

Article

Deep Learning Activation Layer-Based Wall Quality Recognition Using Conv2D ResNet Exponential Transfer Learning Model

Bubryur Kim ¹, Yuvaraj Natarajan ^{1,2}, Shyamala Devi Munisamy ³, Aruna Rajendran ³, K. R. Sri Preethaa ^{1,2,*}, Dong-Eun Lee ^{4,*} and Gitanjali Wadhwa ²

¹ Department of Robot and Smart System Engineering, Kyungpook National University, 80, Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea

² Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore 641407, India

³ Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai 600062, India

⁴ School of Architecture, Civil, Environment and Energy Engineering, Kyungpook National University, 80, Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea

* Correspondence: k.r.sripreethaa@kpri.ac.in (K.R.S.P.); dolee@knu.ac.kr (D.-E.L.)

Abstract: Crack detection is essential for observing structural health and guaranteeing structural safety. The manual crack and other damage detection process is time-consuming and subject to surveyors' biased judgments. The proposed Conv2D ResNet Exponential model for wall quality detection was trained with 5000 wall images, including various imperfections such as cracks, holes, efflorescence, damp patches, and spalls. The model was trained with initial weights to form the trained layers of the base model and was integrated with Xception, VGG19, DenseNet, and ResNet convolutional neural network (CNN) models to retrieve the general high-level features. A transfer deep-learning-based approach was implemented to create a custom layer of CNN models. The base model was combined with custom layers to estimate wall quality. Xception, VGG19, DenseNet, and ResNet models were fitted with different activation layers such as softplus, softsign, tanh, selu, elu, and exponential, along with transfer learning. The performance of Conv2D was evaluated using model loss, precision, accuracy, recall, and F-score measures. The model was validated by comparing the performances of Xception, VGG19, DenseNet, ResNet, and Conv2D ResNet Exponential. The experimental results show that the Conv2D ResNet model with an exponential activation layer outperforms it with an F-score value of 0.9978 and can potentially be a viable substitute for classifying various wall defects.

Keywords: deep learning; Conv2D; activation layer; transfer learning; F-score

MSC: 68T10; 68T45



Citation: Kim, B.; Natarajan, Y.; Munisamy, S.D.; Rajendran, A.; Sri Preethaa, K.R.; Lee, D.-E.; Wadhwa, G. Deep Learning Activation Layer-Based Wall Quality Recognition Using Conv2D ResNet Exponential Transfer Learning Model. *Mathematics* **2022**, *10*, 4602. <https://doi.org/10.3390/math10234602>

Academic Editors: Jakub Nalepa and Jonathan Blackledge

Received: 12 October 2022

Accepted: 2 December 2022

Published: 5 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cracks are the first signs of civil structure degradation, and they can arise for various reasons, including structural foundation displacement, shrinkage and extension, uneven mix, bloated soil, and overloaded environmental and manufacturing disasters. Crack detection and recognition activities can be conducted automatically or manually and are subjected to human experts visually analyzing and evaluating the structure [1]. Manual inspection involves fetching a schematic of a crack and recording the circumstances of the abnormalities. Usually, manual inspection methods take a long time, rely on the observer, are potentially sensitive to the inspector's insight, and lack descriptive methodology [2]. Automatic inspection methods provide a coherent solution that reduces subjectivity and replaces manual observation with the human eye [3]. Automatic crack identification has

been established considering slow, biased, and old human inspection processes for fast and efficient surface defect assessment [4]. An image vision computer system is introduced to address this shortcoming, with the goal of instantly and reliably transforming image or video input into meaningful intelligence. It contains structural components, recognizes modifications in a reference image, and quantifies local and global visual damage [5]. Automating the process may significantly reduce computational costs and allow regular inspection intervals. The detection of cracks and disintegration of a bridge is examined as a part of automation using the rapid Haar transform [6]. Crack detection filters of various sizes were developed to locate cracking spots in inspection imagery. A semi-automatic strategy incorporating Sobel and Laplacian operators was used to identify crack edges, and a graph search algorithm was used to obtain cracks depending on the user input [7]. A principal-component-analysis-(PCA)-based approach combined with linear structural analysis has been reported to identify linear structural fractures in concrete bridge decks with the highest classification accuracy [8]. One of the most critical structural inspection and maintenance methods is vision-based technology, which utilizes essential diagnostic imaging devices, such as sensors [9]. Miscommunication deflection assessment, steel corrosion recognition, and spalling diagnosis are some of the most recent advancements in vision-based inspection and testing [10]. As a result, this technique has some constraints in terms of real-world applications as creating an automated system that can encompass all unpredicted possibilities for fast perceptible damage remediation in the physical world is difficult [11]. Deep learning (DL) has recently been recognized as one of the most potent remedies to this challenge [12]. In addition, machine learning (ML) strategies for DL models based on neural networks containing numerous hidden units have evolved with numerous benefits for a better solution [13].

Structural inspections across long, nonintrusive distances based on high precision have been performed using optics and computer vision advancements. The fundamental shortcoming of image-processing-based approaches is the lack of consistency and vibration among the fracture pixels [14]. DeepLab V2 was used to detect several cracks in the images. Automated inspection has achieved an expected improvement owing to the rapid evolution of ML. ML algorithms can acquire feature representations and execute confidence intervals without requiring human model training, which is performed using conventional techniques [15]. Data collection, feature extraction, and categorization are all conventional machine learning algorithms for pavement crack detection. Shallow convolutional-neural-network-(CNN)-based architectures have been implemented to identify surface cracks and achieve greater accuracy with efficient computational costs [16,17]. Deep CNN systems use a multilayer neural net to retrieve significant characteristics from the input data [18]. Numerous analyses of ML-based crack detection methods have revealed that the classifier may not produce reliable results if the derived features do not identify actual cracks [19]. The Hessian matrix has been used both to accentuate cracks over blebs or staining and to modify the thickness fluctuation of cracks during image pre-processing [20]. Probabilistic relaxation is employed to determine cracks coarsely, eliminate noise, and accomplish adaptive thresholding for their operation [21]. Transfer learning methods allow CNNs to be used without incurring high computing costs or needing a prior understanding of the working functionality of CNN layers. Visual Geometry Group's VGGNet [22], Microsoft's ResNet [23], and Inception-V3 [24] are some of Google's transfer learning design models that employ photographic data as input. Models built over the hybridization of ML algorithms have been found to increase the performance of vision-based systems over traditional algorithms. Hybridizing optimization techniques with support vector machines (SVM), such as fuzzy logic, k-nearest neighbours, artificial neural networks (ANN), and evolutionary algorithms, led to substantial improvements in recognition accuracy [25]. Pavement Crack Detection Net (PCDNet) eliminates more localized distortion, detects smaller cracks, and interprets information at a significantly faster rate than other methods [26]. Optimal intelligence technology was used to examine the wall quality parameters concerning many components of the dynamic situations of the retaining wall [27]. In addition, using ant

colony optimization (ACO), these criteria resulted in the best design solution. A hybrid ANN-ACO algorithm utilizes various parameters under different structural conditions, and dynamic loads remarkably impact structural models [16]. Ensemble models built by stacking the best-performing ML algorithms have proven to produce efficient results in predictive models [28]. A pre-processed CNN classifier of VGG-16 with ResNet-50 detects fractures in image and inception models for object localization incorporated class activation monitoring (CAM) for excellent optimal detection [29]. Optimizing the hyperparameters of CNN architectures, such as VGG16 and Resnet, proved to achieve greater accuracy in object identification and classification [30]. Unmanned aerial vehicles (UAVs) and camera systems were implemented to identify cracks and trigger U-Net in pixel-wise classification for feature and flaw identification with various feature sets. Texture-based video processing methods handle local binary patterns (LBP) using SVM and Bayes decision theories. For noisy and complicated bridge photographs, wavelet features were retrieved from the scene using a sliding-window-texture-analysis-based technique [31]. A deep CNN-based damage locating method used DenseNet to identify the damaged and undamaged steel frames from the images provided as inputs. The model outperformed the MobileNet and ResNet architectures [32]. Cracks are detected by the trained ConvNet, SVM, and boosting methods on a sliding window, where the SVM sometimes fails to distinguish the crack from the background. CNN using traditional Canny and Sobel edge detection methods can scan larger pixel images and exhibit excellent robustness during the training process. The AdaBoost classifier was used for pre-processing the crack image, and DL techniques were used for crack-detection in the image data [33]. An EfficientNet-based transfer learning model was developed to detect and classify surface cracks in high-rise buildings using UAV. Microcrack detection is achieved by solving the binary classification problem of cracks using autoencoders and softmax regression [34]. CNN-based order forensics framework for detecting picture operator chains is described. The two-stream Framework captures both local noise residual data and manipulating artifacts proof. The model may automatically detect alternated detection features straight from picture data and is suggested explicitly for forensically recognizing a chain consisting of two image operators [35]. To directly extract features from the photos, dual-filtering CNN base was designed. It treats each resampling parameter as a separate class, followed by the formulation of resampling parameter estimation and reformulates it as a multi-classification problem [36]. A reliable blind watermarking system based on 3D convolutional neural networks that can extract and integrate watermark images into animated GIF files was proposed [37].

