CWRepViT-Net: An Encoder-Decoder Deep Learning Framework with RepViT Blocks for Crop Weed Semantic Segmentation in Soybean Fields through their Life Journey

Masoomeh Gomroki[1], Dilshan I. Benaragma [1], Christopher J. Henry[2], Nasem Badreldin[3] & Robert H. Gulden[1]

1Department of Plant Science, University of Manitoba, Winnipeg Manitoba Canada

2Department of Computer Science, University of Manitoba, Winnipeg Manitoba Canada

3Department of Soil Science, University of Manitoba, Winnipeg Manitoba Canada

With the rising global population, the demand for agricultural products is increasing, necessitating advanced technologies for efficient crop management. This study focuses on precision weed management using drone imagery and Deep Learning (DL) during early soybean growth stages. Drone images were collected at six intervals (21,26, 33, 39, 45 and 52 days after seeding) to distinguish between five classes: soil, soybean, volunteer canola, broadleaf weeds, and grassy weeds. A deep learning model, Crop-Weed-RepViT-Net (CWRepViT-Net), based on an encoder-decoder architecture using RepViT and Modified UNet blocks, was developed for semantic segmentation. The five-step framework includes data preprocessing, training the network , segmentation, evaluation, and prediction. CWRepViT-Net achieved over 95% accuracy and a Kappa coefficient of 0.91, demonstrating high reliability in early-stage crop-weed differentiation.