



Automatic concrete slump prediction of concrete batching plant by deep learning

Sarmad Idrees^{a,1}, Joshua Agung Nugraha^{b,1}, Shafaat Tahir^b, Kichang Choi^b, Jongeun Choi^{a,**}, Deug-Hyun Ryu^c, Jung-Hoon Kim^{b,*}

^a School of Mechanical Engineering, Yonsei University, South Korea

^b School of Civil and Environmental Engineering, Yonsei University, South Korea

^c R&D Center, Eugene Corporation, South Korea

ARTICLE INFO

Keywords:

Concrete batching plant
Deep learning
Concrete slump
Quality inspection
Construction safety

ABSTRACT

The workability of fresh concrete is highly important in terms of construction quality and safety. Slump tests are required every 120 m³, yet automated monitoring for each concrete batch remains unavailable in the actual concrete batching plant. To mitigate this issue, we propose an automatic slump prediction method based on the VGG16 neural network by analyzing the video from the final discharge hopper of the batching plant. Additionally, Explainable AI (XAI) is adopted to evaluate and validate our automatic concrete quality inspection approach. Iteratively examining XAI outputs and applying necessary adjustments in data preprocessing helps to achieve better overall performance. The proposed video classification method performed by averaging over the image-level predictions can classify the concrete into four slump classes with an average precision of 85% and an average F1 score of 87%. This demonstrates the possibility of continuous quality evaluation for all concrete produced in the concrete batching plant.

1. Introduction

Over the past few years, machine learning such as deep learning has been one of the highly implemented algorithms in various industries for its high performance. It has also received significant interest in the construction industry as it is capable of automating processes in the construction industry in the whole building lifecycle from planning to maintenance stages (Baduge et al., 2022; Xu et al., 2021; Moein et al., 2023), which improve the performance and productivity of the construction processes. Machine learning has been implemented in many areas of construction such as, cement compressive strength evaluation (Yu et al., 2023), roads damage prediction (Lee et al., 2023), and optimal mix design predictions (Shahrokhishahraki et al., 2024), and retrieving unsafe behavior (Fang et al., 2022).

Numerous studies apply machine learning to concrete-related issues, a prevalent building material. Several focus on methods for detecting cracks in hardened concrete (Ali et al., 2021), predicting compressive strength of specific mixture design (Khan and Abbas, 2023; Han et al.,

2024), predicting shear strength of reinforced concrete (Badra et al., 2022; Chou et al., 2022), and predicting concrete shrinkage (Hilloulin and Tran, 2023). Many machine learning and deep learning approaches have shown promising results and performances for crack detection problems. Deep learning outperforms traditional regression in accurately predicting the complex mechanical properties of concrete by utilizing the existing datasets (Xu et al., 2021). However, the application of machine learning in monitoring the fresh properties of concrete is not widespread.

Concrete is one of the most used materials in the construction industry which contributes to 40–45% of construction materials (Zhang et al., 2022). A concrete batching plant is a facility that mixes dry aggregates, cement, admixtures, and water to produce ready-made concrete mixes and feeds them to truck mixers. The workability of fresh concrete, which attributes to the ease of forming the fresh concrete into a desirable shape without losing the homogeneity and performance, is one of the important aspects that requires quality control. The quality of the fresh concrete produced is basically inspected with conventional

* Corresponding author.

** Corresponding author.

E-mail addresses: jongeunchoi@yonsei.ac.kr (J. Choi), junghoon@yonsei.ac.kr (J.-H. Kim).

¹ These authors contributed equally to this work.

testing methods such as slump and flow tests. For example, in a slump test, fresh concrete is filled into a slump cone, then the cone is turned upside down, and the height of the concrete 'slump' is measured. The fresh properties of concrete are pivotal, influencing numerous aspects of quality and construction, including surface characteristics and form-work pressure. These properties play a critical role in determining the overall performance and durability of concrete structures (Nilimaa, 2022). However, it is not practically possible to test every batch produced because testing is time-consuming and costly as it requires additional human operators. In the USA, the samples for concrete tests should be taken once a day or once every 150 cubic yards of concrete (about 15–20 trucks) according to ACI 318 Section 5.6 (Building Code Requirements for Structural Concrete). According to Korean Construction Specification KCS 14 20 10 about General Concrete, the quality inspection of fresh concrete requirement is once per day or once every 120 m³ (20 trucks). In the construction site, a ready-mixed concrete truck will be rejected if the concrete does not meet the required slump value, which could significantly impact the construction project resources, time and cost due to concrete disposal. Because the moisture content of the dry material in the material hopper and the quality of the individual raw materials change over time, precise dosing control of raw materials in a concrete batching plant cannot always guarantee consistent properties of fresh concrete. The state of each material could change significantly, especially when refilling the bins with new raw materials from different production sites.

This paper introduces a state-of-the-art continuous automated concrete quality control method for predicting concrete slump value after the production process at concrete batching plant. By ensuring the concrete quality at the final discharge hopper of concrete batching plant, the method helps minimize the variability of concrete quality and eliminate the need for manual slump tests. The contributions of the paper are as follows:

- To the best of the authors' knowledge, this is the first automated monitoring technology for predicting the fresh concrete slump value in the process line of concrete batching plant.
- This paper presents an innovative solution for assessing the quality of all concrete produced at batching plants, marking a first in the industry.
- Our approach utilizes transfer learning and predictions from individual frame images for the final output, circumventing the limited training data for video-level models. Moreover, we detail how Explainable AI (XAI) can iteratively refine the model.
- Our method offers a non-contact, inline monitoring system that operates continuously. The deep learning-based slump prediction technique serves as an automated quality sensing tool throughout the production line, with applications extending to quality control. It signifies a pivotal advancement in automation for the ready-mixed concrete sector.

The remainder of the paper is organized as follows: Section 2 provides literature review of fresh concrete slump measurement process and the necessity of the development of automatic slump monitoring technology. Section 3 presents the data acquisition and preparation. Section 4 explains about the deep learning model of Visual Geometry Group 16 (VGG16), prediction mechanism, and accuracy improvement by XAI. Section 5 shows the results and the accuracy for the deep learning model. Lastly, Section 6 concludes the proposed automated method as the viable solution for previously impossible continuous concrete quality monitoring.

2. Literature review

2.1. Quality monitoring for workability of fresh concrete

Concrete with high quality has good workability which includes

deformability, water retention, and segregation resistance (Yang et al., 2021). ACI Standard 116R-90 defines workability as the property of freshly mixed concrete, which determines the ease and homogeneity for the concrete to be mixed, placed, consolidated, and finished. Evaluating concrete workability before casting is crucial to ensuring structural integrity, safety, and construction productivity (Kim and Park, 2018). The most common and widely used test in estimating concrete workability is the slump test which has been used since 1922. The conventional slump test method requires significant resources in terms of time and cost. Operators must conduct the test on-site daily or for every 20 trucks. There is also a chance of measurement errors and wrongly recorded or manipulated after the tests have been conducted. As the slump test is always conducted immediately before placement to ensure that the concrete has good workability, if the concrete has undesirable specifications or properties, the concrete will be disposed and wasted, which will highly impact the construction schedule and cost. Even though the mix design is determined, the wide variety of raw materials may change and make a difference in the workability of the finished concrete product.

