Landslide Detection with Unmanned Aerial Vehicles

Hoai Nam Vu
Naver AI Lab
Posts and Telecommunications
Institute of Technology
Ha Noi, Viet Nam
namvh@ptit.edu.vn

Huong Mai Nguyen
University of Oulu
Oulu, Finland
huonguyen20@student.oulu.fi

Cuong Duc Pham
IVS Technology
Ha Noi, Viet Nam
cuongpd@ivstech.co.kr

Anh Dat Tran
Naver AI Lab
Posts and Telecommunications
Institute of Technology
Ha Noi, Viet Nam
trandat2128@gmail.com

Khanh Nguyen Trong
Naver AI Lab
Posts and Telecommunications
Institute of Technology
Ha Noi, Viet Nam
khanhnt@ptit.edu.vn

Cuong Pham
Naver AI Lab
Posts and Telecommunications
Institute of Technology
Ha Noi, Viet Nam
cuongpv@ptit.edu.vn

Viet Hung Nguyen
Posts and Telecommunications
Institute of Technology
Hanoi, Vietnam
nvhung_vt1@ptit.edu.vn

***Abstract*—Landslide is one of the most dangerous disasters, especially for countries with large mountainous terrain. It causes a great damage to lives, infrastructure and environments, such as traffic congestion and high accidents. Therefore, automated landslide detection is an important task for warning and reducing its consequences such as blocked traffic or traffic accidents. For instance, people approaching the disaster area can adjust their routes to avoid blocked roads, or dangerous traffic signs can be positioned in time to warn the traffic participants to avoid the interrupted road ahead. This paper proposes a method to detect blocked roads caused by landslide by utilizing images captured from Unmanned Aerial Vehicles (UAV). The proposed method comprises of three components: road segmentation, blocked road candidate extraction, and blocked road classification, which is leveraged by a multi-stage convolutional neural network model. Our experiments demonstrate that the proposed method can surpass over several state-of-the art methods on our self-collected dataset of 400 images captured with an UAV.**

***Index Terms*—Convolutional neural network, UAV, Landslide detection.**

I. INTRODUCTION

The monsoon climate countries, characterized by high temperature and humidity, can inherit a diversity of organisms, fertile soil, which brings many advantages for agriculture. However, they also suffer from many natural disasters such as heavy rains, storms, and floods every year as the other side of the problem. These natural factors combine with anthropogenic processes such as deforestation, especially at the upstream areas, can cause grave hazards, for example landslide. When it happens, roads are blocked, which causes hurdles not only in the traffic flow but generate various traffic problems in the form of congestion [1]. Sometimes this kind of disaster can lead to the risk of human life because there are people who got stuck between damaged roads.

For tropical countries, every year there are dozens of floods in the high mountain. The main cause of the flooding is heavy rainfall on steep hill areas. In addition, the low coverage of vegetation leads to rainwater accumulated quickly and flash floods happening in downstream areas. Floods occur very quickly, with high intensity after heavy rains. They often occur without clearly signs, causing the collapse of soil and trees. The people will not be able to deal with the abnormal happening in the flood.

The devastating pace of floods is enormous, which can wash away everything in its path. Among the damages caused by the flood, one of the most terrible is that roads will be destroyed at various locations, which is unable to predict accurately before it happens. This makes the travel of resident is suspended.

Therefore, it is necessary to have suitable method to detect landslides and draw a map of damaged road positions in mountainous areas to give timely information to local government and local residents. We are interested in using data captured by Unmanned Aerial Vehicles (UAV) to detect blocked road caused by landslide. In this paper, we propose a method that bases on two CNN models and one abnormal object extraction module at the middle to analyse the landslide image. The proposed method can be lightweight and deployed to embedded boards on UAV.

The remainder of this paper is structured as follows. Section 2 discusses relevant previous studies. Section 3 presents our method. The experimental evaluation is presented in Section 4, and finally, some concluding remarks and a brief discussion are provided in Section 5.

