MDPI

Using Landsat-5 for Accurate Historical LULC Classification: A Comparison of Machine Learning Models

Denis Krivoguz [1], Sergei G. Chernyi [2,3,*], Elena Zinchenko [2], Artem Silkin [2] and Anton Zinchenko [2]

1 Department of the "Oceanology", Southern Federal University, 340015 Rostov-on-Don, Russia; krivoguz@sfedu.ru
2 Department of Cyber-Physical Systems, St. Petersburg State Marine Technical University, Leninsky Prospect, 101, 198262 St. Petersburg, Russia; eltel85@bk.ru (E.Z.)
3 Department of Electrical Equipment of Ships and Automation of Production, Kerch State Maritime Technological University, 298309 Kerch, Russia
* Correspondence: sergiiblack@gmail.com

Abstract: This study investigates the application of various machine learning models for land use and land cover (LULC) classification in the Kerch Peninsula. The study utilizes archival field data, cadastral data, and published scientific literature for model training and testing, using Landsat-5 imagery from 1990 as input data. Four machine learning models (deep neural network, Random Forest, support vector machine (SVM), and AdaBoost) are employed, and their hyperparameters are tuned using random search and grid search. Model performance is evaluated through cross-validation and confusion matrices. The deep neural network achieves the highest accuracy (96.2%) and performs well in classifying water, urban lands, open soils, and high vegetation. However, it faces challenges in classifying grasslands, bare lands, and agricultural areas. The Random Forest model achieves an accuracy of 90.5% but struggles with differentiating high vegetation from agricultural lands. The SVM model achieves an accuracy of 86.1%, while the AdaBoost model performs the lowest with an accuracy of 58.4%. The novel contributions of this study include the comparison and evaluation of multiple machine learning models for land use classification in the Kerch Peninsula. The deep neural network and Random Forest models outperform SVM and AdaBoost in terms of accuracy. However, the use of limited data sources such as cadastral data and scientific articles may introduce limitations and potential errors. Future research should consider incorporating field studies and additional data sources for improved accuracy. This study provides valuable insights for land use classification, facilitating the assessment and management of natural resources in the Kerch Peninsula. The findings contribute to informed decision-making processes and lay the groundwork for further research in the field.

Keywords: machine learning; LULC; Landsat; classification

1. Introduction

Land Use and Land Cover (LULC) is a concept that describes the utilization and coverage of the Earth's surface [1]. Understanding the state and changes of LULC is critically important for sustainable management of natural resources and environmental protection [2]. In recent years, satellite imagery, aerial photography, and geographic information systems (GIS) have been widely used for LULC analysis. The use of machine learning for LULC classification based on this data is an important tool for LULC analysis and resource management [3]. Knowledge of LULC is also important for decision-making in various fields such as urban planning, agriculture, forestry, ecology, and land-use planning [4]. Determining LULC based on satellite data and GIS analysis allows for an overall picture of land use in a region and for analyzing changes in LULC over time [5]. Such data can be used to identify trends in regional development, plan land use, assess environmental consequences, and monitor changes in the natural environment.

With the rapid development of technology in remote sensing, GIS, and machine learning, LULC analysis is becoming more accurate and effective. Based on the results of such analysis, more precise strategies for land-use management can be developed, which is crucial for sustainable regional development [6].

LULC analysis in arid ecosystems is critically important for understanding and predicting changes in these ecosystems [7,8]. Arid ecosystems are vulnerable ecosystems that are under pressure from human activities, such as land-use change, deforestation, increased grazing lands, and climate change [9]. Changes in LULC can lead to serious consequences for arid ecosystems, such as soil fertility reduction, soil erosion, biodiversity loss, and deterioration of the quality of life of local communities. Therefore, to manage and preserve arid ecosystems, it is necessary to have an accurate understanding of the state of LULC [10].

Remote sensing plays an important role in LULC analysis. Remote sensing allows obtaining information about the Earth's surface without physical contact with it [5,11]. For LULC analysis, remote sensing is used to obtain spectral data, such as radio emissions, light reflection, and heat, from which information about the composition of the Earth's surface and its changes can be obtained [12].

With remote sensing, information about soil types, vegetation cover, land use, the presence of water bodies, and much more can be obtained. These data allow for the evaluation of changes in LULC in a specific area over time, which can be used to predict future ecosystem changes. Thus, remote sensing becomes a necessary tool for planning and managing land use in arid ecosystems, where LULC can have a significant impact on water balance and soil cover.

The search for optimal ways to classify LULC most accurately has been considered in several studies. Jozdani [13] carried out an experimental comparison among different architectures of DNNs (i.e., regular deep multilayer perceptron, regular autoencoder, sparse, autoencoder, variational autoencoder, convolutional neural networks), common ensemble algorithms (Random Forests, Bagging Trees, Gradient Boosting Trees, and Extreme Gradient Boosting), and SVM to investigate their potential for urban mapping using a GEOBIA approach. Jamali [14] aim to evaluate eight machine learning algorithms for image classification implemented in WEKA and R programming language. The aim of [15] was to compare performance of the classification methods, that are Rule Based classifier and Support Vector Machine, of Planetscope and Worldview-3 satellite images in order to produce land use/cover thematic maps. Six machine-learning algorithms, namely random forest, support vector machine, artificial neural network, fuzzy adaptive resonance theory-supervised predictive mapping, spectral angle mapper and Mahalanobis distance were examined in [3]. Ghayour [16] used different kernel functions and hidden layers for SVM and ANN algorithms, respectively. In this objective of [17] at finding out how two composition methods and spectral–temporal metrics extracted from satellite time series can affect the ability of a machine learning classifier to produce accurate LULC maps. For the objective of [18], two datasets were collected at two different urban locations using two different UASs. Basheer [19] aim to evaluate the LULC classification performance of two commonly used platforms (i.e., ArcGIS Pro and Google Earth Engine) with different satellite datasets (i.e., Landsat, Sentinel, and Planet) through a case study for the city of Charlottetown in Canada.

From the other side, classification of historical LULC data was reviewed in several studies. Aim of [20] were to produce historical LULC maps during the 1988–2016 period for spatial and temporal analysis, forecast future LULC until 2040 by using the Markov model, and identify the impact of LULC on urbanization. Drummond [21] describe historical land-use and land-cover (LULC) maps for the northern Colorado urban Front Range. Hoque [22] design four land use/land cover (LULC) scenarios, such as business-as-usual development (BAUD), economic development priority (EDP), ecological protection priority (EPP), and afforestation development priority (ADP), through a Cellular Automata-Markov (CA-Markov) model, and their effects on ecosystem service values (ESVs) were predicted, using historical LULC maps and ESV coefficients of the Lower Meghna River Estuary, Bangladesh.

