

USE OF MACHINE LEARNING TECHNIQUES FOR RAPID DETECTION, ASSESSMENT AND MAPPING THE IMPACT OF DISASTERS ON TRANSPORT INFRASTRUCTURE

P. M. Kikin^{1*}, A. A. Kolesnikov², A. M. Portnov³

¹ Peter the Great St.Petersburg Polytechnic University (SPbPU), St.Petersburg, Russian Federation - it-technologies@yandex.ru

² Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation - alexeykw@yandex.ru

³ Moscow State University of Geodesy and Cartography, Moscow, Russian Federation - portnov@miigaik.ru

KEY WORDS: Road Network, Machine Learning, Artificial Neural Networks, Unet, UAV, Rapid Mapping

ABSTRACT:

Road traffic infrastructure plays a key role in emergency management. It allows to evacuate people from the affected area in the shortest possible time, as well as to organize rapid emergency response. However, disasters often cause the destruction of roads, railways and pedestrian routes, which can significantly affect the evacuation plan and availability of facilities for emergency services, which increases the response time and thereby increases the losses. Therefore, it is very important to quickly provide emergency services with necessary post-disaster maps, created on the principles of rapid mapping. Change detection based on geospatial data before and after damage can make rapid and automatic assessment possible with reasonable accuracy and speed. This research proposes a new approach for detecting damage and detecting the state and availability of the road network based on the satellite imagery data, unmanned aerial vehicles (UAVs) and SAR using various methods of image analysis. We also provided an assessment of the resulting combined mathematical model based on neural networks and spatial analysis approaches.

1. INTRODUCTION

Natural disasters as well as major man made incidents are an increasingly serious threat for civil society. Earthquakes, landslides, collapse and debris flow, floods etc. usually cause damage or inaccessibility to roads network. The roads network, considered as transportation lifeline, has a critical impact on rescue and reconstruction missions after earthquake.

So, effective, fast and coordinated disaster management crucially depends on the availability of real-time roads conditions information in the affected area. Emergency managers require timely and accurate information on areas affected by disasters to prioritize relief efforts and plan mitigation measures against damage (Ge L., et al. 2015). Road damage information, allows decision makers to obtain the transport accessibility and arrange relief routes. Thus, the extraction and assessment of road damage information is quite necessary.

However, in situation assessment from the ground is usually time-consuming and of limited effect, especially when dealing with large or inaccessible areas. A rapid mapping based on remote sensing data can enable fast and effective assessment and analysis of medium to large scale disaster situations with a reasonable accuracy.

Due to recent advancements in remote sensing technology, various data sources are provided which covers abundant ground information and can be acquired from satellite, aircraft, UAV, or many other platforms.

Nevertheless, traditional road damage assessment methods, based on experts' experience to interpret and recognize remote sensing images visually are time-consuming and the accuracy

sometimes can be easily affected by subjective feelings (Wang et al. 2015).

Geomatics semi-automatic techniques, traditionally based on change detection approaches that compute relevant features from a pair of multi-temporal images collected before and after a disaster, suite much better for this purpose and provide better accuracy and performance. These methods analyze the relationship between texture, spatial, or intensity changes and the degree of damage observed after the event. Damage mapping is generally performed by setting appropriated thresholds, often set by following an expert's experience or based on already known ground truth information (Adriano Bruno, et al., 2019). These approaches show acceptable success for damage assessment (Liu, W. et al., 2013, Gokon, H. et al., 2015, Karimzadeh, S. et al., 2018).

However, their transferability is generally restricted due to their site-specific thresholds and the need for an appropriate set of pre- and post-event imagery data. Considering that a suitable pair of pre- and post-event data might not be available.

In the same time usually these techniques supposed to use just one source of data, while in a time-critical disaster situation, utilization of multiple data sources is particularly desirable. Each additional data source provides extra features which can increase possible mapping efficiency and accuracy. For example, using of social media data may complement the remote sensing one, resulting in better assessment accuracy.

For example, even in high resolution SAR images roads can often be confused with other targets such as railway tracks, rivers or even tree hedges (Henry Corentin et al. 2018).