In order to address the problem and reduce the number of resources to perform the slump test as concrete quality monitoring process, especially in the big-scale construction project, the implementation of effective autonomous slump test method is needed. Concrete quality monitoring process can be applied on three stages as illustrated in Fig. 1. Those are concrete production stage, concrete delivery stage using ready-mixed concrete truck and construction stage when the concrete slump test is performed. The innovative sensor-based methods have been implemented to monitor the slump value during the delivery stage of concrete to the project site, and inspection stage at the jobsite. For instance, the Center for Transportation Infrastructure and Safety at Missouri University of Science and Technology (Khayat and Libre, 2014) introduced the VERIFI system. This system estimates concrete slump in a truck mixer by using a regression equation that correlates hydraulic pressure sensor values with slump values while the drum rotates. The VERIFI system is advantageous to implement during the delivery stage in the ready-mixed truck since the feedback controller reduce the difference in concrete fresh properties from the concrete batching plant to the construction site by automatically adding water or admixtures to maintain the slump at target value. Malekipour et al. also present the method to measure the ready-mixed concrete slump value during transportation using the parameters of hydraulic pressure, drum rotation speed, and concrete weight inside the truck mixer along with its calibration methods for different trucks (Malekipour and Moodi, 2021). In countries such as South Korea, adding water during concrete delivery is prohibited by regulations, rendering active slump control systems during delivery impractical. The fundamental assumption at the delivery stage is that high-quality concrete should be received from the concrete batching plant, and the primary objective of quality control at this stage is to maintain the slump value with minimal loss until it reaches the construction site. Adjusting the quality at the delivery stage by receiving low-quality concrete from the concrete batching plant implies the use of excessive water and admixtures, which can increase the likelihood of concrete segregation at the construction site and may not achieve the required concrete strength. At the construction stage, concrete workability is generally managed using conventional slump test. Workers manually measure and record slump height or flow diameter using a ruler and a stopwatch, but there is a risk of easy manipulation, and a shortage of inspection personnel makes it impossible to conduct comprehensive inspections for all incoming concrete at job sites. In response to this need, a new digital system with advanced algorithms has been developed to digitally perform the slump test or flow test using a depth sensor (Kim and Park, 2018) or stereo camera (Tuan et al., 2021). This non-contact sensing method, which does not rely on human measurements, can be used for on-site inspections, helping to prevent the use of low-quality concrete at the construction stage. However, in order to fundamentally ensure construction safety, prevent resource

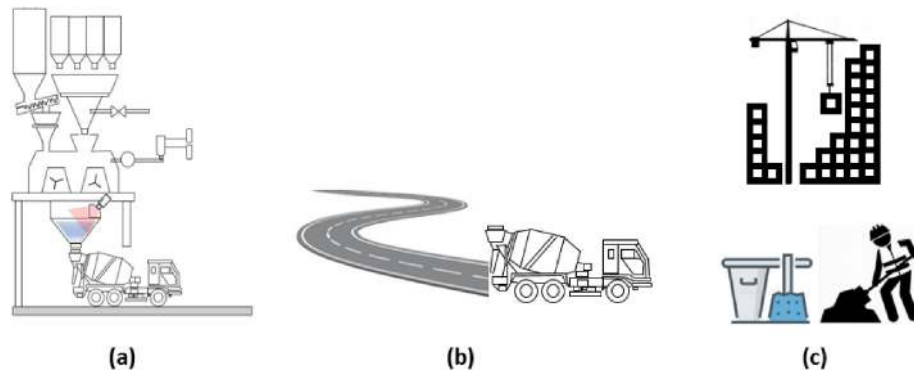


Fig. 1. Three stages of concrete quality monitoring: (a) Concrete production stage, (b) Concrete delivery (in-transit) stage using ready-mixed truck, and (c) Construction stage (Workability test).

waste, and reduce economic and social costs, it is most important to secure quality through comprehensive inspection at the concrete production stage. Automated digital technologies and systems have been identified as crucial strategies in the development of smart and sustainable concrete construction practices. These innovations are central to enhancing efficiency and reducing environmental impact in the construction industry (Nilimaa et al., 2023). Therefore, this paper focuses on the development of a new technology capable of automatically monitoring concrete slump during the concrete production stage.

2.2. Monitoring of concrete slump at the concrete production stage

Image processing and deep learning techniques can be applied at the concrete production stage, using visual data from cameras attached to processing equipment like mixers or hoppers (see Fig. 2) to estimate and analyze concrete's fresh properties. Juez et al. (2017) proposed an image analysis technique to detect the various phases of the concrete during the mixing process, in which a digital GoPro camera was installed on top of a laboratory-scale mixer to take pictures of the concrete surface. Yang et al. (2021) proposed a deep learning-based method to predict the slump flow values of concrete during the mixing process, where the mixing process videos were recorded by placing the camera in the mixing station. The workability of fresh concrete was predicted by implementing a Long Short-Term Memory (LSTM) based model at a mixing station where the dynamic flow agitated by the rotating blades of the mixer was observed. Although previous studies have demonstrated the feasibility of monitoring of fresh properties during the concrete production stage by installing a camera on a mixer in a laboratory environment (Yang et al., 2021; Juez et al., 2017), but in an actual concrete batching plant where the batches of concrete are produced all day long, it is practically difficult to install a camera on the mixer because the camera is easily contaminated. Dust particles generated during the material supply process and fresh concrete splashed by the rotating blade during the mixing process easily contaminate the camera,

resulting in frequent cleaning of the camera lens or protective glass. This maintenance problem can mitigate the effectiveness of introducing an automated monitoring system in the concrete batching plant. Additionally, quality trends in video can be easily affected in such a dusty environment of a mixer, which can affect the predictive accuracy of machine learning in a real-world environment.

In a concrete batch plant, the best place to install a camera to predict the slump of the concrete without worrying about contamination is the discharge hopper. After the concrete is mixed, it is transferred to the discharge hopper by opening of the valve. When the truck mixer arrives, the valve at the bottom of the discharge hopper is opened to discharge the fresh concrete. This paper focuses on the prediction of concrete slump from the video stream while the concrete is discharged from this hopper. This is a challenging topic as the difference between the slump classes cannot be visibly examined by human eyes. The application of deep learning techniques to predict concrete slump in the discharge hopper of a batching plant is a crucial technology that enables quality inspection for all produced fresh concrete.

3. Materials and methods

The study concentrated on developing a deep learning model to classify concrete slump values in a batching plant using video data, as depicted in Fig. 2. A digital camera was installed at the discharge hopper to capture the concrete flow. The dataset consisted of videos categorized into four slump classes. Transfer learning with the VGG16 architecture was employed due to limited data availability. Model performance was enhanced iteratively using Explainable AI (XAI) techniques, such as Class Activation Mapping (CAM), to refine data preprocessing. The final model achieved higher accuracy by focusing on informative features while disregarding irrelevant elements. Prediction at the Video-level with voting system surpassed image-level classification in accuracy, offering a more effective method for concrete slump classification. The study's findings suggest the effectiveness of deep learning models in



Fig. 2. Various Images were acquired from the placement of the camera in the concrete batching plant to monitor the fresh concrete: (a) Images acquired by the camera at the mixer and (b) Images acquired by the camera at the discharge hopper.

automating concrete quality assessment processes in industrial settings, as depicted in Fig. 5.

