Developments in the Built Environment 21 (2025) 100587

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.sciencedirect.com/journal/developments-in-the-built-environment

Multi-level optimisation of feature extraction networks for concrete surface crack detection

Faris Elghaish[a], Sandra Matarneh[b], Farzad Pour Rahimian[c], Essam Abdellatef[d], David Edwards[e,i], Obuks Ejohwomu[f,i], Mohammed Abdelmegid[g], Chansik Park[h,*]

[a] School of Natural and Built Environment, Queen's University Belfast, Northern Ireland, UK
[b] Department of Civil Engineering, Engineering, Al-Ahliyaa Amman University, Amman, Jordan
[c] Centre for Sustainable Engineering, Teesside University, Middlesbrough, UK
[d] Department of Electrical Engineering, Engineering, Sinai University, El-Arish, 45511, Egypt
[e] Department of the Built Environment, Birmingham City University, Birmingham, B4 7XG, UK
[f] School of Engineering, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[g] School of Civil Engineering, University of Leeds, Leeds, UK
[h] Department of Architectural Engineering, Chung-Ang University, Seoul, South Korea
[i] CIDB Centre of Excellence, University of Johannesburg, Johannesburg, 2092, South Africa

ARTICLE INFO

Keywords:
Deep learning
Concrete surface
Cracks
Xception
ResNet101
Multi-level optimisation

ABSTRACT

With the increasing utilisation of deep learning (DL) for detecting and classifying distress in concrete surfaces, the demand for accurate and precise models has risen. This study proposes a novel empirical approach of multilayer optimisation for two prominent DL models, namely ResNet101 and Xception, to classify distress in concrete surfaces. Both models were trained using 20,000 images depicting various types of cracks and tested with another set of 20,000 images. Four algorithms (Sequential Motion Optimisation (SMO), shuffled frog-leaping algorithm (SFLA), grey wolf optimisation (GWO), walrus optimisation (WO)) were then applied to enhance classification accuracy. After evaluating the DL models' overall performance, the four algorithms were grouped into two layers. The first layer comprised SMO, SFLA, GWO and their combined application. Subsequently, the second stage implemented the WO optimiser to enhance performance further. The outcomes demonstrated a substantial positive impact on the accuracy of both CNN models. Specifically, ResNet101 achieved 98.9% accuracy and Xception reached 99.2% accuracy. In the accuracy breakdown, ResNet101 achieved 97.6% accuracy and Xception achieved 98.3% accuracy in the first stage, compared to 87.4% for Xception and 83.1% for ResNet101 before optimisation. Given that this approach achieves over 99% accuracy in detecting cracks on concrete surfaces, it offers a significant improvement in the efficiency and cost-effectiveness of structural health surveys for large buildings. Furthermore, it provides structural engineers with precise data to accurately determine and implement the required maintenance actions.

1. Introduction

As the most frequently used construction material in the world, concrete is widely employed in buildings, bridges, highways, dams and other structures (Gu et al., 2016). Concrete buildings are affected by external conditions and stresses, resulting in local deterioration (Qu et al., 2021). Cracks are a common type of damage in concrete construction products (Zhang et al., 2023). Internal and external variables such as heat expansion, overload, restraint or chemical reactions can all cause concrete cracking (Larosche and Delatte, 2009). Leaving these minor defects to grow will impact the structure's durability and safety. Consequently, rapid and accurate health monitoring or detection is essential for the resilience of concrete structures. Manual inspection is the most often employed inspection and detection approach for structural surface cracks but is subject to the disadvantages of being labour-intensive and time-consuming, which can lead to inspector visual fatigue and detection errors (Liu et al., 2024a; Xiong et al., 2024). To improve structural surface crack detection and the limits of current inspection methods, both academic and industry practitioners have focused on automating the crack detection procedure with high

* Corresponding author.
E-mail address: cpark@cau.ac.kr (C. Park).

https://doi.org/10.1016/j.dibe.2024.100587
Received 9 September 2024; Received in revised form 29 November 2024; Accepted 1 December 2024
Available online 15 December 2024

accuracy and real-time capacity. Given the rapid advancement of deep learning (DL) and its concomitant capability to integrate nonlinear classification with automated feature extraction, numerous predictive models featuring deeper network architectures have been suggested for the purpose of detecting surface cracks in concrete (Li et al., 2022a). For example, Cha et al. (2017) employed a convolutional neural network (CNN) and sliding window methodologies to detect and localise cracks within concrete images. Xu et al. (2019b) proposed an enhanced Faster R-CNN for automatic diagnosis and localisation of seismic damage in reinforced concrete columns (such as concrete cracking, spalling, reinforcement exposure and reinforcement buckling), with an overall average accuracy of 80%. Jang et al. (2019) presented a novel approach for evaluating concrete cracks that integrates DL with a hybrid image-based method. The hybrid images (produced by combining vision and infrared cameras) were subsequently fed into a DL network to achieve recognition outcomes. Andrushia et al. (2022) and Andrushia et al. (2021) utilised a hybrid CNN-Long Short-Term Memory (LSTM) model and the Ripplet transform respectively, to detect postfire concrete surface cracks.

Nevertheless, DL models are often criticised for their substantial training expenses and lack the capability to provide incremental updates (Yang et al., 2021). To overcome these restrictions, the concept of transfer learning (TL) has emerged by leveraging knowledge from a pre-trained model's to learn on another set of data (Liu et al., 2024b). Yu et al. (2022c) developed a vision-based automated approach to identify the surface condition of concrete structures. This method incorporates TL, decision-level image fusion and cutting-edge pre-trained convolutional neural networks (CNNs). Han et al. (2022) proposed a novel image-based detection approach for surface cracks to minimise a potential safety accident. The method consisted of four steps *viz.*: super-pixel segmentation, detection of damage areas, pixel-level identification and fracture location. First, utilising simple linear iterative clustering, the unmanned ariel vehicle (UAV) images were pre-segmented into suitable and universal sizes. Then, separate crack segmentation data sets for detection and pixel-level identification were created, on which the YOLO (You Only Look Once) V3 and DeepLab V3+ models were trained. Finally, a panoramic crack location and presentation approach was proposed by combining images with UAV flight records. Zhang et al. (2023) proposed a hybrid probabilistic deep learning method based on TL for classifying concrete surface damage. By integrating Bayesian inference into a deep convolutional network, their approach quantifies uncertainty in crack detection, achieving an accuracy of 0.9909. This method enhances recognition accuracy, reaching 0.9249 in complex images, even under noisy conditions.

Although pre-trained DL models are widely integrated and studied in the field of structural health monitoring, they have hitherto (and continue to) evolve in different ways. For example, several existing studies presented a comparison of the performance of different existing pre-trained CNNs for crack detection in concrete images (BaniMustafa et al., 2023; Paramanandham et al., 2022; Yu et al., 2022b, 2022c). Other studies fine-tuned existing pre-trained CNNs to augment concrete crack detection accuracy (Çelik and König, 2022; Li et al., 2022b). Henceforth, scant research has focused on optimising the pre-trained CNNs using different techniques to boost crack detection accuracy (Asadi Shamsabadi et al., 2022b; Chang et al., 2022; Miao and Srimahachota, 2021).

Consequently, this study applies hybrid optimisation techniques to enhance the detection and classification capabilities of pre-trained models to identify concrete surface cracks. The concatenation technique provides a novel means of combining pre-trained CNNs based on the TL technique to build a highly accurate model when compared to other available methods in the literature.

Six algorithms, including AlexNet, VGG19, GoogleNet, ShuffleNet, ResNet101, and Xception, were tested to detect cracks on concrete surfaces based on their depth, size, and parameter count (in millions) as indicated by prior studies. After evaluating their performance, ResNet101 and Xception were chosen to conduct the proposed hybrid two-stage optimisation process. Models developed underwent training using 20,000 images depicting various types of cracks and were tested with another set of 20,000 images. Four algorithms (*viz.:* spider monkey optimisation (SMO), shuffled frog-leaping algorithm (SFLA), grey wolf optimisation (GWO), walrus optimisation (WO)) were then applied to enhance accuracy. After assessing the overall performance of ResNet101 and Xception, the four algorithms were grouped into two layers. The first layer comprised SMO, SFLA and GWO, and their combined application. Subsequently, the second stage implemented the WO optimiser to further enhance performance. These optimised networks were specifically designed to improve accuracy when dealing with large datasets comprising 40,000 images, achieving an accuracy of 98.9% for Optimised ResNet-101 and 92.2% for Optimised Xception.

After evaluating the impact of the proposed multi-level optimisation during different development phases, this study also presents a framework for multi-layer optimisation showing the method of combining different optimisers using a concatenated technique and then applying additional algorithms in a new layer to maximise performance. This flexible framework can be used with different CNN models and for different datasets.

2. Conceptual background

2.1. Convolutional neural networks (CNNs) for concrete surface crack detection: training from scratch

Contemporary research has attempted to overcome the limitations of conventional visual inspection and sensor-based methods (including temperature effects and uncertainties), through the implementation of computer vision-based approaches, with a particular emphasis on DL methods (Deng et al., 2020). For example, Yang et al. (2018) and Alipour et al. (2019) demonstrated that a fully convolutional neural network (FCNN) can accurately classify concrete surface images at the pixel level, even when they contain various types of cracks. FCNN is an innovative CNN that replaces fully connected layers (dense layers) with 1×1 convolutions to undertake the same function as dense layers. FCNN utilises both the upsampled conclusion of the final layer, which matches the size of the input image, as well as the information gathered from earlier layers to do a pixel-wise classification. Similarly, Ni et al. (2019) developed a fully autonomous 'machine vision' method to extract crack contours using neural feature fusion and pixel-level classification.

Yu et al. (2021) suggested, a DL model called YOLOv4-FPM as an extension of the YOLOv4 model to enable UAV real-time detection of concrete bridge cracks. Focal loss is utilised to optimise the loss function in YOLOv4-FPM, thereby enhancing precision and overcoming the difficulties posed by complex backgrounds. The experimental findings indicated that YOLOv4-FPM achieves a mean average precision (MAP) of 0.976, surpassing YOLOv4 by 0.064. Cha et al. (2018) developed a structural visual inspection method that utilises a Faster Region-based Convolutional Neural Network (Faster R-CNN) to detect different types of defects in near real-time. To achieve this, a database was created consisting of 2366 images that are labelled for five different types of damage. Types of damage include concrete crack, steel corrosion at two levels (medium and high), bolt corrosion and steel delamination. Subsequently, the architecture of the Faster R-CNN undergoes modifications and is then trained, validated and tested utilising this database. Analysis results indicate that the average precision (AP) ratings for the five damage kinds are 90.6%, 83.4%, 82.1%, 98.1% and 84.7% accordingly – while the mean AP is 87.8%.

Tan et al. (2021) implemented a CNN-based version of the YOLOv3 algorithm to identify defects in sewage pipes autonomously. The research focused on improving the network architecture, bounding box estimation, data augmentation techniques and loss function. Analysis results indicated that the suggested model outperformed Faster R-CNN and YOLOv3. It acquired a MAP value of 92%, surpassing previously attained accuracy levels, which focused on the automated identification of sewage pipe defects. Zhang et al. (2021) employed a combined approach of a one-dimensional convolutional neural network (1D-CNN) and a long-short term memory (LSTM) technique to identify cracks in concrete bridge decks. The developed model exhibits accuracy rates of 99.05%, 98.9% and 99.25% for the training, validation and testing data, respectively. Kang and Cha (2022) introduced STRNet, an architectural framework based on deep CNNs, designed for real-time pixel-level segmentation of concrete cracks in complicated scenarios. STRNet is capable of processing RGB images/videos with a testing input size of 1024×512. The framework consists of a novel encoder based on STR modules, a decoder with attention and coarse upsampling, a standard convolutional operator, a learnable swish nonlinear activation function and batch normalisation. The network is trained using 1203 images and augmented extensively through synthesis; it is then evaluated using 545 testing images of different resolutions (1280×720, 1024×512). The network achieves precision, recall, F1 score and mIoU (mean intersection over union) values of 91.7%, 92.7%, 92.2% and 92.6% respectively. Xu et al. (2022) introduced a method for segmenting cracks in concrete using convolution-deconvolution feature fusion using holistically stacked networks. The proposed network employs channel attention to filter the original features recovered by the VGG-16 network, to replicate a reliable crack response and prevent interference from invalid and inaccurate feature replies. Emergent findings reported from multiple tests demonstrate that the suggested network provides superior recognition capability and enhanced resilience by capturing linear topological structures.

