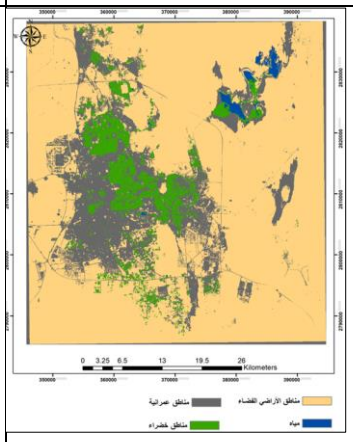
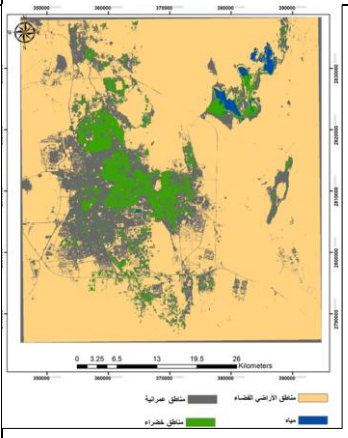
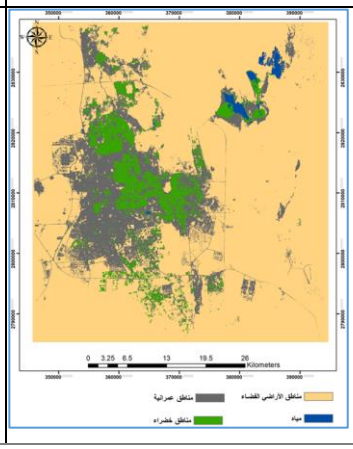


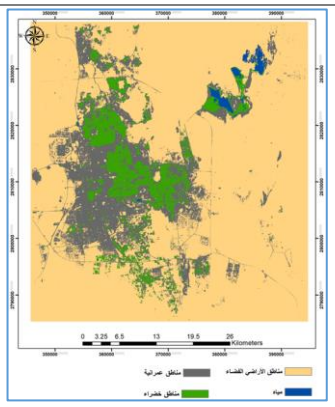
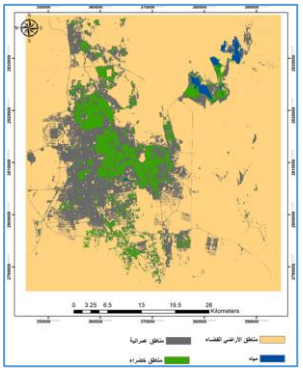
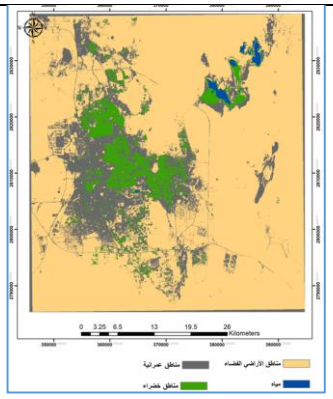
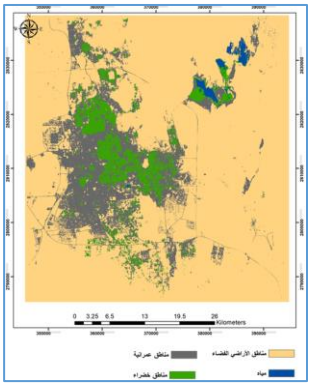
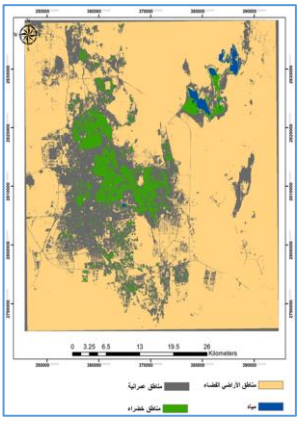
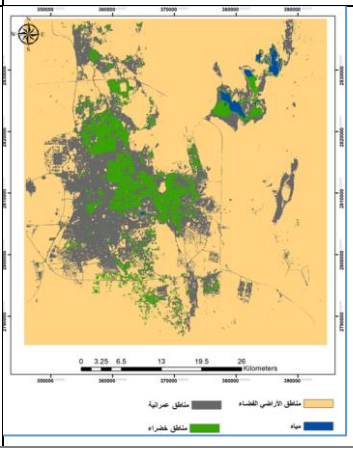
المصدر: من إعداد الباحثة.

شكل (٥). اقتطاع عينات التدريب للأصناف المختلفة بمعدّل ٥ عينات في الصنف.

٨-٢- التصنيف وتقييم الدقة:

بعد تنفيذ التصنيف المراقب بتطبيق خوارزميات SVM وRF على المرئية Landsat-8 لمنطقة الدراسة، نوضح في شكل (٦) خرائط استخدام الأراضي والغطاء الأرضي لمنطقة الدراسة الناتجة عن التصنيف وفق السيناريوهات المقترحة.

السيناريو	خوارزمية RF	خوارزمية SVM
٥ عينات تدريب		
١٠ عينات تدريب		
١٥ عينة تدريب		

٢٠ عينة تدريب		
٢٥ عينة تدريب		
٣٠ عينة تدريب		

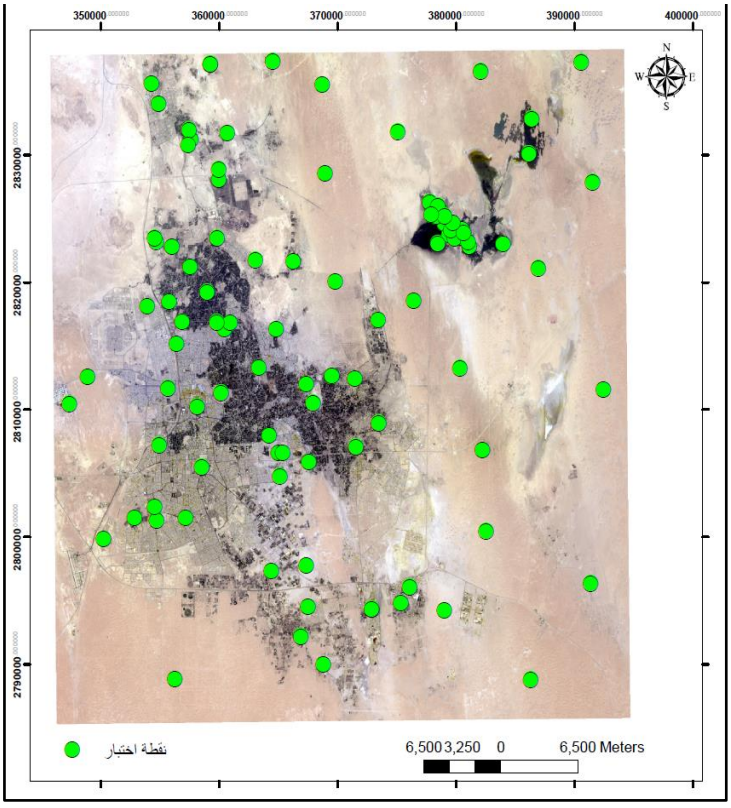
المصدر: من إعداد الباحثة.

