DOI: 10.63550/ICEIP.2025.34.68.067

Methods for Constructing Graphs of Transport Networks Using GIS from Remote Sensing Data

Martynova Natalia[a], Budarova Valentina[b]
Industrial University of Tyumen, 38 Volodarskogo Street, Tyumen, Russia
natali.cherdanceva@mail.ru, budarova@bk.ru

Keywords: Remote sensing of the earth, Geographic information systems, Road network graphs, Machine learning, Transport modeling.

Abstract: The article discusses the problem of automated construction of graphs of road networks based on remote sensing data and geographic information systems. An instrumental and mathematical approach to the formalization and modeling of transport networks is presented, which allows taking into account the topological and metric characteristics of road objects. Methods for identifying and classifying road segments using machine learning algorithms and geostatistical analysis are described. An analysis of the applicability of various sources of remote sensing data and their integration with geoinformation technologies is carried out. Examples of the implementation of the approach are given and the prospects for its application in the tasks of transport planning, infrastructure monitoring and automated mapping are considered. The results obtained demonstrate the effectiveness of the developed algorithms and the possibility of their use in modern intelligent systems for managing traffic flows.

1 INTRODUCTION

The development of digital technologies and remote sensing of the Earth (ERS) opens up new opportunities for the analysis of transport networks. Graph models of road networks are the basis for navigation, transport planning and flow modeling.

Studies have shown that machine learning methods, such as convolutional neural networks, can achieve up to 95% accuracy in the classification of road features (Shao S. et al., 2022). The use of remote sensing data in combination with geographic information systems (GIS) simplifies the automated extraction of transport arteries, which is confirmed by the work of Jia Y. et al. (2023), where the use of Sentinel-2 multispectral images made it possible to achieve high detail mapping of the road network.

2 OVERVIEW OF EXISTING METHODS

In modern research (Shao S. et al., 2022). various methods are used to extract road networks from remote sensing data, including:

- Image classification using machine learning;
- Segmentation methods based on the spectral characteristics of coverages;
- Vectorization of road objects using structure recognition algorithms;
- Combined approaches that take into account the spatial characteristics of road elements.

Use of convolutional neural networks (CNNs) and transformers for segmentation of road objects (Shao S. et al., 2022). Studies show that the combination of optical satellite imagery and deep learning can improve the accuracy of road detection by up to 95%, and the integration of Sentinel-2 data and deep learning algorithms (Jia Y. et al., 2023) allows for

[a] https://orcid.org/0000-0001-9603-5563
[b] https://orcid.org/0000-0002-7749-5932

more accurate identification of transport routes in complex urban environments.

Segmentation methods based on the spectral characteristics of coverages, including multispectral and hyperspectral image analysis. Studies show that the combination of Sentinel-2 data and spectral classification algorithms (e.g., Support Vector Machine and Random Forest) can achieve an accuracy of over 90% of pavement identification (Tan Y.-C., 2024). In addition, the use of neural network segmentation methods, such as U-Net, has demonstrated high efficiency in identifying the boundaries of road objects in satellite images (Bastani et al., 2018).

Bastani et al. (2018) presented the RoadTracer method, based on deep learning, which allows you to efficiently reconstruct the topology of road networks from satellite images. Studies by Mattyus et al. (2017) also demonstrate that the use of a combination of CNN and CRF (Conditional Random Fields) improves the accuracy of road feature reconstruction by taking into account spatial coherence.

Studies by Zhou J et al. (2018) demonstrate that the use of graph neural networks (GNNs) for the analysis of road networks can improve the accuracy of reconstructing complex transport structures. At the same time, studies by Yu Bin et al. (2024) show that the integration of LiDAR data with multispectral imagery and spatial clustering techniques significantly improves the identification and classification of road features. These approaches are especially effective in complex terrain and dense urban areas, where standard segmentation methods show a decrease in accuracy.

