TRAINING OF NEURAL NETWORKS TO DECIPHER THE ROAD NETWORK ACCORDING TO SPACE IMAGERY RECEIVED BY THE "RESURS-P"

N. N. Kasatikov[a], Sh. M. Umarov[b,*], A.D. Fadeeva[a], S. A. Tolmachev[c]

[a] Moscow Aviation Institute, Moscow (MAI), Russia - nick925@yandex.ru
[b] Dept. of Photogrammetry, Moscow State University of Geodesy and Cartography (MIIGAiK), Moscow, Russia - shapi0890@mail.ru
[c] Russian Technological University (RTU MIREA), Moscow, Russia

Commission II, WG II/8

KEY WORDS: Neural networks, Special processing of satellite images, "Resurs-P", GroMobile development, Geoportal, Digital modern map, GCPs

ABSTRACT:

Our team has developed a neural network for road recognition on our digital twin, aimed at enhancing transportation-related applications. The neural network is trained on large datasets of road images and utilizes various deep learning architectures and techniques to improve its accuracy and reliability. The embedded neural network can recognize different road features, such as lane markings, road signs, and obstacles, and can identify the location and direction of the road. The integration of this neural network in our digital twin can help optimize transportation-related operations, reduce accidents, and improve overall traffic flow. The developed neural network architecture and training methodology, as well as its performance evaluation on various datasets, are presented in this paper. Additionally, the paper discusses the future directions for research in this area and the potential of the developed neural network for other applications in the digital twin domain.

1. INTRODUCTION

At the time of writing, cases of occurrence using artificial intelligence. Such structures include cartography. Unfortunately manual labor is recognized as the core of its task. We also offer partial automation of creating a digital landscape model using modern technologies. The article will create a geoportal using neural networks to detect roads.

Technology of using neural networks in the field of mapping Every year. By 2022, the market is expected to more than double to $73.91 billion with an average growth of 19.2% (Y . LeCun, Y. Bengio, and G. Hinton, 2015). When writing the material, we relied on geoservices (geoportals), which analyze the Earth's surface, geovisualization, remote sensing of the Earth, GPS, GIS (in particular, photogrammetric services), geodesy.

A modern map is, in other words, a digitized interpretation of the terrain image. The process of creating maps depends on two facts. Initial - obtaining a graphic image of the terrain: a snapshot of space or terrain with GPS-referencing and geographic data. Special - detection of images of objects in the image and saving in digital format (lines, points of view, particular examples) with reference to coordinates: latitude, longitude, height above sea level. Most of these tasks, in our opinion, should be excluded by neural networks. But still, in the field of cartography and geosciences, often even in simple processes, human influence remains too important at the moment.

Rendering vector maps based in a repository of space accumulations and measurements on the ground for the sale of processed vector geodata to companies involved in printing paper maps and creating cartographic services (both B2C and B2B - in the form of an SDK with payment on request to the API) (A . Krizhevsky, I. Sutskever, and G. Hinton, 2012). The described market, a few years ago, received a serious returbation, due to the transfer from paper maps to digital samples and the approval of B2B services under the restrictions of Yandex and 2Gis, which allow you to quickly create new B2B / B2C platforms based on a ready-made SDK that work with user / operator geolocation. However, raw vector data is consistently bought by companies (hydro technical, educational and other services) that require large amounts of data rendering and compliance with industrial and government requirements. There are also few global players in this segment of B2B/B2C services: the main ones are Google Maps, TomTom, Here Maps and OpenStreetMap.org (open data) (D. Silver, A.et al., 2016). There are also a number of flagship companies like Yandex in Russia, which readers of their B2B business often sell data, but jumped to B2C. Drawing up contracts for topographic maps, Base Vector Map, DEM expensive procedure, bids reach billion, small private company coming soon with preferred use of off-the-shelf B2B solutions powered by Google Maps SDK instead of buying raw data, selling and even more so self-designing the SDK. Basic data for free services is mostly, but limited in a number of functions that attract new companies that subsequently find solvent small businesses or customers (I. Goodfellow, Y. Bengio, and A. Courville, 2016; S. Hochreiter and J. Schmidhuber, 1997).

