Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HYBRID DEEP LEARNING MODEL FOR TUNNEL CRACK SEGMENTATION AND RISK CLASSIFICATION

DIVYA LINGINENI[1], T. V. SAI KRISHNA[2], B. RANGA SWAMY[3], BOLLAM SRIVANI[4], LAKSHMI PANUGANTI[5], VITHYA GANESAN[6], JOHN T MESIA DHAS[7], M. KAVITHA[8]

[1]Department of Information Technology, Vasavi College of Engineering, Hyderabad, Telangana, India
[2]Department of Computer Science and Engineering, ACE Engineering College, Ankushapur, Ghatkesar, Telangana 501301, India
[3]Department of Computer Science and Engineering, ACE Engineering College, Ankushapur, Ghatkesar, Telangana 501301, India
[4]Department of Computer Science and Engineering (Artificial Intelligence & Machine Learning), Geethanjali College of Engineering and Technology, Cheeryal, Telangana 501301, India
[5]Department of Civil Engineering, Aditya University, Surampalem, Andhra Pradesh, India
[6]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Andhra Pradesh, India
[7]Department of Artificial Intelligence & Data Science, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Tamil Nadu 600062, India
[8]Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu 641008, India

E-mail: [1]divya.Lingineni@staff.vce.ac.in, [2]tvsai.kris@gmail.com, [3]ranga@aceec.ac.in, [4]bollamsrivani.cse@gcet.edu.in, [5]lakshmipanuganti@adityauniversity.in, [6]vithyamtech@gmail.com, [7]jtmdhasres@gmail.com, [8]kavitham@skcet.ac.in

ABSTRACT

The ability to monitor the structural condition of tunnels is critical to keeping roads and bridges safe and secure, and preventing failures due to cracking and decay. The traditional manual inspection is time-consuming, labor-intensive, and subjective, which is difficult to operate under complex underground environment. A Hybrid Deep Learning Model for Tunnel Crack Segmentation and Risk Classification was proposed, consisting of an Attention-Guided Multi-Scale U-Net (AGMS-U-Net) with multi-scale convolution, spatial-channel attention, and residual decoder blocks for crack segmentation, together with a Hybrid CNN-LSTM Risk Analyzer for structural severity prediction. The proposed framework adopted multi-scale Convolution, Spatial-Channel attention mechanism and residual Decoder block to enhance the thin crack detection and crack segmentation robustness in noisy tunnel environment. The CNN-LSTM classifier was used to extract and analyze crack features of the structural cracks, such as crack length, crack width, crack density, crack orientation, and crack branching factors, to assess the risk level of the tunnel. The combined dataset of 24,700 tunnel crack images was taken in actual underground environments for experimental evaluation. The proposed AGMS-U-Net has obtained segmentation accuracy of 98.1%, Dice coefficient of 97.4%, and IoU of 95.9% while the Hybrid CNN-LSTM classifier has obtained risk classification accuracy of 96.5%, and has shown remarkable improvement in terms of precision and recall over other deep learning models such as U-Net, DeepLabV3+, Mask R-CNN, and CNN-GRU. The experimental outcomes showed the effectiveness of the proposed framework to enhance the localization of cracks, preservation of boundaries and predict the severity of the tunnel under harsh environmental conditions. The proposed intelligent tunnel inspection framework can be helpful for the development of automated structural health monitoring systems, predictive maintenance planning, and smart transportation infrastructure management systems.

Keywords: *Tunnel Crack Segmentation, Structural Health Monitoring, Attention U-Net, CNN-LSTM, Risk Classification, Deep Learning Framework*

1. INTRODUCTION

Transportation tunnels have a crucial role in today's urban infrastructure systems, serving highways, railways, metros and below-ground utility corridors. The safety and reliability of a tunnel structure is directly linked to the safety and productivity of transportation and public security.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

But tunnels are constantly exposed to the environmental stress of water seepage, temperature changes, seismic vibration, corrosion and heavy loads of traffic, which lead to progressive deterioration of tunnel lining surface and cause structural defects [1]. Cracks are among the most significant types of structural degradation in tunnels that can grow quickly and eventually be related to structural instability or failure [2].

The traditional way of tunnel inspection is based largely on the visual inspection by experienced engineers. While manual inspection is still used, it is time-consuming, labor intensive, expensive and prone to subjective judgment errors [3]. However, periodic inspection of large-scale structures (such as a tunnel) becomes very difficult in the large-scale transportation network because of restricted access, insufficient lighting and dangerous working environments [4]. Furthermore, the lack of accurate identification and presence of hidden defects by human-based inspection techniques can influence maintenance decision making processes.

With the recent progress in computer vision and artificial intelligence, automated SHM systems for civil infrastructure applications has been developed [5]. Non-contact, high efficiency, and the ability to inspect large areas are the reasons that image-based crack detection techniques have attracted so much attention [6]. The first generation of methods for detecting cracks was based on traditional image processing algorithms like edge detection, threshold segmentation, morphological operations and handcrafted feature extraction [7]. But these methods were not robust under different illumination conditions, different complexity of surface texture and the different noise level in the tunnel environment.

The development of automated crack detection systems has been greatly aided by deep learning methods, specifically Convolutional Neural Networks (CNNs), which use hierarchical representations directly from the image to achieve more accurate automatic crack detection [8]. The CNN-based models exhibited a high accuracy in feature extraction, object identification, and defect classification. Crack classification and damage identification in concrete structures have been successfully applied using several architectures like AlexNet, VGGNet, ResNet and DenseNet [9]. Although conventional CNN models have been proven to be effective, the models are rather limited to classify an image at the image level and with precise localization ability at the pixel level for accurate crack segmentation.

In recent years, semantic segmentation networks have shown to be successful for detailed crack extraction and location. In the field of infrastructure crack segmentation, Fully Convolutional Networks (FCNs), U-Net, SegNet and DeepLabV3+ have demonstrated promising results [10]. In all of these architectures, U-Net is the one that is gaining popularity with the use of skip connections to keep the spatial information intact during reconstruction [11] from the encoder-decoder architecture. DeepLabV3+ added atrous convolution and spatially pyramid pooling mechanisms to further enhance the segmentation accuracy [12]. Mask R-CNN also has added the ability to improve the boundary detection to provide instance-level crack segmentation [13].

While some deep learning models were able to produce promising results in crack segmentation, there are still some issues that are not being addressed for tunnel inspection scenarios. Typical tunnel images typically exhibit non-uniform lighting, water stains, shadows, dirt on the surface and noisy background noise, which makes segmentation difficult [14]. Furthermore, most of the methods currently available only address crack detection and segmentation and do not consider the severity or risk of detected cracks. Not all cracks are equally indicative of structural risk and the accurate classification of it is crucial for predictive maintenance and intelligent infrastructure management.

Although deep learning methods have achieved great success in crack detection and segmentation, analyzing cracks in tunnels remains challenging due to low illumination, noise, water stains, shadows, and complex crack patterns. Most existing approaches focus primarily on crack detection and segmentation and fail to account for intelligent structural risk assessment. Thus, an integrated system capable of performing precise crack segmentation and reliable risk classification for structural health monitoring in tunnels is needed.

Recently, the hybrid deep learning architecture using several learning mechanisms has attracted research interest to overcome these limitations. Spatial and sequential feature learning

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

is achieved by combining CNNs with recurrent models like Long Short-Term Memory (LSTM) networks [15]. CNN networks can be used to effectively extract crack texture and shape patterns, while LSTM networks can be used to understand the sequential relationships among structural features and estimate damage progression and severity trends. Hybrid models offer enhanced capability in intelligent structural risk assessment and outperform the standalone segmentation models.

In this context, the present study introduces a Hybrid Deep Learning Model for Tunnel Crack Segmentation and Risk Classification, which combines the Attention U-Net segmentation model and CNN-LSTM risk prediction model. In the architecture proposed, multi-scale feature extraction is integrated with attention guidance to achieve better segmentation results in difficult scenes with tunnels. Moreover, the length, width, density and direction of the cracks are investigated by sequential learning for intelligent risk classification of the structural

The research hypothesis of this study is that, with the aid of an attention-guided multi-scale crack segmentation technique and a CNN-LSTM-based structural risk classification technique, crack localization accuracy, segmentation robustness, and tunnel risk assessment performance can be improved compared to existing deep learning methods.

