Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2025) 4   0-4 42

ECCOMAS
Proceedia

COMPDYN 2025
10th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Rhodes Island, Greece, 15-18 June 2025

FROM PIXELS TO PERFORMANCE: DEEP LEARNING FOR CONCRETE STRUCTURAL HEALTH MONITORING

Afaq Ahmad[1], Mati Ullah[2] and Vagelis Plevris[2]

[1] Department of Built Environment, Oslo Metropolitan University
Pilastered 35, Oslo 0166, Norway
e-mail: afaq.ahmad@oslomet.no

[2] Department of Civil and Environmental Engineering, College of Engineering, Qatar University
P.O. Box: 2713, Doha, Qatar
e-mail: matiullah@qu.edu.qa, vplevris@qu.edu.qa

Abstract

Accurate structural health monitoring (SHM) of concrete structures depends heavily on the reliable analysis of material components, particularly aggregate distribution. Traditional image processing methods often fall short in segmenting aggregates due to low contrast with the surrounding cement paste and the heterogeneous nature of concrete. This study proposes a deep learning–based approach using the DeepLabV3+ semantic segmentation framework with a MobileNetV2 backbone to address these challenges. A comprehensive, high-resolution dataset was developed from concrete specimens with varying mix designs and curing conditions, enabling the model to be trained from scratch. Quantitative results show that the proposed method achieves significantly higher accuracy than conventional segmentation techniques, with an average Intersection over Union (IoU) of 88.67% and an F1-score of 92.14%. Qualitative comparisons further demonstrate the model's ability to accurately distinguish aggregates under varying visual conditions. The framework's efficiency, adaptability, and precision make it a promising tool for automated inspection and predictive maintenance in SHM applications. This work advances the integration of deep learning into civil engineering workflows and paves the way for intelligent, data-driven infrastructure management.

Keywords: Concrete Aggregate Segmentation, Structural Health Monitoring (SHM), Deep Learning, Semantic Segmentation, DeepLabV3+.

ISSN:2623-3347 © 2025 The Authors. Published by Eccomas Proceedia.
Peer-review under responsibility of the organizing committee of COMPDYN 2025. doi: 10.7712/120125.127 9.25577

1 INTRODUCTION

Structural Health Monitoring (SHM) has emerged as a vital component in the lifecycle management of civil infrastructure, providing continuous or periodic evaluation of structural conditions to detect damage, assess performance, and ensure safety [1]. By enabling early diagnosis of deterioration and facilitating predictive maintenance, SHM systems support extended service life and cost-effective infrastructure management [2]. Modern SHM strategies combine sensor technologies, data acquisition systems, and analytical tools to monitor key parameters such as strain, displacement, temperature, and load effects across various structural materials [3].

In recent years, the integration of Artificial Intelligence (AI) and machine learning into SHM has opened new avenues for automation, accuracy, and real-time assessment [4, 5]. AI techniques—particularly deep learning—offer the ability to extract complex patterns from large datasets, enabling intelligent interpretation of sensor data, imagery, and signal information. Applications of AI in SHM range from anomaly detection and damage classification to predictive modeling and condition-based maintenance planning [6, 7]. These approaches significantly reduce the reliance on manual inspections and subjective judgment, improving both reliability and efficiency in structural diagnostics.

Among structural materials, concrete remains a cornerstone of civil engineering, forming the backbone of modern infrastructure and construction. However, the heterogeneous and composite nature of concrete makes its condition assessment especially challenging. Parameters such as workability, homogeneity, compressive strength, and surface integrity are critical for ensuring structural performance and durability. Among these parameters, compressive strength is particularly important and is typically assessed using destructive or non-destructive testing methods [8]. However, these conventional techniques often involve costly equipment, are labor-intensive, and may be impractical for in-situ or widespread real-time applications [9].

Another essential aspect of concrete evaluation is the analysis of aggregate distribution and proportion, which significantly influences mechanical behavior. Determining the ratio of aggregate particles to cement paste is fundamental to assessing mixture quality, but it depends on accurate segmentation of aggregates from the surrounding matrix. This task is inherently difficult due to the low color contrast between aggregates and cement paste and the limited availability of annotated image datasets [10, 11]. Although some studies have employed colorants to enhance contrast, such methods are unsuitable for real-world, non-invasive applications. Consequently, there is a pressing need for advanced, automated, and robust analysis techniques to support concrete mix design and performance prediction [12].

Traditional image processing methods often struggle to capture the complex and variable nature of aggregates, leading to inconsistencies in segmentation results. In contrast, recent advancements in deep learning (DL) have enabled more accurate and adaptable solutions through semantic segmentation techniques. This study evaluates the performance of traditional image processing methods alongside DL-based semantic segmentation approaches for extracting aggregate features from cross-sections of cylindrical concrete samples.

The primary objective is to develop a DL–based segmentation framework capable of supporting real-time property assessment for SHM applications. By harnessing the power of state-of-the-art semantic segmentation models, this research aims to address the limitations of traditional methods and provide a more robust, scalable solution for aggregate analysis. The results of this work are expected to enhance quality control in construction and improve predictive modeling in structural engineering.

The remainder of this paper is structured as follows: Section 2 provides a review of relevant literature. Section 3 details the dataset development, materials, and methods. Section 4 describes the model training process. Section 5 presents the results and discussion, and Section 6 concludes the study and outlines potential future research directions.