It is possible to conditionally distinguish exceptional types of objects on the map: landscape, water bodies, forests, sights, roads and organizations (places). Part of the information (contours of roads, buildings, water bodies) easily absorbs the map without being tied to the terrain. The second part (the name of streets, roads, organizations, houses) can be collected only by highlighting it on the spot.

The most options for creating topographic maps:

1. Manual tracing of the contours of objects on satellite images from GPS - binding

* Corresponding author.

Publisher's note: Copernicus Publications has not received any payments from Russian or Belarusian institutions for this paper.

2. Manual GPS recording –tracks
3. Manual recognition of the detection of low-observable areas made by UAVs/drones (aerial photography)

All processes of creating and replenishing cards can be divided into manual and promising automatic. In the first case, some large objects like areal rivers or vegetation elements are determined automatically, and some structures and roads are drawn automatically. Next, use neural networks and machine learning.

2. NEURAL NETWORKS. EDUCATION

Neural networks can be used for image recognition tasks such as identifying country roads and highways. The process can be divided into several stages:

> Data collection. The first step is to collect an image dataset containing country roads and highways. These images can be obtained from various public datasets or can be collected by creating a custom dataset through manual image collection.
> Pre-processing. before feeding images to the neural network, they need to be pre-processed. This may include tasks such as resizing images, normalizing pixel values, and converting them to a format that can be read by a neural network.
> Training. After pre-processing the data, it can be used to train the neural network. The network can be trained using a supervised learning approach where it is provided with images of country roads and highways and their corresponding labels (i.e. whether the image depicts a country road or a highway). During training, the neural network learns to identify features that distinguish two classes of images.
> Validation. Once a network has been trained, it needs to be validated to ensure that it can generalize to new images. This includes testing the network on a set of images it hasn’t seen before and evaluating its performance.
> Deployment. Once the network has been trained and validated, it can be deployed for use in real applications. For example, it can be used in a self-driving car to determine if the car is on a country road or on a highway, which can then be used to customize the car’s driving behavior.

In general, neural networks can be used to identify country roads and highways by learning to recognize patterns and features in images that are specific to each type of road. With enough training data and proper preprocessing and validation, a neural network can achieve a high level of accuracy on this task.

Automation of manual labor occurs using conventional computing technologies. When constructing rivers, the operator places the beginning of the road and leads to the path where it ends with the mouse, which takes time. When using neural networks, the program itself will choose the destination of the roads and conduct them in a short time. This operation will often greatly save the operator’s labor, from one section of the map to another, and the operator cannot release the mouse until the oil has accumulated.

In this regard, the task of restoring vector layers to create a map comes to the fore. This needs to be divided into several regions:

1. Road recognition on a satellite image
2. Vector layer correction
3. Saving a layer

2.1 Architecture

There are several neural network architectures that can be used for road recognition problems. The suitability of a particular architecture depends on factors such as the complexity of the task, the size of the data set, and the available computing resources. Here are some examples of neural network architectures that are commonly used for road recognition:

> Convolutional Neural Networks (CNNs): CNNs are well suited for image recognition tasks such as road recognition. They are designed to explore the spatial hierarchy of objects by applying convolutional filters to the input image. CNNs are commonly used in a supervised training setup where they are trained on labeled on-road and off-road image datasets.
> Recurrent Neural Networks (RNNs): RNNs are suitable for tasks involving sequential data, such as recognizing traffic patterns from video footage. They can capture time dependencies by using a memory mechanism to store information about previous time steps. RNNs are often used in conjunction with CNNs to process both spatial and temporal characteristics.
> Fully Convolutional Networks (FCNs): FCNs are a type of CNN that can process images of arbitrary size and produce pixel-by-pixel predictions. They are often used for semantic segmentation tasks where the goal is to label each pixel in an image with a class label. FCN can be trained on on-road and off-road image datasets to create a pixmap of the road.
> Deep Residual Networks (ResNets): ResNets are a type of CNN that use residual connections to improve the flow of information through the network. They are designed to solve the problem of vanishing gradients that can occur in very deep neural networks. ResNets can be used for road recognition problems by training them on labeled road and off-road imagery datasets.
> EfficientNets : EfficientNets are a family of CNNs designed to achieve state-of-the-art performance in image recognition tasks while using fewer resources than other networks. They use a compound scaling method to balance the number of parameters, FLOPS, and precision. EfficientNets can be used for road recognition tasks when computing resources are limited.