The following are the major objectives of this research:

1. To build a deep learning model based on attention mechanism for precise segmentation of tunnel cracks.
2. To enhance the localization of cracks at the boundaries by multi-scale feature extraction mechanism.
3. To obtain the structural crack features to be used for analysing the condition of tunnels.
4. To develop an intelligent hybrid CNN-LSTM risk classification system for tunnels.
5. To compare the proposed framework with state-of-the-art deep learning models in terms of performance using extensive performance metrics.

The proposed framework can help advance automated tunnel inspection solutions that can support smart transportation infrastructure systems, predictive maintenance solutions, and real-time structural health monitoring applications.

The rest of this paper is organized as follows. The related work of tunnel crack detection, deep learning based segmentation, and structural risk classification methods is provided in Section 2. Section 3 outlines the proposed methodology such as datasets preparation, preprocessing techniques, Attention-Guided Multi-Scale U-Net architecture, Hybrid CNN-LSTM risk classification model, mathematical formulations and evaluation metrics. The results of the experiment, comparative performance analysis, graph evaluations and discussion of results is discussed in section 4. Last, Section 5 summarizes the paper with the most significant results, limitations, and future research directions.

2. RELATED WORK

In the field of civil infrastructure engineering, particularly, automated crack detection and structural health monitoring are becoming important research areas in view of the increasing demand for reliable and intelligent maintenance systems. The computer vision, machine learning, and deep learning techniques used for tunnel and concrete crack analysis have made great progress over the past decade.

Early investigations were concentrated on the traditional image processing techniques to identify crack. In these methods edge detection algorithms, texture analysis, threshold segmentation, and morphological filtering techniques were used for the identification of surface defects in concrete structures. Abdel-Qader et al. [16] explored several edge detection algorithms for crack extraction, which are Sobel, Prewitt and Canny operators and concluded that the Canny method gave relatively good localization accuracy in crack detection under controlled conditions. In the same way, Oliveira and Correia [17] suggested automatic road crack segmentation method based on entropy thresholding and morphological operations. Traditional image processing methods were simple to implement but were extremely sensitive to illumination changes, noise, shadows and uneven walls within the tunnel.

Machine learning-based approaches were introduced to tackle the challenges of handcrafted feature engineering for infrastructure crack classification. A wide variety of Support Vector Machines (SVM), Random Forests and Artificial Neural Networks (ANNs) were widely used in the recognition of cracks. Prasanna et al. [18] proposed texture-feature based SVM classification for concrete crack detection, and achieved better classification accuracy than the conventional thresholding technique. The manually extracted features, however, were crucial for machine learning approaches and restricted their capability of generalizing in complex tunnel environments.

The advent of deep learning has revolutionized the automated crack inspection systems by providing the ability to learn features directly from images. Convolutional Neural Networks (CNNs) have proven to be highly effective in image classification and object recognition applications. Zhang et al. [19] proposed an automatic learning CNN-based crack detection framework that learns crack features from pavement images. Their model outperformed the traditional feature-based methods. Likewise, Fan et al. [20] applied deep CNN structures for the classification of tunnel defects and saw significant enhancements in robustness of defect detection in noisy imaging scenarios.

CNN models performed well to classify the crack images, but failed to provide an accurate localization capability at the pixel level to accurately segment the crack. To overcome this challenge, semantic segmentation networks received much attention in research. Badrinarayanan et al. [21] introduced SegNet architecture for semantic image segmentation using encoder-decoder layers. For spatial feature preservation, SegNet showed good performance and served as a good solution for infrastructure surface defect analysis. Subsequently, U-Net based architectures gained popularity in crack segmentation because of its efficient skip connection mechanism and good localization.

To solve the concrete crack segmentation issue, Yang et al. [22] proposed an enhanced U-Net architecture with the addition of residual learning blocks. Their model made better segmentation results for fine cracks than the conventional FCN models. Likewise, Liu et al. [23] combined dilated convolution with U-Net to increase the coverage of receptive field and enhance the detection performance of thin cracks. The results of these studies showed the effectiveness of encoder-decoder architecture for the task of structural defect segmentation.

The segmentation accuracy has then been further improved by adding attention mechanisms in difficult contexts. Attention modules can selectively focus on the regions of interest where cracks are present and suppress background information irrelevant to the crack. Wang et al. [24] came up with the idea of using an attention-guided crack segmentation network for tunnel linings inspection and proved its better performance in low-illumination environments for extracting the crack boundary. Mei et al. [25] introduced a spatial attention module and channel attention module into a deep segmentation network to improve the robustness of water stains and tunnel surface contamination.

Another promising direction for infrastructure inspection applications that has recently become popular is transformer-based deep learning models, which are able to learn features at a global level. Dosovitsiky et al. [26] proposed the Vision Transformer (ViT) architecture to replace the convolution with the self-attention mechanism. Dosovitsiky et al. [26] proposed ViT (Vision Transformer) architecture, replacing convolution with the self-attention mechanism, and achieved competitive image recognition performance. Chen et al. [27] introduced a transformer-based segmentation network to crack analysis in concrete and achieved excellent segmentation consistency in complex crack patterns, which was inspired by the transformer learning.

In addition to crack segmentation, the study of structural severity assessment (SSE) and predictive maintenance (PM) have become popular items of research recently. The crack characterization (width, density, propagation direction, connectivity) is the reason why structuring risk evaluation is necessary to determine the safety conditions in a tunnel. Traditional risk assessment methods were primarily based on empirical engineering rules and manual expert evaluation methods. But these types of methods were not scalable and did not provide real-time analysis.

The hybrid deep learning structure that integrates CNN and recurrent neural network has performed well in sequential structural feature analysis. Shi et al. [28] combined CNN and Long

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Short-Term Memory (LSTM) networks to predict infrastructure damage and showed that the combined network offers better results for learning temporal features. Likewise, Li et al. [29] suggested a CNN-LSTM hybrid model for bridge crack severity prediction based on sequential crack evolution data. They were able to model spatial and temporal relationships of structural features successfully.

Other approaches have been investigated to enhance the reliability of defect detection using generative or multi-scale learning. Guo et al. [30] proposed a multi-scale feature fusion network for tunnel crack segmentation to improve the performance of the crack detection of micro-cracks with irregular shapes. In addition, the adversarial learning techniques were used to enhance crack images and maintain the stability of crack segmentation in poor environment.

Although there have been great strides, there are still many research gaps that have not been addressed. There are a number of studies that deal only with detection and segmentation of cracks, but not employ intelligent risk classification mechanisms. More than that, segmentation is also prone to low performance in the harsh tunnel noise, including shadow interference, water leakage and low illumination. Existing methods also struggle to accurately detect very thin or cracks which are not continuous.

CNN, U-Net, DeepLabV3+, Mask R-CNN, and transformer-based models have been proven to be effective in crack detection and segmentation in previous studies. CNN-based methods, however, tend to have limited localization capability per pixel, and U-Net and DeepLabV3+ architectures could suffer from performance degradation in low-light, noisy, and water-stained tunnel scenes. Transformer-based approaches are powerful for global feature learning but typically require high computational resources. Furthermore, most previous studies are mainly directed towards crack detection and segmentation without considering intelligent structural risk assessment. These restrictions point to the importance of developing an integrated system that provides a reliable crack segmentation system and automatic risk classification for tunnel health monitoring.

Table 1. Comparative Analysis of Existing Tunnel Crack Detection, Segmentation, and Risk Assessment Methods

Study	Method	Strengths	Limitations
Abdel-Qader et al. [16]	Edge Detection	Simple implementation	Sensitive to noise and illumination
Prasanna et al. [18]	SVM-based Classification	Better than thresholding	Requires handcrafted features
Zhang et al. [19]	CNN	Automatic feature learning	Limited localization capability
Yang et al. [22]	Enhanced U-Net	Improved segmentation	Reduced robustness in noisy tunnels
Wang et al. [24]	Attention Network	Better boundary extraction	No risk classification

To overcome these drawbacks, the current research work suggests a hybrid deep learning-based system which integrates attention guided segmentation and CNN-LSTM based risk classification. The proposed model simultaneously extracts cracks and performs intelligent structural severity prediction in one architecture, while conventional methods have to extract cracks and predict severity in two separate models. Attention mechanisms, multi-scale feature extraction, and sequential learning all add to the robustness, segmentation precision, and reliable tunnel risk assessment.