Hufkens [23] use a combination of historical (1958) aerial photography and contemporary remote sensing data to map long-term changes in the extent and structure of the tropical forest surrounding Yangambi (DR Congo) in the central Congo Basin. Yao [24] demonstrate an innovative learning method of convolutional neural network (CNN) to identify landuse and land cover (LULC) patterns and extract features to disaggregate socio-economic factors by using remote sensing imageries at 30 m spatial resolution. The contribution of Leta [25] was to assess the temporal and spatial LULC dynamics of the past and to predict the future using Landsat images and LCM (Land Change Modeler) by considering the drivers of LULC dynamics. The aim of Firozjaei [26] were to evaluate the historical impacts of mining activities on surface biophysical characteristics, and for the first time, to predict the future changes in pattern of vegetation cover and land surface temperature (LST). The transition potential maps and the transition probability matrices between LULC types were provided by the support vector machine algorithm and the Markov chain model, respectively, to project the 2021 and 2040 LULC maps by Jalayer [27]). As a case study Mäyrä [28] use U-Net to automatically extract fields, mires, roads, watercourses, and water bodies from scanned historical maps, dated 1965, 1984 and 1985 for the 900 km study area in Southern Finland.

The primary objective of this research is to investigate the historical classification of land use and land cover (LULC) in the absence of comprehensive field studies, relying solely on limited data sources such as topographic maps, scientific articles, and cadastral data. This study aims to assess the feasibility and reliability of utilizing these data sources for LULC classification and understanding the temporal dynamics of land use patterns.

The novelty of this paper lies in its innovative approach to historical land use and land cover (LULC) classification, particularly in the context of limited data sources and the absence of extensive field studies. While existing research often relies heavily on comprehensive field data, this study explores the feasibility and reliability of utilizing alternative sources such as topographic maps, scientific articles, and cadastral data for LULC classification.

Unlike many previous studies that prioritize high-resolution remote sensing data and extensive field surveys, this research ventures into a less-explored territory by demonstrating that valuable insights can still be extracted from limited data sources. By doing so, it challenges the traditional norms of LULC classification methodologies and opens up new possibilities for regions or scenarios where field studies may be constrained.

Moreover, this study contributes to the broader scientific discourse by addressing the nuanced complexities associated with historical LULC patterns. The comparison of classified results with existing historical records and alternative data sources offers a novel perspective on the reliability and quality of classification outcomes. By meticulously examining the potential errors and limitations arising from the absence of extensive field research, the paper provides a valuable guide for researchers and practitioners seeking to embark on similar historical LULC studies.

2. Materials and Methods

2.1. Research Area

The Kerch Peninsula (Figure 1) is located on the southeastern coast of the Sea of Azov and the northeastern coast of the Black Sea, within the Crimean Peninsula [29]. The climate in this region is moderately continental, with an average temperature of around 11 degrees Celsius per year, and the most favorable period for research in this area is spring and autumn [30]. The territory of the Kerch Peninsula has many geological, hydrological, landscape, and ecological peculiarities that may affect the quality of remote sensing data and the results of land use and land cover (LULC) classification [31]. For example, there are several mud lakes in the area that may affect land zoning and vegetation distribution. Moreover, different types of landscapes, such as steppes, forest steppes, rocky outcrops, and sand dunes, can also affect the results of LULC classification [32].

Figure 1. Study area location.

The Kerch Peninsula was chosen for the study for several reasons. Firstly, it is a region with a unique ecological situation located on the border of three natural zones: forest, steppe, and subtropical. This leads to a high diversity of plant and animal species. Secondly, the Kerch Peninsula has problems related to unauthorized land use. For example, in some areas, illegal deforestation occurs, leading to worsening environmental conditions. Therefore, studying the state of land cover in this peninsula may help identify problem areas and develop measures to address them. Thirdly, the Kerch Peninsula is essential for tourism development in the region. It is famous for its landscapes, artificial and natural attractions, as well as unique flora and fauna. Studying the state of land cover will help identify the territories that are most suitable for tourist use and form recommendations for nature conservation in these areas.

2.2. Description of Algorithms Used in the Research

In this research, several machine learning algorithms were employed for land use and land cover classification in the Kerch Peninsula. The algorithms utilized in the study encompassed diverse approaches to effectively discern and categorize different land use and land cover classes based on limited data resources.

Deep Neural Network (DNN). This algorithm is a powerful tool for pattern recognition and classification tasks [9,33–35]. It comprises multiple layers of interconnected neurons that allow for non-linear feature extraction and representation [36,37]. The DNN utilized

five layers with 128, 64, 32, 16, and 8 neurons, respectively, along with the Rectified Linear Unit (ReLU) activation function to introduce non-linearity. The Adam optimizer with a learning rate of 0.001 was employed for efficient weight updating. Dropout regularization (rate: 0.5) was used to prevent overfitting.

Random Forest. This ensemble learning technique constructs multiple decision trees during the training process and outputs the mode of their predictions [38–40]. The algorithm utilized 100 trees in the forest, each with a maximum depth of 20 nodes. The minimum number of samples required to split an internal node was set to 5, and the minimum number of samples required to be at a leaf node was set to 2.

Support Vector Machine (SVM). SVM is a supervised learning algorithm used for classification and regression tasks [41–43]. The Radial Basis Function (RBF) kernel with a parameter value of 0.1 was employed. The regularization parameter (C) was set to 100 to control the trade-off between maximizing the margin and minimizing the classification error.

AdaBoost. This boosting algorithm combines multiple weak learners (e.g., decision trees) sequentially to build a strong model. The study utilized 50 base models, each being a decision tree with a depth of 2.

2.3. Classification Performance Evaluation

To evaluate the performance of a classification model, metrics such as accuracy, precision, recall, and F1-score can be used [44,45]. Accuracy is the proportion of correctly classified objects and is calculated using the formula:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision is the proportion of objects that were correctly classified as positive relative to all objects that were classified as positive:

$$\text{precision} = \frac{TP}{TP + FP}$$

Recall is the proportion of objects that were correctly classified as positive relative to all objects that are positive:

$$\text{recall} = \frac{TP}{TP + FN}$$

The F1-score is the harmonic mean between precision and recall and can be calculated using the following formula:

$$F1 = 2 \cdot \frac{\text{precision} - \text{recall}}{\text{precision} + \text{recall}}$$

where TP is the number of true positive classifications, TN is the number of true negative classifications, FP is the number of false positive classifications, and FN is the number of false negative classifications.

2.4. Landsat Data

Landsat-5 is an American satellite system developed and managed by NASA, which was launched in 1984 for monitoring the Earth's surface. The satellite is equipped with instruments for capturing images in a wide range of spectral wavelengths, including visible, infrared, and thermal ranges. Landsat-5 data has been used for monitoring changes in the Earth's surface, including forests, water bodies, mountain ranges, and other natural and man-made objects. Landsat-5 images provide information on soil quality, vegetation, and other surface aspects, making them very important for research related to changes in the Earth's surface.

Landsat-5 data was obtained using Google Earth Engine and the geemap library [46], which allow for fast and efficient processing and analysis of numerous images from various sources for a wide range of scientific tasks, including LULCC analysis.