Although CNNs are common and effective in classification, they require a sheer volume of labelled image data when training from scratch, which is time-consuming and costly, with many structural health monitoring authorities needing more technologies for collecting data (Elghaish et al., 2022). Furthermore, the task of designing the CNN architecture to attain optimal outcomes is extremely challenging, as numerous hyperparameters (such as the convolutional quantity, fully connected layers and pooling) have an impact on the CNNs' efficacy (Liu et al., 2022). To tackle these challenges, the principles of TL and fine-tuning prove to be advantageous.

2.2. *Pre-trained convolutional neural networks for concrete surface crack detection: transfer learning*

As indicated earlier, retraining a pre-trained CNN model using a new dataset is referred to as transfer learning (TL and fine-tuning (Gao and Mosalam, 2018). This involves utilising one of the pre-existing CNN models as a starting point and adapting it to the new dataset (Asadi Shamsabadi et al., 2022a). A pre-trained model's capacity for feature extraction and learning prediction rules is employed in the TL process (Piyathilaka et al., 2020). This approach facilitates the creation of a more cost-effective and computationally expedient classifier model when compared to training a network from scratch (Li et al., 2022b). Due to the unique feature, the concept of TL and using pre-trained weights has gained momentum in many domains, including asphalt pavement crack detection (Matarneh et al., 2024).

Zhang et al. (2020) proposed a single-stage detector based on YoloV3 that is both quicker and simpler to implement. To enhance YoloV3, the TL technique with fully pretrained weights and batch normalisation was implemented. The introduction of focal loss improved the precision of YoloV3 with a detection accuracy of up to 80%. Li et al. (2019) employed model-based TL as an approach to initialise the parameters of the FCNN during the training procedure. In this study, a database consisting of 2750 images of concrete structures was created. Each image has dimensions of 504×376 pixels and contains various types of damage, such as cracks, spalling, efflorescence and holes. These images were labelled manually to identify and categorise four types of damage. Subsequently, the architecture of the FCNN is altered, trained, validated and assessed utilising this database (Li et al., 2019). Emergent findings indicated that there is a pixel accuracy (PA) of 98.61%, a mean pixel accuracy (MPA) of 91.59%, a mean intersection over union (MIoU) of 84.53%, and a frequency weighted intersection over union (FWIoU) of 97.34% (Li et al., 2019).

Liu et al. (2019) presented a trained U-Net with high efficacy that could detect concrete cracks in raw images under different conditions (e.g. brightness, disturbed background, etc.). The model's precision, developed by training on a dataset of 57 photos, has a high level of accuracy and can reach 0.9 even in complicated cases. The VGG-16 network was fine-tuned by modifying its architecture, namely by lowering the number of completely connected layers from three to two. Additionally, the final classification layer for surface crack detection was adjusted to consist of a Softmax layer with seven tags. The suggested model's mean accuracy of 83.76% outperformed the detection results of AlexNet, GoogLeNet and ResNet. Qu et al. (2022) conducted another study in which a CNN and a transformer named CrackT-net were proposed for crack segmentation. The authors used richer features (RF) UNet++ and polarised self-attention to enhance feature representation capabilities. In addition, they replaced the last feature extraction layer by the transformer to enhance the model's performance. Results showed that the proposed method proved effective with F-score values of 0.856, 0.700 and 0.637 on three data sets.

Rao et al. (2021) introduced an automated vision-based detection technique that utilises CNN models and a non-overlapping window to identify crack/non-crack states of concrete building images. A collection of crack/non-crack concrete structures was created which consisted of 32,704 training patches, 2074 validation patches and 6032 test patches. The performance of this approach was assessed by employing 15 cutting-edge CNN models, including the number of parameters needed for training, the area under the curve, and the inference time. The proposed model achieved a detection accuracy of over 95% and a precision of over 87% for the majority of CNN models in identifying cracks. Hang et al. (2023) developed a DL semantic segmentation network with an attention mechanism to detect concrete cracks. To differentiate between pixels that represent cracks and pixels that do not, this study introduced a new model called AFFNet. This model utilises a ResNet101 backbone that has been pretrained on ImageNet. The model developed was assessed using the test dataset, achieving a mean intersection over union (MIoU) of 84.49%.

Asadi Shamsabadi et al. (Asadi Shamsabadi et al., 2022a) introduced a hybrid architecture that combines U-Net and a vision transformer (ViT) for concrete crack detection. A ViT structure was applied to the output feature maps of U-Net to address the inherent bias of CNNs, specifically the limited ability to extract contextual data. Nevertheless, the structure heavily depends on a substantial number of convolutional layers to extract features; an approach which is susceptible to information loss. However, results showed that with TL and the differentiable intersection over union (IoU) loss function, the encoder-decoder network equipped with ViT could achieve an enhanced real-world crack segmentation performance. Wang and Su (2022) proposed a conventional transformer-based network called SegCrack for detecting cracks in concrete. They employed the online hard example mining technique to enhance the network's performance. To decrease the computational cost of SegCrack, the sequence length is lowered by a reduction factor R. SegCrack demonstrated a precision of 96.66%, recall of 95.46%, F1 score of 96.05%, and mean intersection over union of 92.63% when evaluated on the test set. The crack images utilised for training exhibit a significantly greater number of background pixels compared to crack pixels, which could potentially lead to an imbalance in the categories. To tackle this problem, Wu et al. (2024) introduced a dual attention transformer network called PCTNet, which is designed specifically for segmenting concrete cracks at the pixel level. PCTNet combines the strengths of CNNs and transformer networks to accurately segment cracks by completing self-attention calculations on both local and global features. Furthermore, the PCTNet network incorporates TL by leveraging pretraining on the ADE20K dataset to enhance performance and expedite the convergence of the network. PCTNet demonstrated a significant decrease of up to 64% in computational expenditure in comparison to transformer networks, while outperforming both convolutional and transformer networks in performance. It achieved a precision of 95.89%, recall of 93.77%, F1-score of 94.8%, and mIoU of 90.53%.

2.3. Optimised convolutional neural networks - feature selection optimisers

Several studies introduced the use of various optimisers using different CNN architectures; their approaches vary depending on the CNN architecture, datasets and optimiser type. Gao and Mosalam (2018) introduced the concept of a Structural ImageNet and demonstrated its application in four initial baseline recognition tasks *viz.*: component type identification, spalling condition check, damage level evaluation and damage type determination. To avoid overfitting, the authors utilised TL with the visual geometry group network (VGGNet) to extract features using feature extraction and fine-tuning techniques, revealing various potentials of deep TL in image-based structural damage recognition. TL was utilised on the AlexNet model by Rajadurai and Kang (2021) after adjusting the weights of the architecture, modifying the classification layer to accommodate two output classes (no-cracks and cracks), and enhancing the image datasets by introducing random rotation angles. The AlexNet model was trained using the stochastic gradient descent with momentum optimiser after fine-tuning. Evaluation of the trained AlexNet model's performance involved the utilisation of precision, recall, accuracy and F1 measures. The training process yielded an accuracy of 99.9% and a loss of 0.1% with a learning rate of 0.0001 and 6 epochs. After being trained, the AlexNet model achieved a prediction accuracy of 99.9%.

Chen et al. (2019a) revealed that the utilisation of the Adam optimiser and batch normalisation (BN) approach resulted in rapid convergence of the classification model and the attainment of a peak accuracy of 99.71%. Dimensionality reduction was undertaken utilising the spider monkey optimisation (SMO) approach. A deep neural network (DNN) (based on SMO) was then fed with the reduced dataset and generated classification results with accuracies of 99.4% and 92%. Additionally, precision values of 99.5% and 92.7%, recall values of 99.5% and 92.8%, and F1 scores of 99.6% and 92.7% were reached. Notably, these results were attained with the shortest training time. Guernine and Kimour (2021) proposed an enhanced version of the GWO algorithm to effectively explore a designated space of possibly suitable CNN architectures while simultaneously optimising their hyperparameters. Samma et al. (2021) utilised a skilled two-layer optimiser to refine a pre-trained VGG19 model. They indicated that the comparative experiment conducted on the two-layer optimiser offers evidence that supports its higher performance when compared to well-established optimisers such as AOA, WGA, RLMPSO, PSO and CLPSO in the context of VGG-19 filter selection.

A method used to improve the ability of the trained model to make generalisations is the modified chicken swarm algorithm (ECSA). ECSA was adopted by Yu et al. (2022a) to optimise the meta-parameters of the deep CNN model. Experimental results of this study suggest that the ECSA algorithm outperforms the CSA and PSO algorithms in optimising the meta-parameters of deep CNNs. More precisely, ECSA exhibits superior precision in optimisation and a higher rate of convergence. Elhariri et al. (2020) investigated the impact of using multi-objective optimisation approaches using a hybrid filter-wrapper strategy for selecting features in the context of identifying fracture severity.

Gaur et al. (2023) employed pre-trained models, namely VGG16, AlexNet and Inception V3 architectures to extract characteristics from concrete images at three distinct scales. Subsequently, the characteristics are concatenated into a feature tensor. A total of 5760 pictures out of the 7200 were utilised for training purposes, while 1440 were kept aside for model validation. Emergent findings indicated that the classification groups of crazing, cracks and pop-outs exhibited the lowest precision rates throughout the testing phase, with rates of 95% and 98% respectively. Conversely, the remaining classification groups achieved precision rates of 100%. The proposed model demonstrates its effectiveness with an overall accuracy of 98.31% and an F1-score of 0.98. Amjad et al. (2023) created a comprehensible empirical prediction model for self-healing concrete using multivariate linear regression (MVLR) and other methodologies, such as DNN, extreme gradient boosting (XGBoost), and CNN. While the XGBoost model achieved the highest R2 value of 0.95 and had the fewest statistical mistakes, the MVLR model provided a more easily understandable approach.

While several studies have applied the optimisation of the CNN concept, most have investigated only one or two optimisers and reported their impact on the test CNN architecture. This study will evaluate four optimisers using two different pre-trained CNN architectures to determine the best optimiser for concrete crack detection.

3. Methodology and logic

The philosophical approach of positivism and deductive reasoning were adopted to empirically test models developed as per this deterministic research (Howden-Chapman et al., 2023; Owusu-Manu et al., 2022). Chen et al. (2019b) assert that the progression of a machine learning/DL model encompasses distinct stages, comprising problem selection, data collection, model development, validation, impact assessment and deployment. The datasets utilised in this study were sourced from (Özgenel and Sorguç, 2018) and consist of two subsets: negative (without cracks) and positive (with cracks) images, used for image classification. Each subset contains 20,000 images, resulting in a total of 40,000 images, all standardized to 227x227 pixels with RGB channels. These images were derived from 458 high-resolution source images, each originally sized at 4032x3024 pixels, collected from various METU Campus buildings. The high-resolution images demonstrate significant variability in surface finish and illumination conditions, adding diversity to the dataset. Notably, the images were captured approximately 1 m away from the surfaces, with the camera positioned directly facing the target. Despite variations in the concrete surfaces, including finishes such as exposed concrete, plastering, and paint, all images were taken on the same day under similar illumination conditions. This consistency in lighting helps minimise variability due to environmental factors while still reflecting the diversity of surface finishes, moreover, no data augmentation techniques, such as random rotation, flipping, or tilting, were applied (Özgenel and Sorguç, 2018). The development of the optimisation solution in this present study initiated with the exploration of search algorithms to identify effective models for detecting concrete surface cracks. Six CNN models were selected based on their depth, size and parameter count (in millions), informed by prior studies (Dorafshan et al., 2018) and according to the performance in detecting and classifying cracks from concrete surfaces, two models were selected, which are Xception and ResNet101.

