

Comparison of Support Vector Machine and Random Forest Algorithms in Developing Wheat Cultivation Maps (Case Study: Hashtroud County, East Azerbaijan Province)

Sayyad Asghari Sareskanrud^{1✉}, Hamid Soleimani Youzband², Aboozar Sadeghi³

1. Corresponding author, Department of Physical Geography, Faculty of Social Sciences, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: s.sayyad@uma.ac.ir
2. Department of Physical Geography, Faculty of Social Sciences, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: Hamidsoleimani6606@gmail.com
3. Department of Physical Geography, Faculty of Social Sciences, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: aboozarsadeghi@uma.ac.ir

Article Info	ABSTRACT
Article type: Research Article Article history: Received: 11 July 2024 Revised: 15 October 2025 Accepted: 11 September 2024 Published: 21 November 2025 Keywords: Wheat, Machine Learning, Support Vector Machine, Random Forest, Cultivation Map.	Objective: Cereals are among the most important sources of dietary protein, with wheat being a significant cereal crop with high protein content. Currently, the rapid and excessive population growth, coupled with the perceived scarcity of resources available to meet basic human needs, poses one of the greatest challenges facing the world. The foundation of proper planning and management in agricultural affairs relies on having accurate and up-to-date statistics and information on agricultural capacities. Hashtroud is one of the southern counties of East Azerbaijan Province and serves as the center of Hashtroud County. This city is located between the geographical coordinates of 36°45' to 37°24' north latitude and 46°25' to 47°24' east longitude. The lack of research in the field of creating wheat cultivation maps using machine learning algorithms in Iran and the study area prompted us to undertake research in this area, aiming to develop wheat cultivation maps in the study area using the selected machine learning methods. Methods: In this study, Sentinel2-L2A satellite images were first acquired, and the Normalized Difference Vegetation Index (NDVI) was extracted using the set of images containing terrestrial reflectance data. Then, using the R programming language in the Jupyter Notebook environment, Support Vector Machine (SVM) and Random Forest (RF) classification algorithms were applied to the images. Finally, the outputs of both algorithms were analyzed in ArcMap software, and the final maps were generated. Results: The evaluation of the generated maps' output revealed that the Random Forest algorithm, with an overall accuracy of 93% and a Kappa coefficient of 87%, performed better and more effectively compared to the Support Vector Machine algorithm, which had an overall accuracy of 90% and a Kappa coefficient of 82%. This preference is due to the higher accuracy and Kappa coefficient of the Random Forest algorithm compared to other algorithms, indicating a greater agreement with reality and higher precision in predictions. Conclusions: The analysis of the findings from this research indicated that each algorithm has its own specific strengths and weaknesses. The Support Vector Machine (SVM) algorithm is often used due to its simple structure and effective performance in many classification tasks. However, in this study, it performed worse compared to the other algorithm, the Random Forest. The Random Forest algorithm, known for its ability to combine different models and reduce the effect of overfitting, typically provides accurate results. Nevertheless, its high computational complexity can be problematic in larger-scale applications.

Cite this article: Asghari Sareskanrud, S., Soleimani Youzband, H., & Sadeghi, A. (2025). Comparison of Support Vector Machine and Random Forest Algorithms in Developing Wheat Cultivation Maps (Case Study: Hashtroud County, East Azerbaijan Province). *Journal of Geography and Planning*, 31 (93), 57-79.
<http://doi.org/10.22034/gp.2024.62456.3278>



© The Author(s).

DOI: <http://doi.org/10.22034/gp.2024.62456.3278>

Publisher: University of Tabriz.

Introduction

Cereals are considered one of the most important sources of dietary protein, with wheat being a major high-protein cereal crop. Due to its high yield, the ability to be cultivated in most parts of the world, excellent baking properties, nutritional and industrial value, and superior gluten quality for bread production, no other cereal can compete with wheat. The imbalance between the supply of food and the growth of the population in many countries, especially in Africa and the Middle East, is a serious issue. Given the population increase and the limited arable land, expanding the cultivated area is not feasible, and thus, increasing the yield per unit area must be planned. Therefore, timely and accurate mapping of wheat production on a large spatial scale has significant implications for global food security warnings, optimizing agricultural land use, and decision-making. One of the modern and up-to-date methods for this purpose is the use of remote sensing techniques, which enable continuous monitoring and land-use changes across various resources. Considering the current conditions and the opinions of experts from relevant organizations, the lack of accurate and sufficient data on wheat cultivated areas and changing cropping patterns underscores the necessity of this research. Hashtroud County in East Azerbaijan Province was selected as the study area to leverage modern remote sensing technology based on satellite images for rapid and cost-effective identification and determination of cultivated areas. Ultimately, the lack of research on developing wheat cultivation maps using machine learning algorithms in Iran and the study area prompted us to conduct this research to create wheat cultivation maps in the study area using the desired machine learning methods.

Materials and Methods

In this study, Sentinel2-L2A satellite images were initially downloaded, and the Normalized Difference Vegetation Index (NDVI) was extracted using the set of images containing ground reflectance data. Subsequently, to perform classifications based on the Support Vector Machine (SVM) and Random Forest (RF) algorithms, this set of images was imported into various machine learning libraries of the R programming language in the Jupyter Notebook environment. After extracting the study area from the downloaded images, supervised classifications (SVM and RF) were applied to them. Finally, the validation stages of the classifications were performed in the Rstudio software, and the overall accuracy and kappa coefficient values were calculated. After observing, analyzing, and interpreting the results, the best map from each algorithm was selected and compared. The final stage of creating the output maps was done in ArcMap software, and the area of the wheat fields was also determined.

Results

In this section, we will examine and compare the results obtained from different algorithms to determine the best algorithm based on the criteria of accuracy and kappa coefficient. The aim of this comparison is to select the algorithm that demonstrates the best performance on the data under review. The comparison criteria for the algorithms include two main factors: accuracy and kappa coefficient. Accuracy represents the percentage of correct predictions out of the total predictions. This criterion indicates the algorithm's ability to correctly identify samples. The kappa coefficient is a measure for evaluating the agreement between the algorithm's predictions and reality, eliminating the effect of chance. The kappa coefficient ranges from -1 to 1, with values closer to 1 indicating higher agreement. According to the tables, it can be observed that the Random Forest algorithm, with an overall accuracy of 93% and a kappa coefficient of 87%, performed better and more appropriately compared to the Support Vector Machine algorithm, which had an overall accuracy of 90% and a kappa coefficient of 82%. This selection is due to its higher accuracy and greater kappa coefficient, indicating higher agreement with reality and greater accuracy in predictions.

Conclusion

In this study, two algorithms, Support Vector Machine (SVM) and Random Forest (RF), were examined. The results from these algorithms showed that each has its own strengths and weaknesses. The Support Vector Machine algorithm, with an accuracy of 90% and a kappa coefficient of 82%, demonstrated acceptable performance. This algorithm is used in many classification problems due to its simple structure and adequate efficiency. However, in this study, it performed weaker compared to the other algorithm, the Random Forest. The Random Forest algorithm, with an accuracy of 93% and a kappa coefficient of 87%, showed excellent performance. This algorithm typically provides accurate results due to its ability to combine different models and reduce the effect of overfitting. Nevertheless, its high computational complexity can be problematic in larger applications. Overall, comparing the results of this study with similar research indicates that the Random Forest algorithm generally performs better than the Support Vector Machine algorithm in classifying agricultural data. These findings confirm that the RF algorithm, due to its high capability in combining different models and reducing overfitting, can be used as an effective method in analyzing agricultural data. On the other hand, the SVM algorithm, with its decent accuracy and efficiency, remains an efficient method for many classification problems, though it may sometimes underperform compared to RF. Given the importance of wheat as a high-protein cereal crop and its critical role in meeting global food needs, using accurate and effective methods in analyzing data related to this crop is of great significance. The results of this study show that using the Random Forest algorithm can help improve the accuracy and efficiency of analyses, making it a powerful tool for agricultural management and planning.

مقایسه الگوریتم‌های ماشین‌بردار پشتیبان و جنگل تصادفی در تهیه نقشه کشت گندم (مطالعه موردی: شهرستان هشتروند استان آذربایجان شرقی)

صیاد اصغری سراسکانرود^۱✉، حمید سلیمانی یوزبند^۲، ابوذری صادقی^۳

۱. نویسنده مسئول، استاد، گروه جغرافیای طبیعی، دانشکده علوم اجتماعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: s.sayyad@uma.ac.ir

۲. دانشجوی کارشناسی ارشد، گروه جغرافیای طبیعی، دانشکده علوم اجتماعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: hamidsoleimani6606@gmail.com

۳. دانشجوی دکتری، گروه جغرافیای طبیعی، دانشکده علوم اجتماعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: aboozarsadeghi@uma.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۳/۰۴/۲۱ تاریخ بازنگری: ۱۴۰۴/۰۷/۲۳ تاریخ پذیرش: ۱۴۰۳/۰۶/۲۱ تاریخ انتشار: ۱۴۰۴/۰۸/۳۰ کلیدواژه‌ها: گندم، یادگیری ماشین، ماشین‌بردار پشتیبان، جنگل تصادفی، نقشه کشت.	هدف: غلات از مهم‌ترین منابع پروتئین غذایی به حساب می‌آیند و گندم یکی از محصولات غله‌ای مهم با پروتئین بالا است. در حال حاضر رشد روزافزون و بیش از حد جمعیت و احساس کمبود منابع در دسترس برای رفع نیازهای اساسی و مهم انسان‌ها، یکی از بزرگترین مشکلات روبه‌روی جهانیان است. پایه و اساس برنامه‌ریزی و مدیریت صحیح در امورات کشاورزی را به داشتن آمار و اطلاعات دقیق و به‌روز از ظرفیت‌های کشاورزی می‌توان دانست. هشتروند یکی از شهرستان‌های جنوبی استان آذربایجان شرقی و مرکز شهرستان هشتروند است. این شهر در بین مختصات جغرافیایی ۳۶ درجه و ۴۵ دقیقه الی ۳۷ درجه و ۲۴ دقیقه عرض شمالی و ۴۶ درجه و ۲۵ دقیقه الی ۴۷ درجه و ۲۴ دقیقه طول شرقی واقع شده‌است. مسئله عدم انجام پژوهش در زمینه تهیه نقشه کشت گندم با الگوریتم‌های یادگیری ماشین در کشور ایران و منطقه مورد مطالعه، ما را وادار به انتخاب و انجام پژوهش در این زمینه کرد تا نقشه کشت گندم و در منطقه مورد مطالعه با روش‌های مدنظر یادگیری ماشین تهیه کنیم. روش پژوهش: در این پژوهش، ابتدا تصاویر ماهواره Sentinel2-L2A دانلود شده و با استفاده از مجموعه تصاویر دارای اطلاعات بازتاب زمینی، شاخص پوشش گیاهی نرمال شده استخراج شد. سپس با استفاده از زبان برنامه‌نویسی R در محیط Jupyter Notebook الگوریتم‌های طبقه‌بندی ماشین‌بردار پشتیبان و جنگل تصادفی روی تصاویر اعمال شدند و در آخر خروجی‌های هر دو الگوریتم در نرم‌افزار Arcmap تجزیه و تحلیل شده و از نقشه‌های نهایی خروجی گرفته شد. نتایج: ارزیابی خروجی نقشه‌های تهیه شده نشان داد که الگوریتم جنگل تصادفی با میزان دقت کلی ۹۳ درصد و ضریب کاپای ۸۷ درصد عملکرد بهتر و مناسب‌تری نسبت به الگوریتم ماشین‌بردار پشتیبان با دقت کلی ۹۰ درصد و ضریب کاپای ۸۲ درصد داشته است. این انتخاب به دلیل دقت بالاتر و ضریب کاپای بیشتر آن نسبت به سایر الگوریتم‌ها است که نشان‌دهنده توافق بیشتر با واقعیت و دقت بالاتر در پیش‌بینی‌ها می‌باشد. نتیجه‌گیری: تحلیل یافته‌های این تحقیق نشان داد که هر الگوریتم دارای نقاط قوت و ضعف خاص خود می‌باشد. الگوریتم ماشین‌بردار پشتیبان به دلیل ساختار ساده و کارایی مناسب در بسیاری از مسائل طبقه‌بندی مورد استفاده قرار می‌گیرد. با این حال، در این پژوهش در مقایسه با دیگر الگوریتم‌ها یعنی الگوریتم جنگل تصادفی عملکرد ضعیف‌تری داشت. الگوریتم جنگل تصادفی نیز به دلیل توانایی در ترکیب مدل‌های مختلف و کاهش اثر بیش‌برازش معمولاً نتایج دقیقی ارائه می‌دهد. با این حال، پیچیدگی محاسبات بالای آن می‌تواند در کاربردهای بزرگ‌تر مشکل‌ساز باشد.