The characteristics of Landsat-5 data include a pixel resolution of 30 m and a maximum dynamic range value of 255. An important feature of Landsat-5 data is their ability to be reused multiple times, as the satellite continuously scans the Earth's surface and records data at different points in time.

Data about channels are useful for image analysis and processing, as each channel contains information about different characteristics of the Earth's surface. Specifically, channels 1–3 correspond to the visible spectrum, while channels 4–7 correspond to the infrared spectrum. They also differ in resolution, which allows them to be used for various image processing tasks (Table 1).

Table 1. Technical characteristics of the Landsat-5 satellite bands.

Band	Name	Wavelength Center, nm	Resolution, m
1	Blue	0.45–0.52	30
2	Green	0.52–0.60	30
3	Red	0.63–0.69	30
4	Near Infrared	0.76–0.90	30
5	Shortwave Infrared 1	1.55–1.75	30
6	Shortwave Infrared 2	2.08–2.35	30
7	Mid Infrared	10.40–12.50	60

Taken together, Landsat-5 data represent a valuable source of information for analyzing changes in the Earth's surface, including land use classification and changes in vegetation cover, making them an important tool for many scientific and practical applications.

Landsat data were collected using geemap package [46] as a median values over 1990–1994 period for the Kerch peninsula.

2.5. Data Collection

Various sources of data were utilized in this study, including cadastral data and materials published in scientific literature. Landsat-5 images acquired during the period from 1990 to 1994 were also used to classify the territory. However, field data that could serve as a basis for analyzing land use and land cover changes in the Kerch Peninsula were not available at that time.

Therefore, the use of cadastral data and materials published in scientific literature was the only available method for obtaining information about the territory and its changes. In addition, archival cartographic data from topographic maps by the "Genshtab" (Soviet topographic maps) for 1990–1991 [47] were used. Various archival high spatial resolution satellite data were also used for manual classification of LULC samples using Google Earth Pro Historical Imagery, which allowed for the creation of training and testing datasets for subsequent classification of the Kerch Peninsula territory.

As a result, 12580 different sample points were obtained, which were assigned LULC class labels used in this study (Figures 2 and 3). They were further divided into training and testing sets in a 7:3 ratio, which corresponded to 8387 points for the training set and 4193 for the testing set.
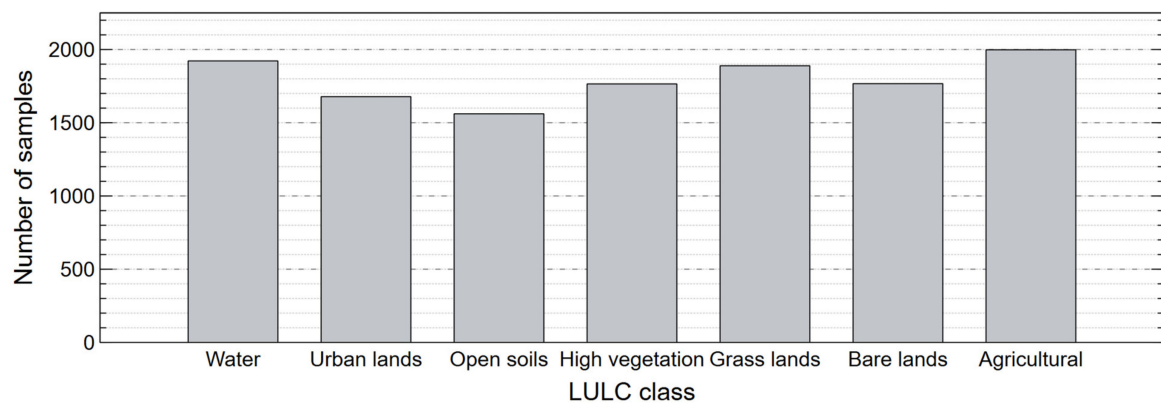
Figure 2. Distribution of sample points for the training and testing datasets for each land use class.
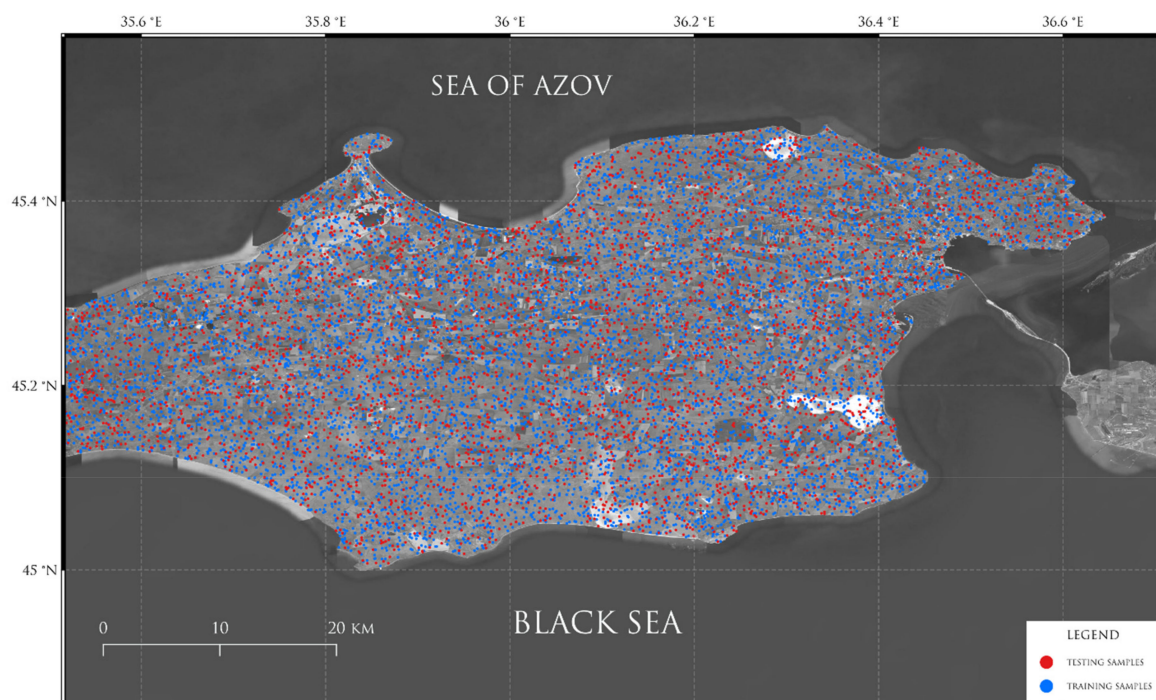
Figure 3. Spatial distribution for the training and testing datasets.

2.6. LULC Classes

In this study, 7 LULC classes were used, which included Agricultural areas, Water, Bare vegetation, Herbaceous vegetation, Forest vegetation, Urban areas, and Bare lands (Table 2).

Table 2. Short description of LULC classes.

Class	Description
Water	Areas covered by water bodies such as lakes, rivers, and reservoirs.
Urban lands	Developed areas characterized by buildings, infrastructure, and human settlements.
Open soils	Areas of exposed soil or bare land without significant vegetation cover.
High vegetation	Regions with dense and thriving vegetation, such as forests, woodlands, or dense vegetation cover.

