

Analysis of machine learning algorithms' performance in multitemporal image classification for agricultural management (Case study: Qazvin plain)

Masoud Soltani¹, Bahareh Bahmanabadi^{2*}

¹ Assistant Professor, Department of Water Science and Engineering, Faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran

² Ph.D., Department of Water Science and Engineering, Faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran

Extended Abstract

Introduction

Due to the rapid increase in global population and the intensifying challenges posed by climate change, there is an urgent need for sustainable farming methods and effective natural resource management. Agriculture constitutes the foundation of economic stability and food security, therefore, approaches that promote optimal resource utilization while causing minimal environmental impacts are essential. Achieving this objective necessitates the ability to carry out rapid and precise crop mapping, which plays pivotal role in planning agricultural operations, yield estimation, and ensuring sustainability.

Traditionally, crop type monitoring and mapping have relied on field-based approaches. However, they are often labor-intensive, time-consuming, and expensive, especially when applied on a large scale. In contrast, remote sensing approaches, empowered by advanced satellite imagery and machine-learning technologies, offer a practical alternative for large-scale monitoring and detailed classification of land cover and crop types. Among the most promising tools are the Sentinel-1 and Sentinel-2 satellites, which provide high spatial, temporal, and spectral resolution data, making them well-suited for agricultural applications.

Sentinel-1 generates radar imagery, which is especially useful for monitoring vegetation structure even under cloudy conditions. Sentinel-2, on the other hand, supplies high-resolution optical imagery with numerous spectral bands, ideal for vegetation study. By integrating these datasets, complementary physical and spectral characteristics of the land surface can be effectively captured.

Various ML algorithms exist. Herein, the performance of three of the most widely used algorithms: Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost). Evaluating the relative strengths of each algorithm in classification provides essential insights into their applicability across diverse agricultural settings.

Materials and Methods

This study was conducted in the Qazvin Irrigation Network, an agriculturally significant region of about 80,000 hectares in Iran. The study area includes different types of land cover, such as wheat, alfalfa, fallow land, urban areas, and bare soil, representing the most commonly cultivated crops in the region. To address these challenges, data from the Sentinel-1 and Sentinel-2 satellites were utilized. The former acquires reliable radar data that captures surface roughness and soil moisture even under cloudy conditions. Meanwhile, the latter delivers high-resolution optical imagery with many spectral bands, which is excellent for studying vegetation health and structure. Several necessary data-preprocessing steps were carried out to ensure that accurate classifications were developed. Atmospheric and sensor noise reduction was accomplished via radiometric and geometric corrections, respectively, for radar and optical imagery. Derived spectral indices such as NDVI, SAVI, and LAI aided the detection of vegetation characteristics under study by enhancing separability at spectral levels. Furthermore, temporal fusion was performed by combining images taken at different times to account for the phenological changes in vegetation over the growing seasons.

Finally, three machine-learning algorithms were implemented for classifying the preprocessed satellite images: RF, SVM, and XGBoost. The reasons for choosing RF, an ensemble-based approach, include its robustness to noise and its ability to handle complicated datasets. On the other hand, SVM was adopted because it optimizes classification boundaries through its kernel-based feature. XGBoost is a highly accurate advanced gradient boosting technique that can support large-scale computing with low expenses. The dataset was then divided into a training and testing set in a 70/30 ratio to prevent overfitting and ensure the model's reliability. Classification accuracy was assessed based on overall accuracy and the kappa coefficient, while the Jeffries-Matusita (*JM*) test quantified spectral separability between land cover classes.

Results and Discussion

The results demonstrated that integrating optical and radar data significantly improved classification accuracy. Among the three algorithms, RF outperformed the others, achieving an overall accuracy of 93.98% and a kappa coefficient of 0.996. These results highlight RF's ability to handle spectrally overlapping classes and complex datasets effectively.

The XGBoost algorithm also performed well, achieving an overall accuracy of 93.94%. However, its performance was slightly hindered by its inability to distinguish between classes with similar spectral characteristics, such as wheat and alfalfa. While providing reasonable results, SVM achieved a lower overall accuracy of 83.79%, mainly due to its susceptibility to spectral overlap.

The *JM* test revealed that certain classes, such as wheat and alfalfa, exhibited low spectral separability. This limitation underscores the importance of integrating radar data and spectral indices to enhance differentiation. The study also highlighted the potential of temporal data fusion to capture phenological changes, further improving classification performance.

Conclusion

This study indicated the potential of integrating multi-source remote sensing data and machine learning algorithms for crop classification in semi-arid regions. The RF algorithm proved the most accurate and robust method, showing its adaptability to the heterogeneous and complex nature of the datasets. XGBoost and SVM are also very promising, but their performance could be improved further with advanced parameter tuning.

Future research should investigate the application of more advanced techniques, such as CNNs and deep learning frameworks, to improve classification accuracy further. A deeper understanding of the dynamics of crops and land use changes can be achieved by including multi-temporal and multi-spectral datasets.

The results of such a study can have substantial implications for sustainable agriculture and resource management. In this context, remote sensing and machine learning technologies offer means to address critical challenges related to food security and environmental conservation in the most climate-vulnerable regions.

Keywords: Sentinel-1, Sentinel-2, Machine learning, Random Forest, XGBoost, Support Vector Machine.

Article Type: Research Article

Acknowledgement

We would like to express our sincere gratitude to the Imam Khomeini International University for the financial and logistical supports that significantly contributed to this research project.

Conflicts of interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data Availability Statement:

The datasets are available upon a reasonable request to the corresponding author.

Authors' contribution

Masoud Soltani: Guiding, Data Collection, Conceptualization, Editing the paper, Controlling the results; **Bahareh Bahmanabadi:** Programming, Conceptualization, methodology, Writing - Initial draft preparation, performing software/statistical analysis

*Corresponding Author, E-mail: b.bahmanabadi@alumni.ut.ac.ir

Citation: Soltani, M. & Bahmanabadi, B. (2025). Analysis of machine learning algorithms' performance in multitemporal image classification for agricultural management (Case study: Qazvin plain). *Water and Soil Management and Modeling*, 5(3), 54-73.

doi: 10.22098/mmws.2025.16425.1532

Received: 23 December 2024, Received in revised form: 10 March 2025, Accepted: 15 March 2025, Published online: 23 September 2025.

Water and Soil Management and Modeling, Year 2025, Vol. 5, No. 3, pp. 54-73.

Publisher: University of Mohaghegh Ardabili

© Author(s)





تحلیل عملکرد الگوریتم های یادگیری ماشین در طبقه بندی تصاویر چند زمانه در مدیریت کشاورزی (مطالعه موردی: دشت قزوین)

مسعود سلطانی^۱، بهاره بهمن آبادی^{۲*}

^۱ استادیار، گروه علوم و مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران
^۲ دکتری، گروه علوم و مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران

چکیده

با افزایش جمعیت و تغییرات اقلیمی، مدیریت کشاورزی به منظور تضمین امنیت غذایی نیاز به اطلاعات صحیح و دقیق با پیوستگی زمانی و مکانی بالا دارد. با بهره گیری از فناوری سنجش از دور به همراه الگوریتم های یادگیری ماشین، امکان طبقه بندی محصولات کشاورزی با صحت زیادی فراهم شده است که می تواند نقش به سزایی در پیش بینی وضعیت محصولات و نیازهای منابع ایفا کند. این مطالعه با هدف ارزیابی صحت طبقه بندی تصاویر ماهواره ای سنتینل ۱ و ۲ برای دوره های کشت بهاره و پاییزه در سال ۱۴۰۳-۱۴۰۲ در منطقه دشت قزوین انجام شده است. در این پژوهش، عملکرد الگوریتم های جنگل تصادفی، ماشین بردار پشتیبان و تقویت گرادیان شدید مورد بررسی قرار گرفت. تصاویر ماهواره ای سنتینل ۱ و ۲ برای تحلیل پنج کلاس اصلی (گندم، یونجه، آیش، مناطق شهری و اراضی بایر) که محصولات غالب در منطقه هستند را در دو دوره کشت، با استفاده از شاخص های SAVI، NDVI و LAI پردازش شدند. داده های آموزش/تست با نسبت ۳۰/۷۰ تقسیم و نتایج طبقه بندی ها با شاخص هایی چون صحت کلی و ضریب کاپا ارزیابی شد. از آزمون جفریس-ماتوسیتا (JM) برای بررسی تفکیک پذیری طیفی کلاس ها استفاده شد. نتایج ارزیابی نشان داد که الگوریتم جنگل تصادفی با صحت کلی ۹۸/۹۳ درصد و ضریب کاپای ۰/۹۹ بهترین عملکرد را در تفکیک کلاس ها ارائه داد. الگوریتم تقویت گرادیان شدید نیز با صحت کلی ۹۳/۹۴ درصد عملکرد قابل قبولی داشت، اما برای کلاس هایی با شباهت طیفی، صحت کمتری نسبت به جنگل تصادفی نشان داد. در مقابل، روش ماشین بردار پشتیبان به دلیل حساسیت به هم پوشانی طیفی، در جداسازی کلاس ها عملکرد ضعیف تری داشت و در نهایت روش طبقه بندی جنگل تصادفی با صحت کلی و ضریب کاپای بیش از ۰/۹۹ بهترین عملکرد را داشت. آزمون JM نشان داد که برخی کلاس ها مانند گندم و یونجه جداسازی ضعیفی دارند، اما استفاده از داده های راداری و شاخص های طیفی توانسته جداسازی را بهبود بخشد. استفاده از ترکیب داده های اپتیک و راداری به همراه الگوریتم های یادگیری ماشین، ابزاری مؤثر برای مدیریت منابع کشاورزی است.

واژه های کلیدی: سنتینل ۱ و ۲، جنگل تصادفی، ماشین بردار پشتیبان، تقویت گرادیان شدید.

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: b.bahmanabadi@alumni.ut.ac.ir

استناد: سلطانی، مسعود. و بهمن آبادی، بهاره (۱۴۰۴). تحلیل عملکرد الگوریتم های یادگیری ماشین در طبقه بندی تصاویر چند زمانه در مدیریت کشاورزی (مطالعه موردی: دشت قزوین). مدل سازی و مدیریت آب و خاک، ۵(۳)، ۵۴-۷۳.
doi: 10.22098/mmws.2025.16425.1532

تاریخ دریافت: ۱۴۰۳/۱۰/۰۳، تاریخ بازنگری: ۱۴۰۳/۱۲/۲۰، تاریخ پذیرش: ۱۴۰۳/۱۲/۲۵، تاریخ انتشار: ۱۴۰۴/۰۷/۰۱

مدل سازی و مدیریت آب و خاک، سال ۱۴۰۴، دوره ۵، شماره ۳، صفحه ۵۴ تا ۷۳.

ناشر: دانشگاه محقق اردبیلی
© نویسندگان



۱- مقدمه

با توسعه اجتماعی و اقتصادی، رشد سریع جمعیت و تأثیرات تغییرات اقلیمی جهانی و گسترش شهرنشینی، توسعه کشاورزی پایدار با چالش‌های قابل‌توجهی مواجه شده است (Li et al., 2019). بنابراین، دستیابی به اطلاعات به‌موقع، کارآمد، صحیح و دقیق درباره مساحت کشت و توزیع مکانی محصولات زراعی برای مدیریت دولتی در تدوین سیاست‌های غذایی کشاورزی، بهینه‌سازی منابع زمین، تنظیم ساختارهای کشت و تضمین امنیت غذایی ملی ضروری است (Dong et al., 2020). پوشش اراضی کشاورزی، در مقایسه با سایر انواع پوشش‌های زمینی، از تغییرپذیری بسیار بالایی در بازه‌های زمانی کوتاه‌مدت برخوردار است. این تغییرات عمدتاً ناشی از تحولات فنولوژیکی گیاهان زراعی است که از مرحله کاشت تا برداشت، چندین بار دچار تغییر می‌شوند. همچنین، در بسیاری از موارد، زمین‌های کشاورزی در طول یک سال زراعی بیش از یک بار زیر کشت محصولات مختلف قرار می‌گیرند، که این موضوع باعث تنوع و تغییرات مداوم در نوع پوشش اراضی می‌شود (Gili et al., 2022).