2.1 Methods for automated detection of road networks and graph construction

Based on Earth remote sensing data, the road mask can be extracted automatically, which is necessary for further construction of the road network graph by classifying images. There are several types of classifications:

1) Controlled classification methods

Methods for selecting objects of this type of classification are based on samples of reference image fragments for training. In a 1992 paper, "Classification of multispectral remote sensing data using a back-propagation neural network," Heermann P.D. and Khazenie, N. described a backpropagation algorithm based on the spectral and contextual information of image pixels. Later, in a 2011 scientific article "Automated road network extraction using artificial neural network", Indian scientists Kirthika, A.M., Mookambiga, A. described a method for classifying an image using a backpropagation neural network. To identify the roads, the spectral features of the image pixels were used, as well as additional image texture parameters: contrast, entropy and uniformity for each pixel. An example of the result obtained is shown in Figure 1.

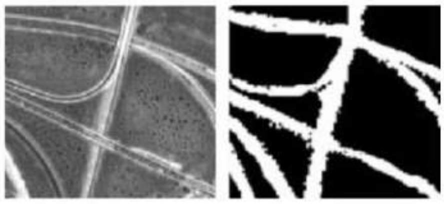

Figure 1: Fragment of a selected section of the road network Kirthika A.M., Mookambiga A.

2) Unsupervised classification methods

Unsupervised classification methods do not need training samples, which have many advantages in solving classification problems. However, the accuracy of the methods is lower than that of controlled classification methods in general.

In 2012, Unsalan C. and Sirmacek B. used a method of extracting roads based on probability and graph theory. The system has three main modules: probabilistic extraction of the centerline of the road, a module of extracting the shape of the road and the formation of a road network based on graph theory. These three modules can be used sequentially or alternately, depending on the specific situation. Many experiments have been conducted using different imagery, such as GeoEye, Ikonos and QuickBird, to show the advantages and disadvantages of the system in Figure 2. Fragment of the extracted road. Unsalan C. and Sirmacek B. An example of research results is presented. It can be seen that the system has good stability and accuracy when highlighting the road network.

Figure 2: Fragment of the extracted road. Unsalan C. and Sirmacek B.

In 2014, Zelang Miao et al. proposed a semi-automatic method for extracting the centerline of the road from the extraction of high-spatial resolution images, the essence of which is the initial sections of the road for their subsequent connection, based on the creation of a traffic probability map and the subsequent operation of threshold division of the image into classes of roads and non-roads, after which a model for estimating the density of the core is built (Fig. 1). 3.).

Figure 3: Road detection by semi-automatic method. (Zelang Miao and others).

3) Active Contour Model

The basic idea of the active contour model (snake) was described by Michael Kass et al. in 1988 and is as follows: it takes some control points to form a certain shape as a template, and uses the elastic deformation of the pattern (spline) to match it to local elements in the image.

Anil P.N., S. Natarajan. (2010) presented a new approach using an active contour model to extract a road. First, a softened median filter was used to remove noise by preserving the edges. The operator then enters the starting points on the road, and a "snake" model is used to extract the road. Figure 4 shows the results of road network extraction.

Figure 4: Extraction of the road network using the Anil P.N., S. Natarajan algorithm.

2.2.1. Object-Oriented Classification Method

Object-oriented classification (OOC) is a method of processing remote sensing (ERS) data, in which not a single pixel is analyzed, but a group of pixels combined into meaningful objects based on their spectral, textural and geometric characteristics. This approach allows for spatial connectivity of the data and reduces the noise associated with pixel classification methods.

The OOC process begins with the segmentation stage, when the original image is divided into segments (superpixels) taking into account the similarity of spectral and textural features. The parameters of compactness, smoothness, and color homogeneity can be used as segmentation criteria. After segmentation, the resulting objects are analyzed, including the calculation of their attribute characteristics, such as the average brightness of the channels, shape, size, orientation, contrast, and other spatial indicators.