2 LITERATURE REVIEW

Recent research on concrete aggregate segmentation has primarily relied on traditional image processing techniques, largely due to the lack of comprehensive, annotated datasets required to train and validate DL models effectively. Yang and Buenfeld [13] proposed an image analysis method combining grey-level thresholding, filtering, and binary operations to separate aggregates from cement paste in BSE images of concrete. The technique enables accurate phase segmentation, allowing reliable measurement of porosity and anhydrous content. Chudasama et al. [14] used image segmentation for extracting homogeneous regions and enabling image analysis. They proposed a segmentation technique combining Fuzzy Canny edge detection with morphological operations to improve boundary accuracy. The method enhances feature extraction for applications like medical imaging and object recognition, demonstrating effective region separation through optimized edge-based processing. Coenen et al. [15] developed a semi-supervised framework incorporating prior knowledge to segment concrete aggregates effectively. This paper addresses label imbalance in semi-supervised semantic segmentation by proposing a novel framework that includes a shared encoder, a supervised primary decoder, and an auxiliary decoder for unlabeled data. Enhanced by class distribution priorities and auto-encoder regularization, our method outperforms purely supervised and standard consistency training on concrete aggregate segmentation, demonstrating improved accuracy and robustness.

The role of deep learning in material analysis is growing rapidly. Cha et al. [16] presents a CNN-based approach for automated concrete crack detection, thereby eliminating the need for traditional image processing techniques. Trained on 40K images, the model achieves 98% accuracy and adapts via sliding-window analysis for high-resolution images. Testing under challenging conditions demonstrates superior performance over edge-detection methods, proving its robustness for real-world infrastructure inspections.

Zhou et al. [17] evaluates deep learning crack segmentation methods across various image types, introducing the public FIND dataset, which contains raw, filtered, and fused images. Benchmarking nine CNN architectures reveals performance gains from data fusion, identifying optimal models and modalities. The study provides actionable insights for crack detection research and practical applications in infrastructure inspection.

Deep learning has emerged as a powerful tool in computer vision, particularly for complex image analysis tasks. Unlike traditional methods, DL operates as an end-to-end system, processing raw image data directly to produce final outputs. The introduction of Fully Convolutional Networks (FCNs) [18, 19] has significantly advanced semantic segmentation by enabling pixel-wise classification through variable-sized input images and up sampling via deconvolution layers. This architecture enables precise segmentation at the pixel level, overcoming many of the limitations of earlier methods. The ability of FCNs to process unstructured image data and produce detailed, structured outputs makes them highly suitable for the demanding task of aggregate segmentation in concrete, with promising implications for both material science and SHM applications [18, 19].

3 MATERIALS AND METHODS

3.1 Database

The research methodology involves three main phases: (i) development of a high-quality image dataset; (ii) training of a Deep Convolutional Neural Network (DCNN)-based semantic segmentation model; and (iii) post-processing and morphological analysis of segmented aggregates [20-24].

To generate a diverse and representative dataset, nine concrete mixtures were prepared by combining three cement:sand:aggregate ratios (1:3:6, 1:2:4, 1:1.5:3) with three water–cement (*w*/*c*) ratios (0.4, 0.5, 0.6), following ACI 318 [25] and ASTM C94 [26] standards. Additionally, supplementary cementitious materials (SCMs) such as silica fume (SF) and fly ash (FA) were incorporated at 0%, 15%, and 25% replacement levels. These SCMs enhanced both the contrast in color between paste and aggregates for improved image segmentation, and the workability of mixtures, potentially influencing aggregate dispersion patterns.

Figure 1 presents the workflow for specimen preparation and imaging. The process includes four steps: (1) mixing concrete in designated ratios, (2) casting the samples in cylindrical molds (150 mm diameter × 300 mm height), (3) cutting each cured cylinder into three equal slices after 14 and 28 days of curing, and (4) imaging the cross-sectional surfaces under standardized conditions. Forty-eight cylindrical specimens were prepared in total.

To eliminate surface irregularities often found at the top and bottom of concrete specimens, only the four interior faces of each sliced cylinder were imaged, as depicted in Figure 2. This ensured consistency and prevented misrepresentation due to material segregation or settling during the casting process.