Subsequently, an iterative optimisation process commenced, initially validating the effectiveness of these selected algorithms through empirical experiments. Following this validation, researchers introduced four established algorithms to assess their capabilities when integrated with both CNN models.

After individually evaluating the performance of these four algorithms, the study employed the concatenation technique (Noreen et al., 2020). This technique facilitated the amalgamation of features from different layers, establishing connections within the network to retain and reuse information from earlier layers to significantly enhance gradient flow and preserve essential features during the training process. Three optimisers were combined namely SMO, SFLA and GWO to increase the detection and classification parameters including accuracy and F-Score percentages.

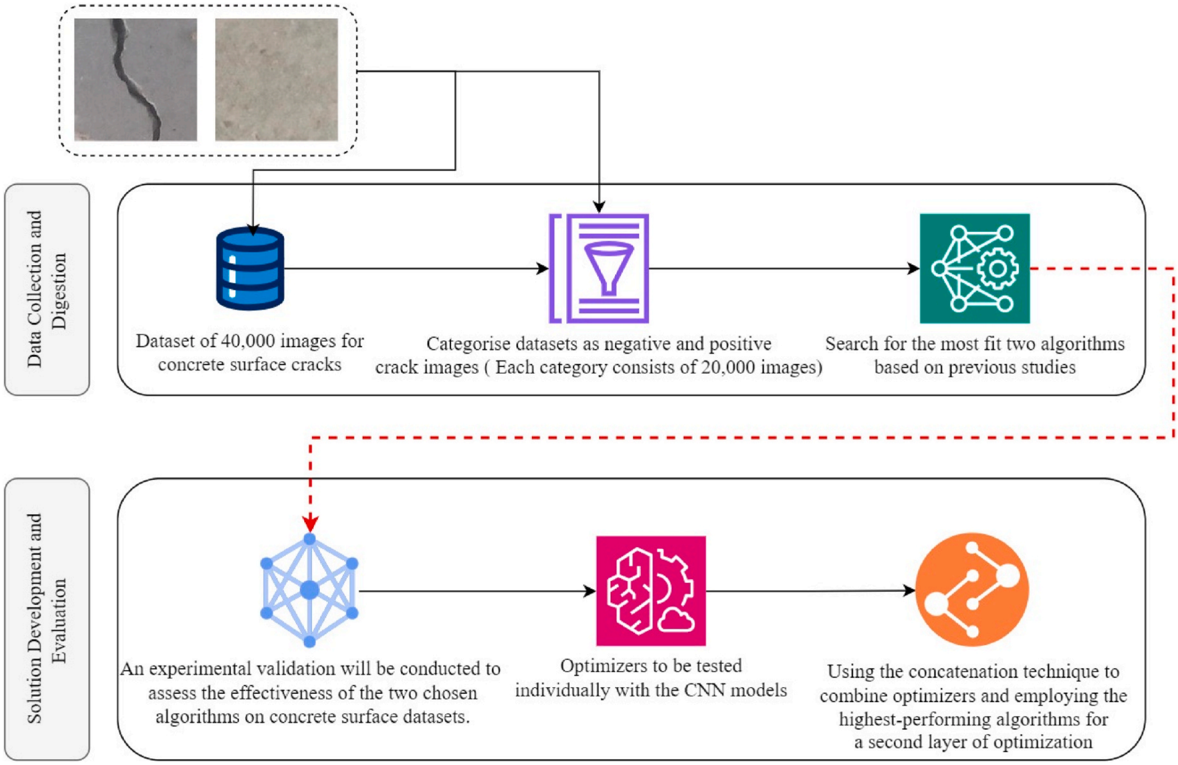

Fig. 1. Research methods and logic.

3.1. Rationale of selecting optimisation algorithms

SMO, SFLA, GWO, and WO were selected for this study based on their suitability for addressing the specific challenges posed by the dataset of images used for detecting cracks on concrete surfaces across all models. These algorithms were chosen after a thorough review of optimisation methods commonly used in similar applications, such as crack detection on concrete and pavement surfaces, with a focus on their ability to handle high-dimensional, noisy environments and achieve convergence precision.

Each algorithm brings distinctive strengths: SMO excels in balancing exploration and exploitation, which is critical for avoiding local optima (Jiang and Tsai, 2021); SFLA is robust in solving multimodal optimisation problems, making it suitable for complex datasets with diverse crack patterns (Ding et al., 2020); GWO effectively mimics social hierarchy and collective hunting behaviours (Kumaran et al., 2018), ensuring efficient search processes; and WO demonstrates strong performance in navigating nonlinear search spaces, which are often encountered in image-based crack detection tasks (Houssein et al., 2024). These attributes align closely with the requirements of image database optimisation tasks, such as feature selection and parameter tuning, where adaptability and precision are critical.

While these algorithms may not universally outperform others in all optimisation problems, they were selected over alternatives like PSO or GA due to their demonstrated efficiency and adaptability in addressing the specific complexities of this study's dataset. By leveraging these capabilities, the study achieved satisfactory results without resorting to more computationally expensive or highly specialised algorithms that might have provided only marginal improvements. This pragmatic approach balances practicality and computational efficiency, emphasising the contextual advantages of these methods within the scope of crack detection on concrete surfaces, rather than asserting their superiority in broader scenarios.

3.2. Evaluation metrics

Three performance metrics-accuracy, sensitivity and F-score—were employed to evaluate the models (Varoquaux and Colliot, 2023). According to ALJUHNI (ALJUHNI),Xu et al. (2019a) accuracy refers to the overall effectiveness of the model in correctly identifying the presence or absence of cracks. It is defined as the ratio of correctly classified instances, whether crack-positive or crack-negative, to the total number of instances. Mathematically, it is expressed as:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

Where: TP (True Positives) are cases where cracks are correctly identified, TN (True Negatives) are cases where non-cracked surfaces are accurately classified, FP (False Positives) are cases where non-cracked surfaces are incorrectly classified as cracked, and FN (False Negatives) are cases where cracks are present but not identified.

Sensitivity (also known as Recall or True Positive Rate) in crack detection measures the model's ability to correctly identify all cracked surfaces. It is crucial in ensuring that no cracks go undetected and is calculated as:

$$\text{Sensitivity (Recall)} = TP / (TP + FN)$$

This metric is especially important in safety-critical applications, where missing a crack could lead to severe structural failures.

The F-Score (or F1-Score) provides a balanced evaluation of the model's precision and recall in crack detection. It considers both the accuracy of predicting cracks (precision) and the ability to detect all cracks (recall). It is defined as:

$$\text{F-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

Where precision is given by:

$$\text{Precision} = TP / (TP + FP)$$

Subsequently, the optimal optimiser was selected to enhance the

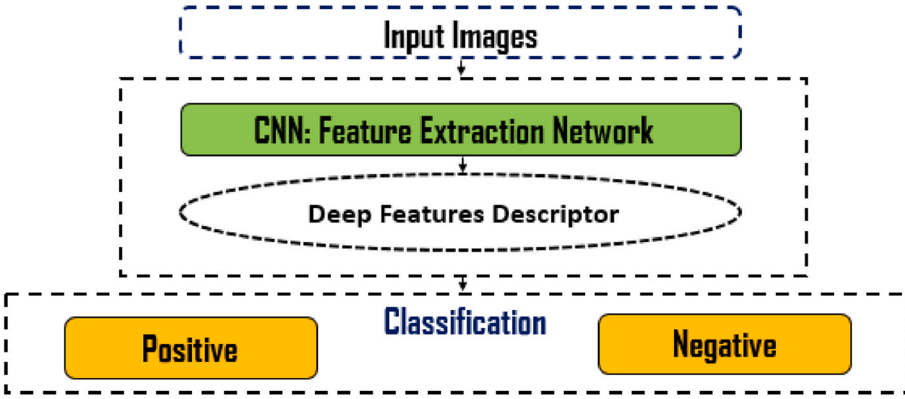

Fig. 2. Deep features descriptor.

Table 1
Comparison between ResNet-101 and Xception pre-trained CNN model characteristics.

CNN	Depth	Size	#Parameters (millions)
AlexNet	8	227 MB	61
VGG19	19	535 MB	144
GoogleNet	22	27 MB	7
ShuffleNet	50	5.4 MB	1.4
ResNet-101	101	167 MB	44.6
Xception	71	85 MB	22.9

overall performance of the optimised model from the initial stage. Fig. 1 depicts the selection process for models, optimisers and the development of the fully connected optimised DL model.

4. Results and analysis

4.1. Comparative performance analysis of unoptimised six CNN models

Fig. 2 portrays using six pre-trained CNN DL models, AlexNet, VGG19, GoogleNet, ShuffleNet, ResNet101 and Xception, for extracting features from concrete images to discern cracks as positive or negative dichotomous groupings. These models were chosen due to their extensive architecture, intricate information flow, computational efficiency and adaptability across diverse datasets, ensuring high performance in identifying cracks accurately.

Table 1 presents the architectural structure of the six chosen deep learning models, highlighting their characteristics, including depth, size and parameters. This design enables these networks to effectively capture intricate data features and patterns. Notably, both models exhibit a substantial parameter count, empowering them to learn complex representations from the data. The six mentioned models undergo initial testing without the use of optimisers for crack detection across a substantial dataset comprising 40,000 concrete crack images. Subsequently, a two-stage optimisation approach was employed. The first stage involves the individual and collective utilisation of four optimisers—SMO, SFLA, GWO and their combined application. Based on these outcomes, the second stage implements WO optimisers to enhance performance further.

Fig. 3 demonstrates that Xception achieved a higher F-score at 88.17% and accuracy 93.69% and the second most accurate model is ResNet101 at 84.58% F-score with accuracy 92.87%%. This underscores the potential of both models for optimisation towards greater accuracy and sensitivity. Additionally, Xception exhibited greater accuracy in identifying positive values (images depicting concrete surface cracks), surpassing ResNet101 by 3.59%. Considering the close performance of both models, the planned two-stage optimisation process as conducted.

Fig. 4 shows the confusion matrix of the six pre-trained CNN models based on the values mentioned in Table 1 and Fig. 3. It displays the four components of the confusion matrix: True Positive (TP), True Negative

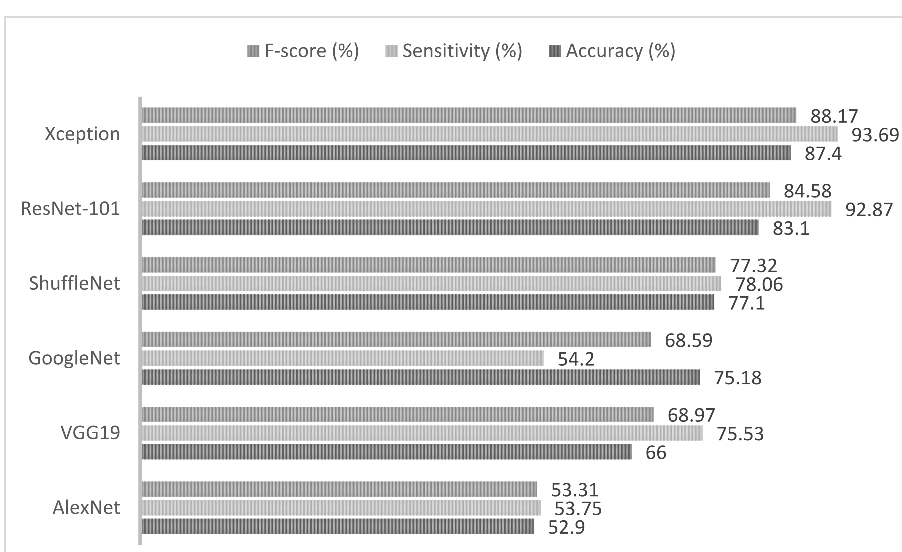

Fig. 3. Graphical comparison between six pre-trained CNN Models without applying de-noising CNN approach in terms of accuracy, sensitivity and F-score.