We chose the U-Net architecture to create a neural network. Net has been shown to be very effective for object segmentation in images, including hydrographic features. This is due to the fact that it is able to capture both the global and local context in an image, allowing for more accurate predictions(A. Graves, A.-r. Mohamed, and G. Hinton, 2013). The U-Net architecture has several advantages for implementing hydrographic feature detection, including:

> Reliability: U-Net resistant to changes in the appearance of hydrographic elements, such as changes in lighting or shadows. This is because it can learn features at multiple scales, allowing both fine and coarse detail to be captured.
> Efficient Training: U-Net trains effectively even with limited training data. This is because it uses skipped connections between the encoder and decoder, which allows it to learn features at different scales and reduces the need for a large number of training samples (Y. LeCun, L. Bottou, A.et al., 1998).
> Flexibility: The U-Net is flexible and can be adapted to

a range of hydrographic detection tasks, including the detection of rivers, lakes, and other water bodies.

Overall, U-Net is a powerful and efficient deep learning architecture for hydrographic feature detection, and its ability to process complex and variable image data makes it a valuable tool for a wide range of applications in hydrography and related fields.

3. EXPERIMENTAL PROCESSING

In the elite world, it's hard to underestimate. The rapid development of digital technologies and the development of broadband Internet allows you to receive enormous amounts of digital information. Governments and large companies are using resources to develop data processing using high-performance computing clusters. Cloud computing centers make it possible to work with the so-called large Data) . One type of data directed to this vast area is Earth remote sensing data.

The development of digital technologies based on the collection of neural networks is used in the processing of Earth remote sensing data. In this paper, we will highlight the issue of the applicability of neural networks in deciphering the road network based on satellite imagery. Satellite image , we work on "Resurs-P" measurements , is the standard for displaying the territory. Unlike OpenStreetMap (OSM) services , a satellite image can show the real picture. The territory of the Kaliningrad region and the Republic of Mordovia was chosen for the research. An archive of remote sensing materials is used to choose from. The study is composed of two months:

1. Stage of photogrammetric processing
2. Neural network training stage

3.1 Photogrammetric data processing

At the first stage, we carried out photogrammetric processing of 43 satellite imagery routes, revenues from the "Resurs-P" satellites. All paths up scaling of a high-resolution panchromatic channel with multispectral up scaling levels (Pansharpening). The acquisition of the resulting color images has a resolution of 1 meter. For orthorectification in both areas, DEM with a step of 10 meters were prepared in advance.

Photogrammetric processing consisted of a single orthorectification of each route in the TOVI software. During orthorectification, ground control points and check points were involved. Acceptance of Measured Calculation Assistance, Necessary Correction to Original RPC - event model. The RPC-model describes the relationship between the normalized spatial coordinates of the desired point in space (X,Y,Z) and its normalized coordinates in the image system (x,y) as a ratio of polynomials(1,2)(Dial G., Grodecki J. Block, 2003):

$$x = \frac{P(a, X, Y, Z)}{P(b, X, Y, Z)}; \tag{1}$$

$$y = \frac{P(c, X, Y, Z)}{P(d, X, Y, Z)}; \tag{2}$$

where: P = Polynomial function

When considering processing, the authors use the recommendations on the number of points and the correction model, for the most part in the article (Umarov Sh.M., Sudorgin A.S., 2020).

Refinement of the original RPC-model, which is an approximation of a rigorous implementation model, was produced using various types of geometric corrections. In our work, we applied possible correction types: shift (3,4), affine transformation (5,6), and polynomial correction of the second degree (7,8).

$$x' = a_0 + x; \tag{3}$$

$$y' = b_0 + y; \tag{4}$$

$$x' = a_0 + a_1x + a_2y; \tag{5}$$

$$y' = b_0 + b_1x + b_2y; \tag{6}$$

$$x' = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2; \tag{7}$$

$$y' = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2; \tag{8}$$

where: x, y = image coordinates
x', y' = corrected coordinates of objects in the image

Well proved in assembly on small scenes on the area. For long routes with terrain detection, polynomial corrections of the second degree are used.

