



Pixel-level classification for bridge deck rebar detection and localization using multi-stage deep encoder-decoder network

Habib Ahmed, Chuong Phuoc Le, Hung Manh La^{*}

Advanced Robotics and Automation Lab, Department of Computer Science and Engineering, University of Nevada Reno, USA

ARTICLE INFO

Keywords:

Structural health monitoring (SHM)
Non-destructive evaluation
Ground penetrating radar (GPR)
Rebar detection and localization
Deep convolutional neural networks
Deep encoder-decoder networks
Pixel-Level classification

ABSTRACT

Structural Health Monitoring (SHM) and Nondestructive Evaluation (NDE) of civil infrastructure have been active areas of research for the past few decades. The traditional inspection methods for civil infrastructure, mostly relying on the visual inspection are time-consuming, labor-intensive, error-prone, and often provide subjective results. In the wake of rising costs for infrastructural maintenance, the time factor, safety issues, and the error-prone nature of human inspection methods, there is an increased need for the development of automated methods for bridge inspection and maintenance. The purpose of this research is to provide a novel Deep Learning-based approach for rebar detection and localization within bridge decks. The proposed system is trained using Ground Penetrating Radar (GPR) data from 8 real bridges in the United States. The results have been discussed in terms of qualitative and quantitative aspects with considerable potential and various issues that need to be explored in future works. Due to the similarity in the type of parabolic signatures present in other GPR-related applications, this technique can be generalizable to other applications. The proposed approach for rebar detection and localization has considerable implications for the civil experts in general and GPR community in particular.

1. Introduction

With each passing year, the importance of utilizing timely and cost-effective methods for non-destructive evaluation (NDE) and inspection of infrastructure is emphasized in the wake of catastrophic incidents related to the destruction of civil infrastructures (Ahmed et al., 2019a, 2019b, 2020a, 2020b, 2021, 2022; Kaur et al., 2016; Ahmed and La, 2021; La et al., 2013a, 2013b, 2014a, 2014b, 2015, 2017, 2019). Out of the different infrastructure-related incidents, throughout the past years, the destruction of bridges has been discussed in some of the recent studies (Penn, 2018; Kirk and Mallett, 2022; Wright, 2012; Briaud et al., 2019). Negligence and lack of timely evaluation have led to recent bridge-related disasters in different parts of the world, which have resulted in considerable loss of lives, and destruction of civil infrastructure and property. Fig. 1 shows the destruction of the arch bridge in Taiwan in 2019, which led to the fall of the bridge deck within the body of water onto fishing boats passing from underneath the bridge (News, 2019). Although this incident was following in the wake of a typhoon, the actual cause of bridge destruction could not be fully unearthed (News, 2019). Many recent bridge-related accidents are highlighting

and stressing the importance of cost-effective, time-efficient, routine, and automated methods for bridge inspection and monitoring.

Existing studies related to rebar detection and localization using Ground Penetrating Radar (GPR) data have several limitations, as outlined in a recent study (Ahmed et al., 2020a). The existing studies have utilized a block-based approach for learning and classification between rebar and non-rebar regions within the larger B-scan images. The block-based approach examines portions of images for the presence or absence of rebar hyperbolic signatures. At the same time, variations in the intensity of hyperbolic signatures, presence of noise artefacts, and reflective signals cause challenges towards effective rebar detection and localization within existing block-based approaches (Ahmed et al., 2020a). Therefore, the existing block-based methods cannot be used to provide reliable performance in real-time, practical robotic applications for NDE of bridges. It is, for this reason, the development of a rebar detection and localization approach using pixel-level classification has been proposed in this research. Consequently, leveraging the Deep Encoder-Decoder framework will allow effective pixel-level rebar and non-rebar classification. This study will examine the superior performance of state-of-the-art pixel-level Deep Encoder-Decoder Networks.

^{*} Corresponding author.

E-mail address: hla@unr.edu (H.M. La).

<https://doi.org/10.1016/j.dibe.2023.100132>

Received 23 November 2022; Received in revised form 19 February 2023; Accepted 19 February 2023

Available online 24 February 2023

2666-1659/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

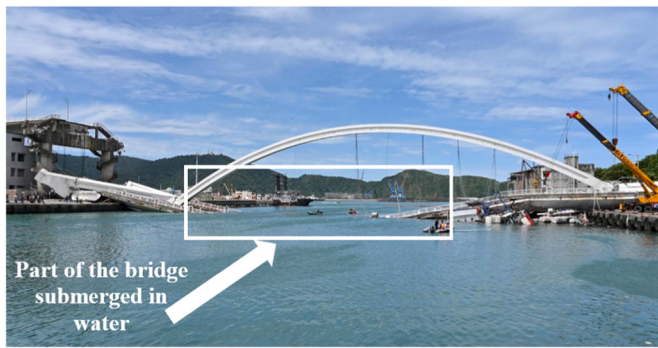


Fig. 1. Bridge destruction causing the arch bridge's deck to fall and submerge in the body of water in Taiwan (News, 2019).

This paper has been divided into five sections. This particular section discussed the motivation towards furthering the existing state-of-the-art for bridge deck evaluation and maintenance. Section 2 has been dedicated to the discussion related to existing research conducted in the field of civil infrastructure evaluation in relation to rebar detection and localization. Section 3 highlights the proposed methodology for the development of a novel Multi-stage Deep Encoder-Decoder-based system for rebar detection and localization. Section 4 will provide results and evaluate the performance of the proposed system for rebar detection and localization using qualitative and quantitative metrics. In section 5, which is the final section, an overall conclusion and recommendations for future research will be provided.

2. Related works

In this section, the primary focus will be on exploring the different learning-based approaches for rebar detection and localization with particular emphasis towards Deep Learning-based approaches (See Ahmed et al., 2020b) for an exhaustive evaluation of literature starting with relevant earliest research in the 1960s up till the present time). A number of different machine learning-based methods have been leveraged for rebar detection and localization in the past (e.g. Support Vector Machine (SVM) (Kaur et al., 2016), Naïve Bayes classifier (Gibb and La, 2016), Radon Transform (Wang et al., 2020), Hough Transform (Capineri et al., 1998; Windsor et al., 2005, 2014). It is only in the recent few years that focus has shifted towards leveraging Deep Learning-based methods for rebar detection and localization (Dinh et al., 2019; Ahmed et al., 2019a, 2019b; Besaw and Stimac, 2015). Study by Dinh et al. (2019) proposed the usage of 24-layer deep CNN model for rebar classification. The different steps for rebar detection and localization include time-zero correction, migration, filtering and thresholding, estimation of weighted centroid, and image classification using CNN network (Dinh et al., 2019). The use of Residual Neural Networks (ResNet-50) has also been proposed in recent studies related to rebar detection and localization (Ahmed et al., 2019a, 2019b). The preliminary examination of results using GPR data from real-world bridges has shown that ResNet models with varying network depths (e.g. ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152) provide increased accuracy and generalizability, which can allow rebar detection systems to accurately classify data from new bridges (Ahmed et al., 2019a, 2019b, Ahmed et al., 2020a). A number of challenges have also been discussed in sufficient detail, which prevent accurate and reliable detection and localization of rebar signatures (Ahmed et al., 2020a). Another study implemented the multi-objective genetic algorithm for classification of rebar images (Harkat et al., 2016). The use of improved Mask R-CNN-based method with distance-guided Intersection-over-Union (DGIoU) was proposed in (Hou et al., 2021).

The research related to rebars and their detection has spread from being solely rebar detection and localization within bridges to other civil

structures containing steel rebars. For example, research by Li et al. (2021) leveraged the You Only Look Once-version 3 (YOLO-v3) detector algorithm for counting rebar on construction sites. The performance of the rebar counting algorithm was verified and compared with some of the state-of-the-art Deep Learning models (e.g. Faster R-CNN, Mask R-CNN, and different variants of YOLO) (Li et al., 2021). A semi-automatic genetic algorithm-based method for rebar detection and localization was developed with data obtained from different tunnel structures. Another recent study explored the efficacy of utilizing point cloud data for assessing rebar diameter on construction sites during the manufacturing and construction stages of infrastructure development (Kim et al., 2021). As, discussed in (Ahmed et al., 2020a), the issue of signal interference has been further explored by (Xiang et al., 2021) using cascaded frequency filters to recognize rebar signatures on a collection of synthetic and on-site data from building structures (e.g. shear wall, columns, slabs).

Another study by Liu et al. (2020) made use of a Single Shot Multi-box Detector (SSD) with a 13-layered VGG-16 backbone for the detection and localization of rebar signatures in concrete slabs and walls. Several different learning-based algorithms (e.g. Naïve Bayes, Nearest Neighbors, Classification trees, Support Vector Machine (SVM)) were used and their performance was compared for accurately detecting the diameter of different-sized rebars (Kim and Lee, 2018). The research work for rebar detection and localization has been extended using a mixture of synthetic and real data from highways in (Lei et al., 2019). In this study, Faster R-CNN was deployed for rebar detection with Double Clustering Seeking Estimate Algorithm (DCSE), and Column-based Traverse Filter Points (CTFP) were used for rebar localization (Lei et al., 2019). In the proceeding discussion, the emphasis will be focused on different Deep Encoder-Decoder Networks developed and utilized in the prior studies, and the novel proposed model for rebar detection and localization developed in this research.

3. Methodology

In this section, some of the different aspects of the proposed system methodology will be discussed. In the first sub-section, some of the existing studies utilizing Deep Encoder-Decoder-based networks will be discussed to highlight the background of the proposed method leveraging multi-stage Deep Encoder-Decoder network. In the second sub-section, the focus will be towards data collection for creating the dataset leveraged in this study for pixel-level rebar segmentation using the proposed Deep Encoder-Decoder framework. In the third sub-section, a brief discussion regarding image pre-processing functions and data annotation process will be provided. In the fourth and final sub-section, the details of the proposed Deep Encoder-Decoder Network will be outlined.

3.1. Background on Deep Encoder-Decoder Networks

The usage of Deep Encoder-Decoder Networks has gained increased importance within diverse fields in the past few years. Out of the different Deep Encoder-Decoder frameworks proposed in the recent studies, some of the most popular models include, but are not limited to, SegNet, PSPNet, DeepCrack, UNet, and DeepLab (Badrinarayanan et al., 2017; Liu et al., 2019, 2020; Ronneberger et al., 2015). The different frameworks for Deep Encoder-Decoder Networks range from fully-supervised, semi-supervised to completely unsupervised. Some of the works have introduced the usage of Deep Encoder-Decoder networks in the field of image processing, background subtraction, and semantic segmentation to image compression, image-to-image transfer, video deblurring, and image captioning (Chen et al., 2019; Chen et al., 2020; Zhou et al., 2019a; Zhou et al., 2018; SpoorthiG. and Gorthi, 2019; Nakazawa and Kulkarni, 2019; Ma et al., 2020; Yuan et al., 1061). Table 1 provides information regarding the architectural details for the different Deep Encoder-Decoder Networks proposed for the diverse

Table 1

Different Deep Encoder-Decoder models used in recent studies, along with their various properties.

