

geoprzestrzennych Big Data wymaga skutecznego rozwiązania wyzwań związanych z przetwarzaniem, analizą i interoperacyjnością danych, a także inwestycji w rozwój zaawansowanych technik.

**Słowa kluczowe:** teledetekcja, big data, dane geoprzestrzenne, geoinformatyka

## **Review of Artificial Intelligence Tools for Processing Ground-based LiDAR Data in Precision Forestry**

### **Przegląd narzędzi sztucznej inteligencji do przetwarzania naziemnych danych LiDAR w leśnictwie precyzyjnym**

**Maksymilian Kulicki**<sup>1,2\*</sup>, Carlos Cabo<sup>3</sup>, Tomasz Trzciniński<sup>2,4</sup>, Janusz Będkowski<sup>1</sup>, Krzysztof Stereńczak<sup>2,5</sup>

<sup>1</sup> Institute of Fundamental Technological Research Polish Academy of Sciences, Pawińskiego St. 5B, 02-106 Warsaw, Poland

<sup>2</sup>IDEAS NCBR, Chmielna 69 St., 00-801 Warsaw, Poland

<sup>3</sup>Department of Mining Exploitation and Prospecting, University of Oviedo, Mieres, Spain

<sup>4</sup>Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland

<sup>5</sup>Department of Geomatics, Forest Research Institute, 3 Braci Leśnej St., Sękocin Stary, 05-090 Raszyn, Poland

\*Corresponding author: maksymilian.kulicki@ideas-ncbr.pl

#### **Abstract**

Ground-based LiDAR tools, both static and mobile, are getting increasingly popular in forestry research. They enable detailed and accurate mapping of individual trees, not available with airborne scanning. With the advent of such high-detail data collection, there emerges a crucial need for appropriate data processing techniques. A promising approach involves the use of Artificial Intelligence methods, in particular Machine Learning and Deep learning.

This work is a systematic review of such AI methods used to process ground-based LiDAR data for forest inventory purposes. We identified three core tasks solved by AI: semantic segmentation, individual tree segmentation, and tree species classification, along with several other tasks.

Our results show that field is rapidly growing, driven by the increasing popularity of Deep learning methods since 2022. Deep learning approaches give better results than traditional machine learning models on a variety of tasks. However, while the current research leverages many state-of-the-art DL architectures, many methods are overlooked. Many advanced AI models, particularly instance segmentation architectures for identifying individual trees, are still overlooked. Additionally, Graph Neural Networks, a promising class of models, remains largely unexplored. The field also lags behind in areas of semi-supervised and self-supervised learning, approaches which can make use of unlabeled data. A major hurdle for progress is the lack of openly shared code and data, hindering research reproducibility. This issue is amplified by the lack of standardized evaluation metrics. The field could benefit from large benchmark datasets that would provide rich training data and a fair ground for comparison.

In summary, the review shows the increasing potential of using AI for forest inventory with ground-based LiDAR. It could be further enhanced by standardized evaluation, open code sharing and exploring new AI paradigms.

**Keywords:** forestry, artificial intelligence, machine learning, deep learning

#### **Abstrakt**

Naziemne skanery LiDAR, zarówno statyczne, jak i mobilne, stają się coraz bardziej popularne w leśnictwie. Umożliwiają one szczegółowe i dokładne mapowanie pojedynczych drzew, niedostępne w przypadku skanowania z powietrza. Wraz z pojawieniem się tak szczegółowych danych, pojawia się kluczowa potrzeba odpowiednich technik przetwarzania. Obiecujące podejście obejmuje wykorzystanie metod sztucznej inteligencji, w szczególności uczenia maszynowego i głębokiego uczenia.

Niniejsza praca stanowi systematyczny przegląd takich metod wykorzystywanych do przetwarzania naziemnych danych LiDAR do inwentaryzacji lasów. Zidentyfikowaliśmy trzy podstawowe zadania rozwiązywane przez sztuczną inteligencję: segmentację semantyczną, segmentację poszczególnych drzew i klasyfikację gatunków drzew, a także kilka innych zadań.