II. Related Works

Road segmentation plays a pivotal role in many computer vision-based traffic surveillance applications [2][3][4][5]. In this paper, it is one of the most important steps to detect the region of interest (ROI) for the classification stage afterward. In the literature, many algorithms have been proposed to improve segmentation performance. For example, Liu *et al.* [6] has introduced a super-pixel segmentation that is based on color and space distance, especially apply to aerial images. The features collected from super-pixel-based has gained an outperformance, compared to other local-features algorithms with 95.1% in scene recognition accuracy. In particular, Austria, Denmark, Rumania, and other European countries have put many concerns and effort into high crash road segment detection. So far, by pointing out the deficiencies of previous proposes in identifying high crash roads with fixed and floating segments, the dynamic segmentation in [7] has improved the percentage of counted accidents based on the wavelet theory [8]. Regarding the computational heavy of segmentation caused by pixel-wise operation, Martha *et al.* [9] have proposed a method that support to optimize the segmentation technique. The success of this paper can be applied in multiscale landslides, thus, easily allow for differently sized features to be identified. However, the maximum value for recognition accuracy can be reached is 77.7%, which is not feasible to apply in real-time detection. Besides, the different threshold-based methods (K-means clustering [10] or knowledge-based)can also affect the efficiency of the segmentation model.

With the rapid development of deep learning methods in the computer vision domain, most studies have now agreed on the fact that convolutional neural networks (CNNs) can be an optimal solution for powerful image classification and object recognition [11] [12]. While working with landslide detection, we are processing images acquired from flying objects such as a drone. Relating to landform recognition task, object-based approach has been used widely in drone's images [9, 13] and different characteristic of images such as colour of shadow [14] can also be used. Besides, splitting the image into grids and predicting on each tile, [15] has formulated the multi-labelling classification within a conditional random field (CRF) framework to exploit simultaneously spatial contextual information and cross-correlation between labels. To improve the accuracy of the model working with UAV images, [16] has been proposed to increase the depth of the network, which is on the other hand negatively affects the computing. Silva *et al.* [17] also combined the Mask R-CNN [18] and transfer learning with ResNet [19] 50 and 101 to improve the landslide detection performance. This work has ended up with promising evaluation results with 100% of precision. However, the combination of Mask R-CNN for segmentation and ResNet101 for classification will bring a huge delay while running the system in the real-time domain. To deal with this problem, we utilize the MobileNet [20] for our classification model because of its compactness. Since with less computation power and fewer parameters to run, the MobileNet still maintains a competitive accuracy compared to other state-of-the-art models in classification. The use of it as a backbone ensures a solid lightweight decision-making system operating in real-time. In particular, MobileNetV3 [21] was used in our study to reduce the latency and optimize computational time with piece-wise linear analogue (RELU6) instead of sigmoid.

For the specific case of image-related data, [22] has proposed an excellent work under the combination of bag-of-visual-word (BoVW), probabilistic latent semantic analysis (pLSA), and k-nearest neighbors classifier (k-NN). The basic principle is to divide the image into sub-grids, apply CSIFT (color SIFT) on each patch to find out the local features, and quantize it into words to BoVW representation. The sub-image is then classified by the pLSA model and k-NN classifier as landslide or non-landslide. The success of the work has created a simple and computationally efficient but still robust detection system.

Finally, as the growing need in using more deep convolutional networks to solve practical applications, many studies have been carried out on landslide detection over the past decade. According to that, a detailed comparison between different machine learning and deep learning methods for landslide detection has been given out in [23].

III. Methodology

Our proposed method is shown in Fig. 1, which consists of three main steps: road segmentation, blocked road candidate extraction, and classification. The first step aims to segment the road pixel from the input image. Then, the segmented road image will be the input of the next step, where we extract blocked road candidates by using edge detection and heuristic rules. Finally, the blocked road candidates are classified into real blocked roads or not.

A. Segmentation

In the segmentation step, we use SD-Unet [24] to distinguish the road pixel from the background pixel. SD-Unet is a petite version of Unet, which achieves higher performance while requiring 8x fewer computations, leading to faster inference. In the proposed work, except the first layer, Gadosey et al. has replaced all the standard convolutional layers in U-net with depth-wise separable layers. By doing this, the number of parameters and required computations in the U-Net model is much reduced. However, the use of depth-wise layer can lead to achieving lower performance compared to standard convolution layers. Therefore, different normalization and standardization method are also applied to recovered the accuracy loss.