Study	Application	Encoder	Decoder	Data	Performance
SegNet	Image Segmentation	VGG16	VGG16	SUN RGBD	mIOU = 60.1%
ROI	ROI Compression	FMTN	IMTN	MSRA-B	F1 = 83.00%
DHEDN	Image Caption	CNN/LSTM	SF-LSTM	MS-COCO	L1 = 71.3%
FCESNet	Background Subtraction	Conv LSTM	DeConv LSTM	CDnet	F1 = 90.4%
W-SegNet	Image Segmentation	Conv/Decomp.	DeConv.	VOC2012	IoU = 39.1%
Br-GAN	Image-Image Transfer	Conv + RB	RB + DeConv	CityScape	IoU = 22.3%
CDEDnet	Medical Segmentation	DeepLabV3	DeepLabV3	CHAOS	Acc. = 99.46%
Dense-Xnet	Medical Segmentation	FC-DenseNet	FC-DenseNet	PET	F1 = 72.63%
CPCE	Image Denoising	Conv.	DeConv.	MGH	PSNR = 30.14
C-DeepIED	Medical Segmentation	Conv/LSTM	DeConv/LSTM	CBV imgs.	Acc. = 99.10%
CE-net	Medical Segmentation	DAC	RMP	DRIVE	Acc. = 94.5%
RIED-net	Image Segmentation	Residual Inception	Residual Inception	Mayo	SSIM = 0.962
Res-Dyad	Remote Sensing	Conv. Residual	DeConv. Residual	LandUse	Prec. = 99.4%
MSFg-Net	Object Detection	FGNet	FGNet	CDnet-2014	F1 = 84.3%
DAB-Net	Image Segmentation	DAB	PAD	CityScape	mIoU = 66.4%
MyRF-Net	Background Subtraction	Conv.	DeConv.	CDnet-2014	F1 = 0.9514

applications.

The application of the Deep Encoder-Decoder network has broadened to a wide range of different research areas. For pixel-level crack detection, a number of the studies have developed different Deep Encoder-Decoder frameworks, namely DeepCrack (Liu et al., 2020) and optimized Encoder-Decoder framework utilizing switch module, UNet and DeepCrack in (Liu et al., 2019, 2020). Another research area that has received attention in the application of Deep Encoder-Decoder networks is segmentation within images related to remote sensing and aerial imagery (Salem et al., 2019; Chen et al., 2020; Pan et al., 2019). In the outdoor urban environment, several applications have been developed using variants of the Deep Encoder-Decoder Networks. Scene text detection and verification for Chinese textual understanding has been discussed in one of the studies (Gao et al., 2020). Another method for road scene understanding using a Deep Encoder-Decoder network with VGGNet as backbone was developed in (Zhou et al., 2019b). The development of MSFgNet has been proposed for the effective moving object detection in video-based data from outdoor environments (Cook et al., 2015). For semantic segmentation of video data from urban environments, the development of Depth-wise Asymmetric Bottleneck (DAB-Net) was discussed in one of the recent works (Neumann, 2017). A video-based foreground extraction method for traffic and surveillance applications was developed using a novel MvRF-CNN Deep Encoder-Decoder network (Lee et al., 2017).

Within the field of medical imaging, different variants of the Deep Encoder-Decoder Networks have been proposed, namely CDED-net with Deeplab V3+ as the backbone for boundary segmentation in medical imaging (He et al., 2016), Sclera-Net for Sclera segmentation with Residual Encoder-Decoder networks (Naqvi and Loh, 2019), Multiple Sclerosis segmentation from MRI images using U-Net-based Encoder-Decoder network (Salem et al., 2019), low-dose CT image segmentation using Residual Encoder-Decoder-based CNN (RED-CNN) (Chen and Cohn, 2010), segmentation and classification of Coronary Microvascular Disease using Deep fully-convolutional Encoder-Decoder framework (Pan et al., 2019), Lymphoma segmentation from full-body PET/CT scan images using DenseX-Net (Chen and Cohn, 2010), Deep Residual Inception-based Encoder-Decoder Network for medical imaging analysis using different imaging databases (Kim et al., 2021), and Melanoma detection using Deep Encoder-Decoder framework (Adegun and Viriri, 2019). Some of the other fields in which existing and novel Deep Encoder-Decoder networks have been proposed include 2D phase unwrapping in power signals (Lee et al., 2017), defect detection in semiconductor manufacturing (Flint et al., 1061), desert seismic noise suppression (Gucunski et al., 2015a) and smoke density estimation (Yuan et al., 1061), to name a few applications.

3.2. Dataset

There is a dearth of publicly available GPR data that can be used in order to develop or validate systems for bridge inspection in particular and other applications in general. Consequently, the data from this study has been collected by one of the authors of this research (La et al., 2015, 2017). The GPR data has been acquired from several different actual bridges in the United States using Robot-Assisted Bridge Inspection Tools (RABIT), which is shown in Fig. 2. A significant part of the GPR data used in this research is one small segment of the overall GPR data collected from the inspection and evaluation performed on 40 different bridges in the United States between the period 2013 and 2014 (Ahmed et al., 2020a, 2020b; La et al., 2017). It can be seen from Table 2 that the bridge data has been taken from different types of bridges (e.g., suspension, beam, truss, girder). A portion of the GPR data has also been used in previous studies (Ahmed et al., 2019a, 2019b, 2020a). The physical dimensions vary considerably, ranging from the largest bridge in the dataset (i.e., Broadway Bridge, AR) spanning to around 2786 ft. The smallest in terms of length that has been used in this study belongs to the Dove Creek Bridge, BC, which spans for around 50 ft. Table 2 outlines the crucial properties of the different bridges in terms of the bridge name, geographical location, and physical properties of the different bridges. Table 2 also highlights the number of images acquired from the different bridges.

Despite the usage of data in prior studies, there are marked

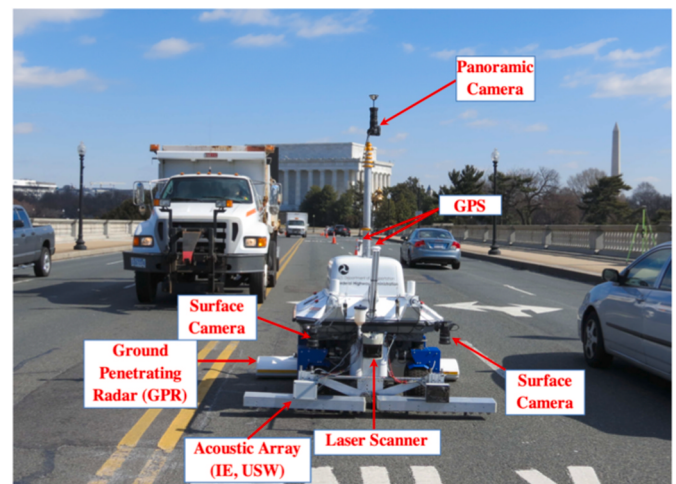


Fig. 2. Robotics Assisted Bridge Inspection Tool (RABIT) (La et al., 2015, 2017; Gucunski et al., 2015b) is the robotic platform that was used for data collection in this study.

Table 2

Data from different real bridges that have been used in this studies. Data has been collected from the varying states of the United States using the RABIT platform (La et al., 2015, 2017; Gucunski et al., 2015b), and using GPR carts provided by the GSSI company.

Bridge Location	Bridge Type	Bridge Dimensions (ft.)	Number of Images
1. Galena Creek, NV	Twin Span Arch	1,726.5 × 62.0	135
2. East Helena, MT	Concrete Tee-Beam	66.9 × 40.0	185
3. Kendall Pond, NH	Girder	78.1 × 44.0	70
4. Piscataqua, ME	Through-Arch	4,503 × 98	40
5. Broadway, AR	Arch	2,786 × 40	160
6. Fordway, NH	Beam	131 × 23	195
7. Dove Creek Rd., MO	BC Beam	50 × 45	150
8. Baxterville Bridge, CO	Lost-through Truss	117 × 15.4	110
Total			1,055

differences in how the data has been leveraged in this study compared with prior studies (Kaur et al., 2016; Ahmed et al., 2019a, 2019b, 2020a). With respect to annotation, in contrast to prior studies (using block-based annotation), this study utilized pixel-level annotation for training and validation of the proposed model. With respect to the image sizes, in contrast to prior studies (image sizes of 50 × 50 pixels, 100 × 100 pixels and 250 × 250 pixels), the image size used for training and validation of the proposed model in this study is 768 × 768 pixels. The detailed specifications of the computer system used for the training and validation of the proposed rebar detection and localization system are given as follows: Ubuntu 18.04 LTS, 32 GB memory, 350 GB hard disk, Intel® Core i7-8700 CPU with 3.2 GHz clock speed and NVIDIA® GeForce® GTX 1080 TI Graphical Processing Unit (GPU). For the purpose of training and validation of the proposed system for rebar detection and localization, Tensorflow, PyTorch and Keras libraries have been used within Python programming language framework. With regards to the network parameters for training of the proposed models, the number of epochs is set to 100, value of batch size is specified to 8 and learning rate is set to 0.001. The data is divided between training and validation sets based on “leave-one-out” approach, such that out of the total data from eight bridges, training of the proposed framework is conducted on seven bridges and validation is performed on data from one bridge (this same process is repeated for all bridges, i.e. training/validation cycle is performed eight times and average results for all metrics are highlighted). This process is used to perform validation on all of the bridges to assess the performance of the proposed system for rebar detection and localization. The use of this approach allows the researchers to assess the ability of the proposed system to provide reliable performance on unseen data.

3.3. Image pre-processing functions

A number of different pre-processing functions have been used in this study, which will be discussed in this sub-section. The original data is converted into image format (e.g. JPEG, PNG), which can be used for the training and validation of Deep Learning-based frameworks. The different pre-processing functions that have been used require manual operations by the researcher. Some of these functions include cropping, resizing, and modification of image brightness, contrast, and color balance to ensure that the rebar signatures can be effectively highlighted within the diverse bridge data in a uniform manner irrespective of the bridge data being analyzed.

The type of data annotation used in this study is pixel-based annotation (individual pixels are classified as either belonging to rebar signature or background). The prior studies leveraged block-based annotation (regions of the image are taken and separately classified as either rebar if rebar signatures are present or background if no rebar

signatures are present) of the rebar dataset. One of the benefits of using pixel-level annotation is that the output is able to classify each pixel of the input image as either belonging to one class or another. For the case of block-based annotations, there are always ambiguous regions containing sections of rebar and background simultaneously, which cannot be accurately classified as belonging to one class or the other, leading to reduced performance of learning-based systems. However, since, annotation is a manual process; the overall time and effort required for performing pixel-level annotation is more when compared to block-based annotation.