استناد: اصغری سراسکانرود، صیاد؛ سلیمانی یوزبند، حمید؛ و صادقی، ابوذری (۱۴۰۳). مقایسه الگوریتم‌های ماشین‌بردار پشتیبان و جنگل تصادفی در تهیه نقشه کشت

گندم (مطالعه موردی: شهرستان هشتروند استان آذربایجان شرقی). *جغرافیا و برنامه‌ریزی*، ۳۱ (۹۳)، ۵۷-۷۹

<http://doi.org/10.22034/gp.2024.62456.3278>



مقدمه

غلات از مهم‌ترین منابع پروتئین غذایی به حساب می‌آیند و گندم یکی از محصولات غله‌ای مهم با پروتئین بالا است (آلتینباک و دوپونت^۱، ۲۰۰۳). هیچ غله‌ای به علت بازدهی زیاد و امکان کشت گندم در بیشتر نقاط جهان و نیز قابلیت پخت و خاصیت تغذیه‌ای و صنعتی و کیفیت فوق‌العاده گلوتن آن برای تولید نان نمی‌تواند با آن رقابت کند (رجب‌زاده، ۱۳۸۰). گندم که به‌طور متوسط فراهم‌کننده‌ی ۲۲ درصد کل پروتئین در رژیم غذایی انسان است، بیشترین سطح زیرکشت و تولید را در دنیا به خود اختصاص داده‌است (فائو^۲، ۲۰۱۳). گندم احتمالاً نخستین گیاه زراعی است که توسط انسان کشف شده است. این فرایند به طور احتمالی در حدود ۱۲-۱۸ هزار سال قبل از میلاد آغاز شده است. قدیمی‌ترین نمونه گندم کشت شده، در حفاریات اطراف رودخانه نیل کشف شده و به حدود ۴۰۰۰ سال قبل از میلاد مسیح باز می‌گردد. براساس بررسی‌ها، لینه گندم را متعلق به کوه‌های اورال و اودیسه آن را مرتبط با سیسیل می‌داند. در حال حاضر رشد روزافزون و بیش از حد جمعیت و احساس کمبود منابع در دسترس برای رفع نیازهای اساسی و مهم انسان‌ها، یکی از بزرگترین مشکلات روبه‌روی جهانیان است. پایه و اساس برنامه‌ریزی و مدیریت صحیح در امورات کشاورزی را به داشتن آمار و اطلاعات دقیق و به‌روز از ظرفیت‌های کشاورزی می‌توان دانست (عبدالله‌زاده و نصیری، ۱۳۸۷). آگاهی از چگونگی تناسب و انطباق فعالیت‌های کشاورزی هر منطقه با شرایط آب و هوایی، لازمه هرگونه فعالیت کشاورزی است (مجرد و همکاران، ۱۳۹۶). گندم تقریباً ۲۰ درصد از کل کالری و پروتئین مورد نیاز رژیم غذایی را برای ۴/۵ میلیارد نفر با سطح برداشت ۲۱۵/۹ میلیون هکتار تامین می‌کند (فائو، ۲۰۱۹). با رشد بی‌سابقه‌ی جمعیت جهان به حدود ۹ میلیارد نفر در سال ۲۰۵۰، تولید گندم مستلزم دو برابر شدن مصرف آن تقاضا است (نلسون و همکاران، ۲۰۱۰). پیش‌بینی می‌شود که ۲۵ تا ۷۰ درصد این مهم محقق شود (تیلمن و همکاران، ۲۰۱۱). با این حال، در سال‌های اخیر، ۳۷ درصد از مناطق برداشت جهانی گندم، رکود عملکرد را تجربه کرده‌اند و توانایی تولید کشاورزی برای رضایت از افزایش تقاضای جهانی را به چالش کشیده‌اند (ری و همکاران، ۲۰۱۲). نبود تناسب میان عرضه مواد غذایی و رشد جمعیت در بسیاری از کشورهای جهان به‌ویژه در آفریقا و خاورمیانه مسئله‌ای جدی است (کومو و همکاران، ۲۰۱۷). با توجه به افزایش جمعیت و محدود بودن اراضی قابل کشت، افزایش سطح زیرکشت مقدور نبوده و افزایش مقدار عملکرد در واحد سطح اراضی قابل کشت باید برنامه‌ریزی شود (گاشاو و همکاران، ۲۰۱۸؛ مینتا و همکاران، ۲۰۱۸؛ سیدمحمدی و همکاران، ۲۰۱۹). بنابراین، نقشه‌برداری به موقع و دقیق از تولید گندم در ابعاد فضایی بزرگ پیامدهای مهمی برای هشدار امنیت غذایی جهانی، بهینه‌سازی کاربری اراضی کشاورزی و تصمیم‌گیری دارد (نلسون و بورچ فیلد، ۲۰۲۱). مطالعات متعددی پیش‌بینی عملکرد محصول را با استفاده از شاخص‌های ماهواره‌ای، متغیرهای آب و هوایی و ویژگی‌های خاک انجام داده‌اند (چای و همکاران، ۲۰۱۹؛ چائو و همکاران، ۲۰۲۱؛ ژئونگ، کو و یئوم^{۱۲}، ۲۰۲۲). روش‌های سنجش از دور اغلب به روابط کالیبره شده بین شاخص‌های گیاهی (VIS) و میزان محصول بستگی دارد (لای و همکاران، ۲۰۱۸؛ پنگ و همکاران، ۲۰۲۰). با این حال، VIS اندازه‌گیری‌های غیرمستقیم عملکرد را فراهم می‌کند (رمبولد و همکاران، ۲۰۱۳). متغیرهای آب و هوایی به جای اغلب شاخص‌های مبتنی بر ماهواره،

¹ Altenbach, Dupont² FAO³ Nelson and et.al⁴ Tilman and et.al⁵ Ray and et.al⁶ Kumm and et.al⁷ Gashaw and et.al⁸ Minta and et.al⁹ Nelson and Burchfield¹⁰ Cai and et.al¹¹ Cao and et.al¹² Jeong, Ko and Yeom¹³ Lai and et.al¹⁴ Peng and et.al¹⁵ Rembold and et.al

امکان دریافت مستقیم اطلاعات زیست محیطی در مورد وضعیت رشد محصول را فراهم می‌کنند (چای و همکاران^۱، ۲۰۱۹؛ لوو و همکاران^۲، ۲۰۲۱). تخمین سطح زیرکشت محصولات کشاورزی اغلب به روش‌های تخمین کارشناسی، برآورد از طریق فهرست برداری و استفاده از تکنولوژی‌های جدید (سنجش از دور، سامانه اطلاعات جغرافیایی و ...) صورت می‌پذیرد. از لحاظ دقت، روش تخمین کارشناسی دارای دقت بسیار کمی می‌باشد و نمی‌توان از آن در مسائل مهم کشاورزی استفاده کرد. در روش برآورد از طریق فهرست برداری، هرچند سطح زیرکشت محصول تخمین زده می‌شود ولی این روش قادر به نمایش پراکندگی محصول نمی‌باشد و همچنین هزینه‌ی بالای نیروی انسانی یکی دیگر از عوامل محدود کننده این روش می‌باشد. طبقه‌بندی کاربری اراضی یکی از مهم‌ترین موضوعات مطرح در بحث سنجش از دور است که صحت طبقه‌بندی و الگوریتم بکار رفته در طبقه‌بندی نقش مهمی را در کاربرد و ارزش نقشه‌های ایجاد شده، ایفا می‌کند (عبادی و همکاران، ۱۴۰۳). در ایران برای بدست آوردن آمار مربوط به کشاورزی، اغلب از روش‌های کارشناسی و روش‌های سنتی استفاده می‌شود. نتایج حاصل از پژوهش‌های انجام گرفته نشان می‌دهند که این روش‌ها دارای خطاهای زیادی می‌باشند. درحالی که استفاده از تکنولوژی‌های جدید و تصاویر ماهواره‌ای یک روش برای تخمین سطح زیرکشت اراضی کشاورزی محسوب می‌شود که اخیراً از اهمیت بسیار بالایی برخوردار شده است. داده‌های سنجش از دوری فرصت‌های مناسبی برای تحلیل داده‌های مکانی فراهم می‌کند و کاربردهای مناسبی برای بررسی پوشش اراضی و سطح زیرکشت و تغییرات آن‌ها را دارا می‌باشد (ریاحی و همکاران، ۱۳۹۸). در واقع با استفاده از تصاویر ماهواره‌ای و داده‌های سنجش از دوری می‌توان بدون مراجعه مداوم به زمین‌های کشاورزی با صرف زمان و هزینه‌ی کمتر نسبت به سایر روش‌ها، سطح زیرکشت اراضی را با دقت بهتری برآورد کرد. سامانه‌های سنجش از دوری به دلیل خاصیت متناوب بودن اخذ تصاویر یک ناحیه، تصویربرداری در طول موج‌های مختلف در یک زمان و امکان پردازش و تفسیر سریع این اطلاعات، به‌طور گسترده در تجزیه و تحلیل داده‌های کشاورزی استفاده می‌شوند (موسوی، عباس‌زاده و جانعلی‌پور، ۱۳۹۹). براساس پژوهش‌های انجام گرفته، می‌توان روش‌های نقشه‌برداری محصولات کشاورزی را بر اساس تفاوت در منابع داده به سه دسته تقسیم کرد: روش نقشه‌برداری سنجش از دور اپتیک، روش نقشه‌برداری سنجش از دور مایکروویو و روش تلفیقی. بیشتر این روش‌ها خود به چهار زیر گروه تقسیم می‌شوند: روش‌های مبتنی بر یادگیری ماشین، سری زمانی تصاویر، استفاده از شاخص‌های گیاهی و پردازش‌های شی‌گرایی (ژائو و همکاران^۳، ۲۰۲۱). در این راستا پژوهشگران داخلی مطالعات نسبتاً کمی را انجام داده‌اند ولی پژوهشگران خارجی به کرات این موضوع را مورد بررسی قرار داده‌اند.

یکی از فرضیاتی که این پژوهش قصد تحلیل آن را دارد این است که کدام یک از الگوریتم‌های یادگیری ماشین (ماشین‌بردار پشتیبان و جنگل تصادفی) برای تهیه نقشه کشت گندم در منطقه مورد مطالعه مناسب‌تر است؟ طبق مطالعاتی که در این زمینه انجام شده است انتظار می‌رود الگوریتم ماشین‌بردار پشتیبان در نهایت دقت بهتری نسبت به الگوریتم جنگل تصادفی از خود ارائه بدهد.

از روش‌های نوین و به‌روز مورد توجه بدین منظور، استفاده کردن از تکنیک‌های سنجش‌ازدور می‌باشد که پایش مداوم و تغییر کاربری اراضی در منابع مختلف را میسر کرده است (نوروزی، صانعی و رزقی، ۱۳۹۷). داده‌های سنجش‌ازدور به جهت ارائه اطلاعات به‌روز، پوشش‌های تکراری، سنجش در محدوده‌های طیفی متفاوت، ارزان بودن، تنوع اشکال داده، سنجش کمی و رقومی پدیده‌های سطح زمین، در ارزیابی منابع طبیعی از اهمیت بالایی برخوردار می‌باشد (رضایی مقدم و همکاران، ۱۳۹۲). و همچنین به دلیل دارا بودن خصوصیات همچون چندزمانه بودن، چندطیفی بودن، دارا بودن قدرت تفکیک مکانی و رادیومتریکی مناسب و گوناگون، دید گسترده و یکپارچه، قابلیت تفکیک عوارض و پدیده‌های مختلف کشاورزی (سطح، نوع و الگوهای زمانی و مکانی کشت) را دارد (عبدالله‌زاده و نصیری، ۱۳۸۷). باتوجه به شرایط موجود و بررسی نظر کارشناسان سازمان‌ها و ارگان‌های وابسته، نبود اطلاعات و آمار دقیق و کافی از اراضی سطح زیرکشت گندم و تغییر الگوی کشت به ضرورت این پژوهش می‌افزاید.

¹ Cai and et.al² Luo and et.al³ Zhao and et.al

که شهرستان هشت‌رود از توابع استان آذربایجان شرقی به‌عنوان منطقه مورد مطالعه انتخاب شود تا بتوان با تکیه بر فناوری نوین سنجش‌ازدور مبتنی بر تصاویر ماهواره‌ای در زمان اندک و هزینه کمتر در شناسایی و تعیین سطح زیرکشت اقدام نمود. بدیهی است وجود چنین اطلاعاتی در آگاهی تغییر کاربری مزارع کشت‌گندم، میزان سطح زیرکشت و تخمین محصول، مدیریت صحیح برداشت و ذخیره‌سازی محصولات و ارزیابی خسارت ناشی از تگرگ، رعد و برق، باد و باران می‌تواند اهمیت داشته باشد. لذا بر اهمیت این موضوع می‌افزاید که از داده‌های جدید و در دسترس جهت بروزرسانی و تهیه نقشه‌های کشت‌گندم در سازمان‌ها و ارگان‌های مربوطه استفاده شود. درنهایت، مسئله عدم انجام پژوهش در زمینه تهیه نقشه کشت‌گندم با الگوریتم‌های یادگیری ماشین در کشور ایران و منطقه مورد مطالعه، ما را وادار به انتخاب و انجام پژوهش در این زمینه کرد تا نقشه کشت‌گندم و در منطقه مورد مطالعه با روش‌های مدنظر یادگیری ماشین تهیه کنیم.