3.1. Data

3.1.1. Data acquisition

The input dataset for the deep learning method is obtained using a digital camera, DC-S4236THRX (IDIS, South Korea), which is installed at the top of the final discharge hopper on the concrete batching plant. The camera offers a resolution of 1920×1080 pixels and features an IP67 protection rating. Although previous studies have been conducted with a camera placed on top of the laboratory scale mixer (Yang et al., 2021; Juez et al., 2017), as raw material is poured during the mixing process of the concrete batching plant, there is a high amount of dust particles that could disrupt the quality of the recordings and making it difficult to apply the deep learning method. Therefore, as the first deep learning method applied to the actual concrete batching plant, we propose a way to place the camera on top of the hopper with a cleaner environment and better recording.

The camera will record the characteristics and flow of fresh concrete as a video stream while the concrete is discharged from this hopper into a ready-mixed concrete truck. The camera is installed at a specific location with an opening gate to protect it, ensuring that the flow and particles of fresh concrete do not easily affect the quality of the video obtained. The schematic of the concrete batching plant, the discharge hopper, and the camera can be illustrated in Fig. 3. The video-stream is recorded for around 10 s during the discharging process of the fresh concrete from the hopper.

The videos obtained from the Eugene Corporation's Cheonan concrete batching plant in South Korea are recorded in two separate days with the various distribution of slump values. Each video is categorized and tagged based on the production specification code into four slump classes, viz.: S80, S120, S150, and S180. In particular, S80 represents a concrete mixture with an 80 mm slump value and a lower water-to-cement ratio, while S180 represents a concrete mixture with a 180 mm slump value and a higher water-to-cement ratio. According to BS EN 12350-2, slump grades based on the slump test are divided into four categories, ranging from S1 to S4 (BS EN and 12350-2, 2019). In South Korea, the specifications for ready-mixed concrete are typically provided in accordance with KS F 4009 (KS F 4009, 2016). In the category of Ordinary Concrete with the maximum aggregate size of 25 mm, slump values are typically specified as 80 mm, 100 mm, 120 mm, 150 mm, and 180 mm. Among these, the predominant slump specifications produced at concrete batching plants are 80 mm, 120 mm, 150 mm, and 180 mm. Fig. 4 shows examples of four slump classes of concrete material contained in a final discharge hopper. When carefully examining the snapshots of videos from four slump values, it is apparent that S80

concrete aggregate has the roughest surface and less water content than the other slump values, while S180 has the smoothest surface and most water content. From the video, we can see the fresh concrete with visibly well-mixed cement paste and aggregates, along with hopper wall, and level bars, which indicated the fresh concrete level during the discharging process. Distinguishing between these four slump values visually is extremely challenging for a person. Therefore, in this paper, we strive to build a deep learning model that classifies the fresh concrete into one of four slump classes.

3.1.2. Data preparation

We divide the dataset on the basis of videos rather than images, into training and testing sets as shown in Table 1. The dataset contains a total of 208 recordings for four slump values (i.e., S80, S120, S150, and S180), and each video sequence is approximately 10 s. Table 1 represents the distribution of concrete specifications produced over two days. Here, it can be observed that the number of data points for the 80 mm slump value is relatively small compared to other data. This is due to the infrequent use of concrete with low flowability at construction sites because it is difficult to pour, and it requires more effort to remove air voids with a vibrator. Images are then generated from the video by sampling it with 10 frames per second. Hence the total number of corresponding images are approximately 20,800. The test video sequences were never exposed to the network during the training phase. The images were resized to $224 \times 224 \times 3$ to adhere to the input dimension specifications of the pre-trained VGG16 network. This resizing is performed even after cropping the images for model refinement, a decision guided by the analysis of XAI results, as discussed in Section 5. In addition, random horizontal flipping was also performed to increase the number of images as data augmentation.

3.2. Methods

3.2.1. Deep learning model structure

Deep learning is a branch of machine learning based on artificial neural network algorithms inspired by how the human brain is organized and functions. In an image-based deep learning approach, it is essential to learn the feature representation capable of extracting information from the images to generalize the data distribution (Alzubaidi et al., 2021). Convolutional Neural Networks (CNNs) are considered the preferred choice to execute the feature extraction process for image dataset due to their ability to extract meaningful visual information. Various CNN architectures have been developed to achieve better performance, including Visual Geometry Group (VGG) (Simonyan and Zisserman, 2014), which has models like VGG16 comprising 16 convolutional layers, ResNet (He et al., 2015a), GoogleNet (Szegedy et al., 2015), DenseNet (Huang et al., 2017), and more. Each backbone

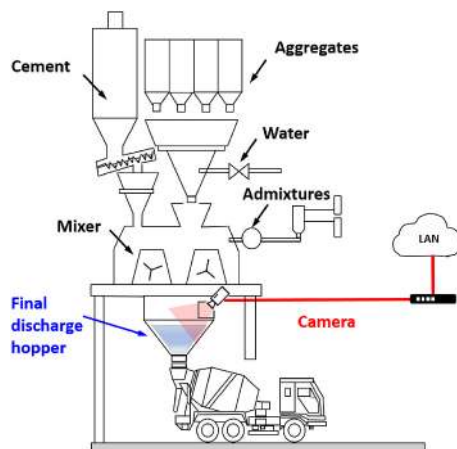


Fig. 3. The schematic and photo of concrete batching plant (camera placement in the discharge hopper).

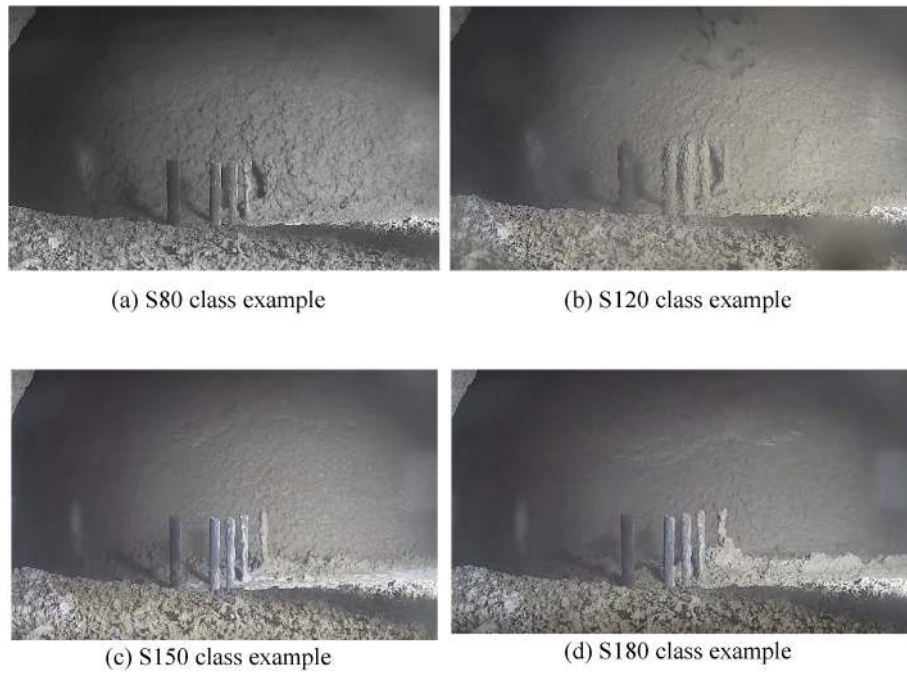


Fig. 4. Examples of four slump classes of concrete material produced in a concrete batching plant: (a) S80 class, (b) S120 class, (c) S150 class, and (d) S180 class.