The data annotation process is carried out separately for the two stages of the proposed model, which can allow both the stages to fulfill their respective roles towards detecting the rebar layer and rebar signatures respectively. For annotating GPR images for the training and validation of the rebar layer in stage 1 of the proposed model, the annotators manually label the rebar layer pixels as 1 and other pixels as 0 so that stage 1 is able to separate rebar layer from other parts of the image. For annotating GPR images for the training and validation of the rebar signature in stage 2 of the proposed model, the annotators manually label the rebar layer pixels as 1 and other pixels as 0, so that the stage 2 is able to separate rebar signature pixels from other parts of the image.

3.4. Proposed multi-stage deep encoder-decoder framework for rebar detection and localization

The proposed model of the Deep Encoder-Decoder Network has been inspired by studies related to Deep Encoder-Decoder Networks developed in the recent past (some of which have been covered in the prior section). Fig. 3 outlines the model for the two-stage Deep Encoder-Decoder Network proposed in this study. In the next discussion, the terms ‘block’ and ‘modules’ will be used interchangeably to refer to the essential building elements given in Fig. 3. The proposed network has two main parts, namely the Rebar Layer Identification Framework (RLIF) is Stage 1 and the Rebar Signature Localization Framework (RSLF) is Stage 2. The input data is pre-processed using a number of different functions to ensure that the raw images are cropped, resized and adjusted to ensure optimal performance of the proposed system. These two stages have been explicitly defined, as they highlight the novelty of the proposed approach in comparison with recent studies (Kaur et al., 2016; Gibb and La, 2016; Ahmed et al., 2019a, 2019b, 2020b). The details regarding each part of the proposed network will be discussed in sufficient detail in the proceeding sub-section. The output from both the stage 1 and 2 are concatenated together using a pixel-wise AND operation, which ensures that the different reflective signals and noise artefacts can be reduced to get the final output result.

In order to effectively highlight the motivation and rationale for this approach, it is essential to shed light on the different challenges and limitations highlighted towards development of effective rebar detection and localization systems as outlined in (Ahmed et al., 2020a). Out of the different issues discussed in (Ahmed et al., 2020a), one of the major issue relates to the presence of parabolic reflective anomalies in parts of the GPR image that are visually similar to the actual rebar signatures. These reflective signals are visible in the lower parts of the Fig. 4 (a). The frequency of occurrence of these reflection decreases the overall performance of the rebar detection and localization system. In order to mitigate these reflective signals from increasing the false positive rate for the proposed rebar detection and localization system, a two-stage approach has been proposed, which separates rebar layer and rebar signature using two separate stages of the proposed framework. It is important to understand that these reflective signals are present below the actual rebar signatures. This is an important insight that can allow us to leverage the rebar layer (this is the layer in which the actual rebar signatures are present) for developing an effective rebar detection and localization system. This approach has not been utilized in earlier studies, which makes the findings relevant and useful in terms of the

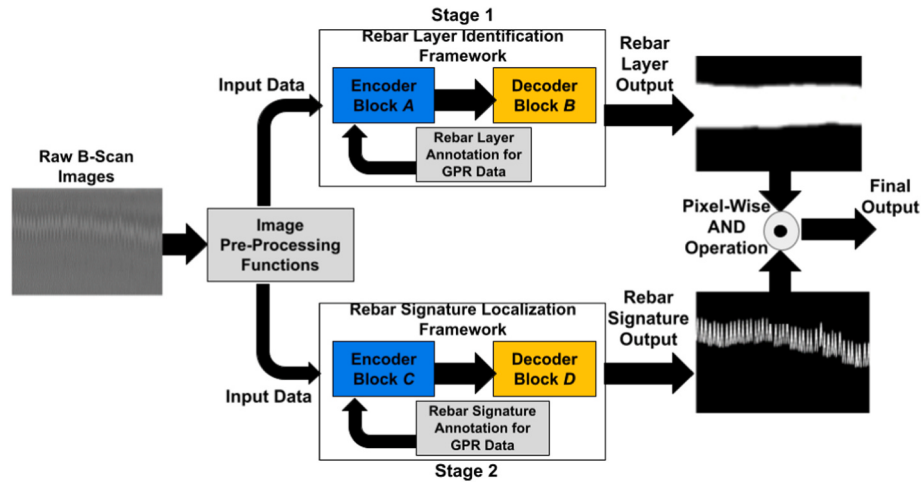


Fig. 3. Overview of the Proposed Model of the Novel Multi-Stage Deep Encoder-Decoder Network for Rebar Detection and Localization. The input image of the B-scan after undergoing pre-processing operations is put through the First and Second Stages of the Deep Encoder Decoder Network.

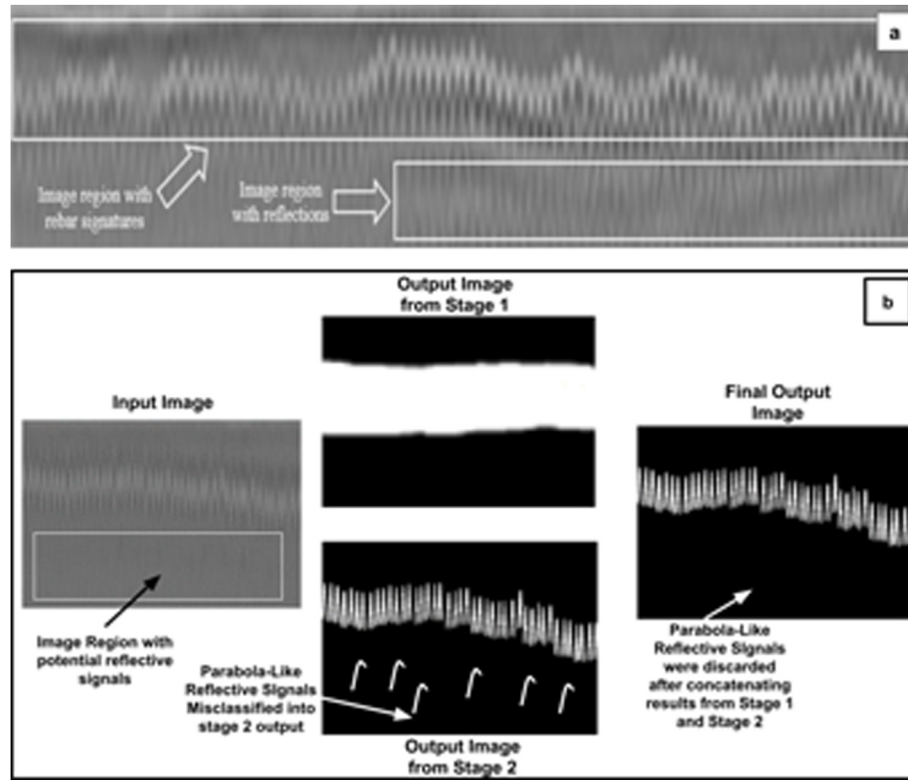


Fig. 4. Rationale for developing a two-stage approach with rebar layer and rebar signature separation and concatenation in the final stages of the proposed model with (a) the presence of rebar reflective signals in the GPR B-scan images (Ahmed et al., 2020a) and (b) a look at the different inputs and outputs of the different stages in terms of how the presence of rebar reflective signals can be minimized in the final output by leveraging a two-stage approach.

utilization of rebar layer as an important feature for rebar detection and localization. Fig. 4 (b) highlights the different inputs and outputs of the different stages of the proposed framework in terms of how the presence of rebar reflective signals can be minimized in the final output by leveraging a two-stage-based approach. In the first stage, the rebar layer is separated; this part contains all of the rebar signatures. The second stage separates all rebar parabola signatures from the pre-processing images; but the potential presence of rebar reflective signals can lead to false positive rates at this stage of the framework. In order to minimize the false positive rates and eliminate the reflective signals from stage 2 output, the outputs of stage 1 and 2 are concatenated together using

pixel-wise AND operation. Using this operation, the final output is able to retain the parabola signals present within the rebar layer while eliminating the reflective signals that exist below the rebar layer. In this manner, based on findings from (Ahmed et al., 2020a), this approach is able to improve on one of the various issues affecting the performance of prior rebar detection and localization methods.

3.4.1. Rebar Layer Identification Framework

As the name suggests, the Rebar Layer Identification Framework (RLIF) primarily deals with the visual differentiation between pixels belonging to image regions containing the rebar signatures and

background pixels. These regions or collections of pixels can also be termed the ‘rebar layer.’ The Encoder Block A and Decoder Block B constitute the RLIF. Similarly, the Encoder Block C and Decoder D constitute the RSLF, which will be discussed in the following subsection. Although images can contain reflection signals similar to parabolic rebar signatures, most B-scan images contain a single rebar layer to specify the underground depth at which the rebar is visually present within the B-scan images. For this reason, the RLIF focuses on highlighting a single layer, where most of the rebar parabolic signatures are present. This will be possible by providing a pixel-level annotation of the pre-processed GPR image data into either rebar layer or background pixels, making it a binary classification problem. The architecture of the proposed RLIF will be similar to SegNet, which is shown in Fig. 5, such that the Encoder-Decoder model used in RLIF will consist of a single Encoder and Decoder modules. Several different architectures (e.g., UNet (Ronneberger et al., 2015) and PSPNet (He et al., 2015), and MobileNet (Howard et al., 2017) encoder used as part of the overall SegNet architecture) will be tested and evaluated for the construction of RLIF.

3.4.2. Rebar Signature Localization Framework

The Encoder Block C and the Decoder Block D, which are given in Fig. 3, constitute the Rebar Signature Localization framework (RSLF). The architecture of the proposed RSLF is also similar to SegNet, which is shown in Fig. 5, such that the Encoder-Decoder model used in RSLF will consist of a single Encoder and Decoder modules. However, the type of Encoder-Decoder architectures used in RSLF and RLIF are different in terms of their internal network-level characteristics. At the same time, the type of pixel-level annotation performed for the two framework is also different. As, the first stage of the proposed framework (i.e. RLIF) seeks to highlight the rebar layer and the second stage (i.e. RSLF) attempts to identify the individual rebar parabolic signatures, preferably present within the rebar layer. It can be seen in Fig. 5 that the initial stages of the SegNet architecture constitute the Encoder block and the final stages belong to the Decoder block. Two main types of blocks have been used in this study, including the Deep Encoder block with different network layers connected and pooling layers. The Deep Decoder block is similar in construction to the Deep Encoder block. The only difference is that instead of the pooling layers, the up-sampling layers are used to ensure that the output from the encoder can be resized to the actual image size provided at the input of the proposed model.