پیشینه پژوهش

۱. مطالعات داخلی

صدوقی و همکاران (۱۳۹۹) با استفاده از تصاویر سری‌زمانی سنجنده‌های OLI و ETM+ به شناسایی و بررسی تغییرات سطح زیرکشت محصولات زراعی روستای حسین آباد میش‌مست با استفاده از شاخص پوشش گیاهی NDVI در الگوریتم درخت تصمیم‌گیری پرداختند. نتایج آن‌ها نشان داد در سال ۲۰۰۵ محصولات زراعی شامل گندم، جو و یونجه بوده و در سال ۲۰۰۹ و ۲۰۱۴ ذرت و سیفی‌جات به محصولات زراعی منطقه اضافه شده است. عیبات و همکاران (۱۴۰۰) با استفاده از تصاویر ماهواره‌ای لندست ۸ به برآورد سطح زیرکشت محصولات کشاورزی در شهرستان شوشتر برای سال ۱۳۹۸ پرداختند. با استفاده از شاخص گیاهی NDVI و محصولات زراعی در مراحل مختلف رشد و با توجه به تقویم زراعی آن‌ها، نقشه الگوی کشت محصولات این منطقه استخراج شد. نویدی و همکاران (۱۴۰۱) به ارزیابی تناسب اراضی برای کشت‌گندم در دشت‌های آبی ایران با استفاده از مطالعات خاک، طبقه‌بندی اراضی در سطح کشور و با توجه به تصاویر ماهواره‌ای و نقشه‌های کاربری و در نظر گرفتن پهنه‌های زراعی-اقلیمی در دشت‌های آبی در سراسر کشور به تعیین توان اراضی برای کشت آبی گندم پرداختند. نتایج آن‌ها نشان داد ۱/۴ میلیون هکتار در کلاس مناسب، ۲/۰۸ میلیون هکتار در کلاس نسبتاً مناسب، ۱/۴ میلیون هکتار در کلاس تناسب کم و ۵۲۴ هزار هکتار در کلاس نامناسب در حال حاضر و ۵۰۳ هزار هکتار در کلاس نامناسب دائمی می‌باشند. اصغری سراسکانرود و همکاران (۱۴۰۳) در پژوهشی به تهیه نقشه کشت‌برنج براساس خصوصیات فنولوژیکی با استفاده از سری زمانی تصاویر سنتینل ۱ پرداختند. نتایج نشان داد که استفاده از سری زمانی بازپراکنش داده‌های راداری متناسب با فنولوژی در طبقه‌بندی‌ها باعث افزایش دقت طبقه‌بندی می‌شود. همچنین نتایج نشان داد که استفاده از تصاویر سنتینل ۱ به همراه GEE کارایی بالایی در جهت نظارت بر اراضی شالیزار در مناطق شمالی بدلیل وجود ابر خواهد داشت.

۲. مطالعات خارجی

نصرالله و همکاران^۱ (۲۰۱۸) در مطالعه‌ای تحت عنوان "یک رویکرد نوآورانه برای نقشه‌برداری از مناطق کشت‌گندم با استفاده از تصاویر با رزولوشن بالای Sentinel-2" به تجزیه و تحلیل دقیق نقشه‌برداری مزارع گندم زمستانی در منطقه دشت بقاع در لبنان به وسیله رویکرد نقشه‌برداری گندم ساده و مؤثر (SEWMA)^۲ پرداختند. نتایج این مطالعه نشان داد که مساحت کشت گندم در این منطقه در سال ۲۰۱۷ نسبت به سال ۲۰۱۶ کاهش یافته و این روش SEWMA دقت بالایی در تشخیص مساحت‌های کشت‌گندم در سال‌های مختلف دارد. این روش امکان تطبیق با سیستم‌های حمایت مالی و کمپانسیون مزارع را نیز فراهم می‌کند. این رویکرد SEWMA یک روش ساده و مؤثر برای نقشه‌برداری مزارع گندم زمستانی در لبنان است که به دولت اطلاعات حیاتی برای مدیریت بهتر منابع کشاورزی و حمایت مالی فراهم می‌کند. اسلام^۳ (۲۰۲۱) با استفاده از داده‌های سری

^۱ Nasrallah and et.al

^۲ Statewide Enhanced Wetland Mapping and Assessment

^۳ Islam

زمانی ALOS و MODIS به نقشه‌برداری شالیزارهای برنج بنگلادش پرداخت. نتایج نشان داد که داده‌های ALOS بدلیل وضوح بالا نسبت به MODIS که دارای وضوح متوسط هستند می‌توانند نقشه‌هایی با دقت نسبتاً بالا را تهیه کنند. هودایت و پاتل^۱ (۲۰۲۲) با استفاده از تصاویر سنتینل ۲ به بررسی سطح زیرکشت محصول در ناحیه پوربا مدینیپور بنگال غربی با روش‌های یادگیری ماشین: الگوریتم نزدیک‌ترین همسایه (KNN) و جنگل تصادفی (RF)^۲ پرداختند. نتایج نشان داد که سطح زیرکشت عمده تابستانه (برنج بورو، سبزیجات و تاک فوفل) شناسایی شد که نقشه‌های استخراج شده دارای دقت کلی ۹۷/۱۶ درصد و ۹۷/۲۲ درصد در طبقه‌بندی نزدیک‌ترین همسایه و ۹۵/۹۹ درصد و ۹۶/۰۸ درصد در جنگل تصادفی بودند. فاما و همکاران^۳ (۲۰۲۲) با استفاده از روش‌های یادگیری ماشین: ماشین بردار پشتیبان (SVM)^۴، جنگل تصادفی (RF) و درختان طبقه‌بندی و رگرسیون (CART) به نقشه‌برداری کرت‌های برنج در کشور سنگال پرداختند. نتایج نشان داد که دقت کلی و ضریب کاپا به ترتیب برای SVM برابر با ۹۴/۵ درصد و ۹۶/۲ درصد، برای CART^۵ برابر ۹۶/۵ درصد و ۹۷/۶ درصد و برای RF برابر با ۹۷/۱ درصد و ۹۸ درصد بود. لو و همکاران^۶ (۲۰۲۲) در مطالعه‌ای تحت عنوان "استفاده از الگوریتم‌های یادگیری عمیق برای نقشه‌برداری دقیق سیستم تولید گندم جهانی" از یک چارچوب عمومی به نام "سیستم نقشه‌برداری تولید گندم جهانی" (GWPMs) استفاده کردند. این چارچوب با استفاده از مدل‌های مبتنی بر داده، توزیع مکانی مناطق برداشت گندم و تخمین عملکرد آن را در هشت کشور اصلی تولید کننده گندم در سراسر جهان نقشه‌برداری می‌کند. نتایج این تحقیق نشان داد که GWPMs قادر است نقشه‌های دقیقی از مناطق کشاورزی با توانایی پیش‌بینی عملکرد محصولات با دقت بالا تولید کند. این تحقیق به درک و تحقیقات در زمینه امنیت غذایی جهانی و توسعه سیستم‌های کشاورزی پایدار کمک می‌کند و می‌تواند به بهبود مدیریت منابع زمین و زراعت در سطح جهانی کمک کند. فنگ و همکاران^۷ (۲۰۲۴) در مطالعه‌ای با ترکیب روش‌های مبتنی بر پیکسل و شیء‌گرا به نقشه‌برداری از توزیع مکانی گندم زمستانه در استان شاندونگ پرداختند. نتایج نشان داد که این روش با دقت کلی ۹۲.۲ درصد و ضریب کاپای ۰.۸۴، عملکرد بهتری نسبت به روش‌های سنتی دارد و تأثیر ناهمواری‌های توپوگرافی بر توزیع گندم را به خوبی نشان می‌دهد.

روش پژوهش

۱. محدوده مورد مطالعه

هشترود یکی از شهرستان‌های جنوبی استان آذربایجان شرقی و مرکز شهرستان هشترود است. این شهر در بین مختصات جغرافیایی ۳۶ درجه و ۴۵ دقیقه الی ۳۷ درجه و ۲۴ دقیقه عرض شمالی و ۴۶ درجه و ۲۵ دقیقه الی ۴۷ درجه و ۲۴ دقیقه طول شرقی واقع شده‌است. محدوده حوزه طبیعی شهر هشترود از شمال به شهرستان بستان‌آباد، از شرق به شهرستان میانه، از جنوب به شهرستان چاراویماق و از غرب به شهرستان مراغه منتهی می‌شود. این شهر را به دلیل جریان ۸ رودخانه آبدوغموش، آجی‌چای، قرانقو، قلعه‌چای، قوری‌چای، لیلان‌چای، امیرلو و قره‌قیه در منطقه هشترود نامیدند. هشترود دارای آب‌وهوای کوهستانی با زمستان‌هایی سرد و تابستان‌هایی معتدل است. میانگین سردترین دمای آن ۳- درجه سانتی‌گراد و مربوط به ماه ژانویه (دی و بهمن) و میانگین گرم‌ترین دمای آن ۳۴+ درجه سانتی‌گراد و مربوط به ماه اوت (مرداد و شهریور) است. پیشه اکثر ساکنان هشترود کشاورزی و مهم‌ترین تولیدات آن حبوبات، غلات، میوه‌جات، تره‌بار، زردآلو، گردو، گندم و جو و گیاهان علوفه‌ای می‌باشد. ارتفاع عمومی زمین در این شهرستان از ۱۳۰۰ متر تا ۳۲۰۰ متر متغیر است. قسمت اعظم اراضی منطقه را زمین‌هایی تشکیل می‌دهند که ارتفاع عمومی آن‌ها از ۱۵۰۰ متر الی ۲۰۰۰ متر می‌باشد (شکل ۱).

¹ Hudait and Patel

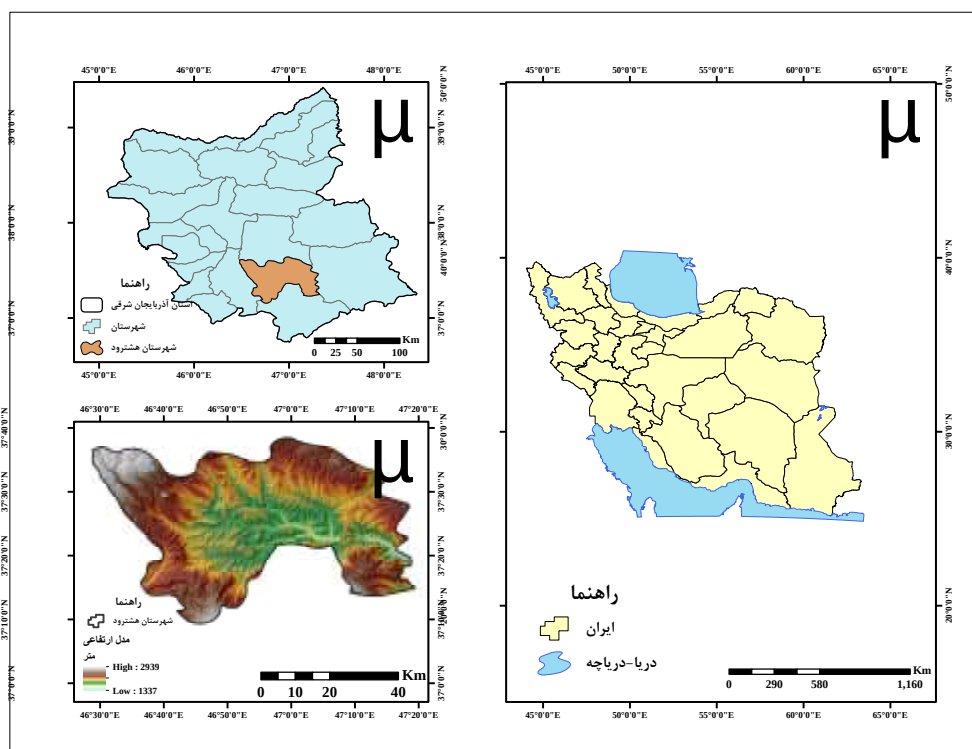
² Random Forest

³ Fama and et.al

⁴ Support Vector Machine

⁵ Classification And Regression Trees

⁶ Luo and et.al



شکل ۱. منطقه مورد مطالعه

۲. مواد و روش

در این پژوهش، ابتدا تصاویر ماهواره Sentinel2-L2A که تاریخ آن‌ها در جدول ۱ ذکر شده است، تهیه شدند.