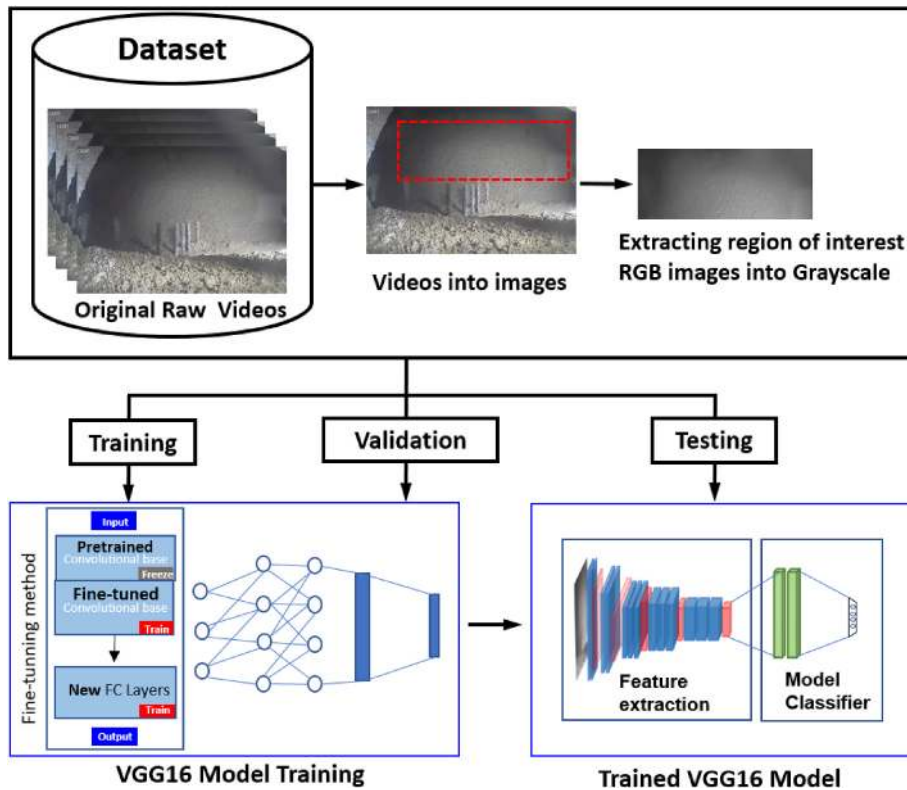


Fig. 5. VGG16 model architecture for predicting the class of concrete slump.

architecture differs in terms of parameter size, computational complexity, and efficiency. To date, the capability to perform large-scale image recognition tasks is considered as effective as the newer and denser convolutional networks (Simonyan and Zisserman, 2014; Elharrouss et al., 2022).

In traditional deep learning, a vast of dataset is required for better generalization of data distribution. Transfer learning is able to reduce

the dependence on large target domain datasets and massive computational resources by transferring the knowledge from the source domain to a target domain to improve performance on a target domain task (Ribani and Marengoni, 2019). Transfer learning has been adopted in various application-based vision tasks with a limited amount of training data. In addition, it has been widely used in the construction domain for diverse applications such as the estimation of time dependent strength of

Table 1

Dataset training and test distribution.

Slump value	Training set videos	Test set videos	Total videos
S80	18	9	27
S120	54	24	78
S150	44	11	55
S180	37	11	48
Total	153	55	208

specific soil (Kang et al., 2024), detection of several kinds of defects in sewers (Situ et al., 2023), and object detection in construction sites for visible dust (Wang et al., 2023). The transfer learning technique is applied with VGG16 to deal with the limited data by exploiting the features extracted from the large publicly available image dataset to reduce the training time with better overall accuracy.

In this study, we propose the VGG16 network as the backbone architecture for feature extraction and image classification tasks. VGG16 is a convolutional neural network consisting of 16 layers with trainable parameters and other layers, such as the max pool layer (see Table 2). VGG16 architecture primarily consists of three types of layers: a stack of convolution layers using various filters to extract features from the images, a max-pooling layer to decrease the size of the image by extracting important features, a flattened layer to create 1D tensors from batches of features, and lastly three fully connected layers in which the first 2 fully connected layers have dense units of 4096 neurons and a fully connected softmax layer. In all networks, the fully connected layers have the same configuration. The softmax activation is used in the final classification layer to forecast the probability of each class.

The overall procedure for estimating the concrete slump for a test video on concrete discharge at a concrete batching plant is shown in Fig. 5. The images are first extracted from the original video at a rate of 10 per second. In the initial stage, the region of interest is selected, excluding parts that appeared irrelevant to the observation of fresh concrete properties. After cropping the region of interest, grayscale version of the RGB image is created. The dataset is then divided into three parts: training, validation, and testing. The VGG16 model is imported, modified, and evaluated after being trained using the dataset. Pre-trained weights for the VGG16 architecture are frozen before model training. The model architecture is fine-tuned, the new fully linked layers are added, and the future model inference and temporal results are produced when the optimal model is synthesized. Although we trained the network using images as a part of our model, our objective is to predict the concrete slump class for an entire video sequence.

3.2.2. Transfer learning for concrete slump classification

Training convolutional neural networks from scratch is particularly challenging due to the substantial data requirements and the lengthy training time involved. To address this challenge and the limitations posed by the limited number of datasets in training the network to predict slump values, we have employed transfer learning in this study. Transfer learning is a machine learning technique that leverages knowledge gained while solving one problem to address another related problem. It operates by freezing certain parameters, altering the output layer, and fine-tuning weights, thus enabling the direct utilization of a well-trained model. The convolutional layer weights are initialized from a publicly available pre-trained network for classification. The

convolutional layers were initialized with weights trained on the ImageNet dataset. The last two fully connected layers were initialized with random weights (i.e., Kaiming initialization (He et al., 2015b)) to learn better representation of our target domain dataset. Kaiming initialization was used because it generates an initialization strategy that enables deep models to converge by carefully examining the non-linearity of rectified linear units (ReLU). Subsequently, the entire network was trained with our dataset without freezing any network layer. Our model was trained with a learning rate of 0.005, an image batch size of 32, and cross-entropy as the loss function. Stochastic gradient descent was used as an optimizer, and the model was trained for 15 epochs.

3.2.3. Prediction of concrete slump class for a video via voting

The selected VGG16 network can be divided extractor and classifier. The last fully connected layers are referred to as a classifier, which learns to map the extracted high-level features to individual classes. Thus, the last fully connected layers significantly improve the network's classification accuracy. After extracting features from the backbone architecture, a classifier is implemented to map the extracted features to desired outcome results. This classifier can be a logistic regression classifier, support vector machines (SVM), random forest classifier, or a neural network. In this paper, we adopted a feedforward network as a classifier to predict the slump value. In image-based classification, the accuracy is determined for each image regardless of its video class. On the other hand, video-based classification involves a voting process where predictions from image-level classifications within a single video sequence are averaged (see Fig. 6). To elaborate, we select a video sequence from the test set and perform slump value classification on each image within that sequence. The class that receives the highest number of predicted slump class votes from the image-level frames within the video is considered the predicted class for that particular video.

