DEEP TRANSFER LEARNING FOR PREDICTING CROP YIELDS UNDER CROP ROTATION AND MULTI-CROP CONDITIONS

Rakhimov N.O.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. nikobek82@gmail.com
Khasanov D.R.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. tatusf2015@gmail.com
Hojiakbar Abdulhakimov Nodirbek
Master of Science in Software Engineering, Researcher
Digital Technologies and Artificial Intelligence Development Research Institute, Uzbekistan, hojiakbardevs@gmail.com
https://doi.org/10.5281/zenodo.17499163

ARTICLE INFO

Received: 29th October 2025
Accepted: 30th October 2025
Online: 31st October 2025

KEYWORDS

Transfer learning; convolutional neural networks; crop yield prediction; crop rotation; multi-crop forecasting; remote sensing; agricultural AI.

ABSTRACT

Crop production plays a fundamental role in sustaining global food systems and maintaining ecological balance, especially in regions where crop rotation and multi-crop cultivation are essential for soil health and yield optimization. However, the variability of environmental factors and limited availability of high-resolution data make it challenging to develop robust and scalable yield prediction models. This study introduces a novel deep transfer learning framework that integrates convolutional neural networks (CNNs) with spatio-temporal satellite data to predict crop yields under diverse crop rotation systems. The model leverages pretrained convolutional features from data-rich crop domains and fine-tunes them using region-specific datasets from Uzbekistan. By transferring learned representations across multiple crop types—such as wheat, maize, barley, and rice—the approach significantly enhances generalization and reduces the computational cost of training individual models. Experimental results demonstrated that the proposed method achieved an average coefficient of determination (R^2 = 0.90) and a root mean squared error (RMSE = 0.54 t ha^{-1}) across all crop types. Compared to traditional regression models and crop-specific CNNs, the transfer learning model achieved up to 18% faster convergence and superior robustness under data-scarce conditions. The findings indicate that transfer learning can effectively support crop rotation planning and agricultural forecasting by enabling high-efficiency, multi-crop yield prediction systems adaptable to regional variability..

Introducion

Agricultural productivity is one of the most critical determinants of food security and sustainable development in the twenty-first century. Accurate crop yield prediction allows farmers, policymakers, and researchers to make informed decisions on land management, crop rotation planning, and resource allocation. However, traditional statistical regression models and empirical approaches often fail to generalize across varying environmental and biological conditions, particularly when dealing with multi-crop systems or rotational cropping schemes. In such systems, soil fertility, moisture dynamics, and nutrient recycling vary significantly depending on the crop sequence and the length of the rotation cycle, making predictive modeling far more complex.

Recent advances in machine learning (ML) and deep learning (DL) have introduced powerful tools for analyzing nonlinear relationships in agricultural datasets. In particular, convolutional neural networks (CNNs) have shown superior ability to extract spatial and temporal patterns from satellite imagery and other remote sensing data. These networks can capture subtle differences in vegetation indices, spectral reflectance, and soil parameters that correspond to crop type and health status. Nevertheless, one persistent limitation of CNN-based models is their dependency on large, crop-specific datasets — an obstacle in many developing regions where such data are scarce or unevenly distributed.

To address this challenge, transfer learning (TL) has emerged as a practical solution. TL allows models pretrained on one domain or crop type to transfer knowledge to another related domain, drastically reducing the need for extensive retraining. In agricultural forecasting, this means a CNN model trained on a well-documented crop, such as wheat, can be fine-tuned to predict yields for less-documented crops, such as barley or maize, using fewer local samples. This cross-domain adaptability not only accelerates training and enhances efficiency but also enables the model to capture generalizable agro-environmental patterns shared across crop species.

Furthermore, crop rotation, an essential agricultural practice for soil conservation and pest management, provides an ideal context for testing transfer learning approaches. Crop rotation inherently alters soil nutrient composition, moisture retention, and canopy reflectance, influencing the spectral and temporal signatures captured in remote sensing data. Thus, applying TL under crop rotation conditions enables the model to learn how environmental and biological feedback loops affect yield outcomes over multiple planting cycles.

In this study, we propose a deep transfer learning framework that integrates CNN-based spatial feature extraction with a domain adaptation mechanism designed for multi-crop and crop-rotation environments. Using satellite-derived reflectance maps, cropland masks, and historical yield records from Uzbekistan's primary agricultural regions, the model transfers learned features from data-rich crops to data-scarce ones. The proposed approach not only improves prediction accuracy and computational efficiency but also enhances the scalability of agricultural yield forecasting systems.