Literature studies reviewed that manual inspection of structures is challenging, time-consuming and provides biased results. Several research studies have implemented different image-based methods, machine learning and deep learning algorithms to enable the automatic monitoring of building structures. The performance of ML algorithms on image data could have been more appreciated as it depends on many features and encompasses complex feature engineering tasks. The DCNN-based models were employed for damage detection and classifications, but not all models performed efficiently due to insufficient data, overfitting, and vanishing gradient problems. The models' competency was enhanced by customising the convolutional layers, hybridising, ensembling, and transfer-learning techniques. The ResNet, DenseNet, VGG, and Xception models were observed to perform efficiently on detecting the structural damages such as concrete cracks, and steel bar damages.

To overcome the issues discussed in previous studies, this study proposes an automatic vision-based crack-identification system based on DL to identify crack portions from a large dataset of images acquired in the environment. The main breakthrough is the establishment of neural network-based classification models for various structural environments. Using various cameras and visual equipment, such as drones, this technology intends to ease routine inspections of concrete buildings and increase the speed of the diagnosis process of precise crack distribution while retaining accuracy. First, a series of images captured under a combination of structural, meteorological, and photographic circumstances were

gathered, allowing for easy classification of the images using search keywords. A transfer learning approach was established to minimize time and money while constructing a DL model.

2. Research Methodology

The Conv2D ResNet exponential model was fitted with a dataset of images for each wall defect, such as cracks, holes, efflorescence, damp patches, and spalls. The dataset was collected through publicly available dataset repositories, such as the kaggle dataset or Structural Defects Network (SDNET) 2018 [38]. The model was trained with 80% of the 5000 images and tested with 20% of the images. The research methodology used in this study is shown in Figure 1. In stage 1, existing crack detection methods for building walls were explored and analysed. In Stage 2, a novel Conv2D ResNet exponential model was designed to detect the damage class of the building wall. The dataset consisted of a collection of wall images with different defects, such as cracks, holes, efflorescence, damp patches, and spalls. Training and testing of the proposed model were performed using an 80:20 wall quality dataset. Stage 3 evaluated the proposed model using the test data and compared Conv2D with existing models, such as DenseNet, VGG19, and Xception, of several activation layers, such as softplus, softsign, Relu, elu, and tanh. In stage 4, the performance of the proposed model was analysed using metrics such as precision, recall, F-score, and accuracy. The architecture of the DL activation layer is illustrated in Figure 2.

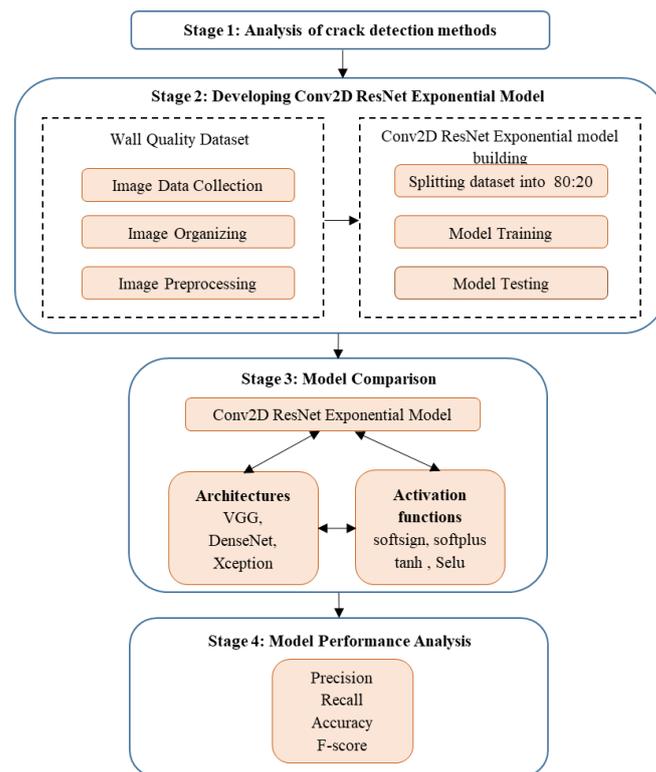


Figure 1. Research methodology flow.

The workflow of the Conv2D ResNet exponential model is illustrated in Figure 3. An open access dataset (SDNET2018) published by Utah State University is used for implementation. In this proposed research work, the Snip&Sketch annotation tool was used for extracting the region of interest from the images available in the dataset. The Conv2D ResNet exponential model was trained with 4000 images and tested with 1000 images. The initial base model was pre-trained with initial weights using ImageNet, and Conv2D was designed with CNN models such as Xception, VGG19, DenseNet, and ResNet to extract the general high-level features. The base model was added to the custom layers, developed

using a transfer-based DL approach, and the performance was analysed on wall quality prediction. The Xception, VGG19, DenseNet, and ResNet models were fitted with different activation layers, such as softplus, softsign, tanh, selu, elu, and exponential, and were evaluated using model loss, precision, accuracy, recall, and F-score measures.

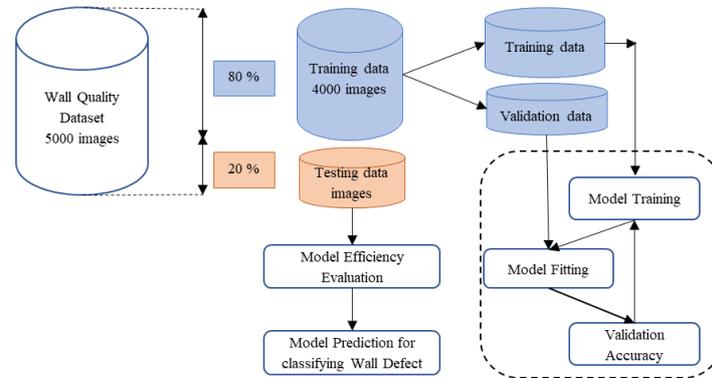


Figure 2. Deep learning activation layer model architecture.

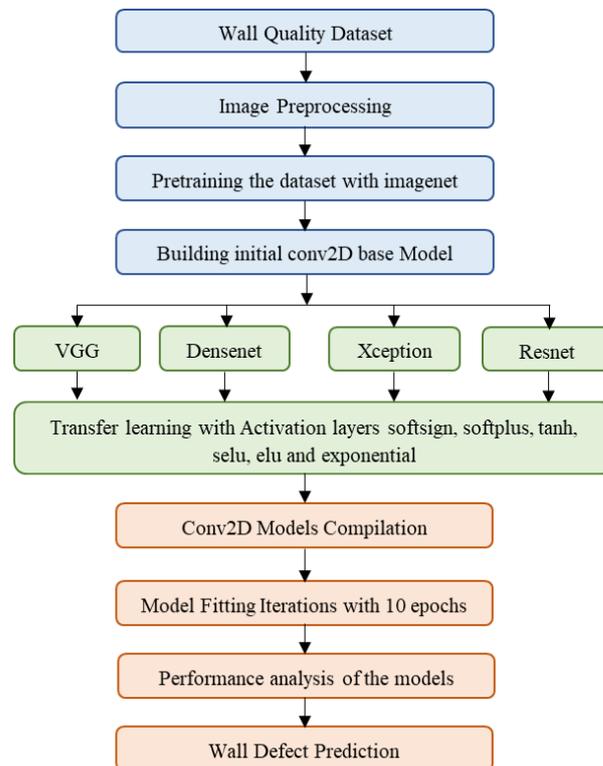


Figure 3. Workflow of Conv2D ResNet exponential model.

As an overview of novelty, Conv2D ResNet exponential model was built over the Conv2D ResNet base model, along with the transfer learning custom activation layers that classify the wall defects more effectively with high accuracy. The wall quality dataset was fitted with the Conv2D ResNet model, which acted as a base model, thereby learning the types of wall defects from 90% of the training data. The acquired knowledge from Conv2D ResNet was transferred to refit the model by integrating the layer with the exponential activation function, which identified the wall defects in a single image, thereby validating the transfer learning by transporting the knowledge from the base model to the custom layer that enhanced the accuracy.

3. Implementation Setup

A dataset with 1000 images was collected for various defects, such as cracks, holes, efflorescence, damp patches, and spalls. Wall cracks signify trouble with the living area’s groundwork. Once wall cracks are detected in a residence, it generally means that the foundation is transitioning. Cracks in walls are caused by the contraction and relaxation of construction materials due to temperature and water content oscillations. A hole in a wall is also a significant defect that affects the quality of a building structure over time. Efflorescence is merely the deposition of salts on the surface of the aggregate, which is usually white. Once dry, the efflorescence consists of a white cover outside the concrete wall. When absorbed water and salts vaporize, they appear on the walls as crystallized patches or a layer of white powder. Damp patch condensation occurs when warm humid air within the same room is exposed to a cold interior wall or surface. Then, it quickly compresses the air back into the water. This evaporation then condenses on the interior wall surface, causing damp patches on the wall. A spall in the wall refers to the discoloration, clamping, fading, imploding, or flaking of concrete or brickwork, especially where the surface components have been destroyed. A spall can occur because of moisture absorption, combustion, or mechanical processes. The wall quality dataset contains 1000 images for each defect, such as holes, cracks, efflorescence, damp patches, and spalls, and is represented as follows in Equations (1)–(6):

$$Wall = H_{1000} + C_{1000} + E_{1000} + D_{1000} + S_{1000}, \tag{1}$$

$$H_{1000} = \lfloor U_{h=1}^{1000} \left\{ \sum_{i=1}^{255} \sum_{j=1}^{255} H_{ijh} \right\} \rfloor, \tag{2}$$

$$C_{1000} = \lfloor U_{c=1}^{1000} \left\{ \sum_{i=1}^{255} \sum_{j=1}^{255} C_{ijc} \right\} \rfloor, \tag{3}$$

$$E_{1000} = \lfloor U_{e=1}^{1000} \left\{ \sum_{i=1}^{255} \sum_{j=1}^{255} E_{ij_e} \right\} \rfloor, \tag{4}$$

$$D_{1000} = \lfloor U_{d=1}^{1000} \left\{ \sum_{i=1}^{255} \sum_{j=1}^{255} D_{ij_d} \right\} \rfloor, \tag{5}$$

$$S_{1000} = \lfloor U_{s=1}^{1000} \left\{ \sum_{i=1}^{255} \sum_{j=1}^{255} S_{ijs} \right\} \rfloor, \tag{6}$$

where H_{ij_h} represents the wall images with holes, C_{ij_c} represents the wall images with cracks, E_{ij_e} represents the wall images with efflorescence, D_{ij_d} represents the wall images with damp patches, and S_{ij_s} represents the wall images with spall. The wall quality dataset images used for implementation are shown in Figure 4.