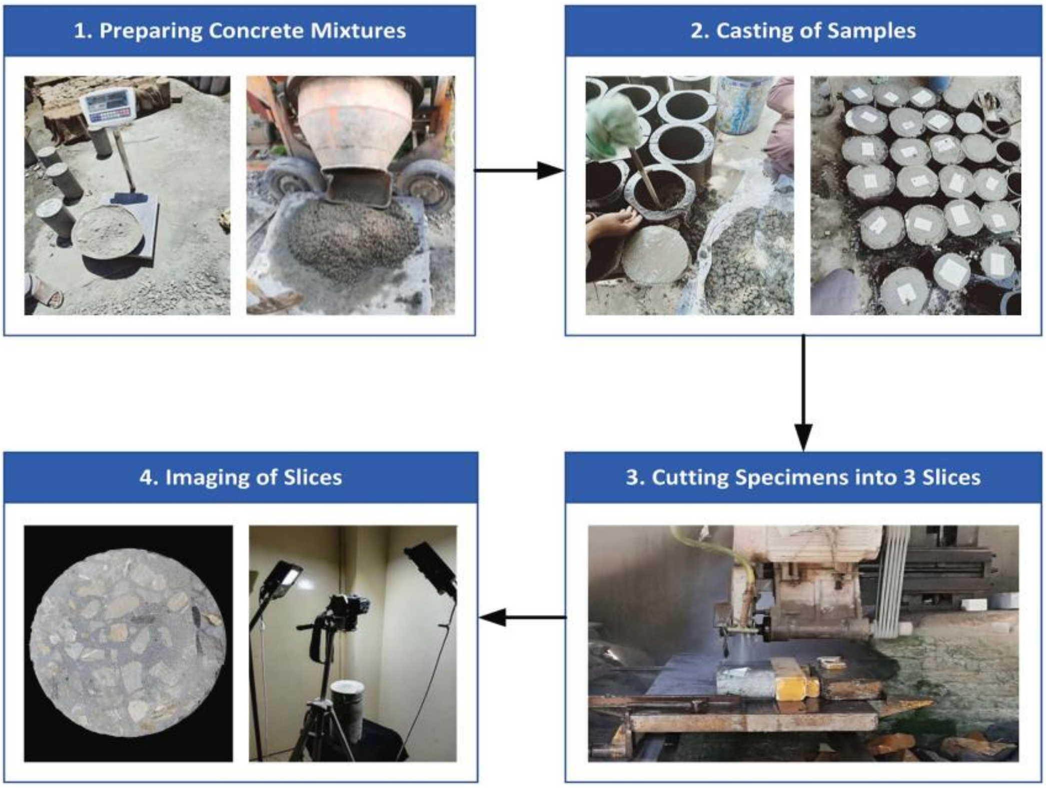

Figure 1: Experimental workflow for dataset generation, including (1) preparation of concrete mixtures, (2) casting of cylindrical specimens, (3) cutting each specimen into three equal slices, and (4) imaging of internal cross-sections under controlled lighting conditions [27].

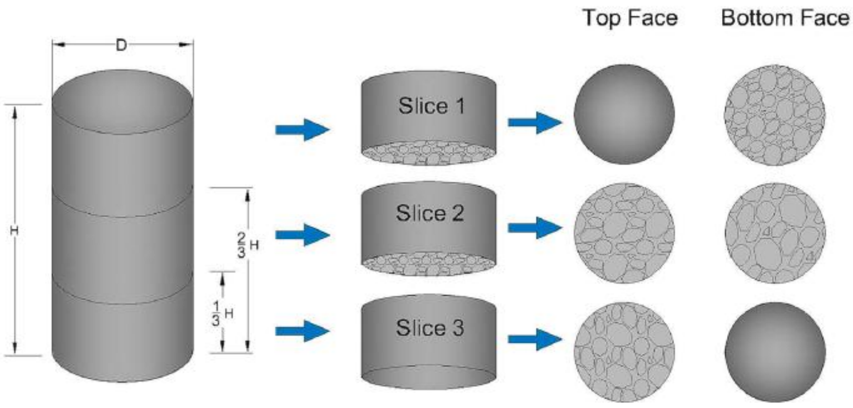

Figure 2: Schematic of specimen slicing and face selection.

3.2 Image Acquisition and Processing

Images were captured in a controlled environment using a 2.00 m × 2.50 m enclosure equipped with 2000 lux LED illumination to minimize shadows and glare. A 24-megapixel Nikon DSLR camera was used to acquire 192 high-resolution images (6000 × 4000 pixels), which were systematically cropped into 512 × 512-pixel sub-images. After filtering out irrelevant backgrounds and edges, a refined dataset of 9,572 images was compiled. Aggregates accounted for 41.4% of the total labeled pixel distribution.

Figure 3 shows an example of a full concrete slice alongside its corresponding manual annotation, while Figure 4 demonstrates the segmentation of detailed image patches extracted from these slices. The latter highlights the pixel-wise mapping between the raw images and their hand-labeled masks—critical for training accurate semantic segmentation models.

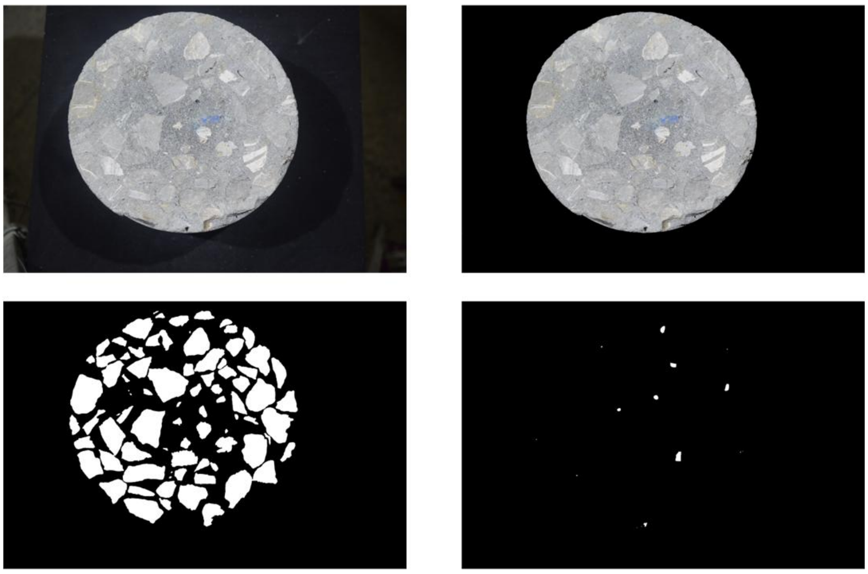

Figure 3: Example of a full concrete cylinder slice with ground truth aggregates.