The next step is to classify objects using machine learning algorithms, including Support Vector Machine (SVM), Random Forest, or neural networks. Each object is assigned a certain class in accordance with the training sample. The final step is post-processing of the results, aimed at eliminating classification errors and clarifying class boundaries.

Advantages of the OOC method:

- Takes into account spatial and spectral characteristics, improving classification accuracy.
- Reduces the impact of noise compared to pixel methods.
- Allows you to highlight complex structures such as road networks, buildings, plant communities.

Disadvantages:

- Sensitivity to segmentation parameters that determine the final result.
- High computational complexity compared to pixel methods.

2.2.2. Neural Network Approach

The algorithm for automated road identification and construction of a road network graph using neural networks consists of several consecutive stages (Figure 5).

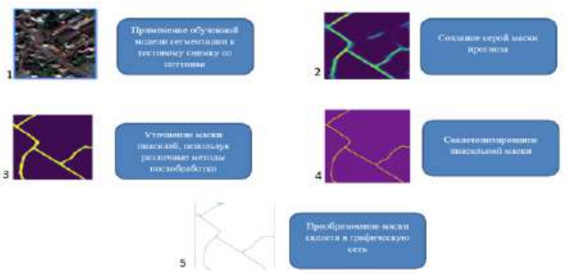

Figure 5: Algorithm for automated road identification and construction of a road network graph.

First, data is collected and pre-processed, including the use of high-resolution satellite imagery, such as Sentinel-2 and Landsat-8, as well as aerial imagery. At this stage, image correction (radiometric, atmospheric, geometric), noise filtering and data normalization are performed.

The next step is to highlight road objects using neural networks. For this purpose, convolutional neural networks (CNNs) are used, in particular, the U-Net and DeepLabV3+ architectures, trained on specialized data sets (for example, SpaceNet, RoadTracer). These models allow you to segment road features in imagery and then optimize them and eliminate false classifications.

After segmentation, road features are vectorized when pixel segments are converted to linear vector structures. It uses the Zhang-Suen skeletonization algorithm to determine the centerlines of the roads, as well as filtering the data to eliminate false segments.

A road network graph is built from vectorized features. Intersections and road endings are defined as graph vertices, and the sections connecting them are defined as edges. In addition, the edges are assigned attributes such as road width, pavement, direction of travel, and capacity. Graph neural networks (GNN) and conditional random field (CRF) methods are used to correct topological errors.

After that, the graph is analyzed and optimized: its connectivity is checked, isolated nodes are eliminated, and errors in intersections are corrected. The shortest routes are also analyzed using the Dijkstra and A* algorithms, and the central nodes of the road network are identified through graph centrality metrics.

At the last stage, the results are validated by comparing the resulting graph with real roadmaps, such as OpenStreetMap. The quality of the model is evaluated using the IoU (Intersection over Union), F1-score, and Precision-Recall metrics.

The finished graph is exported to GIS-compatible formats (Shapefile, GeoJSON, PostGIS) and can be used for transport modeling, infrastructure monitoring, and logistics.

2.2.3. Construction of road graphs using GIS tools

The process of constructing graphs of road networks in geographic information systems (GIS) is based on the processing of data on transport infrastructure, including geometric and attributive characteristics of roads. The quality of the resulting graph is determined by the completeness of the input data, including turn limits, maximum speeds, number of lanes, and other parameters that affect navigation and traffic flow analysis.

Одним из наиболее эффективных инструментов для построения графов дорог является QGIS в link to the PostgreSQL database and the pgRouting extension. The source data can be obtained from OpenStreetMap (OSM) and uploaded to PostgreSQL using the osm2pgsql utility, developed by the community led by Sarah Hoffman and Paul Norman. This tool converts OSM data into a manageable format using a structure of nodes and paths, where closed paths are polygons.