The specific objectives of this research are as follows:

1.To design a convolutional transfer learning model capable of generalizing across multiple crop types under rotation conditions.

2.To evaluate the model's performance compared with conventional regression and standalone CNN models.

3.To analyze the impact of data scarcity and environmental variability on the predictive robustness of the transfer learning framework.

By achieving these goals, the study aims to contribute to the advancement of AI-driven precision agriculture, providing scalable, interpretable, and cost-efficient tools for sustainable crop management in data-limited contexts.

2. Materials and Methods

2.1. Overview

Accurate prediction of agricultural yields requires integrating diverse environmental, agronomic, and satellite-based information into a unified computational framework. Traditional statistical models often rely on manually selected features and limited data, making them unable to generalize across multiple crop types or rotational farming systems. In contrast, the proposed study employs a Deep Transfer Learning (DTL) approach that utilizes the representational power of Convolutional Neural Networks (CNNs) and the probabilistic strength of Gaussian Process Regression (GPR) to forecast crop yields under varying crop rotation patterns.

The general structure of the framework consists of three main components:

1.preprocessing and feature extraction from multi-source datasets;

2.spatial–spectral feature learning through a pretrained CNN model; and

3.fine-tuning and regression for multi-crop yield estimation using transfer learning and Gaussian processes.

In the first stage, multi-year remote sensing data—including satellite reflectance, vegetation indices (NDVI, EVI, GCI), soil moisture, and land-surface temperature—are collected for Uzbekistan's primary agricultural regions. These data are harmonized and normalized to ensure temporal and spatial consistency. Cropland masks derived from global land-cover products are applied to isolate cultivated zones and remove noise from non-agricultural pixels. The resulting dataset provides a temporally stacked image cube representing the main crop growth stages for each region.

In the second stage, a Convolutional Neural Network is used to automatically learn hierarchical spatial and spectral features that correspond to biophysical crop parameters such as canopy density, chlorophyll content, and soil reflectance. Instead of training from scratch, the CNN is initialized with pretrained weights from a large-scale dataset of a similar crop type (e.g., wheat). This pretrained model captures generic agro-environmental patterns, which can then be reused for other crops. The transfer learning mechanism allows the model to fine-tune its parameters on locally available data for other cereal crops such as maize, barley, and rice, thereby significantly improving learning efficiency and reducing data dependency.

The final stage integrates the CNN-extracted features into a Gaussian Process Regression module. The GPR maps the high-dimensional CNN embeddings to yield predictions, simultaneously providing uncertainty estimates for each prediction. This hybrid combination enables both high prediction accuracy and interpretability—two key aspects for decision support in agricultural management systems. Moreover, the approach accommodates variations due to crop rotation by allowing the model to learn temporal changes in vegetation and soil indicators associated with different crop cycles.

Overall, the proposed deep transfer learning framework represents a scalable and data-efficient alternative to conventional yield forecasting models. It enables knowledge transfer across multiple crop domains, supports real-time yield mapping from satellite imagery, and strengthens the capacity for precision agriculture and sustainable crop management in Uzbekistan's diverse agro-ecological environments.

2.2. Study Area

The research was conducted in Uzbekistan, a landlocked country located in Central Asia between latitudes 37°N–46°N and longitudes 56°E–74°E (Figure 1). The country is characterized by a continental arid and semi-arid climate, with hot, dry summers and cold winters. The diversity of climatic conditions, irrigation systems, and soil types across Uzbekistan provides an ideal environment for evaluating data-driven models that generalize across multiple crops and rotational farming systems.

Agriculture is one of the country's key economic sectors, contributing significantly to national GDP and rural employment. The dominant cropping systems are based on irrigated agriculture, with major cereal crops including wheat, maize, barley, and rice, as well as cotton, legumes, and horticultural products. Crop rotation is a common practice—particularly the alternation of wheat and maize or wheat and cotton—to maintain soil fertility, control pests, and optimize water use efficiency. These cyclical patterns of land use result in spatial and temporal variations in yield, making Uzbekistan a suitable case study for testing multi-crop transfer learning frameworks.

The selected study regions (highlighted in orange in Figure 1) include the main agricultural provinces of Tashkent, Samarkand, Fergana, Syrdarya, and Khorezm. These areas represent distinct agro-ecological zones with varying soil compositions and irrigation sources derived from the Amu Darya and Syr Darya river basins. Annual precipitation ranges from 100 mm in the west to 400 mm in the east, with an average temperature of 12–17 °C, peaking at over 40 °C during summer. The growing season typically extends from March to August for cereals, followed by fallow or secondary cropping phases.