* Corresponding author

Using just aerial imagery is challenging due to tree occlusions, building shadows, and varied atmospheric and ground conditions that can cause discontinuities in predictions, limiting their applicability for routing problems (Yoni Nachmany et al. 2019).

So multisource data fusion can increase road segmentation accuracy. In order to use all these heterogeneous data sources together and to achieve better detection, assessment and mapping accuracy, extracting more features from the same data, it seems very prospective to use different machine learning techniques in combination with spatial analysis methods. Machine learning techniques show outstanding results in image classification tasks, that was proved many times since 2012 while ImageNet Large Scale Visual Recognition Challenge. The advances of machine learning technologies, together with satellite remote sensing data, have recently brought much attention to their applicability to damage recognition.

The use of convolutional neural networks trained on large datasets to capture more spatial context of road networks dates back to Mnih and Hinton in 2010 (Mnih et al. 2010), and in recent years, better models (Zhang et al. 2018) and datasets (Shenlong et al. 2016) have been presented for road detection, as well as innovative metrics and approaches that optimized for connectivity.

In this respect the goal of this research is to explore the possibility of using machine learning methods for rapid road damage detection, assessment and mapping employing optical remote sensing data (including satellite and unmanned aerial vehicles (UAV) imagery data) and synthetic aperture radar (SAR) data.

Developing approach will be estimated with a specific focus on:

- speed
- accuracy
- scalability for different disaster situations and territories
- data quality and quantity dependency

2. MATERIALS

Multiple earth observation data of post-disaster events, caused road network damage were available in the case of the 28 September 2018 tsunami in Indonesia. Taken datasets correspond to freely accessible data from:

- Sentinel-1;
- Sentinel-2;
- ALOS-2;
- Landsat-8.

2.1 ALOS-2 data

The L-band ALOS-2 PALSAR-2 SAR satellite imagery is administered by the Japan Aerospace Agency (JAXA). For our research we used two sets of post-event high-resolution SAR data captured on 1 and 3 October. All data were acquired in StripMap observation mode (SM2) with ground sampling distance (GSD) of approximately 5 m after preprocessing and provided in a product level 1.5, SAR amplitude image in a GeoTIFF format.

2.2 Sentinel 1 data

As a part of the European Union's Earth observation program, Copernicus, two C-band SAR satellites were launched in 2014 and 2016. These satellites continuously provide new medium resolution acquisition of the entire globe. The default acquisition modes are the interferometric wide (IW) swath mode for land, with a resolution of 5 m by 20 m, whereas maritime regions are

acquired in the extra wide (EW) swath mode, translating to a resolution of 20 m by 40 m. In general, new data are acquired every six days for Europe and every twelve days for the rest of the Earth. All data are dual-polarized, interferometric wide swath acquisitions, processed to single look complex images. This way, in addition to a simple amplitude-based analysis, the interferometric coherence of two acquisitions can be computed and studied. In reaction to the disaster, data acquisition of Sentinel-1 in this area resumed on 5 was used.

2.3 Sentinel-2 data

Sentinel-2 is a satellite multispectral Earth observation mission operated by the European Space Agency (ESA) as a part of the EU Copernicus Programme. The Sentinel-2 imagery consists of 13 bands in the visible, near-infrared, and shortwave-infrared (VNIR-SWIR) range with a field of view of 290 km and multiple ground sampling distances (GSDs) of 10 m, 20 m, and 60 m. A revisit cycle of 5 days is achieved by a constellation with two twin satellites. We used post-event datasets acquired on 17 September and 2 October, respectively. Blue, green, red, and near-infrared bands that have a 10-m GSD were used in our analysis to make the spatial resolution consistent with the other datasets.

2.4 Landsat-8 data

Landsat-8 is an American Earth observation satellite launched on February 11, 2013. Landsat 8 scene size is 170 km x 185 km. It provides nine spectral bands, including a pan band. Landsat 8 satellite captured an image of this area, on October 2, 2018. Blue, green, red, and near-infrared bands were used as well.