با توجه به پویایی پوشش اراضی کشاورزی، بررسی سطح زیر کشت محصولات کشاورزی از روش‌های میدانی و سایر روش‌ها برای جمع‌آوری داده‌های زمینی استفاده می‌شود که این فرآیند پیچیده، زمان‌بر و پرهزینه است. با پیشرفت و توسعه فناوری‌های سنجش از دور مانند ماهواره‌ها، پهپادها و سایر تصاویر هوایی امکان ثبت مناطق وسیع بدون مداخله مستقیم فراهم شد (Ashourloo et al., 2022). این فناوری‌ها امکان پایش محصولات در مناطق کشاورزی گسترده و ارزیابی سلامت محصولات را فراهم کردند (Li et al., 2022). چرا که، جزئیات دقیق‌تری را در گستره‌های وسیع و در طول زمان فراهم می‌کنند. دسترسی به تصاویر با وضوح مکانی، زمانی، طیفی و رادیومتری بالا از ماهواره‌ها به‌طور قابل‌توجهی دامنه کاربرد این فناوری‌ها را در کشاورزی با بهبود دقت و صحت، افزایش داده است (Seifi Majdar and Ghassemian, 2017; Li et al., 2022). با توسعه سریع فناوری سنجش از دور، کیفیت و سرعت به‌روزرسانی داده‌های سنجش از دور به‌طور قابل‌توجهی بهبود یافته است و داده‌های چندمنبعی سنجش از دور به‌طور گسترده‌ای در زمینه‌های کشاورزی، جنگلداری، دریایی، حفاظت محیط زیست و سایر حوزه‌ها مورد استفاده قرار گرفته است (Sun et al., 2020). از طرفی، طبقه‌بندی تصاویر سنجش از دور همواره یکی از موضوعات پژوهشی مورد توجه در کاربرد فناوری سنجش از دور می‌باشد، که به استفاده از

داده‌های سنجش از دور برای تهیه نقشه‌های کاربری یا پوشش زمین اشاره دارد (Tuia et al., 2011). اما، ویژگی‌های پویا و متغیر اراضی کشاورزی باعث می‌شوند که در تصاویر ماهواره‌ای، مزارع مربوط به محصولات مختلف در برخی از زمان‌های خاص تقویم زراعی، علی‌رغم تفاوت در نوع محصول، شباهت‌های طیفی از خود نشان دهند. از سوی دیگر، حتی زمین‌های کشت یک محصول خاص مانند گندم نیز ممکن است به‌دلیل عواملی مانند روش آبیاری (آبی یا دیم)، نوع رقم کشت‌شده، زمان‌های مختلف کاشت و شرایط متفاوت خاک و آبیاری، در تصاویر ماهواره‌ای به شکل‌های گوناگونی ظاهر شوند. این تنوع و پیچیدگی، تحلیل و طبقه‌بندی پوشش اراضی کشاورزی را به چالشی مهم تبدیل کرده‌است (Gili et al., 2022).

با پیشرفت فناوری، ادغام فناوری‌های سنجش از دور با الگوریتم‌های یادگیری ماشین و هوش مصنوعی توانایی آن‌ها را در تحلیل مجموعه داده‌های بزرگ و پیچیده بهبود بخشیده است. این امر به پیش‌بینی دقیق‌تر در مورد بازده محصولات و نیاز آبی کمک کرده است (Sarker, 2021). از طرف دیگر، استفاده از قابلیت یادگیری از داده‌ها، الگوریتم‌های یادگیری عمیق را برای انجام وظایف طبقه‌بندی محصولات مناسب کرده و راه‌حلی مناسب برای کشاورزی مدرن ارائه می‌دهد که در این بین، روش‌های نظارت‌شده طبقه‌بندی، که بر داده‌های چندزمانه و شاخص‌های پوشش گیاهی تکیه دارند، نتایج قابل‌توجهی را در تفکیک انواع محصولات کشاورزی ارائه کرده‌اند (Khosravi, 2024; Yadav et al., 2024).

الگوریتم‌های یادگیری ماشین به‌طور گسترده در نقشه‌برداری محصولات کشاورزی مورد استفاده قرار گرفته‌اند. از جمله روش‌های متداول می‌توان به الگوریتم‌های جنگل تصادفی^۱ (Phalke et al., 2020; d'Andrimont et al., 2021)، شبکه‌های عصبی^۲ (Yaramasu et al., 2020)، ماشین‌های بردار پشتیبان^۳ (Zheng et al., 2015) و تقویت گرادیان شدید^۴ (Chen and Guestrin, 2016) اشاره کرد. این الگوریتم‌ها با یادگیری ویژگی‌های متمایز کلاس‌های هدف از نمونه‌های آموزشی، داده‌های جدید را طبقه‌بندی می‌کنند.

از جمله پژوهش‌های انجام شده در زمینه طبقه‌بندی تصاویر در جهان و ایران، Saei Jamalabad et al. (2018)، با تولید ویژگی‌های جدید از جمله گرادیان طیفی باندها و شاخص‌های گیاهی تصاویر چندزمانی و اختلاف آن‌ها و انتخاب ویژگی‌های بهینه، عملکرد مدل جنگل تصادفی را بهبود یافت، به‌طوری‌که،

¹ Random Forest (RF)

² Neural Network

³ Support Vector Machine (SVM)

⁴ Extreme Gradient Boosting (XGBOOST)

هم‌چنین، در پژوهش Salehi Shafa et al. (2022)، از مدل شبیه‌سازی-بهینه‌سازی با استفاده از GMS، شبکه عصبی مصنوعی (ANN) و الگوریتم ژنتیک چندهدف (NSGA-II) برای بهینه‌سازی بهره‌برداری از آبخوان دشت شهریار توسعه داده شد. در این پژوهش با استفاده از هوش مصنوعی نشان داد که برنامه‌ریزی برای الگوی کشت و کاهش مصرف آب کشاورزی، منجر به افزایش قابل توجه تراز و بیلان آب زیرزمینی شده است. در هند، Padma VVL et al. (2024) به بررسی کارایی الگوریتم‌های ماشین بردار پشتیبان، نقشه‌برداری زاویه طیفی^۵ و جنگل تصادفی در طبقه‌بندی محصولات با استفاده از داده‌های چندطیفی Planet Scope پرداختند. نتایج نشان داد که ماشین بردار پشتیبان با مرزگذاری دقیق، نقشه‌برداری زاویه طیفی با رویکرد مشابهت طیفی و جنگل تصادفی با یادگیری ترکیبی به‌طور مؤثر قادر به تمایز محصولات هستند. در ایران، Asadi and Shamsoddini (2024)، به بررسی نقش باندهای لبه قرمز در تفکیک محصولات زراعی با استفاده از ترکیب تصاویر اپتیک و رادار پرداخته و نشان داد که جنگل تصادفی بهترین عملکرد را در طبقه‌بندی داشت. در ترکیه نیز Simsek, (2025) مقایسه‌ای از عملکرد الگوریتم‌های جنگل تصادفی، شبکه عصبی مصنوعی، ماشین بردار پشتیبان و تقویت گرادیان شدید در طبقه‌بندی چندزمانه محصولات کشاورزی با استفاده از تصاویر سنتینل ۲ در ناحیه چوکورووا انجام داده است. نتایج نشان داد که الگوریتم تقویت گرادیان شدید با صحت کلی ۹۲/۱۴ درصد بهترین عملکرد را داشته و مشکلات طبقه‌بندی برای محصولات با مراحل فنولوژیکی مشابه، باعث کاهش صحت شد.

با توجه به اقلیم نیمه‌خشک با بارش کم، تأخیر بالا و تغییرات دمایی قابل توجه استان قزوین، این استان همواره با چالش‌هایی در کشاورزی فاریاب مواجه می‌باشد. طبقه‌بندی محصولات زراعی در چنین شرایطی اهمیت زیادی دارد زیرا به مدیریت بهینه منابع آبی، پیش‌بینی عملکرد محصولات و بهینه‌سازی استفاده از منابع کمک می‌کند. هم‌چنین، این طبقه‌بندی با امکان برنامه‌ریزی بهتر برای تغییرات اقلیمی مانند جابجایی الگوهای بارندگی و دما، باعث افزایش کارایی استفاده از زمین و کاهش اثرات محیط‌زیستی می‌شود. علاوه بر این، طبقه‌بندی محصولات زراعی می‌تواند مدیریت آفات و بیماری‌ها را بهبود بخشد و پایداری محصولات کشاورزی از جمله گندم و ذرت که در دشت قزوین رایج هستند را تضمین کند. هدف پژوهش پیش‌رو استفاده از الگوریتم‌های

صحت کلی جداسازی گندم و جو توسط روش جنگل تصادفی به‌طور میانگین تا ۱۴/۵ درصد و ضریب کاپای آن، ۵/۲ درصد افزایش یافت. در مطالعه‌ای Viskovic et al. (2019)، به بررسی استفاده از تصاویر چندطیفی و چندزمانه ماهواره‌ای برای طبقه‌بندی محصولات کشاورزی با روش‌های یادگیری ماشین پرداخته است. نتایج نشان داد که الگوریتم جنگل تصادفی با صحت ۸۴/۲۰ درصد و ضریب کاپا ۰/۸۱، عملکرد بهتری نسبت به سایر الگوریتم‌ها داشت.

Pourhasan et al. (2021) با استفاده از روش طبقه‌بندی یادگیری عمیق برای تفکیک محصولات با استفاده از تصاویر سری زمانی لندست ۸ و شاخص NDVI^۱ ارائه دادند که صحت بالاتری نسبت به سایر روش‌های مرسوم مانند جنگل تصادفی و ماشین بردار پشتیبان داشت. این روش می‌تواند در برنامه‌ریزی‌های کلان برای مدیریت اراضی و برآورد نیازهای کشاورزی کشور مفید باشد. در پژوهشی که در حوزه نازل‌چای طی بازه‌های زمانی ۲۶ ساله با استفاده از تصاویر ماهواره‌ای و روش طبقه‌بندی نظارت‌شده "احتمال بیشینه"^۲ انجام شد، صحت کلی نقشه‌های تولیدشده به‌ترتیب برای سال‌های ۱۳۷۱، ۱۳۹۲ و ۱۳۹۷ برابر با ۹۴/۱۴ درصد، ۹۷/۰۴ درصد و ۹۲/۷۸ درصد و ضرایب کاپا نیز به‌ترتیب ۰/۸۶، ۰/۹۲ درصد و ۰/۹۰ درصد بود که دقت مناسب نتایج را تأیید می‌کند. این یافته‌ها بیانگر تغییرات قابل توجه کاربری اراضی در این زیرحوضه و افزایش چشمگیر باغات پس از سال ۱۳۹۲ هستند که ضرورت بازنگری در مدیریت منابع آب و خاک را نشان می‌دهد (Esmailnezhad and Zeinalzadeh, 2020).

Moumni and Lahrouni. (2021) به بررسی طبقه‌بندی محصولات کشاورزی در منطقه نیمه‌خشک دشت حوض مراکش با استفاده از ترکیب تصاویر ماهواره‌ای سنتینل ۱ و ۲ پرداخته است. نتایج نشان داد که ترکیب داده‌های باند C و طیفی بهبود قابل توجهی در صحت طبقه‌بندی محصولات (صحت کلی ۸۹ درصد و ضریب کاپا ۰/۸۵) نسبت به داده‌های منفرد ایجاد کرده است. هم‌چنین، Asgari and Hasanlou (2023) از شش شاخص طیفی در سه روش یادگیری ماشینی نظارت‌شده، از جمله جنگل تصادفی، گرادیان تقویت‌شده^۳ و نزدیک‌ترین همسایه^۴ برای نقشه‌برداری محصول استفاده کردند. نتایج طبقه‌بندی از طریق نمونه‌های آزمایشی که صحت کلی رضایت‌بخشی را برای غالب‌ترین انواع محصول در مزارع مختلف نشان داد.

^۱ Normanized Difference Vegetation Index

^۲ Maximum Likelihood

^۳ Gradient Boost

^۴ K-Nearest Neighborhood

^۵ Spectral Angle Mapper

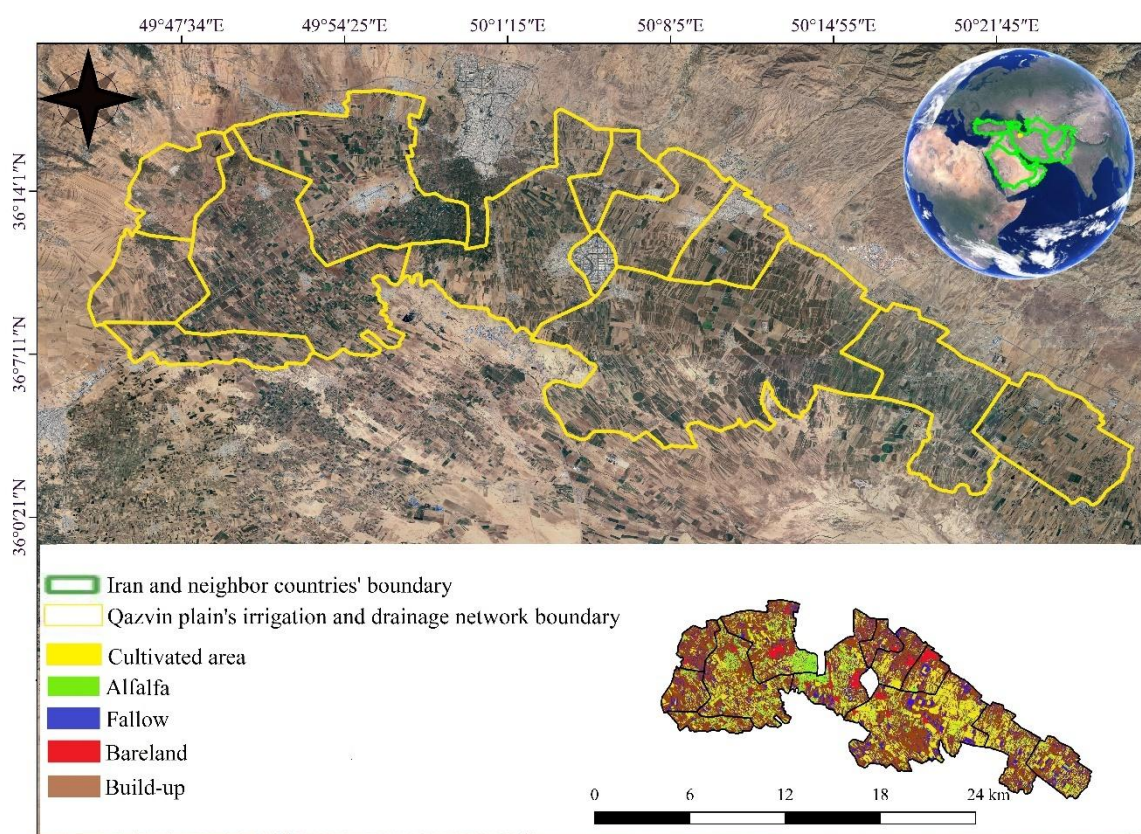
به صرفه برای روش های سنتی و میدانی پیشنهاد می دهد. ارائه مسیرهای پژوهشی آینده، شامل بهره گیری از الگوریتم های پیشرفته یادگیری عمیق و داده های مکمل چندزمانه، نیز افق های جدیدی را برای بهبود صحت و کاربردهای این حوزه باز می کند.