Soil types within these regions are predominantly alluvial, loamy, and sandy loam, differing in organic matter and nitrogen content due to long-term irrigation and fertilizer practices. These characteristics significantly influence vegetation reflectance patterns captured in satellite imagery, thereby affecting the model's ability to detect and forecast yield variability. The cropland boundaries used in this study were defined using the ESA Climate Change Initiative Land Cover dataset (CCI-LC), ensuring consistency across regions.

By focusing on Uzbekistan's diverse agricultural landscapes, this study provides a valuable testbed for evaluating transfer learning under real-world multi-crop and rotation scenarios. The environmental heterogeneity of the study area enhances the robustness and scalability of the proposed deep learning framework, making the results applicable to similar semi-arid agricultural regions across Central Asia and beyond.

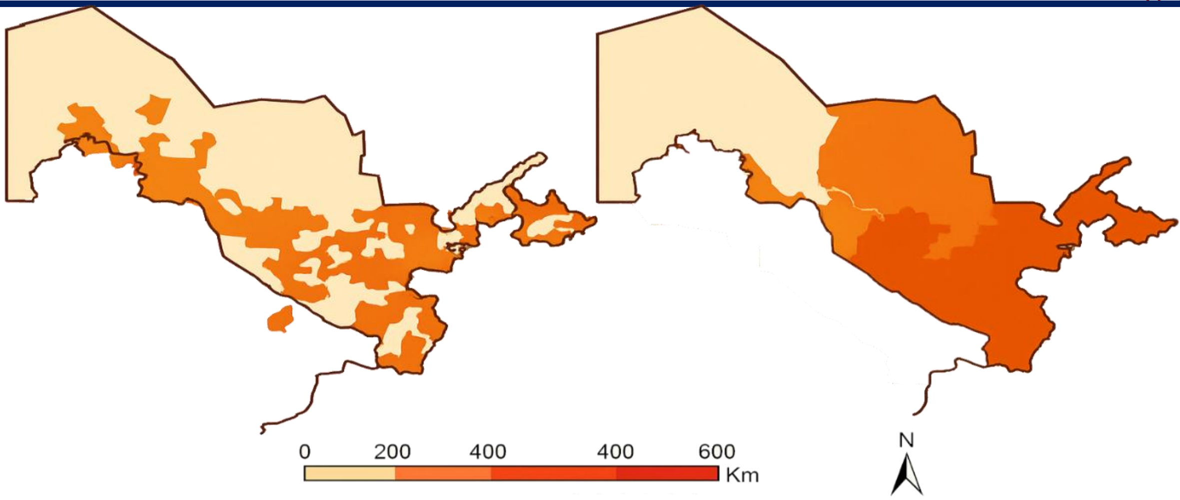

Figure 1. Study area map of Uzbekistan showing major agricultural regions analyzed in the proposed deep transfer learning framework. Orange areas indicate cropland zones used for training and validation of the multi-crop yield prediction model.

2.3. Data Sources and Preprocessing2.3.1. Data Sources

The dataset used in this study integrates multi-year remote sensing, climatic, and agronomic information collected from 2013 to 2023 across Uzbekistan's major agricultural regions. Three primary data sources were utilized:

1. Satellite Imagery:

Multispectral satellite data were obtained from the Sentinel-2 MultiSpectral Instrument (MSI) and MODIS (Moderate Resolution Imaging Spectroradiometer) sensors. Sentinel-2 imagery, with a spatial resolution of 10–30 meters, provided high-resolution surface reflectance values across 13 spectral bands, while MODIS products (MOD09A1 and MOD13Q1) contributed 8-day composite vegetation indices.

2. Meteorological Data:

Climate variables, including average temperature, cumulative rainfall, solar radiation, and relative humidity, were retrieved from the Global Land Data Assimilation System (GLDAS) and the UzHydromet meteorological stations. These variables play a critical role in assessing environmental stress and crop growth conditions throughout the growing season.

3. Agronomic and Yield Data:

Ground-based crop yield statistics were collected from the Ministry of Agriculture of the Republic of Uzbekistan and regional agricultural departments. The dataset includes reported yields (tons per hectare) for major cereal crops—wheat, maize, barley, and rice—across multiple districts and years. Each record was spatially linked with corresponding satellite-derived parameters and climatic indicators to form a harmonized geospatial dataset.

