

DETECTION OF VISIBLE BOUNDARIES FROM UAV IMAGES USING U-NET

B. Fetai^{1,*}, M. Račić¹, J. Tekavec¹, A. Liseč¹

¹ University of Ljubljana, Faculty of Civil and Geodetic Engineering, Jamova cesta 2, 1000 Ljubljana, Slovenia –
(bujar.fetai, matej.racic, jernej.tekavec, anka.lisec)@fgg.uni-lj.si

Commission I, ICWG I/II

KEY WORDS: U-Net, visible boundary, detection, UAV, cadastral mapping

ABSTRACT:

Peoples land rights are secure if they are registered in a formal cadastral system. More than 70% of global land rights are not registered in any formal cadastral system. The contemporary efforts are on accelerating the cadastral mapping process as a basis of defining land rights boundaries. Proposed surveying techniques are indirect ones – delineation of visible parcel boundaries from remote sensing imagery. This research aims at automatizing the procedure of visible boundary delineation from Unmanned Aerial Vehicle (UAV) imagery through deep learning. U-Net architecture was selected to train the model and predict visible boundaries. The model was trained on an available edge detection dataset, which was the closest to our domain problem. The model was tested on a tiled UAV images. The U-Net architecture was implemented in Keras and written in Python, running on top of the TensorFlow library. The training was done through Google Colaboratory. The evaluation metrics of the trained model indicated 0.95 overall accuracy. The average percentage of correctly detected visible boundaries was almost 80% for the tiled UAV images. This percentage is very satisfying since the model was trained on everyday imagery which is very different from UAV ones. The automatic boundary detection by using U-Net is applicable mostly for rural areas where the visibility of the boundaries is continuous. In cases where the boundaries are not visible, manual delineations are still required.

1. INTRODUCTION

Registering land rights in a formal cadastral system substantially contributes to increasing landowners tenure security. On a global scale, more than 70% of land rights are not part of any land administration or cadastral system (Enemark et al., 2014). The contemporary challenge is to accelerate and complete cadastral mapping, particularly in developing countries with low cadastral coverage. Cadastral mapping is considered the initial step when establishing a cadastral system, and it serves as a basis for defining the land units, i.e., the boundaries that the land rights concern. To accelerate cadastral mapping, the indirect surveying techniques are proposed – delineation of cadastral boundaries from high-resolution remote sensing imagery (Williamson, 2010). The direct or ground-based surveying techniques are considered as slow and expensive (Enemark, 2009; Williamson, 2010).

The application of image-based cadastral mapping holds on findings that many cadastral boundaries coincide with visible natural or human-made objects and can be easily detectable from remote sensing imagery (Luo et al., 2017). Especially, detection of visible boundaries from data acquired with sensors on Unmanned Aerial Vehicles (UAVs), have gained increasing popularity in cadastral applications. This is due to the high boundary delineation potential in urban and rural areas (Colomina, Molina, 2014; Crommelinck et al., 2016). Furthermore, UAVs can be used for both creation and updating of cadastral maps (Manyoky et al., 2012; Ramadhani et al., 2018; Koeva et al., 2018). Even though most of the visible cadastral boundaries can be detected from remote sensing imagery, many case studies reported manual delineation of cadastral boundaries (Crommelinck et al., 2016). The contemporary boundary delineation approach aims to simplify and accelerate image-based cadastral mapping through semi-automatic or automatic

detection and extraction of visible boundaries from images acquired with a high-resolution optical sensor.

Only a limited number of studies have investigated the automatic approach for cadastral boundary delineation. Mainly, tailored object-based workflows using detection algorithms were applied to automatize the cadastral mapping procedure. Investigating the technical transferability of object-based workflows is a continuing trend, especially within UAV-based cadastral mapping. For instance, both the gPb contour detection and Envi Feature Extraction (FX) module and have proven almost 80% correctness of automatically extracted visible cadastral boundaries (Crommelinck et al., 2017; Fetai et al., 2019). Considering state-of-the-art methods for automatic boundary detection in cadastral mapping, deep learning is becoming highly prominent in cadastral applications (Ma et al., 2019).