According to the experimental result from [24], we hereby also change the core of the UNet model with depth-wise separable convolutional layers instead of the normal ones and add the group normalization after each ReLU activation. The

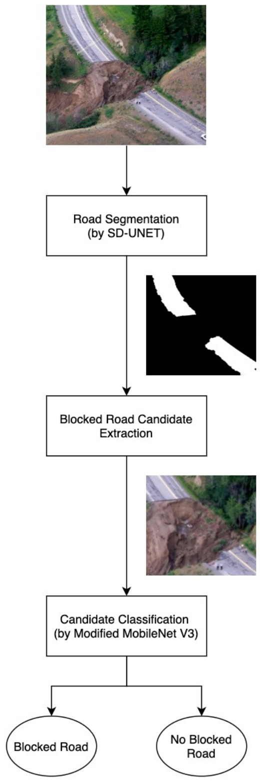

Fig. 1. Flowchart of our proposed method

elements in one SD-UNet block include 2 depth-wise separable convolution, 2 activation layers, and 1 group normalization.

Since the original SD-Unet was well trained and tested on the BRATs dataset, in this study, we trained the network with mapillary dataset [25] and fine-tuned it to fit with our customization. The model was evaluated on our self-collected dataset, which consists of 400 images. The dataset is manually annotated by LabelBox. The landslide image data was collected with an UAV (drone) at the time of heavy rains and the catastrophic floods happened.

The segmentation model has been trained to predict on every pixel of the image to determine if it is a background or foreground. Therefore, the output image of the network will be a binary array, in which, each pixel is binarized into 1 or 0 and displayed as black (0) or white (255) as the intensity value.

By doing the road segmentation as a pre-stage in the classification process, we can improve our classifier performance in detecting the landslide. This can be explained in [26] [27] and discussed further in the section IV.

B. Blocked road candidate extraction (Region of Interest - ROI)

The purpose of this stage is to narrow down the area of interest. In particular, we only focus on the region, which contains the unconnected parts of the road. By focusing on just a small area instead of a whole image, the redundant information is reduced. Therefore, the next step classifier can achieve better accuracy. In addition, we can minimize the time computational cost for our real-time detection system.

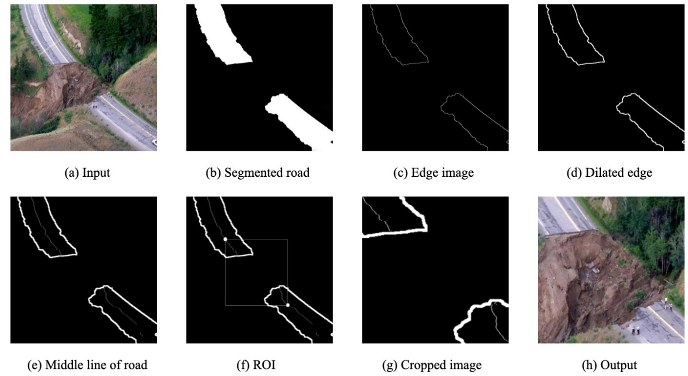

Fig. 2. Blocked road candidate extraction procedure

To detect ROI, the segmented road image from the first step will go through the Canny algorithm to detect the edges of the segmented road. As we can see from the illustration (Fig. 2c), the detected line is quite thin and blurred. However, the performance of the classifier is strongly dependent on the quality of the input image. Therefore, we enhanced the image quality by applying the dilation operation on the edge to make it bolder and clear.

The next step is to form the middle line of the road (Fig. 2e). This process requires all the middle points, which are calculated from edges to concatenate them afterward. In order to locate the disrupted area, we need to have at least 2 points on 2 different road parts. In particular, the points must be ones of the middle points we found from the previous step. With the coordinates, we can easily decide the unique rectangle to display the ROI. This can be referred to Fig. 2f. Finally, the original image is cropped with the same position as the interested region before getting the final input of the classifier (Fig. 2h).

