

# 基于改进TransUNet的沥青路面病害检测研究

程文党<sup>1</sup>, 彭福刚<sup>2</sup>, 刘文江<sup>3</sup>, 侯珂<sup>1</sup>

<sup>1</sup>山东交通学院轨道交通学院, 山东 济南

<sup>2</sup>山东科洋工控自动化有限公司, 山东 济南

<sup>3</sup>山东交通学院航空学院, 山东 济南

收稿日期: 2023年12月19日; 录用日期: 2024年1月11日; 发布日期: 2024年1月22日

## 摘要

本文研究了基于图像分割算法的沥青路面病害检测问题, 提出了一种基于TransUNet的分割网络, 用于沥青路面病害的检测和识别。该网络在TransUNet的骨干网络中引入了ECA注意力机制, 以增强特征的表达能力和自适应性。本文使用Dice评价系数作为评价指标, 对路面裂缝和坑洞进行分割, 并在公开的路面病害数据集上进行了实验。实验结果表明, 所提出的TransUNet-RD模型mIoU和Dice等方面优于原始的TransUNet模型和UNet模型, 证明了其在沥青路面病害检测方面的有效性和优越性。

## 关键词

道路病害, 图像分割, TransUNet, 注意力机制

# Research on Asphalt Pavement Disease Detection Based on Improved TransUNet

Wendang Cheng<sup>1</sup>, Fugang Peng<sup>2</sup>, Wenjiang Liu<sup>3</sup>, Ke Hou<sup>1</sup>

<sup>1</sup>School of Rail Transportation, Shandong Jiaotong University, Jinan Shandong

<sup>2</sup>Shandong Keyang Industrial Control Automation Co., Ltd., Jinan Shandong

<sup>3</sup>School of Aeronautics, Shandong Jiaotong University, Jinan Shandong

Received: Dec. 19<sup>th</sup>, 2023; accepted: Jan. 11<sup>th</sup>, 2024; published: Jan. 22<sup>nd</sup>, 2024

## Abstract

This paper studies the problem of asphalt pavement disease detection based on image segmentation algorithms, and proposes a segmentation network based on TransUNet for asphalt pavement disease detection and recognition. The network introduces the ECA attention mechanism in the back-bone network of TransUNet, to enhance the feature expression ability and adaptability. This

paper uses the Dice coefficient as the evaluation index, segments the pavement cracks and potholes, and conducts experiments on the public pavement disease dataset. The experimental results show that the proposed TransUNet-RD model is superior to the original TransUNet model and UNet model in terms of mIoU and Dice, proving its effectiveness and superiority in asphalt pavement distress detection.

## Keywords

Road Diseases, Image Segmentation, TransUNet, Attention Mechanism

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

沥青路面是一种常见的路面材料，具有耐久性、防水性和降噪性等优点。但是，沥青路面也会受到自然环境、交通荷载、施工质量等因素的影响，产生裂缝、坑洞、车辙、沉陷等病害，影响路面的使用寿命和行车安全。因此，沥青路面病害的及时检测，对于制定养护方案和保障道路运营具有重要意义。

传统的沥青路面病害检测方法主要依赖于人工巡查和评估，这种方法不仅耗时耗力，而且容易受到人为因素的干扰，导致检测结果的主观性和不一致性。近年来，随着计算机视觉技术的发展，利用图像处理和机器学习的方法，自动化地对沥青路面病害进行检测和识别，成为了一种新的研究方向。图像分割是一种将图像划分为若干个具有相同或相似特征的区域的技术，是沥青路面病害检测的关键步骤。图像分割的目的是将图像中的病害区域与正常区域分开，从而为后续的病害分类和评估提供基础。

近年来，随着深度学习技术的发展，出现了基于深度学习的缺陷检测方法。这些方法可以分为两类，根据它们的网络体系结构进行分类。第一类是双阶段检测算法，代表有区域卷积神经网络[1] (R-CNN)、Fast R-CNN [2]和 Faster R-CNN [3]等。R-CNN 利用滑动窗口生成候选框域，但这种方法步骤繁琐且计算量大。为了解决 R-CNN 中耗时的选择性搜索，Faster R-CNN 引入了区域候选网络来代替选择性搜索，如 Alzraiee 等[4]将 Faster R-CNN 用于道路表面损伤。研究证实，双阶段检测方法可用于道路等表面缺陷的检测，但中间过程复杂，需要大量重复性实验，效率不高，难以适应现场检测要求。