جدول ۱. تصاویر استفاده‌شده

ردیف	نام سنجنده	تاریخ شمسی	تاریخ میلادی
۱	Sentinel2-L2A	۲۹ فروردین ۱۴۰۲	۱۸ آوریل ۲۰۲۳
۲	Sentinel2-L2A	۱۸ اردیبهشت ۱۴۰۲	۸ می ۲۰۲۳
۳	Sentinel2-L2A	۲۸ اردیبهشت ۱۴۰۲	۱۸ می ۲۰۲۳
۴	Sentinel2-L2A	۲ خرداد ۱۴۰۲	۲۳ می ۲۰۲۳
۵	Sentinel2-L2A	۱۲ خرداد ۱۴۰۲	۲ ژوئن ۲۰۲۳
۶	Sentinel2-L2A	۱ تیر ۱۴۰۲	۲۲ ژوئن ۲۰۲۳
۷	Sentinel2-L2A	۶ تیر ۱۴۰۲	۲۷ ژوئن ۲۰۲۳
۸	Sentinel2-L2A	۱۶ تیر ۱۴۰۲	۷ ژوای ۲۰۲۳
۹	Sentinel2-L2A	۳۱ تیر ۱۴۰۲	۲۲ ژوای ۲۰۲۳

در مرحله بعد، نمونه‌های آموزشی و صحت‌سنجی کلاس گندم به کمک داده‌های GPS گارمین مدل eTrex از نیمه دوم فروردین ۱۴۰۲ تا اواخر تیر ۱۴۰۲ اخذ شد. همینطور تصاویر Google Earth Pro مربوط به سال ۱۴۰۲ نیز جهت ثبت سایر کلاس‌ها به کار گرفته شد. پیش‌پردازش تصاویر در نرم‌افزار Arcmap صورت گرفت و از مجموعه تصاویر دارای اطلاعات بازتاب زمینی، شاخص پوشش گیاهی نرمال شده طبق رابطه ۱ استخراج شد.

$$NDVI = (P_{NIR} - P_{RED}) / (P_{NIR} + P_{RED}) \quad (1)$$

که در آن RED نشانگر انعکاس در باند قرمز و NIR نشانگر انعکاس در باند مادون قرمز نزدیک می‌باشد. به‌طور کلی مقادیر NDVI مابین ۰ و ۱ نشانگر سطوح دارای پوشش گیاهی و مابین -۱ و ۰ نشانگر آب و ابر می‌باشد. در ادامه، برای انجام طبقه‌بندی‌ها براساس الگوریتم‌های ماشین‌بردار پشتیبان و جنگل تصادفی، این مجموعه تصاویر به کتابخانه‌های مختلف یادگیری ماشین زبان برنامه‌نویسی R در محیط Jupyter Notebook وارد شدند و پس از برش منطقه مورد مطالعه از تصاویر دانلود شده، طبقه‌بندی‌های نظارت‌شده SVM و RF بر آن‌ها اعمال شد. در آخر، مراحل صحت‌سنجی طبقه‌بندی‌ها در نرم‌افزار Rstudio انجام گرفت و مقادیر صحت کلی و ضریب کاپا به‌ترتیب طبق روابط ۲ و ۳ محاسبه شدند. بعد از مشاهده و تجزیه و تحلیل و تفسیر نتایج، مناسب‌ترین نقشه از هر الگوریتم انتخاب شده و با یکدیگر مقایسه شدند. در نهایت مراحل ایجاد خروجی نقشه‌ها در نرم‌افزار Arcmap انجام شد و مساحت مزارع گندم نیز به‌دست آمد.

$$OA = \frac{\sum_{i=1}^k E_{ii}}{N} \quad (2)$$

که در آن k نشانگر تعداد کلاس‌ها، N تعداد کل پیکسل‌های معلوم و E_{ii} اعضای قطری ماتریس خطا می‌باشد. این المان تفاوت‌هایی که بین کلاس‌ها وجود دارد به‌علت برخورد یکسان با همه کلاس‌ها، در نظر نمی‌گیرد (اصغری سراسکانرود و همکاران، ۱۳۹۹).

$$\hat{K} = \frac{N \sum_{i=1}^k x_{ii} - \sum_{i=1}^k (x_{i+} \times x_{+j})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+j})} \quad (3)$$

که در آن، N تعداد پیکسل‌های واقعیت زمینی، x_{i+} مجموع عناصر سطر i ام و x_{+j} مجموع عناصر ستون j ام می‌باشد. پس از بررسی روش‌ها و تعیین اهداف پژوهش و آشنایی با منطقه مورد مطالعه، تعداد ۵ کلاس آب، مناطق شهری، کشت آبی و باغات، مناطق بایر و گندم در منطقه مورد مطالعه تعیین و نمونه‌های آموزشی آن‌ها با رویکرد تصادفی برداشت شد. منظور آب در این تحقیق پهنه‌های آبی است که می‌تواند شامل آب سدها، تالاب‌ها، آب رودخانه و ... باشد که در مقالات خارجی به عنوان Water Bodies شناخته می‌شود. باید توجه داشت با توجه به مقیاس تصاویر مورد استفاده و زمان اخذ تصاویر، عمدتاً آب رودخانه‌ها قابل تشخیص نبوده است اما آب رودخانه قرنقو با دقت قابل مشاهده است.

۳. ماهواره سنتینل ۲^۱

ماهواره سنتینل-۲ بخشی از برنامه کوپرنیک اتحادیه اروپا است که برای نظارت بر سطح زمین طراحی شده است. این ماهواره‌ها به صورت دو جفت (Sentinel-2A و Sentinel-2B) کار می‌کنند و تصاویر با رزولوشن بالا را از سطح زمین تهیه می‌کنند. سنتینل-۲ دارای ۱۳ باند طیفی در محدوده‌های مرئی، مادون قرمز نزدیک و مادون قرمز کوتاه‌موج است. رزولوشن مکانی تصاویر بین ۱۰، ۲۰ و ۶۰ متر بسته به باند مورد استفاده متغیر است (جدول ۲). تکرار زمانی تصاویر هر ۵ روز یک بار است که این ویژگی آن را برای کاربردهای محیط‌زیستی، کشاورزی و مدیریت بحران بسیار مفید می‌کند.

¹ Sentinel-2

جدول ۲. مشخصات باندهای ماهواره ستینل ۲

باند	عنوان	طول موج (نانومتر)	وضوح مکانی (متر)
B1	آتروسل ساحلی	۴۴۳	۶۰
B2	آبی	۴۹۰	۱۰
B3	سبز	۵۶۰	۱۰
B4	قرمز	۶۶۵	۱۰
B5	لبه قرمز	۷۰۵	۲۰
B6	لبه قرمز	۷۴۰	۲۰
B7	لبه قرمز	۷۸۳	۲۰
B8	مادون قرمز نزدیک	۸۴۲	۱۰
B8a	لبه قرمز	۸۶۵	۲۰
B9	بخار آب	۹۴۰	۶۰
B10	مادون قرمز طول موج کوتاه- سیروس	۱۳۷۵	۶۰
B11	مادون قرمز طول موج کوتاه	۱۶۱۰	۲۰
B12	مادون قرمز طول موج کوتاه	۲۱۹۰	۲۰

۳. الگوریتم ماشین بردار پشتیبان

اخیرا توجه ویژه‌ای به ماشین‌های بردار پشتیبان به عنوان یک روش طبقه‌بندی شده‌است. ماشین بردار پشتیبان با تمرکز بر موارد آموزشی که در لبه توصیف‌گرهای کلاس قرار می‌گیرند، به دنبال یافتن ابرصفحه جداکننده بهینه بین کلاس‌ها می‌باشد. به این موارد آموزشی بردارهای پشتیبان گفته می‌شود. از موارد آموزشی غیر از بردارهای پشتیبان صرف نظر می‌شود. به این ترتیب نه تنها یک ابرصفحه بهینه جاگذاری می‌شود بلکه از نمونه‌های آموزشی کم‌تری نیز به‌طور موثر استفاده می‌شود. بنابراین یک طبقه‌بندی با دقت بالا و مجموعه آموزشی کوچک به دست می‌آید (مرسیر و لنون^۱، ۲۰۰۳). این ویژگی بسیار مفید است، به ویژه برای مجموعه داده‌های سنجش از دوری و به‌طور خاص برای تجزیه و تحلیل تصویر شیء‌گرا که در آن نمونه‌های شیء‌گرا کم‌تر از نمونه‌های پیکسل مبنا هستند.

۴. الگوریتم جنگل تصادفی

این الگوریتم‌ها سریع و آسان برای پیاده‌سازی هستند. پیش‌بینی‌های بسیار دقیقی ارائه می‌دهند و می‌توانند با تعداد بسیار زیادی متغیر ورودی بدون بیش‌برازش (Overfitting) کار کنند. در واقع، آن‌ها به عنوان یکی از دقیق‌ترین روش‌های یادگیری عمومی موجود محسوب می‌شوند. در رویکرد برایمن، هر درخت در مجموعه ابتدا با انتخاب تصادفی یک گروه کوچک از مختصات ورودی (همچنین به نام ویژگی‌ها یا متغیرها) در هر گره شکسته می‌شود و سپس با محاسبه بهترین شکست بر اساس این ویژگی‌ها در مجموعه آموزشی، درخت شکل می‌گیرد. از روش CART (برایمن و همکاران^۲، ۱۹۸۴) برای رشد درخت به حداکثر اندازه، بدون قلمه‌زدن، استفاده می‌شود. این روش تصادفی‌سازی زیرفضایی با متد بگینگ (برایمن، ۱۹۹۶؛ بوهمان و یو^۳، ۲۰۰۲؛ بوجا و اشتوتسله^۴، ۲۰۰۶؛ بیو، چرو و گویادر^۵، ۲۰۱۰) ترکیب می‌شود تا هر بار که یک درخت جدید شکل می‌گیرد، مجموعه داده آموزشی با جایگزینی دوباره نمونه‌برداری شود.

¹ Mercier and Lennon

² Breiman

³ Buhlmann and Yu

⁴ Buja and Stuetzle

⁵ Biau, Cérou and Guyader

نتایج

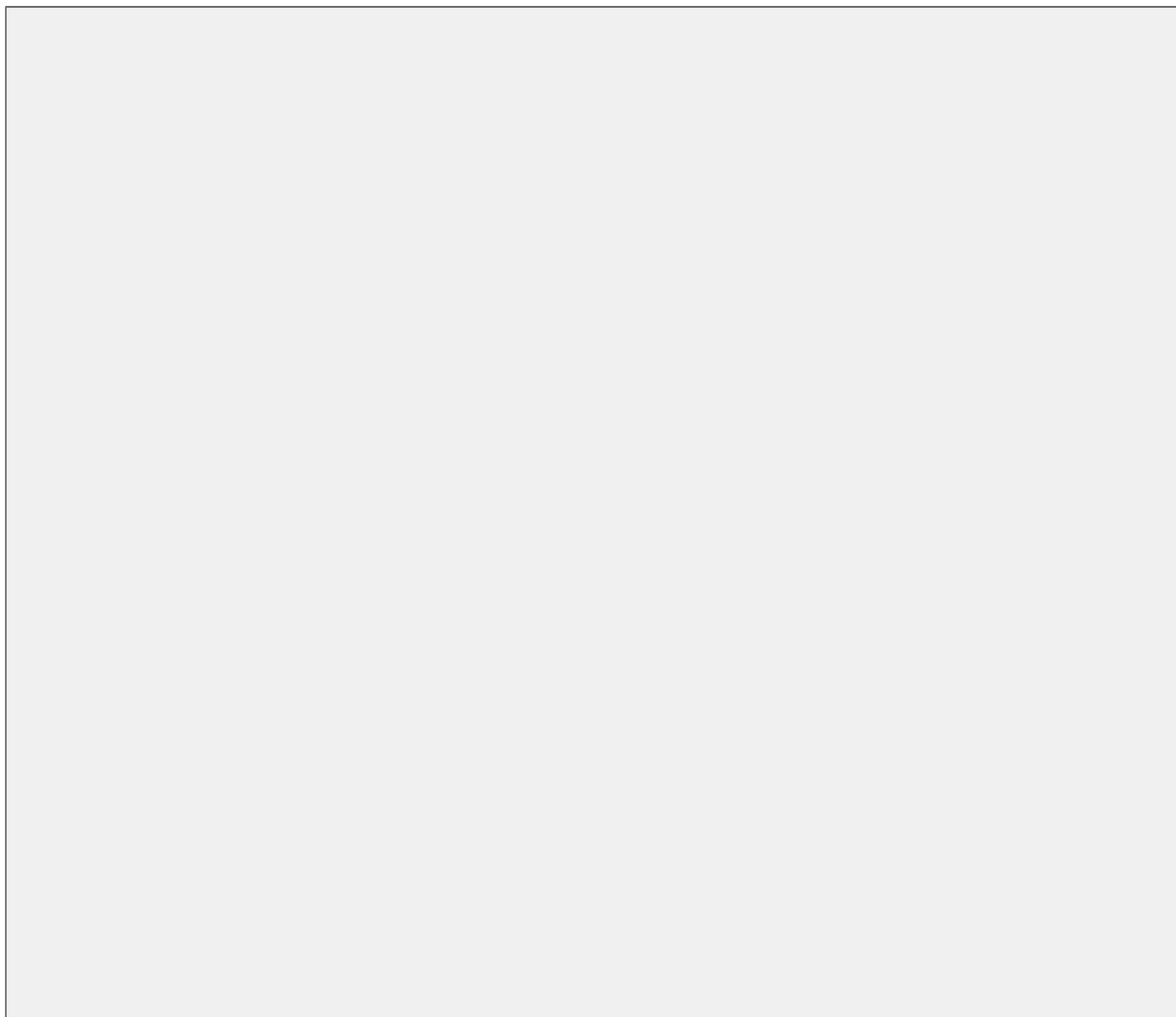
بررسی و تجزیه و تحلیل نتایج و ارزیابی دقت از ملزومات هر کار پژوهشی می‌باشد. در این بخش داده‌های مورد استفاده و نتایج بدست آمده مورد تجزیه و تحلیل قرار گرفته است و در آخر با توجه به دقت‌های نقشه‌های بدست‌آمده بهترین الگوریتم تعیین شده است.