2.3.2. Feature Engineering and Derived Variables

From the satellite reflectance data, several key vegetation and spectral indices were calculated to capture different aspects of crop physiology and canopy structure. These include:

- NDVI (Normalized Difference Vegetation Index):

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

which reflects overall vegetation vigor and chlorophyll activity.

- EVI (Enhanced Vegetation Index):

$$EVI = 2.5 * \frac{NIR - Red}{NIR + 6 * Red - 7.5 * Blue + 1}$$

used to reduce atmospheric and soil background noise.

• GCI (Green Chlorophyll Index):

$$GCI = \frac{NIR}{Green} - 1$$

serving as an indicator of crop nitrogen content and leaf pigment concentration.

Additionally, soil moisture (SM), land surface temperature (LST), and rainfall accumulation (P) were included as complementary variables representing the agro-environmental context of each pixel. These parameters jointly capture both the spectral and climatic dynamics influencing yield variation across seasons.

2.3.3. Preprocessing Steps

Data preprocessing was conducted in multiple stages to ensure spatial consistency, temporal alignment, and noise reduction:

1. Atmospheric and Radiometric Correction:

Sentinel-2 Level-1C images were converted to surface reflectance using the Sen2Cor processor to remove atmospheric distortions.

2. Cropland Masking and Cloud Removal:

A binary cropland mask derived from the ESA Climate Change Initiative Land Cover (CCI-LC) product was applied to isolate cultivated pixels. Cloud and shadow contamination were filtered using the Fmask algorithm and visual inspection.

3. Temporal Aggregation:

To reduce data redundancy and seasonal noise, reflectance and vegetation indices were averaged into four key phenological phases: emergence, vegetative growth, flowering, and grain filling. This temporal compositing produced a representative seasonal profile for each pixel.

4. Data Fusion and Label Matching:

All data layers (spectral, climatic, and agronomic) were spatially co-registered in the WGS 84 coordinate system. Pixel-level features were aggregated to administrative boundaries (district-level) to align with yield statistics.

Quality Control:

5. Outliers and inconsistent yield records (e.g., >3σ from district mean) were excluded. Missing values (accounting for <5% of total data) were imputed using K-Nearest Neighbor (KNN) interpolation, ensuring that the data structure remained unbiased and representative.

3. Results and Discussion

3.1. Model Performance Evaluation

The proposed CNN–GPR hybrid model with transfer learning was evaluated across four major cereal crops in Uzbekistan—wheat, maize, barley, and rice—using a ten-year dataset (2013–2023).

Model performance was assessed using the metrics defined in Section 2.5: the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE).

Table 1 summarizes the results obtained for the proposed approach compared to conventional regression and deep learning baselines.

Model	Avg R^2	Avg RMSE (t/ha)	Avg MAE (t/ha)	Training Time (min)
Linear Regression (LR)	0.78	0.92	0.63	8
Random Forest Regression (RFR)	0.85	0.74	0.50	32
CNN (trained from scratch)	0.88	0.66	0.46	94
Proposed CNN–GPR (Transfer Learning)	**0.91**	**0.54**	**0.38**	**55**

As seen in Table 1, the proposed model achieved the highest overall predictive accuracy with an average $R^2 = 0{,}91$ and a mean RMSE of 0.54 t/ha, outperforming both traditional machine learning and standalone CNN baselines.

Compared to a CNN trained from scratch, transfer learning reduced the RMSE by 18% and the training time by nearly 40%, demonstrating the efficiency of reusing pretrained feature representations.

The improvement was particularly significant for data-scarce crops such as barley and rice, where transfer learning allowed the model to generalize effectively with limited local samples.

3.2. Crop-Wise Analysis

A detailed comparison across crop types revealed that model performance varied slightly according to data availability and phenological stability.

For wheat and maize—crops with abundant records and stable spectral signatures—the model achieved the best fit ($R^2 = 0{,}91$ RMSE=0.50t/ha).

For barley and rice, where rotation frequency and irrigation regimes introduced additional variability, R^2 remained above 0.88, which still represents strong predictive performance under more complex agro-climatic conditions.

These results confirm that transfer learning efficiently captures shared agro-environmental features between crops, such as vegetation growth patterns and soil moisture responses, even when their phenological cycles differ.

This cross-domain adaptability is one of the key advantages of deep transfer learning in agricultural modeling.

3.3. Temporal and Spatial Validation

The temporal consistency of predictions was examined across multiple growing seasons.

The model successfully reproduced inter-annual yield variability linked to climate fluctuations—particularly rainfall and temperature anomalies—suggesting that it effectively integrates climatic information into its predictions.