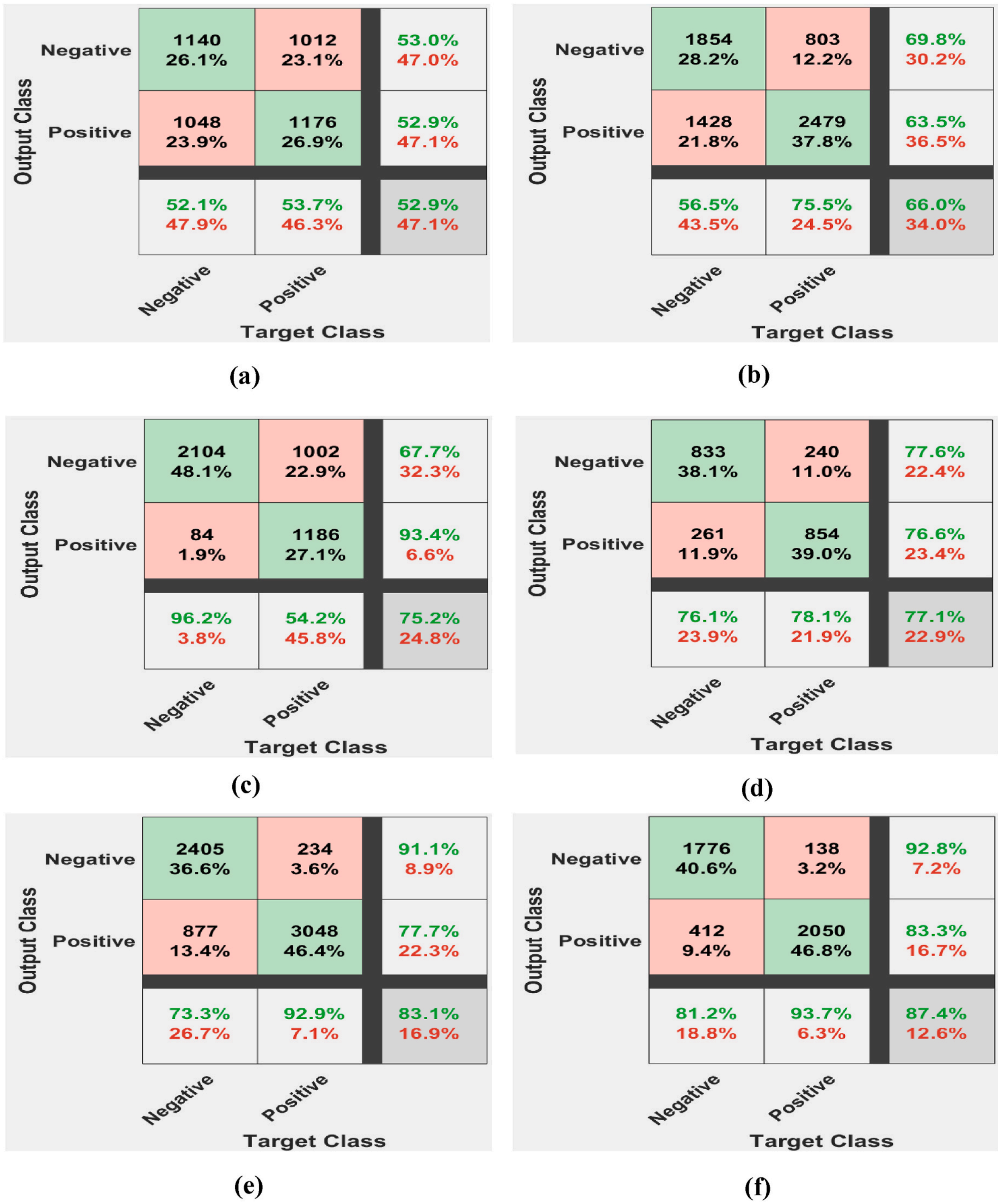

Fig. 4. Confusion matrices of various pre-trained CNN Models used for classification of concrete surface crack images: (a) AlexNet, (b) VGG19, (c) GoogleNet, (d) ShuffleNet, (e) ResNet-101, and (f) Xception.

(TN), False Positive (FP), and False Negative (FN). From chart (f) of the Xception model, it can be observed that TP is 40.6% and TN is 46.8%, with very small FP and FN values, both around 12.6%. Therefore, the accuracy of Xception is the highest among the six models, approximately 87.4%.

4.2. First optimisation layer

Four optimisers were utilised to enhance the performance of both CNN models – refer to Table 2.

Table 2 shows that all optimisers increased the F-score by at least

Table 2
Performance metrics for 1st optimisation layers.

CNN	Optimiser	Accuracy (%)	Sensitivity (%)	F-score (%)
ResNet-101	**SMO**	89.3	82.4	90.35
	SFLA	90.6	99.97	91.39
	GWO	89.4	82.5	90.7
	WO	92.5	95.38	92.71
Xception	**SMO**	93.7	98.69	94.01
	SFLA	96.2	92.99	96.37
	GWO	94	92.32	93.87
	WO	96.4	93.27	96.52

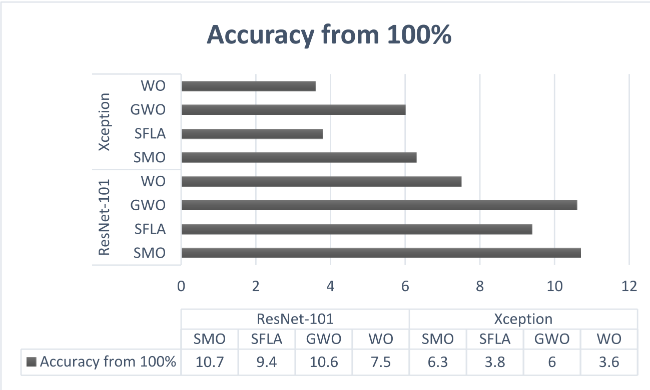

Fig. 5. Accuracy deviation from 100% for the first optimisation layer.

5.7% compared to the performance in Table 1 before optimisation. The exceptional performance of Xception paired with the WO optimiser is particularly noteworthy, achieving an impressive F-score of 96.52%. While all optimisers contributed significantly to performance enhancement, WO stands out for achieving the highest accuracy, recording 86.31% for ResNet and 96.4% for Xception. WO was implemented using the concatenation technique for the second optimisation stage, alongside three additional optimisers. This strategic integration will further refine the models and leverage the collective strengths of the optimisers to enhance the overall performance.

Fig. 5 shows the deviation of the two models with all optimizers from 100% according to F-score values from 100%. This is an important indicator to be used in future research when selecting optimal optimizers with one of these large pre-trained models.

Figs. 6 and 7 show the confusion matrices for both ResNet 101 and Xception with the four optimisers. These matrices display all indicators: TP, TN, FP, and FN for each optimiser, supporting the future selection of individual optimisers with CNN models. For example, chart (d) in Figs. 6 and 7 shows that the WO optimiser has the highest sum of TP and TN

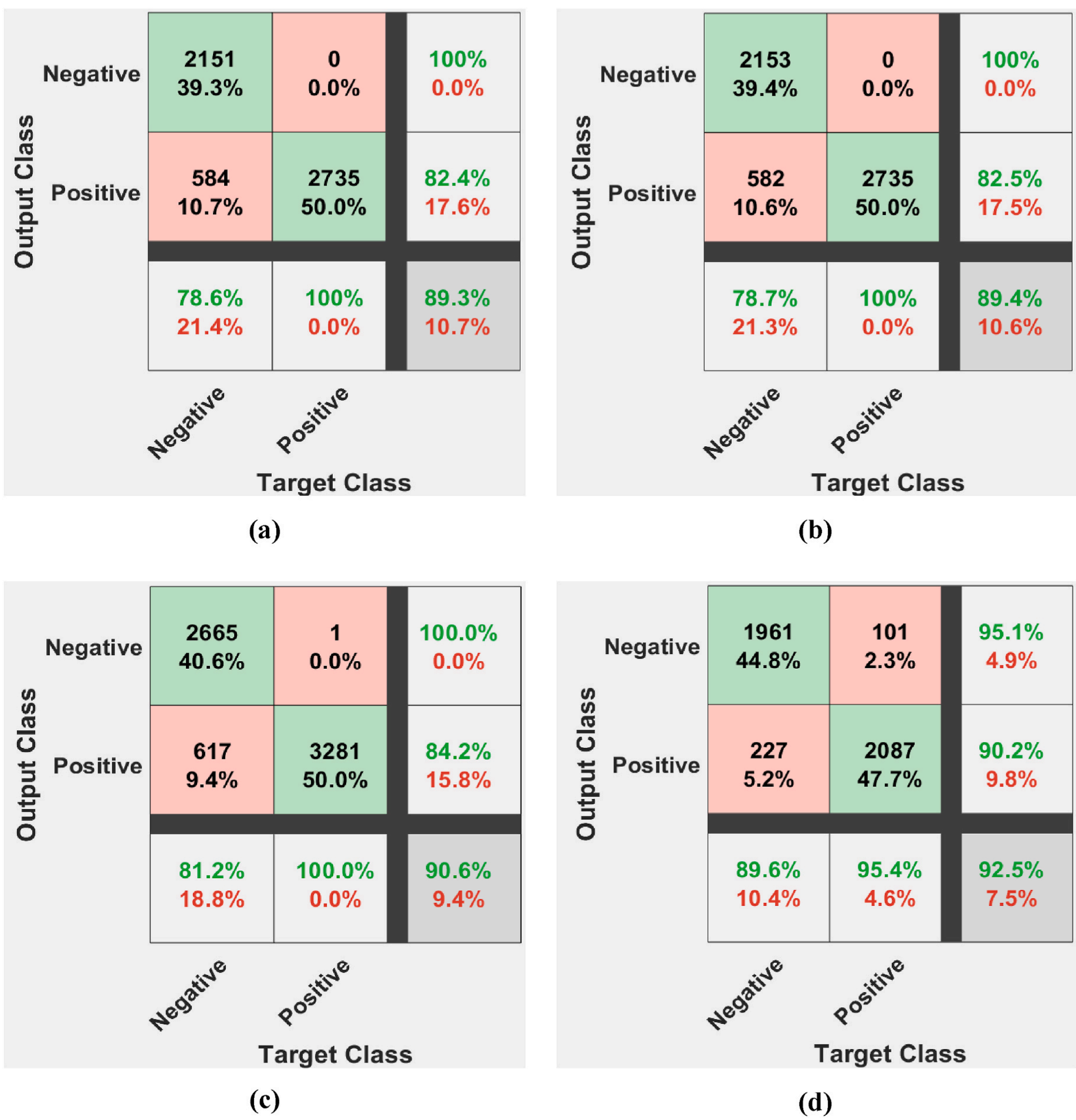

Fig. 6. Confusion matrices of concrete surface crack images classification using ResNet-101 CNN with different optimizers; (a) SMO, (b) GWO, (c) SFLA, and (d) WO.