4. Results and discussion

This section will discuss the salient features of the proposed network and its performance, along with their various implications for future research. The proposed method has two main parts, namely the Rebar Layer Identification framework (RLIF) and Rebar Signature Localization Framework (RSLF). For the first stage of the proposed network (i.e., RLIF), three major Deep Encoder-Decoder Networks have been used, namely the UNet (Ronneberger et al., 2015), PSPNet (He et al., 2015) with two variants (e.g. PSP-50 and PSP-101), and SegNet with MobileNet encoder module (Howard et al., 2017). For the case of the second stage of the proposed network (i.e., RSLF), three main Encoder-Decoder networks (e.g., UNet (Ronneberger et al., 2015), PSPNet (He et al., 2015), and SegNet (Badrinarayanan et al., 2017)) have been used, along with variations in the Encoder modules to find the most suitable Architecture-Encoder pair in terms of different qualitative and quantitative performance metrics. Some of the different Encoder modules leveraged within the context of the different Architectures include VGG-16, VGG-19, ResNet-50, and ResNet-Xception.

It is important to understand that the quantitative and qualitative analyses of the proposed approach for rebar detection and localization will be fundamentally different in nature. It is due to the diverse nature of the analysis of the results that the researchers can be able to better appreciate and gain a deeper examination of the workings of the proposed approach. Within the quantitative analyses of results, the focus is towards utilizing the different statistical performance metrics (e.g. mIoU, Precision, Recall, Dice Loss) to assess the overall feasibility of the proposed approach for rebar detection and localization. Therefore, the quantitative analyses deal with comparing and analyzing the statistical results obtained for different metrics in response to the utilization of different base architectures and encoder modules. In the qualitative analyses of results, the focus is on the visual quality of the output images provided by the different base architectures and encoder modules used in this study. In order to accomplish that, the researchers manually compare the difference between the ground truth and output images, such that the closer the output images are to the ground truth, the better the overall qualitative assessment of the proposed method.

4.1. Quantitative analysis

In this section, the primary emphasis will be on examining the statistical evaluation of the different aspects of the proposed method for

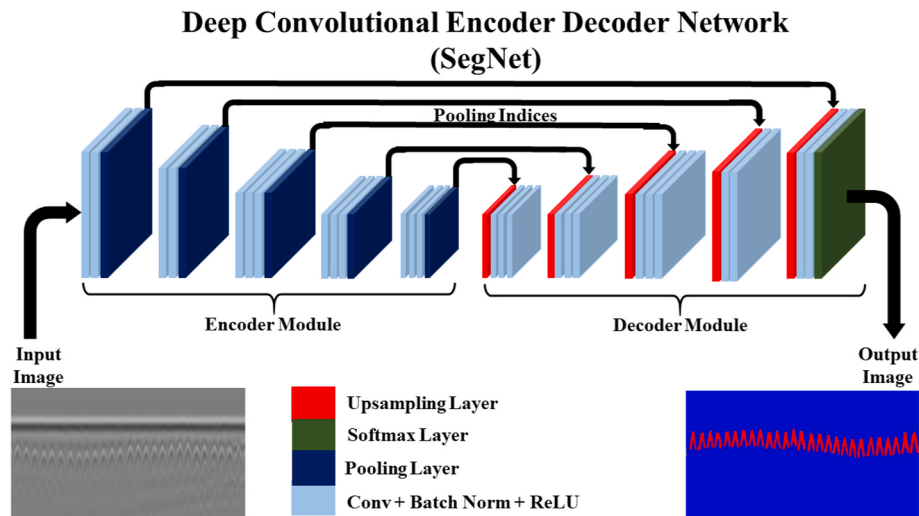


Fig. 5. Architectural Framework for SegNet (Badrinarayanan et al., 2017) with one Encoder and one Decoder module, which has gained considerable attention in the recent past. The RLIF will make use of different widely-deployed Deep Encoder Decoder Networks, which are similar in construction to the SegNet (Badrinarayanan et al., 2017).

rebar detection and localization. Tables 3 and 4 highlight the overall quantitative performance of the different Architectures and Encoder modules for the two stages of the proposed system for rebar detection and localization. Several different performance evaluation metrics have been used for assessing the performance of the Deep Encoder-Decoder networks in the different applications in the recent past (He et al., 2015; Howard et al., 2017; Ronneberger et al., 2015; He et al., 2015; Badrinarayanan et al., 2017). The different metrics used in this study include Dice Loss, mean-Intersection-over-Union (mIoU), Precision, and Recall. The lower values for Dice Loss are more suitable as they correspond to the level of loss incurred for the different combinations of network frameworks used in the proposed system for rebar detection and localization. For all the other performance metrics (e.g. mIoU, Precision, and Recall), the higher values correspond to improved performance of the proposed rebar detection and localization system.

For the case of stage 1 of the proposed framework, four different types of base networks have been used, namely UNet, PSPNet-50, PSPNet-101, and SegNet framework. Out of the different Deep Encoder-Decoder networks utilized in the first stage, the most promising results have been outlined by the SegNet with MobileNet Encoder. Many of the original Deep Encoder-Decoder architectures (e.g. UNet, PSPNet-50 and PSPNet-101) used in stage 1 utilize default encoder modules in order to limit the level of complexity and variables being used in this study. The MobileNet framework utilized as an encoder within the SegNet framework provides a lightweight Deep Encoder-Decoder network from the different models utilized in stage 1 of the proposed rebar detection and localization system. For the different encoder modules and base architectures used in the second stage of the proposed network, the highest performance has been highlighted by combining the SegNet framework with the ResNet-Xception encoder module. When comparing the performance of PSPNet with 50 and 101 layers in the first stage of the proposed system, increasing the complexity and number of layers has an overall negative effect on the performance of the rebar detection and localization system. Compared to these two frameworks, the complexity and number of layers for UNet and MobileNet are limited. However, as it can be seen in Tables 3 and 4, these two networks (e.g., MobileNet and UNet) can provide a higher level of performance with the different combinations of Architecture-Encoder pairs leveraged at the second layer of the proposed framework for rebar detection and localization system.

A different combination of base architecture and encoder modules was used for stage 2 of the proposed rebar detection and localization system. It is important to understand that PSPNet with different number of layers does not support the usage of different encoder modules. The different base architectures used in the second stage of the proposed

system include SegNet, UNet, and PSPNet. The different encoder modules utilized include VGG-16, VGG-19, ResNet-50, and ResNet-Xception. In terms of the number of layers, the different encoder modules can be ranked from the lowest to the highest number of layers as follows: VGG-16, VGG-19, ResNet-Xception, and ResNet-50. In terms of improved performance, the most crucial combination of Architecture-Encoder pairs at the second stage of the proposed framework is SegNet-ResNet-Xception (where SegNet is the base architectural framework and ResNet-Xception is the encoder module). Of the different encoder modules used, the most effective one can be classified as ResNet-Xception, which has shown improved performance when leveraged within different base architectural frameworks. ResNet-Xception encoder module with the SegNet framework at the second stage and SegNet framework with MobileNet encoder module at the first stage have the highest values for the different metrics (e.g., Dice Loss, mIoU, Precision, and Recall) are 12.20%, 93.57%, 97.43%, and 96.62%. All other values for the different frameworks at the first and second stage have comparatively lower values of mIoU, Precision, and Recall, as well as higher values for Dice Loss, as can be seen in Tables 3 and 4

The performance of the proposed system cannot be directly compared with the majority of the existing studies in the field of rebar detection and localization conducted with an emphasis on bridge inspection in particular. The primary reason for this fact is that earlier studies utilize block-based techniques, which make use of different metrics, such as accuracy and loss (Ahmed et al., 2019a, 2019b, 2020a, 2020b). These metrics cannot be used for the current study since it leverages pixel-based methods for classification, such as individual pixels are classified as either belonging to rebar or non-rebar classes. The performance evaluated using these metrics (e.g., mIoU, Dice Loss, Precision, and Recall) can be considered a more reliable and accurate reflection of the actual performance of the proposed rebar detection and localization system.

4.2. Qualitative analysis

In this sub-section, the primary emphasis will be on examining the non-statistical evaluation of the different aspects of the proposed method for rebar detection and localization. The section will be further sub-divided into two parts due to the multi-layer nature of the proposed framework for the rebar detection and localization. The first sub-section will outline the different insights and qualitative characteristics of the different output results from the first stage of the proposed framework. The second sub-section will discuss the different qualitative features of the different output results from the second stage of the proposed framework for rebar detection and localization.

Table 3

Results are shown for the dataset from eight different bridges. The results are shown for a different set of base architectures and encoder pairs for the second stage of the proposed framework. The most promising results for the different results at stage 1 and 2 of the proposed framework has been given in bold fonts.

Stage 1 Base Model	Stage 1 Encoder	Stage 2 Base Model	Stage 2 Encoder	Dice Loss (%)	mIoU (%)	Precision (%)	Recall (%)
UNet	Default	SegNet	VGG16	25.15	85.73	88.43	82.26
"	"	"	VGG19	22.19	87.24	88.55	86.72
"	"	"	ResNet-50	16.68	90.65	90.53	90.06
"	"	"	Xception	15.45	92.48	93.06	92.50
"	"	UNet	VGG16	20.14	89.50	85.29	86.63
"	"	"	VGG19	19.24	90.57	88.33	88.21
"	"	"	ResNet-50	18.91	91.74	89.66	88.37
"	"	"	Xception	19.25	89.50	89.90	88.40
"	"	PSPNet	N/A	32.17	86.63	85.75	84.28
PSPNet-50	"	SegNet	VGG16	23.53	84.58	81.77	80.08
"	"	"	VGG19	20.06	86.64	84.17	82.21
"	"	"	ResNet-50	19.21	89.83	84.45	86.33
"	"	"	Xception	18.75	90.48	91.06	90.50
"	"	UNet	VGG16	23.16	83.25	81.76	81.24
"	"	"	VGG19	21.77	85.36	83.62	80.14
"	"	"	ResNet-50	19.15	87.51	82.50	80.19
"	"	"	Xception	19.20	87.70	82.20	80.55
"	"	PSPNet	N/A	32.27	82.25	80.37	79.92

Table 4

Results are shown for the dataset from eight different bridges. The results are shown for a different set of base architectures and encoder pairs for the second stage of the proposed framework. The most promising results for the different results at stage 1 and 2 of the proposed framework has been given in bold fonts.