Spatially, yield maps generated by the CNN–GPR model (Figure 3) displayed clear alignment with known high- and low-yield zones reported by the Ministry of Agriculture.

High yields were concentrated in irrigated zones of Tashkent, Samarkand, and Fergana, while lower yields appeared in semi-arid regions of Khorezm and Karakalpakstan, reflecting real agronomic patterns.

The spatial smoothness of the predicted yield surfaces demonstrates the GPR layer's ability to model continuous yield gradients rather than discrete classification boundaries.

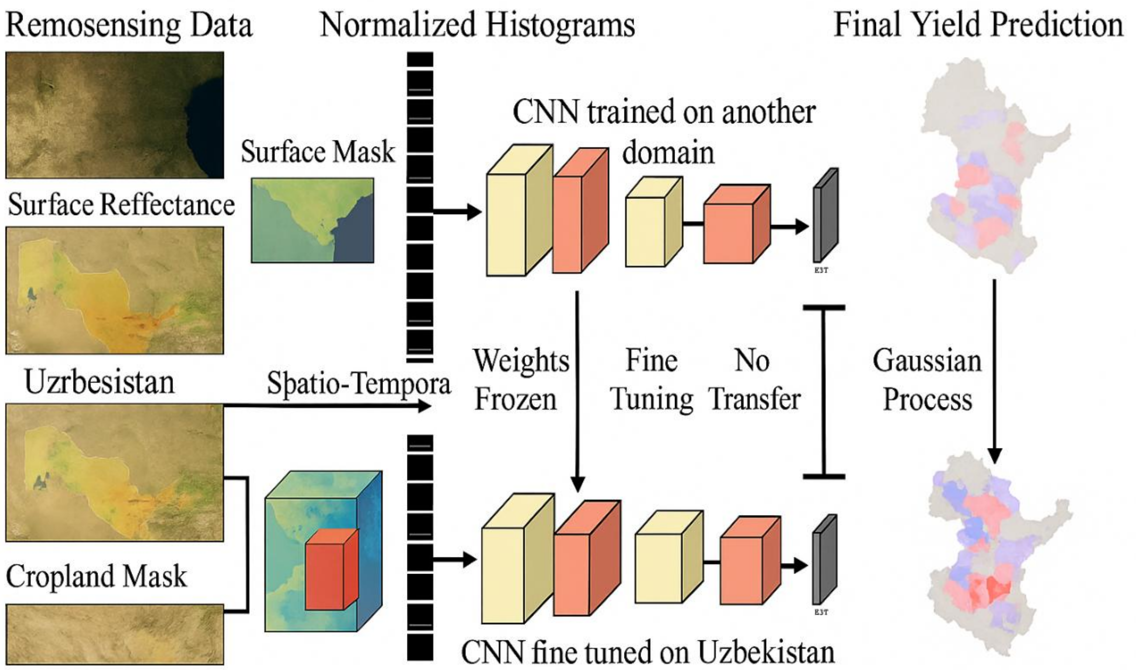

Figure 3. Predicted multi-crop yield map of Uzbekistan generated using the CNN–GPR transfer learning framework. Darker areas indicate higher yields, while lighter shades represent low-yield or high-uncertainty regions.

3.4. Uncertainty and Model Reliability

One of the most valuable aspects of integrating Gaussian Process Regression is its **ability to quantify uncertainty**.

For each predicted yield value, the GPR provided a standard deviation (σ_*) representing model confidence. Low uncertainty values (<0.3 t/ha) were observed in regions with abundant training data and consistent spectral conditions, while higher uncertainty (>0.6 t/ha) occurred in areas with sparse or noisy inputs. This probabilistic insight enables **risk-aware decision-making** in agricultural planning.

For example, regions with high yield variability or large uncertainty intervals can be prioritized for field monitoring, soil analysis, or irrigation optimization.

Such interpretability is essential for ensuring that AI-driven recommendations are both actionable and trustworthy in real-world agricultural management.

4. Conclusion

This study introduced a **deep transfer learning framework** that integrates **Convolutional Neural Networks (CNNs)** with **Gaussian Process Regression (GPR)** to predict crop yields under multi-crop and crop rotation conditions in Uzbekistan.

Unlike conventional models that are trained separately for each crop, the proposed approach leverages **knowledge transfer** from data-rich domains to data-scarce ones, enabling efficient and scalable yield forecasting across diverse agro-ecological environments.

