学 位 論 文 内 容 の 要 旨

Dissertation Abstract

博士 (環境科学)　　　　氏　名　JUAN XIAO

学 位 論 文 題 名

Title of dissertation

Enhancing crop monitoring using unmanned aerial vehicle (UAV) images and deep learning
(ドローン画像とディープラーニングを用いた作物モニタリングの強化)

Effective crop monitoring is essential for maximizing crop yield, promoting sustainable agriculture, and ensuring food security. Unmanned Aerial Vehicle (UAV) is an important platform for precision agriculture that provides images with precise and accurate spatial and spectral information about crops. Deep learning (DL) techniques have demonstrated significant potential in feature extraction and non-linear mapping. Therefore, this research aimed to enhance crop monitoring by integrating UAV images and DL techniques. Firstly, the research used UAV, DL, and geographic information systems (GIS) technology to monitor corn growth performance across different management practices. A literature review was then conducted on the SpatioTemporal Fusion (STF) of remote sensing images to reveal the potential of UAV image prediction for crop monitoring. Finally, this research proposed a DL-based STF model for centimeter-scale UAV image prediction to provide multi-temporal UAV images for enhancing crop monitoring.

Two experimental corn fields were divided into four plots to assess the effects of varying corn management practices (i.e., seeding schedule, planting depth, and fertilization method) on corn growth performance. RGB and multispectral cameras were mounted on UAVs to collect corn field images. Plant height, Normalized Difference Vegetation Index (NDVI), Normalized Difference Red-Edge Index (NDRE), plant density, and plant volume were mapped based on the UAV images. YOLOv5 model and UAV images were investigated for

counting corn plants. Additionally, the Otsu thresholding method was evaluated as an automatic method for extracting plant height, NDVI, and NDRE values from the background. YOLOv5 and Otsu thresholding were efficient and accurate for automatically counting corn plants and extracting corn heights as well as VIs, respectively. The emergence rates of corn seeds were 40%, 33%, 41%, and 62% in plots A, B, C, and D, respectively. Variations in corn field management practices significantly affected the emergence rate, with fertilizer application close to seeds emerging as the optimal practice for achieving higher emergence rates across experimental plots. This study used UAVs and DL techniques to provide precise information and valuable insights into corn field practices, which can help farmers to optimize corn cultivation. The techniques applied in this study could be extrapolated to improve cultivation processes for other crops.

Results from the literature review demonstrate that more than 100 STF models have been proposed during the past two decades. However, many of these models were not practically applied. Therefore, this research provides a comprehensive review of the conception, development, challenges, and applications of different STF categories. Specifically, the focus is on examining the architectures and strategies of DL-based STF models that use convolutional neural networks (CNNs). The number of these CNN-based STF models has increased sharply since 2018. The findings of this review can be used to help guide the selection and design of STF models to promote their applications. In addition, future directions for STF modeling were proposed, such as the requirement of benchmark datasets, the promotion of the practical applications, and the inclusion of centimeter-scale images acquired by UAVs for frequent precise monitoring.

State-of-the-art STF models fuse images from various satellites, not satisfying the demand for precise crop monitoring. The review demonstrated that DL-based STF models have a high potential for higher-resolution image fusion. To this end, this research proposed an end-to-end DL-based STF model, named UAV-Net, that can predict centimeter-scale UAV images. UAV-Net exhibits an encoder-decoder architecture with Modified ResNet (MResNet), Feature Pyramid Network (FPN), and decoder modules. The encoder used MResNet modules to extract input features, while the FPN module performs a multiscale fusion of the features extracted by the encoder before reconstructing UAV images using transposed convolution in the decoder module.

Through the comparison and ablation experiments, this study evaluated the efficacies of the MResNet modules with 18, 34, and 50 layers, along with the FPN module of UAV-Net. The experimental results on real-world datasets demonstrated that UAV-Net adequately predicted UAV images both visually and quantitatively. The mean value of four metrics, namely structural similarity index measure (SSIM), correlation coefficient (CC), spectral angle mapper (SAM), and root mean square error (RMSE), for the fused UAV image on August 7 are 0.8824, 0.7595, 0.0301 and 0.0334, respectively. Similarly, for the fused image on July 31, the mean values are 0.7623, 0.7263, 0.0571 and 0.0419. These results demonstrate the accuracy and precision of the predicted centimeter-scale images, indicating the potential of UAV-Net in various precision environmental monitoring applications.

Overall, this research presents the significant potential of using UAV images and DL techniques for precise and efficient crop monitoring, thereby making valuable contributions to the achievement of food security. The methods applied in this research can be extrapolated and adopted to enhance cultivation processes for various crops. In particular, the innovative STF model proposed in this research enables the prediction of centimeter-scale UAV images, thereby extending the applications beyond agriculture to various Earth observation monitoring. These applications include areas such as forest and ecological monitoring, where the accurate and high-resolution images produced by UAV-Net can greatly facilitate improved source management practices and decision-making processes.