شكل (٦). خرائط استخدام الأراضي والغطاء الأرضي LCLU الناتجة عن تطبيق كل السيناريوهات

لتقييم دقة خرائط استخدام الأراضي الناتجة عن السيناريوهات السابقة لعدد عينات التدريب، تم أولاً توليد نقاط الاختبار وفق السناريوهات التالية:

- توليد آلي لـ ٢٠ نقطة اختبار موزعة بمعدل ٥ نقاط لكل صنف.
- توليد آلي لـ ٤٠ نقطة اختبار موزعة بمعدل ١٠ نقاط لكل صنف.
- توليد آلي لـ ٦٠ نقطة اختبار موزعة بمعدل ١٥ نقطة لكل صنف.
- توليد آلي لـ ٨٠ نقطة اختبار موزعة بمعدل ٢٠ نقطة لكل صنف.
- توليد آلي لـ ١٠٠ نقطة اختبار موزعة بمعدل ٢٥ نقطة لكل صنف.
- توليد آلي لـ ١٢٠ نقطة اختبار موزعة بمعدل ٣٠ نقطة لكل صنف.

تم توزيع نقاط الاختبار بشكل متساو وموحد داخل كل صنف وذلك بالاستعانة بالأمر EQUALIZED_STRATIFIED_RANDOM المتوافر داخل النافذة **Create Accuracy Assessment Points** في البرنامج ArcGIS 10.8، ثم تم الاقتطاع الآلي للقيم الموافقة لها على المرئيات المصنفة الناتجة عن السيناريوهات الخاصة بعدد عينات التدريب. بعد ذلك، تم اقتطاع القيم المرجعية الموافقة لهذه النقاط من المرئية SAS Planet. نوضح في شكل (٧) توزع هذه العينات وفق السيناريو الخاص بـ ٢٥ نقطة اختبار في الصنف.



المصدر: من إعداد الباحثة.

شكل (٧). توزع نقط الاختبار وفق سيناريو ٢٥ نقطة في الصنف.

بعد اشتقاق خرائط استخدام الأراضي وفق السيناريوهات كلها، تم تقييم دقة هذه الخرائط وذلك باستخدام مؤشر دقة التصنيف الكلية OA والتي تُستخدم لتقييم دقة التصنيف مع مراعاة الاتفاق العرضي. نوضح في جدول (٣) نتائج تقييم دقة تصنيف المرئية باستخدام المعامل كابا k وفق الخوارزميتين SVM وRF، كما يوضح جدول (٤) نتائج تقييم دقة التصنيف للمرئية Landsat باستخدام الدقة الكلية للتصنيف OA وذلك وفق الخوارزميتين SVM وRF.

جدول (٣). تقييم دقة التصنيف باستخدام المعامل كابا

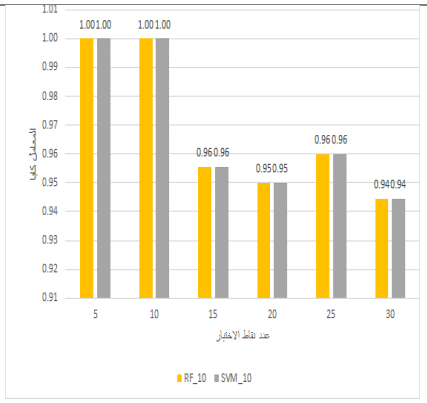
	SVM						RF					
اختبار	5	10	15	20	25	30	5	10	15	20	25	30
تدريب												
5	0.93	0.97	0.93	0.93	0.95	0.91	0.87	0.80	0.82	0.80	0.80	0.80
10	1.00	1.00	0.96	0.95	0.96	0.94	1.00	1.00	0.96	0.95	0.96	0.94
15	1.00	1.00	0.96	0.95	0.96	0.93	1.00	1.00	0.96	0.95	0.96	0.94
20	1.00	0.97	0.93	0.93	0.95	0.92	1.00	0.97	0.93	0.93	0.95	0.92
25	1.00	0.97	0.93	0.93	0.95	0.92	1.00	1.00	0.96	0.95	0.96	0.96
30	1.00	0.93	0.91	0.92	0.93	0.91	1.00	1.00	0.96	0.95	0.95	0.96

جدول (٤). تقييم دقة التصنيف باستخدام الدقة الكلية **OA**

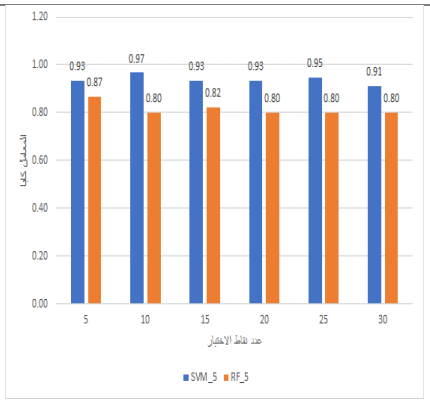
	SVM						RF					
اختبار	5	10	15	20	25	30	5	10	15	20	25	30
تدريب												
5	0.95	0.98	0.95	0.95	0.96	0.93	0.90	0.85	0.87	0.85	0.85	0.85
10	1.00	1.00	0.97	0.96	0.97	0.96	1.00	1.00	0.97	0.96	0.97	0.96
15	1.00	1.00	0.97	0.96	0.97	0.95	1.00	1.00	0.97	0.96	0.97	0.96
20	1.00	0.98	0.95	0.95	0.96	0.94	1.00	0.98	0.95	0.95	0.96	0.94
25	1.00	0.98	0.95	0.95	0.96	0.94	1.00	1.00	0.97	0.96	0.97	0.97
30	1.00	0.95	0.93	0.94	0.95	0.93	1.00	1.00	0.97	0.96	0.97	0.96

أولاً: تحليل تأثير عدد عينات التدريب

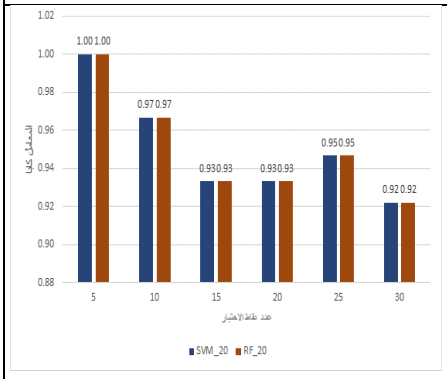
يوضح الشكلان (٨) و(٩) نتائج التحليل بناءً على عدد عينات التدريب وفقاً للسيناريوهات الستة المقترحة وذلك لتكوين رؤية واضحة حول تأثير زيادة عدد عينات التدريب على دقة خرائط استخدام الأراضي الناتجة عن الخوارزميتين SVM و RF.