Nasze wyniki pokazują, że dziedzina ta szybko się rozwija, napędzana rosnącą popularnością uczenia głębokiego od 2022 roku. Podejścia oparte na uczeniu głębokim dają lepsze wyniki niż tradycyjne modele uczenia maszynowego. Chociaż obecne badania wykorzystują wiele nowoczesnych architektur, wiele metod jest pomijanych. Część zaawansowanych modeli, w szczególności architektury segmentacji które można by zastosować do identyfikacji pojedynczych drzew, jest nadal pomijanych. Ponadto grafowe sieci neuronowe, obiecująca klasa modeli, pozostaje w dużej mierze niezbadana. Dziedzina pozostaje również w tyle w obszarach semi-supervised i self-supervised learning, metod które mogą wykorzystywać dane pozbawione ludzkich anotacji. Ważną przeszkodą dla rozwoju jest brak otwartego udostępniania kodu i danych, co utrudnia odtwarzalność badań. Problem ten potęguje brak ustandaryzowanych metod ewaluacji. Dziedzina mogłaby skorzystać z dużych, publicznych zbiorów danych, które zapewniłyby źródło danych treningowych i podstawę do porównań.

Podsumowując, praca pokazuje rosnący potencjał wykorzystania sztucznej inteligencji do inwentaryzacji lasów za pomocą naziemnego LiDAR-u. Potencjał ten można dodatkowo zwiększyć poprzez standaryzację ewaluacji, otwarte udostępnianie kodu i eksplorowanie nowych paradygmatów sztucznej inteligencji.

**Słowa kluczowe:** leśnictwo, sztuczna inteligencja, uczenie maszynowe

## **Assessment of the accuracy of winter wheat yield models based on remote sensing data**

### **Ocena dokładności modeli wzrostu pszenicy ozimej na podstawie danych teledetekcyjnych**

**Radosław Gurdak**<sup>1\*</sup>, Katarzyna Dąbrowska-Zielińska<sup>2</sup>, Jędrzej Bojanowski<sup>3</sup>

<sup>1</sup>Polish Space Agency, Trzy Lipy 3 St., 80-172 Gdańsk, Poland

<sup>2</sup>Institute of Geodesy and Cartography, Modzelewskiego 27 St., 02-679 Warsaw, Poland

<sup>3</sup>CloudFerro S.A., Fabryczna 5 St., 00-446 Warsaw, Poland

\*Corresponding author: radoslaw.gurdak@polsa.gov.pl

#### **Abstract**

The importance of agriculture should be seen primarily in the context of meeting the world's nutritional needs. Despite the measures taken, the percentage of people who experienced moderate or high food insecurity is systematically growing. The presented research has been performed and designed under the PhD studies at the University of Warsaw. The goal of the PhD thesis is to create an algorithm that allows to estimate the yield of winter wheat based on satellite data using machine learning methods - Support Vector Regression (SVR), Random Forest (RF) and AquaCrop Model.

During 2016-2019 crop growing season a series of ground measurements of soil vegetation parameters have been performed in Poland agricultural areas. Field measurements were conducted at agriculture study site which belongs to Joint Experiment for Crop Assessment and Monitoring (JECAM-GEOGLAM).

The research assessed the range of variability of various LAI prediction methods: empirical model, Sentinel Application Platform (SNAP), soCAP and LAI products from the Copernicus Land Monitoring website, calculated on the basis of Proba-V satellite images. The results obtained from the images from the Sentinel-2 satellite showed the best accuracy (RMSE=0,52). The root mean square error (RMSE) for the empirical model, SNAP and soCAP was the lowest in all phases. There is a significant decrease in the effectiveness of the prediction of the LAI after the use of 300-meter decade-long products calculated on Proba-V satellite (RMSE=1,84).

The final stage of the work was the use of estimated LAI values from satellite imagery and meteorological data in SVR, RF and AquaCrop models in order to predict yields. The results obtained with the use of RF calculated