۱. بررسی نمودار زمانی تغییرات شاخص NDVI

NDVI میزان و وضعیت پوشش گیاهی را نشان می‌دهد. استفاده از این شاخص دلیل حساسیت آن به تراکم پوشش گیاهی می‌باشد. از این شاخص می‌توان در تخمین بازتابش شاخص سطح برگ، سطح زیرکشت پوشش سطح زمین و شدت رشد توده زنده گیاهی استفاده کرد. نکته‌ای که در تصاویر رادار باید به آن توجه کرد این است که حتی در یک محصول معین در کشاورزی، خواص دی‌الکتریک، زبری سطح و ساختار تاج معمولاً یکسان نیستند و می‌توانند توزیع مجدد را در کلاس تغییر دهند. براساس جدول ۲ نقشه‌های حاصل طبقه‌بندی شده است. باتوجه به شکل ۲ و جدول ۳، بیشترین مقدار بیشینه NDVI مربوط به تصویر با تاریخ ۳۱ تیر ماه ۱۴۰۲ با مقدار ۰/۹۷۲۴ که نمایانگر اوج سبزیگی می‌باشد و کمترین مقدار بیشینه NDVI مربوط به تصویر با تاریخ ۱۶ تیر ماه ۱۴۰۲ با مقدار ۰/۸۴۵۶ می‌باشد. همچنین بیشترین مقدار کمینه NDVI مربوط به تصویر با تاریخ ۱۶ تیر ماه ۱۴۰۲ با مقدار ۰/۲۹۶۷- و کمترین مقدار کمینه NDVI مربوط به تصویر با تاریخ ۱ تیر ماه ۱۴۰۲ با مقدار ۰/۶۹۹۴- می‌باشد. نیز بیشترین و کمترین مقدار میانگین NDVI به ترتیب مربوط به تصاویر با تاریخ‌های ۲ خردادماه و ۱ تیرماه ۱۴۰۲ با مقادیر ۰/۲۷۸۱ و ۰/۱۲۰۱ می‌باشد.

جدول ۲. طبقه‌بندی شاخص NDVI

بازه	طبقه‌بندی شاخص NDVI
۰ تا -۰/۳	خاک مرطوب
-۰/۳ تا -۱	آب
۰ تا ۰/۱	بدون پوشش گیاهی
۰/۱ تا ۰/۳	پوشش گیاهی ناچیز با زمینه خاک
۰/۳ تا ۰/۵	پوشش گیاهی با تراکم متوسط
۰/۵ تا ۰/۷	پوشش گیاهی با تراکم بالا
۰/۷ تا ۱	پوشش گیاهی با تراکم خیلی زیاد



شکل ۲. نقشه‌های حاصل از اعمال شاخص پوشش گیاهی بر روی تصاویر
جدول ۳. مقدار NDVI مربوط به تصاویر

تاریخ تصاویر	بیشترین مقدار	کمترین مقدار	میانگین
۲۹ فروردین	۰/۸۸۲۸	-۰/۵۷۴۲	۰/۱۵۴۳
۱۸ اردیبهشت	۰/۹۴۰۴	-۰/۵۵۹۳	۰/۱۹۰۵
۲۸ اردیبهشت	۰/۹۲	-۰/۵	۰/۲۱
۲ خرداد	۰/۹۳۰۹	-۰/۳۷۴۶	۰/۳۷۸۱
۱ تیر	۰/۹۳۹۶	-۰/۶۹۹۴	۰/۱۲۰۱
۶ تیر	۰/۹۳۴۶	-۰/۶۴۷۰	۰/۱۳۸۸
۱۶ تیر	۰/۸۴۵۶	-۰/۲۹۶۷	۰/۲۷۴۴
۳۱ تیر	۰/۹۷۲۴	-۰/۴۶۵۵	۰/۲۵۳۷

۲. بررسی پارامترهای الگوریتم ماشین‌بردار پشتیبان

الگوریتم ماشین‌بردار پشتیبان یکی از قدرتمندترین روش‌های یادگیری ماشین برای طبقه‌بندی و رگرسیون است که در این پروژه برای طبقه‌بندی تصاویر ماهواره‌ای سنتینل-۲ استفاده شده است. دو پارامتر کلیدی این الگوریتم شامل گاما^۱ (γ) و هزینه^۲ (C) هستند. پارامتر گاما میزان تأثیر یک نمونه آموزشی بر مرز تصمیم‌گیری را تعیین می‌کند؛ مقادیر پایین گاما به معنای تأثیر زیاد هر نمونه و مقادیر بالا به معنای تأثیر کم هر نمونه است. پارامتر هزینه نیز میزان جریمه برای خطاهای طبقه‌بندی را مشخص می‌کند؛ مقدار بالای C سعی در کاهش خطاها دارد، اما ممکن است منجر به بیش‌برازش شود. در این کد، از روش جستجوی شبکه‌ای^۳ برای پیدا کردن بهترین مقادیر گاما و هزینه استفاده شده است. برای این منظور، مقادیر گاما در بازه 10^{-6} تا 10^{-1} و مقادیر هزینه در بازه 10^{-1} تا 10^1 مورد بررسی قرار گرفتند. ابتدا با استفاده از تابع `tune.svm` بهترین مقادیر برای این پارامترها انتخاب شدند. این تابع مدل‌های مختلف SVM را با ترکیب‌های مختلف از مقادیر گاما و هزینه آموزش داده و عملکرد هر مدل را ارزیابی می‌کند تا بهینه‌ترین ترکیب را پیدا کند. بهترین مقادیر یافت شده برای پارامترهای گاما و هزینه به ترتیب برابر با 10^{-3} و 10^0 بودند. سپس، مدل SVM با استفاده از این پارامترهای بهینه آموزش داده شد. این تنظیمات به مدل کمک کردند تا با دقت بالا و بدون بیش‌برازش، تصاویر ماهواره‌ای را طبقه‌بندی کند. نتایج طبقه‌بندی نهایی با استفاده از این مدل بهبود یافته و دقت و صحت بالاتری را نسبت به مقادیر پیش‌فرض نشان دادند. این فرآیند بهینه‌سازی پارامترها به طور قابل توجهی کارایی و عملکرد مدل SVM در طبقه‌بندی تصاویر ماهواره‌ای را افزایش داد.

۳. بررسی پارامترهای الگوریتم جنگل تصادفی

این الگوریتم از ترکیب چندین درخت تصمیم‌گیری برای بهبود دقت و کاهش بیش‌برازش استفاده می‌کند. در اینجا از ۱۰۰۰ درخت تصمیم‌گیری (پارامتر `ntree`) برای آموزش مدل استفاده شده است تا دقت طبقه‌بندی افزایش یابد و اثر نویز کاهش یابد. همچنین، پارامتر اهمیت ویژگی‌ها (`importance = TRUE`) فعال شده است تا مشخص شود کدام ویژگی‌ها بیشترین تأثیر را در طبقه‌بندی دارند. داده‌های آموزشی با استفاده از تابع `raster::extract` استخراج و سپس با استفاده از تابع `randomForest` مدل جنگل تصادفی آموزش داده شده است. پس از آموزش، مدل با استفاده از تابع `predict` برای طبقه‌بندی تصاویر ماهواره‌ای به کار رفته است. نتایج طبقه‌بندی به صورت فاکتور^۴ ذخیره شده و کلاس‌های مختلف نام‌گذاری شده‌اند تا تفسیر و استفاده از نتایج آسان‌تر شود. این تنظیمات به مدل کمک کرده‌اند تا با دقت بالا و قابلیت اطمینان مناسب، تصاویر ماهواره‌ای را طبقه‌بندی کند و دقت کلی ۹۳.۲۱٪ را در این پروژه به دست آورد.

۴. بررسی نتایج الگوریتم ماشین‌بردار پشتیبان

باتوجه به نقشه حاصل از الگوریتم ماشین‌بردار پشتیبان (شکل ۳) و ماتریس خطای مربوطه در دسته‌بندی "آب"، از مجموع نمونه‌ها، ۲۸۰۷۴ نمونه به درستی تشخیص داده شده‌اند، در حالی که ۲۱۳ نمونه به اشتباه به "کشت آبی و باغات"، ۷۷۵ نمونه به "مناطق بایر" و ۵۵ نمونه به "گندم" تخصیص داده شده‌اند. در دسته‌بندی "مناطق شهری"، ۳۶۲۹ نمونه به درستی تشخیص داده شده و ۹۰۵ نمونه به اشتباه به "کشت آبی و باغات"، ۵۰۶۹ نمونه به "مناطق بایر" و ۱۱ نمونه به "گندم" تخصیص داده شده‌اند. برای "کشت آبی و باغات"، ۱۴۱۹۰ نمونه به درستی تشخیص داده شده و ۱۳۱ نمونه به اشتباه به "آب"، ۱۸ نمونه به "مناطق بایر" و ۴۰۷ نمونه به "گندم" تخصیص داده شده‌اند. در دسته‌بندی "مناطق بایر"، ۱۳۹۷۸۴ نمونه به درستی تشخیص داده شده و ۹۷ نمونه به اشتباه به "آب"، ۶۵۶ نمونه به "مناطق شهری"، ۱۴۲ نمونه به "کشت آبی و باغات" و ۵۹۱۶ نمونه به "گندم" تخصیص داده شده‌اند. نهایتاً در دسته‌بندی "گندم"، ۱۷۹۷۷ نمونه به درستی تشخیص داده شده و

^۱ Gamma

^۲ Cost

^۳ grid search

^۴ factor

۱۳۹ نمونه به اشتباه به "آب"، ۴ نمونه به "مناطق شهری"، ۱۵۶۹ نمونه به "کشت آبی و باغات" و ۴۲۷۱ نمونه به "مناطق بایر" تخصیص داده شده‌اند (جدول ۴). همچنین این الگوریتم برای دسته‌بندی داده‌ها با استفاده از معیارهای مختلف ارزیابی شد. حساسیت برای دسته‌های "آب"، "مناطق شهری"، "کشت آبی و باغات"، "مناطق بایر" و "گندم" به ترتیب ۰/۹۸۷۱، ۰/۸۴۶۱۲، ۰/۸۳۳۷۷، ۰/۹۲۷۶ و ۰/۷۳۷۷۹ بود که نشان‌دهنده نسبت درست شناسایی نمونه‌های مثبت واقعی است. تخصصی بودن الگوریتم در شناسایی این دسته‌ها به ترتیب ۰/۹۹۴۴، ۰/۹۷۱۴۷، ۰/۹۹۷۱۸، ۰/۹۰۸۱ و ۰/۹۶۸۴۶ بود که میزان درستی تشخیص نمونه‌های منفی را نشان داد. ارزش پیش‌بینی مثبت که نشان‌دهنده نسبت درست نمونه‌های مثبت پیش‌بینی شده به کل نمونه‌های مثبت پیش‌بینی شده است، برای این دسته‌ها به ترتیب ۰/۹۶۴۲، ۰/۳۷۷۴۷، ۰/۹۶۲۲۹، ۰/۹۵۰۱ و ۰/۷۵۰۲۹ بود. همچنین، ارزش پیش‌بینی منفی برای این دسته‌ها به ترتیب ۰/۹۹۸۰، ۰/۹۹۶۷۷، ۰/۹۸۵۸۰، ۰/۸۶۹۱ و ۰/۹۶۶۳۹ بود که نشان‌دهنده نسبت درست نمونه‌های منفی پیش‌بینی شده به کل نمونه‌های منفی پیش‌بینی شده است. شیوع دسته‌ها نیز به ترتیب ۰/۱۳۲۹، ۰/۰۲۰۰۴، ۰/۰۷۹۵۲، ۰/۶۵۳۷ و ۰/۱۱۳۸۴ بود. میزان شناسایی که نشان‌دهنده نسبت نمونه‌های درست شناسایی شده به کل نمونه‌ها است، به ترتیب ۰/۱۳۱۲، ۰/۰۱۶۹۶، ۰/۰۶۶۳۰، ۰/۶۰۶۴ و ۰/۰۸۳۹۹ بود. شیوع شناسایی نیز به ترتیب ۰/۱۳۶۰، ۰/۰۴۴۹۲، ۰/۰۶۸۹۰، ۰/۶۳۸۲ و ۰/۱۱۱۹۵ بود. در نهایت، دقت متعادل که ترکیبی از حساسیت و تخصصی بودن است، برای این دسته‌ها به ترتیب ۰/۹۹۰۷، ۰/۰۸۷۹، ۰/۹۱۵۴۸، ۰/۹۱۷۸ و ۰/۸۵۳۱۲ بود (جدول ۵). این نتایج نشان‌دهنده عملکرد الگوریتم در دسته‌بندی‌های مختلف است. صحت کلی و ضریب کاپای اختصاصی نیز ۹۰ درصد و ۸۲ درصد بود (جدول ۶).