The accuracy of determining the position in the images obtained as a result of single orthorectification does not exceed 2.5 meters. The achieved accuracy allows you to apply these orthophthros in creating documents of a scale of 1: 25000. Residual errors at the check points of one of the pictures are shown on (Figure 1). Obtained as a result of a single orthorectification stages were further processed in the PHOTOMOD "GeoMosaic" As a result, orthomosaics were obtained for the entire Kaliningrad region and the central part of the Republic of Mordovia. The resulting orthophoto coverages were loaded by a specially created geoportal to continue research on the second topic.

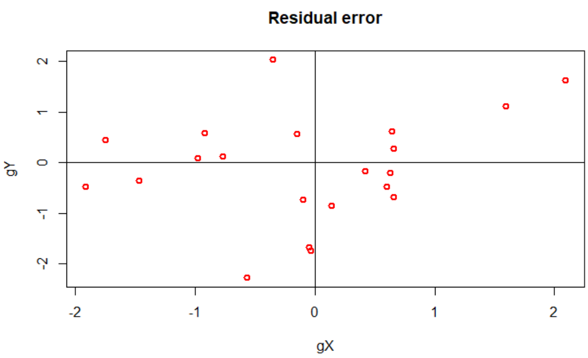

Figure 1: Residual error (meters)

3.2 Masks on "Resurs-P"

More about creating masks for a dataset from "Resurs-P" reference images. Masks are binary images that are used to indicate which parts of an image are in use and which are not. Computer review review typically uses image segmentation in different areas or objects. Here is how we can create masks for datasets from reference images.

First, we collected a set of reference images representing the different classes or objects that we planned to segment. The initial masks we observe using image download software such as Adobe Photoshop or GIMP. The mask is a binary image, in pixels, of the object of interest at times being white, and the pixels, at that time the background, being black. After that, we used them to train segmentation models. The segmentation model will predict masks for new images based on the reference images you provide. We then assumed that the masks could be predicted by the image loading software to make sure they were accurate.

In machine weight estimation, these are the values assigned to an individual attribute that determine their relative severity in the model. Weight management can help improve model performance and reduce overfitting. Here is how we controlled the weight.

Before we started training the models, we normalized the data to make sure all the features are available at the same scale. This helped prevent functions with large values from dominating the models. We have set the initial weights: when weights are added, the models fit randomly. We set the initial weights manually to improve the performance of the models. Once you have used regulation techniques such as L1 and L2 regulation that can help control weight by punishing weight. This helped prevent the model from being prevented and degraded. The next step is to adjust the weight during training: during training, we adjusted the weight depending on the performance of the model. For example, some features do not contribute much to the model and we can reduce their weight to reduce their impact on predictive models. Likewise, if the dependency is highly correlated, we reduced our weight there too to prevent overfitting.

4. APPLICATION ON THE GEOPORTAL

Uploaded images to the digital twin of memory are expensive in online mode. Our digital twin is written in the Leflet open source library . This open source library was chosen for its convenience and "friendly" interface, unlike other products. Our geoportal has a neural network module that allows you to quickly detect roads and rivers (Figure 2).

Figure 2: The program recognizes the road

It is written in Python, with the presence of the TensorFlow library , Keras . The knowledge base, as it was written above, was transferred in the "Resurs-P" images. At the same time, it can also work on Google images of the area. After processing the program, in the selected area of the territory, the vector layers of the atmosphere (Figure 3).

Figure 3: Choosing a site of wealth

As we see on the territory of Kaliningrad, the neural network does not always accurately estimate the edge of the road, and the operator who detects the calculation on the map detects additional learning of income. After completion by the operator, the file can be saved in .myth /.the middle of the expected work with it in GIS devices. Saving the road layer allows you to transfer data to GIS Panorama 11.11.5 on the sit map format. In the future, this data is converted for a map of the card format and is already converted and processed with it.

4.1 Testing on desktop and mobile devices

Testing of the site was carried out on most popular and popular files, in particular Mozilla, Firefox, Microsoft Edge, Windows 7, Windows 10, Linux Astra.