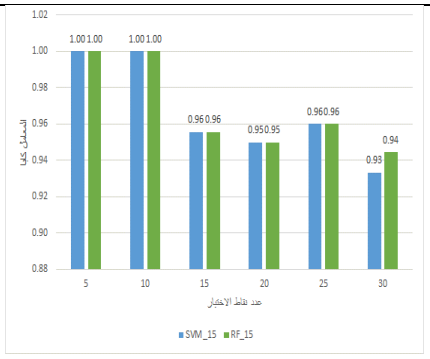
حالة ١٠ عينات تدريب في الصنف الواحد



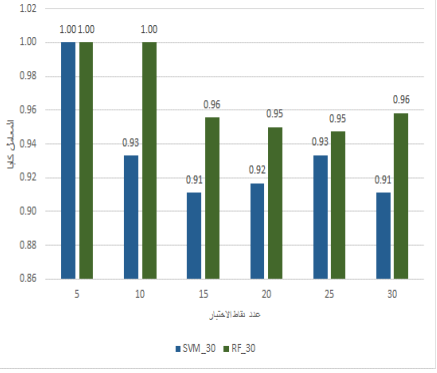
حالة ٥ عينات تدريب في الصنف الواحد



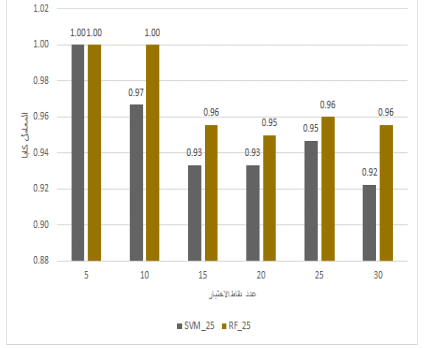
حالة ٢٠ عينة تدريب في الصنف الواحد



حالة ١٥ عينة تدريب في الصنف الواحد



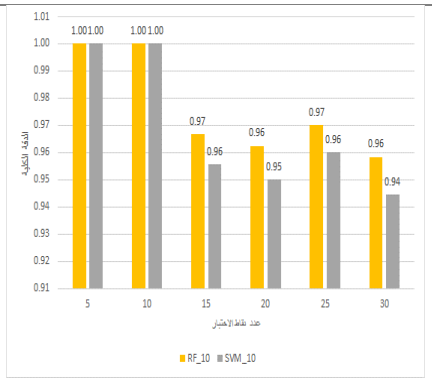
حالة ٣٠ عينة تدريب في الصنف الواحد



حالة ٢٥ عينة تدريب في الصنف الواحد

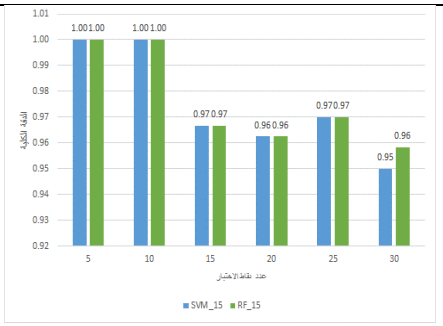
المصدر: من إعداد الباحثة.

شكل (٨). المعامل كابا للتصنيف كتابع لعدد عينات التدريب.

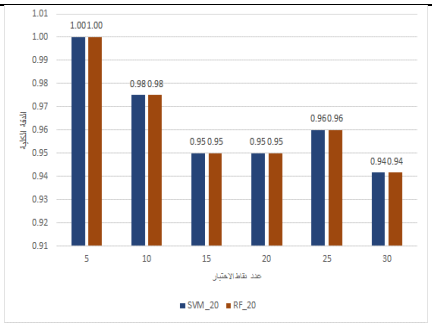


حالة ٥ عينات تدريب في الصنف الواحد

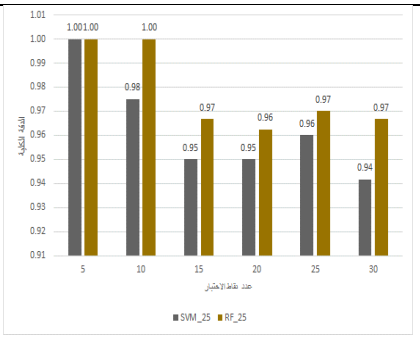
حالة ١٠ عينات تدريب في الصنف الواحد



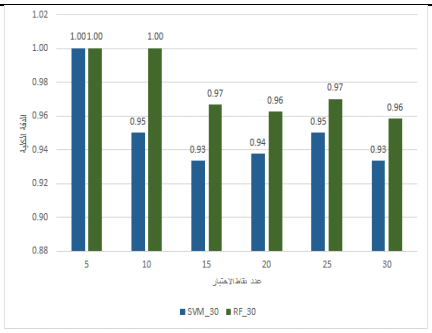
حالة ١٥ عينة تدريب في الصنف الواحد



حالة ٢٠ عينة تدريب في الصنف الواحد



حالة ٢٥ عينة تدريب في الصنف الواحد



حالة ٣٠ عينة تدريب في الصنف الواحد

المصدر: من إعداد الباحثة.

شكل (٩). الدقة الكلية للتصنيف كتابع لعدد عينات التدريب.

من شكل (٨) نلاحظ أنه بالنسبة الخوارزمية SVM:

١. عند التدريب باستخدام ٥ عينات تراوحت قيمة كابا بين ٠,٩٣ و٠,٩٧ عند استخدام من ٥ إلى ٣٠ نقطة اختبار وأقل قيمة لكابا كانت عند استخدام ٣٠ نقطة اختبار (٠,٩١).

٢. عند التدريب باستخدام ١٠ عينات كانت قيمة كابا ثابتة عند ١ حتى الوصول إلى استخدام ١٠ نقاط اختبار ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٤).

٣. عند التدريب باستخدام ١٥ عيّنة كانت قيمة كابا ثابتة عند ١ حتى الوصول إلى استخدام ١٠ نقاط اختبار ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٣).

٤. عند التدريب باستخدام ٢٠ عيّنة كانت قيمة كابا ثابتة عند ١ حتى الوصول إلى استخدام ٥ نقاط اختبار ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٢).

٥. عند التدريب باستخدام ٢٥ عيّنة نجد أن النمط مشابه للتدريب باستخدام ٢٠ عيّنة مع أقل قيمة لكابا عند استخدام ٢٠ و٣٠ نقطة اختبار (٠,٩٢).

٦. عند التدريب باستخدام ٣٠ عيّنة بدأت قيمة كابا عند ١ ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩١).

أما بالنسبة للخوارزمية RF فنلاحظ ما يلي:

١. عند التدريب باستخدام ٥ عينات تراوحت قيمة كابا بين ٠,٨ و٠,٨٧ وأقل قيمة لكابا كانت عند استخدام ١٥ إلى ٣٠ نقطة اختبار (٠,٨).

٢. عند التدريب باستخدام ١٠ عينات كانت قيمة كابا ثابتة عند ١ حتى الوصول إلى استخدام ١٠ نقاط اختبار ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٤).

٣. عند التدريب باستخدام ١٥ عيّنة كان النمط مشابهاً لحالة التدريب باستخدام ١٠ عينات مع أقل قيمة لكابا عند استخدام ٣٠ نقطة اختبار (٠,٩٤).

٤. عند التدريب باستخدام ٢٠ عيّنة كانت قيمة كابا ثابتة عند ١ حتى الوصول إلى استخدام ٥ نقاط اختبار ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٢).

٥. عند التدريب باستخدام ٢٥ عيّنة نجد أن النمط مشابه للتدريب باستخدام ٢٠ عيّنة مع أقل قيمة لكابا عند استخدام ٢٠ و٣٠ نقطة اختبار (٠,٩٦).

٦. عند التدريب باستخدام ٣٠ عيّنة بدأت قيمة كابا عند ١ ثم انخفضت قليلاً. أما أقل قيمة لكابا فكانت عند استخدام ٣٠ نقطة اختبار (٠,٩٦).

من التحليلات السابقة نستنتج أنه عند تطبيق الخوارزمية SVM تتحسن قيمة المعامل كابا بشكل ملحوظ بزيادة عدد عينات التدريب كما أن قيمة كابا تصل إلى ١ مع عينات التدريب من ١٠ إلى ٣٠ عيّنة. من ناحية أخرى نجد تناقصاً طفيفاً في الدقة عند زيادة عدد عينات الاختبار. وعند تطبيق الخوارزميةRF نجد أنها تظهر أداءً مستقراً، وتحقق معامل كابا مساو لـ ١ مع عينات التدريب من ١٠ إلى ٣٠ كما أن قيمة كابا تكون أقل عند استخدام ٥ عينات تدريب، لكنها تتحسن بسرعة مع زيادة عينات التدريب. بمعنى آخر، استفادت الخوارزميتان SVM وRF من زيادة عدد عينات التدريب، لكن RF أظهرت أداءً مستقراً وثابتاً بشكل أكبر مع زيادة عدد عينات التدريب مقارنةً بـ SVM. من ناحية أخرى، تصير دقة كلتا الخوارزميتين أعلى عند زيادة عدد عينات التدريب وتكون مستقرة عند استخدام عينات التدريب من ١٠ فأكثر.