Recent evidence indicates that deep learning ensures higher accuracy on delineating visible boundaries rather than a few of object-based methods (Xia et al., 2019). The study from Crommelinck et al. (2019) reported that Convolutional Neural Network (CNN), namely VGG19 architecture, provides a more automated and more accurate approach for visible boundary delineation compared to Random Forest (RF) machine learning approach. Also, the model based on VGG19 architecture provided more promising loss and accuracy metrics compared to other CNN architectures - ResNet, Inception, Xception, MobileNet, and DenseNet. In line with this, improving the accuracy of automatic visible boundary detection remains a major challenge in a contemporary cadastral mapping. Another CNN architecture that has not been sufficiently explored for such a purpose in cadastral applications is U-Net.

U-Net was initially developed for biomedical image segmentation. The architecture is designed to work with fewer

* Corresponding author

training images and still provide precise segmentations (Ronneberger et al., 2015). In general, it is claimed that the main challenge with CNN architectures is a large amount of training data preparation and computational requirements. Thus, providing thousands of UAV training data can be seen as a limitation for the detection of visible boundaries with CNNs. Considering this, the main objective of this study is to explore the potential of U-Net, based on online available training samples, as a deep learning-based detector for visible boundaries.

2. MATERIALS AND METHODS

2.1 U-Net

The network is symmetric and contains two main parts, which gives it the U-shaped architecture (Figure 1). The left part, contracting path, is a typical convolutional network that consists of repetitive usage of convolutions (3x3 convolutions), each followed by a Rectified Linear Unit (ReLU) and a max-pooling operation (2x2 convolutions). During the contraction path, contextual information, i.e., the depth of the images is increased while the spatial information, i.e., the extent of images is reduced. The purpose of the contraction path is to capture the context of the input image. The right part, expansive path merges the contextual information and spatial information through a sequence of up-convolutions (2x2 transposed convolution) and concatenations the information with the corresponding cropped contextual map from the contracting path. During the expansion path, the extent of the image is going to be upsized to its original size. The expanding path aims to enable precise localization combined with the contextual information from the contracting path (Ronneberger et al., 2015).

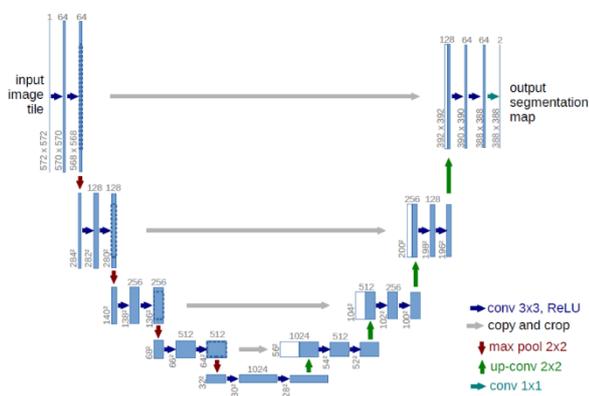


Figure 1. The original architecture of U-Net

2.2 Training approach and dataset

In this research, U-Net was implemented in high-level neural network API Keras (François et al., 2015) written in Python, running on top of the TensorFlow library. The implementation of U-Net in Keras was done by modifying and referencing to Cieslik (2017) repository. The training of the model was done through Google Colaboratory, which provided a stronger GPU, more memory, and efficient calculations.

The generalized workflow of the training approach, which aims to detect visible boundaries or edges from tiled UAV images, is presented in Figure 2.

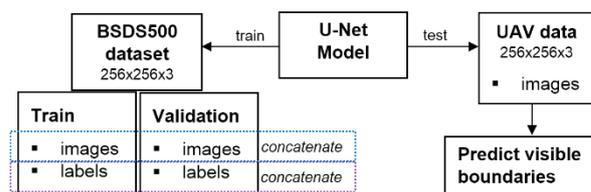


Figure 2. The generalized workflow of the trained model based on U-Net architecture

In deep learning, CNN models can be trained in two approaches: from scratch or via transfer learning (Wani et al., 2020). The model was trained from scratch based on BSDS500 - dataset. Later, the trained model was applied to the tiled UAV images - testing dataset. The tiled UAV images represent rural areas - as it is assumed that the number of visible parcel boundaries relevant for cadastral mapping is higher compared to the dense urban ones (Luo et al., 2017).