2.5 UAV data

Since we were not able to find any publicly available dataset for this disaster and in generally any dataset, containing unavailable roads (to our knowledge), we decided to develop our own, by searching Google Images for relevant images, using the following keywords: [Disaster | Landscape] + "aerial view" + "drone". We selected only images originating from aerial views, most likely captured by UAV. Images related to damaged, but still available for off-road transport roads, as well as images of ordinary roads were collected from many other different sources. Finally, we made a dataset consists of 97 images with unavailable roads, and 1000 with damaged and ordinary roads each one.

Since it is going to use supervised machine learning, all the training imagery data were annotated with ground-truth data. Ground-truths are binary images showing if a pixel belongs to the any type of road or to the background. To do this the Intel Computer Vision Annotation Tool was used. The roads were identified as either:

- unavailable;
- damaged;
- undamaged;

with the help of Google Earth optical images. With 'damaged' class we marked roads, that do not have critical damage and still available for off-road transport. With 'unavailable' class we marked completely unavailable roads.

Each road type was assigned in a specific label color. The masks for all road types were merged into a binary ground truth, which was then smoothed.

3. METHODS

Within this work, we proposed an approach designed for roads semantic segmentation using different remote sensing data: UAV, satellite and SAR.

Our workflow was divided into three main phases.

In the first preprocessing step, the raw remote sensing data was calibrated and converted to geocoded images.

During the second step, segmentation, to evaluate different deep-learning approaches, we built a few models for pixel-based road state recognition used:

- optical satellite data separately
- SAR satellite data separately
- optical and SAR data together
- UAV data separately

In the last step, mapping, we built a road state map from the outputs of developed models.

3.1 Pre-processing

Satellite SAR data: The processing procedures for the Sentinel-1 and ALOS-2 PALSAR-2 datasets were almost the same and were performed using Scanex SAR processor software. Slave images were coregistered to a single master image, with the additional steps of debursting and merging of subswaths for Sentinel-1. Each dual-pol acquisition was then despeckled individually using NL-SAR (Deledalle, C., et al., 2015). In addition, pairs of subsequent acquisitions were used to compute coherence estimates, again using NL-SAR for the estimation. The benefit of nonlocal methods is twofold: despeckling improves the robustness for subsequent classification, and they provide a less biased coherence estimate (Schmitt, M., et al., 2019) compared to estimators with a smaller window. Finally, all products were geocoded using the 3 s SRTM DEM.

Satellite optical data: We performed atmospheric and terrain correction on the two Sentinel-2 Level 1C datasets using Scanex Image processor.

UAV optical data: To increase UAV dataset, we employed data augmentation techniques (Krizhevsky, A., et al., 2012) to enlarge artificially the number of training images using label-preserving transformations, such as translations, transposing and reflections, and altering the intensities of the RGB channels. In this way, at every run of our training procedure, each image from the training dataset was randomly transformed before used as input to the model.

3.2 Segmentation

Fully Convolutional Neural Networks (FCNN) architecture: FCNN have proven to be very effective in areas such as image classification and recognition. It involves 4 following main operations:

- Convolution
- Non Linearity (ReLU)
- Pooling or subsampling
- Classification

The first element, a DCNN encoder, analyzes the images and outputs a cluster of predictions. The image data is gradually down-sampled, proportionately becoming more meaningful. The second element, a decoder, applies up- sampling operations to

restore the spatial properties of the predictions until the predictions share the same size as the input image. It is often done using bilinear interpolation or fractionally strided convolutions, also called deconvolutions (M. D. Zeiler et al., 2010). For classification tasks, the DCNN output is classified by fixed-size fully-connected layers, the network's bottleneck, imposing a maximum input size upstream. For segmentation tasks, FCNNs remove this input size constraint by replacing the fully-connected layers by convolutional layer.

Convolution is a mathematical operation on two functions to produce a third function that expresses how the shape of one is modified by the other. The convolution operation can be considered as the feature detector. It defines a certain number of $N \times N$ filters which slides over the image and by computing the dot product creates 'Feature Map'. Different values for such filters produce different feature maps. The CNN will learn the values of the filters used in the convolutional layers on its own but it needs to provide it with a specific parameters like filter size, number of filters, etc. Practically a big number of filters allows to extract more image features and improves the network prediction ability. Thus, it is necessary to set three important parameters before performing the convolution step: depth, stride and zeropadding.