ومن شكل (٩) نلاحظ أنه بالنسبة الخوارزمية SVM:

١. تتراوح الدقة الكلية بين ٠,٩٣ و٠,٩٨عند تدريب النموذج على مجموعة بيانات بعدد ٥.

٢. عند تدريب النموذج على مجموعة بيانات بعدد ١٠ إلى ٣٠ تكون الدقة الكلية مستقرة بشكل كبير وتصل إلى ١,٠٠ في بعض الحالات والخوارزمية تظهر أداءً ممتازاً عند أعداد البيانات الأكبر حيث تصل الدقة إلى ١,٠٠ عند زيادة عدد عينات التدريب إلى ١٠ فأكثر.

أما بالنسبة للخوارزمية RF فنلاحظ ما يلي:

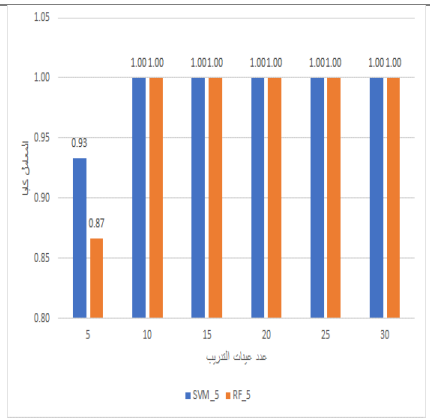
١. تتراوح الدقة الكلية بين ٠,٨٥ و٠,٩٠ عند تدريب النموذج على مجموعة بيانات بعدد ٥.

٢. عند تدريب النموذج على مجموعة بيانات بعدد ١٠ إلى ٣٠ تكون الدقة الكلية مستقرة بشكل كبير وتصل إلى ١,٠٠ في بعض الحالات والخوارزمية تظهر أداءً ممتازاً عند أعداد البيانات الأكبر حيث تصل الدقة إلى ١,٠٠ عند زيادة عدد عينات التدريب إلى ١٠ فأكثر.

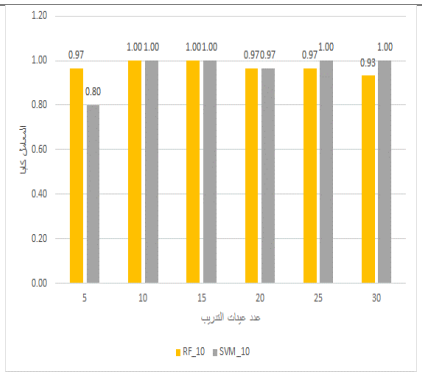
من التحليلات السابقة نستنتج أنه في حالة التدريب على مجموعة بيانات بعدد صغير (٥) تتفوق SVM على RF حيث تظهر نتائج SVM دقةً أعلى من RF، أما في حالة التدريب على مجموعة بيانات بعدد أكبر (١٠ إلى ٣٠) فإن الخوارزميتين تقدمان أداءً ممتازاً حيث الدقة تصل إلى ١,٠٠ في العديد من الحالات. وأخيراً، يمكن القول بأن كلتا الخوارزميتين مناسبتان للعمل على أعداد كبيرة من عينات التدريب، بينما SVM قد يكون الأفضل للأعداد الأصغر.

ثانياً: تحليل تأثير عدد نقاط الاختبار

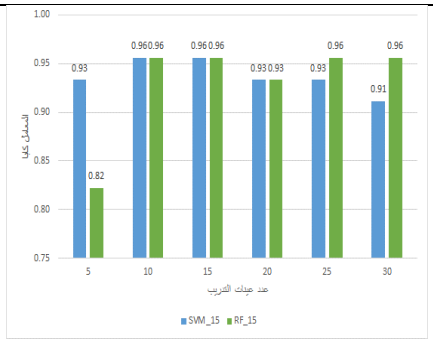
يوضح الشكلان (١٠) و(١١) نتائج التحليل بناءً على عدد نقاط الاختبار وفقاً للسيناريوهات الست المقترحة وذلك لتكوين رؤية واضحة حول تأثير زيادة عدد نقاط الاختبار على دقة خرائط استخدام الأراضي الناتجة عن الخوارزميتين SVM و RF.



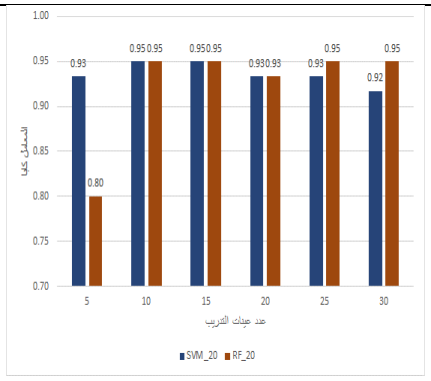
حالة ٥ نقاط اختبار في الصنف الواحد



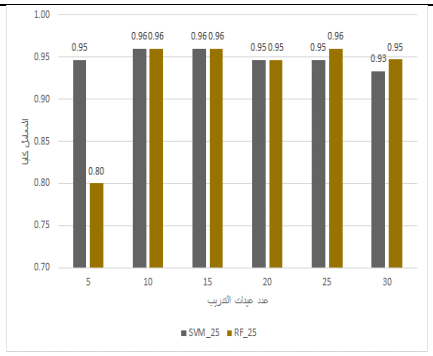
حالة ١٠ نقاط اختبار في الصنف الواحد



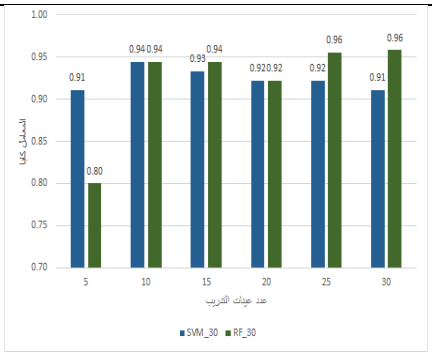
حالة ١٥ نقطة اختبار في الصنف الواحد



حالة ٢٠ نقطة اختبار في الصنف الواحد



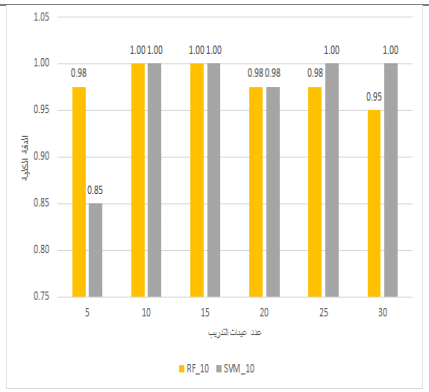
حالة ٢٥ نقطة اختبار في الصنف الواحد



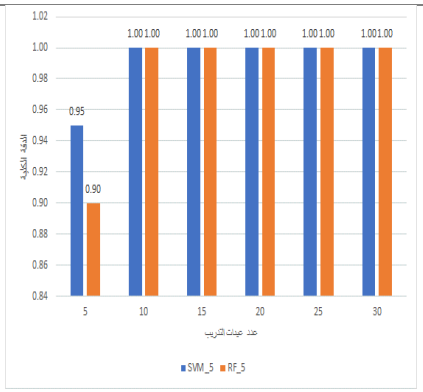
حالة ٣٠ نقطة اختبار في الصنف الواحد

المصدر: من إعداد الباحثة.