The wall quality dataset was pre-trained with the ImageNet by substituting the weights and trained with convolutional neural network models, such as Conv2D Xception, DenseNet, VGG19, and ResNet models, to extract the essential features from the image. Equation (7) represents the Gaussian function applied for feature extraction. The parameter “r” denotes the variance of the Gaussian function.

$$Feature(i, j, r) = \frac{1}{\sqrt{2\pi r}} \exp\left(-\frac{(i^2 + j^2)}{2r^2}\right), \tag{7}$$

The gaussian orientation function used for image filtering is shown in Equation (8).

$$Orient(i, j, r, \theta) = Feature_{ii} \cos^2(\theta) + 2 \text{Feature}_{ij} \cos(\theta) \sin(\theta) + Feature_{jj} \sin^2(\theta), \tag{8}$$

where $Feature_{ii}$, $Feature_{ij}$, and $Feature_{jj}$ denote the second derivatives of the Gaussian function, as represented in Equations (9)–(11).

$$Feature_{ii}(i, j, r) = \frac{(i^2 - r^2) \exp\left(\frac{-((i^2+j^2))}{2r^2}\right)}{\sqrt{2\pi r^5}}, \tag{9}$$

$$Feature_{jj}(i, j, r) = \frac{(j^2 - r^2) \exp\left(\frac{-((i^2+j^2))}{2r^2}\right)}{\sqrt{2\pi r^5}}, \tag{10}$$

$$Feature_{ij}(i, j, r) = \frac{ij \exp\left(\frac{-((i^2+j^2))}{2r^2}\right)}{\sqrt{2\pi r^5}}. \tag{11}$$

The input images in the wall quality dataset were processed with four-layer Conv2D layers—convolution filtering, sigmoid filter, linear transformation, and linear sigmoid—to generate the final output and are denoted in Equations (12)–(14).

$$Z = Feature_{ij} * Filter, \tag{12}$$

$$Con[R, C] = Feature_{ij} * Kernel[R, C], \tag{13}$$

$$Con[R, C] = \sum_j \sum_k kernel[j, k] Feature[R - j][C - k], \tag{14}$$

where “ R, C ” represents the rows and columns of the input image matrix.

The input images in the wall quality dataset were trained with Conv2D and designed with CNN models, such as Xception, VGG19, DenseNet, and ResNet, to extract the general high level features, which are represented in Equations (15)–(18).

$$DenseNet = (5 CP + (TL(6 + 12 + 48) + CL(32)) * 2 DB) \tag{15}$$

$$VGG19 = (5 CP + (CRM(2 + 3 + 4 + 5) + FCL(3)) \tag{16}$$

$$\begin{aligned} Xception = & ((Entry(2 CP + 3 * (2 SC + 1 MP))) \\ & + Middle(8 * (3Relu + 3SC)) \\ & + Exit((1CP + 2Relu + 2SC) + 1FCL) \end{aligned} \tag{17}$$

$$ResNet = (62 CP + (CL(2 + 4 + 8 + 16)) * 3 PL) \tag{18}$$

where $CP, TL, CL, DB, CRM, FCL, SC, MP$, and PL represent the convolution and pooling layer, transition layer, classification layer, dense block layer, convolution relu max pooling layer, fully connected layer, separable convolution layer, max pooling layer, and normal pooling layer, respectively.

Apply the sigmoid function to the above equation, as shown in (19) and (20).

$$ASig = Sigmoid(Z), \tag{19}$$

$$Sigmoid(Z) = \frac{1}{(1 + e^{-Z})}. \tag{20}$$

The linear transformation was applied to the above layer to process the third layer of the CNN, as in (17).

$$ZLinear = Weight^T \times ASig + Bias. \tag{21}$$

The final output is given by applying a linear sigmoid, as given in (18).

$$Output = Sigmoid(ZLinear). \tag{22}$$

The proposed Conv2D ResNet exponential model is built with exponential activation layer and is denoted in (23).

$$\text{Conv2DResNetexponentialmodel} = \text{Exponential}\{\text{Resnet}(\text{Output})\} \quad (23)$$

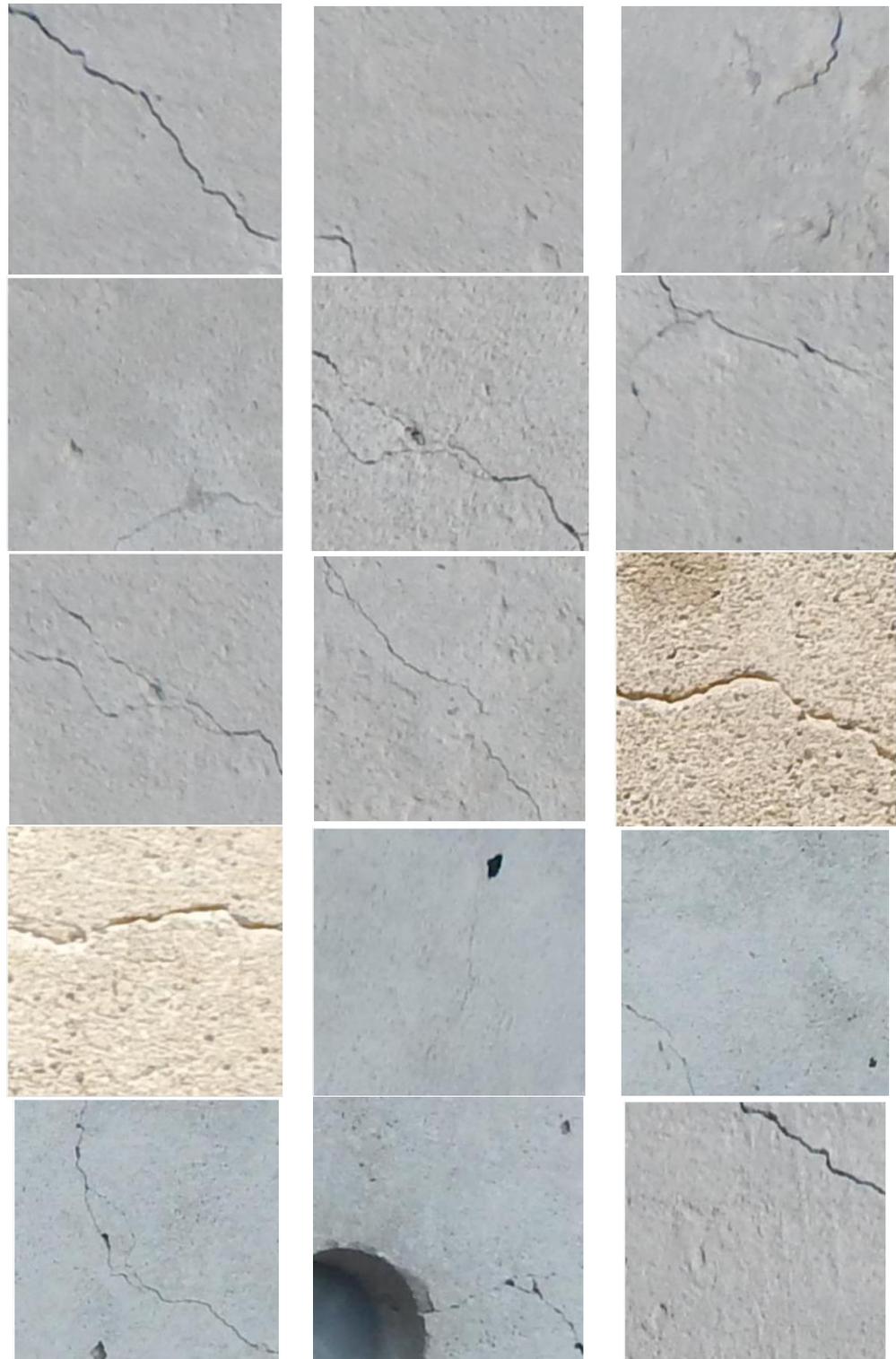


Figure 4. Wall quality dataset with cracks, holes, efflorescence, damp patches, and spall defects.