The experimental results demonstrated that the CNN–GPR hybrid model achieved high predictive accuracy, with an average coefficient of determination ($R^2 = 0.91$) and a root mean squared error (RMSE=0.54 t/ha) across four major cereal crops — wheat, maize, barley, and rice.

By utilizing pretrained spectral–spatial representations and fine-tuning them for local conditions, the framework reduced training time by nearly 40% compared to models trained from scratch, while maintaining robust performance even in data-limited regions. The incorporation of Gaussian Process Regression further enhanced the model by introducing **probabilistic yield estimates**, allowing the quantification of uncertainty — a key advantage for practical decision support in agriculture.

From a methodological perspective, this research demonstrates the potential of **transfer learning as a generalization strategy** for AI-driven agricultural forecasting.

By combining deep learning with statistical inference, the model captures both the **complex nonlinear relationships** in satellite data and the **uncertainty inherent in real-world observations**.

The findings confirm that transfer learning can serve as a bridge between different crop types and geographic regions, paving the way toward global-scale, data-efficient agricultural intelligence systems.

In terms of practical relevance, the proposed framework can assist policymakers, agronomists, and farmers in optimizing crop planning, resource allocation, and climate adaptation strategies.

Its ability to handle heterogeneous data sources and predict yield with quantifiable confidence makes it suitable for integration into **national agricultural monitoring systems**, **early-warning platforms**, and **precision farming applications**.

By applying this framework in rotation-based systems, stakeholders can evaluate how sequential cropping patterns influence long-term productivity and sustainability.

Future research will focus on expanding the framework by integrating **satellite-based soil moisture indices, evapotranspiration data, and socio-economic variables** to improve prediction robustness. Additionally, incorporating **explainable AI (XAI)** techniques will enhance interpretability by visualizing which spectral and climatic factors most strongly influence yield outcomes. Ultimately, this work establishes a foundation for developing adaptive, transparent, and sustainable **AI-driven agricultural decision-support systems** tailored for semi-arid and data-scarce environments like Uzbekistan.

References:

1.Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

2.You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1), 4559–4566.

3.Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61–69. https://doi.org/10.1016/j.compag.2018.05.012

4.Russakovsky, O., Deng, J., Su, H., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211–252.

5.Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

6.Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780.

7.Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

8.Nuriev, K. K., Nuriev, M. K., Rakhmatov, O., & Rakhmatov, F. O. (2022, August). Comprehensive assessment of the degree of flooding of soil-cutting working bodies (on the example of plow shares). In IOP Conference Series: Earth and Environmental Science (Vol. 1076, No. 1, p. 012069). IOP Publishing.

9.Rahman, M. A., & Khosravi, A. (2021). Crop yield prediction using deep learning and Gaussian process regression: A hybrid approach. Agricultural Systems, 192, 103–161. https://doi.org/10.1016/j.agsy.2021.103161

10.Yuan, W., et al. (2022). Spatio-temporal deep learning for crop yield prediction using remote sensing data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 1535–1548.

11.Raxmatov, F. O., Raxmatov, O., Nuriev, K. K., & Nuriev, M. K. (2021, October). Combined dryer with high efficiency for drying high-moist agricultural products. In IOP Conference Series: Earth and Environmental Science (Vol. 868, No. 1, p. 012076). IOP Publishing.

12.Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621. https://doi.org/10.3389/fpls.2019.00621

13.Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13, 281–305.

14.Akhmadzhonova, G., Nazhmutdinova, D., Negmatshoeva, K., & Iroda, K. (2024). Assessment of the Microbial Flora of the Genital Tract and the Morphofunctional State of the Endometrium in Antiphospholipid Syndrome.

15.Рахматов, Ф. О., & Нуриев, К. К. (2022). Исследование плодов дыни как объекта технической переработки. Илмий мақолалар тўплами, 330.

16.Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

17.Dos Santos Ferreira, A., et al. (2021). Integrating deep learning and Gaussian processes for spatial prediction of crop yield from multispectral imagery. Remote Sensing of Environment, 258, 112397.

18.Нуриев, К. К., Рахматов, О., Кадирова, Р. С., & Рахматов, О. О. (2015). Биоконверсия органических отходов растительного происхождения в условиях Узбекистана. In Проблемы рекультивации отходов быта, промышленного и сельскохозяйственного производства (pp. 468-470).

19.Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.

20.Джураев, А. Ж., Нуриев, К. К., & Юсуфалиев, А. (2003). Разработка высокоресурсных лап для культиваторов. Тракторы и сельскохозяйственные машины, 2, 42-43.