شکل ۳. نقشه حاصل از الگوریتم ماشین بردار پشتیبان

جدول ۴. ماتریس خطای مربوط به الگوریتم ماشین‌بردار پشتیبان

پیش‌بینی (Prediction)	مرجع (Reference)					
	گندم	مناطق بایر	کشت آبی و باغات	مناطق شهری	آب	
	۵۵	۷۷۵	۲۱۳	۰	۲۸۰۷۴	آب
	۱۱	۵۰۶۹	۹۰۵	۳۶۲۹	۰	مناطق شهری
	۴۰۷	۱۸	۱۴۱۹۰	۰	۱۳۱	کشت آبی و باغات
	۵۹۱۶	۱۲۹۷۸۴	۱۴۲	۶۵۶	۹۷	مناطق بایر
	۱۷۹۷۷	۴۲۷۱	۱۵۶۹	۴	۱۳۹	گندم

جدول ۵. آمار براساس کلاس‌های طبقه‌بندی‌شده الگوریتم ماشین‌بردار پشتیبان

گندم	مناطق بایر	کشت آبی و باغات	مناطق شهری	آب	
۰/۷۳۷۷۹	۰/۹۲۷۶	۰/۸۳۳۷۷	۰/۸۴۶۱۲	۰/۹۸۷۱	حساسیت
۰/۹۶۸۴۶	۰/۹۰۸۱	۰/۹۹۷۱۸	۰/۹۷۱۴۷	۰/۹۹۴۴	تخصصی بودن
۰/۷۵۰۳۹	۰/۹۵۰۱	۰/۹۶۲۲۹	۰/۳۷۷۴۷	۰/۹۶۴۲	ارزش پیش‌بینی مثبت
۰/۹۶۶۳۹	۰/۸۶۹۱	۰/۹۸۵۸۰	۰/۹۹۶۷۷	۰/۹۹۸۰	ارزش پیش‌بینی منفی
۰/۱۱۳۸۴	۰/۶۵۳۷	۰/۰۷۹۵۲	۰/۰۲۰۰۴	۰/۱۳۲۹	شیوع
۰/۰۸۳۹۹	۰/۶۰۶۴	۰/۰۶۶۳۰	۰/۰۱۶۹۶	۰/۱۳۱۲	میزان شناسایی
۰/۱۱۱۹۵	۰/۶۳۸۲	۰/۰۶۸۹۰	۰/۰۴۴۹۲	۰/۱۳۶۰	شیوع شناسایی
۰/۸۵۳۱۲	۰/۹۱۷۸	۰/۹۱۵۴۸	۰/۹۰۸۷۹	۰/۹۹۰۷	دقت متعادل

جدول ۶. صحت کلی و ضریب کاپای اختصاصی الگوریتم ماشین‌بردار پشتیبان

۰/۹۰۴۸	صحت
۰/۸۲۵۵	ضریب کاپا

همچنین مساحت و تعداد پیکسل‌های مربوط به هر کدام از کلاس‌های طبقه‌بندی‌شده (آب، مناطق شهری، کشت آبی و باغات، مناطق بایر و گندم) در روش ماشین‌بردار پشتیبان طبق جدول ۷ می‌باشد.

جدول ۷. مشخصات کلاس‌ها در الگوریتم ماشین‌بردار پشتیبان

کلاس‌ها	تعداد پیکسل	مساحت (کیلومتر)
آب	۴۰۰۸۲۴	۴۰/۰۸۲۴
مناطق شهری	۱۱۰۹۵۷۶	۱۱۰/۹۵۷۶
کشت آبی و باغات	۹۶۷۱۷۲	۹۶/۷۱۷۲
مناطق بایر	۱۴۹۱۰۲۷۲	۱۴۹۱/۰۲۷۲
گندم	۵۴۶۶۳۵۲	۵۴۶/۶۳۵۲
مجموع	۲۲۸۵۴۱۹۶	۲۲۸۵/۴۱۹۶

۵. بررسی نتایج الگوریتم جنگل تصادفی

باتوجه به نقشه حاصل از الگوریتم جنگل تصادفی (شکل ۴) و ماتریس خطای مربوطه در دسته‌بندی "آب"، از مجموع نمونه‌ها، ۲۸۳۰۳ نمونه به درستی تشخیص داده شده‌اند، در حالی که ۱۲۵ نمونه به اشتباه به "کشت آبی و باغات"، ۱۳۶ نمونه به "مناطق بایر" و ۶ نمونه به "گندم" تخصیص داده شده‌اند. در دسته‌بندی "مناطق شهری"، ۳۸۱۶ نمونه به درستی تشخیص داده شده و ۴۶۲ نمونه به "مناطق بایر" تخصیص داده شده‌اند. برای "کشت آبی و باغات"، ۱۶۱۲۲ نمونه به درستی تشخیص

داده شده و ۳ نمونه به اشتباه به "آب"، ۴ نمونه به "مناطق بایر" و ۴۹۳ نمونه به "گندم" تخصیص داده شده‌اند. در دسته‌بندی "مناطق بایر"، ۱۳۴۹۴۷ نمونه به درستی تشخیص داده شده و ۴ نمونه به اشتباه به "آب"، ۴۶۹ نمونه به "مناطق شهری"، ۱۵۹ نمونه به "کشت آبی و باغات" و ۷۵۵۰ نمونه به "گندم" تخصیص داده شده‌اند. نهایتاً در دسته‌بندی "گندم"، ۱۶۳۱۷ نمونه به درستی تشخیص داده شده و ۱۳۱ نمونه به اشتباه به "آب"، ۴ نمونه به "مناطق شهری"، ۶۱۳ نمونه به "کشت آبی و باغات" و ۴۳۶۸ نمونه به "مناطق بایر" تخصیص داده شده‌اند (جدول ۸). همچنین این الگوریتم برای دسته‌بندی داده‌ها با استفاده از معیارهای مختلف ارزیابی شد. حساسیت برای دسته‌های "آب"، "مناطق شهری"، "کشت آبی و باغات"، "مناطق بایر" و "گندم" به ترتیب ۰/۹۹۵۱، ۰/۸۸۹۷۲، ۰/۹۴۷۲۹، ۰/۹۶۴۵ و ۰/۶۶۹۶۶ بود که نشان‌دهنده نسبت درست شناسایی نمونه‌های مثبت واقعی است. تخصصی بودن الگوریتم در شناسایی این دسته‌ها به ترتیب ۰/۹۹۸۶، ۰/۹۹۷۸۰، ۰/۹۹۷۴۶، ۰/۸۸۹۶ و ۰/۹۷۳۰۳ بود که میزان درستی تشخیص نمونه‌های منفی را نشان داد. ارزش پیش‌بینی مثبت که نشان‌دهنده نسبت درست نمونه‌های مثبت پیش‌بینی شده به کل نمونه‌های مثبت پیش‌بینی شده است، برای این دسته‌ها به ترتیب ۰/۹۹۰۷، ۰/۸۹۲۰۱، ۰/۹۶۹۹۲، ۰/۹۴۲۸ و ۰/۷۶۱۳۰ بود. همچنین، ارزش پیش‌بینی منفی برای این دسته‌ها به ترتیب ۰/۹۹۹۳، ۰/۹۹۷۷۴، ۰/۹۹۵۴۶، ۰/۹۲۹۹ و ۰/۹۵۸۲۱ بود که نشان‌دهنده نسبت درست نمونه‌های منفی پیش‌بینی شده به کل نمونه‌های منفی پیش‌بینی شده است. شیوع دسته‌ها نیز به ترتیب ۰/۱۳۳۹، ۰/۰۲۰۰۴، ۰/۰۷۹۵۲، ۰/۶۵۳۷ و ۰/۱۱۳۸۴ بود. میزان شناسایی که نشان‌دهنده نسبت نمونه‌های درست شناسایی شده به کل نمونه‌ها است، به ترتیب ۰/۱۳۲۲، ۰/۰۱۷۸۳، ۰/۰۷۵۳۳، ۰/۰۶۳۰۵ و ۰/۰۷۶۲۴ بود. شیوع شناسایی نیز به ترتیب ۰/۱۳۳۵، ۰/۰۱۹۹۹، ۰/۰۷۷۶۶، ۰/۰۶۶۸۷ و ۰/۱۰۰۱۴ بود. در نهایت، دقت متعادل که ترکیبی از حساسیت و تخصصی بودن است، برای این دسته‌ها به ترتیب ۰/۹۹۶۹، ۰/۹۴۳۷۶، ۰/۹۷۲۳۸، ۰/۹۲۷۰ و ۰/۸۲۱۳۴ بود (جدول ۹). این نتایج نیز نشان‌دهنده عملکرد الگوریتم در دسته‌بندی‌های مختلف است. صحت کلی و ضریب کاپای اختصاصی نیز ۹۳ درصد و ۸۷ درصد بود (جدول ۱۰).

37°0'N



شکل ۴. نقشه حاصل از الگوریتم جنگل تصادفی

جدول ۸. ماتریس خطای مربوط به الگوریتم جنگل تصادفی

پیش‌بینی (Prediction)	مرجع (Reference)					
	گندم	مناطق بایر	کشت آبی و باغات	مناطق شهری	آب	
	۶	۱۳۶	۱۲۵	۰	۲۸۳۰۳	آب
	۰	۴۶۲	۰	۳۸۱۶	۰	مناطق شهری
	۴۹۳	۴	۱۶۱۲۲	۰	۳	کشت آبی و باغات
	۷۵۵۰	۱۳۴۹۴۷	۱۵۹	۴۶۹	۴	مناطق بایر
	۱۶۳۱۷	۴۳۶۸	۶۱۳	۴	۱۳۱	گندم

جدول ۹. آمار براساس کلاس‌های طبقه‌بندی شده الگوریتم جنگل تصادفی

گندم	مناطق بایر	کشت آبی و باغات	مناطق شهری	آب	
۰/۶۶۹۶۶	۰/۹۶۴۵	۰/۹۴۷۲۹	۰/۸۸۹۷۲	۰/۹۹۵۱	حساسیت
۰/۹۷۳۰۳	۰/۸۸۹۶	۰/۹۹۷۴۶	۰/۹۹۷۸۰	۰/۹۹۸۶	تخصصی بودن
۰/۷۶۱۳۰	۰/۹۴۲۸	۰/۹۶۹۹۲	۰/۸۹۲۰۱	۰/۹۹۰۷	ارزش پیش‌بینی مثبت
۰/۹۵۸۲۱	۰/۹۲۹۹	۰/۹۹۵۴۶	۰/۹۹۷۷۴	۰/۹۹۹۳	ارزش پیش‌بینی منفی
۰/۱۱۳۸۴	۰/۶۵۳۷	۰/۰۷۹۵۲	۰/۰۲۰۰۴	۰/۱۳۲۹	شیوع
۰/۰۷۶۲۴	۰/۶۳۰۵	۰/۰۷۵۳۳	۰/۰۱۷۸۳	۰/۱۳۲۲	میزان شناسایی
۰/۱۰۰۱۴	۰/۶۶۸۷	۰/۰۷۷۶۶	۰/۰۱۹۹۹	۰/۱۳۳۵	شیوع شناسایی
۰/۸۲۱۳۴	۰/۹۲۷۰	۰/۹۷۲۳۸	۰/۹۴۳۷۶	۰/۹۹۶۹	دقت متعادل

جدول ۱۰. صحت کلی و ضریب کاپای اختصاصی الگوریتم جنگل تصادفی

صحت	۰/۹۳۲۱
ضریب کاپا	۰/۸۷۱۲

همچنین مساحت و تعداد پیکسل‌های مربوط به هر کدام از کلاس‌های طبقه‌بندی شده (آب، مناطق شهری، کشت آبی و باغات، مناطق بایر و گندم) در روش جنگل تصادفی طبق جدول ۱۱ می‌باشد.

جدول ۱۱. مشخصات کلاس‌ها در الگوریتم جنگل تصادفی

کلاس‌ها	تعداد پیکسل	مساحت (کیلومتر)
آب	۲۳۸۰۸۱	۲۳/۸۰۸۱
مناطق شهری	۳۶۳۰۶۶	۳۶/۳۰۶۶
کشت آبی و باغات	۱۰۶۱۵۲۸	۱۰۶/۱۵۲۸
مناطق بایر	۱۵۶۹۱۳۸۹	۱۵۶۹/۱۳۸۹
گندم	۵۵۰۰۱۳۲	۵۵۰/۰۱۳۲
مجموع	۲۲۸۵۴۱۹۶	۲۲۸۵/۴۱۹۶

۶. مقایسه نتایج طبقه‌بندی و تعیین بهترین الگوریتم

در این پژوهش، دو الگوریتم ماشین‌بردار پشتیبان و جنگل تصادفی برای طبقه‌بندی تصاویر ماهواره‌ای سنتینل-۲ مورد استفاده قرار گرفتند و نتایج آن‌ها ارزیابی شد. هر دو الگوریتم با تنظیم پارامترهای بهینه آموزش داده شدند و عملکرد آن‌ها با استفاده از معیارهای دقت کلی و ضریب کاپا مقایسه گردید. ماشین‌بردار پشتیبان با پارامترهای بهینه گاما (γ) و هزینه (λ) به دقت کلی ۹۰٪ و ضریب کاپای ۸۲٪ دست یافت. این مدل با استفاده از فرآیند جستجوی شبکه‌ای بهینه‌سازی شد و توانست عملکرد مناسبی در طبقه‌بندی تصاویر ماهواره‌ای ارائه دهد. در مقابل، الگوریتم جنگل تصادفی با استفاده از ۱۰۰۰ درخت تصمیم‌گیری و فعال‌سازی اهمیت ویژگی‌ها ($\text{importance} = \text{TRUE}$) عملکرد بهتری نشان داد و دقت کلی ۹۳.۲۱٪ و ضریب کاپای ۸۷.۱۲٪ را به دست آورد. این مدل به دلیل استفاده از تعداد زیادی درخت و ترکیب نتایج آن‌ها، توانست دقت و صحت بالاتری در پیش‌بینی کلاس‌ها ارائه دهد. مقایسه نتایج ماتریس خطا نیز نشان داد که الگوریتم جنگل تصادفی در اکثر کلاس‌ها عملکرد بهتری نسبت به ماشین‌بردار پشتیبان داشت. به‌ویژه در کلاس‌هایی که تنوع بیشتری داشتند، جنگل تصادفی توانست با دقت بیشتری این کلاس‌ها را طبقه‌بندی کند. همچنین، از نظر معیارهای حساسیت^۱ و ویژگی^۲، الگوریتم جنگل تصادفی عملکرد بهتری نسبت به SVM داشت. حساسیت بالاتر در برخی کلاس‌های خاص و ویژگی بالاتر در اکثر کلاس‌ها، نشان‌دهنده قدرت بالاتر جنگل تصادفی در مدیریت داده‌های پیچیده و تنوع‌دار است. با توجه به دقت بالاتر و ضریب کاپای بیشتر، الگوریتم جنگل تصادفی به عنوان بهترین الگوریتم برای این پروژه انتخاب شد. این الگوریتم توانست با مدیریت بهتر داده‌های بزرگ و پیچیده و کاهش بیش‌برازش، نتایج قابل اطمینان‌تری ارائه دهد. به طور کلی، جنگل تصادفی به دلیل دقت و قابلیت اطمینان بیشتر در طبقه‌بندی تصاویر ماهواره‌ای سنتینل-۲ برتری داشت و به عنوان الگوریتم برتر در این پروژه تعیین گردید.