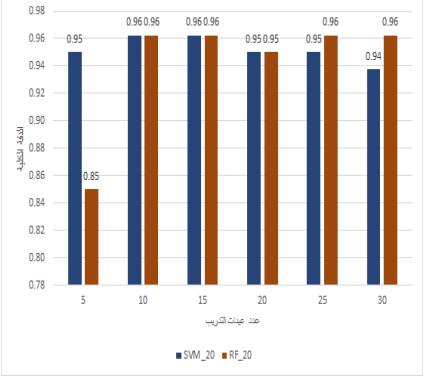
شكل (١٠). المعامل كابا للتصنيف كتابع لعدد نقاط الاختبار.



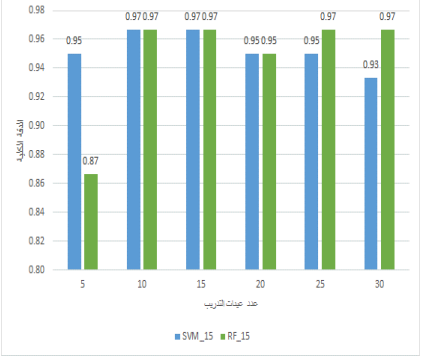
حالة ١٠ نقاط اختبار في الصنف الواحد



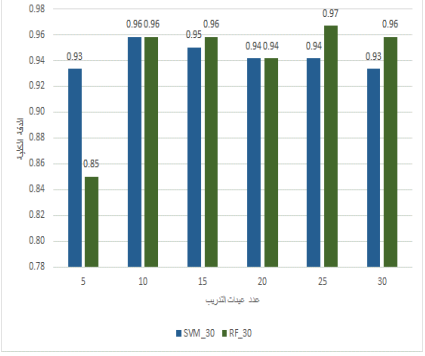
حالة ٥ نقاط اختبار في الصنف الواحد



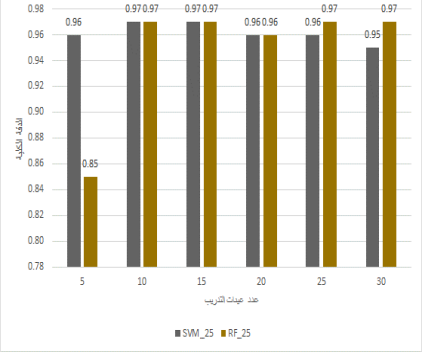
حالة ٢٠ نقطة اختبار في الصنف الواحد



حالة ١٥ نقطة اختبار في الصنف الواحد



حالة ٣٠ نقطة اختبار في الصنف الواحد



حالة ٢٥ نقطة اختبار في الصنف الواحد

المصدر: من إعداد الباحثة.

شكل (١١). الدقة الكلية للتصنيف كتابع لعدد نقاط الاختبار.

من شكل (١٠) نلاحظ أنه بالنسبة الخوارزمية SVM:

١. عند خمس عينات تدريب، تتراوح قيمة كابا بين ٠,٩١ و٠,٩٧، مع أعلى قيمة عند ١٠ نقاط اختبار (٠,٩٧).

٢. عند ١٠ عينات تدريب، وصلت قيمة كابا إلى ١,٠٠ عند ٥ و١٠ نقاط اختبار، ثم انخفضت تدريجياً إلى ٠,٩٤ عند ٣٠ نقطة اختبار.

٣. عند ١٥ و٢٠ و٢٥ و٣٠ عينة تدريب بدأت قيمة كابا من ١,٠٠ ثم انخفضت إلى ٠,٩١ عند ٣٠ نقطة اختبار.

أما بالنسبة للخوارزمية RF فنلاحظ ما يلي:

١. عند خمس عينات تدريب، تراوحت قيمة كابا بين ٠,٨٠ و٠,٨٧، وبقيت مستقرة عند معظم أعداد نقاط الاختبار.

٢. عند ١٠ إلى ٣٠ عينة تدريب، وصلت قيمة كابا إلى ١,٠٠ عند ٥ و١٠ نقاط اختبار، ثم انخفضت تدريجياً إلى ٠,٩٤-٠,٩٦ عند ٣٠ نقطة اختبار.

نتيجة للتحليلات السابقة نجد أن خوارزمية SVM تظهر أداءً ممتازاً عند أعدد عينات التدريب الأكبر (١٠ إلى ٣٠) مع قيمة كابا قريبة من ١,٠٠، ولكننا نلاحظ أيضاً تراجعاً بسيطاً في الأداء مع زيادة نقاط الاختبار.

أما بالنسبة للخوارزمية RF فنلاحظ أنها تظهر فعاليةً جيدة عند أعداد عينات التدريب الأكبر ولكنها أقل استقراراً عند التدريب على مجموعة بيانات صغيرة (٥). كما نلاحظ أن أداء هذه الخوارزمية أقل بكثير عند التدريب على مجموعة بيانات صغيرة حيث كانت قيمة كابا تكون أقل بكثير من حالة الخوارزمية SVM.

في كل الأحوال، تظهر كلتا الخوارزميتين فعالية عالية عند أعداد عينات التدريب الأكبر، مع تفوق SVM في حالات عدد عينات التدريب الصغيرة.

من شكل (١١) نلاحظ أنه بالنسبة للخوارزمية SVM:

١. عند خمس عينات تدريب، تتراوح قيمة الدقة الكلية بين ٠,٩٣ و٠,٩٨، مع أعلى قيمة عند ١٠ نقاط اختبار (٠,٩٨).

٢. عند ١٥ و٢٠ و٢٥ و٣٠ عينة تدريب، تستمر قيمة الدقة الكلية في الانخفاض مع زيادة نقاط الاختبار، لكنها تظل قريبة من ١,٠٠.

أما بالنسبة للخوارزمية RF

١. عند خمس عينات تدريب، تتراوح قيمة الدقة الكلية بين ٠,٨٥ و٠,٩٠، وتبقى ثابتة عند القيمة ٠,٨٥ عند معظم أعداد نقاط الاختبار.

٢. عند ١٠ إلى ٣٠ عينة تدريب تصل قيمة الدقة الكلية إلى ١,٠٠ عند ١٠ نقاط اختبار، ثم تنخفض تدريجياً إلى ٠,٩٦ عند ٣٠ نقطة اختبار.

نستنتج من التحليلات السابقة أن الخوارزمية SVM فعالة عند أعداد عينات التدريب الأكبر (١٠ إلى ٣٠) مع قيمة الدقة الكلية قريبة من ١,٠٠. ومع زيادة نقاط الاختبار نلاحظ تراجعاً بسيطاً في أداء هذه الخوارزمية. من ناحية أخرى، نجد أن خوارزمية RF فعالة عند أعداد عينات التدريب الأكبر (١٠ إلى ٣٠) مع قيمة الدقة الكلية قريبة من ١,٠٠، كما نلاحظ أن فعالية الخوارزمية أقل بكثير عند التدريب على مجموعة بيانات صغيرة (٥) حيث تكون قيمة كابا أقل بكثير من حالة SVM.

في كل الأحوال، تظهر كلتا الخوارزميتين فعالية عالية عند أعداد عينات التدريب الأكبر، مع تفوق SVM في حالات عدد عينات التدريب الصغيرة.