In the mobile version of the site (Figure 4), the work is observed in the same way by surface adjustments. The mobile version of the site works in the same way as it does on all browsers (testing was carried out on Safari) and in the production of files for iOS and android.

Small features of the site that are present in versions:

1. Search for territory by word
2. Cache in each session, until the site is restored, recognized territories will be shown
3. Convenient tools

In addition, a complete 3D characterization of the world has been created. Uninterested in the group of suspicions on the suspicion of the default-specified height. Large objects were blocked using efficient services (Mapbox) and inserted into the program.

CONCLUSIONS

Methods described to date, artificial intelligence seeks to help a person in the field of remote sensing data processing. Manual

https://doi.org/10.5194/isprs-archives-XLVIII-2-W3-2023-109-2023 | © Author(s) 2023. CC BY 4.0 License.

Figure 4: Work mobile version

labor accounts for approximately 80% of the mapping industry. This area of income depends on the person and his work. Our solution allows not only to simplify its work, but also to significantly improve the creation of CCM files. The application also works in the mobile version. The task of the double double map is to show the real frequency of the region, for this the module with algorithms created by our neural network was used, and the results presented in our study are an invaluable help. The use of neural digital networks allows you to quickly and efficiently find the specified types of objects. In addition, the built-in data storage function allows you to work with internal external applications.

Neural networks can be used for various tasks such as predictive maintenance, anomaly detection, and optimization. For example, a neural network can analyze data from sensors on a proposed facility to predict when maintenance is unavoidable or detect any anomalies in its assignment. By analyzing data from multiple sources, neural networks select a comprehensive view of an asset’s performance and help account for its performance.

Creating your own geoportal helps to properly configure its operation, both for users of ordinary computers and laptops, and for users who prefer mobile communications.

Saving data in a convenient format allows you to use it both in applications and in other open resources. Overall, neural networks can help make digital twins more accurate, efficient, and effective in their applications. As digital twins become more common at various scales, the role of neural networks in their design and implementation is expected to increase significantly, as described in our development.

REFERENCES

Y. LeCun , Y. Bengio , and G. Hinton. ”Deep learning.” Nature, 521(7553), 2015, pp. 436-444.

A. Krizhevsky , I. Sutskever , and G. Hinton. ”ImageNet classification with deep convolutional neural networks.” Advances in Neural Information Processing Systems, 2012, pp. 1097-1105.

D. Silver, A. Huang, CJ Maddison, A. Guez , L. Sifre , G. Van Den Driessche , J. Schrittwieser , I. Antonoglou , V. Panneershelvam , M. Lanctot , S. Dieleman , D. Grewe , J. Nham , N. Kalchbrenner , I. Sutskever , T. Lillicrap , M. Leach, K. Kavukcuoglu , T. Graepel, and D. Hassabis. ”Mastering the game of Go with deep neural networks and tree search.” Nature, 529(7587), 2016, pp. 484-489.

I. Goodfellow , Y. Bengio , and A. Courville. ”Deep learning.” MIT Press, 2016.

S. Hochreiter and J. Schmidhuber. ”Long short-term memory.” Neural computation, 9(8), 1997, pp. 1735-1780.

A. Graves, A.-r. Mohamed, and G. Hinton. ”Speech recognition with deep recurrent neural networks.” 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 6645-6649.

Y. LeCun , L. Bottou , Y. Bengio , and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 1998, pp. 2278-2324.

G. Dial , J.Grodecki. Block Adjustment of High-Resolution Satellite Images Described by Rational Polynomial // Photogrammetric Engineering Remote Sensing. 2003. V. 69 (1): 59–68

Sh.M.Umarov, A.S. Sudorgin. Comparison of various types of geometric transformations in the geodetic orientation of high spatial resolution panchromatic images obtained by ” Resurs-P” satellites. Izvestiya vuzov Geodeziya i aerofotosyemka ”. Izvestia vuzov ”Geodesy and Aerophotosurveying ”. 2020, 64(1): 54-62. [In Russian]. DOI: 10.30533/0536-101X-2020-64-1-54-62.

Revised April 2023