۲- مواد و روش ها

۲-۱- منطقه مورد مطالعه

شبکه آبیاری دشت قزوین با وسعتی معادل ۸۰۰۰۰ هکتار ناخالص به صورت نواری است به طول ۹۴ کیلومتر که از محدوده زیاران شروع شده و به اراضی کهک تاکستان منتهی می شود (شکل ۱). مساحت خالص اراضی تحت پوشش شبکه آبیاری طالقان (قزوین) ۶۰۰۰۰ هکتار است که حدود ۷۲ درصد آن را اراضی درجه ۱ و ۲ تشکیل می دهد. الگوی کلی کشت منطقه ۵۰ درصد کشت پائیزه است، که بیش تر به محصول گندم و جو اختصاص دارد، ۱۵ درصد کشت بهاره و ۳۵ درصد آیش در نظر گرفته می شود (Darvish Hendi and Amiri Tokaldany, 2024).

یادگیری ماشین در طبقه بندی تصاویر ماهواره ای برای دوره های کشت بهاره و پاییزه و بررسی صحت سه روش منتخب ماشین بردار پشتیبان، جنگل تصادفی و تقویت گرادیان شدید در منطقه شبکه آبیاری قزوین می باشد. نوآوری این پژوهش در چندین جنبه کلیدی خلاصه می شود. ترکیب داده های اپتیکی سنتینل ۲ و راداری سنتینل ۱ به عنوان رویکردی نوآورانه، توانسته قدرت تفکیک پذیری طیفی و زمانی را افزایش داده و بخشی از محدودیت های استفاده از داده های تک منبعی را برطرف کند. همچنین، این پژوهش با تمرکز بر منطقه ای با اقلیم نیمه خشک و تنوع بالای پوشش زمین، چارچوبی کاربردی برای تحلیل و مدیریت منابع کشاورزی در مناطق مشابه ارائه داده است. استفاده از شاخص های طیفی، در کنار داده های چندزمانه و آزمون جفریس-ماتوسیتا برای ارزیابی تفکیک پذیری طیفی، صحت و عمق تحلیل را افزایش داده و رویکردی کم تر بررسی شده در پژوهش های مشابه را ارائه کرده است. علاوه بر این، با نشان دادن کاربرد ترکیب فناوری های سنجش از دور و یادگیری ماشین در مدیریت منابع کشاورزی، این پژوهش جایگزینی عملی و مقرون



شکل ۱- موقعیت جغرافیایی شبکه آبیاری و زهکشی دشت قزوین و کاربری اراضی آن
Figure 1- Geographical location of Qazvin Plain irrigation and drainage network, and its land use

۲-۲- داده‌های مورد استفاده و پردازش آن‌ها

در مطالعه کنونی، به منظور تهیه نقشه طبقه‌بندی تصاویر، از ترکیب تصاویر اپتیک سنتینل ۲ و تصاویر راداری سنتینل ۱ در سال ۲۰۲۳-۲۰۲۴ استفاده شد. با توسعه فناوری ماهواره‌ای، در سال ۲۰۱۴ اولین ماهواره از سری ماهواره‌های سنتینل به فضا پرتاب شد. این سری شامل سنتینل ۱ (راداری) و سنتینل ۲ (چندطیفی) است که به دلیل قدرت تفکیک مکانی بالا (۱۰ متر) و تناوب تصویربرداری مناسب (۵ روز برای سنتینل ۲ و ۱۲ روز برای سنتینل ۱)، به همراه رایگان بودن داده‌ها، به محبوبیت زیادی در مطالعات کشاورزی دست یافته‌اند. این ماهواره‌ها در زمینه‌های مختلفی از جمله نقشه‌برداری محصولات کشاورزی و ارزیابی تغییرات کاربری اراضی در سراسر جهان کاربرد دارند. با توجه به ویژگی ماهواره سنتینل ۲، باندهای طیفی در محدوده مرئی، مادون قرمز نزدیک و مادون قرمز میانی با توان تفکیک مکانی‌های متفاوت ۱۰، ۲۰ و ۶۰ متری انتخاب شدند که شامل باندهای ۲ تا ۷ و باندهای ۸، ۸A و باندهای ۱۱ و ۱۲ می‌شدند. لازم به ذکر است که تصحیحات لازم برای تصاویر این سنجنده در محیط گوگل ارث انجین صورت گرفت و در نهایت سه شاخص طیفی NDVI^۲، SAVI^۳ و LAI^۴ محاسبه و روی تصاویر به صورت یک لایه مجزا، روی تمامی پیکسل‌های تصاویر در دو دوره کشت بهاره و پاییزه اعمال شد. ترکیب شاخص‌های گیاهی با روش‌های طبقه‌بندی تصاویر، مانند استفاده از تصاویر پهپادها یا داده‌های چندطیفی، صحت تحلیل را افزایش داده و کاربردهای پیشرفته‌ای همچون تخمین زیست‌توده و پیش‌بینی محصول را ممکن می‌کند. سپس تصاویر چند فصلی GRD سنتینل ۱ در گوگل ارث انجین فراخوانی شد، این داده‌ها برای محدوده مکانی و زمانی مشابه سنتینل ۲ فیلتر شده و سپس، پلاریزاسیون VV که نشان‌دهنده بازتاب سیگنال‌های راداری است، انتخاب شد. به منظور کاهش نویز اسپکل در تصاویر راداری، فیلتر حداکثر اعمال شد. که برای حذف داده‌های پرت و ایجاد تصویری پایدارتر به کارگرفته تنها اطلاعات مرتبط پردازش شوند و صحت تحلیل افزایش یابد (Asadi and rasti et al., 2022; Shamsoddini, 2024). جدول ۱، نشان دهنده ویژگی‌های تصاویر سنتینل ۲ و ۱ است که در این پژوهش مورد استفاده قرار گرفته‌است. در این مطالعه، پنج کلاس اصلی محصولات زراعی، شامل گندم، یونجه، مناطق شهری، اراضی بایر و آیش در دوره کشت پاییزه و محصولات ذرت، یونجه، مناطق شهری، اراضی بایر و آیش برای دو دوره کشت بهاره تعریف شدند. این کلاس‌ها نمایانگر تنوع پوشش

زمین در دشت قزوین بوده و به منظور تحلیل دقیق‌تر، برای هر کلاس نمونه‌های مشخصی براساس نقاط گزارش شده توسط کشاورزان در سامانه گوگل ارث انجین، استخراج شدند. تعداد نمونه‌های هر کلاس در دوره کشت پاییزه به ترتیب عبارت‌اند از: گندم ۲۸۱ نمونه، یونجه ۲۹۹ نمونه، اراضی بایر ۱۴۸ نمونه، مناطق شهری ۱۷۸ نمونه و آیش ۱۰۸ نمونه و در دوره کشت بهاره به ترتیب: ذرت ۸۳۰ نمونه، یونجه ۹۷۷ نمونه، اراضی بایر ۱۵۲ نمونه، مناطق شهری ۱۷۸ نمونه و آیش ۱۰۸ نمونه. در مجموع، ۱۰۱۴ نمونه برای دوره کشت پاییزه و ۲۲۵۴ نمونه برای دوره کشت بهاره در مراحل آموزش و تست مدل جمع‌آوری و مورد استفاده قرار گرفت. لازم به ذکر است که در این پژوهش، استفاده از تصاویر راداری به تمایز میان دو کلاس اراضی بایر و آیش کمک به سزایی کرده است. چراکه، مناطق بایر به طور معمول مناطقی هستند که به طور کلی خالی از هرگونه پوشش گیاهی هستند. اما مناطق آیش می‌توانند به صورت پراکنده یک پوشش گیاهی داشته باشند. به دلیل استفاده از تصاویر راداری حتی نوع خاک هم در میزان بازتابش امواج مؤثر است. و این تفاوت در نوع خاک در مناطق آیش و بایر می‌تواند باعث تمایز میان این دو کلاس شود. برای توزیع نمونه‌ها، از تقسیم‌بندی استاندارد ۷۰ درصد برای آموزش و ۳۰ درصد برای تست استفاده شده است. این توزیع متوازن به کاهش احتمال بیش‌برازش کمک می‌کند و نتایج حاصل از مدل را قابل اطمینان‌تر می‌سازد (Gholamy et al., 2018). انتخاب این درصدها بر اساس استانداردهای علمی در مسائل طبقه‌بندی نظارت‌شده انجام شده است. استفاده از این روش موجب می‌شود که مدل نه تنها بتواند اطلاعات کافی برای یادگیری داشته باشد، بلکه داده‌های مستقل کافی نیز برای ارزیابی عملکرد آن در اختیار باشد. این رویکرد به ایجاد تعادلی میان صحت مدل و قابلیت تعمیم آن کمک کرده و در نهایت، نتایجی دقیق‌تر و معتبرتر ارائه می‌دهد.

۲-۲-۱- داده‌های زمینی

با توجه به اهمیت انتخاب دقیق داده‌های مختصاتی کلاس‌های مورد بررسی در بخش ورودی داده‌ها برای آموزش و ارزیابی الگوریتم‌های طبقه‌بندی، داده‌های مربوط به تاریخ کشت و مختصات زمین‌ها از کشاورزان و کشت و صنعت منطقه در سال ۱۴۰۳ جمع‌آوری شد. در نهایت، ۷۰ درصد از این داده‌ها برای آموزش الگوریتم‌ها و ۳۰ درصد باقیمانده برای ارزیابی عملکرد الگوریتم‌ها استفاده شدند. ارزیابی عملکرد الگوریتم‌ها با استفاده از ماتریس خطا، ضریب کاپا و صحت کلی، انجام شد.

¹ COPERNICUS/S1_GRD

² Normalize Difference Vegetation Index

³ Soil Adjusted Vegetation Index

⁴ Leaf Area Index

جدول 1- مشخصات طیفی و مکانی داده‌های سنجنده سنتینل 2 و 1

Table 1- Spectral and Spatial Characteristics of the Sentinel-2 and Sentinel-1 Data

Satellite	Bands' Name	Bands' Number	Wavelength (nm)	Pixel Size (m)
COPERNICUS/Sentinel-2	Blue	2	492.4	10
	Green	3	559.8	10
	Red	4	664.6	10
	Red Edge 1	5	704.1	20
	Red Edge 2	6	740.5	20
	Red Edge 3	7	782.8	20
	NIR	8	832.8	10
	Red Edge 4	8A	864.8	20
	SWIR 1	11	1613.7	20
	SWIR 2	12	2202.4	20
COPERNICUS/Sentinel-1	VV	-	-	10

۲-۳-۲- ماشین بردار پشتیبان^۱

ماشین بردار پشتیبان یکی از تکنیک‌های پرکاربرد در حل مسائل طبقه‌بندی و رگرسیون است. در الگوریتم ماشین بردار پشتیبان، مسئله طبقه‌بندی با استفاده از مفهوم حاشیه^۲ حل می‌شود. حاشیه به عنوان کم‌ترین فاصله بین مرز تصمیم‌گیری^۳ و نزدیک‌ترین نمونه‌های داده تعریف می‌شود. در این روش، مرز تصمیم‌گیری به گونه‌ای انتخاب می‌شود که حاشیه بیشینه گردد (Neetu and Ray, 2019).