The BSDS500 is one of the few accessible datasets for edge or contour detection which can be used for training CNNs, and at the same time fit our domain problem. Contours are usually defined as object boundaries, which are derived from connecting edges. It consists of 500 everyday images. The images and corresponding contours/boundaries are organized in three subsets, namely, Train (200 images), Validation (100 images), and Testing (200 images). Each image has hand-labelled boundaries (on average 5 annotators), i.e., ~2500 samples if each annotator is considered separately. The BSDS500 dataset is available in (Berkeley, 2011). In this study, the focus was only on Training and Validation images. To increase the number of training samples and at the same time to increase the flexibility in the validation split, the images from training and validation subsets were concatenated. Besides, the target image size was set to 256 pixels in a row and height. The tiled UAV images, with the same image size as training images, were used as testing data.

The UAV images were captured on October 19th, 2018 in the noontime (good weather conditions, clear sky) using DJI Phantom 4 Pro, at 80 m flight altitude. The UAV images were captured with digital camera 1" CMOS 20mp with a focal length of 24mm. The selected rural area included roads, agricultural fields, hedges, and tree groups, which are assumed to indicate cadastral boundaries. The planimetric accuracy assessment of the UAV orthoimage was based on the comparison between surveyed Ground Control Points (GCPs) and the coordinates of GCPs on the UAV orthoimage. The estimated root-mean-square-error (RMSE) was 2.5 cm.

In order, to fit the size of testing images with the training ones the UAV orthoimage was cropped to 256x256 pixel tiles. To increase the field of view it was necessary the original spatial resolution of the UAV orthoimage to be resampled to a larger Ground Sample Distance (GSD), namely from 2 cm to 25 cm. For each tiled UAV image, manually were digitized the ground-truth boundaries. The ground-truth data were buffered to 0.50 cm and converted from vector to raster. The labels contained two classes, namely, 'boundary (1)' and 'no-boundary (0)'. The ground-truth boundaries, i.e., labels for the tiled UAV images were needed to perform the accuracy assessment for the testing dataset.

2.3 Accuracy Assessment

The accuracy assessment in this study mainly investigates two aspects: the evaluation of the U-Net model, and the evaluation of the detection quality for the testing data, i.e., tiled UAV images.

The model was evaluated by monitoring the loss and accuracy of training and validation data. The loss is defined as the sum of errors for each example in training between labels and predictions. To maximize the model's efficiency, the error or loss should be minimized. In this study, binary cross-entropy was used as a loss function and is expressed with the following equation (Alshaiikli et al., 2019):

$$L(y, p(y)) = -\frac{1}{N} \sum_{i=0}^n (y * \log p(y)) + (1 - y) * \log(1 - p(y)) \quad (1)$$

where $p(y)$ – predicted value
 y – the true label
 N – number of samples

To assess the performance of the model based on U-Net, overall accuracy was used as an evaluation metric. The overall accuracy is expressed with the following equation:

$$OA = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

where the definitions of TP, FP, FN, and TN are shown in Table 1, which represents the confusion matrix.

True label	Predicted	
	Boundary (1)	No-boundary (0)
Boundary (1)	True Positive (TP)	False positive (FP)
No-boundary (0)	False negative (FN)	True Negative (TN)

Table 1. Confusion matrix

The same confusion matrix was used to assess the detection quality of visible boundaries on tiled UAV images. The quality of detection was expressed with the error of omission, commission, and kappa coefficient. All predictions with a value < 0.5 were defined as 0 – ‘no-boundary’, and predictions with value ≥ 0.5 , were defined as 1 – ‘boundary’. The calculation of errors was done by using GRASS GIS (version 7.4.2) functionalities.

3. RESULTS AND DISCUSSION

The model was trained based on the original architecture of U-Net, considering the same layer depth and same convolutional layers. The input image size was set to 256x256 pixels. To avoid the resizing of the output image from the max-pooling operation, the padding was set to ‘same’. Also, as an optional function, the dropout rate of 0.5 was used. The Sigmoid was applied as a final activation layer to retrieve the predictions as it is well suited for a problem of binary classification. During the training ‘Adam’ was set as the optimizer, and the learning rate was set to 0.0001.