C. Classification

At this stage, the ROI given from the previous stage becomes the input for the classifier to determine the final output. Suspicious regions can occur in the input image, in which the segmented road is discontinued due to big truck, tree shadow and other image degradation. The suspicious region occurrence reduces the accuracy of the method if we consider all blocked road candidates as real blocked roads. For practical application, we aim to design a small and precise system, which acts as an early incident detector and reduces the number of false alarm cases.

At this stage, the ROI given from the previous stage becomes the input for the classifier to determine the final output. The suspicious region can occur in the input image, in which the segmented road is discontinued due to big truck, tree shadow, and other image degradation. The suspicious region occurrence reduces the accuracy of the method if we consider all blocked road candidates as real blocked roads. For practical application, we aim to design a small and precise

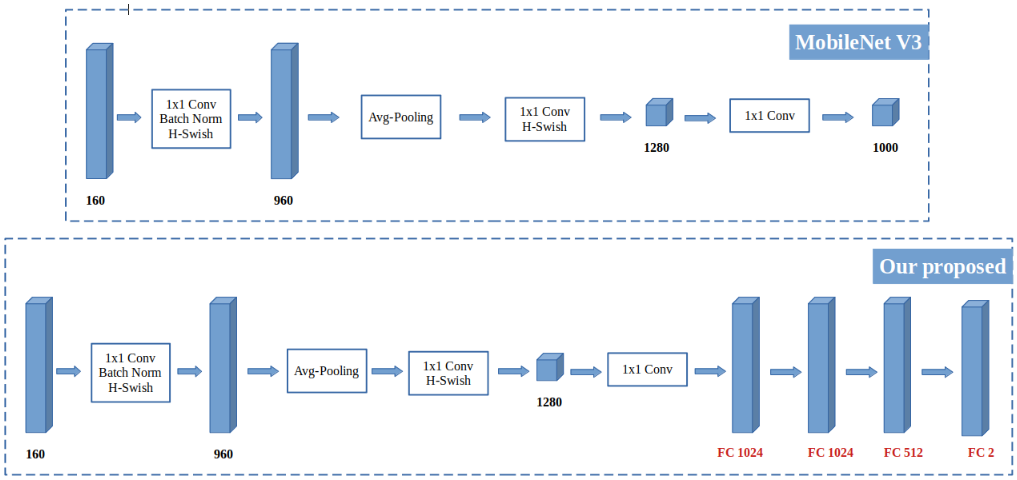

Fig. 3. The change in the last layer from MobileNet V3 to our proposed model

system, which acts as an early incident detector and reduces the number of false alarm cases. In the viewpoint of real-world application, our system has to operate in the context where computational power is limited, therefore, we implemented MobileNet network [20] to do classification. The MobileNet still maintains a competitive accuracy compared to other state-of-the-art models while ensuring real-time performance. In particular, MobileNet V3 [21] was used in our study to reduce the latency and optimize computational time.

By using hard swish (h-swish) at the second half of the model, coupling with the squeeze-and-excitation (SE) block, MobileNetV3 has gained significant improvements in performance compared to the previous version at almost no computational cost. In detail, the new non-linearity activation: h-swish has fixed some weakness of its base by replacing the computationally expensive sigmoid with a piecewise linear analogue (RELU6):

$$swish(x) = x \cdot \sigma(x) \tag{1}$$

$$h - swish(x) = x\frac{RELU6(x+3)}{6} \tag{2}$$

In order to train the model, we first split the data into 2 separated folders: *normal* and *abnormal*. By utilizing the Mobilenet V3 Large, we discarded the last 1000 neurons layer then added up 4 fully connected layers so that the model can learn more sophisticated functions and give more accurate results (Fig. 3). The last dense layer contains 2 nodes which are corresponding to 2 labels. We also did augmentation with the data to prevent the model from overfitting. The transformation includes rotation, zooming, shifting, and flipping.

IV. Experimental Results

In this section, we will show and discuss the results of all experiments in our work. For more specification, we compared the performance amongst models in 3 schemes such as segmentation, classification, and whole method. In order to get a fair comparison between all methods, we have set up the same hardware configuration (CPU: Intel Core i7; RAM: 8GB; GPU: RTX 2070 Super) and dataset (our collected one).