^۱ Sensitivity

^۲ Specificity

بحث و نتیجه‌گیری

گندم یکی از محصولات غله‌ای مهم با پروتئین بالا است (آلتینباک و دوپونت، ۲۰۰۳). هیچ غله‌ای به علت بازدهی زیاد و امکان کشت گندم در بیشتر نقاط جهان و نیز قابلیت پخت و خاصیت تغذیه‌ای و صنعتی و کیفیت فوق‌العاده گلوتن آن برای تولید نان نمی‌تواند با آن رقابت کند (رجب‌زاده، ۱۳۸۰). گندم که به‌طور متوسط فراهم‌کننده‌ی ۲۲ درصد کل پروتئین در رژیم غذایی انسان است، بیشترین سطح زیرکشت و تولید را در دنیا به خود اختصاص داده‌است. در حال حاضر رشد روزافزون و بیش از حد جمعیت و احساس کمبود منابع در دسترس برای رفع نیازهای اساسی و مهم انسان‌ها، یکی از بزرگترین مشکلات روبه‌روی جهانیان است. پایه و اساس برنامه‌ریزی و مدیریت صحیح در امورات کشاورزی را به داشتن آمار و اطلاعات دقیق و به‌روز از ظرفیت‌های کشاورزی می‌توان دانست (عبدالله‌زاده و نصیری، ۱۳۸۷). گندم تقریباً ۲۰ درصد از کل کالری و پروتئین مورد نیاز رژیم غذایی را برای ۴/۵ میلیارد نفر با سطح برداشت ۲۱۵/۹ میلیون هکتار تامین می‌کند (فائو، ۲۰۱۹).

در این پژوهش، دو الگوریتم ماشین‌بردار پشتیبان و جنگل تصادفی مورد بررسی قرار گرفتند. نتایج حاصل از اجرای این الگوریتم‌ها نشان داد که هر الگوریتم دارای نقاط قوت و ضعف خاص خود می‌باشد. الگوریتم ماشین‌بردار پشتیبان با میزان دقت ۹۰ درصد و ضریب کاپای ۸۲ درصد عملکرد قابل قبولی داشت. با این حال، در این پژوهش در مقایسه با دیگر الگوریتم یعنی الگوریتم جنگل تصادفی عملکرد ضعیف‌تری داشت. الگوریتم جنگل تصادفی با میزان دقت ۹۳ درصد و ضریب کاپای ۸۷ درصد عملکرد بسیار خوبی نشان داد. این الگوریتم به‌دلیل توانایی در ترکیب مدل‌های مختلف و کاهش اثر بیش‌برازش معمولاً نتایج دقیقی ارائه می‌دهد. با این حال، پیچیدگی محاسبات بالای آن می‌تواند در کاربردهای بزرگ‌تر مشکل‌ساز باشد. نتایج این پژوهش در بخش بعدی با نتایج چندین پژوهش مشابه مقایسه و بررسی شده است.

در مطالعه‌ای که توسط زارعی و همکاران (۱۳۹۷) انجام شد، الگوریتم SVM با دقت ۸۸ درصد و ضریب کاپای ۸۰ درصد عملکرد مناسبی در طبقه‌بندی داده‌های کشاورزی داشت، اما در مقایسه با الگوریتم RF که دقت ۹۱ درصد و ضریب کاپای ۸۶ درصد را نشان داد، عملکرد ضعیف‌تری داشت. نتایج این پژوهش نیز نشان‌دهنده برتری نسبی الگوریتم جنگل تصادفی بر ماشین‌بردار پشتیبان بود، که مشابه با یافته‌های پژوهش حاضر است. مطالعه دیگری توسط محمدی و همکاران (۱۳۹۸) انجام شد که در آن الگوریتم SVM با دقت ۸۷ درصد و ضریب کاپای ۷۹ درصد و الگوریتم RF با دقت ۹۲ درصد و ضریب کاپای ۸۶ درصد مورد بررسی قرار گرفتند. این پژوهش نیز نشان داد که الگوریتم RF به دلیل توانایی بالا در ترکیب مدل‌های مختلف و کاهش اثر بیش‌برازش، عملکرد بهتری نسبت به SVM دارد. نتایج این مطالعه با یافته‌های پژوهش حاضر سازگار بوده و برتری الگوریتم جنگل تصادفی را تأیید می‌کند. همچنین، در مطالعه‌ای که توسط عزیززاده و همکاران (۱۳۹۹) انجام شد، الگوریتم SVM با دقت ۸۹ درصد و ضریب کاپای ۸۱ درصد و الگوریتم RF با دقت ۹۴ درصد و ضریب کاپای ۸۸ درصد مورد ارزیابی قرار گرفتند. این پژوهش نیز نشان داد که الگوریتم RF به دلیل کاهش اثر بیش‌برازش و توانایی در ترکیب مدل‌های مختلف، عملکرد دقیق‌تری ارائه می‌دهد. در مجموع، مقایسه نتایج پژوهش حاضر با پژوهش‌های مشابه نشان می‌دهد که الگوریتم جنگل تصادفی با توجه به انتخاب تعداد درختان، نوع پوشش و ... که در این مطالعه و مطالعات مشابه انجام گرفته است، عملکرد بهتری نسبت به الگوریتم ماشین‌بردار پشتیبان در طبقه‌بندی داده‌های کشاورزی دارد. این یافته‌ها تأیید می‌کنند که الگوریتم RF به دلیل توانایی بالا در ترکیب مدل‌های مختلف و کاهش اثر بیش‌برازش، می‌تواند به عنوان یک روش موثر در تحلیل داده‌های کشاورزی مورد استفاده قرار گیرد. از سوی دیگر، الگوریتم SVM با دقت و کارایی مناسب خود، همچنان به عنوان یک روش کارآمد در بسیاری از مسائل طبقه‌بندی محسوب می‌شود، اما ممکن است در مقایسه با RF در برخی موارد عملکرد ضعیف‌تری داشته باشد. از نظر کمی، دقت بالاتر جنگل تصادفی نسبت به ماشین‌بردار پشتیبان، نشان‌دهنده توانایی بهتر این الگوریتم در شناسایی دقیق‌تر انواع کاربری اراضی است. برای مثال، در تحلیل ماتریس‌های خطا، جنگل تصادفی در شناسایی کلاس‌هایی با تنوع طیفی بیشتر مانند مناطق شهری و کشت آبی و باغات، نتایج بهتری ارائه داد. در حالی که ماشین‌بردار پشتیبان در طبقه‌بندی کلاس‌های خاصی دچار مشکلاتی شد و در برخی موارد خطاهای بیشتری به نمایش گذاشت، جنگل تصادفی به دلیل تجمع اطلاعات از

درخت‌های متعدد، دقت بهتری ارائه داد. این یافته‌ها نشان می‌دهند که جنگل تصادفی توانایی بهتری در مدیریت داده‌های پیچیده و کاهش تأثیر نویز دارد و می‌تواند در طبقه‌بندی دقیق‌تر کاربری اراضی به‌ویژه در مناطق با تغییرات شدید طیفی مؤثرتر باشد. به طور کلی، دقت بالاتر و کاهش خطاها در الگوریتم جنگل تصادفی به وضوح بیانگر برتری این روش نسبت به ماشین‌بردار پشتیبان در این مطالعه است. پژوهش حاضر به اهداف خود در زمینه طبقه‌بندی دقیق کاربری اراضی با استفاده از الگوریتم‌های یادگیری ماشین به میزان قابل توجهی دست یافته است. با انتخاب و پیاده‌سازی الگوریتم جنگل تصادفی به عنوان بهترین گزینه، توانسته‌ایم به دقت بالاتری در طبقه‌بندی کاربری اراضی دست یابیم و تحلیل‌های دقیق‌تری از داده‌های ماهواره‌ای ارائه دهیم. به طور کلی، این مطالعه نشان‌دهنده قابلیت‌های الگوریتم‌های یادگیری ماشین در مدیریت داده‌های پیچیده و بهبود فرآیندهای طبقه‌بندی است. نتایج به دست آمده، توانسته است اهداف تحقیق را به خوبی محقق سازد و زمینه‌های بهینه‌سازی و پیشرفت‌های آینده را برای استفاده از تصاویر ماهواره‌ای در تحلیل‌های جغرافیایی فراهم آورد. با توجه به اهمیت گندم به عنوان یکی از محصولات غله‌ای مهم با پروتئین بالا و نقش اساسی آن در تامین نیازهای غذایی جهانی، استفاده از روش‌های دقیق و موثر در تحلیل داده‌های مربوط به این محصول از اهمیت بالایی برخوردار است. نتایج این پژوهش نشان می‌دهد که استفاده از الگوریتم جنگل تصادفی می‌تواند به بهبود دقت و کارایی تحلیل‌ها کمک کند و به عنوان یک ابزار قدرتمند در مدیریت و برنامه‌ریزی‌های کشاورزی مورد استفاده قرار گیرد.

منابع

- اصغری سراسکانرود صیاد، جلیلیان روح اله، پیروزی نژاد نوشین، مددی عقیل، یادگاری میلاد. (۲۰۲۰). ارزیابی شاخص های استخراج آب با استفاده از تصاویر ماهواره ای لندست (مطالعه موردی: رودخانه گاماسیاب کرمانشاه). نشریه تحقیقات کاربردی علوم جغرافیایی، سال بیستم، شماره ۵۸، پاییز ۹۹
- اصغری سراسکانرود، صیاد، شریفی طولارود، حسین، سبحانی، بهروز (۱۴۰۳). تهیه نقشه کشت برنج براساس خصوصیات فنولوژیکی با استفاده از سری زمانی تصاویر سنتینل ۱. سنجش از دور و سامانه اطلاعات جغرافیایی و منابع طبیعی، شماره ۲، دوره ۱۵. رجبزاده، ناصر. (۱۳۸۰). "مبانی فناوری غلات". انتشارات دانشگاه تهران.
- رضایی مقدم، محمدحسین؛ ولیزاده کامران، خلیل؛ اندریانی، صغری؛ الماس پور، فرهاد (۱۳۹۳). مقایسه روش های شبکه عصبی و ماشین‌بردار پشتیبان در استخراج نقشه های کاربری و پوشش اراضی با استفاده از تصاویر لندست ۸ (مطالعه موردی: حوضه صوفی چای). جغرافیا و برنامه ریزی، ۱۹(۵۲)، ۱۸۳-۱۶۳.
- ریاحی، ضیائیان فیروزآبادی، عزیزپور، دارویی، پرستو. (۲۰۱۹). تعیین و بررسی سطح زیرکشت محصولات زراعی در ناحیه لنجانان با استفاده از تصاویر ماهواره ای. نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۹(۵۲)، ۱۴۷-۱۶۹.
- سازمان غذا و کشاورزی ملل متحد (فائو). (۲۰۱۹). گزارش سالانه تولید و مصرف گندم در جهان. فائو، رم.
- سیدمحمدی، جعفرزاده، علی اصغر، سرمیدیان، شهبازی، فرزین، قربانی. (۲۰۱۷). مقایسه کارایی روش های AHP، TOPSIS و ریشه دوم در تعیین اولویت کشت گندم، جو و ذرت تحت آبیاری بارانی در دشت مغان. دانش آب و خاک، ۲۷(۲)، ۴۵-۵۹.
- صدوقی، رجایی، طاهر، روحانی. (۲۰۲۱). شناسایی و بررسی تغییرات سطح زیر کشت محصولات زراعی روستای حسین آباد میش مست با استفاده از تصاویر ماهواره ای. علوم آب و خاک (Isfahan University of Technology)، ۲۴(۴)، ۲۳۹-۲۵۴.
- عبادی، الهامه؛ اسفندیاری درآباد، فریبا؛ اصغری سراسکانرود، صیاد؛ مصطفی زاده، رثوف؛ ملانوری، الهام (۱۴۰۳). بررسی تغییرات کاربری اراضی با استفاده از تکنیک شی گرا و پیکسل پایه (مطالعه موردی: حوضه فیروزآباد خلخال). جغرافیا و برنامه ریزی، ۲۸(۸۷)، ۲۷۵-۲۶۱.
- عبداله زاده محمود، برادران نصیری مهران. (۲۰۰۸). تعیین سطح زیر کشت سیب زمینی در شهرستان بروجن با استفاده از سری زمانی تصاویر IRSP6 (همایش).
- عبیات، امانپور، عبیات، عبیات، ماجده. (۲۰۲۲). برآورد سطح زیرکشت محصولات کشاورزی با استفاده از تصاویر ماهواره لندست ۸ (مطالعه موردی: شهرستان شوشتر). به زراعی کشاورزی ۲۴(۲)، ۴۶۵-۴۷۹.
- مجرد، فیروز؛ فرهادی، بهمن؛ الفتی، ژیلا (۱۳۹۶). شبیه سازی عملکرد و تعیین تاریخ کاشت گندم دیم و آبی بر مبنای عوامل اقلیمی و محیطی در استان کرمانشاه با مدل آکواکراپ. جغرافیا و برنامه ریزی شهری، ۲۲(۶۶)، ۲۷۱-۲۵۱.
- موسوی، سید احمد، عباس زاده طهرانی، نادیا، جانعلی پور، میلاد. (۱۳۹۹). برآورد سطح زیر کشت گندم دیم با استفاده از تصاویر ماهواره ای سنتینل ۲- مطالعه موردی: منطقه سجاس رود شهرستان خدابنده، استان زنجان. پژوهش و فناوری محیط زیست، ۵(۷)، ۷۷-۹۰.
- نوروزی، صانعی، رزقی. (۲۰۱۹). شناسایی و تفکیک مزارع برنج با استفاده از روش نیمه خودکار در شمال ایران. زراعت و اصلاح نباتات، ۱۴(۴)، ۱۱-۲۱.
- نویدی، میرناصر، سیدمحمدی، سید جلالی، زین الدینی، فرج نیا، دلسوز. (۲۰۲۲). ارزیابی تناسب اراضی برای کشت گندم در دشت های آبی ایران. پژوهش های خاک، ۳۶(۲)، ۱۲۷-۱۴۵.