Table 2. *Cont.*

Class	Description
Grass lands	Areas dominated by grasses and other herbaceous plants, often used for grazing or agricultural purposes.
Bare lands	Land devoid of vegetation cover, including areas with minimal or no soil and exposed rock surfaces.
Agricultural	Land utilized for agricultural activities, including crop cultivation, farming, or livestock rearing.

3. Results

3.1. Learning Configuration

In this study, the use of various machine learning models for land use and land cover classification in the Kerch Peninsula territory was investigated. Archive field data from 1990, cadastral data, and materials published in scientific literature were used for model training and testing. Landsat-5 images for 1990 provided by USGS were used as the input data.

The system configuration for model training included an ASUS laptop with an Intel i7-11370H processor (4 cores, 3.30 GHz), 16 GB DDR4 RAM, and an NVIDIA GeForce RTX 3050 graphics card (4 GB). The following Python libraries were used for data processing and model training: NumPy 1.19.5, pandas 1.3.3, scikit-learn 1.0, seaborn 0.11.2, and Veusz 3.5.1. All models were trained in Python version 3.9.5.

3.2. Hyperparameter Tuning

In this study, hyperparameter tuning strategies were employed to optimize the performance of machine learning models for land use and land cover classification in the Kerch Peninsula (Table 3). The objective was to enhance the accuracy of the models and achieve superior results on the test dataset. Four popular machine learning algorithms were used: deep neural network (DNN), Random Forest, support vector machine (SVM), and AdaBoost.

Table 3. Hyperparameters for machine learning models used in the study.

Model	Hyperparameters
Deep Neural Network	Number of layers: 5 Number of neurons in each layer: 128, 64, 32, 16, 8 Activation function: ReLU Optimizer: Adam Learning rate: 0.001 Regularization: Dropout (0.5)
Random Forest	Number of trees: 100 Maximum depth of trees: 20 Minimum number of samples required to split an internal node: 5 Minimum number of samples required to be at a leaf node: 2
Support Vector Machine (SVM)	Kernel type: RBF Kernel parameter: 0.1 Regularization parameter: 100
AdaBoost	Number of base models: 50 Type of base model: Decision Tree Depth of trees: 2

The DNN model underwent extensive hyperparameter tuning by varying the number of layers, number of neurons in each layer, learning rate, regularization coefficients, and activation functions. The tuning process utilized a combination of random search and

grid search methods, which allowed the hyperparameter space to be efficiently explored. Cross-validation was employed to evaluate the model's performance during the tuning process, ensuring a robust assessment.

Similarly, the Random Forest model had its hyperparameters tuned on two crucial parameters: the number of trees in the forest and the maximum depth of the trees. Grid search was used to systematically explore different combinations, and the hyperparameter values that yielded the highest accuracy on the test dataset were selected.

For the SVM model, hyperparameter tuning was conducted on the regularization parameter C and the kernel gamma parameter. Grid search was employed to assess various combinations and determine the optimal hyperparameter values that led to the highest classification accuracy.

Lastly, the AdaBoost model had its focus on tuning the number of base models and the learning rate coefficient. Grid search was utilized to assess different combinations, and the hyperparameter values that resulted in the best accuracy on the test dataset were determined.

Throughout the hyperparameter tuning process, rigorous cross-validation techniques were employed to prevent overfitting and ensure that the selected hyperparameter values generalized well to unseen data.

Hyperparameter tuning is an important step in machine learning model training that allows improving the model's quality and achieving better results on test data. The following machine learning methods were used for the LULC classification task: deep neural network, Random Forest, adaboost, and SVM.

3.3. Results of LULC Classification Using Landsat-5 Data

The classification of LULC using Landsat-5 data was performed using four machine learning models: deep neural network, random forest, support vector machine, and AdaBoost. The input data consisted of multispectral imagery. As the result, 4 different LULC maps were produced (Figure 4).
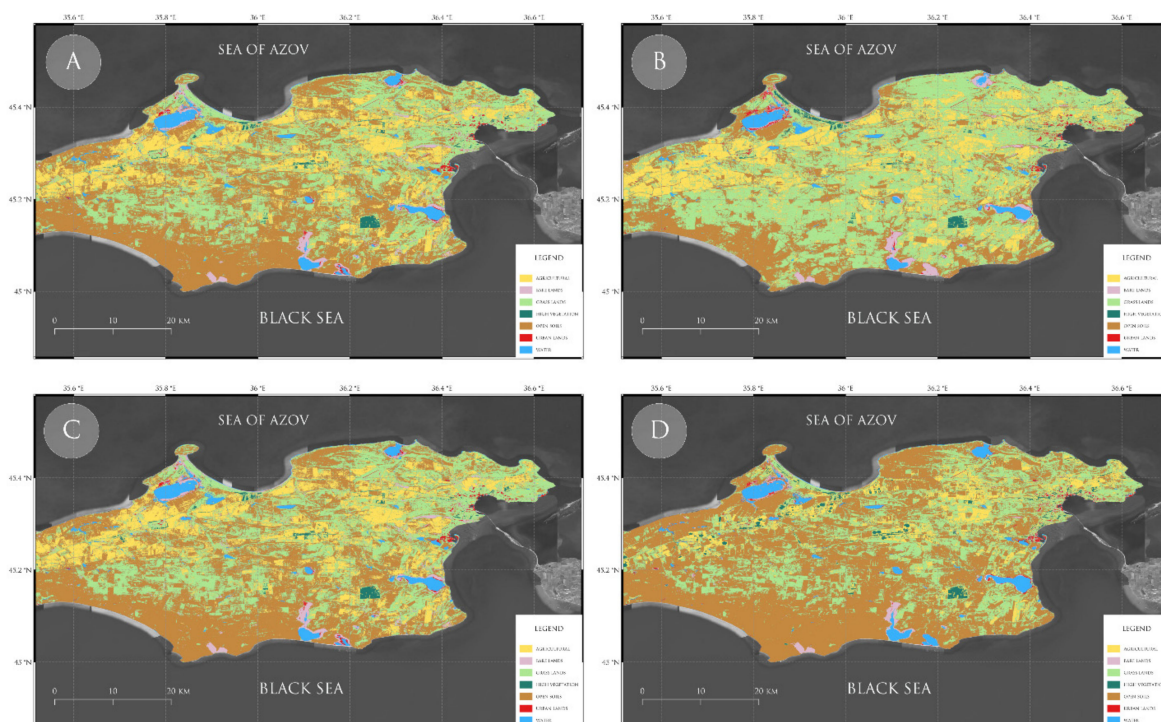
Figure 4. Results of LULC classification using Landsat-5 data for the Kerch Peninsula (1990–1994). (**A**)—DNN, (**B**)—Random Forest, (**C**)—Support Vector Machine, (**D**)—Adaboost.