In the first step of road segmentation, we compare our modified SD-Unet model with other SOTA (state of the arts) methods. We utilize two metrics for the comparison which are Intersection over Union (IoU) and dice coefficient. In particular, IoU reflects the similarity between 2 boxes or 2 areas by calculating the percentage of overlap, hereby, the region is between the ground truth (GT) and the predicted segmentation outputs (PO). This metric is formatted in equation 3. The dice coefficient (equation 4) on the other hand also measures the percentage of overlap, however, it reckons one more time at the overlap region between the ground truth and the prediction.

$$IoU(GT, PO) = \frac{|GT \cap PO|}{|GT \cup PO|} = \frac{TP}{TP + FP + FN} \tag{3}$$

$$Dice(GT, PO) = \frac{2|GT \cap PO|}{|GT \cup PO|} = \frac{2TP}{2TP + FP + FN} \tag{4}$$

As can be seen from Table I, our modified SD-UNet achieved the best performance with the highest dice-coefficient 69% at the inference time of 13ms. Regarding the effectiveness confront, Attention Unet [28] shows its robustness of 69% in IoU and 92% in dice-coefficient. However, the structure complexity leads to their huge computational time of 67 ms. With an IoU value under the average, only achieved 47%, UNet has registered as the worst model for road segmentation. There is also a paradox in this experiment is that the LedNet[29] on the other hand outperformed UNet with 52% in IoU and 85% in dice-coefficient, concurrently it required a longer time to compute an instance (60ms) while it has been proposed as the lightweight version of UNet.

From Table II, we have proved our right choice in model selection with the experiments to compare the MobileNetV3 between different variants of EfficientNet [30] from B0 to B7. The EfficientNet was chosen for comparison because it has been proposed as an enhanced version of MobileNet by

TABLE I
DIFFERENT SEGMENTATION METHOD IN COMPARISON

Model	IoU	Dice Coefficient	Inference time (ms)
UNet	0.47	0.82	54
LedNet [29]	0.52	0.85	60
Attention Unet [28]	**0.69**	0.92	67
SD-UNet	0.65	**0.93**	**13**

fusing with ResNet, EfficietnNet. In addition, its variants have achieved better performance compared to original MobileNet models. At the same time, this also leads to an extra latency in inference time, approximately 2-6 times slower. We choose MobileNet V3 Large with a proper balance between F1-score and inference time.

TABLE II
CLASSIFICATION MODELS COMPARISON AMONG MOBILENET V2, MOBILENET V3 AND EFFICIENTNET VARIANTS

Model	F1-score	Inference time (ms)
MobileNet V2	0.9138	89.1968
MobileNet V3 Small	0.9483	65.0501
MobileNet V3 Large	**0.9655**	**89.0919**
EfficientNet B0	0.9483	117.4593
EfficientNet B1	0.9828	144.8111
EfficientNet B2	0.9828	151.6269
EfficientNet B3	0.9655	186.2684
EfficientNet B4	0.9655	230.3987
EfficientNet B5	0.9655	299.4162
EfficientNet B6	0.9655	377.5594
EfficientNet B7	0.9828	517.8696

The most interesting comparison is made to prove the effectiveness and feasibility of our proposed method. We compare our method with several scenarios where only a single stage is used such as classification, detection, or segmentation. For example, a CNN model for image recognition can be used to identify whether an image contains a blocked road or not. Alternatively, object detection models can also be utilized to detect the occurrence of blocked roads in the input image. Finally, segmentation models with heuristic rules are also able to detect the occurrence of blocked roads in the image.

