



Combining Object-Oriented and Deep Learning Methods to Estimate Photosynthetic and Non-Photosynthetic Vegetation Cover in the Desert from Unmanned Aerial Vehicle Images with Consideration of Shadows

Jie He¹, Du Lyu^{2,3}, Liang He¹, Yujie Zhang¹, Xiaoming Xu⁴, Haijie Yi², Qilong Tian², Baoyuan Liu¹, Xiaoping Zhang¹, Jose Alfonso Gomez⁵, Josef Krasa⁶, Tomas Dostal⁶, and Tomas Laburda⁶

¹Institute of Soil and Water Conservation, Northwest A&F University, Yangling, Shaanxi 712100, China

²Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, China

³Shaanxi Satellite Application Center for Natural resources, XiAn, 710002, China

⁴College of Urban, Rural Planning and Architectural Engineering, Shangluo University, Shangluo 726000, China

⁵Institute for Sustainable Agriculture, CSIC, Cordoba, Spain

⁶Department of Landscape Water Conservation, Czech Technical University, Prague

Soil erosion is a global environmental problem. The rapid monitoring of the coverage changes in and spatial patterns of photosynthetic vegetation (PV) and non-photosynthetic vegetation (NPV) at regional scales can help improve the accuracy of soil erosion evaluations. Three deep learning semantic segmentation models, DeepLabV3+, PSPNet, and U-Net, are often used to extract features from unmanned aerial vehicle (UAV) images; however, their extraction processes are highly dependent on the assignment of massive data labels, which greatly limits their applicability. At the same time, numerous shadows are present in UAV images. It is not clear whether the shaded features can be further classified, nor how much accuracy can be achieved. This study took the Mu Us Desert in northern China as an example with which to explore the feasibility and efficiency of shadow-sensitive PV/NPV classification using the three models. Using the object-oriented classification technique alongside manual correction, 728 labels were produced for deep learning PV/NPV semantic segmentation. ResNet 50 was selected as the backbone network with which to train the sample data. Three models were used in the study; the overall accuracy (OA), the kappa coefficient, and the orthogonal statistic were applied to evaluate their accuracy and efficiency. The results showed that, for six characteristics, the three models achieved OAs of 88.3–91.9% and kappa coefficients of 0.81–0.87. The DeepLabV3+ model was superior, and its accuracy for PV and bare soil (BS) under light conditions exceeded 95%; for the three categories of PV/NPV/BS, it achieved an OA of 94.3% and a kappa coefficient of 0.90, performing slightly better (by ~2.6% (OA) and ~0.05 (kappa coefficient)) than the other two models. The DeepLabV3+ model and corresponding labels were tested in other sites for the same types of features: it achieved OAs of 93.9–95.9% and kappa coefficients of 0.88–0.92. Compared with traditional machine learning methods, such as random forest, the proposed method not only offers a marked improvement in

classification accuracy but also realizes the semiautomatic extraction of PV/NPV areas. The results will be useful for land-use planning and land resource management in the areas.