بنابر نتایج بررسی‌های اولیه انجام شده، با توجه به صحت بالاتر کرنل RBF، نسب به کرنل‌های دیگر در این پژوهش، نوع کرنل مورد استفاده RBF بوده و مقادیر گاما^۴ برابر با 0.0002 تنظیم شده است (Alejo et al., 2006). گاما تأثیر مستقیمی بر انعطاف‌پذیری مرز تصمیم‌گیری دارد؛ مقادیر پایین‌تر گاما معمولاً به مرزهای ساده‌تر و قابل تعمیم‌تر منجر می‌شود، درحالی‌که مقادیر بالاتر گاما ممکن است پیچیدگی بیش‌ازحد ایجاد کنند. افزایش حاشیه در این روش منجر به تغییرات خاصی در انتخاب مرز تصمیم‌گیری می‌شود. این ویژگی به ماشین بردار پشتیبان امکان می‌دهد تا با داده‌های پیچیده و چندبعدی به خوبی عمل کند (Ghodsi et al., 2021).

۲-۳-۳- الگوریتم تقویت گرادیان شدید^۵

الگوریتم تقویت گرادیان شدید به عنوان یکی از پیشرفته‌ترین و کارآمدترین روش‌های یادگیری ماشین، به ویژه در تحلیل داده‌های پیچیده و پرجمع شناخته می‌شود. این الگوریتم که مبتنی بر تقویت گرادیان است، با ترکیب مدل‌های ساده‌ای مانند درخت‌های تصمیم، پیش‌بینی‌های دقیقی ارائه می‌دهد

۲-۳-۳- طبقه‌بندی تصاویر با استفاده از روش یادگیری ماشین

۲-۳-۱- روش طبقه‌بندی جنگل تصادفی^۶

جنگل تصادفی یکی از الگوریتم‌های پرکاربرد و قدرتمند در حوزه یادگیری ماشین است که در گروه یادگیری نظارت‌شده^۷ قرار دارد (Breiman, 2001; Goldblatt et al., 2016). یکی از مزایای کلیدی الگوریتم جنگل تصادفی، توانایی آن در کاهش بیش‌برازش^۸ است. تصادفی‌سازی در انتخاب داده‌ها و ویژگی‌ها باعث می‌شود مدل به داده‌های خاص وابسته نباشد و عملکرد بهتری روی داده‌های جدید نشان دهد. همچنین، با ترکیب پیش‌بینی درخت‌های متعدد، الگوریتم تأثیر نویز^۹ در داده‌ها را کاهش داده و پیش‌بینی‌هایی پایدار و قابل اطمینان ارائه می‌کند. این الگوریتم نیز قادر به مدیریت داده‌های پیچیده و غیرخطی بوده و می‌تواند با تعداد زیادی ویژگی و حتی داده‌های ناقص یا ناکامل کار کند. از دیگر مزایای جنگل تصادفی می‌توان به سادگی در پیاده‌سازی و توانایی آن در کاربردهای گسترده نظیر طبقه‌بندی و رگرسیون اشاره کرد (Neetu and Ray, 2019). این الگوریتم در کاربردهایی نظیر شناسایی الگوها (مانند شناسایی تصاویر)، تخمین مقادیر عددی (مانند عملکرد محصولات کشاورزی) و مسائل علمی مختلف به طور گسترده استفاده می‌شود. توانایی آن در ارائه پیش‌بینی‌های دقیق و تعمیم‌پذیر، همراه با مقاومت در برابر نویز و داده‌های ناقص، باعث شده که به یکی از پرکاربردترین الگوریتم‌ها در یادگیری ماشین تبدیل شود (Ramezani, 2024).

¹ Random Forest

² Supervised Learning

³ Overfitting

⁴ Noise

⁵ SVM

⁶ Margin

⁷ Decision Boundary

⁸ Gamma

⁹ XGBOOST

۲-۴-۱- ماتریس خطا^۱

ماتریس خطا یک ابزار ارزیابی مهم در طبقه‌بندی‌های آماری و یادگیری ماشین است که عملکرد مدل را با بررسی تعداد پیش‌بینی‌های صحیح و نادرست ارزیابی می‌کند (رابطه ۲). این ماتریس به‌ویژه برای ارزیابی مشکلات طبقه‌بندی، به‌خصوص در مواردی که کلاس‌های نابرابر یا کلاس‌های فرعی مختلف وجود دارند، مفید است (Shao et al., 2024).

$$\text{Confusion Matrix} = \begin{bmatrix} \text{False Positive (FP)} & \text{True Positive (TP)} \\ \text{True Negative (TN)} & \text{False Negative (FN)} \end{bmatrix} \quad (2)$$

یک ماتریس خطا دارای چهار جزء اصلی است که عبارتند از: True Positive (TP): تعداد نمونه‌هایی که به‌درستی در کلاس مربوطه پیش‌بینی شده‌اند، False Positive (FP): تعداد نمونه‌هایی که به نادرست در کلاس دیگری پیش‌بینی شده‌اند، False Negative (FN): تعداد نمونه‌هایی که به نادرست در کلاس مربوطه پیش‌بینی نشده‌اند، True Negative (TN): تعداد نمونه‌هایی که به‌درستی در کلاس نادرست پیش‌بینی شده‌اند.

۲-۴-۲- ضریب کاپا^۲

میزان توافق بین پیش‌بینی‌ها و مقادیر واقعی را با در نظر گرفتن پیش‌بینی‌های تصادفی محاسبه می‌کند (Ghodsi et al., 2021; Zhan et al., 2024).

$$\text{Kappa} = \frac{P_0 - P_e}{1 - P_e} \quad (3)$$

طبق رابطه ۳، در آن: P_0 برابر است با تعداد نمونه‌های به‌درستی طبقه‌بندی‌شده تقسیم بر کل تعداد نمونه‌ها و P_e برابر است با مجموع حاصل ضرب‌های تعداد واقعی نمونه‌ها و تعداد پیش‌بینی‌شده در هر دسته، تقسیم بر مجذور کل تعداد نمونه‌ها (Zhan et al., 2024).

(Sahin, 2020). ویژگی‌های این الگوریتم مانند سرعت، مقیاس‌پذیری، و توانایی پردازش داده‌های پرحجم، آن را به زینه‌ای ایده‌آل برای کاربردهای متنوع از جمله تحلیل‌های لحظه‌ای و پیش‌بینی‌های دقیق تبدیل کرده است (Chen et al., 2004).

تقویت گرادیان شدید از تکنیک تقویت گرادیان بهره می‌گیرد که به‌صورت افزایشی مدل‌های ضعیف را ترکیب می‌کند. این فرآیند باعث می‌شود خطای مدل به تدریج کاهش یابد، در حالی که جنگل تصادفی تنها با ترکیب مدل‌های مستقل به‌صورت موازی عمل می‌کند. پژوهش‌ها نشان داده‌اند که تقویت گرادیان شدید معمولاً صحت بیش‌تری نسبت به جنگل تصادفی در مسائل طبقه‌بندی و رگرسیون دارد، چراکه بهینه‌سازی در تقویت گرادیان شدید به‌صورت مرحله‌ای و با استفاده از گرادیان تابع زیان انجام می‌شود (Friedman, 2001).

۲-۴-۳- ارزیابی صحت روش‌های طبقه‌بندی

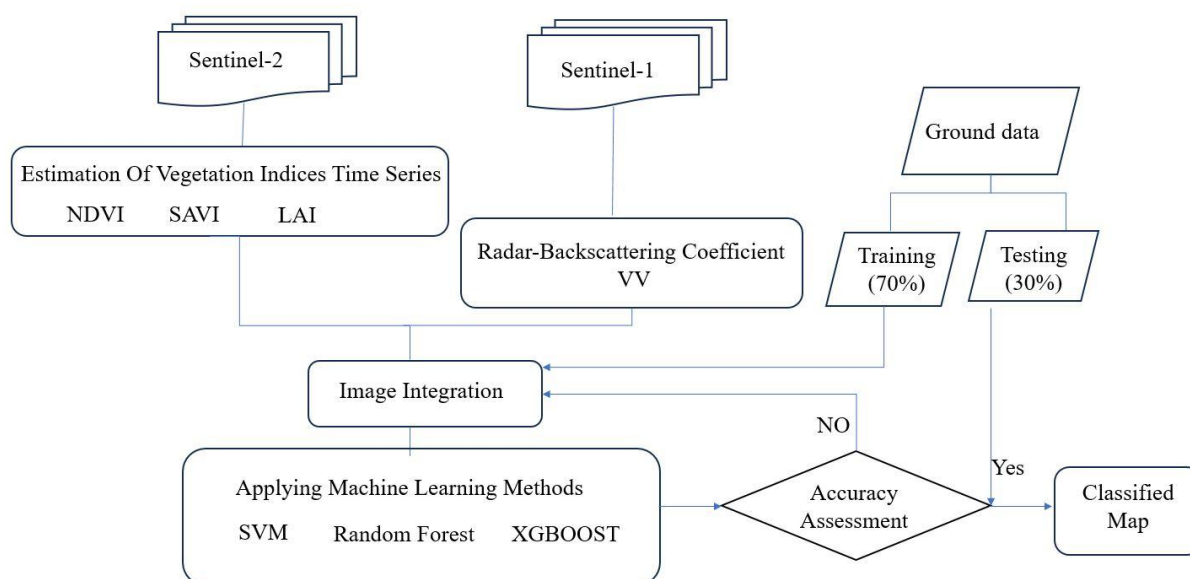
صحت کلی یک معیار ارزیابی است که برای بررسی صحت نتایج طبقه‌بندی به‌کار می‌رود. این شاخص نمایانگر درصد کل داده‌هایی است که به‌درستی طبقه‌بندی شده‌اند و به‌طور کلی از طریق ماتریس خطا محاسبه می‌شود. صحت کلی به‌صورت رابطه ۱، تعریف می‌شود (Li et al., 2014).

$$\text{Overall Accuracy} = \frac{N_{\text{Correct}}}{N_{\text{Total}}} \quad (1)$$

N_{Correct} : تعداد نمونه‌های درستی است که به‌درستی طبقه‌بندی شده‌اند، N_{Total} : کل تعداد نمونه‌ها است.

¹ Confusion Matrix

² Kappa coefficient



شکل ۲- نمودار گردش مراحل مختلف طبقه بندی به منظور تفکیک محصولات زراعی

Figure 2- Flowchart of the classification steps for crop mapping

۳- نتایج و بحث

۳-۱- آزمون تفکیک پذیری

نتایج آزمون تفکیک پذیری جفریس-ماتوسیتا (JM)^۱ برای ارزیابی قابلیت تفکیک طیفی بین شش کلاس کاربری اراضی شامل گندم-جو، یونجه، آیش، مناطق شهری و اراضی بایر در فصل پاییز و شش کلاس کاربری اراضی شامل ذرت، یونجه، آیش، مناطق شهری و اراضی بایر در فصل بهار مورد استفاده قرار گرفت. این آزمون با دامنه عددی بین صفر (عدم جداسازی) و ۲ (جداسازی کامل) نشان داد که برخی کلاس ها دارای جداسازی طیفی کامل بوده و برخی دیگر به دلیل هم پوشانی طیفی، جداسازی پایینی دارند. برای مثال، در فصل پاییز، کلاس های گندم-جو و یونجه با مقادیر JM نزدیک به ۰/۶۵ نشان دهنده هم پوشانی طیفی قابل توجهی هستند، در حالی که کلاس های آیش و اراضی بایر با مقادیر JM بالاتر (بین ۰/۴۸ تا ۱/۰۸) نشان دهنده جداسازی بهتری هستند. کلاس های مناطق شهری و اراضی بایر نیز با مقادیر JM نزدیک به ۲، جداسازی کامل را نشان می دهند. در فصل بهار، کلاس های ذرت و یونجه با مقادیر JM نزدیک به ۰/۷ نشان دهنده هم پوشانی طیفی هستند،

در حالی که کلاس های آیش و اراضی بایر با مقادیر JM بین ۱/۱۲ تا ۱/۵۴ جداسازی بهتری دارند. کلاس های مناطق شهری و اراضی بایر نیز با مقادیر JM نزدیک به ۲، جداسازی کامل را نشان می دهند. تعیین حد آستانه برای شاخص های طیفی و استفاده تصاویر راداری می تواند در شناسایی و تمایز میان کلاس های مختلف، به خصوص مناطق آیش و بایر نقش به سزایی ایفا می کند (Wuyun et al., 2022). مقالات بسیاری از این روش به منظور تفکیک کلاس های مورد بررسی بهره برده اند (Huang et al., 2018; Wang et al., 2018; Sen et al., 2019).

به طور کلی، نتایج این آزمون نشان داد که برخی کلاس ها مانند مناطق شهری و اراضی بایر به دلیل تفاوت های طیفی واضح، به راحتی قابل تفکیک هستند، در حالی که کلاس هایی مانند گندم-جو و یونجه یا ذرت و یونجه به دلیل هم پوشانی طیفی، نیاز به روش های دقیق تری برای تفکیک دارند. جدول ۲، نتایج آزمون تفکیک پذیری جفریس-ماتوسیتا را برای هر کلاس در دو دوره کشت بهار و پاییزه نشان می دهد.