The model was trained with a batch size of 32 for 100 epochs. In addition, the Early Stopping function was set. This function aims to avoid overfitting or underfitting of the trained model, i.e., stops the training once the model performance stops improving on a

hold out validation dataset. The number of steps per epoch was calculated by dividing the total number of training samples with the batch size. The validation split was set to 0.1. The concatenated training data and validation data from BSDS500 resulted in 1469 samples for training and 164 samples for validation. The training duration with Google Collaboratory, for 100 Epochs lasted 82 minutes. This amount of training sample is considered a small dataset, especially for deep learning. However, the loss and accuracy metrics provided interesting results. The results from the training of the model are shown in Figure 3.

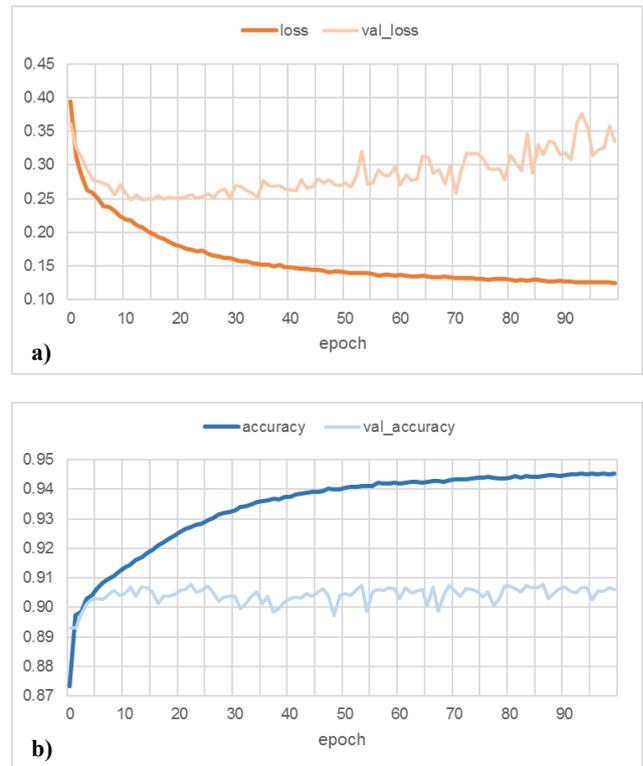


Figure 3. Evaluation of the model trained on U-Net architecture. (a) The Loss. (b) The overall accuracy.

Model train	Accuracy	Loss
at 50 Epochs	0.939	0.142
at 100 Epochs	0.945	0.125

Table 2. The results of the evaluation metrics

The model loss was constantly decreasing from the very first epoch until the end. This is an indicator that the model still is learning on training samples. However, the validation loss until epoch 35, initially was decreasing, and later mostly constant. This was a good sign that the model is not losing the ability to generalize predictions for datasets that were not seen by the model during the training. After epoch 35, the validation loss started very slowly to increase its value, which was an indicator that our model is slightly getting overfitted. Besides, in this study the model is trained on data very different from the testing data by its nature and content. If the network overfitted it would have been a problem to perform well on the data that was not in the training set. In such a case, the model can make accurate predictions for a certain dataset but fails to generalize its learning capacity for another dataset (Wani et al., 2020).

The evaluation metric indicated relatively high accuracy 0.87 from the first epoch and resulting in 0.95 in the last epoch. It is important to point out that from epoch 50 to 100, the accuracy was improving very slightly (Table 2). The relatively high accuracy, from the very first epoch, is mainly due to the imbalanced dataset. The boundaries take up a minimal area of the images compared to the background pixels, and the addressed problem is a single, not multiple class. Furthermore, considering that this study aims to use U-Net architecture to detect visible boundaries, based on the BSDS500 contours, it can be expected that the distribution of pixels per class is highly imbalanced.

The predictions for the tested data, i.e., tiled UAV images were in the range [0-1]. To perform the accuracy assessment, it was required a post-processing step. The prediction maps with a range [0-1] were reclassified. With the re-classification, the predictions < 0.5 were classified as 0 or 'no-boundary', whereas the predictions ≥ 0.5 were classified as 1 or 'boundary'. This was done to match the class values from ground truth data with the predicted class values. The post-processing step can serve as a filtering approach by customizing the class value, which is supposed to be reclassified to 0 or 'no-boundary'. Examples of predictions for UAV tiles are presented in Figures 4, 5, and 6. The results of the accuracy assessment are presented for each example of the tested image in Tables 3, 4, and 5.

