Deep Neural Network based Attention Model for Structural Component Recognition

Paper ID : 143

Sangeeth Dev Sarangi and Bappaditya Mandal

Keele University, Newcastle-under-Lyme, ST5 5BG, United Kingdom

Introduction

- Computer vision and machine learning based concrete structural health inspection/monitoring bring great benefits such as better safety and security for humans, non-contact, at a long distance, rapid, cheap cost and labour, and low interference with the regular functioning of infrastructures.

- The technique of locating and identifying distinctive sections of a structure using structural component recognition is anticipated to be a crucial first step in the automated inspection/management of civil infrastructure.

Aim

- Our research aims to develop novel contextual information in the deep convolutional neural network coupled with an attention model for the automatic recognition of civil structural components on image/video data.

Method

- The proposed DNNAM architecture consists of synchronous dual attention modules (SDAM) and residual modules, which together aid in the extraction of crucial discriminative characteristics from various scales to enhance the performance of both multi-target multi-class and single-class classification

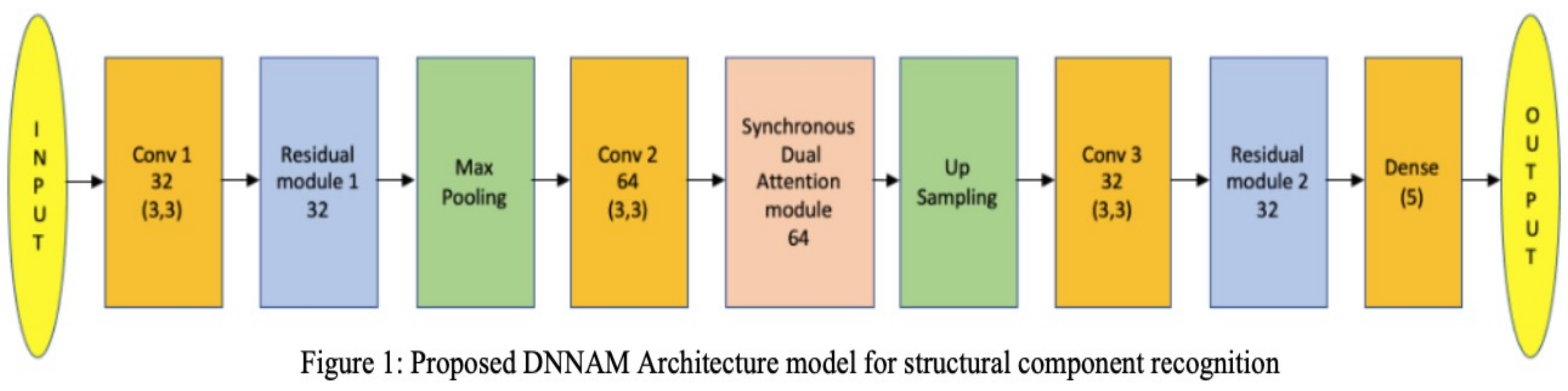

Figure 1: Proposed DNNAM Architecture model for structural component recognition

- The synchronous dual attention module fuses crucial attention operations, including self, spatial, and channel attention synchronously and aggregates their results to highlight discriminative features for multiple target structural component classes.

- The SDAM module is comprised of two modules:
  - Batch of multi-feature attention module
  - Parallel excitation module.

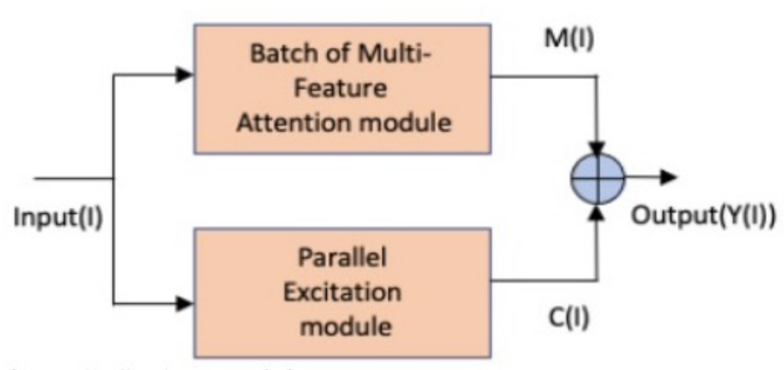

Figure 2: SDAM module

Results

- We have evaluated our algorithms and compared them with the existing methods on the benchmarking dataset for bridge component classification provided by the authors of Narazaki et al., 2017; 2020;

- Our proposed DNNAM model achieves a mean IOU of 65.94% with pixel wise accuracy of 82.85%.

Table 1: Comparison with Pre-existing Benchmarks on the Bridge component classification dataset. We are considering mIOU as more important than PA, Perhaps because of a lack of high-quality ground creation, mIOU is better than PA on this dataset.