Stage 1 Base Model	Stage 1 Encoder	Stage 2 Base Model	Stage 2 Encoder	Dice Loss (%)	mIoU (%)	Precision (%)	Recall (%)
PSPNet-101	"	SegNet	VGG16	29.38	81.55	76.23	77.56
"	"	"	VGG19	26.15	85.40	82.61	82.57
"	"	"	ResNet-50	20.26	88.72	82.63	83.45
"	"	"	Xception	18.41	90.55	91.26	90.59
"	"	UNet	VGG16	21.88	85.65	83.32	85.22
"	"	"	VGG19	24.42	88.58	80.26	80.17
"	"	"	ResNet-50	20.69	90.75	84.13	83.38
"	"	"	Xception	21.24	86.81	82.63	81.50
"	"	PSPNet	N/A	30.16	82.69	82.15	78.86
SegNet	MobileNet	SegNet	VGG16	21.46	91.12	92.50	93.36
"	"	"	VGG19	22.18	91.50	90.07	91.23
"	"	"	ResNet-50	17.37	92.21	95.24	94.45
"	"	"	Xception	12.20	93.57	97.43	96.62
"	"	UNet	VGG16	20.15	90.11	89.83	86.58
"	"	"	VGG19	14.31	92.42	90.50	92.21
"	"	"	ResNet-50	14.11	92.09	90.55	92.60
"	"	"	Xception	15.57	91.33	89.40	88.91
"	"	PSPNet	N/A	27.25	87.55	82.94	80.88

4.2.1. Qualitative analysis: Rebar Layer Identification Framework

In this sub-section, the discussion will deal with the visual results obtained for the first stage of the proposed framework for rebar detection and localization, namely the Rebar Layer Identification Framework (RLIF). In Fig. 6, a number of different results have been highlighted from stage 1 of the proposed framework for rebar detection and localization. The top two images are actual GPR images and ground truth annotated images respectively. The annotation classified between the foreground (i.e. rebar layer) and background (i.e. anything in the image that is not part of the rebar layer). This part highlights a different perspective as compared to quantitative analyses, as there are different visual elements of the results that cannot adequately be discussed in statistical terms.

Out of the different Encoder-Decoder architectures leveraged for the development of the RLIF, some of the Architectures include SegNet, UNet, smaller-version of UNet (i.e., UNet-mini), and PSPNet. The reason UNet-mini results are not highlighted in Table 2 is because they do not provide adequate performance in terms of the different statistical measures (e.g. mIoU, Dice Loss, Precision and Recall). The qualitative results for UNet-mini are shown in Fig. 6 in order to gain a better understanding of the reason for reduced performance. For the case of SegNet architecture in stage 1, MobileNet-v2 was used as the encoder module in order to attempt to reduce the overall size of the two stage framework. The individual encoder-decoder architectures (e.g. SegNet, PSPNet, UNet) are deep networks with considerable complexities and computational overheads.

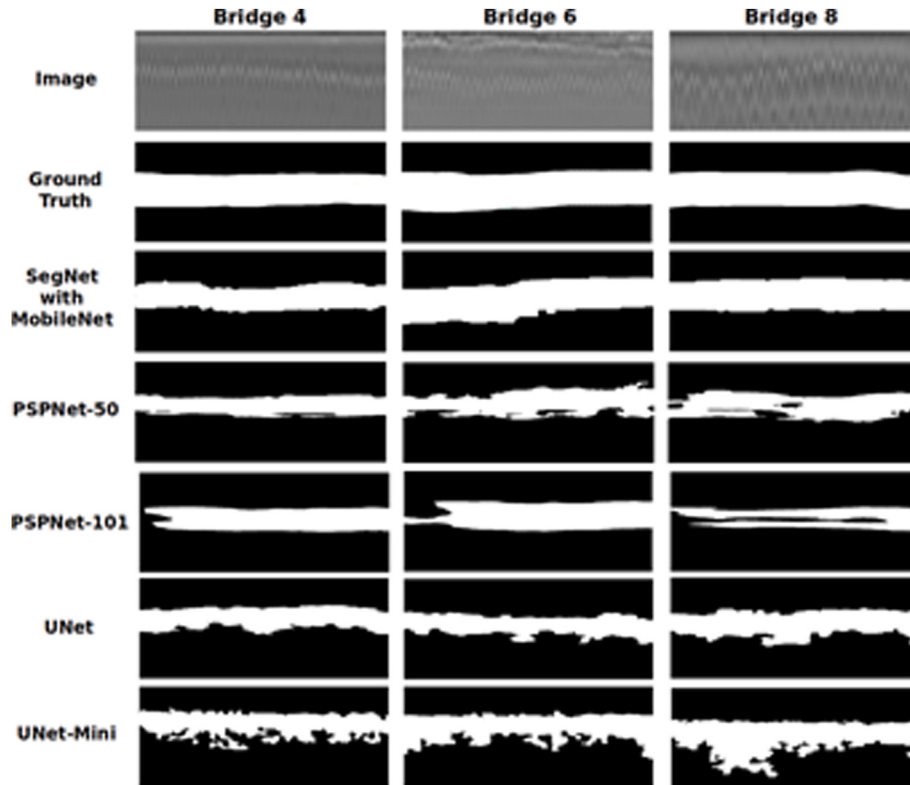


Fig. 6. Results shown for the dataset from three different bridges. The results are shown for different set of base architectures and encoder pairs used in the first stage of the proposed framework.

In order to reduce these overheads, the goal was to introduce a smaller version of Encoder-Decoder network at the first stage for rebar layer identification with the second stage for the rebar signature localization. Another benefit of using a smaller framework for the first stage was that the original version of MobileNet-v2 could not process the image data with dimensions $768 \times 768 \times 3$. In order to work with SegNet architecture with MobileNet-v2 encoder module at the first stage of the framework, the data has to be resized to $256 \times 256 \times 3$, which improves the computational cost of using multiple stages of Deep Encoder-Decoder networks. At the same time, it also increases the performance of the first stage framework leveraging MobileNet-v2 to compare with other Deeper frameworks, (e.g., SegNet, PSPNet and UNet). After passing through the first stage, the results obtained are re-sized back to their original size, so that the results from different Architectures in the first stages can be effectively compared, as shown in Fig. 6. It can be seen in Fig. 6 that SegNet framework with MobileNet-v2 Encoder module is able to provide the most promising results at the stage 1 of the proposed framework. For data from all three bridges, it can be seen that the results are closer to the ground truth in comparison with other frameworks developed for stage 1. Furthermore, the results from SegNet-MobileNet shows the effects of resizing on the output in the form of block-based effects visible at the lower and upper edges of the layer results that were magnified after resizing the images from $256 \times 256 \times 3$ to $768 \times 768 \times 3$ (i.e. the noise and other artefacts that were smaller in the original result images were magnified many times after the images were resized).

For the case of PSPNet-50 and PSPNet-101, the results are more smooth in terms of visual texture. However, there are some issues in terms of patches missing from the different data results shown in Fig. 6. For example, the results from PSPNet-101 bridge 8 data shows some missing portion in the middle of the rebar layer region. Similarly, for PSPNet-101, some minor regions of the rebar layer are missing for the results given for bridge 4 and 6. For the case of PSPNet-50, the results from bridge 4, there are some slight defects at the bottom of the rebar layer region. Similarly, the results from bridge 6 show some minor issues from the top and bottom of the identified rebar layer region. For the results obtained from PSPNet-50, there are some minor missing regions from the middle of the rebar layer region, along with minor issues at the top and bottom of the identified rebar region. The results from PSPNet-101 are more smooth in comparison with PSPNet-50. This shows that the increase in number of layers of the Deep Networks might have some positive impact towards effectively highlighting the rebar layer. At the same time, there is a need to better understand how the first layer networks can better distinguish between features for the foreground (i.e. features belonging to the rebar layer) and background (i.e. features belonging to all other regions of the B-scan images) regions.

For the case of UNet, visually, the output results are much less smooth in comparison to results from the two PSPNet frameworks highlighted. For UNet, the results show some false positive regions, when the output results are compared with the ground truth. The effect of noise and other artefacts are also more pronounced for the case of UNet frameworks. For the case of UNet-mini, which is a smaller, more compact version of the original framework, it can be seen that the output results are much more sensitive towards inaccurately classifying noise and other reflective artefacts as part of the rebar layer region. This phenomenon is much more visible for the case of UNet-mini with data from bridge 8. However, since rebar profiles belong to the upper portion of the rebar layer, which is covered for the majority of the output images from UNet and UNet-mini, the regions of the rebar layer containing the rebar profile signatures are still covered within the output regions.

The overall qualitative analyses of the results from the first stage of the proposed framework for rebar detection and localization has been provided in this sub-section. The main issue highlighted is concerning the false positive regions and the addition of noise and other reflective artefacts below the rebar layer that can be incorrectly classified as actual rebar signatures in the further stages of the framework. This particular

issue is not present in results from many frameworks (e.g. SegNet with MobileNet framework, PSPNet-50, PSPNet-101 and UNet). However, one of the examined frameworks (e.g. UNet-mini) has this particular issue much more pronounced in some of the results. This particular issue will be left for future research to further examine these issues and try to work towards ensuring that the first stage of the framework for rebar detection and localization is able to provide better performance in terms of accurately highlighting the rebar layer region. Since, this particular type of exploration and approach has not been previously used in any of the relevant literature, it is difficult to ascertain the different factors that can affect the accurate detection of rebar layer region. Furthermore, there will also be a need to examine the different network-level characteristics (e.g. number of network layers, type of network layers (pooling, convolution and concatenation layers) and their combination, and network layer dimensions for each layer in the network) that can prevent the inaccurate classification of rebar layer region.

4.2.2. Qualitative analysis: Rebar Signature Localization Framework

In this sub-section, the discussion will deal with the visual results obtained for the second stage of the proposed framework for rebar detection and localization, namely the Rebar Signature Localization Framework (RSLF). Fig. 7 highlights the overall qualitative performance of the different Architectures and Encoder modules for the two stages of the proposed system for rebar detection and localization. The results highlight the qualitative aspects of rebar detection and localization. The information provided on the left-hand side of the images is based on the corresponding architecture's base architecture and encoder modules for the second stage of the rebar detection and localization system.

In Fig. 7, several promising results for the Architecture-Encoder pair have been highlighted, along with some examples of average and low performance results for the other networks leveraged in the second stage of the proposed framework for rebar detection and localization. The primary framework used for the second stage of the proposed framework include UNet (Ronneberger et al., 2015), PSPNet (He et al., 2015), and SegNet (Badrinarayanan et al., 2017). With these base networks, the different encoders were used to examine the effect of different encoders on the overall performance of the rebar detection and localization system. When comparing results for SegNet and UNet, it can be seen that the overall thickness of the rebar signatures is smaller for SegNet results. The results from the UNet framework are closer to the actual ground truth results. ResNet-50, ResNet-Xception, and Inception encoders are not shown here since the increasing number of layers in the encoder does not significantly improve the quality of rebar signatures segmented from the original B-scan images. For the case of Architecture-Encoder pairs with SegNet, the most promising results are revealed for the SegNet framework with ResNet-50 and ResNet-Xception encoder modules. For the case of the UNet framework, both VGG-16 and VGG-19 encoder results are similar in terms of qualitative aspects. However, it is interesting to note that both results are unable to accurately segment some of the rebar signatures for data from bridge 6. However, rebar signatures' overall thickness and quality are closer to the ground truth.