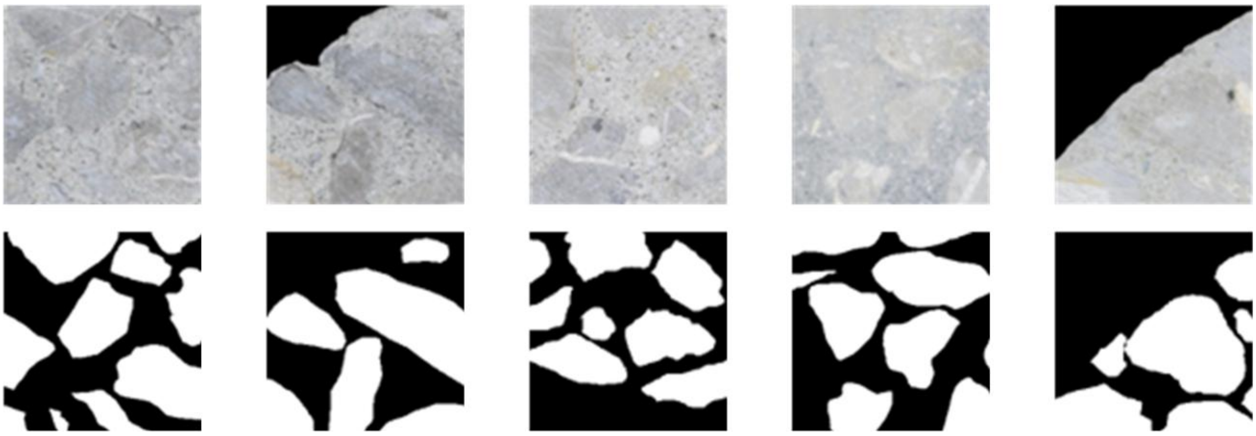

Figure 4: Cropped image patches with their corresponding manually annotated ground truth segmentation.

3.3 Deep Learning Model Architecture

To perform the semantic segmentation of aggregates from concrete imagery, the DeepLabV3+ architecture [27] was adopted, as illustrated in Figure 5. This framework integrates: (1) A DCNN backbone for robust feature extraction; (2) An encoder block that captures multi-scale contextual information using atrous (dilated) convolutions and spatial pyramid pooling; (3) A decoder block that refines spatial details and reconstructs high-resolution segmentation masks.

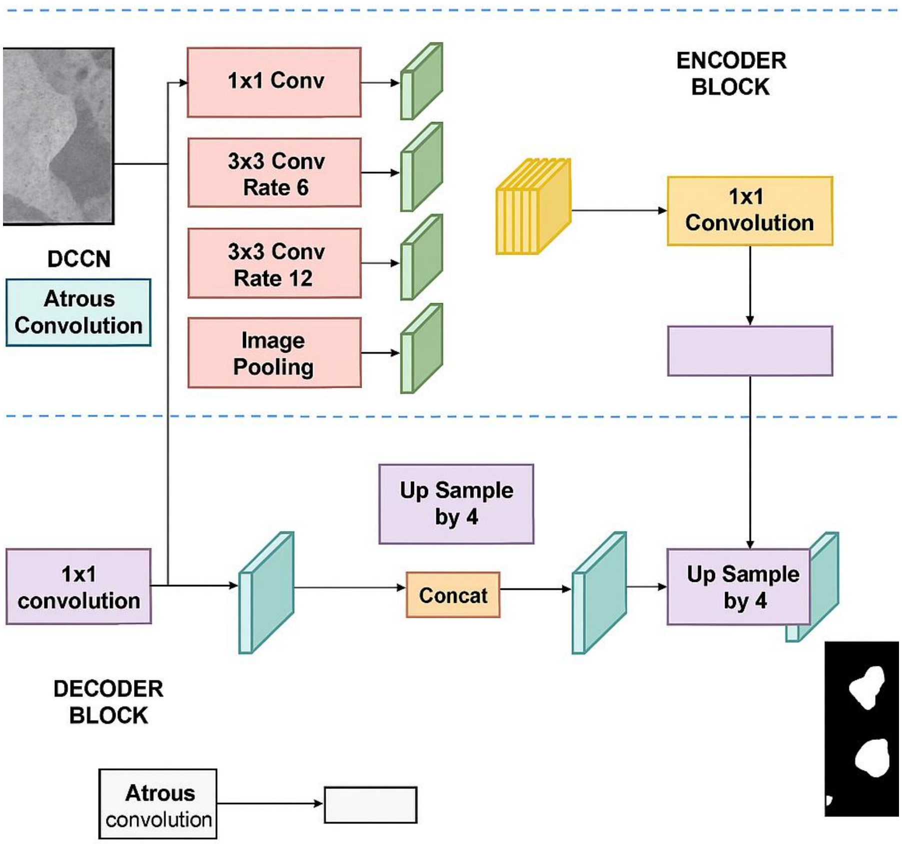

Figure 5: Architecture of the DeepLabV3+ semantic segmentation model.