第二类是单阶段检测算法，代表有 CenterNet [5]和 YOLO [6]等。这类算法不需要先给出目标候选框，而是将目标图像特征提取和预测框的确定结合在一起进行图像的分类估计和区域坐标评估，从而大大提高了检测效率。然而，CenterNet 算法存在明显的缺点，如在小目标和长宽比失衡的情况下表现不佳，并且速度较慢。另外，对于长宽比失衡的目标处理可能会遇到困难。相比之下，YOLO 系列算法，如 YOLOv3 [7]、YOLOv4 [8]和 YOLOv5，在检测速率和精度上都表现出较大的优势。Xue 等人[9]提出了基于 YOLOv3 的路面病害检测方法，可以准确地检测和定位路面裂缝。此外，Rahman 等人[10]提出基于 YOLOv4 的路面病害检测方法，可以同时检测和定位多种类型的路面病害，如裂缝、坑洞和表面破坏。这些基于 YOLO 的方法在路面病害检测方面具有较好的性能。

本文提出了一种基于图像分割算法的沥青路面病害检测方法，该方法采用了 TransUNet 作为图像分割模型，TransUNet 是一种将 Transformer 和 U-Net 结合起来的混合模型，既能够利用 Transformer 的全局自注意力机制来提取图像中的长距离依赖关系，又能够利用 U-Net 的跳跃连接来保留低层次的细节信息，从而实现高效和准确的图像分割。

## 2. 模型介绍及改进

### 2.1. TransUNet 模型框架

TransUNet 是一种将 Transformer 和 U-Net 结合起来的混合模型，用于图像分割任务。它的主要思想是利用 Transformer 的全局自注意力机制来提取图像中的长距离依赖关系，同时利用 U-Net 的跳跃连接来保留低层次的细节信息，从而实现高效和准确的图像分割。

TransUNet 的模型框架如图 1 所示，它由编码器和解码器两部分组成。编码器由一个卷积层和一个 Transformer 层构成，解码器由若干个上采样模块和一个卷积层构成。编码器和解码器之间通过跳跃连接进行特征融合，以增强分割性能。