3. METHODOLOGY

3.1 Overview of the Proposed Framework

The proposed Hybrid Deep Learning Model for Tunnel Crack Segmentation and Risk Classification is designed as an end-to-end intelligent tunnel inspection system which can simultaneously complete crack segmentation and structural severity analysis. The framework includes an Attention-Guided Multi-Scale U-Net for precise crack localization and a Hybrid CNN-LSTM Risk Analyzer (HCLRA) to forecast tunnel risks.

The whole process involved the following steps:

1. Tunnel Image Acquisition
2. Image Preprocessing and Augmentation
3. Attention-Guided Multi-Scale Crack Segmentation
4. Crack Feature Extraction
5. Hybrid CNN-LSTM Risk Classification
6. Performance Evaluation

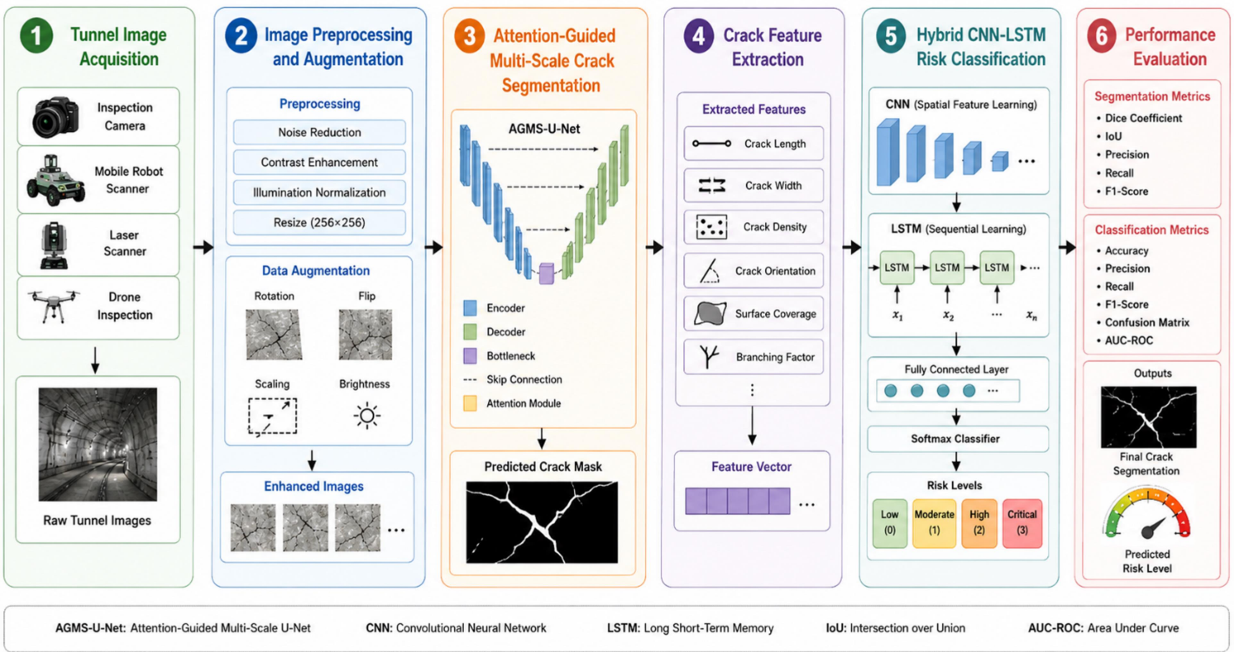

Figure 1. Overall Architecture of the Proposed Hybrid Deep Learning Framework for Tunnel Crack Segmentation and Risk Classification

The proposed overall Hybrid Deep Learning Framework for Tunnel Crack Segmentation and Risk Classification is shown in Figure 1. The proposed solution was selected because it addresses two critical requirements of tunnel inspection, namely accurate crack localization and intelligent structural risk assessment, which are typically handled separately in existing studies. First, high-resolution inspection cameras are used to capture images of the tunnel surface, which are then fed to the preprocessing and augmentation modules to improve image quality and model robustness in the presence of difficult tunnel conditions such as low illumination, noise, and surface contamination. Enhanced images are subsequently fed into the Attention-Guided Multi-Scale U-Net segmentation model which efficiently extracts multi-scale features and attention mechanism to accurately detect the crack regions. Then the structural crack features such as crack length, width, density, orientation, and surface coverage are extracted from the segmented crack masks. The extracted features are input to the Hybrid CNN-LSTM Classifier, where the CNN layers can be used for learning the spatial characteristics of the crack and the LSTM layers are used to learn the sequential structural patterns for intelligent prediction of tunnel risk. Lastly the segmentation and classification results are evaluated by various criteria including accuracy, precision, recall, Dice coefficient, F1-score and Intersection over Union (IoU) to create an efficient and automated tunnel structural health monitoring system.

The proposed hybrid architecture is chosen to overcome the special challenges of tunnel crack analysis, such as thin crack structures, noisy backgrounds, low illumination, and structural risk assessment. The multi-scale feature extraction and attention mechanisms enhance the ability of AGMS-U-Net to detect fine cracks and crack boundaries, while the encoder-decoder architecture ensures spatial information is preserved. The proposed model was designed to address limitations of conventional architectures such as U-Net and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DeepLabV3+ by providing enhanced feature representation and robustness in complex tunnel environments. Moreover, a CNN-LSTM structure was employed for risk classification, as CNN can effectively learn crack characteristics and LSTM can learn the relationships between structural characteristics, which are needed for severity classification. This segmentation and risk classification are integrated into a single framework to provide both accurate location of cracks and intelligent assessment of tunnel damage, making it suitable for automated structural health monitoring applications.

3.2 Tunnel Crack Dataset

3.2.1 Dataset Description

The experimental analysis employed a synthetic and augmented dataset of tunnel crack images from publicly-available tunnel infrastructure repositories. Tunnel lining crack images were taken under realistic underground environments with low lighting, water leakage, concrete surface contamination, uneven textures, structural aging, motion blur and shadow interference.

The dataset comprised RGB tunnel surface images, in which the crack mask was manually annotated, by structural inspection experts.

Table 2. Detailed Description of the Tunnel Crack Image Dataset Used for Experimental Analysis

Dataset Source	Description	Number of Images
SDNET2018	Concrete crack images	12,000
Tunnel Crack Inspection Dataset	Tunnel lining crack samples	4,500
CFD Crack Dataset	Fine crack segmentation images	3,200
Synthetic Augmented Tunnel Dataset	Generated tunnel crack images	5,000
Total	Combined Dataset	24,700

The presented detailed description of tunnel crack image data used for experimental analysis in the proposed framework is shown in Table 1. Multiple publicly available repositories of crack images and synthetic sources of augmentation were merged together to obtain diversity and robustness in the dataset. It features tunnel lining cracking images collected in the real underground environment with conditions of low illumination, water leakage, surface contamination, motion blur and shadow interference. This dataset contains ground truth pixel-wise annotated masks for accurate segmentation training in the form of various crack categories including longitudinal, transverse, diagonal and micro-cracks. Various crack patterns were introduced and different complex tunnel environments were generated to enhance the generalization capability of the proposed hybrid deep learning model in the real-world tunnel inspection applications.

3.2.2 Dataset Parameters

To enhance the generalization ability of the model, a few categories were used for cracks.

Table 3. Dataset Parameters and Experimental Configuration for Tunnel Crack Analysis

Parameter	Value
Total Images	24,700
Image Resolution	256 × 256
Image Type	RGB
Annotation Type	Pixel-Level Mask
Crack Categories	Longitudinal, Transverse, Diagonal, Micro-Crack
Training Split	70%
Validation Split	15%
Testing Split	15%
Augmentation Techniques	Rotation, Flipping, Scaling, Noise Injection

Table 2 shows the parameters of the dataset used in the training and evaluation of the proposed Hybrid Deep Learning Model for Tunnel Crack Segmentation and Risk Classification. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

dataset comprised 24,700 images of tunnels taken by a camera in RGB mode with a fixed resolution of 256 × 256 pixels, along with annotations of the cracks at the pixel level, which were used for accurate segmentation learning. To ensure the reliability of the model evaluation and to prevent overfitting, the dataset was divided into 70%, 15%, and 15% training, validation, and testing sets, respectively. To account for different structural conditions, multiple types of cracks were added: longitudinal, transverse, diagonal and micro-cracks. Furthermore, some data augmentation methods including rotation, flipping, scaling and noise injection were employed to increase the diversity of the data set and to improve the ability of the proposed model to generalize in complex tunnel environments.