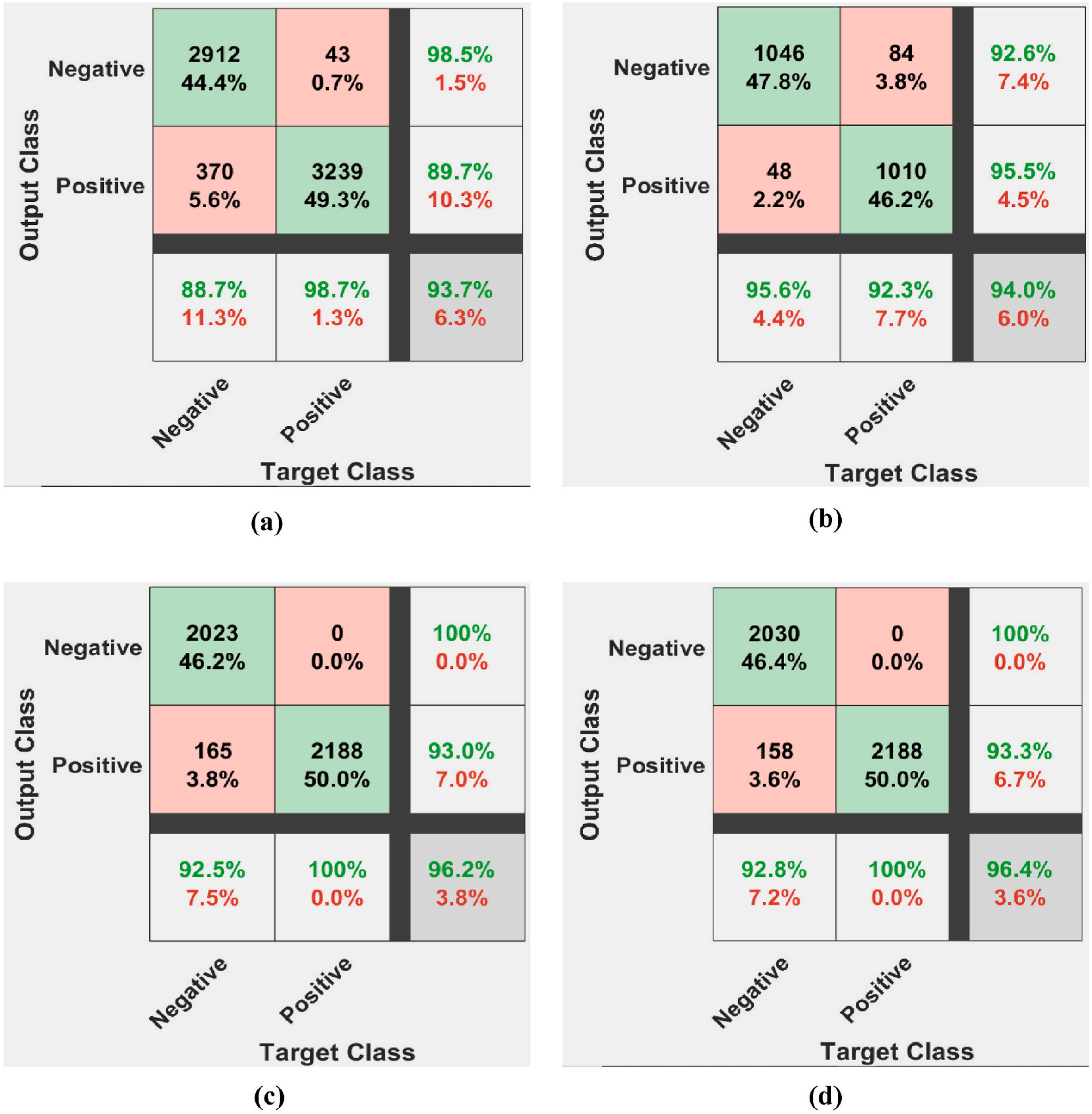

Fig. 7. Confusion matrices of concrete surface crack images classification using Xception CNN with different optimizers; (a) SMO, (b) GWO, (c) SFLA, and (d) WO.

Table 3
Performance of concrete surface crack images classification approach based on concatenation of optimised features using GWO, SMO, and SFLA optimisers (1st optimisation stage).

	Accuracy (%)	Sensitivity (%)	F-score (%)
ResNet-101	97.6	95.46	97.68
Xception	98.3	99.86	98.29

values, with 92.5% for ResNet 101 and 96.4% for Xception.

Given that the WO algorithm has achieved the highest accuracy for both CNN models, with 92.5% for ResNet-101 and 96.4% for Xception, it will not be selected for concentration optimisation. Instead, it will be considered for a second optimisation layer aimed at maximising accuracy further. After implementing the three optimisers (i.e. SMO, SFLA and GWO) simultaneously in a unified process, Table 3 illustrates notable improvements. ResNet101 achieved an accuracy of 97.6%, surpassing the best individual optimisation performance by at least 5.1%. Similarly, Xception exhibited a 1.9 % improvement, reaching an accuracy of 98.3%. These results strongly validate the efficacy of employing multiple optimisers collectively to enhance the performance of CNN feature extraction networks for concrete surface crack detection.

Fig. 8 presents the confusion matrix for the concatenated optimisation, detailing the performance of the three concatenated optimisers applied to ResNet-101 and Xception. The composite indicator in the figure represents the total of True Positive (TP) and True Negative (TN) values, which collectively measure the model's ability to correctly classify both cracked and non-cracked images. This indicator provides a comprehensive assessment of the model's overall classification accuracy, combining its effectiveness in detecting positive cases (images with cracks) and negative cases (images without cracks).

For Xception, the composite indicator is approximately 98.3%, highlighting the model's superior performance in accurately classifying the dataset. This high value underscores the robustness of Xception in managing both positive and negative classifications effectively during the first stage of optimisation, demonstrating its strong suitability for crack detection on concrete surfaces.

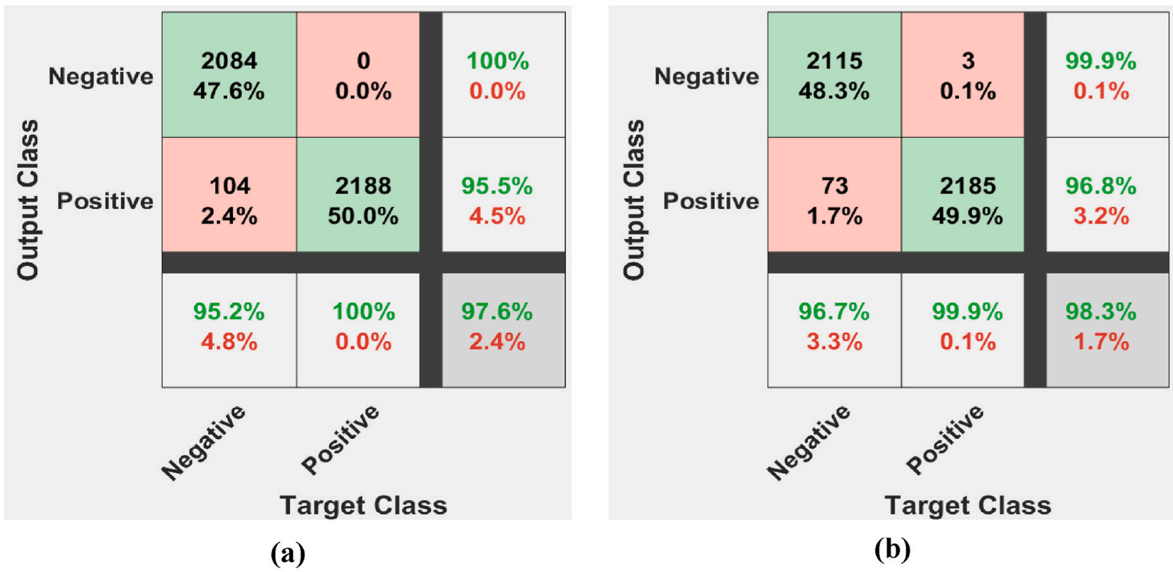

Fig. 8. Confusion matrices of the concrete surface crack images classification approach after the 1st optimisation stage using; (a) ResNet-101 CNN, and (b) Xception CNN.

Table 4
Performance of ResNet-101 and Xception CNN Models after the 2nd optimisation stage.

	Accuracy (%)	Sensitivity (%)	F-score (%)
ResNet-101	98.9	99.54	98.89
Xception	99.2	99.5	99.24

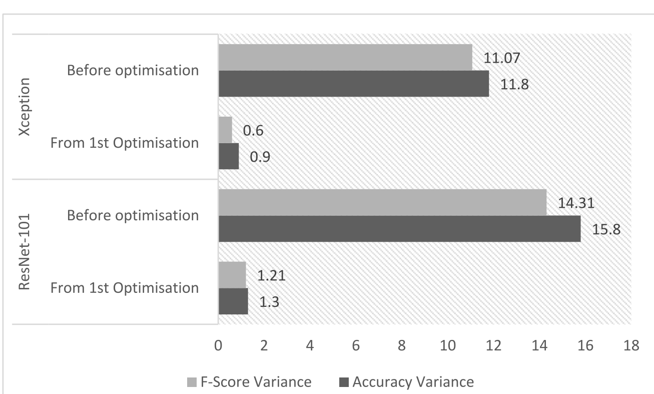

Fig. 9. F-score and Accuracy deviations from 100% between 1st, 2nd and before optimisation stages of optimisation.

4.3. Second optimisation layer

After implementing the WO optimiser, it is evident from Table 4 and Fig. 9 that there has been a significant improvement in all performance parameters. Using the F-score and accuracy as comprehensive indicators, Xception provided an enhancement of 0.6% for F-score and 0.9% for accuracy, while ResNet101 improved by 1.21% for F-score and 1.3% for accuracy. The Xception model achieved an accuracy of 99.2% and an F-score of 99.24%. It can be observed that the second optimisation stage significantly improved both models' performance, as their F-scores are within 1.11%–0.76% from 100%. For detailed performance and validation purposes, Fig. 10 shows the confusion matrix of the results from stage 2.

Fig. 11 provides a comprehensive overview of the optimisation process, illustrating the concatenation technique which technique effectively merges feature maps from multiple layers using three distinct optimisers. These interconnections played a crucial role for both ResNet101 and Xception networks, enabling the retention and utilisation of information from earlier layers. Such mechanisms were instrumental in improving gradient flow and preserving crucial features throughout the training process.

The application of the WO optimiser to both CNN models significantly enhanced performance, resulting in a 0.9% increase in Xception's and a 1.3% increase for ResNet101's accuracy in detecting concrete surface cracks (refer to Fig. 5).

Figs. 12 and 13 display the newly proposed architecture of the multi-layer optimised ResNet 101 and Xception feature extraction networks. The first section illustrates the traditional structure of the convolutional layers, while the second part demonstrates the incorporation of two additional optimisation layers.

4.4. Evaluation and discussion

A comparison between the performance of both CNN models before and after optimisation for detecting concrete surface cracks is conducted. It was observed that the accuracy for both ResNet101 and Xception improved by 15.8% and 11.8% respectively. The sensitivity increased by 6.67% for ResNet101 and 5.81% for Xception. Regarding the F-score, which represents the harmonic average between accuracy and precision, it demonstrates an improvement of 14.31% ResNet101 and 11.07% for Xception – refer to Table 5 and Fig. 14.

The comprehensive results in Fig. 14 present three performance metrics for both models across all development stages, encompassing the capabilities of pre-trained models in detecting concrete surface cracks before optimisation. Initially, it is evident that Xception exhibited superior performance before optimisation, with an F-score of 88.17%, surpassing ResNet101 at 84.58%.

The results demonstrate the incremental improvements achieved under each optimiser, particularly highlighting the substantial enhancements resulting from the WO optimiser for both models. Moreover, the proposed framework, comprising a two-layer optimisation strategy involving the concatenation approach of combining three optimisers (SMO, SFLA and GWO), followed by the application of WO, showcases notable performance lifts.