The non-linearity (ReLU) operation replaces all negative pixels with zero value. This step introduces non linearity to respect the non-linearity of real world datasets. This operation rectifies the feature maps previously computed with only linear operations.

Posterior to this step comes the pooling operation which reduces dimensionality of our feature maps nevertheless keeping the most important information. Pooling helps to reduce the number of parameters of the network and plays a key role in controlling overfitting. It also makes the network invariant to small transformations such as translations as pooling takes the max of a local neighborhood. Pooling applies to each feature map separately.

Among many existing FCNN architectures the Linknet (Chaurasia et al. 2017) was considered the best choice (Lichen Zhou et al. 2018) according to the following advantages in our tasks: fast prediction, lower GPU memory consumption, good accuracy. The architecture of LinkNet is presented in Figure 1. Here, 'conv' means convolution and 'full-conv' means full convolution (J. Long et al. 2015).

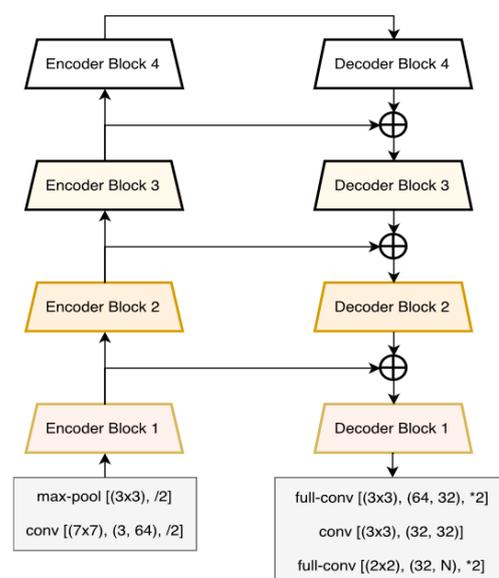


Figure 1. Linknet architecture

3.2.1 Adjusting the FCNNs for road segmentation: The above architecture was modified in order to decrease computational requirements. As the encoder we used pre-trained on ImageNet data resnet34 (Kaiming He et al. 2016) encoder. A sequence of blocks was used as a decoder: 3x3 convolution, upsampling, 3x3 convolution. Hereinafter, “convolution 3x3” is a convolution operation with a 3x3 yard, stride 1 and padding 1.

Experiments have shown that upsampling works no worse than transposed convolution, and there is no need to reduce or increase the dimension of the 1x1 convolution that was implemented in original Linknet, since the learning speed was sufficient. At the 3x3 convolution input, the features from the corresponding resnet layer are received through a skip connection, connected via the depth channel to the output of the previous layer.

As the loss function was chosen:

$$w1 * \text{binary_cross_entropy} + w2 * (1 - \text{dice}) \quad (1)$$

Here weights w1 and w2 both were set up to 0.5, but in the last two epochs, slightly more weight was given to binary cross entropy.

3.2.2 Results evaluation: To evaluate score between the predicted and ground truth binary masks built by neural network it could be simply used F1 score metric. The commonly used Pixel-level F1 score for computer vision segmentation evaluation weights each pixel equally, and a perfect score is only possible if all pixels are classified correctly as either road or background.

But, since routing is our end goal, the F1 metric is suboptimal for given that a slight error in road width is heavily penalized, though a brief break in an inferred road (from a shadow or an overhanging tree, for example) is lightly penalized.

So, to evaluate the results of routing, we used Average Path Length Similarity (APLS) metric proposed by organizers of ‘SpaceNet Road Detection and Routing Challenge’ (Adam Van Etten, 2017), which is based on graph theory and estimates graph similarity matching, focusing on the logical topology (connections between nodes) of the graph. In essence, it sums the differences in optimal paths between ground truth and proposal graphs.