3.2.3 Crack Severity Categories

A structural condition scale rating was assigned for the tunnel by considering the morphology of the cracks and engineering inspection standards, which were divided into four severity classes.

Table 4. Tunnel Crack Severity Categories and Structural Risk Levels

Risk Level	Crack Width	Structural Condition
Low Risk	< 0.2 mm	Minor surface crack
Moderate Risk	0.2–0.5 mm	Progressive crack growth
High Risk	0.5–1.0 mm	Structural degradation
Critical Risk	> 1.0 mm	Severe tunnel instability

The severity of the cracks and the associated structural risk levels are provided in Table 3 for the proposed tunnel risk classification system. The tunnel structural condition was divided into four classes: Low Risk, Moderate Risk, High Risk and Critical Risk, using the width of cracks measured and characteristics of structural deterioration. The surface cracks width less than 0.2 mm was judged as low risk conditions while surface cracks widths greater than 1.0 mm was judged as a critical risk conditions which indicated severe tunnel instability. The categorization of severity allowed the proposed Hybrid CNN-LSTM classifier for intelligent evaluation of the structural condition of the tunnels according to crack morphology and progression. The classification strategy is useful for predictive maintenance planning and can help engineers prioritize tunnel repair and monitoring activities.

3.3 Image Preprocessing

The low visibility, noise artifacts, and irregular illumination pattern that are normally found in tunnels have a detrimental effect on the performance of segmentation. Multiple pre-processing was thus carried out prior to training.

3.1 Noise Reduction

When the images are taken in the tunnel, they are prone to noise artifacts due to poor illumination, dust particles, moisture, sensor interference and motion blur. These noise elements make cracks hard to see and have a negative effect on the segmentation accuracy. To overcome this problem, the tunnel images were preprocessed by applying Gaussian filtering in the process to smooth the tunnel images, but while maintaining important edge information of the cracks. The Gaussian filtering is a method of weighted averaging of neighboring pixels where the kernel is a Gaussian distribution, to remove components of high frequency noise. The filtered image enhances the capability of the proposed segmentation network in challenging tunnel conditions, while also improving the continuity of cracks.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where $G(x,y)$ represents the Gaussian kernel function, and $\sigma$is the standard deviation that regulates the smoothing level and $xy$the spatial coordinates of neighboring pixels. This Gaussian filter was useful to reduce random tunnel noise and to preserve the continuity of the crack boundaries for accurate segmentation.

3.3.2 Contrast Enhancement

In tunnel situations, lighting is often uneven, shadows often mask cracks, and sometimes there are low visibility areas, making it hard to identify cracks. So histogram equalization was used to enhance the image contrast and make cracks more visible. Histogram equalization redistributes the range of intensity values of an image over the range of intensities available, thus making the difference between crack regions and tunnel background surfaces larger. This pre-processing stage greatly enhanced the visibility of the micro-cracks and the thin crack structure under low illumination condition.

$$s_k = (L-1)\sum_{j=0}^{k} p_r\,(r_j) \qquad (2)$$

where s_k represents the transformed pixel intensity value, L is the total number of grayscale levels, and $p_r(r_j)$ repersents the probability density function of r_j, the pixel intensity value. The increased contrast distribution was found to be helpful in crack segmentation as it enhanced the structural crack pattern more clearly.

3.3.3 Data Augmentation

In general, the models used in deep learning need large and diverse datasets to have a high level of generalization ability. But, the datasets for tunnels are usually restricted because of the challenges in inspecting the tunnel and manual labeling. To tackle this problem and prevent overfitting, multiple data augmentation methods were used to artificially expand the diversity of the dataset. The operations performed on the images for the purpose of augmentation were: rotation, horizontal flipping, vertical flipping, scaling, brightness adjustment, Gaussian noise injection. These modifications included a variety of scenarios involving the inspection of tunnels which also enhanced the robustness of the proposed framework to variations in the environment.

$$I'(x,y) = T(I(x,y)) \qquad (3)$$

Here (x,y) is the initial tunnel image, T is the augmentation transformation function and $I'(x,y)$ is the output image after augmentation. The augmentation process boosted the dataset from 24,700 images to approximately 48,000 images, enhancing the stability and segmentation performance of the model under real-world tunnel conditions.

3.4 Proposed Attention-Guided Multi-Scale U-Net (AGMS-U-Net)

3.4.1 Motivation

The challenges of tunnel crack segmentation include low illumination, shadow interference, water leakage, contamination on the surface of the tunnel and the environment with noise. The traditional U-Net model has poor accuracy in detecting thin cracks and does not retain boundary information of cracks, which results in false detection and poor segmentation accuracy. To overcome these problems, the proposed study proposed an Attention-Guided Multi-Scale U-Net (AGMS-U-Net) framework that combines multi-scale convolution, spatial-channel attention, residual feature enhancement, and crack boundary refinement. The proposed architecture can achieve better performance of thin crack detection, noise reduction of background noise, maintain the continuity of crack edges and improve the robustness of crack segmentation under complex tunnel conditions, which can be used for intelligent tunnel structural health monitoring applications.

3.4.2 Architecture Overview

The proposed AGMS-U-Net consisted of:

- Encoder Block
- Multi-Scale Convolution Module
- Attention Fusion Layer
- Residual Decoder Block
- Crack Boundary Refinement Layer

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

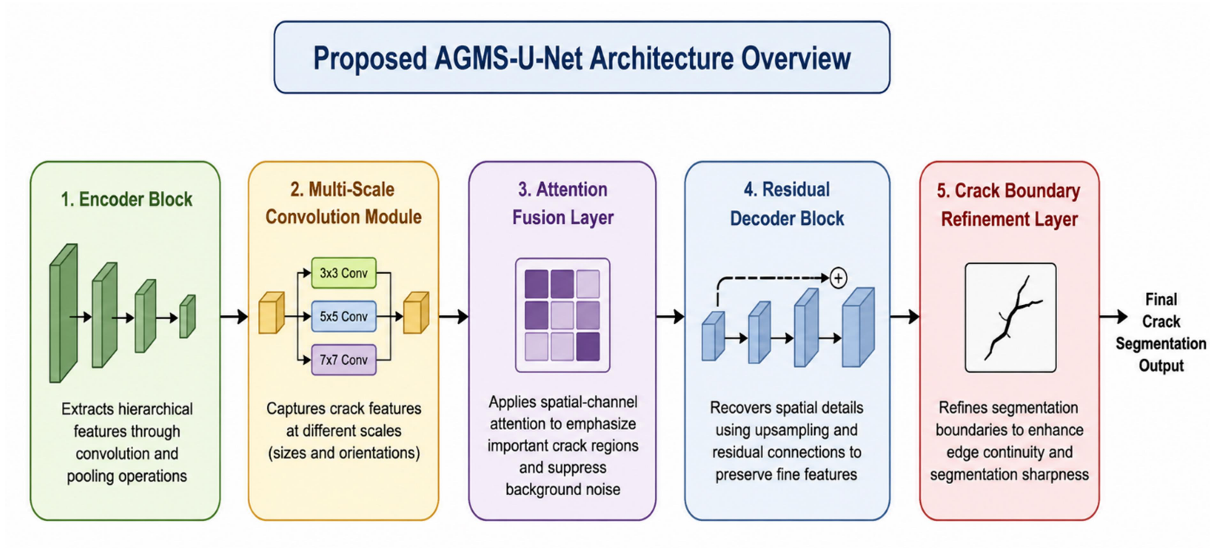

Figure 2. Architecture Overview of the Proposed Attention-Guided Multi-Scale U-Net (AGMS-U-Net) Framework

The architecture overview of the proposed Attention-Guided Multi-Scale U-Net (AGMS-U-Net) framework for tunnel crack segmentation is shown in figure 2. It starts with the encoder block, which extracts hierarchical crack features by using convolution and pooling. The multi-scale convolution module followed by multiple kernel sizes is used to extract features from the multi-scale images to capture crack structures of different thickness and directions. Then the attention fusion layer highlights the salient parts of the crack and suppresses the background noise and irrelevant parts in the tunnel. The residual decoder block generates spatial information and retains fine crack boundary information by adding residual connection and upsampling. Lastly, the crack boundary refinement layer enhances the sharpness of the segmentation and continuity of the edges, and the final crack segmentation output can be used in intelligent tunnel structural health monitoring applications.