This study highlights the use of hybrid optimisation techniques to

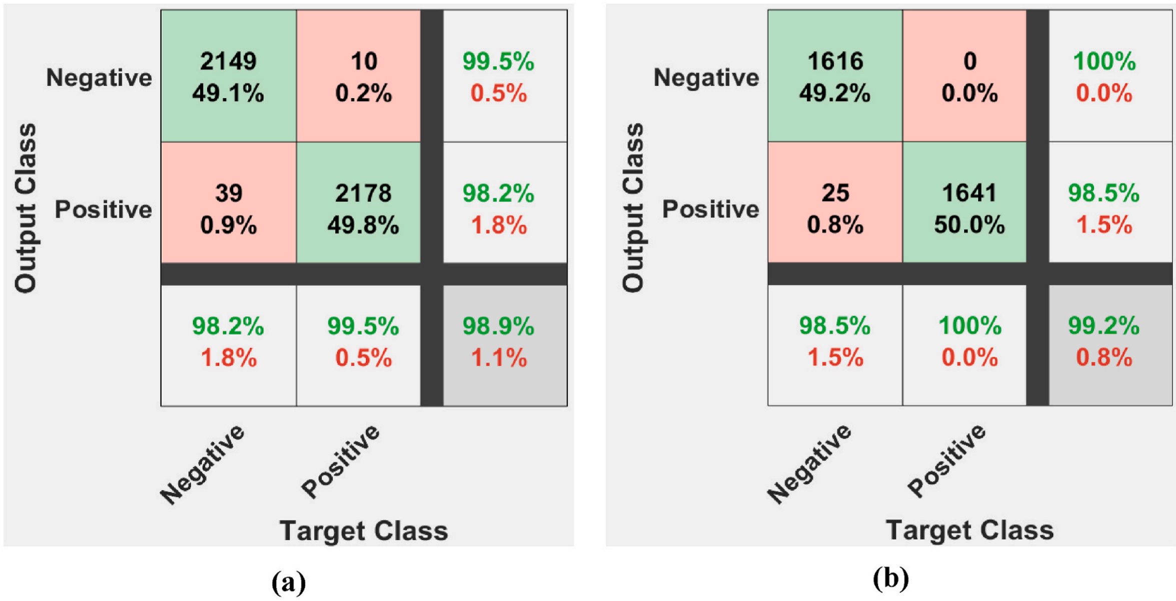

Fig. 10. Confusion matrices of the proposed concrete surface crack images classification approach based on two optimisation stages using; (a) ResNet-101 CNN, and (b) Xception CNN.

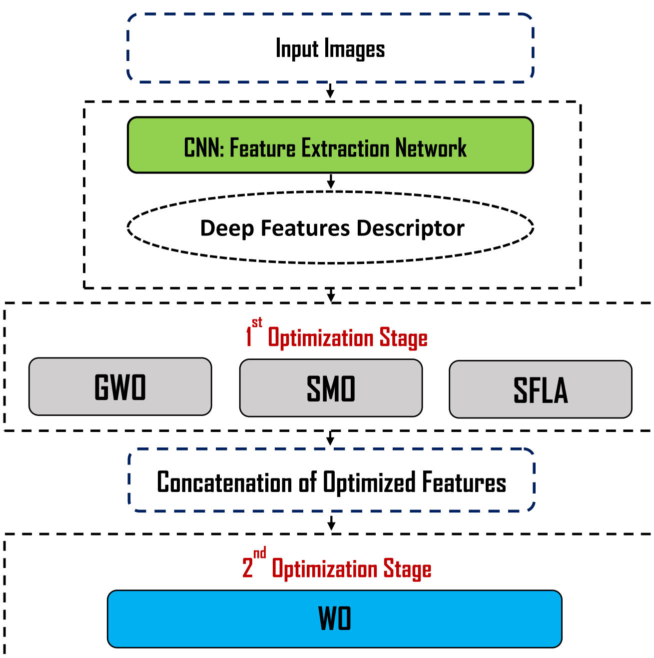

Fig. 11. High-level architecture of the proposed optimisation solution.

enhance pre-trained models' detection and classification capabilities in identifying concrete surface cracks. Six deep learning models were selected based on their depth, size, parameters, and proven effectiveness in previous studies for detecting similar image features. The two best-performing CNN models were then chosen for optimisation. As tested by Akgül (2023), ResNet101 exhibited an accuracy range of 73.54%–83% across different datasets in crack detection from concrete surfaces without optimisation. Moreover, Piyathilaka et al. (2020) evaluated ResNet101 among other pre-trained models for real-time crack detection, confirming its suitability for real-time applications. Xception showcased remarkable performance in structural crack detection, as evidenced by Philip et al. (2023),Sun et al. (2021), achieving accuracies ranging between 79% and 94.5%.

Considering these promising capabilities, both ResNet101 and Xception underwent optimisation on a substantial dataset containing 40,000 images, aimed at improving their crack detection performance across concrete surfaces. The optimisation process comprised two layers. In the first layer, four algorithms (SMO, SFLA, GWO and WO) were individually tested to enhance the models' performance. This resulted in an enhancement of both models' accuracy percentages, ranging from 6.2% to 9.4%. Further refining the optimisation, three algorithms (SMO, SFLA, GWO) were grouped using the concatenation technique. This collective integration led to a minimum 14.5% improvement for ResNet101 and a 10.9% enhancement for Xception. Moving to the second layer of optimisation, Xception saw a 0.9% increase in accuracy percentage, while ResNet101 observed a 1.3% increase. These refinements illustrate the iterative enhancement process, contributing to improved performance for both models in detecting concrete surface cracks.

Fig. 8 illustrates the combined metric percentage, aggregating the three performance metrics—F-score, sensitivity, and accuracy—across different optimisers. This serves as a valuable reference for future research endeavours when selecting optimisers suitable for similar datasets. The impact of optimizers varies between ResNet-101 and Xception, indicating that their effectiveness is model-dependent. While accuracy was the primary focus during the optimisation process in this research, this chart offers a more comprehensive guide for future studies. Notably, optimizers do not have a uniform effect on all CNN models; for instance, SMO has minimal impact on ResNet-101 but shows a significant improvement in Xception, as demonstrated in Fig. 15.

The proposed optimised versions of the two CNN networks facilitate the evaluation of concrete building health conditions, significantly reducing the time and cost required for manual crack inspections. These optimised networks were specifically designed to enhance accuracy when handling large datasets comprising 40,000 images, achieving an accuracy of 98.9% for Optimised ResNet-101 and 99.2% for Optimised Xception.

The proposed solution (and palpable 'accuracy' enhancements demonstrated) could accelerate the adoption of DL for assessing the structural health of buildings. Multi-level optimisation has improved the detection and classification parameters of CNN models, thereby enhancing the reliability and scalability of integrating this technology. The proposed framework and high-level architecture are applicable to

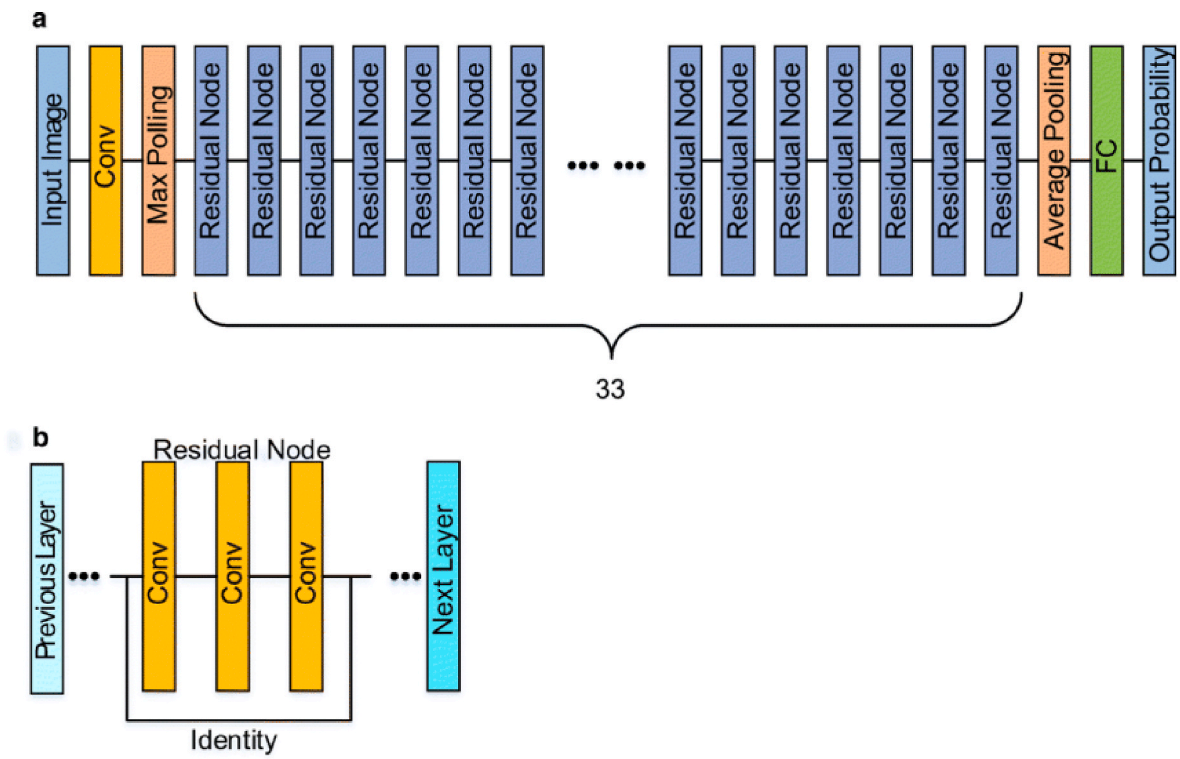

Fig. 12. The comprehensive architecture of ResNet101-based hyper optimisation.

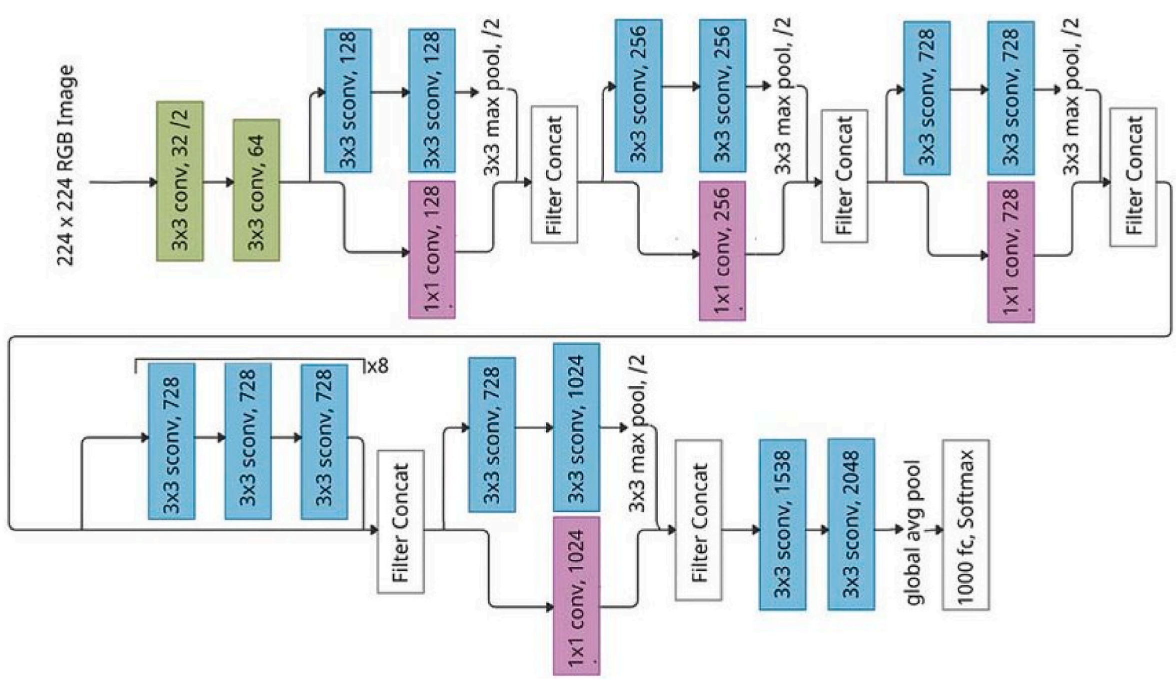

Fig. 13. The comprehensive architecture of Xception-based hyper optimisation.

other CNN models, facilitating an increase in detection and classification parameters. Furthermore, the CNN models (ResNet-101 and Xception) discussed in this paper were specifically trained and tested for concrete cracks. Consequently, the same enhanced models can be applied to different datasets, such as pavement cracks or other objects.