Analyzing the confusion matrices for each of the models, the following conclusions can be drawn. The deep neural network showed (Table 4) high accuracy in classifying Water, Urban lands, Open soils, and High vegetation, with 0.99, 0.95, 0.95, and 0.97 accuracy, respectively. However, the classification of Grass lands, Bare lands, and Agricultural was less accurate, with 0.84, 0.87, and 0.89 accuracy, respectively. The greatest number of errors was made in the classification of Grass lands, where 287 pixels were misclassified. Overall, it can be noted that the neural network performed well in classifying diverse land use types.

Table 4. Confusion matrix for deep neural network.

	Water	Urban Lands	Open Soils	High Vegetation	Grass Lands	BARE LANDS	Agricultural
Water	1871	2	5	1	19	13	11
Urban lands	2	1643	22	0	5	6	0
Open soils	8	100	1442	0	6	5	0
High vegetation	1	0	0	1758	2	2	2
Grass lands	10	5	1	1	1860	10	2
Bare lands	8	3	3	1	14	1715	23
Agricultural	13	0	2	1	26	54	1902

Random Forest demonstrated (Table 5) high accuracy in classifying Water, Urban lands, and Open soils, with accuracies of 0.99, 0.93, and 0.90, respectively. However, significant errors were made in classifying High vegetation and Grass lands, with accuracies of 0.75 and 0.68, respectively. The classification of Bare lands and Agricultural was also not precise, with accuracies of 0.86 and 0.74, respectively. It can be noted that the low accuracy in classifying Grass lands may be related to the difficulty in distinguishing this class from High vegetation and Bare lands, as well as the lack of field research that could help in the precise identification of these classes.

Table 5. Confusion matrix for Random Forest.

	Water	Urban Lands	Open Soils	High Vegetation	Grass Lands	Bare Lands	Agricultural
Water	1844	2	3	18	0	2	53
Urban lands	0	1581	47	8	0	22	20
Open soils	7	69	1391	68	2	21	3
High vegetation	9	17	27	1706	2	3	1
Grass lands	0	6	6	6	1854	16	1
Bare lands	2	3	17	17	14	1701	13
Agricultural	50	10	0	1	1	14	1922

SVM showed (Table 6) the worst results with an accuracy of 0.86. The classification of Urban lands was the most accurate, with an accuracy of 0.95, while the classification of Grass lands was the least accurate, with an accuracy of 0.59. The low accuracy in the classification of Grass lands, as well as Bare lands and Agricultural, may be related to the difficulty in distinguishing these classes from High vegetation and Urban lands, as well as the lack of field research for a more accurate determination of these classes.

Table 6. Confusion matrix for SVM.

	Water	Urban Lands	Open Soils	High Vegetation	Grass Lands	Bare Lands	Agricultural
Water	1910	0	0	1	0	3	8
Urban lands	18	824	57	28	95	107	549
Open soils	6	156	650	48	116	391	194
High vegetation	1	25	27	1583	0	36	93
Grass lands	0	46	38	21	1597	68	79
Bare lands	7	137	142	20	157	1179	125
Agricultural	10	85	52	19	123	57	1662

AdaBoost showed (Table 7) the worst results with an accuracy of 0.58. The classification of Water was the most accurate with an accuracy of 0.96, while the classification of Agricultural was the least accurate with an accuracy of 0.26. Most of the classes were misclassified, which may be due to the insufficient depth of the trees used in the model as well as the lack of data for training the model.

Table 7. Confusion matrix for AdaBoost.

	Water	Urban Lands	Open Soils	High Vegetation	Grass Lands	Bare Lands	Agricultural
Water	1123	68	140	191	223	173	4
Urban lands	140	831	23	107	318	258	1
Open soils	84	9	827	214	135	251	46
High vegetation	22	2	122	1303	52	242	22
Grass lands	129	70	171	212	966	523	18
Bare lands	9	39	27	125	150	1414	3
Agricultural	36	10	106	225	97	46	1478

As a result, the deep neural network showed the best accuracy in LULC classification, reaching 96.2% accuracy, which means high model efficiency. In the confusion matrix, some false classifications can be noticed urban lands were incorrectly assigned to lands with high vegetation, which can be explained by a lack of field research. However, the overall accuracy of the model indicates that it is still effective in LULC classification.

Random Forest also showed good accuracy with 90.5%, but as in the previous model, false classifications were observed. For example, lands occupied by agriculture were incorrectly assigned to lands with high vegetation.

The SVM model showed an accuracy of 86.1%, which is less efficient than the previous models. The confusion matrix also showed some false classifications, especially in classes of agricultural lands and lands with high vegetation. As in previous cases, this may be due to classification being performed on cadastre data, map data, and scientific articles, which could significantly reduce accuracy.

Finally, the AdaBoost model showed the lowest accuracy of 58.4%, indicating that it is less effective in LULC classification than other models. The confusion matrix also showed a significant number of false classifications, including the incorrect identification of lands with high vegetation and lands occupied by agriculture. However, it is necessary to consider the lack of field research and the limitations in the use of cadastre data, map data, and scientific articles, which can reduce model accuracy.

Based on the analysis of the confusion matrices, it can be concluded that the most accurate model is the deep neural network, which is confirmed by its high overall accuracy (0.962). Among the other models, Random Forest showed the highest accuracy (0.905), while SVM and AdaBoost showed lower accuracy (0.861 and 0.584, respectively).

Overall, a high level of accuracy in land use classification is observed for all land use classes. However, there are also certain errors that may be related to the use of only cadastral data, map data, and scientific articles as the source data for classification, as well as the lack of field studies. For example, in the error matrix for the deep neural network, we can observe some false classifications into the „Open soils" class, which may be confused with the "Bare lands" class due to their visual similarity.

Thus, it can be concluded that the use of various machine learning models allows for achieving a high accuracy in land use classification based on cadastral data, map data, and scientific articles. However, for more accurate results, field studies may be required, as well as the use of more diverse data sources.

The Table 8 shows the accuracy, precision, recall, and F1 metrics for each of the LULC classification models. The most accurate model was the deep neural network, which had the highest accuracy (0.962) and high precision, recall, and F1 metrics for all classes. The AdaBoost model also demonstrated high metric values for most classes but showed significantly lower accuracy for the "Bare lands" class. The SVM model has low accuracy for most classes, especially for the "Urban lands" and "High vegetation" classes and is not the best choice for this task. The random forest model demonstrates good accuracy for most classes but has low precision, recall, and F1 metrics for the "Water" and "High vegetation" classes.

Table 8. Accuracy metrics, calculated for each machine learning models.

Model	Accuracy	Precision	Recall	F1
SVM	0.821	0.763	0.556	0.643
DNN	0.962	0.939	0.904	0.921
AdaBoost	0.655	0.710	0.564	0.628
Random Forest	0.814	0.725	0.751	0.737

Given that the research was conducted on the Kerch Peninsula, it can be concluded that the terrain represents a typical sample of landscapes of the Black Sea coast. The largest area is occupied by bare lands (1229 km^2), which are characterized by a low level of vegetation and used for agricultural purposes. Other large land use classes are meadows and pastures (902 km^2) and lands used for agriculture (555 km^2). A small area is occupied by bare lands (85 km^2) and water bodies (81 km^2). Urban areas occupy the smallest area (27 km^2).