^۱ Jeffries-Matusita

جدول ۲- نتایج آزمون تفکیک‌پذیری جفریس-ماتوسیتا

Table 2- Results of the Jeffries-Matusita Separability Test

		Wheat-barley	Alfalfa	Fallow	Bare_land	Build_up
Autumn Cultivation	Wheat-barley	0	0.65	2	2	2
	Alfalfa	0.65	0	2	2	2
	Fallow	2	0.39	0	0.48	1.08
	Bare_land	2	2	0.48	2	2
	Build_up	2	2	1.08	2	0
		Maize	Alfalfa	Fallow	Bare_land	Build_up
Spring Cultivation	Maize	0	0.71	1.12	1.54	1.92
	Alfalfa	0.71	0	2	2	2
	Fallow	1.12	2	0	2	1.42
	Bare_land	1.54	2	2	0	0.41
	Build_up	2	2	1.42	0.41	0

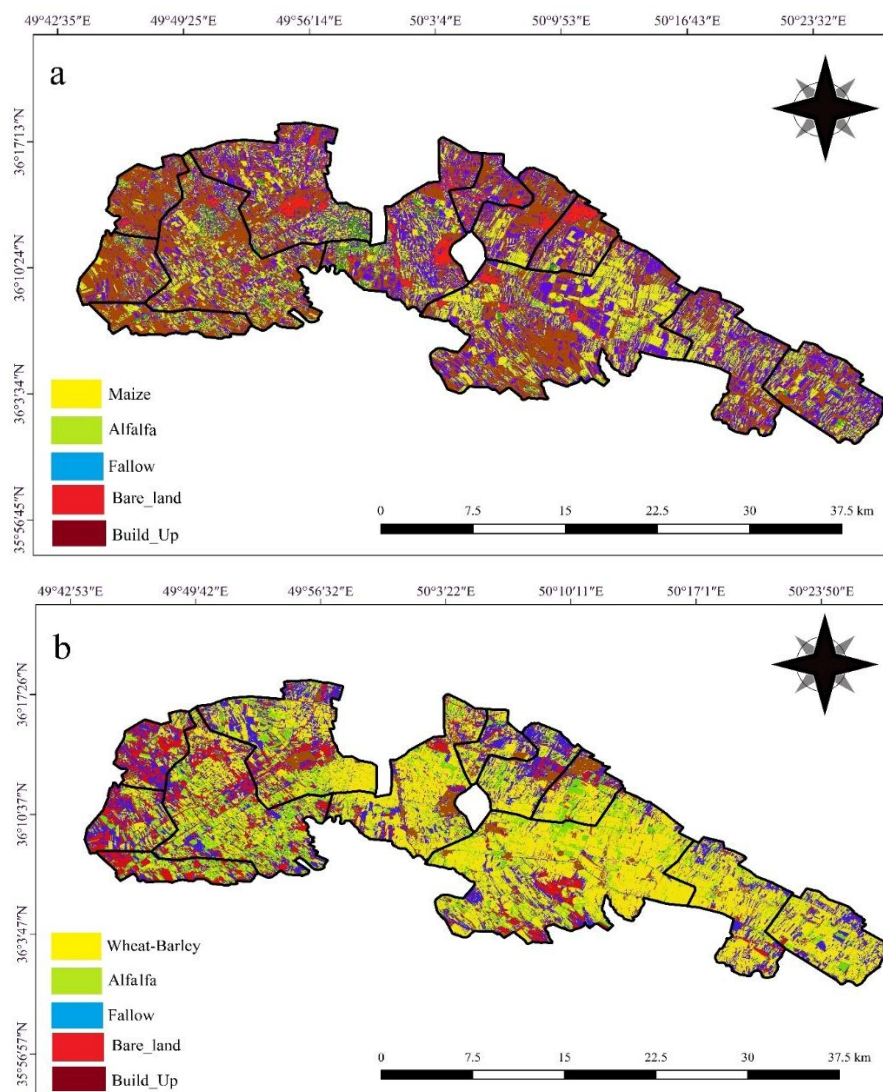
۳-۲- مقایسه عملکرد روش‌های طبقه‌بندی

۳-۲-۱- عملکرد ماشین بردار پشتیبان

نتایج حاصل از ماتریس خطا و شاخص‌های ارزیابی مانند ضریب کاپا و صحت کلی، نشان‌دهنده عملکرد نسبتاً مطلوب ماشین بردار پشتیبان در طبقه‌بندی کلاس‌های مختلف کشاورزی است. با این حال، این روش در جداسازی کلاس‌هایی با هم‌پوشانی طیفی بالا، مانند گندم-جو و یونجه، با چالش‌هایی مواجه است. در بررسی طبقه‌بندی تصاویر برای کلاس‌های پاییزه در شکل ۶ از ۲۴۰ نمونه کلاس گندم و جو، ۲۰۹ نمونه به‌درستی طبقه‌بندی شده است، درحالی‌که ۲۰ نمونه به اشتباه به‌عنوان کلاس یونجه و ۱۱ نمونه در کلاس آیش طبقه‌بندی شده‌اند. این موضوع نشان‌دهنده آن است که ماشین بردار پشتیبان در تشخیص کلاس‌هایی با ویژگی‌های طیفی مشابه، صحت کم‌تری دارد. از طرفی، برای ۲۹۹ نمونه کلاس یونجه، ۲۸۰ نمونه به‌درستی شناسایی شده‌اند که نشان‌دهنده عملکرد نسبتاً خوب این روش در تشخیص این کلاس است.

در مورد کلاس‌های مناطق آیش، بایر و شهری، ماشین بردار پشتیبان عملکرد بهتری نشان داده است. از ۱۰۸ نمونه کلاس مناطق آیش، ۱۰۳ نمونه به‌درستی طبقه‌بندی شده‌اند و ۵ نمونه به اشتباه در کلاس گندم و جو قرار گرفته‌اند. این نتایج نشان‌دهنده آن است که ماشین بردار پشتیبان در تشخیص کلاس‌هایی با ویژگی‌های طیفی مجزا، عملکرد قابل‌قبولی دارد. برای کلاس‌های مناطق بایر و شهری به‌ترتیب ۱۷۷ نمونه از ۱۷۸ و ۱۳۷ نمونه از ۱۴۸ نمونه در کلاس درست قرار گرفته است. طبق جدول ۳، ضریب کاپای ۰/۹۱ نشان‌دهنده توانایی نسبی این روش در جداسازی کلاس‌ها با هم‌پوشانی است. صحت کلی ۹۳/۱۱ درصد نشان می‌دهد که ماشین بردار پشتیبان عملکرد قابل‌قبولی دارد، اما با وجود کاپای بالا، در جداسازی کلاس‌های هم‌پوشانی ضعیف عمل می‌کند.

در جدول ۳، در بررسی صحت این روش برای دوره کشت بهاره، ماشین بردار پشتیبان با صحت کلی ۸۳/۷۹ درصد و ضریب کاپای ۰/۷۵ عملکرد متوسطی داشت. براساس ماتریس خطا (شکل ۶)، عمده چالش‌های این روش در طبقه‌بندی نمونه‌های کلاس‌های با ویژگی‌های طیفی مشابه مشاهده شد. از ۸۳۰ نمونه کلاس ذرت، ۶۱۷ نمونه به‌درستی طبقه‌بندی شدند، اما ۲۱۳ نمونه به اشتباه به‌عنوان کلاس یونجه تشخیص داده شدند. این نتایج بیانگر ضعف ماشین بردار پشتیبان در جداسازی کلاس‌هایی با هم‌پوشانی طیفی بالا است، هرچند این الگوریتم در طبقه‌بندی کلاس‌هایی با ویژگی‌های طیفی مجزا عملکرد نسبتاً مطلوبی نشان داد. مطالعات پیشین نیز نشان داده‌اند که روش ماشین بردار پشتیبان در جداسازی کلاس‌های با شباهت طیفی بالا، مانند پوشش‌های گیاهی متراکم و نیمه‌متراکم، صحت کم‌تری دارد (Shi and Yang, 2015; Wang et al., 2020). مطالعه (Chauhan and Lunagaria, 2025) نیز روش ماشین بردار پشتیبان عملکرد ضعیف‌تری نسبت به سایر روش‌های طبقه‌بندی از جمله جنگل تصادفی داشته‌است. با توجه به این نتایج، می‌توان گفت که ماشین بردار پشتیبان در مواجهه با داده‌هایی که دارای ویژگی‌های طیفی مجزا هستند، عملکرد نسبتاً مطلوبی دارد، اما در مواجهه با داده‌های پیچیده و دارای هم‌پوشانی طیفی، صحت آن کاهش می‌یابد. این موضوع می‌تواند ناشی از ماهیت خطی این الگوریتم باشد که در مواجهه با داده‌های غیرخطی و پیچیده، عملکرد ضعیف‌تری دارد. برای بهبود عملکرد این روش، می‌توان از تکنیک‌هایی مانند ترکیب این روش با سایر الگوریتم‌های یادگیری ماشین استفاده کرد. شکل ۳، نقشه طبقه‌بندی تصاویر برای دو دوره کشت بهاره و پاییزه را به روش ماشین بردار پشتیبان نشان می‌دهد.



شکل ۳- نقشه طبقه بندی تصاویر برای دو دوره کشت بهاره (a) و پاییزه (b) را به روش ماشین بردار تصمیم در منطقه شبکه آبیاری زهکشی دشت قزوین
Figure 3-Classified maps of images for spring (a) and autumn (b) cultivation periods using the SVM method in the irrigation and drainage network of the Qazvin plain

۳-۲-۲- عملکرد جنگل تصادفی

جنگل تصادفی یکی از الگوریتم های قدرتمند یادگیری ماشین است که به طور گسترده در طبقه بندی داده ها و پیش بینی های پیچیده استفاده می شود. این روش مبتنی بر ساختار مجموعه ای^۱ از درخت های تصمیم است که به صورت موازی آموزش دیده و نتایج آن ها برای بهبود صحت و پایداری مدل ترکیب می شوند (Aria et al., 2021). در این مطالعه، عملکرد جنگل تصادفی در طبقه بندی داده های کشاورزی با استفاده از ماتریس خطا (شکل ۶) و شاخص های ارزیابی مانند ضریب کاپا و صحت کلی در جدول ۳، مورد بررسی قرار گرفته است.

در ماتریس خطا جنگل مشاهده می کنیم که این روش عملکرد فوق العاده ای دارد. برای کلاس های پاییزه از ۲۶۰ نمونه کلاس

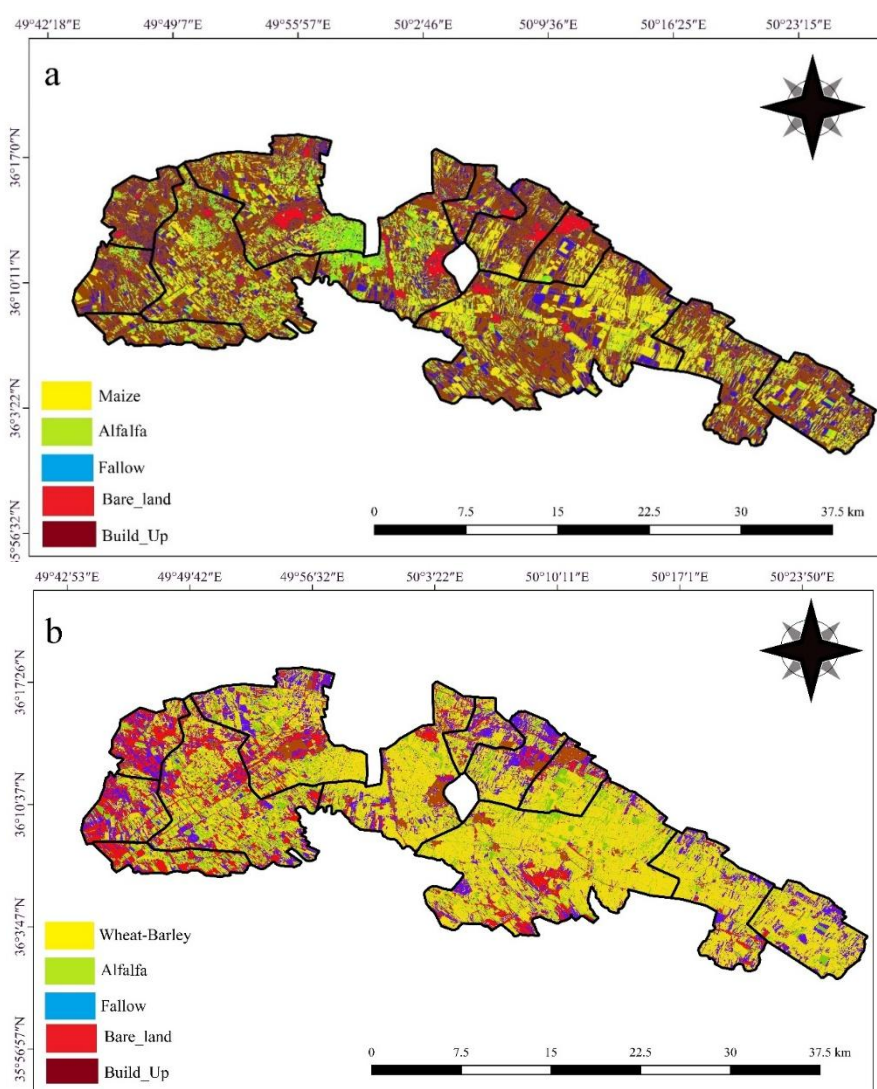
گندم-جو، تنها ۲۰ نمونه در کلاس یونجه قرار گرفته است. این میزان خطا نسبتاً کم است و نشان دهنده توانایی جنگل تصادفی در تشخیص دقیق کلاس ها حتی در حضور هم پوشانی طیفی است. برای کلاس یونجه، از ۲۹۹ نمونه، تنها ۲ نمونه به اشتباه در نظر گرفته شدند. تمامی نمونه های کلاس آیش، مناطق بایر و شهری به درستی طبقه بندی شده اند. از آنجاکه، نمونه های کلاس آیش، مناطق بایر و شهری به درستی طبقه بندی شده اند، تأیید کننده صحت بالای این روش در مدیریت داده های با ویژگی های طیفی مجزا است. براساس نتایج جدول ۳، ضریب کاپا ۰/۹۹ نشان دهنده توانایی بسیار بالای جنگل تصادفی در طبقه بندی دقیق کلاس ها با هم پوشانی است. صحت کلی ۹۹/۶۹ درصد نشان دهنده عملکرد بسیار خوب جنگل تصادفی در طبقه بندی با صحت بالا و پایدار

