ROAD PASSABILITY ESTIMATION USING DEEP NEURAL NETWORKS AND SATELLITE IMAGE PATCHES

Anastasia Moumtzidou, Marios Bakratsas, Stelios Andreadis, Ilias Gialampoukidis, Stefanos Vrochidis, Ioannis Kompatsiaris *

Centre for Research & Technology Hellas
Information Technologies Institute
Thessaloniki, Greece

ABSTRACT

Artificial Intelligence (AI) technologies are getting deeper and deeper into remote sensing and satellite image processing offering value-added products and services in a real-time manner. Deep learning techniques applied on visual content are able to infer accurate decisions about concepts and events in an automatic way, based on Deep Convolutional Neural Networks which are trained on very large external image collections in order to transfer knowledge from them to the considered task. Existing emergency management services focus on the detection of flooded areas, without the possibility to infer if a road from point A to a point B is passable or not. To that end, we propose an automatic road passability service that is able to deliver the parts of the road network which are not passable, using satellite image patches. Experiments and fine-tuning on an annotated benchmark collection indicates the most suitable model among several Deep Convolutional Neural Networks.

Index Terms— Road passability, Deep Convolutional Neural Networks, Crisis Management, Road Network

1. INTRODUCTION

The high applicability of the Artificial Intelligence (AI) has led to the utilization of its technologies in order to develop and advance numerous other fields, among them remote sensing. Applying deep learning techniques on satellite images can offer an automatic identification of concepts or events. More specifically, we are based on Deep Convolutional Neural Networks (DCNNs) that are pre-trained on an external dataset of millions of images and use them to classify satellite imagery, a technique known also as transfer learning.

Our field of application is the Emergency Management applications, a managerial function that seeks to cope with hazards and disasters. While state-of-the-art mainly focuses on the detection of flooded areas in general, we target an explicit problem: starting from a point A to a point B, is a road passable or not due to a flood? Therefore, we introduce a road passability method that can automatically decide whether a roadway depicted in a satellite image is clear and able to be traveled.

The paper is structured as follows. In Section 2 we examine the existing works that are related to the problems of road extraction and flood detection. Section 3 describes the methodology, while Section 4 concerns the experiments and presents the results. Finally, Section 5 concludes and discusses future enhancements.

2. RELATED WORK

Road passability relies on two major sub-problems of remote sensing, being a combination of road extraction and flood detection procedures, with the most recent trends based on the exploitation of neural networks' capabilities. In the following we present the recent advances in both directions.

Road extraction detects road segments, as also defined in [1] where it is proposed to extract the road components from satellite images using Laplacian of Gaussian operator. The image is pre-processed to identify the color space components. At start, a panchromatic and a multispectral image of an area are combined (fused) to obtain more details of the image. Then, objects are identified using HSY color models components. Trying to distinguish roads from sandy regions, hue and luminance may have similar values but can be distinguished using saturation. A morphological method is applied to remove the unwanted objects in the image. In a more recent approach, the work of [3] explores 3 different Fully-Convolutional Neural Networks (FCNNs): FCN-8s with a VGG-19 backbone, Deep Residual U-Net0 and DeepLabv3+ for semantic segmentation. All networks were trained from scratch, where a considerable performance drop is noticed when using weights pretrained on ImageNet, due to the different nature of SAR images compared to optical ones. Adjusting the object segmentation, the task changes from a binary classification to a binary regression model, and instead of predicting each pixel as either road or background, the network weighs how likely it is for each pixel

*This work was supported by EOPEN project, partially funded by the European Commission, under the contract number H2020-776019.

Fig. 1. Part of a Web application that exploits the road passability service, developed for the purposes of H2020-EOPEN project.

to be a road. Due to Object awareness in FCNN, the predicted roads are sometimes disconnected at intersections, requiring re-connection of loose segments. In another work, in order to improve performance at heterogeneous areas (cars, trees on the road) a method on Generative Adversarial Networks (GAN) [6] is proposed to handle road detection. For the segmentation model, the so-called "Segnet" is used to generate a pixel-wise classification map. The GAN defines two models; the generative model, which is used to stimulate the data probability distribution, and the discriminative model, which is used to find whether a sample is coming from the generative model or the ground truth map. The generative and the discriminative models together form an adversarial network. Contrary to these approaches, we aim to infer whether a satellite image patch contains a passable road segment or not, without the need to segment the image patch into "road" and "no-road" regions.