The area studied in this work represents a mixture of urban and rural areas, with a predominance of open land areas (1229 km^2). It is likely that industrial facilities, as well as infrastructure such as roads, airports, recreation areas, etc., are in these areas.

A significant part of the territory is occupied by land use associated with agriculture. Agricultural land is located on an area of 555 km^2. This may indicate that this area plays an important role in supplying cities and settlements with food. Open lands and agricultural lands are usually the object of environmental protection and are important for maintaining biodiversity. However, additional research is needed to determine the degree of human impact on the environment and the presence of potential threats to its conservation.

The accuracy of classification of different models may be related to their ability to process different types of data and noise in the data, as well as their parameters and settings. For example, a deep neural network may be more effective in classifying complex non-linear patterns, while SVM may perform better with linearly separable data. The choice of model and its settings should be based on the specific task and the characteristics of the data. Random forest, SVM, and deep neural network have shown higher accuracy than AdaBoost. This may be due to AdaBoost using datasets that do not meet the assumptions on which the AdaBoost method is based, such as the requirement that the samples are independent and identically distributed, otherwise overfitting may occur [48].

Differences between SVM and AdaBoost may be related to their learning algorithms and model parameters. For example, SVM tries to find the hyperplane that best separates

the classes using a certain distance measure. If the distance between the classes is large, SVM may not achieve high classification accuracy. AdaBoost, on the other hand, uses weighted voting to combine several weak models into one strong model. If inappropriate weak models are selected, this can lead to a decrease in the accuracy of the entire model.

4. Discussion

4.1. Possible Reasons of Classification Results

The superior performance of the deep neural network model with an accuracy of 96.2% can be attributed to its ability to capture complex non-linear patterns in the input data. The DNN's multiple layers and neurons enable it to learn intricate spatial relationships and representations, making it well-suited for the classification of diverse land use types, such as water bodies, urban areas, open soils, and high vegetation.

On the other hand, the Random Forest model achieved an accuracy of 90.5%, showcasing its capacity to handle many trees in the forest and make use of feature importance. However, it faced challenges in distinguishing high vegetation from agricultural lands, which may be attributed to the visual similarities between these classes and the lack of detailed spectral information.

The support vector machine model demonstrated an accuracy of 86.1%, reflecting its effectiveness in dealing with linearly separable data. Nonetheless, it encountered difficulties in accurately classifying certain land use types, particularly agricultural and high vegetation areas. This might be due to the SVM's reliance on the selection of an appropriate kernel function and regularization parameter, which may not be optimal for this specific dataset.

In contrast, the AdaBoost model exhibited the lowest accuracy of 58.4%, indicating that it struggled to achieve robust classifications. The limitations in AdaBoost's performance may be associated with its sensitivity to noisy data and misclassified samples, leading to the suboptimal combination of weak classifiers.

The observed discrepancies in accuracy can also be attributed to the training data sources, which predominantly relied on cadastral data, map data, and scientific articles. The lack of comprehensive field studies may have hindered the accurate identification of certain land use classes, particularly grasslands, bare lands, and agricultural areas, leading to misclassifications in some instances.

Furthermore, the choice of hyperparameters during the tuning process significantly impacted the performance of the models. The successful hyperparameter selection for the DNN, Random Forest, SVM, and AdaBoost models contributed to their respective accuracies, underscoring the importance of rigorous hyperparameter tuning for optimal model performance.

4.2. Possible Causes of Misclassification

One of the primary causes of misclassification can be attributed to the limited scope of data sources used in this study. The reliance on cadastral data, map data, and scientific articles, while providing valuable insights, may not fully capture the complexity and variability of the terrain. Field research, which was not extensively incorporated, could have facilitated a more accurate identification of certain land use classes, especially those with visual similarities such as Grass lands and High vegetation.

Another potential reason for misclassification lies in the nature of the terrain and the unique challenges it presents. The Kerch Peninsula encompasses diverse landscapes with intricate land use patterns, making it difficult for certain models to differentiate between similar classes accurately. For instance, the misclassification of lands occupied by agriculture into High vegetation or vice versa could be attributed to the limited spectral resolution of the Landsat-5 imagery used as input data.

It is essential to acknowledge that, despite the high accuracy achieved by the deep neural network and Random Forest models, misclassifications still occur. These inaccuracies

could be due to the complexity of the classification task and the inherent difficulty in discerning certain land use categories from the available data sources.

4.3. Limitations and Assumptions

The investigation of land use and land cover classification in the Kerch Peninsula using machine learning models has provided valuable insights. However, it is essential to address the limitations and assumptions that might have influenced the study's outcomes.

One significant constraint in this study was the limited availability of data sources. Relying solely on cadastral data, topographic maps, and scientific articles may have introduced inaccuracies due to outdated or incomplete information. Field research data were absent, leading to potential uncertainties in ground truthing and class definitions.

The use of Landsat-5 imagery with a moderate spatial resolution could have affected the accuracy of land cover classification. Some land use classes, especially those with subtle spectral differences, might have been challenging to distinguish accurately.

The quality and accuracy of the input data, such as cadastral records and topographic maps, may have varied across different regions within the study area. Inconsistencies and errors in these datasets could have impacted the overall classification performance.

The choice of machine learning algorithms was based on their commonly used applications but might not be the optimal choice for the specific characteristics of the study area. Other models not considered in this study could potentially yield better results.

Despite these limitations, this study successfully demonstrated the efficacy of deep neural network and Random Forest models for land use and land cover classification in the Kerch Peninsula.

4.4. Future Research and Improvements

In light of the findings and limitations of this study, several ways for future research and improvement can be proposed. Firstly, addressing the issue of limited data sources is crucial for enhancing the accuracy of land use and land cover classification. Incorporating high-resolution remote sensing imagery, such as aerial photographs or satellite data with finer spatial resolution, can provide more detailed information about the landscape, enabling better discrimination between land use classes.

Secondly, conducting comprehensive field studies to validate the classified results would be beneficial. Field surveys would allow for ground-truthing and validation of the model's outputs, reducing misclassifications and improving the overall reliability of the classification results.

Furthermore, exploring the application of ensemble methods, where multiple machine learning models are combined, could lead to improved classification performance. Ensemble methods, such as stacking or blending different classifiers, have been shown to enhance predictive accuracy by leveraging the strengths of individual models and compensating for their weaknesses.

To enhance the generalization ability of the models, transfer learning techniques could be investigated. By leveraging knowledge from pre-trained models on related tasks or domains, transfer learning can help overcome the data scarcity issue and improve the model's ability to classify land use and land cover classes accurately.

Moreover, considering temporal data, such as multi-temporal satellite images, can provide valuable insights into land use dynamics over time. Analyzing land cover changes over different periods can aid in understanding long-term trends and supporting decision-making processes related to land use planning and environmental conservation.