TABLE III
ACCURACY COMPARISON BETWEEN SOTA METHODS

Method	F1 Score
MobileNet	0.822
ResNet	0.852
VGGNet	0.841
YOLO	0.92
Faster RCNN	0.91
SSD	0.893
UNet + rules	0.814
SDNet + rules	0.823
FCN + rules	0.807
Our proposed method	**96.2**

In the three kinds of methods mentioned above, the image classification-based method gives a low accuracy because the blocked road is only a small part of the image, therefore the important information used to distinguish is not focused. While object detection-based methods give high accuracy. However, this kind of method comes with a high false alarm rate. A false alarm occurs when multiple non-blocked road objects are detected in the input image. Segmentation-based methods give low efficiency when the data is highly diverse. Table III shows the detailed results of the methods. Our proposed method can achieve the highest F1 score of 96.2%, which is 4.2% higher than the runner-up method (YOLO). The worst performance is from the FCN segmentation with heuristic rules at F1 score of 0.807. These results have proven that our combination approach is necessary to achieve high performance in detecting blocked roads in the input image.

V. CONCLUSION

We have presented a multistage convoluttional neural architecture for the detection of blocked roads caused by landslide. The network can detect blocked roads with three main steps: road segmentation, blocked road candidate extraction, and blocked road candidate classification. In order to evaluate the effectiveness of the proposed architecture, we collected a dataset of 400 images captured using a UAV. The experiments with several state-of-the art models including MobleNet, ResNet, VGG, YOLO, SSD, Faster RCNN, UNet + rules, SDNet + rules, and FCN + rules show that our proposed method can achieve up to 96% F1-score, which is better than other methods. These results indicate the promising of the proposed method.

ACKNOWLEDGMENT

Hoai Nam Vu was funded by Vingroup Joint Stock Company and supported by the Domestic Master/ PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Vingroup Big Data Institute (VINBIGDATA), code VINIF.2020.TS.103.

REFERENCES

[1] Manh Hung Nguyen, Tuong Vinh Ho, Trong Khanh Nguyen, and Minh Duc Do. Modeling and simulation of the effects of landslide on circulation of transports on the mountain roads. *International Journal of Advanced Computer Science and Applications*, 6(8), 2015. doi: 10.14569/IJACSA.2015.060835. URL http://dx.doi.org/10.14569/IJACSA.2015.060835.

[2] Marcelo Santos, Marcelo Linder, Leizer Schnitman, Urbano Nunes, and Luciano Oliveira. Learning to segment roads for traffic analysis in urban images. In *2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast City, Australia, June 23-26, 2013*, pages 527–532. IEEE, 2013. doi: 10.1109/IVS.2013.6629521.

[3] Nam Vu Hoai, Cuong Pham, Nguyen Manh Dung, and Soonghwan Ro. Detecting and tracking sinkholes using

multi-level convolutional neural networks and data association. *IEEE Access*, 8:132625–132641, 2020.

[4] Cuong Pham and Nguyen Thi Thanh Thuy. Real-time traffic activity detection using mobile devices. In *Proc. of ACM IMCOM*, pages 64:1–64:7. ACM, 2016.

[5] Nam Vu and Cuong Pham. Traffic incident recognition using empirical deep convolutional neural networks model. In *ICCASA*, pages 90–99, 2017.

[6] Hongguang Li, Yang Shi, Baochang Zhang, and Yufeng Wang. Superpixel-based feature for aerial image scene recognition. *Sensors*, 18(1), 2018. ISSN 1424-8220.

[7] Amin Boroujerdian, M. Saffarzadeh, Hassan Yousefi, and Hassan Ghassemian. A model to identify high crash road segments with the dynamic segmentation method. *Accident Analysis Prevention*, 73:274–287, 09 2014. doi: 10.1016/j.aap.2014.09.014.

[8] Stéphane Mallat. *A Wavelet Tour of Signal Processing, 2nd Edition*. Academic Press, 1999. ISBN 978-0-12-466606-1.

[9] Tapas Martha, Norman Kerle, C.J. Westen, V.G. Jetten, and Kumranchat vinod Kumar. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 49:4928 – 4943, 12 2011. doi: 10.1109/TGRS.2011.2151866.

[10] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu. An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):881–892, 2002. doi: 10.1109/TPAMI.2002.1017616.

[11] Lin Du, Xiong You, Ke Li, Liqiu Meng, Gong Cheng, LiYang Xiong, and Guangxia Wang. Multi-modal deep learning for landform recognition. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158:63–75, 12 2019.