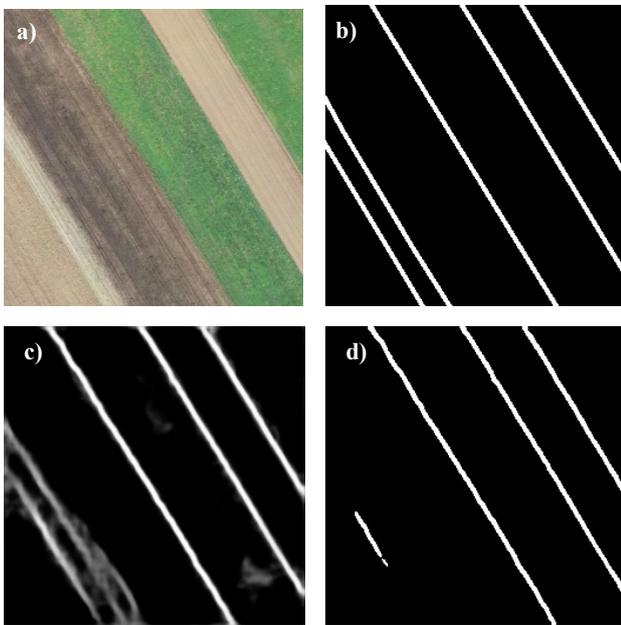


Figure 4. Predictions of the trained model based on U-Net. (a) Example 1 of tiled UAV image. (b) ground truth boundaries. (c) prediction map [0-1]. (d) re-classified prediction map [0,1].

Class	Error of omission [%]	Error of commission [%]	Kappa
Boundary	16.89	39.96	0.68
NoBoundary	2.84	0.88	

Table 3. Accuracy Assessment for Example 1.

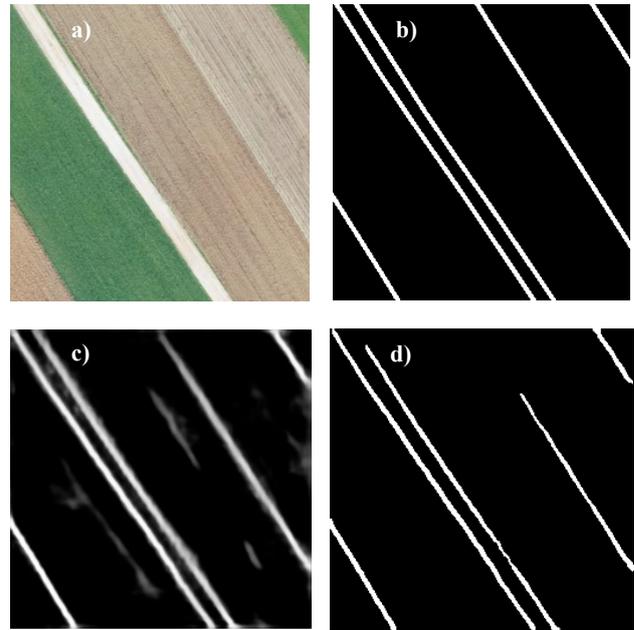


Figure 5. Predictions of the trained model based on U-Net. (a) Example 2 of tiled UAV image. (b) ground truth boundaries. (c) prediction map [0-1]. (d) re-classified prediction map [0,1].

Class	Error of omission [%]	Error of commission [%]	Kappa
Boundary	15.91	27.40	0.76
NoBoundary	1.82	0.92	

Table 4. Accuracy Assessment for Example 2.