The network was trained using the custom-developed dataset described above. The DCNN design, with its end-to-end learning capability and high adaptability to image complexity, enables accurate and efficient segmentation of aggregate particles, particularly in low-contrast and heterogeneous environments typical of real-world concrete samples.

3.4 Base Network MobileNetV2

The DeepLabV3+ architecture incorporates a Deep Convolutional Neural Network (DCNN) backbone to extract semantic and morphological features from input images of concrete slices, facilitating accurate aggregate segmentation [12, 28]. In this study, MobileNetV2 was selected as the backbone network due to its computational efficiency and firm performance in lightweight deep learning tasks, making it well-suited for both desktop and embedded applications [29].

Although designed initially for general-purpose computer vision tasks such as object detection and classification, MobileNetV2 is effectively adapted for semantic segmentation by integrating it within an encoder-decoder framework. Its architecture offers an optimal balance between accuracy and computational complexity, which is critical for training models on large-scale, high-resolution datasets.

As illustrated in Figure 5, the encoder module of the segmentation network consists of several key components:

1. **Feature Extraction:** The input image is passed through multiple convolutional layers of MobileNetV2, generating hierarchical feature maps that capture both spatial structure and contextual information at multiple resolutions.
2. **Atrous (Dilated) Convolutions**: To maintain high spatial resolution and expand the receptive field without increasing the number of parameters, dilated convolutions are applied in the deeper layers of the encoder. This ensures that fine-grained features, such as aggregate boundaries, are preserved.
3. **Atrous Spatial Pyramid Pooling (ASPP)**: The ASPP module aggregates contextual information at multiple scales using parallel atrous convolutions with different dilation rates. This enhances the network's ability to detect aggregates of varying sizes and improves segmentation accuracy across heterogeneous image regions.
4. **Depth-wise Separable Convolutions**: Although not explicitly depicted in Figure 5 MobileNetV2, used as the backbone, relies on depth-wise separable convolutions (a depth-wise convolution followed by a pointwise 1×1 convolution). This architectural design reduces computational complexity while preserving rich feature representation.
5. **Inverted Residual Blocks**: MobileNetV2 introduces inverted residual blocks, which form the backbone's building units. Each block expands the input channels, applies a depth-wise separable convolution, and then projects the result back to a lower-dimensional space using a 1×1 pointwise convolution. These blocks include shortcut (residual) connections, which allow for efficient feature reuse and reduce information loss across layers. This design enables MobileNetV2 to achieve high accuracy with minimal computational cost.

By integrating these components, the MobileNetV2-based encoder in DeepLabV3+ enables accurate, efficient, and scalable semantic segmentation of concrete aggregate images. This approach ensures high-level feature abstraction while preserving critical spatial details necessary for precise pixel-level classification.

3.5 Decoder Block

The decoder block reconstructs high-resolution segmentation masks from the low-resolution feature maps generated by the encoder and the Atrous Spatial Pyramid Pooling (ASPP) module. This is accomplished through a combination of up-sampling operations and convolutional layers, ensuring that the final output dimensions match those of the original input image. The decoder architecture includes the following key components:

- **Skip Connections:** To retain spatial precision and enhance boundary delineation, intermediate feature maps from the encoder are merged with decoder outputs via skip connections. This fusion helps recover fine-grained details that may be lost during down-sampling in the encoder.
- **Up sampling and Refinement**: The decoder up-samples the ASPP output and applies additional 3×3 convolutional layers to refine spatial details.
- **Segmentation Mask Generation**: A final 1×1 convolutional layer outputs the pixel-wise segmentation map, where each pixel is classified into a category (e.g., aggregate or background). This results in a dense mask suitable for morphological analysis.

This decoder design strikes a balance between computational efficiency and segmentation accuracy, making it effective for high-resolution aggregate detection and other image-based structural evaluation tasks.

4 TRAINING OF MODELS

4.1 Training Protocol

The dataset was partitioned into three subsets: 50% for training, 25% for validation, and 25% for testing. This split ensured a balanced evaluation of the model's learning capacity, generalization ability, and predictive performance. All experiments were conducted on a 64-bit Windows machine equipped with an Intel Core i7-8650U processor (1.90–2.11 GHz) and 8 GB RAM.

Although transfer learning is commonly employed to initialize deep learning models using pre-trained weights (e.g., from ImageNet [29]), it was not utilized in this study. Due to the significant domain differences between the ImageNet dataset and the grayscale concrete aggregate images used here, the model was trained from scratch to ensure optimal convergence and adaptation to the specific characteristics of the target data.

The model achieved 85% accuracy within 500 iterations, demonstrating rapid feature learning. Performance stabilized above 90% after 1,000 iterations, with a final accuracy of 95% achieved at 4,200 iterations. The minor fluctuations during training suggest batch variability, while the high final accuracy confirms robust convergence and generalization capability for aggregate segmentation tasks. The loss function exhibited a sharp initial decline, followed by stabilization after ~1,500 iterations, which correlated with improvements in accuracy. While occasional loss spikes occurred from challenging mini-batches, the consistent downward trend confirms effective error minimization. Combined with 95% final accuracy, these results demonstrate the framework's robustness for structural segmentation tasks.