For the case of PSPNet (He et al., 2015), only the results for PSPNet-101 layers have been shown as a reference of relatively inaccurate results with the segmentation of rebar signatures appearing not as parabolic signatures. Instead, the rebars appear as regional blobs with pixel regions for individual rebars intersecting neighboring rebars. Although, this network (i.e. PSPNet-101) can be used for the localization of the rebars. However, the primary issue relates to the instances when the localization results for two neighboring rebars appears as a single region. This particular issue becomes problematic when the distance between two neighboring rebar signatures is reduced, as it can be seen for the output results from bridge 8 for PSPNet-101. Due to the inherent limitation of the PSPNet architecture, it cannot provide flexibility in utilizing multiple different encoder pairs, as is the case with other Deep Encoder-Decoder pairs, e.g., SegNet and UNet. The results for PSPNet with 50 layers architecture were not included, as the results were not

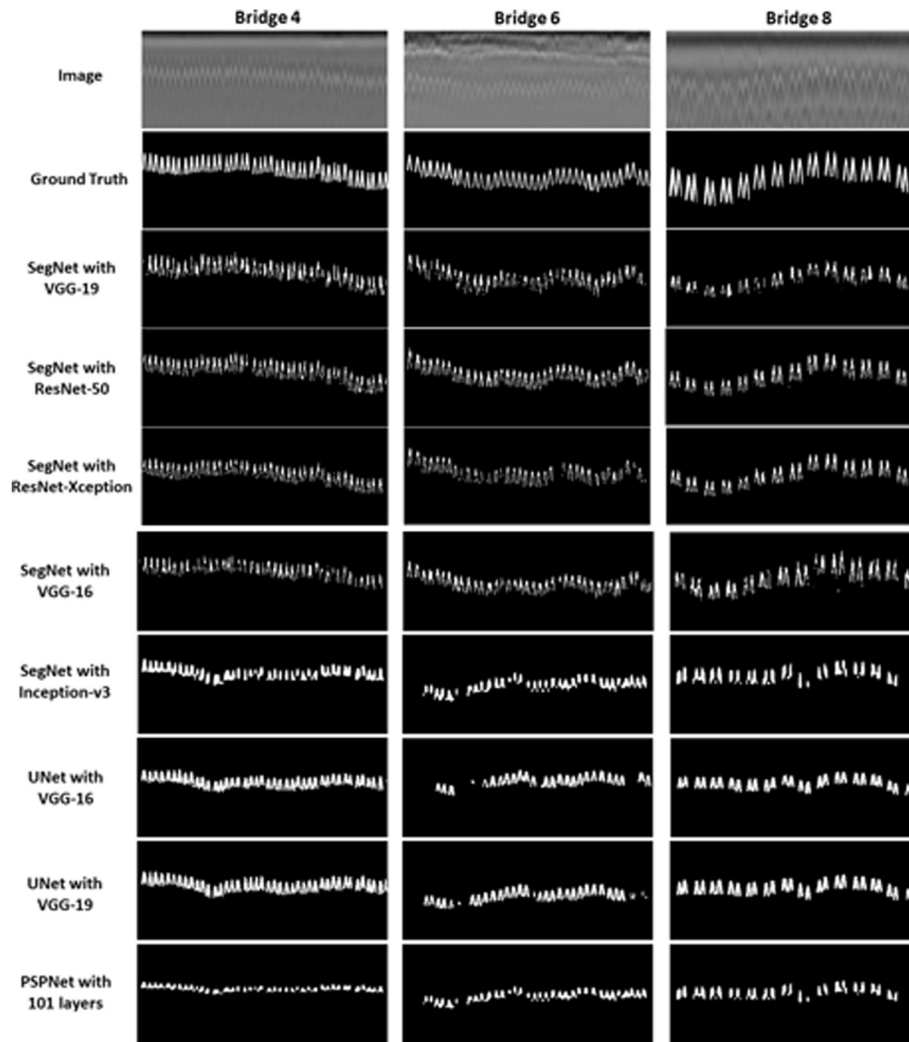


Fig. 7. Results shown for the dataset from three different bridges. The results are shown for different set of base architectures and encoder pairs for the second stage of the proposed framework. The first stage framework in this context is based on SegNet-MobileNet (Howard et al., 2017).

very accurate, and it was not accessible to qualitatively separate individual rebar parabolic signatures. Another set of sub-optimal results is given for the case of SegNet architecture with the Inception-v3 encoder module—both these results are given in Fig. 7 suffer from the same issue, such that the final results are unable to separate the results for individual rebar profiles from the neighboring rebar signatures. In conclusion, it can be seen in Fig. 7 that the most promising results in terms of qualitative aspects include MobileNet-v2 for the first stage of the framework. SegNet with ResNet-50 encoder module or UNet with VGG-16 encoder module gives the best result for the second stage of the proposed framework for rebar detection and localization.

There are a few primary issues that have been highlighted from the qualitative analyses of the results from the second stage of the proposed framework for rebar detection and localization. The first issue is related to the reduced thickness of the output rebar signatures for some output results (e.g. SegNet-ResNet-Xception, SegNet-VGG19 and SegNet-VGG16 Architecture-Encoder pairs). This can be beneficial in cases where the distance between the neighboring rebar signatures is very less. However, it can be a cause for concern, especially in the presence of noise in the images, i.e. if the intensity of noise increases in the image, it can potentially affect the ability of the proposed framework to accurately localize individual rebar signatures. This particular issue needs to be further investigated in future studies in the relevant research area. The second issue is the merger of neighboring rebar signatures for some

output result highlighted (e.g. PSPNet-101). This potential issue can seriously affect the output for bridge data with minimal distance between neighboring rebar signatures. This issue needs to be further explored in future works to fully examine the optimal network characteristics that can cater to the diverse types of data with different physical bridge characteristics (e.g. depth of rebar layer, number of rebars used in construction of the bridge, distance between neighboring rebars, type of material used within bridge deck). The third issue highlighted in this section is the disappearance of individual rebar signatures or failure to accurately classify individual rebar signature by some of the Architecture-Encoder pairs (e.g. UNet-VGG16 and UNet-VGG19) used in the second stage of the proposed framework for rebar detection and localization. Future research should try to explore ways to prevent this issue from affecting the performance of the rebar detection and localization frameworks.

4.3. Comparison between pixel-based and block-based approaches for rebar detection and localization

The pixel-based approach for rebar detection and localization is beneficial, not only for the GPR community, but also for civil experts and non-experts alike. The results discussed in prior sub-sections have demonstrated that the pixel-based approach is more precise and accurate with respect to separating rebar pixels from background pixels

within input GPR images. A comprehensive overview of the different limitations of the block-based approaches for rebar detection and localization has been discussed in (Ahmed et al., 2020a). In order to effectively highlight the results of the block-based and pixel-based approach, Fig. 8 compares the output results provided by the two approaches. The output results for the block-based approach are based on the method developed in (Ahmed et al., 2020a). For the block-based approach, the output is given in the form of a rectangular region-of-interest (ROI), which contains segments of rebar signature, as detected by the developed system. There are some presence of false positive localization results in the lower region due to presence of parabola shaped reflection artefacts in the lower region of the input image. However, further manual inspection by civil experts is still required to visually locate the rebar profile within the rectangular ROIs highlighted by the system. For the case of output generated by the pixel-based approach, the output provided is in the form of pixels highlighted that belong to rebar signature in the input image. By providing a fine-grained output by the system, the pixel-based approach makes it easier for the civil experts to assess whether the results align with the ground truth or not. However, human inspectors will still need to verify the results from both systems at the end. Based on the visual analysis of results outlined by the two approaches, it can be seen that the utilization of pixel-based approaches is beneficial for the GPR community, as utilization of these approaches will allow the GPR experts to gain more accurate, fine-grained and precise output results in comparison to block-based approaches.

4.4. Comparison between robot-based and human-based approaches for bridge inspection

Although, this is not the primary focus of the discussion in this study, it is important to highlight the different ways in which human-based inspection methods differ from automated, robot-based approaches. Previous studies have shed considerable light on the benefits of the automated, robot-based approaches for bridge inspection (Ahmed et al., 2019a, 2019b, 2020a, 2020b, 2021, 2022; Ahmed and La, 2021; La et al., 2013a, 2014a, 2014b, 2014b, 2015, 2017, 2019; La et al., 2014b; Gucunski et al., 2013). In order to effectively compare between the human-based and robot-based approaches for bridge inspection, Table 5 has been provided where different metrics are selected and the pros and cons of each approach has been outlined. Inspection time refers to the time taken for undertaking the inspection activities and data collection from an actual bridge. Detection time refers to the time required for analyzing the collected data. Maneuverability refers to the ability to access different regions of the bridge. Availability refers to the ability to

Table 5

Comparison between the human-based and robot-based inspection approaches with respect to the different metrics.

Metrics	Robot-Based Approaches	Human-Based Approaches
Inspection Time	Low	High
Detection Time	Low	High
Inspection Accuracy	High	Moderate/High
Detection Accuracy	Varies based on Analysis tools used	Varies based on Inspector skills
Initial Costs	Research and Development costs (High)	Training and Skill Acquisition costs (High)
Recurring Costs	Minimal	Hourly Charges for Civil Experts (High)
Maneuverability	Able to reach small spaces and high places (High)	Requires additional gears, equipment and experience (Low)
Risk	Minimal Risk	High Risk of Injury or Death
Replaceable	Easy to Replace	Difficulty in Replacing Skills and Knowledge
Ability to Upgrade	High Time and Cost Factors	High Time and Cost Factors
Availability	Low	Moderate
Adaptability	Low	High

inspect the bridges irrespective of changes in weather (e.g. rain, snow, wind) and time of the day (e.g. morning, evening, night). Adaptability refers to the ability to handle and cope with uncertain situations (e.g. road accidents) that can take place during inspection activities. Despite many advantages in several of the metrics for robot-based approaches, it is important to understand that most of the existing robot-based platforms for bridge inspection are in their initial phases of development and very few examples of practical implementations in the field exist at present (e.g. RABIT platform). The widespread deployment of practical fully-autonomous robots for different routine bridge inspection tasks is a research challenge that is currently being investigated. Therefore, the purpose of comparison drawn in Table 5 is not to favor one approach over another. Since, at present, the emphasis of the existing research in this field is towards developing automated tools and technologies that do not replace human experts, but assist and complement them in promoting efficiency and improved overall performance in the different inspection-related tasks.