3.2.4. Model improvement by the iterative evaluation using XAI

It is important to improve the accuracy and preprocessing techniques in an iterative way during the model validation as shown in Fig. 7. In our approach, an Explainable AI (XAI) method called Class Activation Mapping (CAM) (Zhou et al., 2016) was utilized for validation after the VGG16 model was trained on the training dataset. The trained dataset is validated against the validation dataset after each iteration to demonstrate the accuracy of each iteration. Data processing operations were then conducted to extract more useful information from the dataset. XAI delivers helpful information that can be checked against the domain knowledge for adjustments of the preprocessing during the data processing to improve the workflow and provide desired results.

4. Experiment and results

4.1. Model validation and improvement using XAI

The CAM utilizes the activation maps from a CNN layer for the particular test example to identify the image regions contributing to class prediction. These class activations are projected as a heatmap onto the provided input image to determine the possible image pixels responsible in predicting the particular output. The activation values from the last CNN layer are extracted from our backbone architecture, i.e., VGG16. In this paper, we perform CAM for the purpose of model validation and improvement (see Fig. 8).

Fig. 8 shows the XAI CAM result for trained network at an early stage in the model improvement process. Although the level-bars on the hopper were installed to assist in identifying the height of the concrete mix in the hopper, the CAM results indicate that the model is actually more focusing on the level-bars instead of extracting information from the concrete mix in the hopper. Training with the level-bars, the model may rely on them as a form of clever Hans predictor (Lapuschkin et al., 2019). It is then possible that the concrete mix residual from the

Table 2

VGG16 network model parameters.

Layer type	Number of layers
Convolution layer	13
Max Pooling	5
Fully Connected	3
Dropout	2

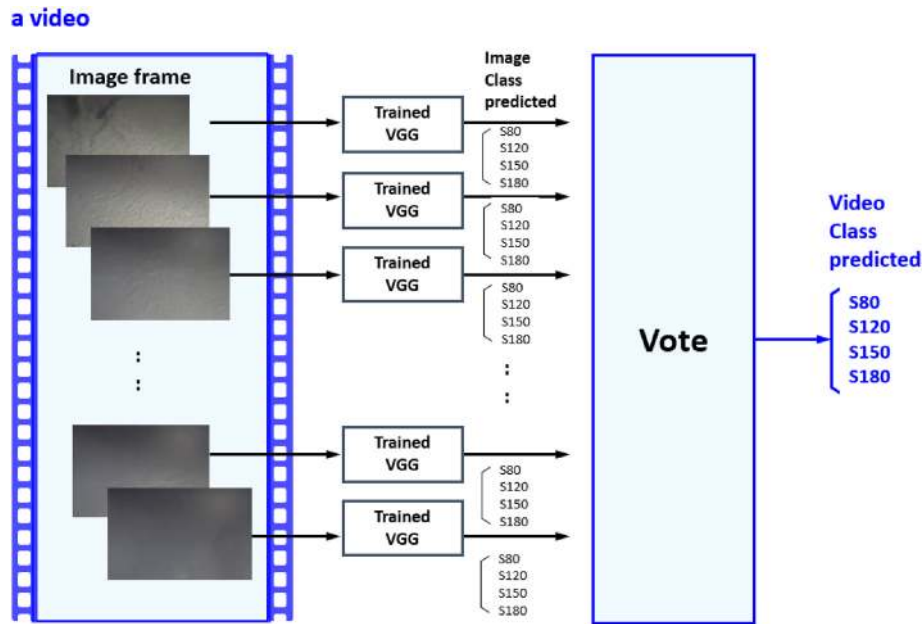


Fig. 6. Prediction of the concrete slump class for a discharge video (a group of images on the left side) via voting over predicted classes of image-level frames.

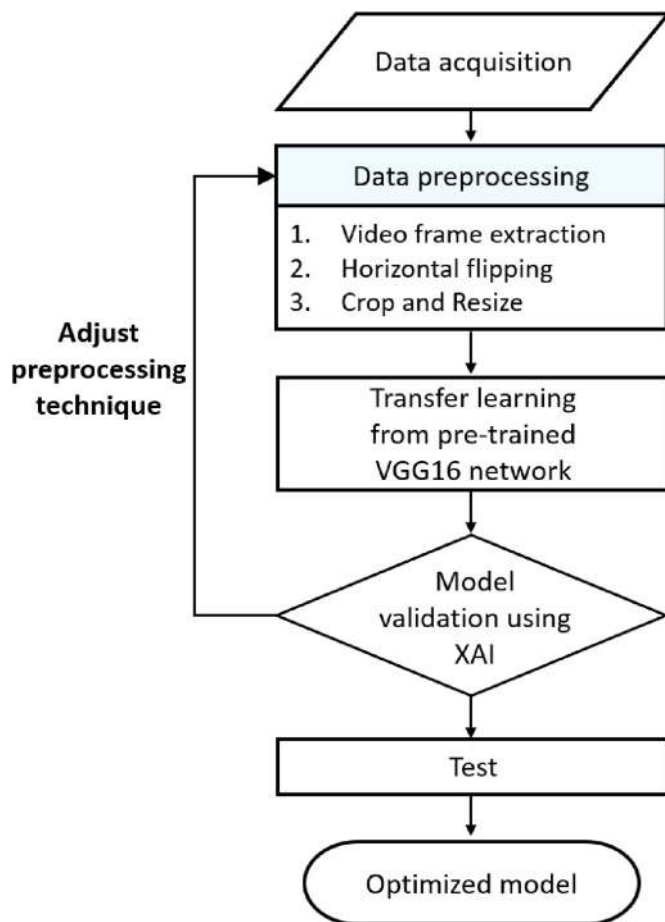


Fig. 7. Model improvement process using XAI for better slump value prediction.

previous concrete batch still remain on the level bars inside the hopper during the production of the next batch of fresh concrete. Hence, this concrete residual could influence the model to predict the false slump

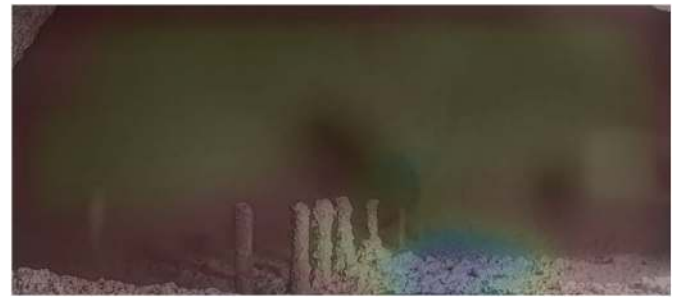


Fig. 8. XAI CAM results for full-image model.

values. The CAM results for the full images model confirm our hypothesis of misleading level bars.