To do that it uses the Average Path Length Similarity (APLS) metric that sums the differences in optimal path lengths between nodes in the ground truth graph G and the proposal graph G’ (Fig. 2).

$$C = 1 - \frac{1}{N} \sum \min \left\{ 1, \frac{|L(a,b) - L(a',b')|}{L(a,b)} \right\}$$

Figure 2. APLS metric. Node a’ is the node in the proposal graph G’ nearest the location of node a in the ground truth graph G. L(a,b) denotes a path distance in the ground truth graph G, and L(a’, b’) denotes the path length between the corresponding nodes in the proposal graph.

In effect, this metric repeats the path difference calculation shown in Figure 4 for all paths in the graph. Missing paths in the graph are assigned the maximum proportional difference of 1.0. The APLS metric scales from 0 (poor) to 1 (perfect).

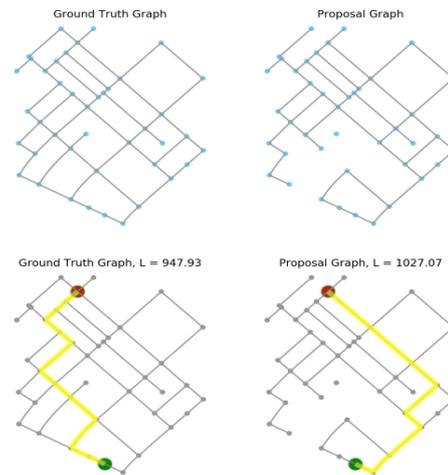


Figure 3. Demonstration of path length difference between sample ground truth and proposal graphs. Upper Left: Ground truth graph. Upper Right: Proposal graph with 30 edges removed. Lower Left: Shortest path between source (green) and target (red) node in the ground truth graph is shown in yellow, with a path length of ~948 meters. Lower Right: Shortest path between source and target node in the proposal graph, with a path length of ~1027 meters; this difference in length forms the basis for our graph similarity metric.

Since our graph have 3 different classes, we can not estimate them all at once. To do this, following the practical disaster emergency services needs, we decided to measure the score of next two graph combinations:

- only undamaged roads;
 - undamaged and damaged roads together;
- thus estimating two types of possible resulting road graphs, one for any vehicle, another only for off-road vehicles, that really makes sense while emergency management.

Additionally, F1 score was also applied to binary masks built by neural network to estimate the initial complexity of classes identification.

3.2.3 Preparing the data: We have observed in our train and test sets the presence of diagonal roads on some images. With mostly horizontal or vertical roads submitted to our CNN, the learning process will be less prone to assess for diagonal roads and it will therefore impact the good classification of our patches. In order to compensate for it, we are applying to each original image four random rotations thus increasing our amount of training data (Figure 4).

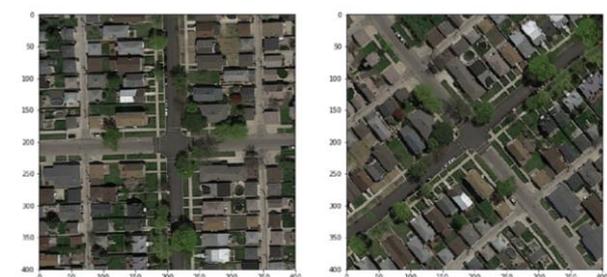


Figure 4. Training image and its rotated variation

Before FCNN training all the data was randomly divided into training and testing datasets, in proportions of 75% and 25% accordingly.

3.3 Mapping

In this step, we transferred the pixel-based classification to the road graph. Building a graph from a pixel probability map is a nontrivial task. The ‘sknw’ library was used for this task. It receives a skeleton at the input, and at the output produces a multigraph. To add vertices in points where the road changes its direction, the Douglas-Pecker algorithm in OpenCV (<http://opencv.org>) – an optimized C/C++ programming library for computer vision, machine learning and robotics – implementation was applied.

4. EXPERIMENTAL RESULTS

Finally, we had 4 pre-trained neural networks, learned using SAR only, optical only, SAR + optical and UAV only data. To estimate the results, we used test dataset consists of 25% of initial data. Crossvalidation techniques were not quite appropriate since the lack of the data.