4. Prescriptive and Predictive Data Analysis of Wall Quality Defects

The pre-training of the wall quality dataset was performed with the ImageNet by substituting the weights and trained with CNN models, such as Conv2D Xception, DenseNet, VGG19, and ResNet models, for extracting the features. The designed base model was fitted to custom layers with various activation layers, such as softplus, softsign, tanh, selu, elu, and exponential, to analyse the performance of identifying the defects in the wall quality. Activation functions play an essential role in developing neural networks. The activation function determines the frequency with which the network structure acquires a training dataset. The activation function at the output layer determines a model's prediction. An activation function is a unit placed at the end or middle of a neural network that determines whether a neuron will be activated. The activation function is a complex nonlinear transformation applied to an input signal. The signal is subsequently processed and provided as an input to the next layer of neurons. Softsign is the activation function of neural networks. Equation (24) provides the mathematical notation for the softsign activation function.

$$\text{softsign}(x) = \frac{(x)}{(1 + |x|)}. \quad (24)$$

The softplus function is a soft equivalent of the ReLU activation function, occasionally used instead of the ReLU in neural networks. Softplus is related to the sigmoid function and is represented by the following mathematical Equation (25).

$$\text{softplus}(x) = \log(1 + e^x). \quad (25)$$

The tanh activation function is a hyperbolic tangent activation function that reflects the sigmoid activation function. The tanh function takes the input as any real value and outputs the value from -1 to 1 . It is represented by the following mathematical Equation (26).

$$\text{tanh}(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}. \quad (26)$$

The scaled exponential linear unit (selu) activation function is implicitly induced with normalization properties. Normalization with selu is conducted as the input value x is less than zero, and the output is the product of x and lambda. When x is zero, the output is equal to 0. If x is less than zero, then the output is the product of lambda and alpha by the x -value minus the alpha value's exponential, multiplied by the lambda value. Equation (27) represents the selu activation function.

$$\text{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leq 0 \end{cases}. \quad (27)$$

The exponential linear unit (Elu) activation function mainly focuses on the positive values, and the alpha value is selected from 0.1 to 0.3, as represented in Equation (28).

$$\text{elu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases}. \quad (28)$$

The exponential activation function indicates the positive-valued function of a real input variable and is represented by Equation (29).

$$\text{exp}(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!}. \quad (29)$$

The model is examined with the activation function discussed above and compiled to analyse performance indices such as step loss, accuracy, validation loss, and validation accuracy. The loss function computes the difference between the actual output of the model and the target outcome. Step loss indicates the loss incurred at each iteration. The

validation loss specifies the loss for the validation dataset, which is part of the training dataset. The wall quality falls under the classification problem, and the loss function is given by Equation (30).

$$Loss = \sum_{i=1}^n |y_{actual} - y_{target}|. \tag{30}$$

Accuracy is the ratio of the number of correct predictions to the total number of predictions and is denoted by Equation (31).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives}. \tag{31}$$

The model was compiled with ten epochs, and prescriptive data analysis was performed by analysing the performance indices for each epoch, as shown in Tables 1–4.

Table 1. Prescriptive data analysis for Conv2D Xception model compilation.

Epochs	Step Loss	Accuracy	Validation Loss	Validation Accuracy
1/10	1.4366	0.4319	0.8906	0.6316
2/10	0.7244	0.7301	0.6696	0.7474
3/10	0.3934	0.8560	0.8088	0.7895
4/10	0.2461	0.9152	0.8654	0.7789
5/10	0.1643	0.9563	0.6566	0.7780
6/10	0.1495	0.9674	0.6766	0.8001
7/10	0.1572	0.9589	0.5416	0.8632
8/10	0.0699	0.9794	0.7579	0.8211
9/10	0.0636	0.9788	0.7390	0.8632
10/10	0.0691	0.9820	1.0562	0.8316

Table 2. Prescriptive data analysis for Conv2D DenseNet model compilation.

Epochs	Step Loss	Accuracy	Validation Loss	Validation Accuracy
1/10	1.4546	0.4219	0.8806	0.6234
2/10	0.7453	0.7201	0.6766	0.7572
3/10	0.3944	0.8350	0.8088	0.7895
4/10	0.3441	0.9262	0.8564	0.7569
5/10	0.1783	0.9673	0.6556	0.7770
6/10	0.1485	0.9774	0.6436	0.8341
7/10	0.1672	0.9689	0.5666	0.8772
8/10	0.0759	0.9884	0.7559	0.8311
9/10	0.0866	0.9898	0.7280	0.8732
10/10	0.0561	0.9920	1.0462	0.8446

Table 3. Prescriptive data analysis for Conv2D VGG19 model compilation.

Epochs	Step Loss	Accuracy	Validation Loss	Validation Accuracy
1/10	1.4556	0.4211	0.8789	0.6324
2/10	0.7463	0.7444	0.6733	0.7452
3/10	0.3954	0.8344	0.8112	0.7555
4/10	0.3461	0.9567	0.8345	0.7669
5/10	0.1744	0.9563	0.6435	0.7760
6/10	0.1465	0.9786	0.6445	0.8871
7/10	0.1678	0.9878	0.5345	0.8342
8/10	0.0766	0.9876	0.7654	0.8321
9/10	0.0882	0.9899	0.7342	0.8652
10/10	0.0555	0.9943	1.0435	0.8896

Table 4. Prescriptive data analysis for Conv2D ResNet model compilation.

Epochs	Step Loss	Accuracy	Validation Loss	Validation Accuracy
1/10	1.5446	0.2319	0.8806	0.6834
2/10	0.4553	0.5701	0.6766	0.7872
3/10	0.9444	0.6350	0.8088	0.7995
4/10	0.4341	0.8862	0.8564	0.7579
5/10	0.7283	0.8673	0.6556	0.7870
6/10	0.4285	0.8774	0.6436	0.8221
7/10	0.6372	0.8689	0.5666	0.8552
8/10	0.7359	0.9384	0.7559	0.8321
9/10	0.0266	0.9798	0.7280	0.8432
10/10	0.0361	0.9880	1.0462	0.8226

The wall quality dataset was pre-trained with the images from ImageNet by substituting the weights and then trained with convolutional neural network models such as Conv2D Xception, DenseNet, VGG19, and ResNet models to extract the essential features from the image. The designed base model was fitted to custom layers with various activation layers, such as softplus, softsign, tanh, selu, elu, and exponential, to analyse the performance defect identification in the wall quality. The performance was analysed in terms of model loss, model accuracy, precision, recall, and F-score, as shown in Tables 5–8 and Figures 5–7. It was observed that the performance of ResNet was comparatively better, with greater accuracy and F-score values, followed by DenseNet, Xception, and VGG19 models. The model performance across different activation layers was studied. It was noted that the models' performance metrics were better with the exponential activation layer than with the other layers. Thus, the Conv2D ResNet model implemented with the exponential activation layer provided an enhanced model for wall quality detection.

Table 5. Performance analysis for Conv2D VGG19 model with various activation layers.

Conv2D VGG19 Model	Precision	Recall	F-Score	Accuracy
Conv2D VGG19—softsign	0.8101	0.8179	0.8127	0.8179
Conv2D VGG19—softplus	0.8196	0.8068	0.8157	0.8169
Conv2D VGG19—selu	0.8255	0.8350	0.8211	0.8350
Conv2D VGG19—elu	0.8496	0.8368	0.8357	0.8469
Conv2D VGG19—tanh	0.8696	0.8684	0.8568	0.8690
Conv2D VGG19—exponential	0.8961	0.8838	0.8868	0.8990

Table 6. Performance analysis for Conv2D Xception model with various activation layers.

Conv2D Xception Model	Precision	Recall	F-Score	Accuracy
Conv2D Xception—softsign	0.8634	0.8656	0.8657	0.8689
Conv2D Xception—softplus	0.8466	0.8489	0.8478	0.8499
Conv2D Xception—selu	0.8566	0.8559	0.8529	0.8550
Conv2D Xception—elu	0.9406	0.9368	0.9357	0.9469
Conv2D Xception—tanh	0.9123	0.9135	0.9168	0.9190
Conv2D Xception—exponential	0.9071	0.9038	0.9056	0.9090

Table 7. Performance analysis for Conv2D DenseNet model with various activation layers.

Conv2D DenseNet Model	Precision	Recall	F-Score	Accuracy
Conv2D DenseNet—softsign	0.8834	0.8756	0.8757	0.8889
Conv2D DenseNet—softplus	0.8566	0.8489	0.8478	0.8499
Conv2D DenseNet—selu	0.8566	0.8569	0.8579	0.8579
Conv2D DenseNet—elu	0.9479	0.9379	0.9257	0.9369
Conv2D DenseNet—tanh	0.9235	0.9188	0.9157	0.9299
Conv2D DenseNet—exponential	0.9123	0.9123	0.9154	0.9157

Table 8. Performance analysis for Conv2D ResNet model with various activation layers.