While the proposed multi-level optimisation has improved the detection and classification metrics of the two models, a limitation exists concerning the size of the training dataset which was divided into 20,000 images for training and 20,000 images for testing. In future research, if the dataset size increases, it is anticipated that the detection and classification metrics will experience further enhancement.

5. Conclusion

This study elucidates upon the mechanism of utilising hybrid optimisation techniques to augment the detection and classification abilities

Table 5
Models performance before and after optimisation.

CNN	Method	Accuracy (%)	Sensitivity (%)	F-score (%)
ResNet-101	**Without optimisation**	83.1	92.87	84.58
	After the 1st optimisation stage	97.6	95.46	97.68
	After the 2nd optimisation stage	98.9	99.54	98.89
	Without optimisation	87.4	93.69	88.17
Xception	**After the 1st optimisation stage**	98.3	99.86	98.29
	After the 2nd optimisation stage	99.2	99.5	99.24

of pre-trained models in identifying concrete surface cracks. Two pre-trained models (*viz.* ResNet101 and Xception) were employed to implement the proposed hybrid optimisation on a dataset comprising 40,000 images. The aim was to enhance their crack detection performance across concrete surfaces, achieved through a two-layer optimisation process. In the first layer, four algorithms (SMO, SFLA, GWO and WO) were individually tested to enhance the models' performance. This led to an improvement in both models' accuracies, ranging from 6.2% to 9.4%. Furthermore, three algorithms (SMO, SFLA and GWO) were combined using the concatenation technique, resulting in a minimum 14.5% improvement for ResNet101 and a 10.9% enhancement for Xception. In the second optimisation layer, Xception observed a 0.9% increase in accuracy percentage, while ResNet101 observed a 1.3% increase. These refinements depict an iterative enhancement process contributing to improved performance in detecting concrete surface cracks for both models. The impact of optimisers exhibits variability between ResNet101 and Xception. Particularly, WO displayed the most significant impact on both models, enhancing Xception's performance by 9% and proving highly beneficial for ResNet101 with a boost of 9.4%. Therefore, WO was selected for use in the second layer of optimisation due to its superior accuracy across both CNN networks. The accuracy of Xception experienced a significant improvement, reaching approximately 99.2%, while ResNet-101 saw a commendable enhancement to 98.8% through the implementation of WO in the second optimisation layer.

These research results represent a significant stride toward achieving a fully automated process for accurately determining the structural health condition of concrete buildings. Such work could improve the performance of structural surveys conducted and potentially eliminate human errors occurring from manual inspections. These palpable benefits could have significant cost and quality implications that could invariably augment profitability of inspections undertaken and/or lead to greater client satisfaction. The impact and value of this work could well extend beyond the initial applications prescribed in this current study. For example, this advancement could pave the way for the adoption of other Industry 4.0 technologies (cf. Newman et al., 2021), such as digital twins and the Internet of Things (IoT). By utilising cameras to capture images of cracks and employing the proposed solutions, it becomes possible to automatically assess both the current and future conditions of the building.

The proposed multilayer optimisation process, encompassing both the first and second layers, has significant potential for implementation across other datasets and pre-trained models. It can enhance model performance in distress detection by either employing the methodology of concentration optimisation or utilising the same algorithms with different pre-trained models or datasets. The flexibility of the approach suggests that it is not limited to ResNet-101 and Xception but can be adapted to other architectures to assess its effectiveness.

Moreover, the methodology's adaptability extends beyond the current application, as it can be employed to detect other types of distress, such as highway pavement distress, in future studies. This versatility underscores the potential of the proposed method to generalise across various models and datasets, offering a robust framework for optimising performance in diverse distress detection scenarios.

CRediT authorship contribution statement

Faris Elghaish: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Sandra Matarneh:** Writing – original draft, Methodology, Data curation. **Farzad Pour Rahimian:** Writing – original draft, Methodology. **Essam Abdellatef:** Software, Methodology. **David Edwards:** Writing – review

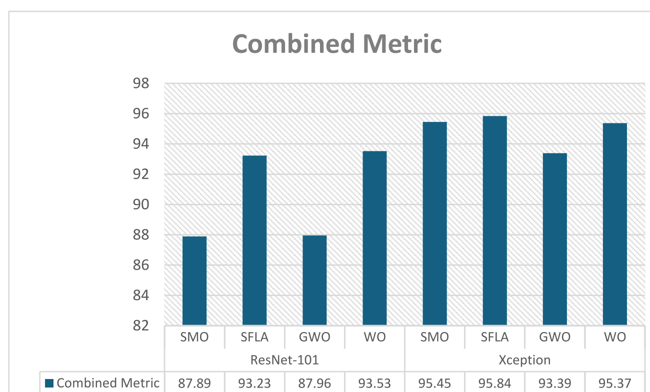

Fig. 15. Combined metric per optimiser for each CNN model.

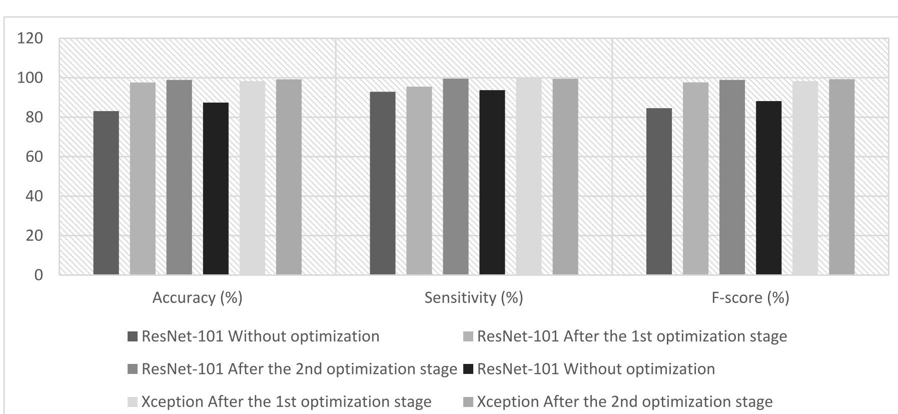

Fig. 14. Graphical Comparison between traditional and optimised ResNet-101 and Xception CNN models.

& editing, Writing – original draft, Visualization, Methodology. **Obuks Ejohwomu:** Writing – review & editing, Visualization, Methodology. **Mohammed Abdelmegid:** Writing – review & editing, Writing – original draft. **Chansik Park:** Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Faris Elghaish reports was provided by Queen's University Belfast. If there are other authors, they 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

Data will be made available on request.

References

Akgül, İ., 2023. Mobile-DenseNet: detection of building concrete surface cracks using a new fusion technique based on deep learning. Heliyon 9 (10), e21097.

Alipour, M., Harris, D.K., Miller, G.R., 2019. Robust pixel-level crack detection using deep fully convolutional neural networks. J. Comput. Civ. Eng. 33 (6), 04019040. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000854.

Amjad, H., Khattak, M.M.H., Khushnood, R.A., 2023. A simplified machine learning empirical model for biomimetic crack healing of bio-inspired concrete. Mater. Today Commun. 37. https://doi.org/10.1016/j.mtcomm.2023.107063.

Andrushia, A.D., Anand, N., Neebha, T.M., Naser, M.Z., Lubloy, E., 2022. Autonomous detection of concrete damage under fire conditions. Autom. ConStruct. 140, 104364. https://doi.org/10.1016/j.autcon.2022.104364.

Andrushia, A.D., Anand, N., Prince Arulraj, G., 2021. Evaluation of thermal cracks on fire exposed concrete structures using Ripplet transform. Math. Comput. Simulat. 180, 93–113. https://doi.org/10.1016/j.matcom.2020.07.024.

Asadi Shamsabadi, E., Xu, C., Rao, A.S., Nguyen, T., Ngo, T., Dias-da-Costa, D., 2022a. Vision transformer-based autonomous crack detection on asphalt and concrete surfaces. Autom. ConStruct. 140, 104316. https://doi.org/10.1016/j.autcon.2022.104316.

Asadi Shamsabadi, E., Xu, C., Rao, A.S., Nguyen, T., Ngo, T., Dias-da-Costa, D., 2022b. Vision transformer-based autonomous crack detection on asphalt and concrete surfaces. Autom. ConStruct. 140. https://doi.org/10.1016/j.autcon.2022.104316.

BaniMustafa, A., AbdelHalim, R., Bulkrock, O., Al-Hmouz, A., 2023. Deep learning for assessing severity of cracks in concrete structures. Int. J. Comput. Commun. Control 18 (1). https://doi.org/10.15837/ijccc.2023.1.4977.

Çelik, F., König, M., 2022. A sigmoid-optimized encoder–decoder network for crack segmentation with copy-edit-paste transfer learning. Comput. Aided Civ. Infrastruct. Eng. 37 (14), 1875–1890. https://doi.org/10.1111/mice.12844.

Cha, Y.-J., Choi, W., Büyüköztürk, O., 2017. Deep learning-based crack damage detection using convolutional neural networks. Comput. Aided Civ. Infrastruct. Eng. 32 (5), 361–378. https://doi.org/10.1111/mice.12263.

Cha, Y.-J., Choi, W., Suh, G., Mahmoudkhani, S., Büyüköztürk, O., 2018. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types, computer-aided civil and infrastructure. Engineering 33 (9), 731–747. https://doi.org/10.1111/mice.12334.

Chang, Z., Wan, Z., Xu, Y., Schlangen, E., Šavija, B., 2022. Convolutional neural network for predicting crack pattern and stress-crack width curve of air-void structure in 3D printed concrete. Eng. Fract. Mech. 271. https://doi.org/10.1016/j.engfracmech.2022.108624.

Chen, K., Yadav, A., Khan, A., Meng, Y., Zhu, K., 2019a. Improved crack detection and recognition based on convolutional neural network. Modelling and Simulation in Engineering 2019, 8796743. https://doi.org/10.1155/2019/8796743.

Chen, P.-H.C., Liu, Y., Peng, L., 2019b. How to develop machine learning models for healthcare. Nat. Mater. 18 (5), 410–414.

Deng, W., Mou, Y., Kashiwa, T., Escalera, S., Nagai, K., Nakayama, K., Matsuo, Y., Prendinger, H., 2020. Vision based pixel-level bridge structural damage detection using a link ASPP network. Autom. ConStruct. 110. https://doi.org/10.1016/j.autcon.2019.102973.

Ding, W., Sun, Y., Ren, L., Ju, H., Feng, Z., Li, M., 2020. Multiple lesions detection of fundus images based on convolution neural network algorithm with improved SFLA. IEEE Access 8, 97618–97631.

Dorafshan, S., Thomas, R.J., Maguire, M., 2018. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Construct. Build. Mater. 186, 1031–1045.

Elghaish, F., Matarneh, S.T., Talebi, S., Abu-Samra, S., Salimi, G., Rausch, C., 2022. Deep learning for detecting distresses in buildings and pavements: a critical gap analysis. Construct. Innovat. 22 (3), 554–579. https://doi.org/10.1108/CI-09-2021-0171.

Elhariri, E., El-Bendary, N., Taie, S.A., 2020. Using hybrid filter-wrapper feature selection with multi-objective improved-salp optimization for crack severity recognition. IEEE Access 8, 84290–84315. https://doi.org/10.1109/ACCESS.2020.2991968.

Gao, Y., Mosalam, K.M., 2018. Deep transfer learning for image-based structural damage recognition. Comput. Aided Civ. Infrastruct. Eng. 33 (9), 748–768. https://doi.org/10.1111/mice.12363.