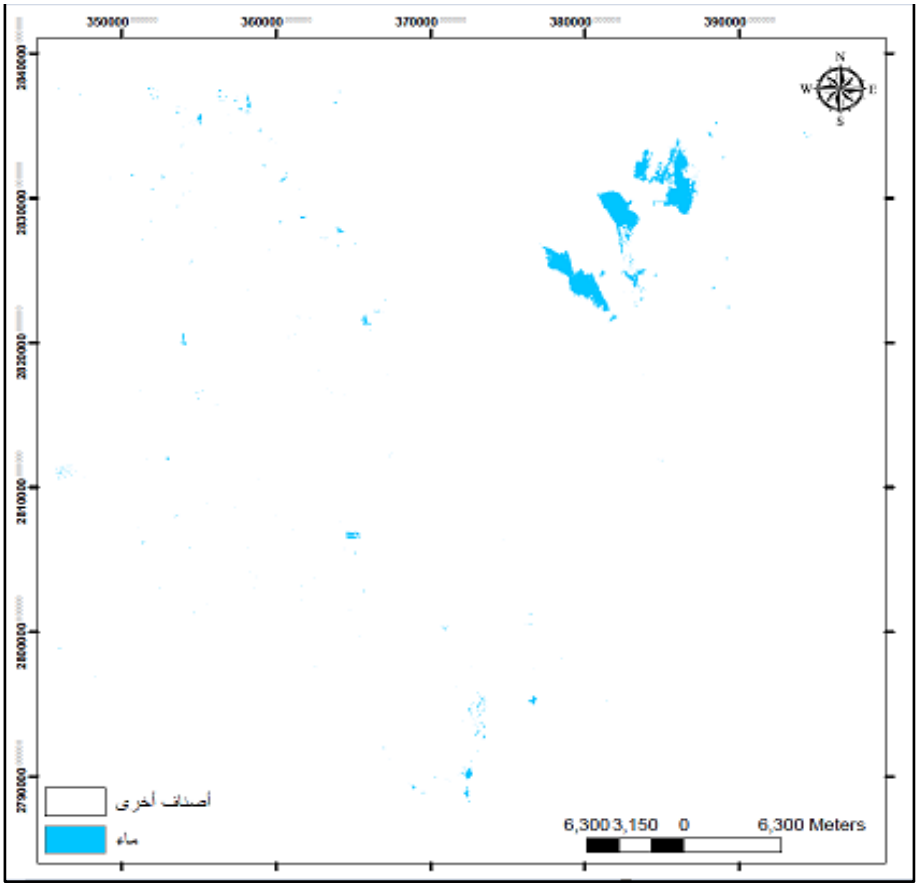
٨-٣- التقييم الهندسي للدقة:

وفي محاولة من أجل الحصول على فكرة أوضح عن تأثير عدد عينات التدريب على فعالية الخوارزميات SVM وRF قمنا بعملية تقييم هندسي للدقة

تقوم على مقارنة مساحات السطوح المائية الناتجة عن التصنيفات باستخدام مختلف أعداد عينات التدريب مع قيمة مساحة هذه الأسطح المحسوبة باستخدام مؤشر المياه المعياري Normalized Difference Water Index (NDWI) (Feyisa et al, 2014). يستخدم هذا المؤشر لتمييز المناطق المائية عن غيرها من الظاهرات الأرضية. إن الانعكاسية الطيفية للماء عاليةً في نطاق الطول الموجي الأخضر Green الذي تتراوح أطوال موجاته من ٠,٥٢ إلى ٠,٦ µm وقليلةً في نطاق الطول الموجي الأحمر القريب الذي تتراوح أطوال موجاته من ٠,٧٦ إلى ٠,٩ µm. تتراوح قيم هذا المؤشر بين −١ و+١ (Xu, 2006)، وتتوافق القيم الموجبة لهذا المؤشر مع المناطق المائية وكلما كانت قيمته أكبر كانت دلالته على هذه المناطق أكبر. ويمكن حساب هذا المؤشر من خلال العلاقة التالية:

$$NDWI = \frac{Green - NIR}{Green + NIR} \qquad (2)$$

تم تطبيق هذا المؤشر على المرئية Landsat-8 التي تغطي منطقة الدراسة بتاريخ الشهر الخامس من العام ٢٠٢٤ فحصلنا على خارطة توزع السطوح المائية في هذه المنطقة والموضحة في شكل (١٣).



المصدر: من إعداد الباحثة.

شكل (١٣). توزع السطوح المائية في منطقة الدراسة.

تم حساب مساحة هذه المسطحات بالاستعانة بالبرنامج ArcGIS 10.8 فوجدنا أنها تساوي ١٩,٦٢٥ كيلومتراً مربعاً. نظراً للدقة العالية للمؤشر NDWI في تحديد المسطحات المائية من المرئيات الفضائية (EL Ashiry, 2024)، فإنه من الممكن أن نعد هذه المساحة مساحةً مرجعية يمكن استخدامها لتقييم فعالية خوارزميات التصنيف. للقيام بذلك، تم حساب مساحة هذه المسطحات انطلاقاً من المرئيات المصنّفة وفق كل السيناريوهات. نوضح في جدول (٥) هذه المساحات.

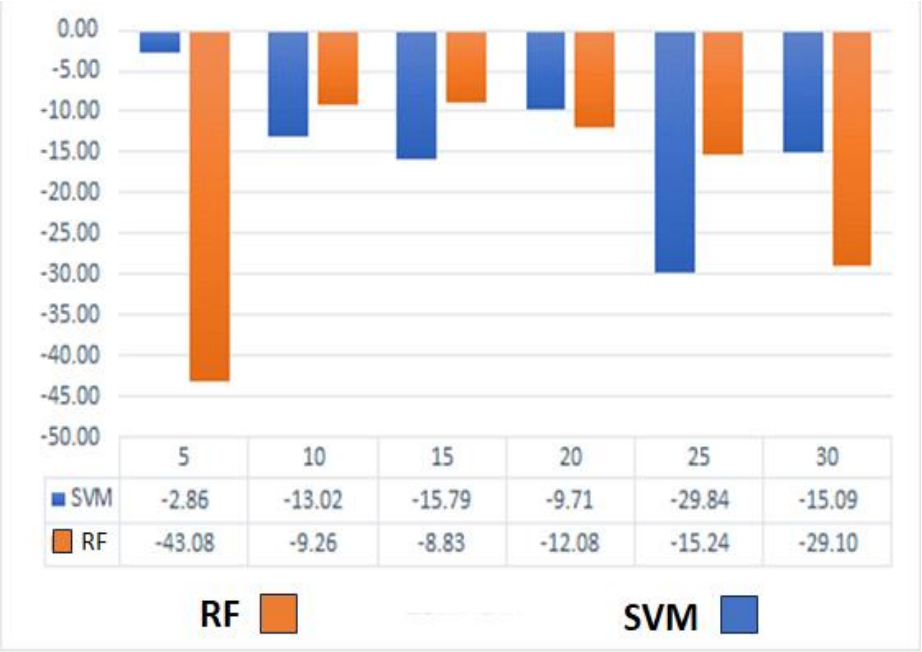
جدول (٥). مساحة المائية انطلاقاً من المرئيات المصنّفة وفق كل السيناريوهات

	خوارزمية SVM	خوارزميةRF
عدد عينات تدريب	المساحة (كيلومتر مربع)	المساحة (كيلومتر مربع)
5	19.06	11.17
10	17.07	17.81
15	16.53	17.89
20	17.72	17.26
25	13.77	16.63
30	16.66	13.91

يتضح في شكل (١٤) النسبة المئوية لتغير المساحة الحقيقة للمسطحات المائية مقارنة بالمساحات الموضحة في جدول (٥) والتي تتبع لعدد عينات التدريب ولخوارزمية التصنيف. تم حساب النسبة المئوية للتغير وفق العلاقة التالية:

$$\Delta A = \frac{(A_{Interpolated} - A_{real})}{A_{real}} \times 100 \qquad (3)$$

حيث: ΔA هي النسبة المئوية لتغير المساحة، A_{real} هي المساحة الحقيقية و $A_{Interpolated}$ هي المساحة الناتجة عن التصنيف.