Conv2D ResNet Model	Precision	Recall	F-Score	Accuracy
Conv2D ResNet—softsign	0.8957	0.8889	0.8888	0.8949
Conv2D ResNet—softplus	0.8661	0.8787	0.8788	0.8799
Conv2D ResNet—selu	0.8586	0.8569	0.8579	0.8579
Conv2D ResNet—elu	0.9348	0.9399	0.9257	0.9379
Conv2D ResNet—tanh	0.9123	0.9166	0.9157	0.9249
Conv2D ResNet—exponential	0.9212	0.9134	0.9978	0.9147

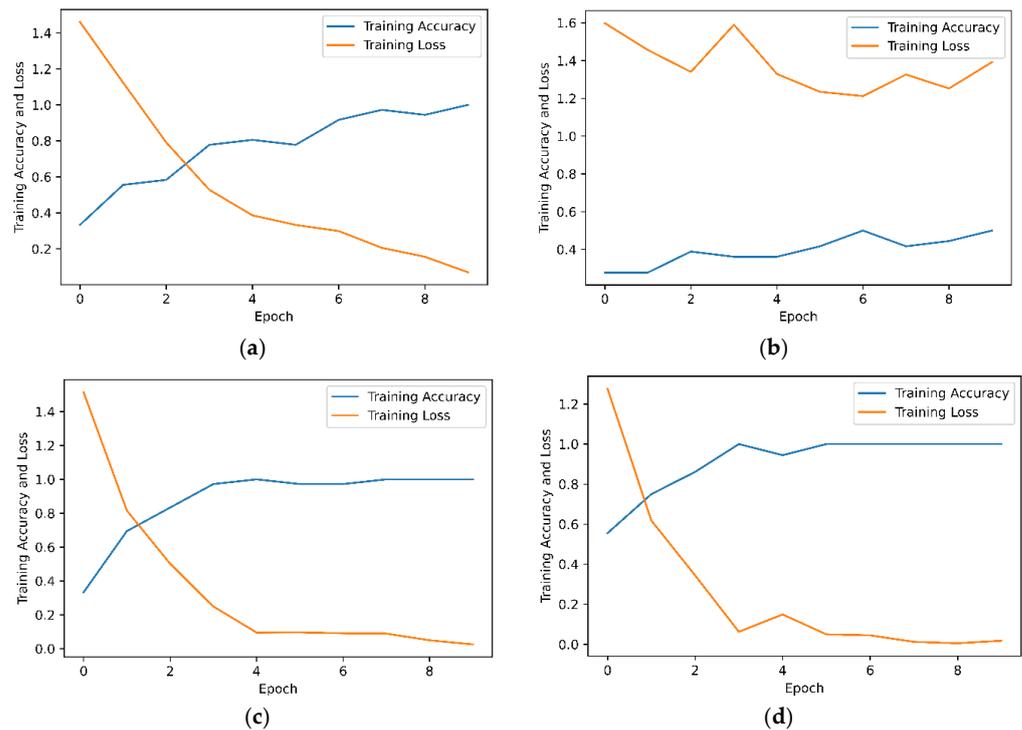


Figure 5. Model loss and accuracy of Conv2D models during training. (a) Xception; (b) VGG19; (c) DenseNet; (d) ResNet.

The performance of the Conv2D Xception, DenseNet, VGG19, and ResNet models using the exponential activation layer was studied using learning curves during the training and validation phases. Figure 5 shows the accuracy and loss curves of all models over each epoch during training phase. The results revealed that accuracy gradually increased over epochs and stabilized after a threshold. The learning loss of the Xception, DenseNet, and ResNet models was more significant during early epochs and dropped as the epoch increased. Figures 6 and 7 showcase the learning curves of the models during the validation phase. It is noted the performance of ResNet exponential model had a consistent increase in accuracy and decrease in loss at each epoch. The other models were found to have declining accuracy at certain epochs.

The wall quality dataset was fitted with the Conv2D ResNet model, which acted as a base model, thereby learning the types of wall defects from 90% of the training data. Now the model had learnt the exact region of interest through which it categorized the wall defect class as cracks, holes, efflorescence, damp patches, or spalls from the more significant number of images. The acquired knowledge from Conv2D ResNet was transferred to refit the model by integrating the layer with the exponential activation function, which identified the wall defects of a single image, thereby validating the transfer learning. When the test image was fitted with the Conv2D ResNet exponential model, the given image was validated with the wall defect class of the actual value and predicted value of the model, which was termed as the True and Guess label, respectively shown in Figures 8–11.

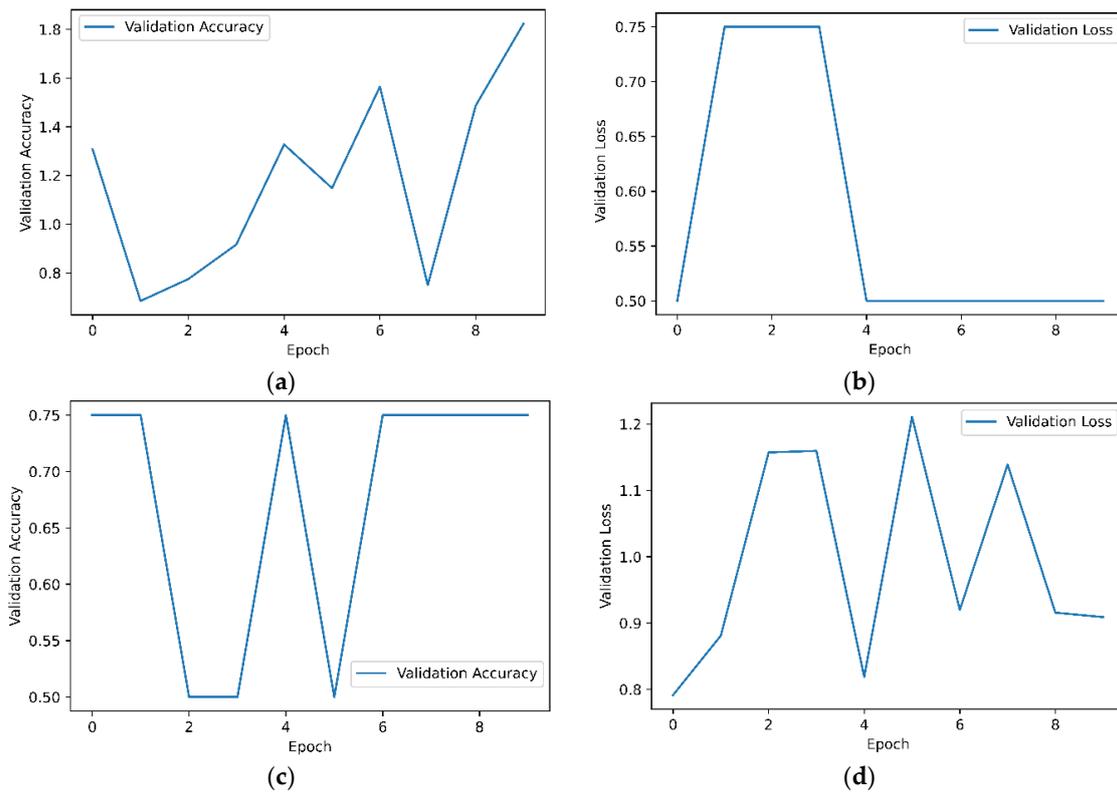


Figure 6. Model accuracy and loss of Conv2D models during cross validation. (a) Model accuracy of Xception; (b) model loss of Xception; (c) model accuracy of VGG19; (d) model loss of VGG19.

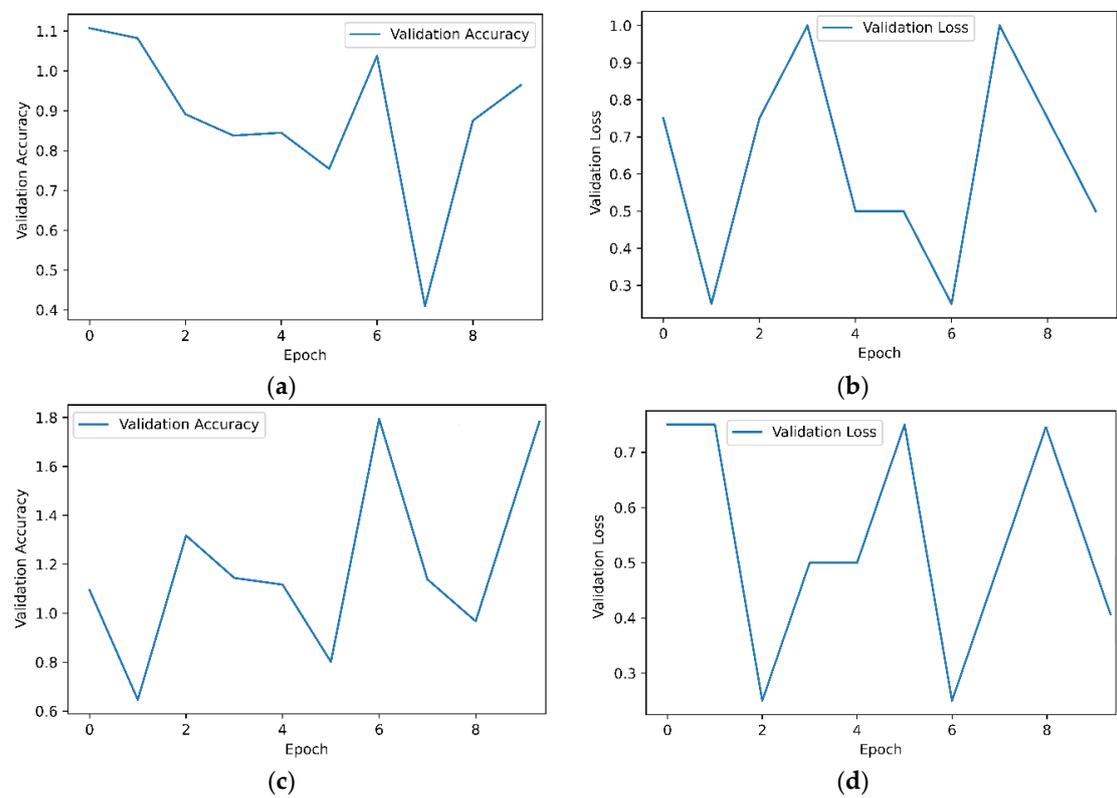


Figure 7. Model loss and accuracy of Conv2D models during cross validation. (a) Model accuracy of DenseNet; (b) model loss of DenseNet; (c) model accuracy of ResNet; (d) model loss of ResNet.