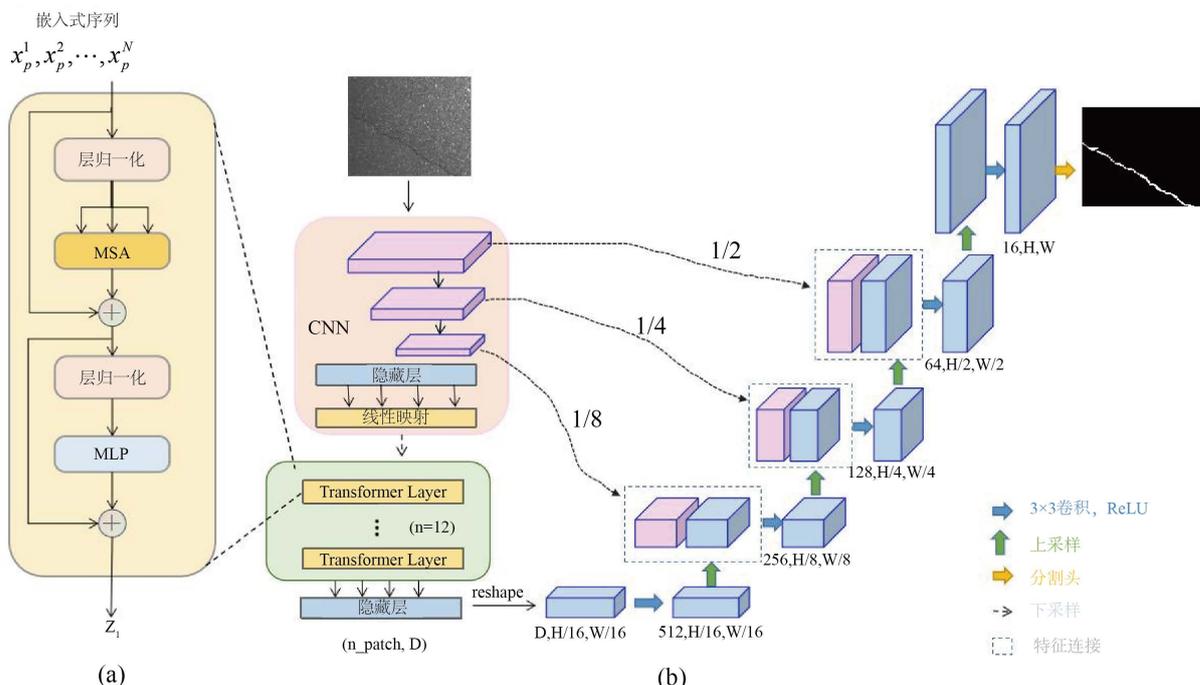


Figure 1. TransUNet model framework

图 1. TransUNet 模型框架

编码器的卷积层用于将输入图像转换为一系列的标记化图像块，每个图像块对应于一个特征向量。这些特征向量作为 Transformer 层的输入序列，经过位置编码和多头自注意力机制，得到编码后的特征序列。Transformer 层的结构和参数设置只是将图像分类的输出层替换为一个线性层，用于输出编码后的特征序列。

解码器的上采样模块由一个转置卷积层和一个卷积层构成，用于将编码后的特征序列逐步上采样到原始图像的分辨率。在每个上采样模块中，通过跳跃连接将编码器的卷积层输出的特征图与上采样后的特征图进行拼接，以融合不同层次的信息。解码器的最后一个卷积层用于输出分割结果，即每个像素的类别概率。

TransUNet 的模型框架简单而有效，能够充分利用 Transformer 和 U-Net 的优点，同时避免了它们的缺点。相比于传统的基于 CNN (Convolutional Neural Network) 的图像分割模型，如 U-Net，TransUNet 能够更好地捕捉图像中的长距离依赖关系，从而提高分割的精度和鲁棒性。相比于纯粹的基于 Transformer

的图像分割模型,TransUNet 能够更好地保留低层次的细节信息,从而提高分割的定位能力和边缘清晰度。

## 2.2. 改进沥青路面病害检测模型

本研究针对沥青路面病害的图像分割问题,提出了一种基于 TransUNet 的分割网络,(TransUNet of Road Detection, TransUNet-RD)。为了进一步提升模型的性能,本研究在 TransUNet 的 backbone 网络中引入了 ECA (Efficient Channel Attention)注意力机制并使用 Dice 评价系数作为评价指标。

### 2.2.1. ECA 注意力机制

ECA 是一种注意力机制,旨在提高网络的表征能力。它不仅减少了参数和计算量,还降低了注意力机制的重复率。与 SE 注意力机制类似,ECA 主要关注通道间的相关性建模,但具有更高的计算效率。