^۱ Ensemble

تصمیم است که به کاهش واریانس و افزایش پایداری مدل منجر می‌شود. در مطالعه دیگری بیان شده است که جنگل تصادفی به دلیل ساختار مجموعه‌ای و انتخاب تصادفی ویژگی‌ها، معمولاً در داده‌های با نویز یا داده‌های نامتوازن بهترین عملکرد را دارد (Chen et al., 2004). این الگوریتم با ساختن چندین درخت تصمیم‌گیری که به صورت هم‌زمان آموزش می‌بینند و نمونه‌گیری از آن‌ها برای پیش‌بینی نهایی، توانایی مقابله با مشکلاتی نظیر خطای نویز و عدم تعادل داده‌ها را دارد (Aria et al., 2021).

شکل ۴، نقشه طبقه‌بندی تصاویر برای دو دوره کشت بهار و پاییزه را به روش جنگل تصادفی نشان می‌دهد.

است. همچنین، در بررسی طبقه‌بندی بهار، جنگل تصادفی با صحت کلی ۹۸/۹۳ درصد و ضریب کاپای ۰/۹۸ بهترین عملکرد را در میان سه روش نشان داد. این الگوریتم توانایی بالایی در طبقه‌بندی داده‌های پیچیده و مدیریت کلاس‌های با هم‌پوشانی طیفی نشان داد. در طبقه‌بندی کلاس ذرت با ۸۳۰ نمونه تنها ۱۱ نمونه در کلاس یونجه طبقه‌بندی شدند. برای کلاس یونجه نیز از ۹۷۷ نمونه، ۸ نمونه به اشتباه در کلاس ذرت قرار گرفتند. تمامی نمونه‌ها در کلاس‌های آیش، مناطق بایر به درستی انتخاب شده‌اند، و برای کلاس مناطق شهری تنها ۱ نمونه در کلاس آیش قرار گرفته است. این صحت ناشی از توانایی جنگل تصادفی در استفاده از رویکردهای مبتنی بر مجموعه و ترکیب چندین مدل درخت



شکل ۴- نقشه طبقه‌بندی تصاویر برای دو کشت بهار (a) و پاییزه (b) را به روش جنگل تصادفی در منطقه شبکه آبیاری زهکشی دشت قزوین

Figure 4- Classified maps of images for spring (a) and autumn (b) cultivation periods using the random forest algorithm in the irrigation and drainage network of the Qazvin Plain

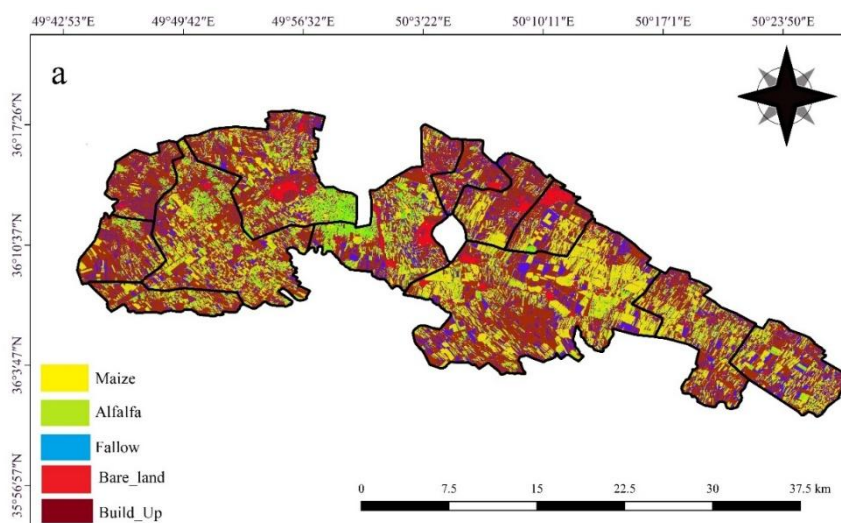
تقویت گرادیان شدید به عنوان یکی از روش‌های قدرتمند یادگیری ماشین، عملکرد قابل قبولی در طبقه‌بندی داده‌ها نشان داده است.

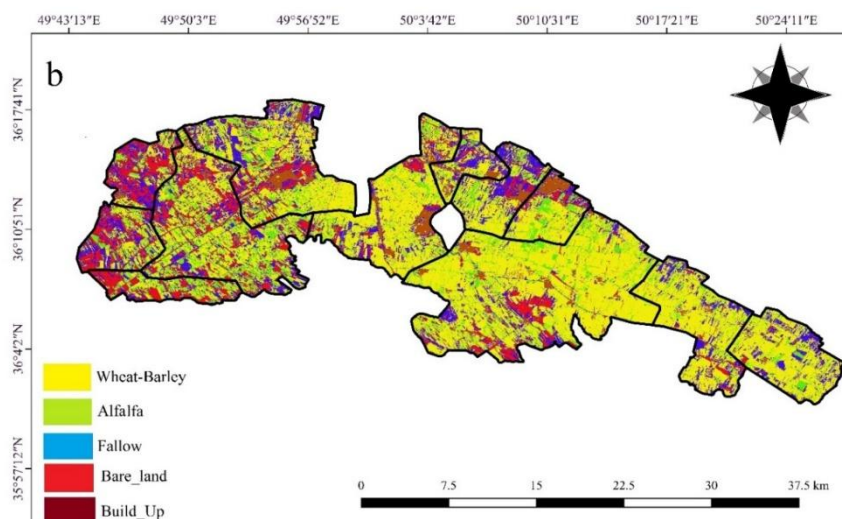
۳-۲-۳- تقویت گرادیان شدید

با این حال، نتایج این مطالعه حاکی از آن است که این روش در مقایسه با جنگل تصادفی عملکرد ضعیفتری دارد (Hastie et al., 2009). این تفاوت احتمالاً ناشی از ساختار متفاوت این دو الگوریتم است. جنگل تصادفی با ترکیب چندین درخت تصمیم‌گیری و کاهش واریانس، عملکرد بهتری در داده‌های پیچیده و با هم‌پوشانی کلاس‌ها دارد، درحالی‌که تقویت‌گرادیان شدید با تمرکز بر کاهش بایاس و بهبود تدریجی مدل، ممکن است در جداسازی کلاس‌های نزدیک به هم با چالش مواجه شود (Friedman, 2001; Ugirumurera et al., 2024). در این مطالعه، ماتریس خطا (شکل ۶)، روش تقویت‌گرادیان شدید نشان می‌دهد که این الگوریتم نیز عملکرد قابل‌قبولی از خود نشان داده‌است، اگرچه عملکرد ضعیف‌تری از جنگل تصادفی نشان داد. در طبقه‌بندی کشت پاییزه، از ۲۷۱ نمونه کلاس گندم-جو، ۲۴۰ نمونه در کلاس درستی قرار گرفته‌اند. برای کلاس یونجه، از ۳۰۰ نمونه تنها ۲ نمونه در کلاس گندم-جو و یک نمونه در کلاس مناطق شهری قرار گرفته‌اند. کلاس‌های یونجه نیز تنها با ۱ نمونه پیکسل خطا رو به رو است، درحالی‌که، کلاس‌های مناطق بایر و شهری به‌درستی طبقه‌بندی شده‌اند. ضریب کاپا ۰/۹۶۵ نشان‌دهنده عملکرد خوب تقویت‌گرادیان شدید است. صحت کلی ۹۷/۳۳ درصد عملکرد این الگوریتم را تأیید می‌کند، اگرچه بهبودهایی برای جداسازی کلاس‌های نزدیک به هم ممکن است مورد نیاز باشد.

در بررسی طبقه‌بندی بهاره، روش تقویت‌گرادیان شدید با صحت کلی ۹۳/۹۴ درصد و ضریب کاپای ۰/۹۰ عملکرد خوبی داشت، اما مشابه با نتایج دوره کشت پاییزه، در مقایسه با جنگل

تصادفی اندکی ضعیف‌تر عمل کرد. چالش‌های اصلی این روش در جداسازی کلاس‌های نزدیک به هم بود. در طبقه‌بندی کلاس ذرت از ۸۳۰ نمونه، ۷۵۸ نمونه به‌درستی انتخاب شده‌اند. در بررسی صحت طبقه‌بندی کلاس یونجه از مجموع ۹۷۷ نمونه تنها ۳۹ نمونه در کلاس ذرت قرار گرفته‌است. این نتایج نشان می‌دهد که تقویت‌گرادیان شدید به‌دلیل ساختار گرادین تقویتی خود در داده‌هایی با پیچیدگی کم‌تر یا میانگین عملکرد مطلوبی دارد، اما برای بهبود جداسازی کلاس‌های با هم‌پوشانی نیاز بررسی و تنظیم پارامترها دارد. پژوهش (Chen and Guestrin, 2016)، نشان داد که الگوریتم تقویت‌گرادیان شدید نسبت به جنگل تصادفی و ماشین بردار پشتیبان با سرعت بالاتر و خطای کم‌تری عمل می‌کند، اما تنظیم پارامترهای مدل برای دستیابی به عملکرد مطلوب ضروری است. هم‌چنین، Fan et al. (2018)، نشان دادند که ترکیب تقویت‌گرادیان شدید با روش‌های انتخاب ویژگی می‌تواند صحت جداسازی کلاس‌ها را بهبود بخشد. بر اساس تحلیل نتایج، جنگل تصادفی بهترین عملکرد را در میان سه روش طبقه‌بندی نشان داد و مناسب‌ترین گزینه برای داده‌های پیچیده با کلاس‌های نزدیک به هم است. روش تقویت‌گرادیان شدید با وجود عملکرد مطلوب، برای دستیابی به نتایج بهتر در کلاس‌های دشوار به تنظیمات بیش‌تر نیاز دارد. روش ماشین بردار پشتیبان در این مقایسه ضعیف‌ترین عملکرد را داشت، که عمدتاً ناشی از عدم توانایی در جداسازی کلاس‌های با شباهت طیفی بالا بود. شکل ۶، نقشه طبقه‌بندی تصاویر برای دو دوره کشت بهاره و پاییزه را به روش تقویت‌گرادیان شدید نشان می‌دهد.