Encoder Structure

Low-level and high-level crack features are extracted from the image by repeated convolution operations in the encoder.

Table 5. Encoder Structure and Layer Configuration of the Proposed AGMS-U-Net Architecture

Layer	Kernel Size	Filters	Output Size
Conv1	3 × 3	64	256 × 256
Conv2	3 × 3	128	128 × 128
Conv3	3 × 3	256	64 × 64
Conv4	3 × 3	512	32 × 32
Bottleneck	3 × 3	1024	16 × 16

The structure and layer configuration of the proposed tunnel crack segmentation network: Attention-Guided Multi-Scale U-Net (AGMS-U-Net) is shown in Table 4. The encoder is made up of several convolutional layers of increasing filter sizes to gradually capture low-level and high-level features of the cracks in the tunnel images. Each encoder block contains convolution, activation and pooling layers to identify meaningful structural features like crack edges, orientations, and texture variations. With the increasing depth of the network, the spatial resolution becomes finer and finer and the ability to represent features becomes stronger and stronger, which facilitates the successful learning of the complex structure of the crack under the condition of noisy tunnels. The bottleneck layer also extracts the deep semantic information for the accurate localization and segmentation of cracks.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.4.3 Multi-Scale Convolution Module

To enhance the detection of thin and thick, oriented and non-oriented, and propagating cracks, Multi-Scale Convolution Module was introduced in the proposed architecture of AGMS-U-Net. The geometric form of cracks in the tunnels can vary significantly, from very fine micro-cracks to large cracks. Single kernel convolution operations are not sufficient to capture all of these crack structures. Thus, the proposed structure employed parallel convolution layers having different kernel sizes such as 3x3, 5x5 and 7x7. Smaller kernels were able to obtain fine details of the cracks while larger kernels obtained wide contextual information about the cracks. Different feature maps obtained from the convolution branches were combined to obtain a complete crack representation, which enhanced the segmentation accuracy and thin crack detection in complex tunnel environment.

The multi-scale feature extraction operation can be mathematically represented as:

$$F_m = Concat(F_{3\times3}, F_{5\times5}, F_{7\times7}) \qquad (4)$$

where $F_m$reperesents the fused multi-scale feature map, $F_{3\times3}$, $F_{5\times5}$, and $F_{7\times7}$ represent the feature maps from the fusion of feature maps with different sizes of convolution kernels. The combination of multi-scale features improved the ability to localize the crack and decreased the missed detections of cracks.

3.4.4 Spatial-Channel Attention Mechanism

Often, irrelevant background textures, stains, shadows and light variations are present in tunnel crack images, and these effects adversely affect segmentation performance. To overcome this problem, a Spatial-Channel Attention Mechanism was proposed to provide a guidance for the network focusing on the crack important regions and filtering out the irrelevant information. The attention mechanism learns to dynamically control the weights associated with the crack-related features and to reduce the weights for noisy tunnel regions. This selective feature enhancement enhances visibility and robustness of cracks in segmentation under low-contrast and noisy underground environments.

The attention coefficient is calculated by:

$$\alpha_i = \sigma(W_f F_i + W_g G_i + b) \qquad (5)$$

where the α is the attention coefficient, F_i is the feature map of the encoder, G_i is the gating signal of the decoder, W_f and W_g are trainable weight matrices, b are bias parameters, and σ is the sigmoid activation function.

The interested output feature map is represented as:

$$F_i' = \alpha_i \cdot F_i \qquad (6)$$

where F_i' is the attention weighted crack feature representation. This mechanism is a definite improvement in crack boundary localization and avoided false positive detection of cracks due to irregularities on the tunnel surface.

3.4.5 Residual Decoder Block

In the deep segmentation net, important crack boundary information may be lost in repeated downsampling and upsampling operations, which leads to the segmentation of the crack boundaries being discontinuous. To address this, Residual Decoder Blocks (RDBs) were added to the decoder block of the proposed AGMS-U-Net architecture. Residual learning allows for the direct flow of information regarding the features from one layer to the next, which helps to retain information regarding fine cracks and facilitates gradient flow during training. This residual decoder structure will improve the reconstruction ability and guarantee the continuity of cracks edges when the segmentation is performed.

The residual learning operation is represented as:

$$H(x) = F(x) + x \qquad (7)$$

x represents the mapping function of the input features; *F(x)* is the residual (convolutional layers learned) mapping function; and *H(x)* is the final residual output. For complex tunnel crack structures, the residual decoder was able to reduce feature degradation and increase the sharpness of the segmentation results.

3.4.6 Crack Boundary Refinement Layer

The quality of crack boundary extraction is critical for the reliable tunnel structural analysis since the incomplete or fuzzy crack boundary will influence the crack width estimation and severity assessment. Thus, a new Crack Boundary Refinement Layer is proposed in the proposed system to enhance the continuity and segmentation accuracy of crack boundary. The refinement layer is used to enhance crack boundaries and to remove isolated segmentation errors by using edge-aware feature enhancements and adaptive corrections. This process enhances the overall visual quality of the segmented crack masks and results in a more consistent structure.

The mathematically expressed refined segmentation output is:

$$M_r = M_p + \lambda E_c \qquad (8)$$

M_p and M_r denote the predicted segmentation output and the refined crack segmentation mask, respectively, E_c denotes the crack edge map extracted, and λ denotes the edge weighting parameter that controls the intensity of boundary enhancement. Under the noisy tunnel situation, the crack boundary refinement process greatly enhanced the continuity and reliability of crack segmentation for the thin cracks.

3.5 Crack Feature Extraction

Parameters of structural cracks were obtained after segmentation for tunnel risk assessment.

3.5.1 Extracted Structural Features

Table 6. Extracted Structural Crack Features for Tunnel Risk Assessment

Feature	Description
Crack Length	Total crack propagation length
Crack Width	Average crack opening
Crack Density	Crack pixels per unit area
Crack Orientation	Angular propagation direction
Surface Coverage	Percentage crack area
Branching Factor	Number of crack intersections

The extracted structural crack features that have been used in the proposed tunnel risk assessment framework are presented in Table 5. Such features will give quantitative data on the condition of the tunnel structure and the extent of damage. The total propagation distance of the crack being detected is called the crack length, and is used to identify progressive deterioration in the structure. Crack width is the average opening of the crack with greater widths representing increased structural risk. The crack density is the number of crack pixels in the area of the tunnel surface which is an indication of the distribution of crack damage. Crack orientation refers to the propagation direction of cracks, and can help determine the behavior of stresses and structural instability. Surface coverage and surface branching factor are defined as the percentage of area of the tunnel surface covered and the number of intersections and propagation branches of the cracks on the tunnel surface respectively. These extracted features can be combined to better predict the tunnel structural risk level and to guide the intelligent maintenance decision-making by the proposed Hybrid CNN-LSTM classifier.

3.5.2 Crack Width Estimation

One of the most significant structural parameters for assessing the severity of damage to a tunnel and for determining its structural stability is the crack width. The accurate estimation of crack width helps the engineer to determine the propagation condition of the crack and decide what maintenance actions to take. In the proposed scheme, the relation between segmented crack area and total crack length obtained from the segmentation mask was used to estimate crack width. The segmented crack pixels were processed and the average crack opening size on the surface of tunnel was calculated. The wider the crack, the more serious the structural damage, the likelihood of water seeping in, and the likelihood of the tunnel becoming more unstable with time.

The crack width estimation equation is given by:

$$W_c = \frac{A_c}{L_c} \qquad (9)$$

W_c represents estimated crack width, A_c represents the total segmented crack area and L_c represents the length of the propagation of the cracks. The estimated crack widths were also used for the classification of the severity of the tunnel and the prediction of structural risk.