To further build the graph of the road network, the osm2po program developed by Carsten Möller is used. It converts the source data into a format suitable for working in PostgreSQL with pgRouting, providing the ability to build optimal routes, model traffic flows, and perform spatial analysis.

Advantages of the method:

- Ability to process large amounts of open OSM data.
- High graph accuracy due to extensive attributive information.
- Integrate with modern GIS such as QGIS, ArcGIS, and more.

Disadvantages:

- Complexity of setting up and processing data.
- The need for pre-cleaning and filtering of OSM data.

2.2 Instrumental and mathematical aspects

Automated construction of road network graphs includes several key stages:

- Data pre-processing – filtering, alignment, correction of georeferencing of satellite images;
- Selection of road objects – the use of machine vision algorithms and geostatistical methods;
- Conversion to a graph model – building a network taking into account the topology of roads, junctions, attribute characteristics;
- Graph optimization and analysis – application of graph theory methods to assess connectivity, determine the shortest routes, and simulate traffic flows.

3 APPLICATION OF GEOGRAPHIC INFORMATION SYSTEMS FOR AUTOMATED IDENTIFICATION OF ROAD NETWORKS AND CONSTRUCTION OF A GRAPH OF THE ROAD NETWORK ON THEIR BASIS

The construction of road graphs was carried out on the territory of the settlements of the Arkhangelsk Region, for subsequent uploading to the GIS ZEMLYA of the Arkhangelsk Region system, in view of the need to use them within the framework of the project "Alcohol Register of the Arkhangelsk Region" to build zones prohibited from the sale of alcoholic beverages.

Identification of roads and construction of the road network graph was carried out by several approaches: Construction of the road network graph using GIS, neural network approach, object-oriented classification method.

3.1. Neural network approach

The use of geographic information systems (GIS) for automated identification of road networks and construction of a road network graph using neural network approaches is an advanced methodology that significantly improves the processes of analysis and optimization of transport infrastructure. In this context, GIS and neural networks can work together effectively to automate complex tasks such as identifying roads, building traffic path graphs, and modeling traffic flows.

To create neural networks, various libraries are used, with the help of which it is possible to create a deep learning neural network in the Python programming language without a full training course.

After researching the libraries, several were chosen for further work - TensorFlow and Keras, as they are easy to learn and suitable for training neural networks aimed at identifying roads.

TensorFlow is a library created by Google that was developed specifically for training neural networks. This library uses a multi-level system of nodes, which allows you to expand the scope of application. Keras is a high-level library that uses TensorFlow as a component. This library is minimalistic in its architecture. It is used to build and train neural networks.

Also, different models are needed to create a neural network. To create a neural network, the U-NET model was used, as it is the most suitable for solving image segmentation problems (Fig. 6). This neural network model was used by the participants who took 1st and 2nd place in the SpaceNet Road Detection and Routing Challenge mentioned above. That is, dividing the digital image into segments (classes) in order to simplify its analysis. The segmentation method is usually used to highlight objects and boundaries (lines, curves, etc.)

Figure 6: Example of image segmentation.

The neural network was trained on the data, which is the input and expected images (Figure 7).

Figure 7: Input and expected image.

To train the neural network, generalized training data in *.tiff format was downloaded from the Road and Building Detection Datasets service (Fig. 7, Fig. 8).

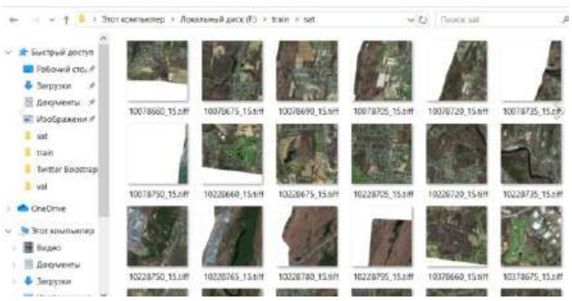

Figure 7: Loaded input data.

Figure 8: Uploaded expected images.

After preparing the data, the model training was run (Fig. 9).