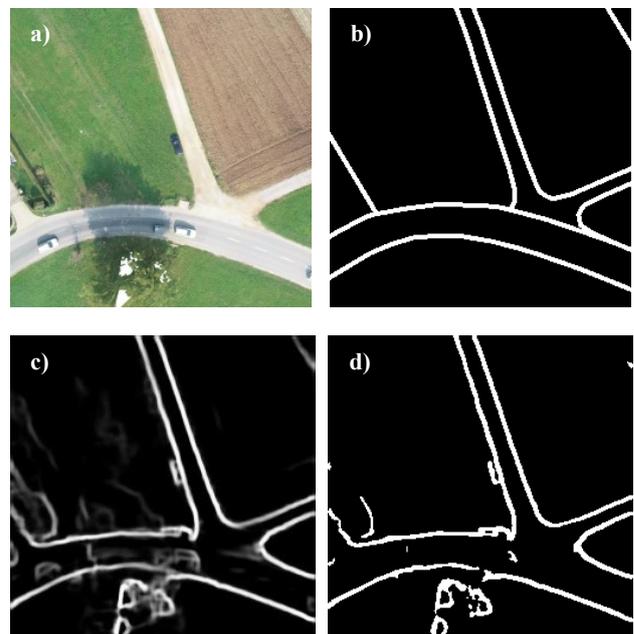


Figure 6. Predictions of the trained model based on U-Net. (a) Example 3 of tiled UAV image. (b) ground truth boundaries. (c) prediction map [0-1]. (d) re-classified prediction map [0,1].

Class	Error of omission [%]	Error of commission [%]	Kappa
Boundary	38.95	36.39	0.59
NoBoundary	2.79	3.10	

Table 5. Accuracy Assessment for Example 3.

From the predictions retrieved on testing data it can be seen that the error of omission and commission differs for three different examples. Also, the errors per ‘boundary’ class should be considered as relevant despite the ‘no-boundary’ class, which represents the background. Due to the imbalanced distribution of pixels per class, the tables of accuracy assessment indicate a very low error of omission and commission for the ‘no-boundary’ class. The average error of omission and commission for 10 tiled UAV images for the class ‘boundary’ is 21.7% and 44.3%, respectively. These values relate to the correctness of 78.3% and completeness of 55.7%. From the provided examples above, it is obvious that the reclassification contributed to decreasing the percentage of completeness.

4. CONCLUSION

This research aims at exploring the potential of original U-Net architecture as a deep learning-based detector for visual land boundaries. The training of the model was based on an available edge dataset. The results show that deep learning-based edge or boundary detection usually are faced with the imbalanced distribution of pixels per class (‘boundary’, ‘no-boundary’). This also influences the overall accuracy of the trained model – which provides relatively high accuracy in the first epochs of the training due to a large number of pixels per tile as ‘no-boundary’. This would be more sensitive when a model is trained from scratch based on remote sensing data for boundary detection problems - rather than when training a model by transfer learning. Training the model in Google Colaboratory was efficient and sustainable for the amount of the data used in this research.

The trained model for the testing data provided very satisfying predictions taking into account that the model was trained on data that are quite different in nature and content from the testing data. The predictions for the testing UAV images resulted in almost 80% of correctly detected visible boundaries. However, the quality of detection is in close relation with the selection of the case study (degree of visible boundaries) and images used for training the network. The automatic boundary detection by using U-Net is applicable mostly for rural areas where the visibility of the boundaries is continuous. In cases where the boundaries are not visible, manual delineations are still required. In general, for developing countries, the automatic detection of visible boundaries might be seen as a promising approach to accelerate cadastral mapping. In countries, with complete cadastral coverage, the same approach might use for automatic revision of existing cadastral maps and automatically define the areas where the updating is needed. Here it has to be emphasized that not all visible boundaries coincide with the real property, i.e., parcels’ boundaries to which rights, restrictions, and responsibilities refer. Therefore the detected visible boundaries have to be verified in the field. Anyhow, the approach can accelerate cadastral mapping by providing the initial dataset on parcel boundaries that are later on verified by the local people.

The research findings have been based on the BSDS500 dataset applied to the original architecture of U-Net. One of the aims of the authors’ further research is to optimize the U-Net model by transfer learning or train from scratch based on remote sensing imagery. Moreover, the sensitivity of different hyper-parameter

settings on the learning optimizer and rate, the depth of the connected layers, and the dropout rate have to be analyzed.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support of the Slovenian Research Agency (research core funding No. P2-0406 Earth observation and geoinformatics).