Flood detection has been a popular problem in the remote sensing community, while nowadays the focus is on the use of Neural Networks, such as in [4], where the Fully-Convolutional Network (FCN), a variant of VGG16 on Gaofen-3 SAR images, is utilized for flood mapping. FCN demonstrates robustness to speckle noise in SAR images. Speckle noise is not filtered, making the deep learning model more universal (data augmentation). To make the model less complex, 7 x 7 kernels are replaced with 3 x 3 kernels greatly reducing conv6 parameters. In [5] the most widely used criteria performances, namely coefficient of determination (R2), sum squared error (SSE), mean squared error (MSE), and root mean squared error (RMSE) are used to optimize the performance of the Artificial Neural Network (ANN). Each method is estimated from the ANN predicted values and the measured discharges (targets). Seven input nodes, each representing flood causative parameters, including rainfall, slope, elevation, soil, geology, flow accumulation, and land use are used during the ANN modeling. There is little variation in maximum and minimum connection weights between the input and the hidden layers nodes except from the rainfall parameter. Rainfall factor is the main factor in the training of the neural network. The sensitivity analysis has shown that the elevation is the most important factor for flood susceptibility mapping. The approach in [8] is based on the segmentation of a single SAR image using self-organizing Kohonen maps (SOMs) and further image classication using auxiliary information on water bodies that could be derived, from optical satellite images. A moving window is applied to process the image and spatial connection between the image pixels is taken into account. Neural networks weights are adjusted automatically using ground-truth training data. In contrast, we propose a unifying approach to infer whether a road is passable or not, due to a severe flood event. We examine state-of-the-art classification methods with transfer learning, aiming to develop an effective road passability estimation service for the case of flooded road networks.

3. METHODOLOGY

3.1. Road passability service

In order to showcase the applicability of the proposed road passability service, we demonstrate a Web user interface that involves a classification service that adopts a DCNN architecture. As seen in Figure 1, the user is presented with a collection of satellite images, which are accompanied by their metadata (i.e., date, location, type). When an image is clicked, it is partitioned to smaller pieces and the classification method is performed to every piece. If a passable road is detected, then a green border appears around the image segment. Otherwise, a red border indicates the detection of a non-passable road. In case that no roads are recognized inside the image, no border is shown. With the results clearly illustrated, one can easily evaluate the effectiveness as well as the usefulness of the service.

3.2. Model selection and implementation

In order to classify satellite images to the class "road passability" we build models by using pretrained Convolutional Neural Networks (CNN). Namely, we experimented with the following models: VGG-19 [7], Inception-v3 [9], and ResNet [2]. VGG was originally developed for the ImageNet dataset by the Visual Geometry Group at the University of Oxford. The model involves 19 layers and it has as input images of size 224 x 224. Inception-v3 is another ImageNet-optimized model. It is developed by Google and has a strong emphasis on making scaling to deep networks computationally efficient, having as input 299 x 299 images. Finally, ResNet-50 is a model developed by Microsoft Research using a structure that uses residual functions to help add considerable stability to deep networks, using as input 224 x 224 images. For each of the aforementioned networks, we performed fine-tuning which involved removing the last pooling layers and replacing it with a new pooling layer with a softmax activation function with size 2 given that our aim is to recognize whether there is evidence of road passability or not.

For the implementation we used TensorFlow[1] and the open-source neural network Python package Keras[2] for developing our models. In general, Keras package simplifies the training of new CNN networks by modifying easily the network structure and the pre-trained weights, freezing the weights in the imported network and eventually training the weights in the newly added layers, in order to combine existing knowledge from the imported weights with the gained knowledge from the domain-specific collection of satellite images with ground-truth annotation on road passability.

4. EXPERIMENTS

4.1. Dataset description

The dataset consists of 1,437 satellite images provided for the MediaEval 2018 Satellite Task "Emergency Response for Flooding Events"[3] - data for "Flood detection in satellite images". These are satellite image patches of flooded areas that were manually annotated with a single label to indicate whether the road depicted is passable or not due to floods. The dataset was randomly split into a training a validation set. The training set contained 1,000 images, while the validation set the remaining 437 images.

4.2. Settings

Several experiments were run in order to find the best performing model. The parameters that were tuned concern the learning rate, the batch size and the optimizer function. Specifically, the values considered for the aforementioned parameters were the following: learning rate values = $\{0.001, 0.01, 0.1\}$, batch size values = $\{32, 64, 128, 256\}$, and the optimizer functions = $\{$Adam, Stochastic Gradient Descent (SGD)$\}$. Finally, the epoch was set to 35 and the loss function considered was the sparse_categorical_crossentropy.

4.3. Results

To evaluate the performance of the different networks we considered accuracy as the evaluation metric. The results of our analysis are shown in Tables 1 and 2 and in general they present the accuracy of the train and the validation set for the four networks (i.e. VGG-19, Inception_v3, ResNet-50, ResNet-101) for two widely used optimizers, i.e. Adam and SGD. Specifically, Table 1 shows how the learning parameter affects the performance of the networks. After a careful observation we can deduce that the networks perform better for the lower values of the learning rate, as they reach an average accuracy of 81.2% and 78.5% for learning rates 0.001 and 0.01 respectively.

In the sequel, we experimented with the batch size parameter and observed the impact on the networks accuracy (Table 2). The conclusion that rises from this experiment is that the increase of the batch size generally improves the accuracy. The best values of accuracy are achieved by ResNet-50 for batch size 256 and Adam optimizer (88.2%) and the Inception_v3 for batch size 128 and Adam optimizer (89.9%) (highlighted in bold). However, the accuracy of the validation set for the Inception_v3 is significantly lower than that of ResNet-50, probably due to over-fitting reasons.