图 2 展示了 ECA 网络结构模型。

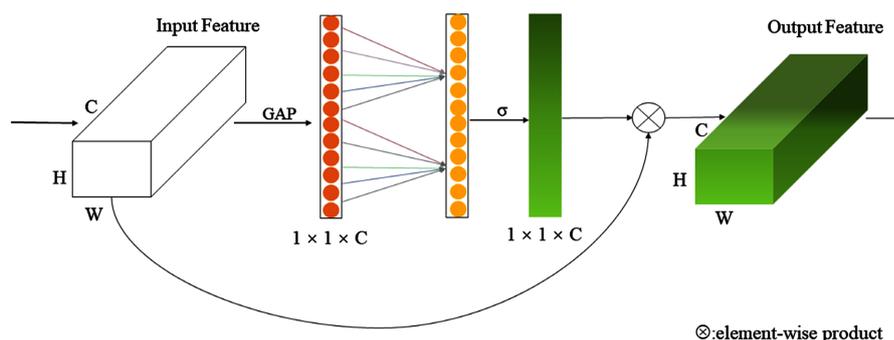


Figure 2. ECA network architecture diagram  
图 2. ECA 网络结构图

ECA 的工作原理如下:首先,对输入的特征图进行全局平均池化操作,将每个通道的特征图转换为一个标量值。这个标量值表示了该通道的全局特征重要性。接下来,使用窗口大小为  $k$  的一维卷积操作,对全局平均池化的结果进行处理。该一维卷积操作可以学习通道间的相关性。它在每个位置上对前后通道的特征进行卷积,生成一个新的通道特征表示。为了确保注意力权重的范围在 0 到 1 之间,一维卷积的输出经过 Sigmoid 函数进行激活。这样可以获得每个通道的注意力权重,用于对原始特征图进行加权。最后,将注意力权重乘以原始特征图,得到经过 ECA 注意力机制加权后的特征。

相较于其他注意力机制,ECA 的特点在于其高效性。通过使用一维卷积而不是全连接层,ECA 可以减少参数数量和计算量,从而降低了注意力机制的重复率。这种优化策略使得在沥青路面病害检测任务中,我们能够以最小的计算代价提高模型的准确率。

### 2.2.2. 添加 Dice 评价系数

为了评估图像分割的效果,我们采用 Dice 评价系数作为性能指标。Dice 评价系数是一种衡量两个集合相似度的方法,通常用于计算预测的分割结果和真实的分割结果之间的一致性。Dice 评价系数的定义如公式(1)所示:

$$\text{Dice}(A, B) = \frac{|A| + |B|}{2 * |A \cap B|} \quad (1)$$

其中,  $A$  和  $B$  是两个集合,  $|A|$  和  $|B|$  分别表示集合  $A$  和  $B$  的元素个数,  $|A \cap B|$  表示集合  $A$  和  $B$  的交集的元素个数。Dice 评价系数的值越接近 1,表示两个集合越相似;越接近 0,表示两个集合越不相似。

为了计算 Dice 评价系数，我们首先需要将预测的分割结果和真实的分割结果转换为二值图像，即将沥青路面病害区域的像素值设为 1，其余的像素值设为 0。然后，我们根据上述公式计算每张图像的 Dice 评价系数，最后求平均得到整个数据集的 Dice 评价系数。

Dice 评价系数作为一种常用的分割性能评价指标，能够全面衡量模型的性能，特别关注模型对路面病害的准确检测。Dice 评价系数的计算方式使得模型更注重准确捕捉目标区域，有助于降低重复率，避免生成过多的重复或错位的预测结果。此外，引入 Dice 评价系数还能促使模型学习更准确的边界信息，提高沥青路面病害检测任务的分割精度。在训练过程中，通过调整损失函数，将 Dice 评价系数纳入优化目标，有助于模型生成更精确、非重复的预测，从而综合提高检测鲁棒性。

### 3. 实验

#### 3.1. 实验环境及数据集

实验在 Windows 操作系统环境下进行，使用的 GPU 是 NVIDIA RTX A5000，显卡大小为 8 GB。设置了 300 个 epochs 进行训练，每个 batch 的大小为 8，使用 AdamW 优化算法进行参数优化，初始学习率设置为 0.01。输入图片的分辨率为  $512 \times 512$ 。

实验数据集是由 Roboflow 网站在 2023 年公开的，用于路面病害检测的数据集。该数据集包含裂缝和坑槽这两种路面病害，共有 2797 张有效缺陷图像。考虑到数据样本较少的情况，使用了数据增强技术对缺陷图像样本进行扩充。通过随机旋转、剪切变换、位置平移、随机缩放等操作，将原始数据扩充为 6005 张有效缺陷图像。为了进行训练和验证，将数据集按照 8:2 的比例随机分为训练集和验证集。

#### 3.2. 评价指标

为了评估 TransUNet-RD 在沥青路面病害分割任务上的性能，本文采用了平均交并比 mIoU (mean Intersection over Union) 作为评价指标。mIoU 是一种常用的语义分割评价指标，它计算了所有类别的交集和并集之比的平均值，反映了分割结果与真实标签的重叠程度。mIoU 的计算公式如公式(2)所示：