Benchmarking Works	mIOU(%)	PA(%)
CNPT - N[1]	50.8	80.3
CPNT - Scene [1]	-	82.4
FCN45 [2]	-	82.3
FCN45 - N [3]	57.0	84.1
FCN45- P [3]	56.9	84.1
FCN45- S [3]	56.6	83.9
SegNet45- N [3]	54.5	82.3
SegNet45 - P [3]	55.2	82.9
SegNet45 - S [3]	55.2	82.9
SegNet45-S - N [3]	55.8	83.1
SegNet45-S - P [3]	55.9	83.3
SegNet45-S - S [3]	55.4	82.7
StructureNet [4]	57.46	89.08
DNNAM	**65.94**	**82.85**

[1](Narazaki et al., 2017) [2](Narazaki et al., 2020) [3](Yeum et al., 2019) [4](Kaothalkar et al., 2022)

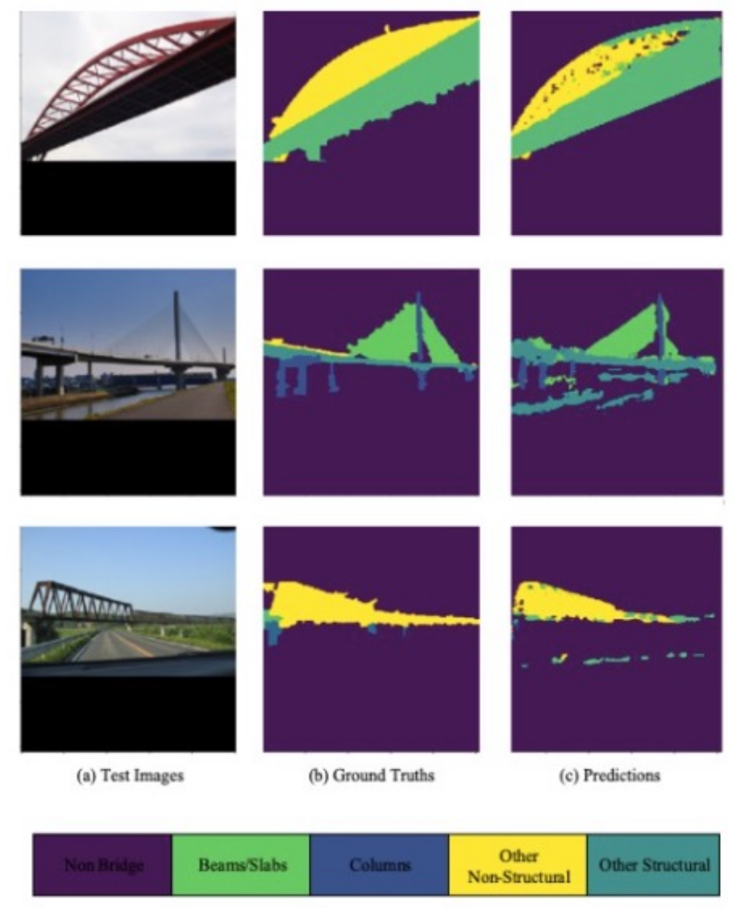

Figure 3 : Segmentation results of our proposed DNNAM model. Our proposed DNNAM model yields a mean IOU of 65.94% with pixel wise accuracy of 82.85%.

Table 2: Assessment of DNNAM on Semantic Augmented Make3D dataset when compared with the baseline model.

Assessment	mIOU (%)	PA (%)
Make3D-S (Liu et al., 2010)		
Baseline ResNet-50	65.83	88.42
DNNAM	**73.42**	**87.47**

- These attention maps assist the network to focus on these regions automatically by emphasising higher weightage on the structural component regions that help to extract robust discriminatory features for their classification.

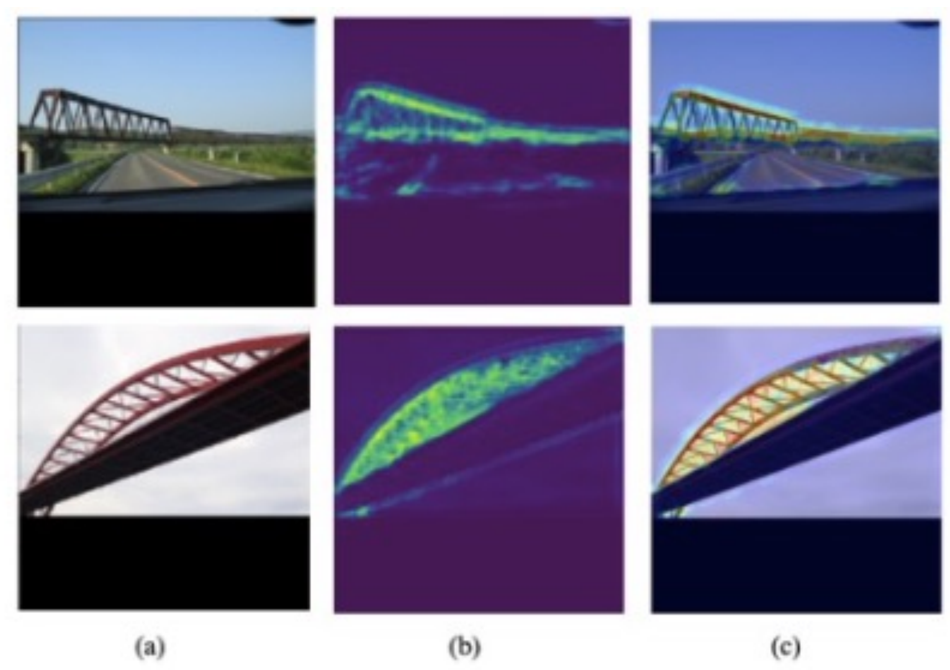

Figure 4 : Attention maps obtained from the proposed DNNAM network for sample images from the Bridge component classification dataset. (a) Original images are followed by their respective (b) attention maps and (c) heatmaps are placed side-by-side.

Conclusion

- Experimental results and ablation studies on benchmarking bridge components datasets show that our proposed architecture outperforms current state-of-the-art methodologies for structural component recognition.

Reference

[1] Narazaki, Y., Hoskere, V., Hoang, T. A., and Spencer, B. F. (2017). Vision-based automated bridge component recognition integrated with high-level scene under- standing.

[2] Kaothalkar, A., Mandal, B., and Puhan, N. B. (2022). Structurenet: Deep context attention learning for structural component recognition.