Figure 9: Fragment of a neural network.

For training, several epochs were set - a full cycle (in the forward and reverse directions) of data passing through the neural network. The more epochs, the more accurate the neural network becomes. 15 epochs were used in the work (Figure 10), due to the large amount of data. The training took place for several hours.

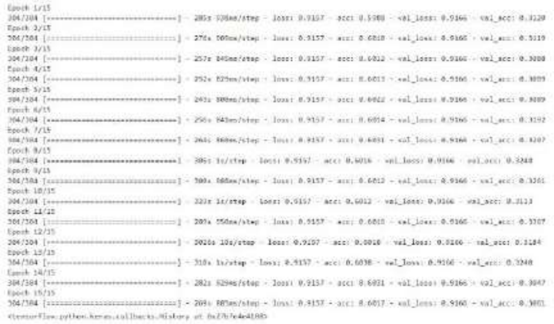

Figure 10: Completed Eras of Learning.

Next, the neural network was applied to an orthophoto of a part of the city of Velsk, Arkhangelsk Region (Fig. 11). Previously, atmospheric, radiometric and geometric corrections were applied to the original image.

Figure 11: Image to be processed.

As a result of the use of a neural network, an intermediate result was obtained (Fig. 32).

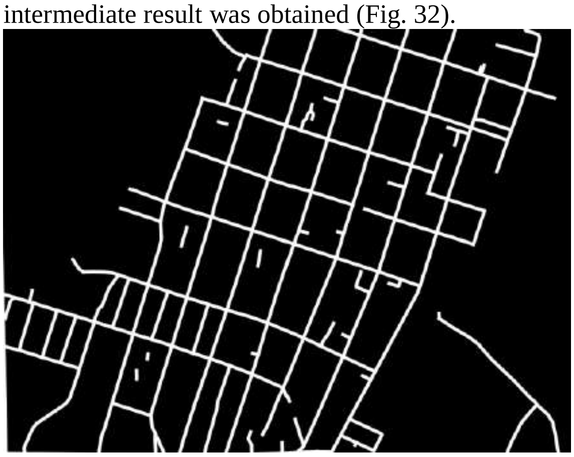

Figure 12: Raster image obtained during the application of a neural network.

As a result, not all roads fell into the class of roads, and there are also "noises" in the resulting image. Then, using the Skeleton script, which was embedded in the neural network, the image was vectorized, as well as edges and nodes of the graph were created (Fig. 33).

Figure 13: The resulting graph of the road network.

4 CONCLUSIONS

The study demonstrates the possibilities of using geographic information systems (GIS) for automated identification of road networks and building graphs based on them, using various approaches: building a graph of a road network using GIS tools, a neural network approach and an object-oriented classification method. Particular attention is paid to the application of the neural network approach using the TensorFlow and Keras libraries, as well as the U-NET model to solve the problems of image segmentation and identification of roads in orthophotos.

The choice of the U-NET model, justified by its successful application in the SpaceNet Road Detection and Routing Challenge, confirms its effectiveness for the tasks of distinguishing objects and boundaries in digital images. The neural network was trained on generalized data in *.tiff format obtained from the Road and Building Detection Datasets service and included 15 epochs due to the large amount of data. This process required significant computing resources and took several hours.

The application of a trained neural network to an orthophoto of a part of the city of Velsk, Arkhangelsk Region, showed that, despite preliminary atmospheric, radiometric and geometric correction of the original image, the results need further refinement. Not all roads were correctly classified, and there was "noise" in the resulting image, indicating the need for better model training and/or data preprocessing.

Using the Skeleton script to vectorize the image and create edges and graph nodes made it possible to automate the process of building a road network. However, the need to refine the results of neural network segmentation indicates that further optimization of the methods and parameters used is required to achieve high accuracy and reliability in identifying road networks.