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{|P_i \cap T_i|}{|P_i \cup T_i|} \quad (2)$$

其中， $N$  是类别的总数， $P_i$  是第  $i$  类的预测结果， $T_i$  是第  $i$  类的真实标签， $|\cdot|$  表示集合的元素个数， $\cap$  表示集合的交集， $\cup$  表示集合的并集。mIoU 的取值范围是  $[0, 1]$ ，值越大，表示分割结果越好。

mIoU 是一种全局的评价指标，它考虑了所有类别的分割效果，能够反映模型的整体性能。但是，mIoU 也有一些局限性，比如它不能区分不同类别的重要性，也不能反映分割结果的定位精度和边缘清晰度。因此，本文还使用了 Dice 评价系数这一评价指标来综合进行评价。

#### 3.3. 实验结果与分析

为了验证本文所提出的算法的有效性，进行了 3 次实验，逐步验证了网络结构变化对性能的影响。这些实验使用了不同的模型进行检测，包括 UNet 模型、TransUNet 模型以及本文提出的在 backbone 网络中添加了 ECA 注意力模块和使用 Dice 评价系数的 TransUNet-RD 模型。这些模型经过训练，并在三个实验中进行了评估。

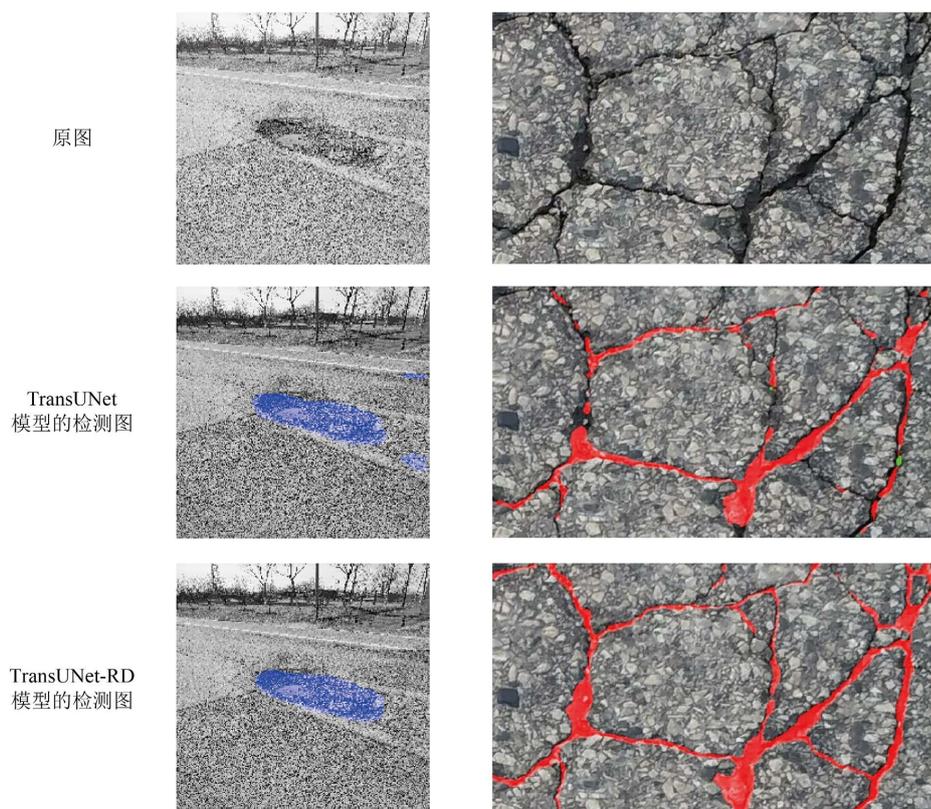
从表 1 中可以看出，改进后的 TransUNet-RD 相较于 TransUNet 的原生模型对两种缺陷的检测效果均有提升。其中，改进后的 TransUNet-RD 相较于 TransUNet 原生模型 mIoU 提升了 5.3%，Dice 评价系数提升了 6.9%；相较于 UNet 模型平均精度 mIoU 提升了 10.2%，Dice 提升了 11.2%。