[12] Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel J. Mollura, and Ronald M. Summers. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Medical Imaging*, 35(5):1285–1298, 2016.

[13] Andre Stumpf and Norman Kerle. Object-oriented mapping of landslides using random forests. *Remote Sensing of Environment*, 115:2564–2577, 10 2011.

[14] Anita Simic Milas, Kristin Arend, Christine Mayer, Martin A. Simonson, and Scudder Mackey. Different colours of shadows: classification of uav images. *International Journal of Remote Sensing*, 38(8-10):3084–3100, 2017.

[15] Abdallah Zeggada, Souad Benbraika, Farid Melgani, and Zouhir Mokhtari. Multilabel conditional random field classification for UAV images. *IEEE Geosci. Remote. Sens. Lett.*, 15(3):399–403, 2018.

[16] Ulzhalgas Seidaliyeva, Manal Alduraibi, Lyazzat Ilipbayeva, and Nurzhigit Smailov. Deep residual neural network-based classification of loaded and unloaded UAV images. In *Fourth IEEE International Conference on Robotic Computing, IRC 2020, Taichung, Taiwan, November 9-11, 2020*, pages 465–469. IEEE, 2020. doi: 10.1109/IRC.2020.00088.

[17] Silvia Liberata Ullo, Amrita Mohan, Alessandro Sebastianelli, Shaik Ejaz Ahamed, Basant Kumar, Ramji Dwivedi, and G. R. Sinha. A new mask R-CNN based method for improved landslide detection. *CoRR*, abs/2010.01499, 2020.

[18] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask r-cnn. pages 2980–2988, 10 2017. doi: 10.1109/ICCV.2017.322.

[19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 770–778. IEEE Computer Society, 2016. doi: 10.1109/CVPR.2016.90.

[20] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *CoRR*, abs/1704.04861, 2017.

[21] Andrew Howard, Ruoming Pang, Hartwig Adam, Quoc V. Le, Mark Sandler, Bo Chen, Weijun Wang, Liang-Chieh Chen, Mingxing Tan, Grace Chu, Vijay Vasudevan, and Yukun Zhu. Searching for mobilenetv3. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 1314–1324. IEEE, 2019. doi: 10.1109/ICCV.2019.00140.

[22] Gong Cheng, Kaiming Li, Tianyun Zhao, Junwei Han, Huihui Li, and Jun Fang. Automatic landslide detection from remote-sensing imagery using a scene classification method based on bovw and plsa. *International Journal of Remote Sensing*, 34:45–59, 01 2013.

[23] Omid Ghorbanzadeh, Thomas Blaschke, Khalil Gholamnia, Sansar Raj Meena, Dirk Tiede, and Jagannath Aryal. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote. Sens.*, 11(2):196, 2019.

[24] Pius Gadosey, Yujian Li, Enock Adjei Agyekum, Ting Zhang, Joying Liu, Peter Yamak, and Firdaous Essaf. Sd-unet: Stripping down u-net for segmentation of biomedical images on platforms with low computational budgets. *Diagnostics*, 10:110, 02 2020. doi: 10.3390/diagnostics10020110.

[25] Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulo, and Peter Kontschieder. The mapillary vistas dataset for semantic understanding of street scenes. In *Proceedings of the IEEE international conference on computer vision*, pages 4990–4999, 2017.

[26] Andrew Rabinovich, Andrea Vedaldi, and Serge Belongie. *Does Image Segmentation Improve Object Categorization?* 01 2007.

[27] Merve YILDIZ ERDEMIR, Taskin Kavzoglu, Ismail

Colkesen, and Emrehan Sahin. An assessment of the effectiveness of segmentation methods on classification performance. 07 2012.

[28] Ozan Oktay, Jo Schlemper, Loic Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas. 04 2018.

[29] Yu Wang, Quan Zhou, Jia Liu, Jian Xiong, Guangwei Gao, Xiaofu Wu, and Longin Jan Latecki. Lednet: A lightweight encoder-decoder network for real-time semantic segmentation. pages 1860–1864, 09 2019.

[30] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *ICML 2019 USA*, pages 6105–6114, 2019.