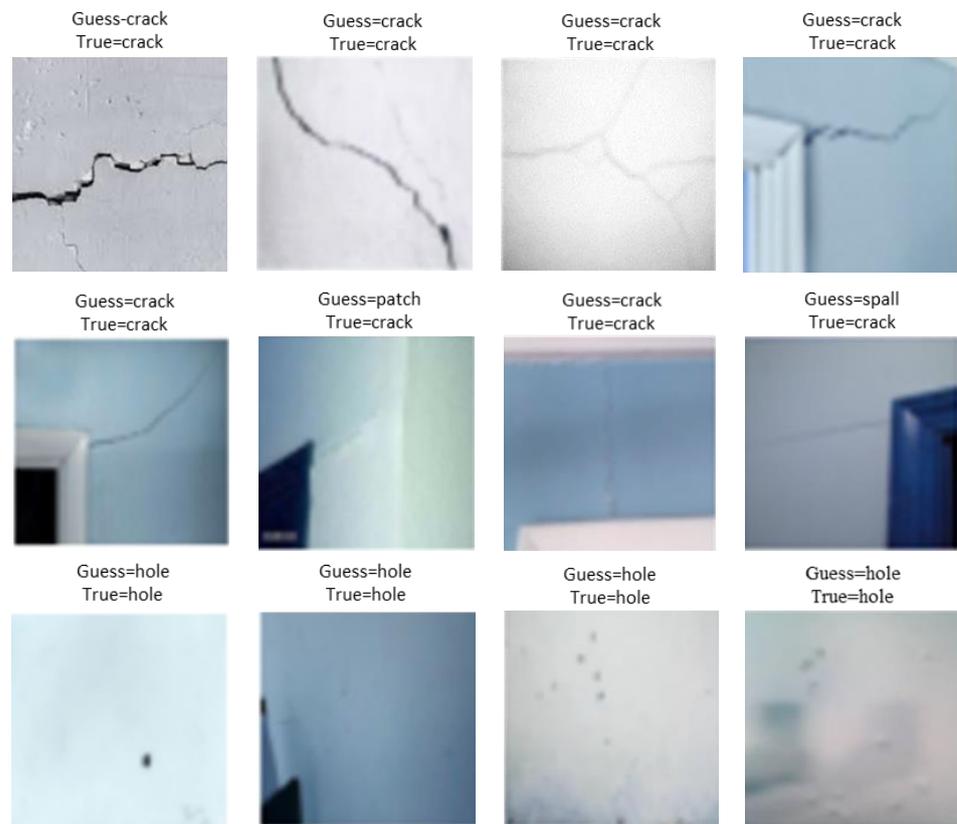


Figure 8. Prediction of wall defects by Conv2D Xception model.

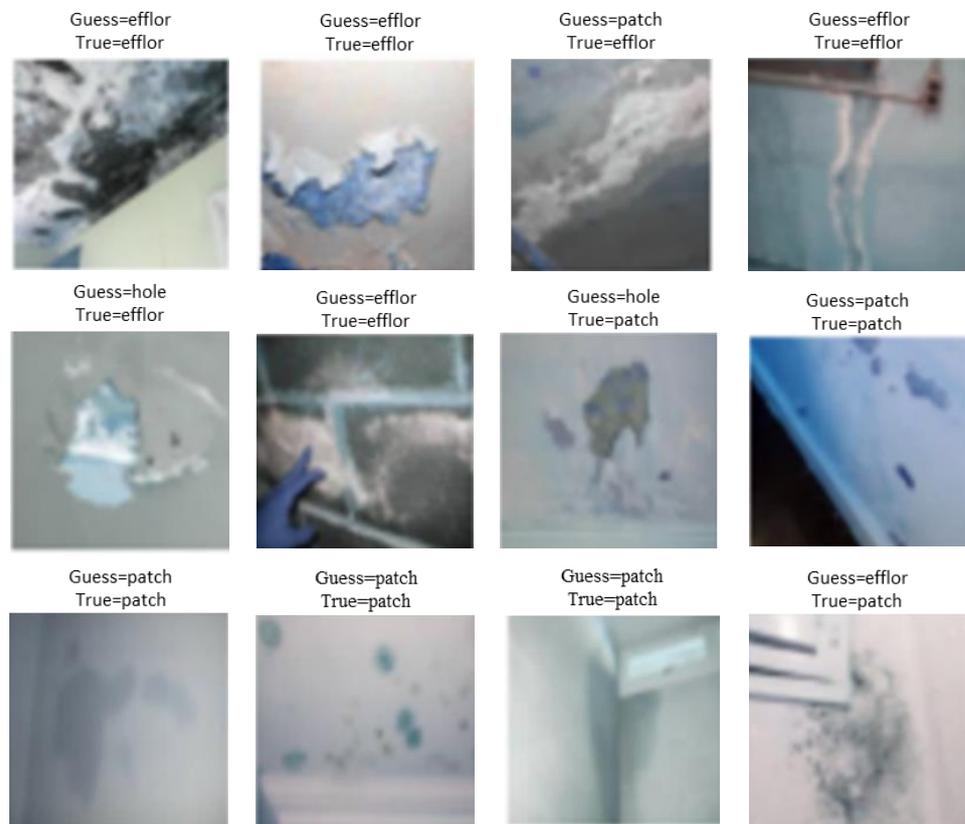


Figure 9. Prediction of wall defects by Conv2D VGG19 base model.

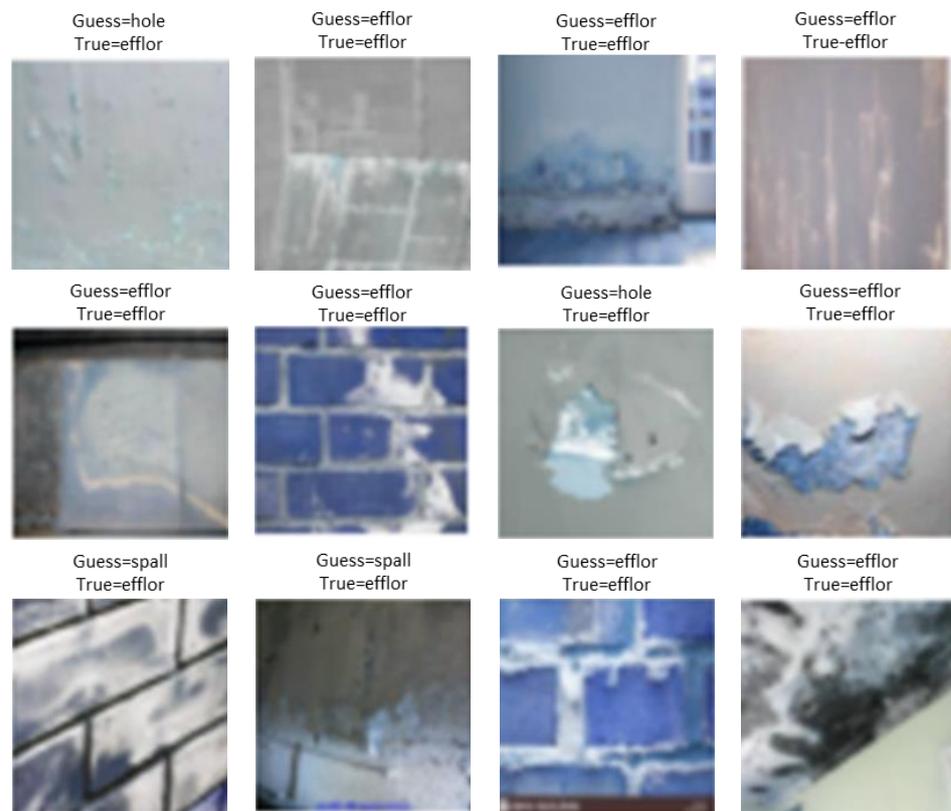


Figure 10. Prediction of wall defects by Conv2D DenseNet base model.

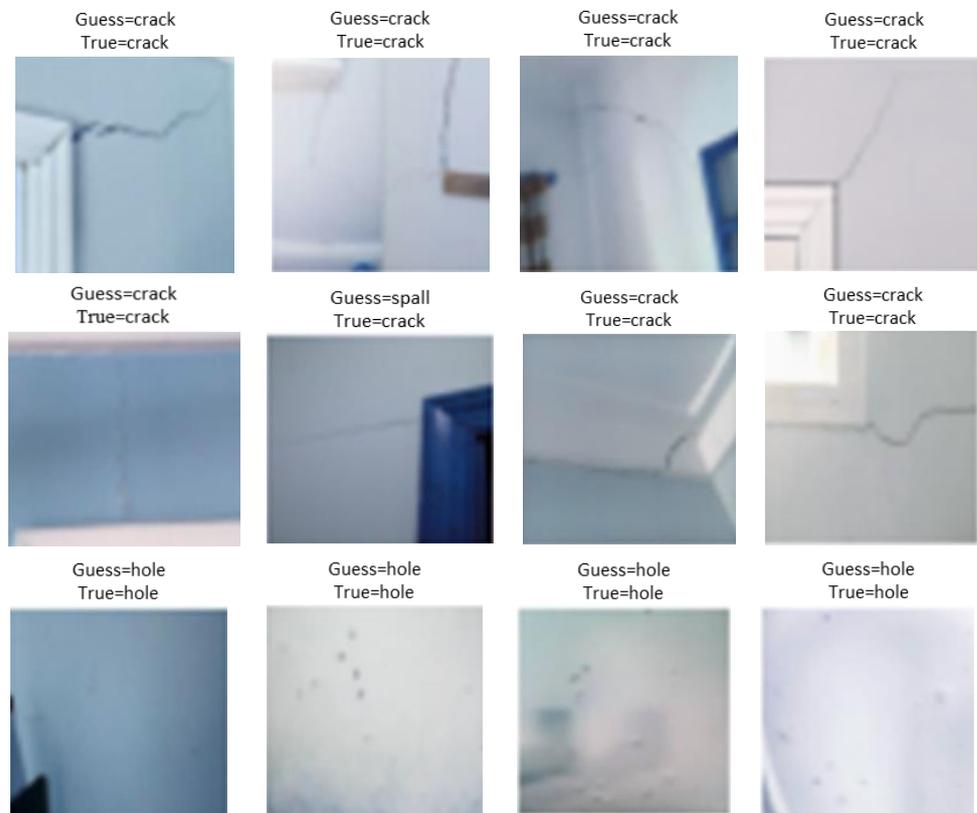


Figure 11. Prediction of wall defects by Conv2D ResNet base model.