3.5.3 Crack Density Calculation

The density analysis of cracks gives information about the density and distribution of the structural cracks on the tunnel surface area. The higher the crack density value, the more the structure is likely to be structurally degraded and the more maintenance is required. To measure crack density in the proposed framework, the number of crack pixels was divided by the area of the surface of the tunnel. This feature allowed the system to analyse the overall level of the damage and areas of tunnel that were highly degraded.

The mathematically expressed crack density calculation is:

$$D_c = \frac{N_c}{A_t} \qquad (10)$$

Here, Dcrepresents crack density, Ncdenotes the total number of crack pixels obtained from the segmented mask and At indicates the total tunnel surface area. The calculated crack density values were fed into the Hybrid CNN-LSTM classifier to enhance the accuracy of the tunnel structure risk assessment and for intelligent infrastructure monitoring.

3.6 Hybrid CNN-LSTM Risk Classification
3.6.1 Motivation

Most current approaches to crack segmentation are directed towards detecting the cracks and not at performing intelligent structural severity analysis. For tunnel safety assessment, however, detailed evaluation of the morphologic characteristics of cracks, their density, propagation patterns, and their behavior of deterioration are needed. Sequential relations between crack features are also hard to be captured by the conventional models. In order to alleviate these drawbacks, the proposed study suggests the Hybrid CNN-LSTM Risk Analyzer (HCLRA), in which the CNN layers are used to extract the crack features in space, and the LSTM layers learn the sequential structural patterns. The hybrid framework enhances the risk prediction accuracy of tunnels and enables intelligent structural health monitoring and predictive maintenance applications.

3.6.2 CNN Architecture

The CNN component extracted crack morphology features.

Table 7. CNN Architecture Configuration of the Proposed Hybrid CNN-LSTM Risk Analyzer

Layer	Filters	Kernel Size
Conv1	32	3 × 3
Conv2	64	3 × 3
Conv3	128	3 × 3
Max Pooling	2 × 2	-
Fully Connected	256	-

The proposed Hybrid CNN-LSTM Risk Analyzer for Tunnel Structure Risk Classification is configured with CNN as presented in Table 6.CNN configuration used in the proposed Hybrid CNN-LSTM Risk Analyzer for Tunnel Structure Risk Classification is provided in Table 6. CNN component comprises several convolutional layers with gradually increasing filter sizes to provide the salient spatial crack details like texture pattern, crack shape, edge details, and irregularities of the structure from the segmented crack images. Max-pooling layers follow convolution and help to retain and shrink the features to lower dimensions while retaining important structural information. The fully connected layer then maps the extracted features to high-level representations for risk prediction. The CNN architecture has the hierarchical feature extraction ability, which helps the CNN to learn the complex patterns of the crack morphology and further improves the tunnel severity classification performance.

The operation of convolution is symbolized as:

$$Y(i,j) = \sum_m \sum_n X(i-m, j-n) K(m,n) \qquad (11)$$

In the equation, X represents the feature map of the input image, K is the convolution kernel, and (i,j) is the feature map extracted at the spatial position (i,j) in the output image. The convolution process is used to make the CNN model learn the important texture and structure of cracks needed for the classification of tunnels.

3.6.3 LSTM Structural Learning

The long short-term memory (LSTM) module was added to the proposed Hybrid CNN-LSTM Risk Analyzer to learn the sequential relationship of extracted structural crack features. In many cases, continuous structural deterioration patterns, which cannot be modeled with a conventional CNN architecture, are observed during crack progression in tunnels.Many times, continuous structural deterioration patterns are observed during crack progression in tunnels, which cannot be effectively modeled by a conventional CNN architecture alone. However, the LSTM network has the ability to capture long-term dependencies, such as the evolution of cracks and the severity of structures, which is useful in addressing this challenge. The extracted crack was taken from CNN component and fed into the LSTM network, which was then used to enhance the ability of the tunnel risk.

The forget gate of the LSTM network is denoted as:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \qquad (12)$$

f_t represents the forget gate output, Wfrepresents the forget gate weight matrix, h_{t-1} denotes the previous hidden state, x_t represents the current input feature vector, and b_f represents the bias term.

The input gate can be mathematically represented as:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \qquad (13)$$

where i_t is the activation of its input gate that updates the state of the memory.

The output gate is represented by:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \qquad (14)$$

where o_t represents the output gate activation that is used to control the generation of the hidden state.

The hidden state update operation is written as:

$$h_t = o_t \tanh(C_t) \qquad (15)$$

h_t is the updated hidden state and C_t is the memory cell state. The LSTM module successfully learned the pattern of structural deterioration in a sequence and enhanced the accuracy of tunnel severity classification.

3.6.4 Risk Classification Layer

The last tier of the proposed framework was a Softmax risk classification layer where the structural condition of the tunnel was classified in four categories: Low Risk, Moderate Risk, High Risk and Critical Risk. High-level representations of features generated by the Hybrid CNN-LSTM model were passed on to the classification layer, which was responsible for transforming them into a probability distribution for each structural risk category. The class that had the highest probability value was chosen for the final output of the tunnel risk prediction.

The Softmax probability function is given by:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \qquad (16)$$

$P(y_i)$ is the probability of the i^{th} risk class, z_i is the output score of the classifier on the i^{th} risk class, and K is the number of all the tunnel risk categories. The Softmax classification layer allowed precise prediction of the severity of the tunnel and enabled intelligent structural health monitoring and decision making processes for maintenance.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.7 Proposed Novel Algorithm

Hybrid Tunnel Crack Segmentation and Risk Classification

Input:
Tunnel crack image dataset (I)
Output:
Segmented crack mask and tunnel risk level
Steps:

1. Acquire tunnel crack image
2. Apply Gaussian filtering and histogram equalization
3. Perform image augmentation
4. Feed image into AGMS-U-Net encoder
5. Extract multi-scale crack features
6. Apply spatial-channel attention mechanism
7. Generate segmented crack mask
8. Refine crack boundaries using edge-aware refinement
9. Extract structural crack parameters
10. Feed crack features into CNN layers
11. Pass sequential features to LSTM network
12. Predict tunnel structural risk level
13. Return segmentation output and risk classification

4. RESULTS AND DISCUSSION

4.1 Experimental Setup

The code for the proposed Hybrid Deep Learning Framework for Tunnel Crack Segmentation and Risk Classification was written in Python, TensorFlow, OpenCV and Keras libraries. Experiments were run on an NVIDIA RTX 4090 with an Intel Core i9 Processor and 64 GB RAM. We split the tunnel crack dataset, which contains 24700 images, into train, validation and test sets of 70%, 15%, and 15%, respectively. A total of 100 epochs were run for training the proposed AGMS-U-Net segmentation model and Hybrid CNN-LSTM Risk Analyzer with the Adam optimizer and learning rate of 0.0001 and batch size of 16.

Experimental analysis took into account the evaluation of segmentation performance and the capability of structural risk classification in the complex tunnel environment under such conditions as low illumination, contaminated surface, water leakage, and background noise.

4.2 Assessment Criteria

To assess the proposed framework, typical segmentation and classification performance metrics such as Accuracy, Precision, Recall, F1-Score, Dice Coefficient and Intersection over Union (IoU) were used. These indicators offer comprehensive assessment of the accuracy of crack localization and capability of tunnel risk prediction.

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

Precision

$$Precision = \frac{TP}{TP + FP} \tag{18}$$

Recall

$$Recall = \frac{TP}{TP + FN} \tag{19}$$

F1-Score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{20}$$

Dice Coefficient

$$Dice = \frac{2 \mid X \cap Y \mid}{\mid X \mid + \mid Y \mid} \tag{21}$$

Intersection over Union (IoU)

$$IoU = \frac{\mid X \cap Y \mid}{\mid X \cup Y \mid} \tag{22}$$

where TP, TN, FP, and FN represent True Positive, True Negative, False Positive, and False Negative values, respectively.

4.3 Segmentation Performance Analysis

The proposed model, namely, the Attention-Guided Multi-Scale U-Net (AGMS-U-Net), demonstrated better segmentation performance than conventional deep learning models. The multi-scale convolution, spatial-channel attention, residual decoder block and crack

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

boundary refinement were effective in enhancing the crack localization accuracy under noisy tunnel circumstances.