References

- Abdollahzadeh, M., Nasiri Mehran Brothers. (2008). *Determining the Potato Cultivation Area in Boroujen County Using Time Series of IRS-P6 Images* (Conference). (In Persian)
- Abiyat, A., Amanpour, A., Abiyat, A., Abiyat, A., Majdeh, M. (2022). *Estimating Agricultural Crop Area Using Landsat 8 Satellite Images (Case Study: Shushtar County)*. *Journal of Agricultural Science*, 24(2), 465-479. (In Persian)
- Asghari Saraskanroud, S., Jalilian, R., Pirouzi Nejad, N., Madadi, A., Yadegari, M. (2020). *Evaluation of Water Extraction Indices Using Landsat Satellite Images (Case Study: Gamasiab River, Kermanshah)*. In Persian)
- Asghari Saraskanroud, Seyed; Sharifi Toularoud, Hossein; Sobhani, Behrouz (2024). Mapping Rice Cultivation Based on Phenological Characteristics Using Sentinel-1 Time Series Imagery. *Remote Sensing and Geographic Information Systems in Natural Resources*, No. 2, Vol. 15. (In Persian)
- Biau, G., Cérou, F., Guyader, A. (2010). *On the Rate of Convergence of the Bagged Nearest Neighbor Estimate*. *Journal of Machine Learning Research*, 11(2).^{۹۹}
- Breiman, L. (1996). *Bagging predictors*. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/bf00058655>
- Bühlmann, P., Yu, B. (2002). *Analyzing bagging*. *Annals of Statistics*, 30(4). <https://doi.org/10.1214/aos/1031689014>
- Buja, A., Stuetzle, W. (2006). *OBSERVATIONS ON BAGGING*. *Statistica Sinica*, 16(2), 323–351. <http://stat.wharton.upenn.edu/~buja/PAPERS/sinica-bagging-buja-stuetzle.pdf>
- Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., Peng, B. (2019). *Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches*. *Agricultural and Forest Meteorology*, 274, 144–159. <https://doi.org/10.1016/j.agrformet.2019.03.010>
- Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., Tao, F. (2021). *Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine*. *European Journal of Agronomy*, 123, 126204. <https://doi.org/10.1016/j.eja.2020.126204>
- Dupont, F., Altenbach, S. (2003). *Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis*. *Journal of Cereal Science*, 38(2), 133–146. [https://doi.org/10.1016/s0733-5210\(03\)00030-4](https://doi.org/10.1016/s0733-5210(03)00030-4)
- Ebadi, Elhameh; Esfandiari Darabad, Fariba; Asghari Saraskanrud, Sayad; Mostafazadeh, Raouf; Melanouri, Elham (2024). "Assessment of Land Use Changes Using Object-Oriented and Pixel-Based Techniques (Case Study: Firoozabad Basin, Khalkhal)." *Geography and Planning*, 28(87), 261-275. (In Persian)
- Feng, Y., Chen, B., Liu, W., Xue, X., Liu, T., Zhu, L., & Xing, H. (2024). Winter Wheat Mapping in Shandong Province of China with Multi-Temporal Sentinel-2 Images. *Applied Sciences*, 14(9), 3940. <https://doi.org/10.3390/app14093940>
- Food and Agriculture Organization of the United Nations (FAO). (2013). *Annual Report on World Wheat Production and Consumption*. FAO, Rome. (In Persian)

- Food and Agriculture Organization of the United Nations (FAO). (2019). *Annual Report on World Wheat Production and Consumption*. FAO, Rome. (In Persian)
- Gashaw, T., Tulu, T., Argaw, M., Worqlul, A. W. (2018). *Modeling the hydrological impacts of land use/land cover changes in the Andassa watershed, Blue Nile Basin, Ethiopia*. *Science of the Total Environment*, 619–620, 1394–1408. <https://doi.org/10.1016/j.scitotenv.2017.11.191>
- Gordon, A. D., Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. (1984). *Classification and regression trees*. *Biometrics*, 40(3), 874. <https://doi.org/10.2307/2530946>
- Hsu, C., Lin, C. (2002). *A comparison of methods for multiclass support vector machines*. *IEEE Transactions on Neural Networks*, 13(2), 415–425. <https://doi.org/10.1109/72.991427>
- Hudait, Manas; Pravin Patel, Priyank; (2022). Crop-type mapping and acreage estimation in smallholding plots using Sentinel-2 images and machine learning algorithms: Some comparisons, *The Egyptian Journal of Remote Sensing and Space Science*, Volume 25, Issue 1, Pages 147-156.
- Islam, Rahedul; (2021). Crop Calender Mapping of Bangladesh Rice Paddy Field with ALOS-2 ScanSAR Data, *Advances in Remote Sensing*, Vol. 10, Pages 115-129.
- Jeong, S., Ko, J., Yeom, J. (2022). *Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea*. *Science of the Total Environment*, 802, 149726. <https://doi.org/10.1016/j.scitotenv.2021.149726>
- Kummu, M., Fader, M., Gerten, D., Guillaume, J. H., Jalava, M., Jägermeyr, J., Pfister, S., Porkka, M., Siebert, S., Varis, O. (2017). *Bringing it all together: linking measures to secure nations' food supply*. *Current Opinion in Environmental Sustainability*, 29, 98–117. <https://doi.org/10.1016/j.cosust.2018.01.006>
- Lai, Y., Pringle, M., Kopittke, P., Menzies, N., Orton, T., Dang, Y. (2018). *An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI*. *International Journal of Applied Earth Observation and Geoinformation*, 72, 99–108. <https://doi.org/10.1016/j.jag.2018.07.013>
- Luo, Y., Zhang, Z., Cao, J., Zhang, L., Zhang, J., Han, J., Zhuang, H., Cheng, F., Tao, F. (2022). *Accurately mapping global wheat production system using deep learning algorithms*. *International Journal of Applied Earth Observation and Geoinformation*, 110, 102823. <https://doi.org/10.1016/j.jag.2022.102823>
- Luo, Y., Zhang, Z., Zhang, L., Cao, J. (2021). *Spatiotemporal patterns of winter wheat phenology and its climatic drivers based on an improved pDSSAT model*. *Science China. Earth Sciences/Science China. Earth Sciences*, 64(12), 2144–2160. <https://doi.org/10.1007/s11430-020-9821-0>
- Mbengue, F., Faye, G., Talla, K., Sarr, M. A., Ferrari, A., Mbaye, M., Dramé, M. S., Sagne, P. (2022). *Evaluation of machine learning classification methods for rice detection using earth observation data: Case of Senegal*. *European Scientific Journal*, 18(17), 214. <https://doi.org/10.19044/esj.2022.v18n17p214>
- Mercier, G., Lennon, M. (2003, July). *Support vector machines for hyperspectral image classification with spectral-based kernels*. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium*. Proceedings (IEEE Cat. No. 03CH37477) (Vol. 1, pp. 288-290). IEEE.

- Minta, M., Kibret, K., Thorne, P., Nigussie, T., Nigatu, L. (2018). *Land use and land cover dynamics in Dendi-Jeldu hilly-mountainous areas in the central Ethiopian highlands*. *Geoderma*, 314, 27–36. <https://doi.org/10.1016/j.geoderma.2017.10.035>
- Mojarad, Firooz; Farhadi, Bahman; Alfti, Jila (2017). "Simulation of Yield and Determination of Sowing Date for Rainfed and Irrigated Wheat Based on Climatic and Environmental Factors in Kermanshah Province Using the AquaCrop Model." *Geography and Urban Planning*, 22(66), 251-271. (In Persian)
- Mousavi, S. A., Abbaszadeh Tehrani, N., Janali Pour, M. (2020). *Estimation of Rainfed Wheat Cultivation Area Using Sentinel-2 Satellite Images: A Case Study in Sajasrood Region, Khodabandeh County, Zanjan Province*. In Persian)
- Nasrallah, A., Baghdadi, N. N., Mhawej, M., Faour, G., Darwish, T., Belhouchette, H., Darwich, S. (2018). *A novel approach for mapping wheat areas using high resolution Sentinel-2 images*. *Sensors*, 18(7), 2089. <https://doi.org/10.3390/s18072089>
- Navidi, M., Mir-Naser, M., Seyed Mohammadi, S., Seyed Jalali, M., Zeinoddini, M., Farajnia, M., Delsouz. (2022). *Land suitability assessment for wheat cultivation in Iran's wetlands*. *Soil Research*, 36(2), 127-145. (In Persian)
- Nelson, G. C., Rosegrant, M. W., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R., You, L. (2010). *Food security, farming, and climate change to 2050: scenarios, results, policy options* (Vol. 172). Intl Food Policy Res Inst.
- Nelson, K. S., Burchfield, E. K. (2021). *Landscape complexity and US crop production*. *Nature Food*, 2(5), 330–338. <https://doi.org/10.1038/s43016-021-00281-1>
- Noroozi, S., Sanaei, M., Razeghi, H. (2019). *Identification and differentiation of rice fields using semi-automatic method in northern Iran*. *Agriculture and Plant Breeding*, 14(4), 11-21. (In Persian)
- Peng, B., Guan, K., Zhou, W., Jiang, C., Frankenberg, C., Sun, Y., He, L., Köhler, P. (2020). *Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction*. *International Journal of Applied Earth Observation and Geoinformation*, 90, 102126. <https://doi.org/10.1016/j.jag.2020.102126>
- Rajabzadeh, N. (2001). *Fundamentals of Cereal Technology*. University of Tehran Press. (In Persian)
- Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). *Recent patterns of crop yield growth and stagnation*. *Nature Communications*, 3(1). <https://doi.org/10.1038/ncomms2296>
- Rembold, F., Atzberger, C., Savin, I. V., Rojas, O. (2013). *Using low resolution satellite imagery for yield prediction and yield anomaly detection*. *Remote Sensing*, 5(4), 1704–1733. <https://doi.org/10.3390/rs5041704>
- Rezai-Moghaddam, Mohammad-Hosseini; Vali-Zadeh Kamran, Khalil; Andriyani, Soghra; Almaspour, Farhad (2014). *Comparison of Neural Network and Support Vector Machine Methods in Extracting Land Use and Land Cover Maps Using Landsat 8 Imagery (Case Study: Sofi-Chay Basin)*. *Geography and Planning*, 19(52), 183-163. (In Persian)

- Riahi, Ziaian Firoozabadi, Azizpour, Darouei, Parasto. (2019). *Determination and Examination of Crop Cultivation Area in Lenjanat Region Using Satellite Images. Journal of Applied Research in Geographic Sciences*, 19(52), 147-169. (In Persian)
- Sadoughi, R., Rajaei, T., Taher, A., Rouhani, A. (2021). *Identification and Investigation of Changes in Agricultural Land Use in Hosseinabad Mishmast Village Using Satellite Images. Water and Soil Sciences (Isfahan University of Technology)*, 24(4), 239-254. (In Persian)
- Seyyed Mohammadi, J., Jafarzadeh, A., Asghar, A., Sarmadian, S., Shahbazi, F., Farzin, F., Ghorbani, M. (2017). *Comparison of Efficiency of TOPSIS, AHP, and Root Square Methods in Determining Priority of Wheat, Barley, and Corn Cultivation under Rainfed Conditions in Moghan Plain. Water and Soil Science Journal*, 27(2), 45-59. (In Persian)
- Tilman, D., Balzer, C., Hill, J., Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, 108(50), 20260–20264. <https://doi.org/10.1073/pnas.1116437108>