Gaur, A., Kishore, K., Jain, R., Pandey, A., Singh, P., Wagri, N.K., Roy-Chowdhury, A.B., 2023. A novel approach for industrial concrete defect identification based on image processing and deep convolutional neural networks. Case Stud. Constr. Mater. 19. https://doi.org/10.1016/j.cscm.2023.e02392.

Gu, X., Jin, X., Zhou, Y., 2016. Basic Principles of Concrete Structures. Springer. ISBN: 3662485656.

Guernine, A., Kimour, M.T., 2021. Optimized training for convolutional neural network using enhanced grey wolf optimization algorithm. Informatica 45 (5).

Han, Q., Liu, X., Xu, J., 2022. Detection and location of steel structure surface cracks based on unmanned aerial vehicle images. J. Build. Eng. 50. https://doi.org/10.1016/j.jobe.2022.104098.

Hang, J., Wu, Y., Li, Y., Lai, T., Zhang, J., Li, Y., 2023. A deep learning semantic segmentation network with attention mechanism for concrete crack detection. Struct. Health Monit. 22 (5), 3006–3026. https://doi.org/10.1177/14759217221126170.

Houssein, E.H., Saeed, M.K., Hu, G., Al-Sayed, M.M., 2024. An efficient improved exponential distribution optimizer: application to the global. Engineering and Combinatorial Optimization Problems. Cluster Computing, pp. 1–36.

Howden-Chapman, P., Bennett, J., Edwards, R., Jacobs, D., Nathan, K., Ormandy, D., 2023. Review of the impact of housing quality on inequalities in health and well-being. Annu. Rev. Publ. Health 44, 233–254.

Jang, K., Kim, N., An, Y.-K., 2019. Deep learning–based autonomous concrete crack evaluation through hybrid image scanning. Struct. Health Monit. 18 (5–6), 1722–1737. https://doi.org/10.1177/1475921718821719.

Jiang, H., Tsai, S.-B., 2021. An empirical study on sports combination training action recognition based on SMO algorithm optimization model and artificial intelligence. Math. Probl Eng. 2021 (1), 7217383.

Kang, D.H., Cha, Y.-J., 2022. Efficient attention-based deep encoder and decoder for automatic crack segmentation. Struct. Health Monit. 21 (5), 2190–2205. https://doi.org/10.1177/14759217211053776.

Kumaran, N., Vadivel, A., Kumar, S.S., 2018. Recognition of human actions using CNN-GWO: a novel modeling of CNN for enhancement of classification performance. Multimed. Tool. Appl. 77 (18), 23115–23147.

Larosche, C.J., 2009. 3 - types and causes of cracking in concrete structures. In: Delatte, N. (Ed.), Failure, Distress and Repair of Concrete Structures. Woodhead Publishing, pp. 57–83. https://doi.org/10.1533/9781845697037.1.57.

Li, D., Liu, J., Hu, S., Cheng, G., Li, Y., Cao, Y., Dong, B., Chen, Y.F., 2022a. A deep learning-based indoor acceptance system for assessment on flatness and verticality quality of concrete surfaces. J. Build. Eng. 51. https://doi.org/10.1016/j.jobe.2022.104284.

Li, S., Zhao, X., Zhou, G., 2019. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network. Comput. Aided Civ. Infrastruct. Eng. 34 (7), 616–634. https://doi.org/10.1111/mice.12433.

Li, Y., Bao, T., Xu, B., Shu, X., Zhou, Y., Du, Y., Wang, R., Zhang, K., 2022b. A deep residual neural network framework with transfer learning for concrete dams patch-level crack classification and weakly-supervised localization. Measurement: Journal of the International Measurement Confederation 188. https://doi.org/10.1016/j.measurement.2021.110641.

Liu, C., Wang, P., Wang, X., Miao, J., 2024a. Autonomous damage segmentation of post-fire reinforced concrete structural components. Adv. Eng. Inf. 61, 102498. https://doi.org/10.1016/j.aei.2024.102498.

Liu, F., Liu, J., Wang, L., 2022. Asphalt pavement crack detection based on convolutional neural network and infrared thermography. IEEE Trans. Intell. Transport. Syst. 23 (11), 22145–22155. https://doi.org/10.1109/TITS.2022.3142393.

Liu, F., Liu, J., Wang, L., Al-Qadi, I.L., 2024b. Multiple-type distress detection in asphalt concrete pavement using infrared thermography and deep learning. Autom. ConStruct. 161, 105355. https://doi.org/10.1016/j.autcon.2024.105355.

Liu, Z., Cao, Y., Wang, Y., Wang, W., 2019. Computer vision-based concrete crack detection using U-net fully convolutional networks. Autom. ConStruct. 104, 129–139. https://doi.org/10.1016/j.autcon.2019.04.005.

Matarneh, S., Elghaish, F., Pour Rahimian, F., Abdellatef, E., Abrishami, S., 2024. Evaluation and optimisation of pre-trained CNN models for asphalt pavement crack detection and classification. Autom. ConStruct. 160, 105297. https://doi.org/10.1016/j.autcon.2024.105297.

Miao, P., Srimahachota, T., 2021. Cost-effective system for detection and quantification of concrete surface cracks by combination of convolutional neural network and image processing techniques. Construct. Build. Mater. 293. https://doi.org/10.1016/j.conbuildmat.2021.123549.

Ni, F., Zhang, J., Chen, Z., 2019. Pixel-level crack delineation in images with convolutional feature fusion. Struct. Control Health Monit. 26 (1), e2286.

Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., Shoaib, M., 2020. A deep learning model based on concatenation approach for the diagnosis of brain tumor. IEEE Access 8, 55135–55144.

Owusu-Manu, D.-G., Quaigrain, R.A., Edwards, D.J., Hammond, M., Hammond, M., Roberts, C., 2022. Energy conservation literacy among households in Sub-Sahara Africa. Int. J. Energy Sect. Manag. 16 (6), 1130–1149.

Özgenel, Ç.F., Sorguç, A.G., 2018. Performance comparison of pretrained convolutional neural networks on crack detection in buildings. Isarc. Proceedings of the International Symposium on Automation and Robotics in Construction, vol. 35. IAARC Publications, pp. 1–8.

Paramanandham, N., Koppad, D., Anbalagan, S., 2022. Vision based crack detection in concrete structures using cutting-edge deep learning techniques. Trait. Du. Signal 39 (2), 485–492. https://doi.org/10.18280/ts.390210.

Philip, R.E., Andrushia, A.D., Nammalvar, A., Gurupatham, B.G.A., Roy, K., 2023. A comparative study on crack detection in concrete walls using transfer learning techniques. Journal of Composites Science 7 (4), 169.

Piyathilaka, L., Preethichandra, D.M.G., Izhar, U., Kahandawa, G., 2020. Real-time concrete crack detection and instance segmentation using deep transfer learning. Engineering Proceedings 2 (1), 91.

Qu, F., Li, W., Dong, W., Tam, V.W., Yu, T., 2021. Durability deterioration of concrete under marine environment from material to structure: a critical review. J. Build. Eng. 35, 102074.

Qu, Z., Li, Y.-x., Zhou, Q., 2022. CrackT-net: a method of convolutional neural network and transformer for crack segmentation. J. Electron. Imag. 31, 023040 - 023040.

Rajadurai, R.-S., Kang, S.-T., 2021. Automated vision-based crack detection on concrete surfaces using deep learning. Appl. Sci. 11 (11), 5229.

Rao, A.S., Nguyen, T., Palaniswami, M., Ngo, T., 2021. Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. Struct. Health Monit. 20 (4), 2124–2142. https://doi.org/10.1177/1475921720965445.

Samma, H., Suandi, S.A., Ismail, N.A., Sulaiman, S., Ping, L.L., 2021. Evolving pre-trained CNN using two-layers optimizer for road damage detection from drone images. IEEE Access 9, 158215–158226. https://doi.org/10.1109/ACCESS.2021.3131231.

Sun, Y., Yang, Y., Yao, G., Wei, F., Wong, M., 2021. Autonomous crack and bughole detection for concrete surface image based on deep learning. IEEE Access 9, 85709–85720.

Tan, Y., Cai, R., Li, J., Chen, P., Wang, M., 2021. Automatic detection of sewer defects based on improved you only look once algorithm. Autom. ConStruct. 131. https://doi.org/10.1016/j.autcon.2021.103912.

Varoquaux, G., Colliot, O., 2023. Evaluating machine learning models and their diagnostic value. Machine Learning for Brain Disorders 601–630.

Wang, W., Su, C., 2022. Automatic concrete crack segmentation model based on transformer. Autom. ConStruct. 139, 104275. https://doi.org/10.1016/j.autcon.2022.104275.

Wu, Y., Li, S., Zhang, J., Li, Y., Li, Y., Zhang, Y., 2024. Dual attention transformer network for pixel-level concrete crack segmentation considering camera placement. Autom. ConStruct. 157, 105166. https://doi.org/10.1016/j.autcon.2023.105166.

Xiong, C., Zayed, T., Abdelkader, E.M., 2024. A novel YOLOv8-GAM-Wise-IoU model for automated detection of bridge surface cracks. Construct. Build. Mater. 414, 135025.

Xu, H., Su, X., Wang, Y., Cai, H., Cui, K., Chen, X., 2019a. Automatic bridge crack detection using a convolutional neural network. Appl. Sci. 9 (14), 2867.

Xu, S., Hao, M., Liu, G., Meng, Y., Han, J., Shi, Y., 2022. Concrete crack segmentation based on convolution–deconvolution feature fusion with holistically nested networks. Struct. Control Health Monit. 29 (8), e2965. https://doi.org/10.1002/stc.2965.

Xu, Y., Wei, S., Bao, Y., Li, H., 2019b. Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network. Struct. Control Health Monit. 26 (3), e2313. https://doi.org/10.1002/stc.2313.

Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., Yang, X., 2018. Automatic pixel-level crack detection and measurement using fully convolutional network. Comput. Aided Civ. Infrastruct. Eng. 33 (12), 1090–1109. https://doi.org/10.1111/mice.12412.

Yang, Y., Zhou, D.-W., Zhan, D.-C., Xiong, H., Jiang, Y., Yang, J., 2021. Cost-effective incremental deep model: matching model capacity with the least sampling. IEEE Trans. Knowl. Data Eng. 35 (4), 3575–3588.

Yu, Y., Rashidi, M., Samali, B., Mohammadi, M., Nguyen, T.N., Zhou, X., 2022a. Crack detection of concrete structures using deep convolutional neural networks optimized by enhanced chicken swarm algorithm. Struct. Health Monit. 21 (5), 2244–2263. https://doi.org/10.1177/14759217211053546.

Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T.N., Zhang, G., 2022b. Vision-based concrete crack detection using a hybrid framework considering noise effect. J. Build. Eng. 61. https://doi.org/10.1016/j.jobe.2022.105246.

Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T.N., Zhang, G., 2022c. Vision-based concrete crack detection using a hybrid framework considering noise effect. J. Build. Eng. 61, 105246. https://doi.org/10.1016/j.jobe.2022.105246.

Yu, Z., Shen, Y., Shen, C., 2021. A real-time detection approach for bridge cracks based on YOLOv4-FPM. Autom. ConStruct. 122. https://doi.org/10.1016/j.autcon.2020.103514.

Zhang, C., Chang, C.C., Jamshidi, M., 2020. Concrete bridge surface damage detection using a single-stage detector. Comput. Aided Civ. Infrastruct. Eng. 35 (4), 389–409. https://doi.org/10.1111/mice.12500.

Zhang, Q., Barri, K., Babanajad, S.K., Alavi, A.H., 2021. Real-time detection of cracks on concrete bridge decks using deep learning in the frequency domain. Engineering 7 (12), 1786–1796. https://doi.org/10.1016/j.eng.2020.07.026.

Zhang, Y., Ni, Y.-Q., Jia, X., Wang, Y.-W., 2023. Identification of concrete surface damage based on probabilistic deep learning of images. Autom. ConStruct. 156, 105141. https://doi.org/10.1016/j.autcon.2023.105141.