5 DISCUSSION

The results of the work emphasize the prospects of using neural network approaches in GIS for automation, in particular:

- Increasing the volume and variety of training data to improve the accuracy of neural network segmentation.
- Development of more efficient data preprocessing and post-processing methods to reduce noise and improve classification quality.
- Explore the possibility of using other neural network models and/or combine different approaches (e.g., neural network approach with object-oriented classification) to improve the reliability of road network detection.
- Automation of the process of calibration and training of neural networks for specific regions and types of roads.
- Integrate the results of automated road network identification with other GIS tools to solve various problems, such as route planning, traffic flow analysis, and accessibility assessment.

In conclusion, the results demonstrate the potential of GIS and neural network approaches to automate the process of identifying road networks and building graphs. Further development of these methods will make it possible to effectively solve various problems related to the planning, management and development of territories, including optimization of transport infrastructure, accessibility analysis and modeling of traffic flows.

ACKNOWLEDGEMENTS

The author is grateful to the Industrial University of Tyumen for providing the opportunity to conduct research. This study was conducted within the research activities of the Department of Geodesy and Cadastral Activities of the Industrial University of Tyumen in specialized scientist laboratories. The author expresses special gratitude to V. A. Shamsutdinov (a leading IT consultant at OOO OCRV) for his valuable contribution to this work.

REFERENCES

Shao, S., Xiao, L., Lin, L., Ren, C., Tian, J., 2022. Road Extraction Convolutional Neural Network with Embedded Attention Mechanism for Remote Sensing Imagery. Remote Sensing, 14(9), 2061.

Jia, Y., Zhang, X., Xiang, R., Ge, Y., 2023. Super-Resolution Rural Road Extraction from Sentinel-2 Imagery Using a Spatial Relationship-Informed Network. Remote Sensing, 15(17), 4193.

Tan, Y.-C., Duarte, L., Teodoro, A. C., 2024. Comparative Study of Random Forest and Support Vector Machine for Land Cover Classification and Post-Wildfire Change Detection. Land, 13(11), 1878.

Bastani, F., He, S., Abbar, S., Alizadeh, M., Balakrishnan, H., Chawla, S., Madden, S., DeWitt, D.J., 2018. RoadTracer: Automatic Extraction of Road Networks from Aerial Images. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 4720-4728.

Máttyus, G., Luo, W., Urtasun, R., 2017. DeepRoadMapper: Extracting Road Topology from Aerial Images. 2017 IEEE International Conference on Computer Vision (ICCV), 3458-3466.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2018. Graph Neural Networks: A Review of Methods and Applications.

Yu Bin, Wang Yuchen, Chen Qihang, Chen Xiaoyang, Zhang Yuqin, Luan Kaiyue, Ren Xiaole, 2024. A review of road 3D modeling based on light detection and ranging point clouds. Journal of Road Engineering, 4(4): 386-398.

Heermann, P.D., Khazenie, N., 1992. Classification of Multispectral Remote Sensing Data Using a Back-Propagation Neural Network. IEEE Transactions on Geoscience and Remote Sensing, 30, 81-88.

Kirthika, A.M., Mookambiga, A., 2011. Automated road network extraction using artificial neural network. 2011 International Conference on Recent Trends in Information Technology (ICRTIT), 1061-1065.

Unsalan, C., Sirmacek, B., 2012. Road Network Detection Using Probabilistic and Graph Theoretical Methods. IEEE Transactions on Geoscience and Remote Sensing, 50(11), 4441–4453.

Zelang Miao, Wang, Wenzhong Shi, Hua Zhang, 1014. A Semi-Automatic Method for Road Centerline Extraction From VHR Images. *IEEE Geoscience and Remote Sensing Letters*№11. C. 1856 - 1860.

Anil P.N., Natarajan, S., 2010. A Novel Approach Using Active Contour Model for Semi-Automatic Road Extraction from High Resolution Satellite Imagery. 2010 Second International Conference on Machine Learning and Computing.