5. CONCLUSION AND FUTURE WORK

In this work we presented our approach in road passability from satellite images using the recent advances in Deep Neural Networks. Tweaking the core settings of the network significant improvement in accuracy can be achieved. Better results appear with lower values of the learning ratio, while increasing the batch size improves accuracy, up to a certain level so as to avoid over-fitting. Additionally, this work highlights the necessity to evaluate alternative ways of fine-tuning pre-trained networks to compare performance differentiation.

Future work includes the combination of our approach with RCNN region proposal neural networks, to inherently perform semantic segmentation, as also described in [10], fusing heterogeneous data sources, to also highlight the road segments, in case they are not available through an external source as a GIS layer or any other format.

[1] https://www.tensorflow.org/

[2] https://keras.io/

[3] http://www.multimediaeval.org/mediaeval2018/multimediasatellite/

Table 1. Neural networks accuracy for different learning rate values.

		Learning rate 0.001		Learning rate 0.01		Learning rate 0.1	
DCNN	Optimizer	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.
VGG-19	Adam	0,85	0,6911	0,5640	0,4851	0,5700	0,5904
VGG-19	SGD	0,8600	0,7140	0,8380	0,7277	-	-
Inception_v3	Adam	0,796	0,6018	0,8120	0,5973	0,4190	0,4348
Inception_v3	SGD	0,7050	0,6590	0,8100	0,6499	0,8040	0,5995
ResNet-50	Adam	0,872	0,6247	0,8400	0,6453	0,5670	0,5973
ResNet-50	SGD	0,789	0,6865	0,8470	0,5538	0,8060	0,6796
ResNet-101	Adam	0,866	0,5515	0,8450	0,4668	0,7470	0,6590
ResNet-101	SGD	0,799	0,6041	0,8700	0,4668	0,8450	0,5858

Table 2. Neural networks accuracy for different batch size values for best performing learning rates.

			Batch size 32		Batch size 64		Batch size 128		Batch size 256	
DCNN	Learning rate	Optimizer	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.
VGG-19	0,01	Adam	0,861	0,7666	0,8610	0,7666	0,8610	**0,7667**	-	-
VGG-19	0,001	SGD	0,876	0,7071	0,8630	0,7117	0,8740	0,7162	-	-
Inception_v3	0,01	Adam	0,788	0,6247	0,8610	0,5789	0,8990	0,5629	0,8800	0,5378
Inception_v3	0,001	SGD	0,792	0,5950	0,8330	0,6224	0,8480	0,5973	0,8550	0,5995
ResNet-50	0,01	Adam	0,833	0,4943	0,8640	0,6957	0,8720	0,7094	0,8820	0,7323
ResNet-50	0,001	SGD	0,804	0,6911	0,8310	0,7094	0,8390	0,7140	0,8390	0,7185
ResNet-101	0,1	Adam	0,86	0,5492	0,8710	0,5126	0,8850	0,5126	0,8890	0,4989
ResNet-101	0,001	SGD	0,789	0,5835	0,8260	0,5995	0,8380	0,5881	0,8390	0,5812

REFERENCES

[1] Reshma Suresh Babu, B Radhakrishnan, and L Padma Suresh. Detection and extraction of roads from satellite images based on laplacian of gaussian operator. In *Emerging Technological Trends (ICETT), International Conference on*, pages 1–7. IEEE, 2016.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[3] Corentin Henry, Seyed Majid Azimi, and Nina Merkle. Road segmentation in sar satellite images with deep fully-convolutional neural networks. *arXiv preprint arXiv:1802.01445*, 2018.

[4] Wenchao Kang, Yuming Xiang, Feng Wang, Ling Wan, and Hongjian You. Flood detection in gaofen-3 sar images via fully convolutional networks. *Sensors*, 18(9): 2915, 2018.

[5] Masoud Bakhtyari Kia, Saied Pirasteh, Biswajeet Pradhan, Ahmad Rodzi Mahmud, Wan Nor Azmin Sulaiman, and Abbas Moradi. An artificial neural network model for flood simulation using gis: Johor river basin, malaysia. *Environmental Earth Sciences*, 67(1):251–264, 2012.

[6] Qian Shi, Xiaoping Liu, and Xia Li. Road detection from remote sensing images by generative adversarial networks. *IEEE access*, 6:25486–25494, 2018.

[7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014. URL `http://arxiv.org/abs/1409.1556`.

[8] Sergii Skakun. A neural network approach to flood mapping using satellite imagery. *Computing and Informatics*, 29(6):1013–1024, 2012.

[9] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[10] Wei Yao, Dimitrios Marmanis, and Mihai Datcu. Semantic segmentation using deep neural networks for sar and optical image pairs. In *Proceedings of the ESA Big data from space conference*, pages 1–4, 2017.