After observing the CAM results, we decided that the dataset should be preprocessed to crop the level bars off and includes only the fresh concrete mix as shown in Fig. 9. The network was again trained with the cropped versions of images excluding the level bars to improve the overall accuracy. Fig. 10 shows the XAI CAM result for re-trained network after adjustment of preprocessing. It can be seen on the heatmap that the model now focuses on the characteristic features of the concrete mix surface on the hopper. As confirmed by the CAM results, the level bars can influence the network to predict the incorrect slump value. The CAM images justify the effectiveness of the image cropping process, as it demonstrates that the model exploits the fresh mix to predict the slump value after training with the cropped images. Hence it is important to crop areas that can be considered static features, such as concrete residues. Section 5.2 introduces the detailed analysis of the



Fig. 9. Cropped images utilized for training improved model.



Fig. 10. XAI CAM results for cropped image model.

results of accuracy before and after the application of XAI.

4.2. Performance evaluation results

The performance comparison results with other backbone architectures, such as ResNet, DenseNet, MobileNet, and ShuffleNet, are outlined in Table 3. Our approach, which integrates transfer learning, a voting strategy for predicting slump values in videos, and enhancements informed by XAI, achieved an exceptional accuracy of 85.45%. Surpassing the next best performing architecture, MobileNet, by a substantial margin of 12.34 percentage points. Notably, even a simple backbone architecture like VGG16 demonstrates superior performance with our approach compared to more recent and advanced backbone architectures. Additionally, we conducted an ablation study to validate our approach of utilizing adjusted images based on evidence from XAI experiments (see Table 4). The results of the ablation study confirm the effectiveness of our approach, which involves transfer learning, adjusted images from XAI, and a voting strategy, in achieving high video-level accuracy.

Furthermore, the model performance over time for a single video recording were tracked in a temporal fashion by collecting images from the testing dataset after the model had been trained and feeding each image into the model for inference separately. The results from the VGG16 model were then plotted on the graph to present the predicted slump values of images on the y-axis over the increasing frame number on the x-axis. The temporal results for four concrete discharge videos with different slump values are depicted in Fig. 11, where each point on the graph represents the predicted slump value for an image frame extracted from the respective video. The temporal result of slump value S80 is shown in Fig. 11(a), where the temporal frame accuracy is almost 95%. The temporal result of slump value S120 is displayed in Fig. 11(b), with a temporal frame accuracy of approximately 65%, and some of the images were falsely predicted as S150 and S80. The temporal result of slump value S150, with temporal frame accuracy around 85%, is shown in Fig. 11(c). The temporal result of slump value S180, whose temporal frame accuracy is around 85%, is shown in Fig. 11(d).

The inaccuracy in predicting the slump class may result from the blurriness of the recordings and there is also some water splashing and dropping near the vision of the camera on some of the recordings so it may affect the recognition performance of the model.

Classification model performance metrics are used to evaluate how well they perform in a specific situation. Precision, recall, and F1-score are among these key metrics. The model precision score is a measure of the proportion of labels that were correctly predicted or positively predicted divided by the total predicted positives, reflecting the model's

Table 3

Performance comparison results with other backbone architectures.

Methods	Test accuracy
ResNet	71.04%
DenseNet	68.75%
MobileNet	73.11%
ShuffleNet	66.67%
Our method (VGG16)	85.45%

Table 4

Ablated performance results for our approach.

Method	Transfer learning	Image-level accuracy	Video-level accuracy
VGG16 (full images)	×	65.34%	72.67%
VGG16 (full images)	✓	69.28%	73.33%
VGG16 (cropped images)	×	75.32%	76.36%
VGG16 (cropped images)	✓	79.52%	85.45%

accuracy or quality in identifying positive instance. Recall calculates how many of the actual positives our model captures by labeling it as positive (true positive) for the specific class. F-score is a crucial metric in machine learning that harmonizes Precision and Recall to measure a model's accuracy effectively. The F1-score is usually used as a single figure that provides comprehensive information about the quality of the model's output.

Tables 5 and 6 display performance metrics for image-based classification and video-based classification for each of individual slump classes. By examining precision, recall, and F1 scores, it can be observed that the proposed video classification method shows an overall improvement compared to image classification. Specifically, the average precision for video classification is higher at 0.85 compared to 0.79 for image classification. Similarly, the average recall for video classification is 0.90, significantly surpassing the 0.81 average recall for image classification. The F1 score, which is the harmonic average of precision and recall, also reflects this enhancement, with an average of 0.87 for video classification compared to 0.80 for image classification. The F1 scores achieved by the proposed video classification method for the respective classes are 0.87, 0.79, 0.92, and 0.88, surpassing those of image classification, which are 0.83, 0.76, 0.80, and 0.79, respectively. This indicates that the video classification approach yields better performance across all examined classes. Looking at individual classes, the recall for both the S150 and S180 classes has been significantly improved in video classification, achieving perfect scores of 1, compared to 0.86 for each in image classification. Particularly for the S150 class, there is also a notable improvement in precision, contributing to the highest F1 score of 0.92 in video classification.

For both image and video classifications, the S120 class consistently exhibits the highest precision but lowest recall among all classes. This pattern occurs because the model tends to learn a bias towards S120, a majority class that is overrepresented in the data distribution. Consequently, the model accurately identifies S120 instances (high precision), but tends to misclassify instances of other classes as S120, resulting in failure to detect actual positive instances of those classes (low recall). However, it is not advisable to artificially alter the dataset distribution by reducing the number of instances to match minority classes, such as S80. Although this may equalize the model's performance across classes, it severely limits the total amount of data, compromising the model's ability to generalize and ultimately reducing overall performance. The data distribution used in this paper represents the data distribution in the real world, reflecting the concrete production statistics of a batch plant according to industry demand. In the future, securing more extensive data collection for each slump class can lead to a balanced dataset. This would mitigate the current issues, allowing the model to generalize the characteristics of each slump class more effectively and representatively.

In summary, the video-level prediction method exhibits superior performance across all metrics when compared to image classification. This suggests that video-level prediction offers a more robust approach for concrete slump classification tasks in the concrete batching plant. This improvement in performance is achieved by averaging out uncertainty in the image-level decisions via voting as proposed in our approach.