Results of class-specific segmentation accuracies according to F1 score are shown in table 1.

| Training data | Undamaged | Damaged | Unavailable |
|---------------|-----------|---------|-------------|
| SAR | 0.45 | 0.35 | 0.31 |
| Optical | 0.81 | 0.65 | 0.52 |
| SAR + optical | 0.84 | 0.64 | 0.57 |
| UAV | 0.93 | 0.73 | 0.35 |

Table 1. Segmentation results F1 score %

Results of road graphs estimation according to APLS score are presented in table 2.

| Training data | Undamaged roads graph | Undamaged + damaged roads graph |
|---------------|-----------------------|---------------------------------|
| SAR | 0.39 | 0.35 |
| Optical | 0.58 | 0.52 |
| SAR + optical | 0.6 | 0.55 |
| UAV | 0.75 | 0.68 |

Table 2. Road graphs APLS score %

It is apparent that using of UAV data for detecting ordinary roads as well, another types of roads should have shown good results. But in case of unavailable roads model trained on UAV data showed very low F1 segmentation score. It was caused by the fact, that training dataset contained more than ten times less training images with unavailable roads and even much less unavailable roads segments.

Another models also showed a decrease in accuracy while dealing with segmentation of non-ordinary roads. The one reason was the same – data unbalancing.

But the main problem while classifying unavailable roads is variability of different reasons of roads’ unavailability and consequently big variation of different visual states of such roads, that require much more training samples than if in case of ordinary roads classification.

Moreover, very often roads may become unavailable due to its overlapping with some other object, e.g. fallen trees, parts of

neighboring infrastructure or even peaces of buildings, that makes it visibly very similar to surroundings.

In the same time if we use models just to achieve the graph of available and particularly available roads, then it doesn’t make sense if we missed the unavailable road with the background. Both means – there is no roads. Although class mismatching with another type of roads reduces validation results usability.

The radar data based FCNN shows the worst result among another’s. We assume two reasons, could influenced this result. The first one is the low accuracy of SAR data in comparison to height changes happens with roads during emergencies. Another one is the fact, that used FCNN had pre-trained on ImageNet encoder. ImageNet data have nothing similar to SAR data, though it caused accuracy decreasing.

5. CONCLUSION

The results show that developed approach based on fully convolutional neural networks using multiple sources of optical remote sensing data and SAR data is suitable for the needs of assessment and mapping of road while emergency management, in case of unavailability of pre-event roads’ mapping data.

It shows very good results if we use UAV data, providing roads graph with 75% and 68% of topological validity and 93% and 73% accuracy of roads’ classification for undamaged roads and particularly available roads accordingly, which are so necessary for emergency management to provide fast and accurate response.

In case of unavailability of UAV data, the combination of SAR and optical data also shows appropriate results, but just slightly better than single optical. Thus using only optical and combination of SAR + Optical data also could be considered applying our approach.

6. DISCUSION AND LIMITATIONS

Current UAV data based FCNN has a very low classification capability for unavailable roads, that makes it impossible for using UAV data for updating the existing accurate data. In order to improve classification capability, it needs to obtain much more UAV images with unavailable roads samples.

In current state the radar data based FCNN shows the worst and inappropriate result, and therefore cannot be used separately for disaster management. It could be significantly improved by training FCNN from scratch, not using pre-trained on ImageNet encoders.

While calculating the accuracy score it needs to take into account if the unavailable road was missed with the background, either with another road class. It would significantly improve the practical significance of the assessment for building road graph from scratch for emergency management.

Despite the case when we achieve the graph of available and particularly available roads if it needs to use our approach for existing road data updating, any misclassification should be taken into account, as we did in current research.