4.2 Performance Metrics

To evaluate the segmentation performance of the proposed model, the Jaccard Index, also known as Intersection over Union (IoU), was employed as the primary metric. This statistical measure quantifies the similarity between two sets by calculating the degree of overlap between the predicted segmentation mask (A) and the ground truth (B).

Mathematically, the Jaccard Index is defined as the ratio of the intersection area to the union area of the two sets:

$$IoU = J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

The resulting value $J(A, B)$ ranges from 0 to 1, where a score of 1 indicates perfect agreement between the predicted and actual segmentation, and 0 denotes no overlap. As a robust metric for binary classification tasks in semantic segmentation, the Jaccard Index provides a reliable basis for assessing model accuracy in detecting and delineating aggregate regions within concrete images.

In addition to the Jaccard Index, we employed three supplementary metrics —precision, recall, and the F1 score —to provide a more comprehensive evaluation of the model's segmentation accuracy. These metrics are widely used in binary classification tasks to assess the balance between false positives and false negatives.

- **Precision** quantifies the proportion of correctly predicted aggregate pixels among all pixels identified as aggregates by the model.
- **Recall** measures the ability of the model to detect all actual aggregate pixels.
- **F1-score** is the harmonic means of precision and recall, offering a balanced metric that accounts for both over-segmentation and under-segmentation.

These metrics are formally defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\text{-}score = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision} \tag{4}$$

Here, TP (True Positives) refers to correctly identified aggregate pixels, FP (False Positives) denotes non-aggregate (suspension) pixels incorrectly labeled as aggregates, and FN (False Negatives) represents actual aggregate pixels that the model did not detect. In our evaluation framework, aggregate pixels were treated as positive instances, while suspension or background pixels were considered negative instances. Together with the Jaccard Index, these metrics provide a robust and interpretable basis for quantifying the model's segmentation performance.

5 RESULTS AND DISCUSSION

The performance of deep learning models is highly dependent on both the quality and quantity of training data. Since the DeepLabV3+ network in this study was trained from scratch, without employing transfer learning, the initialization and tuning of learning parameters played a critical role in achieving accurate segmentation.

The model was trained using the Adam optimizer, configured with exponential decay rates of *$\beta_1 = 0.9$ (for the first moment) and $\beta_2 = 0.999$ (for the* second moment). A StepLR learning rate scheduler was applied, starting with an initial learning rate of 1×10^{-3}, and reducing it by a multiplicative factor of 0.95 after each epoch. The model was trained for 20 epochs, during which convergence was consistently observed. To enhance generalization and reduce the risk of overfitting, we applied standard data augmentation techniques, including random horizontal and vertical flipping, Gaussian noise injection, and HSV color space transformations.

To assess the effectiveness of our approach, we conducted a comparative analysis against traditional image processing methods—namely, Otsu global thresholding, Otsu adaptive thresholding, and K-means clustering—using the same test dataset. Quantitative results are presented in Table 1 demonstrate that the proposed method significantly outperforms these conventional techniques, achieving an average Intersection over Union (IoU) of 88.67% and an F1-score of 92.14%, outperforming all other methods by a wide margin.

Sr. No.	Segmentation Technique	IoU	F1-Score
1	Otsu Global Thresholding	0.4574	0.6134
2	Otsu Adaptive Thresholding	0.4430	0.5994
3	K-means	0.3014	0.4375
4	Proposed Method	0.8867	0.9214

Table 1: Quantitative comparison of segmentation performance using different techniques.