المصدر: من إعداد الباحثة.

شكل (١٤). النسبة المئوية لتغير المساحة الحقيقة للمسطحات المائية كتابع لعدد عينات التدريب وللخوارزمية المطبقة.

يوضح المخطط البياني (شكل (١٤)) الاختلافات في الأداء بين الخوارزميتين عبر نطاقات مختلفة من عينات التدريب (٥، ١٠، ١٥، ٢٠، ٢٥، ٣٠). نلاحظ أنه عند عدد عينات تدريب منخفض تعطي الخوارزمية SVM تغيراً بنسبة -٢,٨٦% أما خوارزمية RF فتعطي تغيراً بنسبة -٤٣,٠٨%، أما عند زيادة عدد عينات التدريب إلى ١٠ و١٥ تعطي خوارزمية SVM عند ١٠ عينات تغيراً بنسبة -١٣,٠٢% و-١٥,٩٧% على الترتيب بينما تعطي خوارزمية SVM عند ١٠ عينات تغيراً بنسبة -٩,٢٦% و-٨,٨٣% على الترتيب. وعند ٢٠ عينة تدريب تعطي خوارزمية SVM تغيراً بنسبة -٩,٧١% وتعطي خوارزمية RF تغيراً نسبته -١٢,٠٨%. وعند زيادة

عدد العينات إلى ٢٥ عينة، تعطي خوارزمية SVM تغيراً بنسبة -٢٩,٨٤% وتعطي خوارزمية RF تغيراً نسبته -١٥,٢٤%، أما عند زيادة عدد العينات إلى ٣٠ عينة فتعطي خوارزمية SVM تغيراً بنسبة -١٥,٠٩% وتعطي خوارزمية RF تغيراً نسبته -٢٩,١٠%.

بناء على ما سبق، نجد أن أداء الخوارزمية RF يكون أسوأ بشكل ملحوظ عند عدد عينات تدريب قليل، ولكن يتحسن بشكل كبير مع زيادة عدد العينات، وعلى العكس، تظهر خوارزمية SVM أداءً أكثر استقراراً وتوازناً عبر جميع نطاقات عينات التدريب. وهذا ما يؤكد النتائج السابقة التي توصلنا إليها عند تقييم دقة التصنيف كتابع لعدد عينات التدريب.

إن النتائج التي توصل إليها البحث تتوافق مع نتائج الأبحاث السابقة التي تناولت تأثير عدد عينات التدريب على دقة التصنيف باستخدام الخوارزميات SVM وRF. نذكر من هذه الأبحاث (Ramezan, et al, 2021) والذي أشار إلى أن زيادة عدد عينات التدريب عادة ما تؤدي إلى تحسين دقة التصنيف. ومع ذلك فإن أداء الخوارزميات يختلف بناءً على طبيعة البيانات وعدد العينات المتاحة. على سبيل المثال، أظهرت الدراسات أن خوارزمية الغابات العشوائية (SVM) تميل إلى الحفاظ على دقة تصنيف عالية حتى مع انخفاض عدد عينات التدريب، بينما تظهر خوارزمية آلات المتجهات الداعمة (RF) حساسية أكبر تجاه تقليل عدد العينات، مما يؤدي إلى تراجع في الأداء عند تقليل عدد العينات. كما نجد في ((Shetty, et al., 2021) نتائج تؤكد صحة النتائج التي توصلنا إليها حيث بين هذا البحث المتعلق بتصنيف البيانات عن بعد، أن أداء خوارزمية SVM يتفوق عادةً على RF عند تقليل عدد عينات التدريب. من

ناحية أخرى، ناقش بحثنا تأثير عدد نقاط الاختبار على دقة التصنيف وهذه هي الثغرة البحثية التي حاولنا سدّها.

٩- خاتمة وتوصيات:

تم في هذا البحث دراسة تأثير عدد عينات التدريب وعدد نقاط الاختبار على دقة خرائط استخدام الأراضي والغطاء الأرضي LULC لواحة الإحساء الواقعة في المنطقة الشرقية من المملكة العربية السعودية والمشتقة من تصنيف مرئية الفضائية من النوع Landsat ٨ باستخدام خوارزميتي SVM وRF. تم في البحث تغيير عدد عينات التدريب بدءاً من خمس عينات وصولاً إلى ثلاثين عيّنة تم قياسها في كل صنف من الأصناف المتوافرة في منطقة الدراسة، ومن ثم القيام بتصنيف المرئية الفضائية عند كل عدد من عينات التدريب وذلك باستخدام الخوارزميتين SVM وRF والحصول على الخرائط LULC الموافقة لكل عدد من أعداد عينات التدريب. فيما يخص تقييم دقة الخرائط LULC الناتجة، فقد تم باستخدام مفهومي الدقة الكلية OA والمعامل كابا وذلك باعتماد عدد متغير من نقاط الاختبار يبدأ بخمس نقاط وينتهي بثلاثين نقطة تم قياسها ضمن كل صنف في المرئيات المصنّفة ومقارنتها مع مقابلاتها على المرئية SAS Planet. من ناحية أخرى، تم تقييم دقة الخرائط LULC الناتجة عبر مقارنة مساحات السطوح المائية الناتجة عن التصنيفات باستخدام مختلف أعداد عينات التدريب مع قيمة مساحة هذه السطوح المحسوبة باستخدام مؤشر المياه المعياري NDWI.

من خلال تحليل النتائج المتعلقة بتقييم دقة التصنيف باستخدام معامل كابا والدقة الكلية للخوارزميات SVM وRF، يمكن استخلاص النتائج التالية:

١- أظهرت نتائج تطبيق الخوارزمية SVM تفوقاً واضحاً في معظم الحالات، سواء عند قياس الدقة الكلية أو معامل كابا.

٢- عند استخدام عدد قليل من عينات التدريب (مثل ٥ عينات)، تحقق خوارزمية SVM دقة عالية وثابتة تتراوح بين ٠,٩٣ و١,٠٠ في معامل كابا وبين ٠,٩٥ و١,٠٠ في الدقة الكلية.

٣- حتى مع زيادة عدد نقاط الاختبار، تحافظ SVM على استقرار أداءها، مما يعكس قدرتها على التعامل مع مجموعات بيانات متنوعة بكفاءة عالية.

٤- على الرغم من أن خوارزمية RF تحقق نتائجاً جيدة، إلا أنها تظهر تذبذباً أكبر في الأداء مقارنة بـ SVM.

٥ - في حالات عينات التدريب القليلة (مثل ٥ عينات)، تكون الدقة ومعامل كابا أقل نسبياً، حيث تتراوح بين ٠,٨٠ و١,٠٠ في معامل كابا وبين ٠,٨٥ و١,٠٠ في الدقة الكلية.

٦- مع زيادة عدد نقاط الاختبار، يظهر تباين واضح في النتائج، مما يشير إلى أن RF قد تكون أقل استقراراً في الأداء.

نستنتج أخيراً أنه عند تساوي عدد عينات التدريب أو عدد نقاط الاختبار، تتفوق خوارزمية SVM بشكل عام على خوارزمية RF من حيث الدقة الكلية ومعامل كابا. تعد SVM أكثر استقراراً وأداءً، مما يجعلها الخيار الأمثل للتصنيف في معظم الحالات. ومع ذلك، يبقى استخدام خوارزمية RF مفيداً في بعض التطبيقات، خاصة إذا تم تحسين معلماتها الفائقة بعناية.