REFERENCES

- Adam Van Etten. SpaceNet Road Detection and Routing Challenge Part I. <https://medium.com/the-downlinq/spacenet-road-detection-and-routing-challenge-part-i-d4f59d55bfce,2017>.
- Adriano Bruno, Xia Junshi, Baier Gerald, Yokoya Naoto, Koshimura Shunichi. (2019). Multi-Source Data Fusion Based on Ensemble Learning for Rapid Building Damage Mapping during the 2018 Sulawesi Earthquake and Tsunami in Palu, Indonesia. *Remote Sensing*. 11. 886. 10.3390/rs11070886.
- Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jos e Garc ia Rodr iguez. A review on deep learning techniques applied to semantic segmentation. *CoRR*, abs/1704.06857, 2017.
- Chaurasia, Abhishek & Culurciello, Eugenio. (2017). LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation.
- Chen L.-C., Papandreou G., Kokkinos I., Murphy K., Yuille A. L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *arXiv preprint arXiv:1606.00915*, 2016.
- Deledalle, C.; Denis, L.; Tupin, F.; Reigber, A.; Jäger, M. NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising. *IEEE Trans. Geosci. Remote Sens.* 2015, 53, 2021–2038, doi:10.1109/TGRS.2014.2352555.
- Dong Liang, Timo Balz, Ziyun Wang, Lianhuan Wei & Mingsheng Liao (2017) Dynamic online visualization of PS-InSAR surface motion estimation results using WebGL, *Remote Sensing Letters*, 8:2, 126-135, DOI: 10.1080/2150704X.2016.1235807
- Ge L., Ng A.H.-M., Li X., Liu Y., Du Z., Liu Q. Near real-time satellite mapping of the 2015 Gorkha earthquake, Nepal. *Ann. GIS.* 2015;21:175–190. doi: 10.1080/19475683.2015.1068221.
- Gokon, H.; Post, J.; Stein, E.; Martinis, S.; Twele, A.; Muck, M.; Geiss, C.; Koshimura, S.; Matsuoka, M. A Method for Detecting Buildings Destroyed by the 2011 Tohoku Earthquake and Tsunami Using Multitemporal TerraSAR-X Data. *IEEE Geosci. Remote Sens. Lett.* 2015, 12, 1277–1281, doi:10.1109/LGRS.2015.2392792.
- Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press. 2016, 800 c. ISBN: 9780262035613
- Henry Corentin, Azimi Seyedmajid, Merkle Nina. (2018). Road Segmentation in SAR Satellite Images With Deep Fully Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*. 10.1109/LGRS.2018.2864342.
- J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- Karimzadeh, S.; Matsuoka, M. Building Damage Characterization for the 2016 Amatrice Earthquake Using Ascending–Descending COSMO-SkyMed Data and Topographic Position Index. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2018, 11, 2668–2682, doi:10.1109/JSTARS.2018.2825399.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 1097-1105
- Lichen Zhou, Chuang Zhang, and Ming Wu. D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- Liu, W.; Yamazaki, F.; Gokon, H.; Koshimura, S. Extraction of Tsunami-Flooded Areas and Damaged Buildings in the 2011 Tohoku-Oki Earthquake from TerraSAR-X Intensity Images. *Earthq. Spectra* 2013, 29, S183–S200, doi:10.1193/1.4000120.
- M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, “Deconvolutional networks,” in *CVPR*, San Francisco, 2010.
- Schmitt, M.; Baier, G.; Zhu, X.X. Potential of Nonlocally filtered Pursuit Monostatic TanDEM-X Data for Coastline Detection. *ISPRS J. Photogramm. Remote Sens.* 2019, 148, 130–141.
- Shenlong Wang, Min Bai, Gell’ert M’attyus, Hang Chu, Wenjie Luo, Bin Yang, Justin Liang, Joel Cheverie, Sanja Fidler, and Raquel Urtasun. *Torontocity: Seeing the world with a million eyes*. *CoRR*, abs/1612.00423, 2016.
- Volodymyr Mnih and Geoffrey E. Hinton. Learning to detect roads in high-resolution aerial images. In *Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, Computer Vision – ECCV 2010*, pages 210–223, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- Wang, J. et al., 2015. Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine. *International Journal of Remote Sensing*, 36(12), pp.3144–3169.2011. *Dengue, Urbanization and Globalization: The Unholy Trinity of the 21st Century, Tropical Medicine and Health*, 39(4), Supplement, pp. 3-11.
- Yoni Nachmany, Hamed Alemohammad; *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019, pp. 83-89
- Z. Zhang, Q. Liu, and Y. Wang. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753, May 2018.