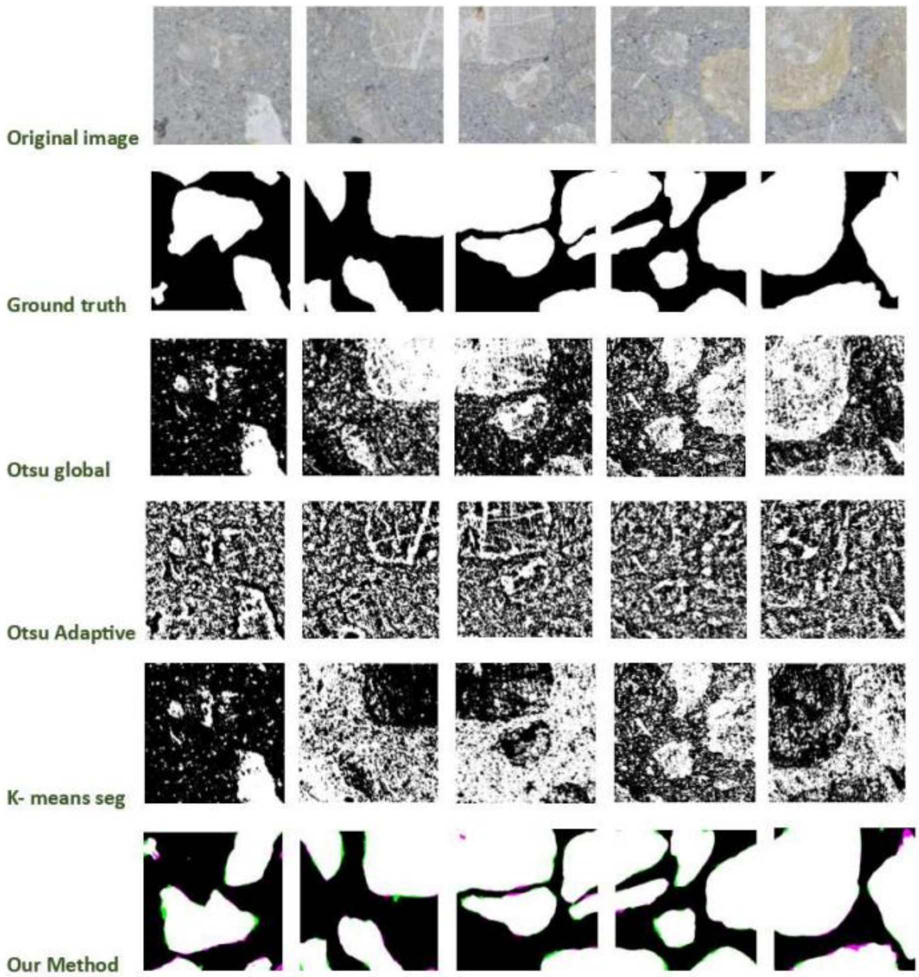

Figure 6: Qualitative comparison of segmentation results

The qualitative comparison in Figure 6 further supports these findings. The visual results show that the proposed deep learning-based approach effectively identifies and separates aggregate particles from the cement paste, even in cases of low color contrast. In contrast, thresholding and clustering-based methods struggle to distinguish these components because they rely solely on pixel intensity values.

The poor performance of traditional methods is rooted in their inability to handle the complex texture, heterogeneous particle shapes, and subtle color variations present in concrete microstructure images. These limitations emphasize the advantage of using deep learning models, which are capable of learning rich, hierarchical features directly from data. Overall, the results demonstrate the superior accuracy, robustness, and adaptability of the proposed method in segmenting concrete aggregates, reinforcing its suitability for structural material analysis tasks.

6 CONCLUSIONS

This study presents a robust deep learning framework for the semantic segmentation of concrete aggregates, utilizing a DeepLabV3+ architecture with a MobileNetV2 backbone. The proposed method addresses one of the critical challenges in Structural Health Monitoring (SHM), accurately identifying aggregate particles within complex concrete microstructures, where traditional thresholding and clustering techniques often fail due to low contrast and heterogeneity.

A comprehensive dataset was developed from carefully prepared and imaged cylindrical specimens with varying mix designs and curing conditions. The resulting dataset enabled the training of a semantic segmentation model from scratch, without relying on transfer learning, thus ensuring domain-specific adaptability. The model was optimized using the Adam optimizer and trained with a learning rate decay strategy, along with extensive data augmentation to promote generalization.

Quantitative evaluations demonstrated the superiority of the proposed deep learning approach over classical image processing methods. The model achieved a mean Intersection over Union (IoU) of 88.67% and an F1-score of 92.14%, significantly outperforming Otsu thresholding and K-means clustering in both accuracy and consistency. Qualitative comparisons further confirmed that our method provided more precise boundary delineation and more reliable segmentation across diverse image samples.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska–Curie grant agreement No. 101153307 (https://cordis.europa.eu/project/id/101153307).

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

REFERENCES

[1] Gharehbaghi, V.R., et al., *A Critical Review on Structural Health Monitoring: Definitions, Methods, and Perspectives.* Archives of Computational Methods in Engineering, 2022. **29**(4): p. 2209-2235 DOI: https://doi.org/10.1007/s11831-021-09665-9.

[2] Farrar, C.R. and K. Worden, *An introduction to structural health monitoring.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2007. **365**(1851): p. 303-315 DOI: https://doi.org/10.1098/rsta.2006.1928.

[3] Mardanshahi, A., A. Sreekumar, X. Yang, S.K. Barman, and D. Chronopoulos, *Sensing Techniques for Structural Health Monitoring: A State-of-the-Art Review on*