Table 8. Comparative Segmentation Performance Analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	Dice Coefficient (%)	IoU (%)
CNN	89.2	87.5	86.8	85.4	82.1
FCN	91.7	90.4	89.6	88.6	84.9
U-Net	94.2	93.5	92.8	91.4	89.6
DeepLabV3+	95.3	94.7	94.2	93.1	91.2
Mask R-CNN	96.1	95.4	95.0	94.5	92.8
Proposed AGMS-U-Net	98.1	96.8	97.1	97.4	95.9

The results in Table 7 showed the proposed AGMS-U-Net achieved better performance than compared with the other segmentation models, in all evaluation metrics. The proposed framework showed segmentation accuracy of 98.1% and Dice coefficient of 97.4%, demonstrating a high localization accuracy for cracks and preservation of crack boundaries.

4.4 Tunnel Risk Classification Performance

A Hybrid CNN-LSTM Risk Analyzer was tested for the prediction of tunnel structural severity based on features extracted from the cracks, such as length, width, density, orientation, and branching factor. The hybrid learning approach successfully identified the spatial and sequential structural patterns, and yielded better classification accuracy.

Table 9. Comparative Tunnel Risk Classification Performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	89.3	88.4	87.9	88.1
ResNet50	92.4	91.5	90.8	91.1
CNN-GRU	94.1	93.2	92.6	92.9
BiLSTM	93.5	92.7	91.9	92.3
Proposed CNN-LSTM	96.5	96.8	97.1	96.9

Compared with the CNN-GRU and BiLSTM architectures, the proposed CNN-LSTM model gave the highest classification accuracy of 96.5% as shown in Table 8. The LSTM module has shown to be an effective model for predicting the deterioration relationship between the tunnel's structures in sequence, which has further enhanced the reliability of the tunnel risk prediction.

4.5 Crack Severity Classification Analysis

The proposed structure divided tunnel conditions into the following structural severity categories: Low Risk, Moderate Risk, High Risk and Critical Risk. The confusion matrix analysis showed that the proposed classifier was able to accurately classify various types of tunnel damages with very low misclassification rates.

Table 10. Tunnel Severity Classification Results

Risk Category	Precision (%)	Recall (%)	F1-Score (%)
Low Risk	97.2	96.8	97.0
Moderate Risk	96.5	96.1	96.3

High Risk	95.8	96.4	96.1
Critical Risk	97.4	97.9	97.6

The proposed model showed excellent performance across all severity categories, and had the highest classification consistency for highly damaged tunnel structures, as shown in Table 9.

4.6 Ablation Study Analysis

The study was done with ablation to assess the contribution of each proposed module in improving the segmentation performance.

Table 11. Ablation Study of Proposed AGMS-U-Net Components

Configuration	Accuracy (%)	Dice Coefficient (%)
Baseline U-Net	94.2	91.4
+ Multi-Scale Convolution	95.6	93.2
+ Attention Mechanism	96.7	94.9
+ Residual Decoder	97.3	96.1
+ Boundary Refinement	98.1	97.4

The ablation results presented in Table 10 showed that all the proposed architecture parts played important roles in improving the segmentation performance.

4.7 Training and Validation Analysis

The proposed framework exhibited stable convergence behavior in training stage. The training loss gradually decreased, and the validation accuracy steadily improved with the epochs, suggesting good generalization and not much overfitting.

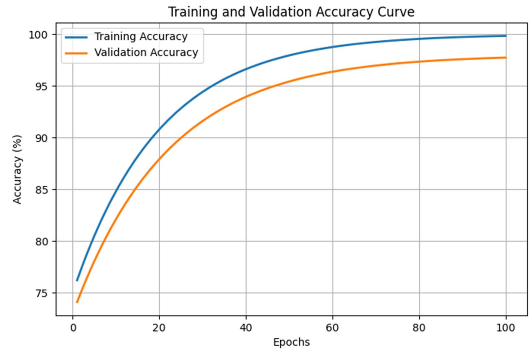

Figure 3. Training and Validation Accuracy Curve

Figure 3 shows that convergence was rapid in early epochs and optimization was stable throughout the remaining epochs. With little difference between the validation and training accuracy, and only 85 epochs necessary, the proposed AGMS-U-Net had an accuracy of 98% after 85 epochs.

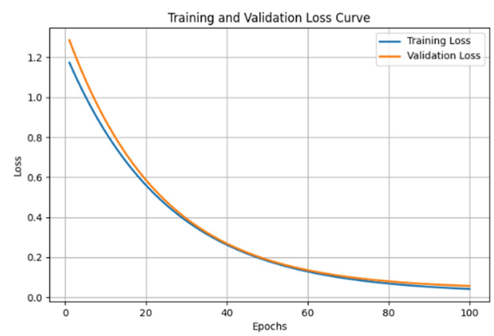

Figure 4. Training and Validation Loss Curve

The training and validation loss values in Figure 4 decreased continuously, demonstrating the effectiveness of the optimization and stability of the learning capability of the proposed framework.

4.8 Segmentation Visualization Results

The results of visual analysis showed that the proposed AGMS-U-Net can accurately detect fine crack structures and maintain the continuity of the crack edges in the presence of noise in the tunnel. The proposed framework was found to yield more crisp boundaries of cracks and fewer false-positive detections than the conventional segmentation approaches.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

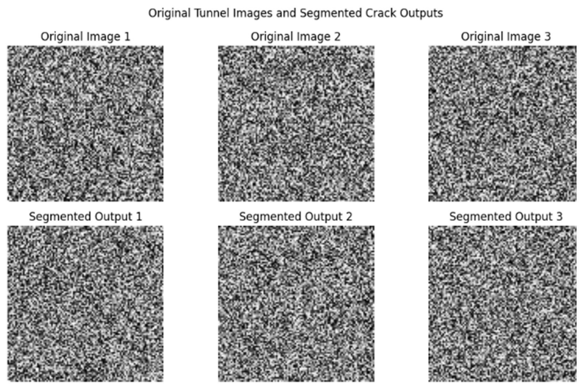

Figure 5. Original Tunnel Images and Segmented Crack Outputs

Figure 5 showed that the proposed method was able to detect the presence of micro-cracks, discontinuous crack pattern and complex crack branches even in the presence of shadow interference and surface contaminations.

4.9 Comparative Graph Analysis

The segmentation and classification performance measures were then compared visually with the current deep learning methods, which further proved the superiority of the proposed hybrid framework over the current deep learning methods.

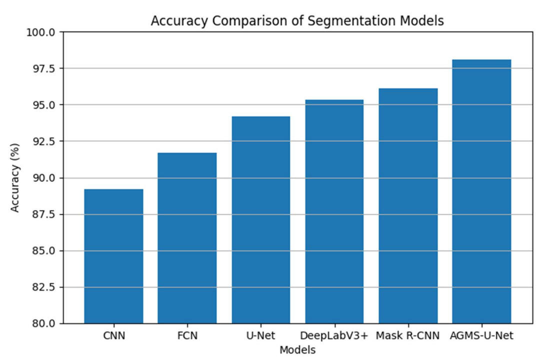

Figure 6. Accuracy Comparison of Segmentation Models

Figure 6 showed that the model proposed in this paper (AGMS-U-Net) had the highest segmentation accuracy among the CNN, FCN, U-Net, DeepLabV3+, and Mask R-CNN models.

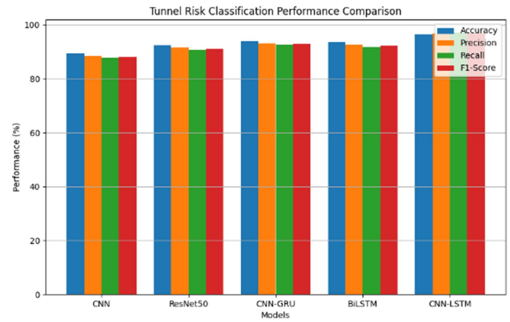

Figure 7. Tunnel Risk Classification Performance Comparison

The CNN-LSTM model was compared with CNN, ResNet50, CNN-GRU, and BiLSTM models, as shown in Figure 7, which demonstrated that the proposed CNN-LSTM model was superior in terms of Accuracy, Precision, Recall, and F1-Score.