Implementation results portrayed that the proposed Conv2D ResNet model with exponential activation layer outperforms with an accuracy of 0.9147, precision of 0.9212, recall of 0.9134, and F-Score of 0.9978 compared with other Conv2D models, such as Xception, VGG19, and DenseNet. The comparative study of learning accuracy during the training and validation phases of the implemented models is represented in Figure 12a,b. The results illustrated the accuracy of each model at every epoch, and it is noticeable that the proposed Conv2D ResNet exponential model steadily gained accuracy at every epoch, whereas the VGG19 model struggled for accuracy during both phases. The DenseNet and Xception models showed instabilities with epochs. The evaluation loss attained by the models is depicted in Figure 12c, which shows that the Conv2D ResNet exponential model had minimal losses at each epoch compared with the DenseNet, Xception and VGG19 models in that order. The Conv2D ResNet exponential model performed efficiently, producing a greater accuracy with minimal loss in wall quality detection.

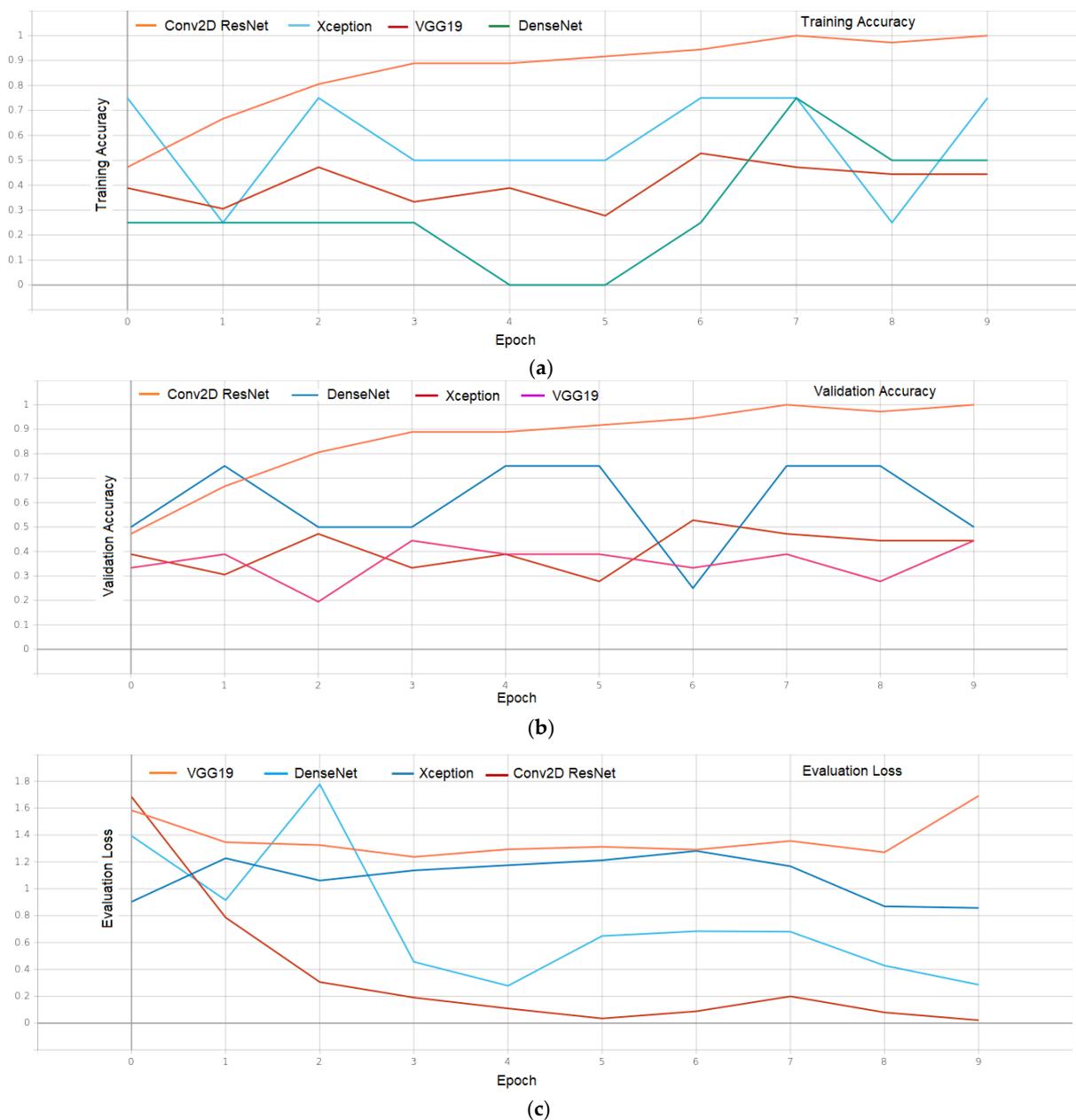


Figure 12. Learning curve comparisons of Conv2D models. (a) Training accuracy; (b) validation accuracy; (c) evaluation loss.

5. Conclusions

This study attempted to analyse the performance of DL models for evaluating the quality of wall structures. This work's main contribution is designing the Conv2D ResNet exponential model-based architecture that classifies wall defects, such as cracks, holes, efflorescence, damp patches, and spalls. A dataset with 5000 images was used to train the proposed model, which achieved the requirements of this research work and outperformed the other Conv2D models. The Conv2D ResNet model with 48 convolution layers, one maxpool, and an average pool layer was implemented in this study. This model served as the base and integrated with the exponential activation layer, improving the classifier's performance in detecting wall defects. The proposed Conv2D ResNet exponential model was further investigated using the performance metrics precision, recall, F-score, and accuracy. The Conv2D ResNet exponential model classified the wall defect type through transfer learning and was also used to analyse the performance of the other CNN model with several activation layers. The wall quality dataset was fitted with the Conv2D ResNet model, which acted as a base model, thereby learning the types of wall defects from 90% of the training data. The acquired knowledge from Conv2D ResNet was transferred to refit the model by integrating the layer with the exponential activation function, which identified the wall defects of a single image, thereby validating the transfer learning by transferring the knowledge from the base model to the custom layer that enhances the accuracy.

This research provides a proper fitting of residual networks to reduce the loss, thereby improving the accuracy, of wall classification with other Conv2D models. The performances of the Xception, VGG19, DenseNet, and ResNet models were fitted with different activation layers such as softplus, softsign, tanh, selu, elu, and exponential, along with transfer learning and analysed using performance evaluation metrics. The dataset used for the proposed Conv2D ResNet exponential model can be used for classifying the defect type in the wall. The same dataset could also be used to identify the defect's depth through object detection methods. However, categorizing the class of defects, such as cracks, holes, efflorescence, damp patches, and spalls in the walls directly related to the characteristics of the wall quality available in the dataset. Once the wall defect class is identified, the respective maintenance procedure can easily be conducted. Implementation results portrayed that the proposed Conv2D ResNet model with exponential activation layer outperforms with an F-Score of 0.997826 compared with other Conv2D models, such as Xception, VGG19, and DenseNet. The potential findings of the proposed Conv2D ResNet exponential model are identifying the appropriate activation layer function that provides the highest accuracy in predicting the type of wall defect. The proposed Conv2D ResNet exponential model improved the overall effectiveness of classifying the wall defects compared with the other deep learning techniques. As an overview of novelty, the Conv2D ResNet exponential model was built over the Conv2D ResNet base model and extended the transfer learning using custom activation layers that effectively classify the wall defects with high accuracy. Despite the Conv2D ResNet exponential model's impressive performance, it is still challenging for researchers to fine-tune the base model hyper-parameters by integrating them with various optimizers. This research's future enhancement focused on validating the accuracy of the wall defect prediction for various combinations of convolutional layers and probabilistic loss functions.