شکل ۵- نقشه طبقه‌بندی تصاویر برای دو دوره کشت بهار (a) و پاییزه (b) را به روش جنگل تصادفی در منطقه شبکه آبیاری زهکشی دشت قزوین
Figure 5- Classified maps of images for spring (a) and autumn (b) cultivation periods using the random forest algorithm in the irrigation and drainage network of the Qazvin Plain

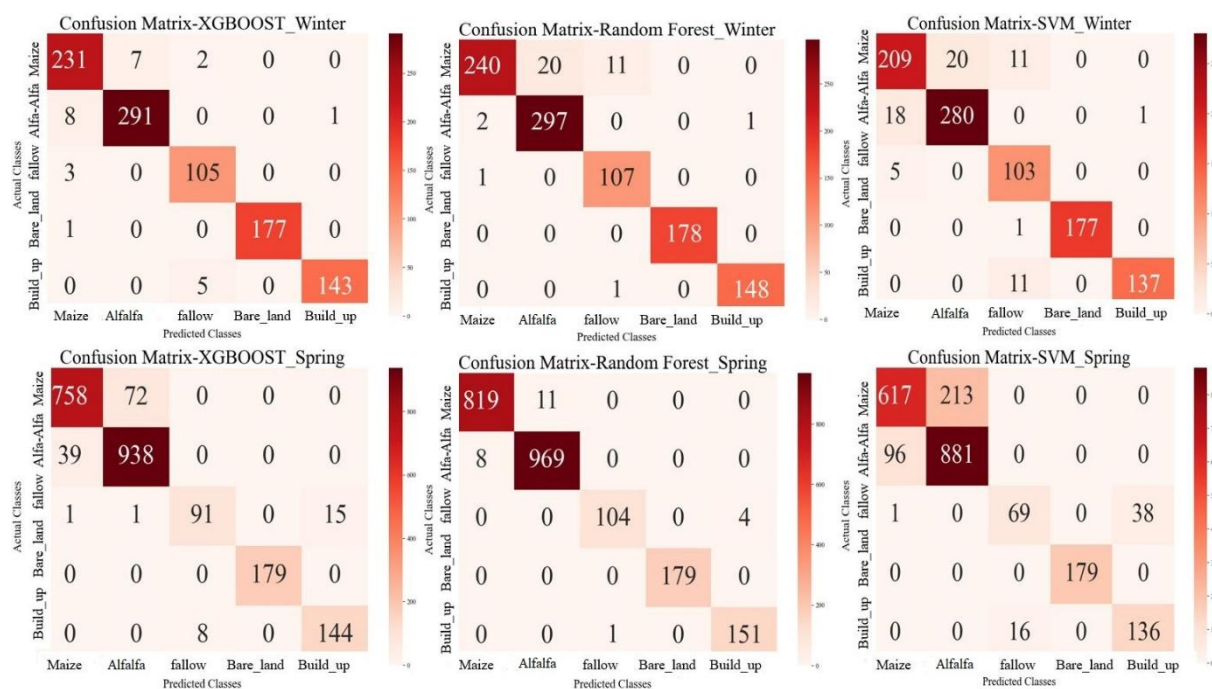
کاپای ۰/۹۱ و ۰/۷۵ نشان می‌دهد، که بیانگر عملکرد ضعیف‌تر این الگوریتم در مقایسه با دو الگوریتم دیگر است. ماتریس خطا نیز تایید کننده این نتایج است. روش تقویت گرادیان شدید، در فصل زمستان بیش‌ترین خطا در طبقه‌بندی کلاس‌های یونجه و ذرت در هر دو فصل مورد بررسی مشاهده می‌شود، البته در بهار عملکرد بهتری از خود نشان داده‌است، روش جنگل تصادفی در هر دو فصل کم‌ترین خطا را در طبقه‌بندی کلاس‌ها نشان می‌دهد که با نتایج صحت کلی و ضریب کاپا مطابقت دارد و به‌عنوان بهترین الگوریتم در این مطالعه شناخته می‌شود. از سوی دیگر، ماشین بردار پشتیبان بیش‌ترین خطا را در طبقه‌بندی کلاس‌ها در هر دو فصل نشان داد اما در نهایت عملکرد ضعیف‌تری نسبت به دو الگوریتم دیگر داشت.

۳-۳- بررسی کلی و مقایسه عملکرد روش‌های طبقه‌بندی

در این مطالعه، نتایج ارزیابی صحت کلی و ماتریس‌های ابهام برای سه الگوریتم طبقه‌بندی شامل ماشین بردار پشتیبان، جنگل تصادفی و تقویت گرادیان شدید در دو فصل پاییز و بهار ارائه شده‌اند. بر اساس جدول ۳، جنگل تصادفی بالاترین صحت کلی را در هر دو فصل با مقادیر ۹۹ و ۹۸ درصد به‌ترتیب برای پاییز و بهار نشان می‌دهد. همچنین، این الگوریتم بالاترین ضریب کاپا را در بین روش‌های دیگر داشته‌است که نشان‌دهنده توافق بسیار خوب بین پیش‌بینی‌ها و داده‌های واقعی است و تقویت گرادیان شدید نیز عملکرد قابل‌قبولی با صحت کلی ۹۶ و ۹۳ درصد و ضریب کاپای ۰/۹۷ و ۰/۹۳ در پاییز و بهار ارائه می‌دهد. در مقابل، SVM کم‌ترین صحت کلی را با مقادیر ۹۳ و ۸۳ درصد و ضریب

جدول ۳- نتایج صحت الگوریتم‌های مختلف طبقه‌بندی داده‌های تست

Table 3- Accuracy results of different classification algorithms applied to test data				
Accuracy assessment		SVM	RF	XGBOOST
Spring Autumn Culti	Overall accuracy	0.93	0.99	0.96
	Kappa coefficient	0.91	0.99	0.97
	Overall accuracy	0.83	0.98	0.93
	Kappa coefficient	0.75	0.98	0.90



شکل ۶- ماتریس خطا روش‌های طبقه‌بندی برای دو فصل کشت بهاره و پاییزه

Figure 6- Confusion matrices of classification methods for spring and autumn cultivation seasons

گردد. همچنین، بهره‌گیری از هسته‌های غیرخطی مانند هسته گوسی در ماشین بردار پشتیبان می‌تواند عملکرد مدل را بهبود بخشد استفاده از منابع داده مکمل مانند شاخص‌های حرارتی، صحت جداسازی کلاس‌ها را می‌تواند افزایش دهند. همچنین، استفاده از الگوریتم‌های مبتنی بر شبکه‌های عصبی مانند CNN¹ و مدل‌های و یا ترکیب روش‌های یادگیری عمیق و کلاسیک (مانند ترکیب CNN و جنگل تصادفی) می‌تواند به افزایش صحت طبقه‌بندی کمک کرده و راه‌حل‌های کارآمدی برای کشاورزی پایدار و مدیریت امنیت غذایی ارائه دهند. علاوه بر این، با توجه به تغییرات اقلیمی و کشاورزی، توسعه مدل‌های پویا که بتوانند تغییرات فنولوژیکی و کاربری اراضی را در طول زمان پیش کنند، ضروری است. این یافته‌ها نه تنها برای مدیریت کشاورزی در منطقه دشت قزوین مفید است، بلکه می‌تواند به عنوان چارچوبی برای مناطق مشابه با شرایط اقلیمی نیمه‌خشک و تنوع بالای پوشش زمین مورد استفاده قرار گیرد. با به کارگیری این روش‌ها، می‌توان به مدیریت بهینه منابع آب، پیش‌بینی عملکرد محصولات و کاهش اثرات محیط زیستی کمک کرد.

سپاسگزاری

نویسندگان از حمایت معاونت پژوهشی دانشگاه بین‌المللی امام خمینی (ره) و بنیاد نخبگان استان قزوین- هیأت اندیشه ورز توسعه

۴- نتیجه‌گیری

این مطالعه با هدف ارزیابی عملکرد الگوریتم‌های یادگیری ماشین در طبقه‌بندی تصاویر چندزمانه ماهواره‌ای سنتینل ۱ و ۲ برای مدیریت کشاورزی در منطقه دشت قزوین انجام شد. نتایج نشان داد که ترکیب داده‌های راداری و اپتیکی ماهواره‌ای به همراه الگوریتم‌های یادگیری ماشین، رویکردی مؤثر برای تفکیک محصولات کشاورزی و مدیریت منابع است. براساس نتایج صحت کلی و ماتریس خطا، روش جنگل تصادفی به دلیل ساختار مجموعه‌ای و توانایی در کاهش واریانس، بهترین گزینه برای طبقه‌بندی داده‌های پیچیده و دارای هم‌پوشانی طیفی است. تقویت گرادیان شدید با وجود عملکرد مطلوب، نیاز به تنظیم دقیق پارامترها دارد تا بتواند در تفکیک کلاس‌های با ویژگی‌های طیفی مشابه بهبود یابد. ماشین بردار پشتیبان به دلیل ماهیت خطی، در مواجهه با داده‌های غیرخطی و پیچیده عملکرد ضعیف‌تری دارد و نیاز به بهبود صحت دارد، درحالی‌که ماشین بردار پشتیبان کم‌ترین صحت و بیش‌ترین خطا را نشان می‌دهد. این یافته‌ها می‌تواند برای پژوهش‌های آینده در انتخاب الگوریتم‌های مناسب برای طبقه‌بندی داده‌های مشابه بسیار مفید باشد و به بهبود صحت و کارایی مدل‌های طبقه‌بندی کمک کند.

پیشنهاد می‌شود برای پژوهش‌های آینده، بهبود تنظیمات مدل با استفاده از روش‌های بهینه‌سازی مانند روش بیزین برای پارامترهای تقویت گرادیان شدید و ماشین بردار پشتیبان بررسی

¹ Convolution neural network

خسروی، ایمان (۱۴۰۳). تهیه نقشه نوع محصول کشاورزی از سری زمانی تصاویر لندست-۸ با استفاده از روش‌های یادگیری ماشین (مطالعه موردی: مرودشت استان فارس). *جغرافیا و برنامه‌ریزی محیطی*، ۳۵(۲)، ۴۵-۶۶.

doi:10.22108/gep.2024.138615.1601

درویش‌هندی، محسن، و امیری تکلدانی، ابراهیم (۱۴۰۲). چالش‌های مدیریت آب کشاورزی در شبکه آبیاری دشت قزوین. *تحقیقات آب و خاک ایران*، ۱۲(۱)، ۱۹۶۲-۱۹۴۵.

doi:10.22059/ijswr.2023.365841.669581

صالحی شفا، نیما، بابازاده، حسین، آقایی، فیاض، صامی، علی، غفوری، محمدرضا، صفوی، مسعود، و پناهدار، علی (۱۴۰۳). تدوین الگوی کشت بهینه به‌منظور مدیریت تغییرات سطح آب زیرزمینی دشت شهریار. *مدل‌سازی و مدیریت آب و خاک*، ۳(۲)، ۲۱۷-۲۳۵.

doi:10.22098/mmws.2022.11792.1169

راستی، سعید، مهدوی فرد، مصطفی، شیخ‌قادر، هدایت، نصیری، ابوزر، و تکتاز، نازنین زهرا (۱۴۰۱). بهبود دقت طبقه‌بندی با ترکیب تصاویر چندفصلی سنتینل ۱ و ۲ به‌منظور تهیه نقشه کاربری اراضی در فضای ابری گوگل ارث انجین (مطالعه موردی: استان گیلان). *جغرافیا و روابط انسانی*، ۵(۳)، ۳۵۷-۳۷۳.

doi:10.22034/gahr.2022.336692.1696

رمضانی اعتدالی، هادی، و احمدی، مژگان (۱۴۰۳). بررسی ارتباط بین شاخص‌های خشکسالی با عملکرد ذرت با استفاده از روش جنگل تصادفی (مطالعه موردی: شبکه آبیاری دشت قزوین). *نیوار*، ۱۲۶-۱۲۷، ۱۳۷-۱۲۷. doi:10.30467/nivar.2024.467444.1299

ساعی جمال آباد، موسی، آبکار، علی اکبر، و مجردی، برات (۱۳۹۷).

طبقه‌بندی گندم زمستانه با استفاده از آنالیز تصاویر بهینه چند زمانی مبتنی بر الگوریتم جنگل تصادفی. *علوم و فنون نقشه برداری*، ۸(۲)، ۱۵۰-۱۳۳. doi:10.1001.1.2322102.1397.8.2.9.8

قدسی، زینب، خیرخواه زرکش، میرمسعود، و قرمزچشمه، باقر (۱۳۹۹).

مقایسه دقت روش‌های ماشین بردار پشتیبان و جنگل تصادفی در تهیه نقشه کاربری اراضی و محصولات زراعی، با استفاده از تصاویر چندزمانه سنتینل-۲. *سنجش از دور و GIS ایران*، ۱۲(۴)، ۹۲-۷۳. doi:10.52547/gisj.12.4.73

References

- Alejo, R., Garcia, V., Sotoca, J.M., Mollineda, R.A., & Sánchez, J.S. (2006). Improving the classification accuracy of RBF and MLP neural networks trained with imbalanced samples. *IDEAL 2006. Lecture Notes in Computer Science, Vol 4224. Springer, Berlin, Heidelberg*. doi:10.1007/11875581_56.
- Aria, M., Cuccurullo, C., & Gnasso, A. (2021). A comparison among interpretative proposals for Random Forests. *Machine Learning with Applications*, 6, 100094. doi:10.1016/j.mlwa.2021.100094.
- Asadi, B., & Shamsoddini, A. (2024). Crop mapping using a combination of Sentinel-1 and 2 images in Ardabil province. *Remote Sensing and GIS*,

پایدار با رویکرد حفاظت از منابع آب زیرزمینی برای تهیه و انتشار آن، قدردانی می‌نمایند.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافعی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش ندارند.

دسترسی به داده‌ها

داده‌ها و نتایج استفاده شده در این پژوهش از طریق مکاتبه با نویسنده مسئول در اختیار قرار خواهد گرفت.

مشارکت نویسندگان

مسعود سلطانی: راهنمایی، جمع‌آوری داده، مفهوم سازی، ویرایش مقاله، انجام نرم افزار/تحلیل آماری، کنترل نتایج؛ **بهاره بهمن آبادی:** برنامه نویسی، مفهوم‌سازی، انجام تحلیل‌های نرم‌افزاری/آماري، نگارش نسخه اولیه مقاله.