**Table 1.** Comparison of the improved TransUNet and original TransUNet effects  
**表 1.** 改进的 TransUNet 与原始 TransUNet 效果对比

Model	Dice	mIoU		
		Crack	Pothole	All
UNet	0.801	0.612	0.774	0.693
TransUNet	0.844	0.688	0.796	0.742
TransUNet-RD	0.913	0.767	0.823	0.795

对验证集中的样本进行随机抽样，路面缺陷检测效果如图 3 所示。从左至右依次为原图、原生 TransUNet 模型的检测图和本文改进的 TransUNet-RD 模型的检测图。从下图可视化结果中可以看到，在该数据集中，原生 TransUNet 模型虽然能够提取出路面病害轮廓，但其检测效果相对较差。在路面坑槽病害检测中存在明显的误检问题，在复杂的路面裂缝病害检测中存在裂缝病害轮廓提取不完整，且泛化能力较差等问题。

相比之下，TransUNet-RD 模型的检测效果明显更为出色。该模型在坑槽和裂缝轮廓提取上表现出较高的精度，同时成功降低了误检率和漏检率。这表明本文所提出的改进模型在路面缺陷检测任务中取得了显著的性能提升。



**Figure 3.** Images of road surface defect detection results  
**图 3.** 路面缺陷检测效果图

综合以上分析，改进后的 TransUNet-RD 模型在沥青路面病害检测方面表现出显著的优势。通过引入

ECA 注意力机制和 Dice 评价系数, 成功提升了模型的精确度和效率, 使其在 mIoU 和 Dice 评价系数等关键指标上均超过了传统的 TransUNet 和 UNet 模型。这些研究结果强调了在路面病害检测领域, 通过创新性的改进, 能够显著提升检测性能。

#### 4. 结语

本研究旨在改进用于检测路面缺陷的 TransUNet 模型的性能。研究提出了几种对模型的修改, 包括在骨干网络中增加 ECA 注意力机制和使用 Dice 评价系数进行评价。实验结果表明, 所提出的 TransUNet-RD 模型在 mIoU 和 Dice 等方面优于原始的 TransUNet 模型。未来的工作将专注于扩展对其他类型的路面缺陷(如龟裂、车辙和修补)的检测能力。总体而言, 对 TransUNet 模型的这些修改能够提供一种更准确、更高效的解决方案, 用于提取路面缺陷。我相信通过持续地研究和改进, 基于深度学习的路面病害检测技术将为道路维护和交通安全带来革命性的改变。

#### 课题来源

交通运输行业重点科技专项, 项目编号: 2022-ZD4-066。

#### 参考文献

- [1] Girshick, R., Donahue, J., Darrell, T., *et al.* (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [2] Girshick, R. (2015) Fast R-CNN. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 07-13 December 2015, 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>
- [3] Ren, S., He, K., Girshick, R., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems*, **28**, 1137-1149.
- [4] Alzraiee, H., Leal Ruiz, A. and Sprotte, R. (2021) Detecting of Pavement Marking Defects Using Faster R-CNN. *Journal of Performance of Constructed Facilities*, **35**, Article 04021035. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0001606](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001606)
- [5] Duan, K., Bai, S., Xie, L., *et al.* (2019) Centernet: Keypoint Triplets for Object Detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-02 November 2019, 6569-6578. <https://doi.org/10.1109/ICCV.2019.00667>
- [6] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [7] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. <https://arxiv.org/abs/1804.02767>
- [8] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. <https://arxiv.org/abs/2004.10934>
- [9] Xue, Y., *et al.* (2020) YOLONG: A YOLO-Based Approach for Road Crack Detection. *IEEE Access*, **8**, 133235-133245.
- [10] Rahman, S.M.M., *et al.* (2021) Deep Learning Based Asphalt Pavement Disease Detection Using YOLOv4. *IET Intelligent Transport Systems*, **15**, 229-236.