5. Conclusion and future works

This paper provides a detailed evaluation of the rebar detection and localization system. Although, some efforts have been dedicated to the development of rebar detection and localization system in the past. Prior

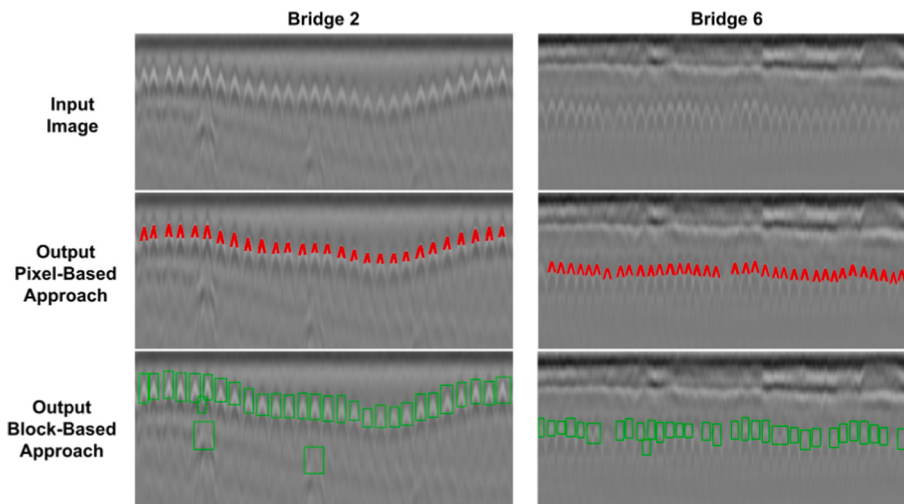


Fig. 8. Comparison between the output results for pixel-based approach used in this study with block-based approach used in (Ahmed et al., 2020a). In the output of the block-based approach, a green bounding box is used to highlight region containing rebar signatures. There is some prevalence of false positive detection (incorrect classification of rebar reflection artefacts as rebar signatures) as well as false negative (unable to correctly classify rebar signatures). There are also instances in which two rebar signatures are collectively classified as a single rebar signature. In the output for the pixel-based approach, individual pixels in the original image belonging to rebar signature class are highlighted with red color to distinguish them from the other pixels. It can be seen that the results are more fine-grained and accurate than block-based approach. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

studies in the rebar detection and localization system have utilized block-based approaches. The proposed system is different in many different ways. In this study, pixel-based classification has been introduced with a multi-stage framework incorporating different Deep Encoder-Decoder networks. Different types of Deep Encoder-Decoder networks with different levels of network complexities and the number of layers have been used for the two stages of the proposed rebar detection and localization system. The most optimal results have been demonstrated using the SegNet-MobileNet Architecture-Encoder pair at the first stage and the SegNet framework with the ResNet-Xception-based encoder module at the second stage. The first stage (namely Rebar Layer Identification Framework) of the proposed system extracted the layer in which most of the rebar signatures are presented. The second stage of the proposed system (namely Rebar Signature Localization framework) is used to extract and localize individual rebar signatures. This paper has provided promising results, which are highlighted in the qualitative and quantitative sub-sections of the results section.

A number of issues have been highlighted in the discussion regarding the different results obtained in the prior sections. In summary, these issues relate to the first and second stages of the proposed framework for rebar detection and localization. For the first stage of the proposed framework, future works should try to reduce the false positive rate of the rebar layer identification, so that the impact of noise can be prevented from affecting the overall efficacy of the proposed system for rebar detection and localization. For the second stage of the proposed framework, future research should try to develop a framework that is generalizable and robust to the diverse types of data characteristics that can potentially affect the performance of rebar detection and localization system. Since, this type of exploration for rebar detection and localization has not been conducted before. Future research will need to further examine and establish the feasibility of using multi-stage framework and inclusion of rebar layer to provide a better performance of rebar detection and localization systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgement

This work is supported by the U.S. National Science Foundation (NSF) under grants NSF-CAREER: 1846513 and NSF-PFI-TT: 1919127, and the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) under Grant No. 69A3551747126 through INSPIRE University Transportation Center, and the Vingroup Joint Stock Company and supported by Vingroup Innovation Foundation (VINIF) under project code VINIF.2020.NCUD.DA094. The views, opinions, findings, and conclusions reflected in this publication are solely those of the authors and do not represent the official policy or position of the NSF, USDOT/OST-R, and VINIF.

The authors would like to thank Duc La and Ethan Chang for their help to label the data.

References

- Adegun, A., Viriri, S.S., 2019. Deep Learning-Based System for Automatic Melanoma Detection. *IEEE Access*, pp. 7160–7173. <https://doi.org/10.1109/ACCESS.2019.2962812>. (Accessed 22 June 2022).
- Ahmed, H., La, H.M., 2021. Steel defect detection in bridges using deep encoder-decoder networks. In: *Proceedings of the 13th International Workshop on Structural Health Monitoring 2022*, pp. 842–849.
- Ahmed, H., Gucunski, N., La, H.M., 2019a. Rebar detection using ground penetrating radar with state-of-the-art convolutional neural networks. In: *The 9th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, pp. 1–6. URL: <https://ara.cse.unr.edu/wp-content/uploads/2014/12/SHMI-I-GPR-Paper-Final-Version-4.pdf>. (Accessed 20 June 2022).
- Ahmed, H., La, H.M., Pekcan, G., 2019b. Rebar detection and localization for non-destructive infrastructure evaluation using deep residual networks. In: *Proceedings of the 14th International Symposium on Visual Computing*, pp. 1–6. https://doi.org/10.1007/978-3-030-33720-9_49. (Accessed 2 June 2022).
- Ahmed, H., La, H.M., Tran, K., 2020a. Rebar detection and localization for bridge deck inspection and evaluation using deep residual network. *Autom. Construct.* 1–28. <https://doi.org/10.1016/j.autcon.2020.103393>. (Accessed 2 June 2022).
- Ahmed, H., La, H.M., Gucunski, N., 2020b. Review of non-destructive civil infrastructure evaluation for bridges: state-of-the-art robotic platforms, sensors and algorithms. *Sensors* 1–38. <https://doi.org/10.3390/s20143954>. (Accessed 2 June 2022).
- Ahmed, H., La, H.M., Tavakoli, A., 2021. Use of deep encoder-decoder network for sub-surface inspection and evaluation of bridge decks. In: *Proceedings of the 13th International Workshop on Structural Health Monitoring, 2022*, pp. 832–841.
- Ahmed, H., Nguyen, S.T., La, D., Le, C.P., La, H.M., 2022. Multi-directional bicycle robot for bridge inspection with steel defect detection system. In: *2022 IEEE International Conference on Intelligent Robots and Systems (IROS)*, pp. 4617–4624.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, pp. 2481–2496. <https://doi.org/10.1109/TPAMI.2016.2644615>. (Accessed 2 June 2022).
- Besaw, L.E., Stimac, P.J., 2015. Deep convolutional neural networks for classifying gpr b-scans. In: *Proceedings of Conference on Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XX. SPIE International Society of Optical Engineering*, Baltimore, MD, pp. 1–8. <https://doi.org/10.1117/12.2176250>. (Accessed 2 June 2022).
- Briaud, J.L., Brandimarte, L., Wang, J., 2019. Probability of scour depth exceedance owing to hydrologic uncertainty. *Georisk* 77–88. <https://doi.org/10.1080/17499510701398844>. (Accessed 22 May 2022).
- Capineri, L., Grande, P., Temple, A.G., 1998. Advanced image processing technique for real-time interpretation of ground penetrating radar images. *Int. J. Imag. Syst. Technol.* 51–59. [https://doi.org/10.1002/\(SICI\)1098-1098](https://doi.org/10.1002/(SICI)1098-1098). (Accessed 2 June 2022).
- Chen, H., Cohn, A.G., 2010. Probabilistic robust hyperbola mixture model for interpreting ground penetrating radar data. In: *Proceedings of IEEE World Congress on Computing Intelligence*, pp. 3367–3374. <https://doi.org/10.1109/IJCNN.2010.5596298>. (Accessed 22 June 2022).
- Chen, S., Lin, H., Yao, M., 2019. Improving the Efficiency of Encoder-Decoder Architecture for Pixel-Level Crack Detection. *IEEE Access*, pp. 186657–186671. <https://doi.org/10.1109/ACCESS.2019.2961375>. (Accessed 22 June 2022).
- Chen, S., Duan, C., Yang, Y., Li, D., Feng, C., Tian, D., 2020. Deep unsupervised learning of 3d point clouds via graph topology inference and filtering. *IEEE Trans. Image Process.* 3183–3199. <https://doi.org/10.1109/TIP.2019.2957935>. (Accessed 12 June 2022).
- Cook, W., Barr, P.J., Haling, M.W., 2015. Bridge failure rate. *J. Perform. Constr. Facil.* 97–109. <https://doi.org/10.26076/bfec-580b>. (Accessed 22 June 2022).
- Dinh, K., Gucunski, N., Doung, T.H., 2019. An algorithm for automatic localization and detection of rebars from gpr data of concrete bridge decks. *Autom. Construct.* 292–298. <https://doi.org/10.1016/j.autcon.2018.02.017>. (Accessed 2 June 2022).
- M. M. Flint, S. L. Fringer, O. Billington, D. Freyberg, N. S. Dffenbaugh, Historical analysis of hydraulic bridge collapses in the continental United States, *J. Infrastruct. Syst.*: doi/abs/10.1061/(ASCE)IS.1943-555X.0000354 [Accessed on 22 June 2022].
- Gao, F., Wu, T., Chu, X., Yoon, H., Xu, Y., Patel, B., 2020. Deep residual inception encoder-decoder network for medical imaging synthesis. *IEEE J. Biomed. Health Inf.* 39–50. <https://doi.org/10.1109/JBHI.2019.2912659>. (Accessed 22 June 2022).
- Gibb, S., La, H.M., 2016. Automated rebar detection for ground penetrating radar. In: *Proceedings of the 12th International Symposium on Visual Computing*, pp. 815–825. https://doi.org/10.1007/978-3-319-50835-1_73. (Accessed 20 June 2022).
- Gucunski, N., Maher, A., Basily, B., La, H.M., Lim, R.S., Parvardeh, H., Kee, S.H., 2013. Robotic platform rabbit for condition assessment of concrete bridge decks using multiple nde technologies. *J. Croatian Soc. Non-Destructive Testing. Croatian Soc. Non-Destructive Testing* 5–12. URL: <https://hrcak.srce.hr/148772>. (Accessed 12 June 2022).
- Gucunski, N., Kee, S.H., La, H.M., Basily, B., Maher, A., 2015a. Delamination and concrete quality assessment of concrete bridge decks using a fully autonomous rabbit platform. *Int. J. Struc. Monit. Maintenance* 19–34. <https://doi.org/10.12989/smm.2015.2.1.019>. (Accessed 2 June 2022).
- Gucunski, N., Kee, S.H., La, H., Basily, B., Maher, A., Ghasemi, H., 2015b. Implementation of a fully autonomous platform for assessment of concrete bridge decks rabbit. In: *Structures Congress 2015*, pp. 367–378.
- Harkat, H., Ruano, A.E., Ruano, M.G., Bennani, S.D., 2016. Gpr target detection using a neural network classifier designed by a multi-objective genetic algorithm. *Appl. Soft Comput.* J. 310–325. <https://doi.org/10.1016/j.asoc.2019.03.030>. (Accessed 2 June 2022).
- He, K., Zhang, X., Ren, S., Sun, J., 2015. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1904–1925. https://doi.org/10.1007/978-3-319-10578-9_23. (Accessed 22 June 2022).
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition. CVPR*, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.