منابع

- اسدی، بهار، و شمس‌الدینی، علی (۱۴۰۳). تفکیک محصولات زراعی با استفاده از ترکیب تصاویر سنتینل-۱ و ۲ در استان اردبیل. *سنجش از دور و GIS ایران*، ۱۶(۳)، ۴۶-۲۵. doi:10.48308/gisj.2023.103095
- اسماعیل‌نژاد، رضا، و زینال‌زاده، کامران (۱۳۹۸). ارزیابی تغییرات کاربری اراضی با استفاده از سنجش از دور و سیستم اطلاعات جغرافیایی در زیرحوضه نالوچای. *مدیریت خاک و تولید پایدار*، ۹(۴)، ۱۵۹-۱۷۲. doi:10.22069/ejsms.2020.16657.1892
- پورحسن، ناهید، شاه‌حسینی، رضا، و سیدی، سید تیمور (۱۴۰۰). ارائه روش طبقه‌بندی مبتنی بر یادگیری عمیق در تفکیک انواع محصولات کشاورزی با استفاده از تصاویر ماهواره‌ای سری زمانی. *علوم و فنون نقشه برداری*، ۱۱(۱)، ۱۴۲-۱۲۹. doi:10.1001.1.2322102.1400.11.1.10.8
- 16(3), 25-46. doi:10.48308/gisj.2023.103095[In Persian]
- Asgari, S., & Hasanlou, M. (2023). A comparative study of machine learning classifier for crop type mapping using vegetation indices. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W1-2022, 79. doi: 10.5194/isprs-annals-X-4-W1-2022-79-2023
- Ashourloo, D., Nematollahi, H., Huete, A., Aghighi, H., Azadbakht, M., Shahrabi, H.S., & Goodarzdasthi, S. (2022). A new phenology-based method for mapping wheat and barley using time-series of Sentinel-2 images. *Remote Sensing of Environment*, 280, 113206. doi: 10.1016/j.rse.2022.113206.

- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. doi: 10.1023/A:1010933404324.
- Chauhan, K.K., & Lunagaria, M.M. (2025). Remote sensing-based crop identification and acreage estimation of Rabi Wheat in Anand, Gujarat. *International Environment and Climate Change*, 15(1), 1-11 doi: 10.9734/ijec/2025/v15i14668.
- Chen, C., Liaw, A., & Breiman, L. (2004). Using random forest to learn imbalanced data. University of California, Berkeley, 110(1-12), 24.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 2nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, doi: 10.1145/2939672.2939785.
- Gili, M.R., Ashourloo, D., Aghighi, H., Matkan, A. A., & SHakiba, A. (2022). Crop classification with deep convolutional neural network based on crop feature. *Environmental Sciences*, 20(4), 37-52. doi: 10.48308/envs.2022.1126
- d'Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M., & van der Velde, M. (2021). From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. *Remote Sensing of Environment*, 266, 112708. doi: 10.1016/j.rse.2021.112708.
- Darvish Hendi, M., & Amiri Tokaldany, E. (2024). Agricultural water management challenges in Qazvin Plain Irrigation Network. *Soil and Water Research*, 54(12), 1945-1962. doi: 10.22059/ijswr.2023.365841.669581 [In Persian]
- Dong, Q., Chen, X., Chen, J., Zhang, C., Liu, L., Cao, X., Zang, Y., Zhu, X., & Cui, X. (2020). Mapping winter wheat in North China using Sentinel 2A/B data: A method based on phenology-time weighted dynamic time warping. *Remote Sensing*, 12(8), 1274. doi: 10.3390/rs12081274.
- Esmailnezhad, r., & Zeinalzadeh, K. (2020). Evaluation of land use changes using remote sensing and GIS in Nazlou Chai sub basin. *Soil Management and Sustainable Production*, 9(4), 159-172. doi: 10.22069/ejsms.2020.16657.1892 [In Persian]
- Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., Xiang, L., & Xiang, X. (2018). Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy Conversion and Management*, 164. doi: 10.1016/j.enconman.2018.02.087.
- Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. doi: 10.1214/aos/1013203451
- Ghodsi, Z., Kheirkhah Zarkesh, M.M., & Ghermezcheshmeh, B. (2021). Comparison of accuracy between Support Vector Machine and Random Forest Classifiers for land use and crop mapping using multi-temporal Sentinel-2 Images. *Iranian Remote Sensing and GIS*, 12(4), 73-92. doi: 10.52547/gisj.12.4.73 [In Persian]
- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. *International Journal of Intelligent Systems Technologies and Applications*, 11(2), 105-111. doi: 10.6148/IJITAS.201806_11(2).0003.
- Goldblatt, R., You, W., Hanson, G., & Khandelwal, A.K. (2016). Detecting the boundaries of urban areas in india: A dataset for pixel-based image classification in google earth engine. *Remote Sensing*, 8(8), 634. doi: 10.3390/rs8080634.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of statistical learning*. 1st Edition: Springer. 536 pages.
- Huang, Y., Reddy, K.N., Fletcher, R.S., & Pennington, D. (2018). UAV low-altitude remote sensing for precision weed management. *Weed Technology*, 32(1), 2-6. doi: 10.1017/wet.2017.89
- Khosravi, I. (2024). Crop mapping from Landsat-8 images time series using Machine-Learning Methods (Case Study: Marvdasht in Fars Province). *Geography and Environmental Planning*, 35(2), 45-66. doi: 10.22108/gep.2024.138615.1601 [In Persian]
- Li, S., Gong, Q., & Yang, S. (2019). A sustainable, regional agricultural development measurement system based on dissipative structure theory and the entropy weight method: a case study in Chengdu, China. *Sustainability*, 11(19), 5313. doi: 10.3390/su11195313
- Li, W., Zhang, H., Li, W., & Ma, T. (2022). Extraction of winter wheat planting area based on multi-scale fusion. *Remote Sensing*, 15(1), 164. doi: 10.3390/rs15010164.
- Li, Y., Zhu, X., Pan, Y., Gu, J., Zhao, A., & Liu, X. (2014). A Comparison of model-assisted estimators to infer land cover/use class area using satellite imagery. *Remote Sensing*, 6(9), 8904-8922. doi: 10.3390/rs6098904.
- Saei Jamalabad, M., Akbari, A., & Mojardi, B. (2018). Winter wheat classification by multi-temporal optimized image analysis based on Random Forest algorithm. *Geomatic Science and Technology*, 8(2), 133-150. dor: 20.1001.1.2322102.1397.8.2.9.8 [In Persian]
- Moumni, A., & Lahrouni, A. (2021). Machine learning-Based classification for crop-type mapping using the fusion of high-resolution satellite imagery in a Semiarid Area. *Scientifica*, 2021(1), 8810279. doi: 10.1155/2021/8810279.

- Neetu, & Ray, S. (2019). Exploring machine learning classification algorithms for crop classification using Sentinel 2 data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 573-578. doi: 10.5194/isprs-archives-XLII-3-W6-573-2019.
- Padma VVL, A., Suhail, M., Lutfullo, I., & Shodiyor, B. (2024). A comparative study of three supervised algorithms for mixed crop classification. *E3S Web of Conference*, Pp. 590-595, doi: 10.1051/e3sconf/202459001004.
- Phalke, A. R., Özdoğan, M., Thenkabail, P.S., Erickson, T., Gorelick, N., Yadav, K., & Congalton, R.G. (2020). Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat, Random Forest, and Google Earth Engine. *ISPRS Photogrammetry and Remote Sensing*, 167, 104-122. doi: 10.1016/j.isprsjprs.2020.06.022.
- Pourhasan, N., Shah-Hosseini, R., & Seydi, S.T. (2021). Deep learning-based classification method for crop mapping using time series satellite images. *Geomatics Science and Technology*, 11(1), 129-142. doi: 10.1001.1.2322102.1400.11.1.10.8 [In Persian]
- Ramezani Etedali, H., & Ahmadi, M. (2024). Investigating the relationship between drought indices and maize yield using random forest method (Case study: Qazvin plain irrigation network). *Nivar*, 126, 127-137. doi: 10.30467/nivar.2024.467444.1299 [In Persian]
- Rasti, S., MahdaviFardn, M., Shaykh Ghaderi, H., Nasiri, A., & Taktaz, N.Z. (2022). Improving classification accuracy by combining multi-season images of Sentinel 1 and 2 in order to prepare a land use map in the cloud space of Google Earth Engine (Case study: Guilan province). *Geography and Human Relationships*, 5(3), 357-373. doi: 10.22034/gahr.2022.336692.1696 [In Persian]
- Sahin, E.K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, 2(7), 1308. doi: 10.1007/s42452-020-3060-1.
- Salehi Shafa, N., Babazadeh, H., Aghayari, F., Saremi, A., Ghafouri, M.R., Safavi, M., & Panahdar, A. (2022). Formulation of an optimized Cropping pattern in order to manage groundwater level changes in Shahriar Plain. *Water and Soil Management and Modelling*, 3(2), 217-235. doi: 10.22098/mmws.2022.11792.1169 [In Persian]
- Sarker, I.H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160. doi: 10.1007/s42979-021-00592-x.
- Seifi Majdar, R., & Ghassemian, H. (2017). A probabilistic SVM approach for hyperspectral image classification using spectral and texture features. *Remote Sensing*, 38(15), 4265-4284. doi: 10.1080/01431161.2017.1317941.
- Sen, R., Goswami, S., & Chakraborty, B. (2019). Jeffries-Matusita distance as a tool for feature selection. *International Conference on Data Science and Engineering (ICDSE)*. doi: 10.1109/ICDSE47409.2019.8971800
- Shao, Z., Ahmad, M.N., & Javed, A. (2024). Comparison of Random Forest and XGBoost classifiers using integrated optical and SAR features for mapping urban impervious surface. *Remote Sensing*, 16(4), 665. doi: 10.3390/rs16040665.
- Shi, D., & Yang, X. (2015). Support Vector Machines for land cover mapping from remote sensor imagery. Pp. 265-279. In *Monitoring and Modeling of Global Changes: A Geomatics Perspective*. SPRINGERREMO. doi: 10.1007/978-94-017-9813-6_13.
- Simsek, F. (2025). Comparison of Agricultural Crop Type Classifications with Different Machine Learning Algorithms (RF-SVM-ANN-XGBoost) by Generating Ground Truth Data from Farmer Declaration Parcels. *Engineering and Geosciences*, 10(2), 207-220. doi: 10.26833/ijeg.1552141.
- Sun, W., Yang, G., Chen, C., Chang, M., Huang, K., Meng, X., & Liu, L. (2020). Development status and literature analysis of China's earth observation remote sensing satellites. *Remote Sensing*, 24(5), 479-510. doi: 10.11834/jrs.20209464.
- 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. doi: 10.1109/JSTSP.2011.2139193.
- Ugirimurera, J., Bensen, E.A., Severino, J., & Sanyal, J. (2024). Addressing bias in bagging and boosting regression models. *Scientific Reports*, 14(1), 18452. doi: 10.1038/s41598-024-68907-5.
- Viskovic, L., Kosovic, I.N., & Mastelic, T. (2019, 19-21 Sept. 2019). Crop classification using multi-spectral and multitemporal satellite imagery with machine learning. 2019 *International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, doi: 10.23919/SOFTCOM.2019.8903738.
- Wang, K., Cheng, L., & Yong, B. (2020). Spectral-similarity-based Kernel of SVM for hyperspectral image classification. *Remote Sensing*, 12(13), 2154. doi: 10.3390/rs12132154.
- Wang, Y., Qi, Q., & Liu, Y. (2018). Unsupervised segmentation evaluation using area-weighted variance and Jeffries-Matusita distance for remote sensing Images. *Remote Sensing*, 10(8), 1193. doi: 10.3390/rs10081193.

- Wuyun, D., Sun, L., Sun, Z., Chen, Z., Hou, A., Teixeira Crusiol, L. G., Reymond, L., Chen, R., & Zhao, H. (2022). Mapping fallow fields using Sentinel-1 and Sentinel-2 archives over farming-pastoral ecotone of Northern China with Google Earth Engine. *GIScience and Remote Sensing*, 59(1), 333-353. doi: 10.1080/15481603.2022.2026638.
- Yadav, S., Sharma, S., & Chaudhary, P. (2024). Review on advancement in machine learning and Deep Learning techniques for crop classification. *SCRS Publication*. doi: 10.56155/978-81-955020-7-3-53.
- Yaramasu, R., Bandaru, V., & Pnvr, K. (2020). Pre-season crop type mapping using deep neural networks. *Computers and Electronics in Agriculture*, 176, 105664. doi: 10.1016/j.compag.2020.105664.
- Zhan, W., Luo, F., Luo, H., Li, J., Wu, Y., Yin, Z., Wu, Y., & Wu, P. (2024). Time-series-based spatiotemporal fusion network for improving crop type mapping. *Remote Sensing*, 16(2), 235. doi: 10.3390/rs16020235.
- Zheng, B., Myint, S.W., Thenkabail, P.S., & Aggarwal, R.M. (2015). A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *Applied Earth Observation and Geoinformation*, 34, 103-112. doi: 10.1016/j.jag.2014.07.002.