Enhanced multi-label classification of surface damage in concrete structures using a refined yolo-v5 with spatial pyramid pooling network integration

Mei-Hsin Chen[1], Ting-Hsuan Lai[2], Tien-Yin Chou[1]*

[1] GIS Research Center, Feng Chia University, Taichung 40724, Taiwan

[2] Ph.D. Program for Infrastructure Planning and Engineering , Feng Chia University, Taichung 40724, Taiwan

The process of identifying surface damage in concrete structures, particularly in bridges, serves as a vital precursor to the expensive and lengthy structural integrity evaluation. The task is made complex by the substantial variability in the physical appearance of concrete, shifts in lighting and weather, the presence of various surface markings, and the potential for diverse types of defects to overlap. To address this issue, we utilize the YOLO-v5 object detection model for the multi-label classification of surface damage in concrete structures, and further enhance it with the proposed YOLO-v5-sppm network structure through the incorporation of our specially developed SPPM layer. The results obtained from the experiments illustrate that the upgraded YOLO-v5-sppm can deliver superior multi-target accuracy.

The SPPM (SPPnet Merge) layer, the foundation of our innovation, is inspired by the SPPnet (Spatial Pyramid Pooling network). This study demonstrates the transformation of the output layer of the original YOLO-v5 from an object detection model to a multi-label classification model by integrating an enhanced SPPM layer. The final detection layer of YOLO-v5 generates three arrays, shaped as (batch_size, anchor_number, image_height/stride, image_width/stride, class_numbers+5), with strides inferred from the YOLO-v5 network structure. As such, the three output arrays from the detection layer possess multiple interrelations, prompting the creation of the SPPM layer using AvgPool3d. Consequently, the SPPM layer can accommodate any input from the detection layer, maintaining a fixed output shape. Lastly, the addition of a linear layer and sigmoid activation facilitates the transition to the multi-label classification model. This approach provides several benefits. Primarily, it allows for the use of pre-trained weights from YOLO-v5, which have been trained using the Microsoft-sponsored COCO dataset. Secondly, the concept of image multi-scale, introduced by YOLO9000, is incorporated into YOLO-v5 training. Despite the initial input setting of 640x640 pixels, the image size fluctuates post multi-scaling. However, the kernel size and stride of the SPPM layer dynamically adapt to the input image size, ensuring that the benefits of multi-scale training are not compromised.

For data collection, we utilized the COncrete DEfect BRidge image dataset (CODEBRIM), which offers multi-label classification for five prevalent concrete defects: Cracks, Spallation, Efflorescence, Exposed Bars, and Corrosion Stain. The accuracy of the models served as the primary metric for result analysis, supplemented by precision and recall assessments. With respect to the experimental setup, the dataset was partitioned into training and validation sets, with a 75:25 split. We employed the Stochastic Gradient Descent (SGD) as the optimizer, set the image_size to 640, batch_size to 16, and trained the models for 300 epochs, across three depth of backbones, namely yolov5s, yolov5m, and yolov5l, arranged in ascending order of network complexity. YOLO-v5 retained the original loss function, while YOLO-v5-sppm adopted Binary Cross-Entropy Loss (BCELoss) as the loss function. The experimental outcomes suggest that the accuracy of YOLO-v5 is reasonably satisfactory, achieving 0.805, 0.833, and 0.824 with yolov5s, yolov5m, and yolov5l backbones respectively. However, the enhanced YOLO-v5-sppm that we proposed exhibits superior accuracy, achieving 0.841, 0.851, and 0.861 across the respective backbones, consistently outperforming YOLO-v5.

Keywords: multi-label classification, yolov5, deep learning, concrete surface damage, spp net