Lastly, investigating the influence of feature engineering on model performance can lead to better feature representations and, subsequently, improved classification accuracy. Utilizing domain knowledge to engineer relevant features, such as vegetation indices, texture measures, or topographic attributes, can enhance the discriminative power of the models.

5. Conclusions

This study investigated the application of various machine learning models for land use and land cover classification in the Kerch Peninsula. The objective was to explore the feasibility of using limited data sources, such as topographic maps, scientific articles, and cadastral data, in the absence of extensive field research. The research aimed to evaluate the performance of four machine learning algorithms: deep neural network (DNN), Random Forest, support vector machine (SVM), and AdaBoost, in accurately classifying different land use and land cover types.

The findings revealed that the deep neural network model exhibited the highest accuracy, achieving an impressive 96.2% accuracy in land use classification. The DNN model demonstrated exceptional performance in classifying water, urban lands, open soils, and high vegetation. However, it faced challenges in accurately classifying grasslands, bare lands, and agricultural areas, which may be attributed to visual similarities between these classes and the lack of specific field data.

The Random Forest model also demonstrated competitive performance with an accuracy of 90.5%. Nevertheless, it encountered difficulties in distinguishing high vegetation from agricultural lands. The SVM model achieved an accuracy of 86.1%, while the AdaBoost model performed the lowest with an accuracy of 58.4%.

The hyperparameter tuning process played a critical role in improving the performance of the machine learning models. By systematically optimizing the hyperparameters using grid search and cross-validation, the models were fine-tuned to achieve their best accuracy on the test dataset.

Overall, this study's contributions lie in the comprehensive comparison and evaluation of multiple machine learning models for land use classification in the Kerch Peninsula. The research demonstrates the potential of utilizing limited data sources, which are more accessible in certain regions, for land use mapping and environmental management. However, it is essential to acknowledge the limitations introduced by the absence of detailed field research, potentially leading to some misclassifications and reduced accuracy in certain classes.

Author Contributions: Conceptualization, D.K. and S.G.C.; methodology, E.Z.; software, A.S.; validation, A.Z., A.S. and D.K.; formal analysis, E.Z., A.Z., A.S., D.K. and S.G.C.; investigation, D.K.; resources, A.S.; data curation, S.G.C.; writing—original draft preparation, A.Z.; visualization, D.K.; funding acquisition, A.S. and E.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmad, F.; Goparaju, L.; Qayum, A. LULC Analysis of Urban Spaces Using Markov Chain Predictive Model at Ranchi in India. *Spat. Inf. Res.* **2017**, *25*, 351–359. [CrossRef]
2. Naikoo, M.W.; Rihan, M.; Ishtiaque, M.; Shahfahad. Analyses of Land Use Land Cover (LULC) Change and Built-up Expansion in the Suburb of a Metropolitan City: Spatio-Temporal Analysis of Delhi NCR Using Landsat Datasets. *J. Urban Manag.* **2020**, *9*, 347–359. [CrossRef]
3. Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad; Pal, S.; Liou, Y.-A.; Rahman, A. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sens.* **2020**, *12*, 1135. [CrossRef]
4. Derdouri, A.; Wang, R.; Murayama, Y.; Osaragi, T. Understanding the Links between LULC Changes and SUHI in Cities: Insights from Two-Decadal Studies (2001–2020). *Remote Sens.* **2021**, *13*, 3654. [CrossRef]
5. Hadi, S.J.; Shafri, H.Z.M.; Mahir, M.D. Modelling LULC for the Period 2010–2030 Using GIS and Remote Sensing: A Case Study of Tikrit, Iraq. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *20*, 012053. [CrossRef]