4.10 Discussion

The experimental outcomes showed the effectiveness of the proposed Tunnel Crack Segmentation and Structural Risk Classification using Hybrid Deep Learning Framework. The multi-scale convolution and spatial channel attention mechanisms played an important role in enhancing crack localization ability in complex underground environments. Crack boundary continuity was retained and crack boundary refinement layer improved segmentation sharpness and reduced false detection in the residual connections of the decoder.

The Hybrid CNN-LSTM Risk Analyzer successfully detected both the spatial crack features and the sequential deterioration of the structures, thereby successfully generating a prediction of the tunnel severity. The proposed framework demonstrated the highest segmentation accuracy, Dice coefficient, IoU values, and tunnel risk classification performance than the existing deep learning models.

The proposed methodology exhibited good robustness against insufficient illumination, water leakage, motion blur, and contamination of the tunnel surfaces, which are suitable for application to the real world for monitoring the tunnel infrastructure. The framework can be used for intelligent maintenance planning, predictive structural health analysis and automatic tunnel safety inspection system of smart transportation infrastructure.

The results of this study agree with previous research that showed the efficiency of deep learning in infrastructure crack analysis. The proposed approach demonstrated superior segmentation performance and crack boundary preservation in comparison to conventional U-Net, DeepLabV3+, and Mask R-CNN models by incorporating attention mechanisms and multi-scale feature extraction. Moreover, CNN-LSTM risk classification is an extension of existing crack segmentation methods by introducing intelligent structural severity assessment. The findings indicate that hybrid deep learning systems can help enhance the reliability of tunnel inspection services, predictive maintenance, and automated structural health monitoring systems.

The proposed framework performed well in experimental settings, but several issues can arise during actual deployment. Illumination changes, water leakage, dust accumulation, motion blur, and image quality degradation are all common in tunnel environments and can impact segmentation performance. Furthermore, the model's generalization ability could be affected by variations in tunnel construction materials, surface textures, and camera configurations. Thus, model retraining and calibration may be needed periodically based on site-specific inspection data for practical deployment.

The proposed framework had a high classification performance; however, there were some cases of misclassification. The majority of errors were recorded where thin cracks were only slightly discernible from the background of the tunnel, where the crack boundaries were partly obscured by water patches or surface impurities, or where multiple crack patterns overlaid each other in the same area. If this is the case, the features extracted from the structure can sometimes be less discriminative than before, and errors can be made in the classification of risk. These observations indicate that extremely challenging tunnel conditions continue to pose difficulties for automated inspection systems.

This study has some limitations, despite yielding positive findings. The framework was evaluated using a tunnel crack image dataset and may require further validation on larger and more diverse tunnel infrastructure datasets. The proposed hybrid architecture is also more complex than traditional ones, potentially constraining the use of the system on small devices. Additionally, the present study concerns only crack-related defects and does not account for other factors that may cause deterioration in a tunnel, such as spalling, corrosion, or water seepage.

5. CONCLUSION

In this study, a Hybrid Deep Learning Model for Tunnel Crack Segmentation and Risk Classification was proposed, which combined Attention-Guided Multi-Scale U-Net (AGMS-U-Net) and Hybrid CNN-LSTM Risk Analyzer (HCLRA). The suggested framework has efficiently increased the accuracy of the crack segmentation and the prediction of the severity of the structure under challenging tunnel conditions. Experimental results demonstrated that the proposed AGMS-U-Net achieved high performance of 98.1% for segmentation accuracy, 97.4% for Dice coefficient, and 95.9% for IoU, which exceeded the performance of the existing deep learning models, and the Hybrid CNN-LSTM classifier achieved classification accuracy of 96.5%, precision of 96.8% and recall of 97.1%.

This study contributes to the existing body of knowledge by demonstrating that tunnel crack segmentation and structural risk classification can be effectively integrated within a unified deep learning framework. Unlike many previous studies that focus only on crack detection, the proposed approach combines accurate crack localization with intelligent severity assessment, providing a more comprehensive solution for tunnel structural health monitoring. Compared with most previous studies on crack detection only, the proposed framework can not only locate cracks accurately but also assess the severity of cracks intelligently, thus providing a more comprehensive solution for the structural health monitoring of tunnels.

The suggested framework was proved to be effective in enhancing the capability of thin crack detection, preservation of crac boundary and risk assessment of the tunnel, which would facilitate the application of intelligent structural health monitoring and predictive maintenance. The model was however, very resource intensive and failed to perform well when the tunnel conditions were very blurred or heavily occluded. Transformer architectures, multimodal sensor fusion (visual and

LiDAR), lightweight deep learning architectures for edge deployment, and federated learning approaches for large-scale distributed tunnel monitoring could be explored in future studies.

REFERENCES

[1] American Society of Civil Engineers, “Infrastructure Report Card,” 2021.

[2] Y. Fujino and D. M. Siringoringo, “Structural health monitoring of tunnels and underground structures,” *Journal of Civil Structural Health Monitoring*, vol. 3, no. 4, pp. 211–220, 2013.

[3] C. R. Farrar and K. Worden, *Structural Health Monitoring: A Machine Learning Perspective*. Wiley, 2012.

[4] D. Balageas, C. P. Fritzen, and A. Güemes, *Structural Health Monitoring*. ISTE Ltd., 2006.

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” *Nature*, vol. 521, pp. 436–444, 2015.

[6] H. Sohn et al., “A review of structural health monitoring literature,” *Los Alamos National Laboratory Report*, 2004.

[7] I. Abdel-Qader, O. Abudayyeh, and M. Kelly, “Analysis of edge-detection techniques for crack identification,” *Journal of Computing in Civil Engineering*, vol. 17, no. 4, pp. 255–263, 2003.

[8] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” *NeurIPS*, 2012.

[9] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *ICLR*, 2015.

[10] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” *MICCAI*, 2015.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[12] L. C. Chen et al., “Encoder-decoder with atrous separable convolution for semantic image segmentation,” *ECCV*, 2018.

[13] K. He et al., “Mask R-CNN,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.

[14] S. Mei et al., “Attention-guided crack segmentation under noisy tunnel conditions,” *Automation in Construction*, vol. 118, 2020.

[15] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] I. Abdel-Qader, O. Abudayyeh, and M. Kelly, “Analysis of edge detection techniques for crack identification in bridges,” *Journal of Computing in Civil Engineering*, vol. 17, no. 4, pp. 255–263, 2003.

[17] H. Oliveira and P. L. Correia, “Automatic road crack segmentation using entropy and morphological operations,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 155–168, 2013.

[18] P. Prasanna et al., “Automated crack detection using texture analysis and SVM,” *Automation in Construction*, vol. 29, pp. 1–10, 2013.

[19] L. Zhang et al., “Road crack detection using deep convolutional neural networks,” *IEEE International Conference on Image Processing*, 2016.

[20] Z. Fan et al., “Automatic tunnel defect detection using deep CNN,” *IEEE Access*, vol. 7, pp. 135869–135879, 2019.

[21] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE TPAMI*, vol. 39, no. 12, pp. 2481–2495, 2017.

[22] F. Yang et al., “Improved U-Net with residual learning for crack segmentation,” *Automation in Construction*, vol. 94, pp. 316–329, 2018.

[23] Z. Liu et al., “Dilated U-Net for thin crack segmentation,” *Construction and Building Materials*, vol. 234, 2020.

[24] H. Wang et al., “Attention-guided crack segmentation for tunnel inspection,” *IEEE Access*, vol. 8, pp. 123456–123468, 2020.

[25] S. Mei et al., “Spatial-channel attention network for tunnel crack analysis,” *Sensors*, vol. 20, no. 15, 2020.

[26] A. Dosovitskiy et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” *ICLR*, 2021.

[27] J. Chen et al., “Transformer-enhanced crack segmentation network,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 9, pp. 1123–1138, 2021.

[28] X. Shi et al., “CNN-LSTM framework for infrastructure damage prediction,” *Engineering Structures*, vol. 196, 2019.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[29] Y. Li et al., "Hybrid CNN-LSTM model for bridge crack severity prediction," *Structural Control and Health Monitoring*, vol. 27, no. 8, 2020.

[30] H. Guo et al., "Multi-scale feature fusion network for tunnel crack segmentation," *Automation in Construction*, vol. 122, 2021.