REFERENCES

- Alshaiikli, T., Liu, W., Maruyama, Y., 2019. Automated Method of Road Extraction from Aerial Images Using a Deep Convolutional Neural Network. *Applied Sciences*, 9 (22), 4825. <https://doi.org/10.3390/app9224825>
- Berkeley, U.C., 2011. Computer Vision Group - Contour Detection and Image Segmentation - Resources. URL <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html> (accessed on 10 March 2020).
- Cieslik, J., 2020. Developing a deep learning edge detector to solve a “toy” problem. URL <http://www.kubacieslik.com/developing-deep-learning-edge-detector-solve-toy-problem/> (accessed on 7 April 2020).
- Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 79–97. <https://doi.org/10.1016/j.isprsjprs.2014.02.013>
- Crommelinck, S., Bennett, R., Gerke, M., Nex, F., Yang, M., Vosselman, G., 2016. Review of Automatic Feature Extraction from High-Resolution Optical Sensor Data for UAV-Based Cadastral Mapping. *Remote Sensing*, 8, 1–28. <https://doi.org/10.3390/rs8080689>
- Crommelinck, S., Bennett, R., Gerke, M., Yang, M., Vosselman, G., 2017. Contour Detection for UAV-Based Cadastral Mapping. *Remote Sensing*, 9, 171. <https://doi.org/10.3390/rs9020171>
- Crommelinck, S., Koeva, M., Yang, M., Vosselman, G., 2019. Application of Deep Learning for Delineation of Visible Cadastral Boundaries from Remote Sensing Imagery. *Remote Sensing*, 11, 2505. <https://doi.org/10.3390/rs11212505>
- Enemark, S., 2009. Land Administration and Cadastral Systems in support of Sustainable Land Governance - a Global Approach, in: Re-Engineering the Cadastre to Support E-Government. Presented at the 3rd Land Administration Forum for the Asia and Pacific Region, PCGIAP, Tehran, Iran, pp. 53–71.
- Enemark, S., Bell, K.C., Lemmen, C., McLaren, R., 2014. *Fit-For-Purpose Land Administration: Joint FIG / World bank publication*. FIG, Copenhagen, Denmark.
- Fetai, B., Oštir, K., Kosmatin Fras, M., Liseč, A., 2019. Extraction of Visible Boundaries for Cadastral Mapping Based on UAV Imagery. *Remote Sensing*, 11(13), 1510. <https://doi.org/10.3390/rs11131510>
- François, C., and others, 2015. Keras. URL <https://github.com/fchollet/keras> (accessed on 20 January 2020)
- GRASS GIS Development Team. GRASS GIS Bringing Advanced Geospatial Technologies to the World, Version 7.4.2;

Open Source Geospatial Foundation: Beaverton, OR, USA, 2018.

Koeva, M., Muneza, M., Gevaert, C., Gerke, M., Nex, F., 2018. Using UAVs for map creation and updating. A case study in Rwanda. *Survey Review*, 50, 312–325. <https://doi.org/10.1080/00396265.2016.1268756>

Luo, X., Bennett, R., Koeva, M., Lemmen, C., Quadros, N., 2017. Quantifying the Overlap between Cadastral and Visual Boundaries: A Case Study from Vanuatu. *Urban Science*, 1, 32. <https://doi.org/10.3390/urbansci1040032>

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166–177. <https://doi.org/10.1016/j.isprsjprs.2019.04.015>

Manyoky, M., Theiler, P., Steudler, D., Eisenbeiss, H., 2012. Unmanned Aerial Vehicle In Cadastral Applications, in: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2011, 57–62. <https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-57-2011>

Ramadhani, S.A., Bennett, R.M., Nex, F.C., 2018. Exploring UAV in Indonesian cadastral boundary data acquisition. *Earth Science Informatics*, 11, 129–146. <https://doi.org/10.1007/s12145-017-0314-6>

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs].

Wani, M.A., Bhat, F.A., Afzal, S., Khan, A.I., 2020. *Advances in Deep Learning, Studies in Big Data*. Springer Singapore, Singapore. <https://doi.org/10.1007/978-981-13-6794-6>

Williamson, I.P. (Ed.), 2010. *Land administration for sustainable development*. 1. ed. ESRI Press Academic, Redlands, Calif.

Xia, X., Persello, C., Koeva, M., 2019. Deep Fully Convolutional Networks for Cadastral Boundary Detection from UAV Images. *Remote Sensing*, 11, 1725. <https://doi.org/10.3390/rs11141725>