6. Alshari, E.A.; Gawali, B.W. Development of Classification System for LULC Using Remote Sensing and GIS. *Glob. Transit. Proc.* **2021**, *2*, 8–17. [CrossRef]
7. Ali, K.; Johnson, B.A. Land-Use and Land-Cover Classification in Semi-Arid Areas from Medium-Resolution Remote-Sensing Imagery: A Deep Learning Approach. *Sensors* **2022**, *22*, 8750. [CrossRef]
8. Mugari, E.; Masundire, H. Consistent Changes in Land-Use/Land-Cover in Semi-Arid Areas: Implications on Ecosystem Service Delivery and Adaptation in the Limpopo Basin, Botswana. *Land* **2022**, *11*, 2057. [CrossRef]
9. Roy, A.; Inamdar, A.B. Multi-Temporal Land Use Land Cover (LULC) Change Analysis of a Dry Semi-Arid River Basin in Western India Following a Robust Multi-Sensor Satellite Image Calibration Strategy. *Heliyon* **2019**, *5*, e01478. [CrossRef]
10. Yonaba, R.; Koïta, M.; Mounirou, L.A.; Tazen, F.; Queloz, P.; Biaou, A.C.; Niang, D.; Zouré, C.; Karambiri, H.; Yacouba, H. Spatial and Transient Modelling of Land Use/Land Cover (LULC) Dynamics in a Sahelian Landscape under Semi-Arid Climate in Northern Burkina Faso. *Land Use Policy* **2021**, *103*, 105305. [CrossRef]
11. Njoku, E.A.; Tenenbaum, D.E. Quantitative Assessment of the Relationship between Land Use/Land Cover (LULC), Topographic Elevation and Land Surface Temperature (LST) in Ilorin, Nigeria. *Remote Sens. Appl. Soc. Environ.* **2022**, *27*, 100780. [CrossRef]
12. Tolentino, F.M.; de Lourdes Bueno Trindade Galo, M. Selecting Features for LULC Simultaneous Classification of Ambiguous Classes by Artificial Neural Network. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100616. [CrossRef]
13. Jozdani, S.E.; Johnson, B.A.; Chen, D. Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. *Remote Sens.* **2019**, *11*, 1713. [CrossRef]
14. Jamali, A. Evaluation and Comparison of Eight Machine Learning Models in Land Use/Land Cover Mapping Using Landsat 8 OLI: A Case Study of the Northern Region of Iran. *SN Appl. Sci.* **2019**, *1*, 1448. [CrossRef]
15. Tuzcu, A.; Taskin, G.; Musaoğlu, N. Comparison of Object Based Machine Learning Classifications of Planetscope and Worldview-3 Satellite Images for Land Use/Cover. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W13*, 1887–1892. [CrossRef]
16. Ghayour, L.; Neshat, A.; Paryani, S.; Shahabi, H.; Shirzadi, A.; Chen, W.; Al-Ansari, N.; Geertsema, M.; Pourmehdi Amiri, M.; Gholamnia, M.; et al. Performance Evaluation of Sentinel-2 and Landsat 8 OLI Data for Land Cover/Use Classification Using a Comparison between Machine Learning Algorithms. *Remote Sens.* **2021**, *13*, 1349. [CrossRef]
17. Nasiri, V.; Deljouei, A.; Moradi, F.; Sadeghi, S.M.M.; Borz, S.A. Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. *Remote Sens.* **2022**, *14*, 1977. [CrossRef]
18. Elamin, A.; El-Rabbany, A. UAV-Based Multi-Sensor Data Fusion for Urban Land Cover Mapping Using a Deep Convolutional Neural Network. *Remote Sens.* **2022**, *14*, 4298. [CrossRef]
19. Basheer, S.; Wang, X.; Farooque, A.A.; Nawaz, R.A.; Liu, K.; Adekanmbi, T.; Liu, S. Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques. *Remote Sens.* **2022**, *14*, 4978. [CrossRef]
20. Akbar, T.A.; Hassan, Q.K.; Ishaq, S.; Batool, M.; Butt, H.J.; Jabbar, H. Investigative Spatial Distribution and Modelling of Existing and Future Urban Land Changes and Its Impact on Urbanization and Economy. *Remote Sens.* **2019**, *11*, 105. [CrossRef]
21. Drummond, M.A.; Stier, M.P.; Diffendorfer, J.J.E. Historical Land Use and Land Cover for Assessing the Northern Colorado Front Range Urban Landscape. *J. Maps* **2019**, *15*, 89–93. [CrossRef]
22. Hoque, M.Z.; Cui, S.; Islam, I.; Xu, L.; Tang, J. Future Impact of Land Use/Land Cover Changes on Ecosystem Services in the Lower Meghna River Estuary, Bangladesh. *Sustainability* **2020**, *12*, 2112. [CrossRef]
23. Hufkens, K.; de Haulleville, T.; Kearsley, E.; Jacobsen, K.; Beeckman, H.; Stoffelen, P.; Vandelook, F.; Meeus, S.; Amara, M.; Van Hirtum, L.; et al. Historical Aerial Surveys Map Long-Term Changes of Forest Cover and Structure in the Central Congo Basin. *Remote Sens.* **2020**, *12*, 638. [CrossRef]
24. Yao, J.; Mitran, T.; Kong, X.; Lal, R.; Chu, Q.; Shaukat, M. Landuse and Land Cover Identification and Disaggregating Socio-Economic Data with Convolutional Neural Network. *Geocarto Int.* **2020**, *35*, 1109–1123. [CrossRef]
25. Leta, M.K.; Demissie, T.A.; Tränckner, J. Modeling and Prediction of Land Use Land Cover Change Dynamics Based on Land Change Modeler (LCM) in Nashe Watershed, Upper Blue Nile Basin, Ethiopia. *Sustainability* **2021**, *13*, 3740. [CrossRef]
26. Firozjaei, M.K.; Sedighi, A.; Firozjaei, H.K.; Kiavarz, M.; Homaee, M.; Arsanjani, J.J.; Makki, M.; Naimi, B.; Alavipanah, S.K. A Historical and Future Impact Assessment of Mining Activities on Surface Biophysical Characteristics Change: A Remote Sensing-Based Approach. *Ecol. Indic.* **2021**, *122*, 107264. [CrossRef]
27. Jalayer, S.; Sharifi, A.; Abbasi-Moghadam, D.; Tariq, A.; Qin, S. Modeling and Predicting Land Use Land Cover Spatiotemporal Changes: A Case Study in Chalus Watershed, Iran. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 5496–5513. [CrossRef]
28. Mäyrä, J.; Kivinen, S.; Keski-Saari, S.; Poikolainen, L.; Kumpula, T. Utilizing Historical Maps in Identification of Long-Term Land Use and Land Cover Changes. *Ambio* **2023**. [CrossRef]
29. Krivoguz, D.; Bespalova, L. Landslide Susceptibility Analysis for the Kerch Peninsula Using Weights of Evidence Approach and GIS. *Russ. J. Earth Sci.* **2020**, *20*, ES1003. [CrossRef]
30. Krivoguz, D.; Bespalova, L. Analysis of Kerch Peninsula's Climatic Parameters in Scope of Landslide Susceptibility. *Bull. KSMTU* **2018**, *574*, 5–11.
31. Krivoguz, D.O.; Burtnik, D.N. Neural Network Modeling of Changes in the Land Cover of the Kerch Peninsula in the Context of Landslides Occurence. *Nauchno-Tekhnicheskiy Vestn. Bryanskogo Gos. Univ.* **2018**, *1*, 113–121. [CrossRef]
32. Krivoguz, D.; Mal'ko, S.; Borovskaya, R.; Semenova, A. Automatic Processing of Sentinel-2 Image for Kerch Peninsula Lake Areas Extraction Using QGIS and Python. *E3S Web Conf.* **2020**, *203*, 03011. [CrossRef]

33. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
34. Montavon, G.; Samek, W.; Müller, K.-R. Methods for Interpreting and Understanding Deep Neural Networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [CrossRef]
35. Krivoguz, D.; Bespalova, L.; Zhilenkov, A.; Chernyi, S. A Deep Neural Network Method for Water Areas Extraction Using Remote Sensing Data. *JMSE* **2022**, *10*, 1392. [CrossRef]
36. Larochelle, H.; Bengio, Y.; Louradour, J.; Lamblin, P. Exploring Strategies for Training Deep Neural Networks. *J. Mach. Learn. Res.* **2009**, *10*, 1–40.
37. Samek, W.; Montavon, G.; Lapuschkin, S.; Anders, C.J.; Müller, K.-R. Explaining Deep Neural Networks and beyond: A Review of Methods and Applications. *Proc. IEEE* **2021**, *109*, 247–278. [CrossRef]
38. Avdeev, B.; Vyngra, A.; Chernyi, S. Improving the Electricity Quality by Means of a Single-Phase Solid-State Transformer. *Designs* **2020**, *4*, 35. [CrossRef]
39. Leo, B. Random Forests. *Mach. Learn* **2001**, *45*, 5–32. [CrossRef]
40. Chistiakov, S. Random Forests: An Overview. *Trans. KarRC RAS* **2013**, *12*, 117–136.
41. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support Vector Machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
42. Noble, W.S. What Is a Support Vector Machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef] [PubMed]
43. Steinwart, I.; Christmann, A. *Support Vector Machines*; Springer Science & Business Medi: Berlin/Heidelberg, Germany, 2008.
44. Bell, J. What Is Machine Learning? In *Machine Learning and the City: Applications in Architecture and Urban Design*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2022; pp. 207–216.
45. Zhou, Z.-H. *Machine Learning*; Springer Nature: Berlin/Heidelberg, Germany, 2021.
46. Wu, Q. Geemap: A Python Package for Interactive Mapping with Google Earth Engine. *J. Open Source Softw.* **2020**, *5*, 2305. [CrossRef]
47. Cruickshank, J.L. The Evolution of Soviet Topographic Maps as Revealed by Their Published Supporting Documentation. *Cartogr. J.* **2021**, *1*, 1–20. [CrossRef]
48. Krivoguz, D.; Bondarenko, L.; Matveeva, E.; Zhilenkov, A.; Chernyi, S.; Zinchenko, E. Machine Learning Approach for Detection of Water Overgrowth in Azov Sea with Sentinel-2 Data. *J. Mar. Sci. Eng.* **2023**, *11*, 423. [CrossRef]