Author Contributions: Conceptualization, B.K. and Y.N.; methodology, K.R.S.P.; software, S.D.M. and K.R.S.P.; validation, B.K., Y.N. and D.-E.L.; formal analysis, A.R. and G.W.; investigation, S.D.M.; resources, A.R. and G.W.; data curation, S.D.M.; writing—original draft preparation, A.R. and K.R.S.P.; writing—review and editing, K.R.S.P. and D.-E.L.; visualization, Y.N. and K.R.S.P.; supervision, Y.N.; project administration, B.K. and D.-E.L.; funding acquisition, B.K. and K.R.S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (NRF-2022H1D3A2A02082296). This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No.2021R1I1A1A01044447).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lacidogna, G.; Piana, G.; Accornero, F.; Carpinteri, A. Multi-technique damage monitoring of concrete beams: Acoustic Emission, Digital Image Correlation, Dynamic Identification. *Constr. Build. Mater.* **2020**, *242*, 118114. [[CrossRef](#)]
2. Liu, P.; Lim, H.; Yang, S.; Sohn, H. Development of a “stick-and-detect” wireless sensor node for fatigue crack detection. *Struct. Health Monit.* **2016**, *16*, 153–163. [[CrossRef](#)]
3. Zhao, S.; Sun, L.; Gao, J.; Wang, J. Uniaxial ACFM detection system for metal crack size estimation using magnetic signature waveform analysis. *Measurement* **2020**, *164*, 108090. [[CrossRef](#)]
4. Yang, X.; Zhou, Z. Design of crack detection system. In Proceedings of the 2017 International Conference on Network and Information Systems for Computers, Shanghai, China, 14–16 April 2017; pp. 147–152. [[CrossRef](#)]
5. Zhang, X.; Wang, K.; Wang, Y.; Shen, Y.; Hu, H. Rail crack detection using acoustic emission technique by joint optimization noise clustering and time window feature detection. *Appl. Acoust.* **2020**, *160*, 107141. [[CrossRef](#)]
6. Amhaz, R.; Chambon, S.; Jerome, I. Automatic Crack Detection on Two-Dimensional Pavement Images: An Algorithm Based on Minimal Path Selection. *Trans. Intell. Transp. Syst.* **2016**, *17*, 2718–2729. [[CrossRef](#)]
7. Amhaz, R.; Chambon, S.; Jerome, I.; Baltazart, V. A new minimal path selection algorithm for automatic crack detection on pavement images. In Proceedings of the 2014 International Conference on Image Processing, Paris, France, 27–30 January 2014. [[CrossRef](#)]
8. Albishi, A.; Ramahi, O.M. Detection of Surface and Subsurface Cracks in Metallic and Non-Metallic Materials Using a Complementary Split-Ring Resonator. *Sensors* **2014**, *14*, 19354–19370. [[CrossRef](#)]
9. Cheon, M.H.; Hong, D.G.; Lee, D. Surface Crack Detection in Concrete Structures Using Image Processing. In *Advances in Intelligent Systems and Computing, Proceedings of the International Conference on Robot Intelligence Technology and Applications 5, Daejeon, Korea, 14–15 December 2017*; Springer: Cham, Switzerland, 2019; p. 751. [[CrossRef](#)]
10. Tedeschi, A.; Benedetto, F. A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices. *Adv. Eng. Inform.* **2017**, *32*, 11–25. [[CrossRef](#)]
11. Sun, H.; Liu, Q.; Fang, L. Research on Fatigue Crack Growth Detection of M(T) Specimen Based on Image Processing Technology. *J. Fail. Anal. Prev.* **2018**, *18*, 1010–1016. [[CrossRef](#)]
12. Kim, B.; Yuvaraj, N.; Sri Preethaa, K.R.; Hu, G.; Lee, D.-E. Wind-Induced Pressure Prediction on Tall Buildings Using Generative Adversarial Imputation Network. *Sensors* **2021**, *21*, 2515. [[CrossRef](#)]
13. Wang, Y.; Huang, Y.; Huang, W. Crack junction detection in pavement image using correlation structure analysis and iterative tensor voting. *IEEE Access* **2019**, *7*, 138094–138109. [[CrossRef](#)]
14. Li, W.; Ju, H.; Susan, L.; Ren, Q. Three-dimensional pavement crack detection algorithm based on two-dimensional empirical mode decomposition. *J. Transp. Eng. Part B Pavements* **2017**, *143*, 2573–5438. [[CrossRef](#)]
15. Kim, B.; Lee, D.B.; Sri Preethaa, K.R.; Hu, G.; Natarajan, Y.; Kwok, K.C.S. Predicting wind flow around buildings using deep learning. *J. Wind. Eng. Ind. Aerodyn.* **2021**, *219*, 104820. [[CrossRef](#)]
16. Kim, B.; Yuvaraj, N.; Sri Preethaa, K.R.; Arun Pandian, R. Surface crack detection using deep learning with shallow CNN architecture for enhanced computation. *Neural Comput. Appl.* **2022**, *33*, 9289–9305. [[CrossRef](#)]
17. Perez, H.; Tah, J.H.M.; Mosavi, A. Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. *Sensors* **2019**, *19*, 3556. [[CrossRef](#)]
18. Zhou, T. Analysis on Construction Quality Control Technology of Reinforced Concrete Shear Wall Structure. *Front. Res. Archit. Eng.* **2018**, *1*, 117–121. [[CrossRef](#)]
19. Spencer, B.F., Jr.; Hoskerea, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222. [[CrossRef](#)]
20. Kumar, P.; Batchu, S.; Swamy, S.N.; Kota, S.R. Real-Time Concrete Damage Detection Using Deep Learning for High Rise Structures. *IEEE Access* **2021**, *9*, 112312–112331. [[CrossRef](#)]
21. Hoang, N.D. Detection of Surface Crack in Building Structure Using Image Processing Technique with an Improved Otsu Method for Image Thresholding. *Adv. Civ. Eng.* **2018**, *2018*, 1–15. [[CrossRef](#)]
22. Lee, B.Y.; Kim, Y.Y.; Yi, S.-T.; Kim, J.-K. Automated image processing technique for detecting and analysing concrete surface cracks. *Struct. Infrastruct. Eng.* **2016**, *9*, 567–577. [[CrossRef](#)]

23. Alam, S.Y.; Loukili, A.; Grondin, F.; Rozière, E. Use of the digital image correlation and acoustic emission technique to study the effect of structural size on cracking of reinforced concrete. *Eng. Fract. Mech.* **2015**, *143*, 17–31. [[CrossRef](#)]
24. Sri Preethaa, K.R.; Sabari, A. Intelligent video analysis for enhanced pedestrian detection by hybrid metaheuristic approach. *Soft Comput.* **2020**, *24*, 12303–12311. [[CrossRef](#)]
25. Talab, A.M.A.; Huang, Z.; Xi, F.; Hai Ming, L. Detection crack in image using Otsu method and multiple filtering in image processing techniques. *Opt.—Int. J. Light Electron Opt.* **2016**, *127*, 1030–1033. [[CrossRef](#)]
26. Mohan, A.; Poobal, S. Crack detection using image processing: A critical review and analysis. *Alex. Eng. J.* **2018**, *57*, 787–798. [[CrossRef](#)]
27. Xu, C.; Gordan, B.; Koopialipoor, M.; Armaghani, D.J.; Tahir, M.M.; Zhang, X. Improving Performance of Retaining Walls Under Dynamic Conditions Developing an Optimized ANN Based on Ant Colony Optimization Technique. *IEEE Access* **2019**, *7*, 94692–94700. [[CrossRef](#)]
28. Kim, B.; Lee, D.-E.; Hu, G.; Natarajan, Y.; Preethaa, S.; Rathinakumar, A.P. Ensemble Machine Learning-Based Approach for Predicting of FRP–Concrete Interfacial Bonding. *Mathematics* **2022**, *10*, 231. [[CrossRef](#)]
29. Yamaguchi, T.; Hashimoto, S. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Mach. Vis. Appl.* **2010**, *21*, 787–809. [[CrossRef](#)]
30. Kim, B.; Yuvaraj, N.; Sri Preethaa, K.R.; Santhosh, R.; Sabari, A. Enhanced pedestrian detection using optimized deep convolution neural network for smart building surveillance. *Soft Comput.* **2020**, *24*, 17081–17092. [[CrossRef](#)]
31. Sharma, S.; Gupta, N.K. A Genetic Approach to Segment and Detect Crack in Rail Track. In Proceedings of the 2019 International Conference on Computing Methodologies and Communication, Erode, India, 27–29 March 2019.
32. Kim, B.; Yuvaraj, N.; Park, H.W.; Sri Preethaa, K.R.; Pandian, R.A.; Lee, D.E. Investigation of steel frame damage based on computer vision and deep learning. *Autom. Constr.* **2021**, *132*, 103941. [[CrossRef](#)]
33. Yusof, N.; Osman, M.; Hussain, Z.; Noor, M. Automated Asphalt Pavement Crack Detection and Classification using Deep Convolution Neural Network. In Proceedings of the 2019 International Conference on Control System, Penang, Malaysia, 29 November–1 December 2019.
34. Yuvaraj, N.; Kim, B.; Sri Preethaa, K.R. Transfer Learning Based Real-Time Crack Detection Using Unmanned Aerial System. *Int. J. High-Rise Build.* **2020**, *9*, 351–360. [[CrossRef](#)]
35. Liao, X.; Li, K.; Zhu, X.; Liu, K.J.R. Robust Detection of Image Operator Chain with Two-Stream Convolutional Neural Network. *IEEE J. Sel. Top. Signal Process.* **2020**, *14*, 955–968. [[CrossRef](#)]
36. Peng, L.; Liao, X.; Chen, M. Resampling parameter estimation via dual-filtering based convolutional neural network. *Multimed. Syst.* **2021**, *27*, 363–370. [[CrossRef](#)]
37. Liao, X.; Peng, J.; Cao, Y. GIFMarking: The robust watermarking for animated GIF based deep learning. *J. Vis. Commun. Image Represent.* **2021**, *79*, 103244. [[CrossRef](#)]
38. Utah State University, University Libraries Home Page. Available online: https://digitalcommons.usu.edu/all_datasets/48/ (accessed on 2 July 2022).