من ناحية أخرى، تم التحقق من النتائج السابقة من خلال تحليل نتائج مقارنة مساحات السطوح المائية الناتجة عن التصنيفات باستخدام مختلف أعداد عينات التدريب مع قيمة مساحة هذه السطوح المحسوبة باستخدام مؤشر المياه

المعياري NDWI. بينت النتائج أن أداء الخوارزمية RF يكون أسوأ بشكل ملحوظ عند عدد عينات تدريب قليل، ولكن يتحسن بشكل كبير مع زيادة عدد العينات، وعلى العكس، تظهر خوارزمية SVM أداءً أكثر استقراراً وتوازناً عبر جميع نطاقات عينات التدريب.

بناءً على النتائج التي تم التوصل إليها يمكن تقديم التوصيات التالية لاختيار واستخدام خوارزميات التصنيف بشكل فعال:

١- يُفضّل استخدام SVM في الحالات التي تتطلب دقة عالية واستقراراً في الأداء. تعتبر SVM الخيار الأمثل عندما يكون لدينا مجموعة بيانات صغيرة أو متوسطة وتحتاج إلى نتائج موثوقة وثابتة.

٢- يمكن استخدام RF عندما تكون البيانات كبيرة العدد ومتنوعة، وعندما يمكن تحسين المعلمات الفائقة بعناية لتحقيق أداء أفضل.

٣- اختبار خوارزميات تصنيف مراقب قائمة على تعلم الآلة مثل خوارزمية الشبكة العصبونية الصناعية Artificial neural network وذلك للوصول إلى نتائج أكثر شموليةً.

١٠- المراجع:

المراجع الأجنبية

1. Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31.
2. Boser, B.E., Guyon, I.M., Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers, in: proceedings of the fifth annual workshop on computational learning theory, ACM New York, NY, USA. 144–152.
3. Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.
4. Chugh, R. S., Bhatia, V., Khanna, K., & Bhatia, V. (2020, January). A comparative analysis of classifiers for image classification. In 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 248-253). IEEE.
5. Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote sensing of environment, 37(1), 35-46.
6. EL Ashiry, A. (2024). Evaluating the Accuracy of Change detection Using Spectral Indices and spectral classification, A Case Study of Fayoum Governorate, Egypt. JES. Journal of Engineering Sciences, 52(4), 212-232.
7. Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. Remote sensing of environment, 140, 23-35.
8. Figueroa, R. L., Zeng-Treitler, Q., Kandula, S., & Ngo, L. H. (2012). Predicting sample size required for classification performance. BMC medical informatics and decision making, 12, 1-10.

9. Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. Remote sensing of Environment, 47(3), 362-368.

10. Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1), 185-201.

11. Fu, Y., Shen, R., Song, C., Dong, J., Han, W., Ye, T., & Yuan, W. (2023). Exploring the effects of training samples on the accuracy of crop mapping with machine learning algorithm. Science of Remote Sensing, 7, 100081.

12. Grabocka, J., Wistuba, M., & Schmidt-Thieme, L. (2016). Fast classification of univariate and multivariate time series through shapelet discovery. Knowledge and information systems, 49, 429-454.

13. Haykin, S. (2009) Neural networks and learning machines, 3rd edn. Pearson, Upper Saddle Rive.

14. Koggalage, R., & Halgamuge, S. (2004). Reducing the number of training samples for fast support vector machine classification. Neural Information Processing-Letters and Reviews, 2(3), 57-65.

15. Laban, N., Abdellatif, B., Ebeid, H. M., Shedeed, H. A., & Tolba, M. F. (2019). Machine learning for enhancement land cover and crop types classification. Machine learning paradigms: theory and application, 71-87.

16. Maulik, U., & Chakraborty, D. (2017). Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques. IEEE Geoscience and Remote Sensing Magazine, 5(1), 33-52.

17. Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39 (9), 2784-2817.

18. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66 (3), 247-259.

19. Myburgh, G., & Van Niekerk, A. (2014). Impact of training set size on object-based land cover classification: A comparison of three classifiers. International Journal of Applied Geospatial Research (IJAGR), 5(3), 49-67.

20. Nemmour, H., & Chibani, Y. (2006). Multiple support vector machines for land cover change detection: An application for mapping urban extensions. ISPRS Journal of Photogrammetry and Remote Sensing, 61(2), 125-133.

21. Ouchra, H., Belangour, A., & Erraissi, A. (2022). Machine learning for satellite image classification: A comprehensive review. In 2022 International Conference on Data Analytics for Business and Industry (ICDABI) (pp. 1-5). IEEE.

22. Rahman, A., Abdullah, H. M., Tanzir, M. T., Hossain, M. J., Khan, B. M., Miah, M. G., & Islam, I. (2020). Performance of different machine learning algorithms on satellite image classification in rural and urban setup. Remote Sensing Applications: Society and Environment, 20, 100410.

23. Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. Remote Sensing, 13(3), 368.

24. Richards, J. A., & Richards, J. A. (2022). Supervised classification techniques. Remote sensing digital image analysis, 263-367.

25. Santamaria-Artigas, A., Vermote, E. F., Franch, B., Roger, J. C., & Skakun, S. (2021). Evaluation of the AVHRR surface reflectance long term data record between 1984 and 2011. International Journal of Applied Earth Observation and Geoinformation, 98, 102317.

26. Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018 (pp. 99-111). Springer Singapore.

27. Shetty, S., Gupta, P. K., Belgiu, M., & Srivastav, S. K. (2021). Assessing the effect of training sampling design on the performance of machine learning classifiers for land cover mapping using multi-temporal remote sensing data and google earth engine. Remote Sensing, 13(8), 1433.

28. Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. Remote sensing of Environment, 62(1), 77-89.

29. Stehman, S. V., & Foody, G. M. (2009). Accuracy assessment. The SAGE handbook of remote sensing, 297-309.

30. Taha, L. E. D. (2016). Classifier ensemble for improving land cover classification. International Journal of Circuits, Systems and Signal Processing, 10.

31. Taufik, A., Ahmad, S. S. S., Ahmad, A. (2016). Classification of landsat 8 satellite data using NDVI thresholds. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 8(4), 37-40.

32. Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. Sensors, 18(1), 18.

33. Tuia, D., Volpi, M., Copa, L., Kanevski, M., & Munoz-Mari, J. (2011). A survey of active learning algorithms for supervised remote sensing image classification. IEEE Journal of Selected Topics in Signal Processing, 5(3), 606-617.

34. Vanbelle, S. (2016). A new interpretation of the weighted kappa coefficients. Psychometrika, 81(2), 399-410.

35. Wang, L., Sousa, W. P., & Gong, P. (2014). Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. International Journal of Remote Sensing, 35(5), 1739-1759.

36. Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, 27(14), 3025-3033.

37. Yang, D., Fu, C.-S., Smith, A.C., Yu, Q. (2017). Open land-use map: A regional land-use mapping strategy for incorporating Open-StreetMap with earth observations. Geo-spatial Information Science., 20930, 269–281.

38. Zhang, H., He, J., Chen, S., Zhan, Y., Bai, Y., & Qin, Y. (2023). Comparing Three Methods of Selecting Training Samples in Supervised Classification of Multispectral Remote Sensing Images. Sensors, 23(20), 8530.

المراجع العربية

١. وزارة الشؤون البلدية والقروية، (٢٠١٩)، تقرير الرؤية العمرانية الشاملة لمحافظة الإحساء، https://saudiarabia.un.org/ar/40086 .