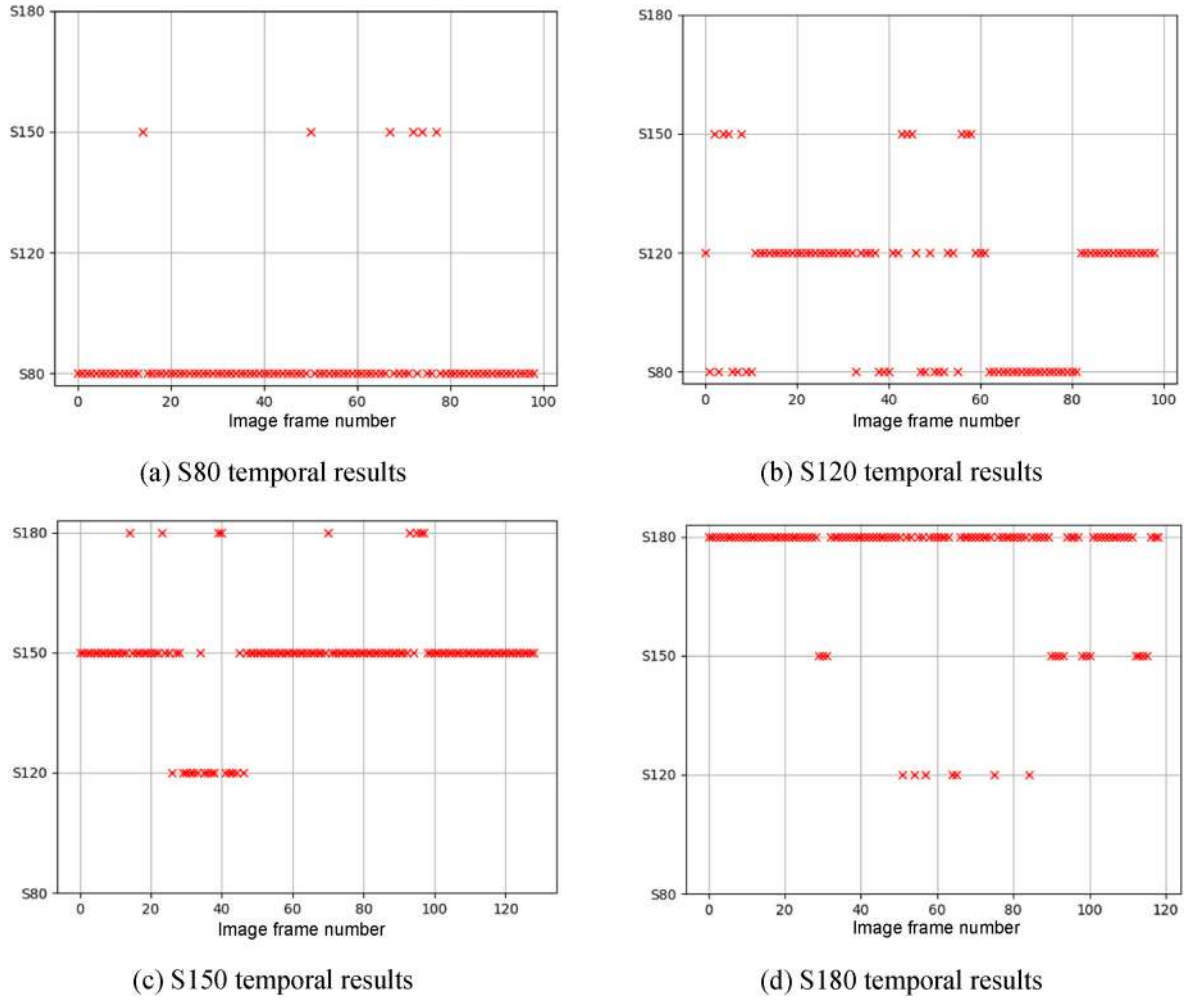


Fig. 11. Temporal results for single video recording on class: (a) S80, (b) S120, (c) S150, (d) S180.

Table 5

Performance metrics for image classification.

Classes	Precision	Recall	F1 Score
S80	0.80	0.86	0.83
S120	0.88	0.67	0.76
S150	0.74	0.86	0.80
S180	0.73	0.86	0.79
Average	0.79	0.81	0.80

Table 6

Performance metrics for video classification.

Classes	Precision	Recall	F1 Score
S80	0.83	0.91	0.87
S120	0.94	0.68	0.79
S150	0.85	1	0.92
S180	0.79	1	0.88
Average	0.85	0.90	0.87

5. Conclusion

In this paper, we propose a deep learning approach that predicts slump values from videos of concrete discharge at batching plants, utilizing the camera installed atop hoppers for practical and viable monitoring. Each video is tagged and categorized based on four slump value

classes of S80, S120, S150, and S180. Images extracted from the acquired video are processed and trained using the VGG16 network through a model improvement process using XAI. The image-level classes for a given video are first predicted by the trained VGG16 network, and the final prediction for the concrete slump is then determined by the highest number of votes for these image-level classes. This approach provides the first automated technique for predicting the slump value of the fresh concrete at the concrete batching plant.

Experimental results show that the proposed vision-based deep learning approach can predict the slump value at the average precision of 85% and the average F1 score of 87%. Even with the limited number of datasets, we show the effectiveness of our approach by predicting the slump value with high precision and accuracy. The proposed automated slump prediction method enables the continuous quality check for all of the produced concrete from the concrete batching plant that otherwise would be practically impossible to inspect all in the construction site, while also help evaluate and maintain the quality of the concrete produced in order to improve the safety of the construction site. Through this technology, it is possible to achieve on-site quality assurance (QA) related to concrete slump requirements and enhance time and cost efficiency in both concrete production and construction stages. In the future, if this monitoring technology is integrated into the material dosing control system of the concrete batching plant as part of a feedback framework, it will facilitate the production of high-quality concrete.

In this study, all recorded videos of the fresh concrete discharge have been tagged based on the production specification code. In future

research, if actual slump values are measured and used as tagging information for the dataset, it could enhance the accuracy of this technology and enable the prediction of continuous slump values. More available datasets can also improve the training of VGG16, which subsequently increases the accuracy of the slump value prediction. This study can be considered as the first step toward automating concrete quality inspection. This deep learning-based method will enhance construction productivity and reduce human dependency. As deep learning technology gains increasing popularity and finds successful applications in many aspects of construction projects, it is imperative to explore more automated methods for predicting concrete workability to establish a safer and more efficient digital quality control method. Future studies are expected to use better deep learning model architectures that directly exploit the temporal information from the video sequence for further improvement of the slump prediction.

CRediT authorship contribution statement

Sarmad Idrees: Writing – original draft. **Joshua Agung Nugraha:** Writing – original draft. **Shafaat Tahir:** Data curation. **Kichang Choi:** Validation. **Jongeeun Choi:** Writing – review & editing. **Deug-Hyun Ryu:** Data curation. **Jung-Hoon Kim:** Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by Eugene Corporation [grant numbers 2023-11-1217]; the National Research Foundation (NRF) Korea [grant numbers NRF-2021R1A2C2013522].