- Hou, F., Lei, W., Xi, J., Xu, M., Luo, J., 2021. Improved mased r-cnn with distance guided intersection over union for gpr signature detection and segmentation. *Autom. Construct.* 1–14. <https://doi.org/10.1016/j.autcon.2020.103414>. (Accessed 2 June 2022).
- Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Efficient Convolutional Neural Networks for Mobile Vision Applications, pp. 1–9. <https://doi.org/10.48550/arXiv.1704.04861>. ArXiv. (Accessed 22 June 2022).
- Kaur, P., Dana, K.J., Romero, F.A., Gucunski, N., 2016. Automated gpr rebar analysis for robotic bridge deck evaluation. *IEEE Trans. Cybern.* 46 (10), 2265–2276. <https://doi.org/10.1109/TCYB.2015.2474747>. (Accessed 2 June 2022).
- Kim, J.-H., Lee, J.-S., 2018. Deep residual network with enhanced upscaling module for super-resolution. In: Proceedings of ICCV, pp. 913–922. URL: https://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w13/Kim_Deep_Residual_Network_CVPR_2018_paper.pdf. (Accessed 22 June 2022).
- Kim, M.-K., Thedja, J.P.P., Chi, H.-L., Lee, D.-K., 2021. Automated rebar diameter classification using point cloud data based machine learning. *Autom. Construct.* 1–14. <https://doi.org/10.1016/j.autcon.2020.103476>. (Accessed 2 June 2022).
- Kirk, R.S., Mallett, W.J., 2013. Highway bridge conditions: issues for congress. Washington, DC: US Congressional Research Service. URL: <https://crsreports.congress.gov/product/pdf/R/R44459/>. (Accessed 22 May 2022).
- La, H.M., Lim, R.S., Basily, B., Gucunski, N., Yi, J., Maher, A., Romero, F.A., Parvardeh, H., 2013a. Mechatronic systems design for an autonomous robotic system for high-efficiency bridge deck inspection and evaluation. *IEEE ASME Trans. Mechatron.* 18 (6), 1655–1664. <https://doi.org/10.1109/TMECH.2013.2279751>. (Accessed 12 June 2022).
- La, H.M., Lim, R.S., Basily, B., Gucunski, N., Yi, J., Maher, A., Romero, F.A., Parvardeh, H., 2013b. Autonomous robotic system for high-efficiency non-destructive bridge deck inspection and evaluation. In: 2013 IEEE International Conference on Automation Science and Engineering (CASE), pp. 1053–1058. <https://doi.org/10.1109/CoASE.2013.6653886>.
- La, H.M., Gucunski, N., Kee, S.-H., Nguyen, L., 2014a. Visual and acoustic data analysis for the bridge deck inspection robotic system. In: The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC), pp. 50–57. <https://doi.org/10.22260/ISARC2014/0008>. (Accessed 22 June 2022).
- La, H.M., Gucunski, N., Kee, S.-H., Yi, J., Senlet, T., Nguyen, L., 2014b. Autonomous robotic system for bridge deck data collection and analysis. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, pp. 1950–1955. <https://doi.org/10.1109/IROS.2014.6942821>. (Accessed 22 June 2022).
- La, H.M., Gucunski, N., Kee, S.-H., Nguyen, L., 2015. Data analysis and visualization for the bridge deck inspection and evaluation robotic system. *Visualization Eng.* 1–16. <https://doi.org/10.1186/s40327-015-0017-3>. (Accessed 22 June 2022).
- La, H.M., Gucunski, N., Dana, K., Kee, S.-H., 2017. Development of an autonomous bridge deck inspection robotic system. *J. Field Robot.* 34 (8), 1489–1504. <https://doi.org/10.1002/rob.21725>. (Accessed 22 June 2022).
- La, H.M., Dinh, T.H., Pham, N.H., Ha, Q.P., Pham, A.Q., 2019. Automated robotic monitoring and inspection of steel structures and bridges. *Robotica* 37 (5), 947–967. <https://doi.org/10.1017/S0263574717000601>. (Accessed 12 June 2022).
- Lee, G.C., Mohan, S.B., Huang, C., Fard, B.N., 2017. A Study of Bridge Failures (1980–2012). The State University of New York, Buffalo, NY, pp. 1–16.
- Lei, W., Hou, F., Xi, J., Tan, Q., Xu, M., Jiang, G., Liu, X., Gu, Q., 2019. Automatic hyperbola detection and fitting in gpr b-scan image. *Autom. Construct.* 1–14. <https://doi.org/10.1016/j.autcon.2019.102839>. (Accessed 2 June 2022).
- Li, Y., Lu, Y., Chen, J., 2021. A deep learning approach for real-time rebar counting on the construction site based on yolov3 detector. *Autom. Construct.* 1–14. <https://doi.org/10.1016/j.autcon.2021.103602>. (Accessed 2 June 2022).
- Liu, Y., Yao, J., Lu, X., Xie, R., Li, L., 2019. Deepcrack: a deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* 139–153. <https://doi.org/10.1016/j.neucom.2019.01.036>. (Accessed 22 June 2022).
- Liu, H., Lin, C., Cui, J., Fan, L., Xie, X., Spencer, B.F., 2020. Detection and localization of rebar in concrete by deep learning using ground penetrating radar. *Autom. Construct.* 1–12. <https://doi.org/10.1016/j.autcon.2020.103279>. (Accessed 2 June 2022).
- Ma, H., Yao, H., Li, Y., Wang, H., 2020. Deep residual encoder-decoder networks for desert seismic noise suppression. *Geosci. Rem. Sens. Lett. IEEE* 17 (3), 529–533. <https://doi.org/10.1109/LGRS.2019.2925062>. (Accessed 22 June 2022).
- Nakazawa, T., Kulkarni, D.V., 2019. Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder decoder neural network architectures in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* 250–257. <https://doi.org/10.1109/TSM.2019.2897690>. (Accessed 22 June 2022).
- Naqvi, R.A., Loh, W.-K., 2019. Sclera-net: Accurate Sclera Segmentation in Various Sensor Images Based on Residual Encoder and Decoder Network. *IEEE Access*, pp. 98208–98228. <https://doi.org/10.1109/ACCESS.2019.2930593>. (Accessed 22 June 2022).
- Neumann, J., 2017. Climate change risks to us infrastructure: impacts on roads, bridges, coastal development and urban drainage. *Clim. Change* 97–109. <https://doi.org/10.1007/s10584-013-1037-4>. (Accessed 2 June 2022).
- News, C., 2019. Bridge Collapse in Taiwan Caught on Video Oil Tanker Falls onto Boats in Fishing Harbor. CBS News. Retrieved from: <https://www.cbsnews.com/news/bridge-collapse-in-taiwan-caught-on-video-oil-tanker-falls-onto-boats-nanfengao-bay-today-2019-10-01/>. (Accessed 22 May 2022).
- Pan, S., Zhang, W., Zhang, W., Xu, L., Fan, G., Gong, J., Zhang, B., Gu, H., 2019. Diagnostic model of coronary microvascular disease combined with full convolution deep network with balanced cross-entropy cost function. *IEEE Access* 7, 177997–178006. <https://doi.org/10.1109/ACCESS.2019.2958825>. (Accessed 22 June 2022).
- Penn, A., 2018. The Deadliest Bridge Collapses in the US in the Last 50 Years. CNN. Retrieved from: <https://www.cnn.com/2018/03/15/us/bridge-collapse-history-trnd/index.html>. (Accessed 22 May 2022).
- Ronneberger, O., Fischer, P., Brox, T., U-net, 2015. Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, Cham, pp. 234–241.
- Salem, M., Valverde, S., Cabezas, M., Pareto, D., Oliver, A., Salvi, J., Rovira, A., Llado, X., 2019. Multiple Sclerosis Lesion Synthesis in Mri Using an Encoder-Decoder U-Net. *IEEE Access*, pp. 25171–25185. <https://doi.org/10.1109/ACCESS.2019.2900198>. (Accessed 22 June 2022).
- Spoorthi, G.E., G. S., Gorthi, R.K.S.S., 2019. Phasenet: a deep convolutional neural network for two-dimensional phase unwrapping. *IEEE Signal Process. Lett.* 54–59. <https://doi.org/10.1109/LSP.2018.2879184>. (Accessed 22 June 2022).
- Wang, S., Hou, X., Zhao, X., 2020. Automatic Building Extraction from High-Resolution Aerial Imagery via Fully Convolutional Encoder-Decoder Network with Non-local Block. *IEEE Access*, pp. 7313–7323. <https://doi.org/10.1109/ACCESS.2020.2964043>. (Accessed 22 June 2022).
- Windsor, C.G., Capineri, L., Falorni, P., 2005. The estimation of buried pipe diameters by generalized hough transform of radar data. *Prog. Electromagnetics Res. Sympos.* 345–349. <https://doi.org/10.2529/PIERS041117130829>. (Accessed 2 June 2022).
- Windsor, C.G., Capineri, L., Falorni, P., 2014. A data pair-labeled generalized hough transform for radar location of buried objects. *Geosci. Rem. Sens. Lett. IEEE* 124–127. <https://doi.org/10.1109/LGRS.2013.2248119>. (Accessed 22 June 2022).
- Wright, L., 2012. Estimated Effect of Climate Change on Flood Vulnerability of US Bridges, Mitigation Adaptation Strategies Global Change, pp. 939–955. (Accessed 22 May 2022).
- Xiang, Z., Ou, G., Rashidi, A., 2021. Robust cascaded frequency filters to recognized rebar in gpr data with complex signal interference. *Autom. Construct.* 1–12. <https://doi.org/10.1016/j.autcon.2021.103593>. (Accessed 2 June 2022).
- C. Yuan, S. Li, H. Cai, V. R. Kamat, Gpr signature detection and decomposition for mapping buried utilities with complex spatial configuration, *J. Comput. Civ. Eng.*. doi:<https://doi.org/10.1061/%28ASCE%29CP.1943-5487.0000764> [Accessed on 22 June 2022].
- Zhou, Z., Rahman, S.M.M., Tajbakhsh, N., Liang, J., 2018. Unet++: a nested u-net architecture for medical image segmentation. *Lect. Notes Comput. Sci.* 11045, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1. (Accessed 12 June 2022).
- Zhou, Y.-F., Jiang, R.-H., Wu, X., He, J.-Y., Weng, S., Peng, Q., 2019a. Branchgan: unsupervised mutual image-to-image transfer with a single encoder and dual decoders. *IEEE Trans. Multimed.* 3136–3150. <https://doi.org/10.1109/TMM.2019.2920613>. (Accessed 2 June 2022).
- Zhou, W., Lv, S., Jiang, Q., Yu, L., 2019b. Deep road scene understanding. *IEEE Signal Process. Lett.* 26 (4), 587–591. <https://doi.org/10.1109/LSP.2019.2896793>. (Accessed 22 June 2022).