References

- Ali, L., Alnajjar, F., Jassmi, H.A., Gocho, M., Khan, W., Serhani, M.A., 2021. Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors* 21 (5), 1688. <https://doi.org/10.3390/s21051688>.
- Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* 8, 53. <https://doi.org/10.1186/s40537-021-00444-8>.
- Badra, N., Haggag, S.A., Deifalla, A., Salem, N.M., 2022. Development of machine learning models for reliable prediction of the punching shear strength of FRP-reinforced concrete slabs without shear reinforcements. *Measurement* 201, 111723. <https://doi.org/10.1016/j.measurement.2022.111723>.
- Baduge, S.K., Thilakarathna, S., Perera, J.S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A., Mendis, P., 2022. Artificial intelligence and smart vision for building and construction 4.0: machine and deep learning methods and applications. *Autom. Construct.* 141, 104440 <https://doi.org/10.1016/j.autcon.2022.104440>.
- BS EN 12350-2, 2019. *Testing Fresh Concrete - Slump Test*. British Standards Institution, London.
- Chou, J.-S., Liu, C.-Y., Prayogo, H., Khasani, R.R., Gho, D., Lalitan, G.G., 2022. Predicting nominal shear capacity of reinforced concrete wall in building by metaheuristics-optimized machine learning. *J. Build. Eng.* 61, 105046 <https://doi.org/10.1016/j.jobe.2022.105046>.
- Elharrouss, O., Akbari, Y., Almaadeed, N., Al-ma'adeed, S., 2022. Backbones-review: feature extraction networks for deep learning and deep reinforcement learning approaches. *arXiv preprint arXiv:2206.08016*. <https://doi.org/10.48550/arXiv.2206.08016>.
- Fang, W., Love, P.E.D., Luo, H., Xu, S., 2022. A deep learning fusion approach to retrieve images of People's unsafe behavior from construction sites. *Dev. Built. Envir.* 12, 100085 <https://doi.org/10.1016/j.dibe.2022.100085>.
- Han, S.H., Khayat, K.H., Park, S., Yoon, J., 2024. Machine learning-based approach for optimizing mixture proportion of recycled plastic aggregate concrete considering compressive strength, dry density, and production cost. *J. Build. Eng.* 83, 108393 <https://doi.org/10.1016/j.jobe.2023.108393>.
- He, K., Zhang, X., Ren, S., Sun, J., 2015a. Deep Residual Learning for Image Recognition. <https://doi.org/10.48550/arXiv.1512.03385> arXiv preprint arXiv:1512.03385.
- He, K., Zhang, X., Ren, S., Sun, J., 2015b. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: 2015 IEEE International Conference on Computer Vision, pp. 1026–1034. <https://doi.org/10.1109/ICCV.2015.123>. Santiago, Chile.
- Hilloulin, B., Tran, V.Q., 2023. Interpretable machine learning model for autogenous shrinkage prediction of low-carbon cementitious materials. *Construct. Build. Mater.* 396, 132343 <https://doi.org/10.1016/j.conbuildmat.2023.132343>.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA. <https://doi.org/10.1109/CVPR.2017.243>.
- Juez, J.M., Artoni, R., Cazaciu, B., 2017. Monitoring of concrete mixing evolution using image analysis. *Powder Technol.* 305, 477–487. <https://doi.org/10.1016/j.powtec.2016.10.008>.
- Kang, G.O., Seo, J., Chang, S., 2024. Application of machine learning algorithm for the estimation of time-dependent strength of basic oxygen furnace slag-treated soil. *Dev. Built. Envir.* 17, 100324 <https://doi.org/10.1016/j.dibe.2024.100324>.
- Khan, M.I., Abbas, Y.M., 2023. Intelligent data-driven compressive strength prediction and optimization of reactive powder concrete using multiple ensemble-based machine learning approach. *Construct. Build. Mater.* 404, 133148 <https://doi.org/10.1016/j.conbuildmat.2023.133148>.
- Khayat, K.H., Libre, N.A., 2014. Automated measurement and control of concrete properties in a ready mix truck with VERIFI. In: *Center for Transportation Infrastructure and Safety*. Missouri University of Science and Technology, Missouri.
- Kim, J.H., Park, M., 2018. Visualization of concrete slump flow using the kinect sensor. *Sensors* 18 (3), 771. <https://doi.org/10.3390/s18030771>.
- KS F 4009, 2016. *Ready-mixed Concrete*, Korean Standards Association (In Korean), Korea Standards Information Center.
- Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W., Müller, K.-R., 2019. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* 10, 1096. <https://doi.org/10.1038/s41467-019-08987-4>.
- Lee, S.Y., Le, T.H.M., Kim, Y.M., 2023. Prediction and detection of potholes in urban roads: machine learning and deep learning based image segmentation approaches. *Dev. Built. Envir.* 13, 100109 <https://doi.org/10.1016/j.dibe.2022.100109>.
- Malekipour, M., Moodi, F., 2021. A novel approach to improve quality of delivered concrete using slump estimations of the ready-mixed concrete (RMC) truck mixer. *J. Build. Eng.* 44, 103361 <https://doi.org/10.1016/j.jobe.2021.103361>.
- Moein, M.M., Saradar, A., Rahmati, K., Mousavinejad, S.H.G., Bristow, J., Aramali, V., Karakouzian, M., 2023. Predictive models for concrete properties using machine learning and deep learning approaches: a review. *J. Build. Eng.* 63, 105444 <https://doi.org/10.1016/j.jobe.2022.105444>.
- Nilimaa, J., 2022. Lateral form pressure induced by SCC. In: *Proceedings of the Nordic Concrete Research XXIV NCR Symposium 2022*, pp. 17–19. Stockholm, Sweden.
- Nilimaa, J., Gamil, Y., Zhaka, V., 2023. Formwork engineering for sustainable concrete construction. *Civil. Eng.* 4, 1098–1120. <https://doi.org/10.3390/civileng4040060>.
- Ribani, R., Marengoni, M., 2019. A survey of transfer learning for convolutional neural networks. In: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, Brazil, pp. 47–57. <https://doi.org/10.1109/SIBGRAPI-T.2019.00010>.
- Shahrokhshahraki, M., Malekipour, M., Mirvalad, S., Faraone, G., 2024. Machine learning predictions for optimal cement content in sustainable concrete constructions. *J. Build. Eng.* 82, 108160 <https://doi.org/10.1016/j.jobe.2023.108160>.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. <https://doi.org/10.48550/arXiv.1409.1556>.
- Situ, Z., Teng, S., Feng, W., Zhong, Q., Chen, G., Su, J., Zhou, Q., 2023. A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods. *Dev. Built. Envir.* 15, 100191 <https://doi.org/10.1016/j.dibe.2023.100191>.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2015.7298594>. Boston, MA, USA.
- Tuan, N.M., Hau, Q.V., Chin, S., Park, S., 2021. In-situ concrete slump test incorporating deep learning and stereo vision. *Autom. Construct.* 121, 103432 <https://doi.org/10.1016/j.autcon.2020.103432>.
- Wang, M., Yao, G., Yang, Y., Sun, Y., Yan, M., Deng, R., 2023. Deep learning-based object detection for visible dust and prevention measures on construction sites. *Dev. Built. Envir.* 16, 100245 <https://doi.org/10.1016/j.dibe.2023.100245>.
- Xu, Y., Zhou, Y., Sekula, P., Ding, L., 2021. Machine learning in construction: from shallow to deep learning. *Dev. Built. Envir.* 6, 100045 <https://doi.org/10.1016/j.dibe.2021.100045>.
- Yang, L., An, X., Du, S., 2021. Estimating workability of concrete with different strength grades based on deep learning. *Measurement* 186, 110073. <https://doi.org/10.1016/j.measurement.2021.110073>.
- Yu, Y., Zhang, C., Xie, X., Yousefi, A.M., Zhang, G., Li, J., Samali, B., 2023. Compressive strength evaluation of cement-based materials in sulphate environment using optimized deep learning technology. *Dev. Built. Envir.* 16, 100298 <https://doi.org/10.1016/j.dibe.2023.100298>.
- Zhang, X., Akber, M.Z., Zheng, W., 2022. Predicting the slump of industrially produced concrete using machine learning: a multiclass classification approach. *J. Build. Eng.* 58, 104997 <https://doi.org/10.1016/j.jobe.2022.104997>.
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016. Learning deep features for discriminative localization. In: 2016 IEEE Conference on Computer Vision and

Pattern Recognition, pp. 2921–2929. <https://doi.org/10.1109/CVPR.2016.31>. Las Vegas, NV, USA.