Performance Criteria and New-Generation Technologies. Sensors, 2025. **25**(5): p. 1424 DOI: https://doi.org/10.3390/s25051424.
[4] Plevris, V. and G. Papazafeiropoulos, *AI in Structural Health Monitoring for Infrastructure Maintenance and Safety.* Infrastructures, 2024. **9**(12) DOI: https://doi.org/10.3390/infrastructures9120225.
[5] Plevris, V., *AI-Driven Innovations in Earthquake Risk Mitigation: A Future-Focused Perspective.* Geosciences, 2024. **14**(9): p. 244 DOI: https://doi.org/10.3390/geosciences14090244.
[6] Bianchi, G., F. Freddi, F. Giuliani, and A. La Placa, *Implementation of an AI-based predictive structural health monitoring strategy for bonded insulated rail joints using digital twins under varied bolt conditions.* Railway Engineering Science, 2025 DOI: https://doi.org/10.1007/s40534-024-00371-3.
[7] Jiménez Rios, A., V. Plevris, and M. Nogal, *Bridge management through digital twin-based anomaly detection systems: A systematic review.* Frontiers in Built Environment, 2023. **9**(1176621) DOI: https://doi.org/10.3389/fbuil.2023.1176621.
[8] Nobile, L., *Prediction of concrete compressive strength by combined non-destructive methods.* Meccanica, 2015. **50**(2): p. 411-417 DOI: https://doi.org/10.1007/s11012-014-9881-5.
[9] Ivanchev, I., *Investigation with Non-Destructive and Destructive Methods for Assessment of Concrete Compressive Strength.* Applied Sciences, 2022. **12**(23): p. 12172 DOI: https://doi.org/10.3390/app122312172.
[10] Pare, S., A. Kumar, G.K. Singh, and V. Bajaj, *Image Segmentation Using Multilevel Thresholding: A Research Review.* Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020. **44**(1): p. 1-29 DOI: https://doi.org/10.1007/s40998-019-00251-1.
[11] Fang, C. and S. Labi, *Image-Processing Technology to Evaluate Static Segregation Resistance of Hardened Self-Consolidating Concrete.* Transportation Research Record, 2007. **2020**(1): p. 1-9 DOI: https://doi.org/10.3141/2020-01.
[12] Song, Y., Z. Huang, C. Shen, H. Shi, and D.A. Lange, *Deep learning-based automated image segmentation for concrete petrographic analysis.* Cement and Concrete Research, 2020. **135**: p. 106118 DOI: https://doi.org/10.1016/j.cemconres.2020.106118.
[13] Yang, R. and N.R. Buenfeld, *Binary segmentation of aggregate in SEM image analysis of concrete.* Cement and Concrete Research, 2001. **31**(3): p. 437-441 DOI: https://doi.org/10.1016/S0008-8846(00)00493-2.
[14] Chudasama, D., T. Patel, S. Joshi, and G.I. Prajapati, *Image Segmentation using Morphological Operations.* International Journal of Computer Applications, 2015. **117**(18): p. 16-19 DOI: https://doi.org/10.5120/20654-3197.
[15] Coenen, M., T. Schack, D. Beyer, C. Heipke, and M. Haist, *Semi-Supervised Segmentation of Concrete Aggregate Using Consensus Regularisation and Prior Guidance.* ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., 2021. **V-2-2021**: p. 83-91 DOI: https://doi.org/10.5194/isprs-annals-V-2-2021-83-2021.
[16] Cha, Y.-J., W. Choi, and O. Büyüköztürk, *Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks.* Computer-Aided Civil and Infrastructure Engineering, 2017. **32**(5): p. 361-378 DOI: https://doi.org/10.1111/mice.12263.
[17] Zhou, S., C. Canchila, and W. Song, *Deep learning-based crack segmentation for civil infrastructure: data types, architectures, and benchmarked performance.* Automation in Construction, 2023. **146**: p. 104678 DOI: https://doi.org/10.1016/j.autcon.2022.104678.

[18] Liu, Z., Y. Cao, Y. Wang, and W. Wang, *Computer vision-based concrete crack detection using U-net fully convolutional networks.* Automation in Construction, 2019. **104**: p. 129-139 DOI: https://doi.org/10.1016/j.autcon.2019.04.005.
[19] Hoang, N.-D. and V.-D. Tran, *Computer vision based asphalt pavement segregation detection using image texture analysis integrated with extreme gradient boosting machine and deep convolutional neural networks.* Measurement, 2022. **196**: p. 111207 DOI: https://doi.org/10.1016/j.measurement.2022.111207.
[20] Qayyum, W., et al., *Predicting characteristics of cracks in concrete structure using convolutional neural network and image processing.* Frontiers in Materials, 2023. **10** DOI: https://doi.org/10.3389/fmats.2023.1210543.
[21] Ahmed, C.F., A. Cheema, W. Qayyum, E. Rana, and A. Ahmad, *Detection of Pavement cracks of UET Taxila using pre-trained model Resnet50 of CNN*, in *1st International Conference on Advances in Civil and Environmental Engineering*. 2022: University of Engineering & Technology Taxila, Taxila, Pakistan.
[22] Ehtisham, R., et al., *Classification of defects in wooden structures using pre-trained models of convolutional neural network.* Case Studies in Construction Materials, 2023. **19**: p. e02530 DOI: https://doi.org/10.1016/j.cscm.2023.e02530.
[23] Ahmad, C.F., et al., *Classification of different size of potholes based on surface area using convolutional neural network.* Discover Applied Sciences, 2024. **6**(9): p. 492 DOI: https://doi.org/10.1007/s42452-024-06207-3.
[24] Qayyum, W., R. Ehtisham, V. Plevris, J. Mir, and A. Ahmad, *Classification of Wall Defects for Maintenance Purposes using Image Processing*, in *9th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering (COMPDYN 2023)*. 2023: Athens, Greece. p. 2529-2539. DOI: https://doi.org/10.7712/120123.10580.21466.
[25] ACI Committee 318, *ACI 318-19: Building Code Requirements for Structural Concrete and Commentary*. 2019.
[26] ASTM, *ASTM C494/C494M-17: Standard Specification for Chemical Admixtures for Concrete*. 2017.
[27] Ullah, M., J. Mir, S.S. Husain, M.L.U.R. Shahid, and A. Ahmad, *Concrete forensic analysis using deep learning-based coarse aggregate segmentation.* Automation in Construction, 2024. **162**: p. 105372 DOI: https://doi.org/10.1016/j.autcon.2024.105372.
[28] Chen, L., W. Shan, and P. Liu, *Identification of concrete aggregates using K-means clustering and level set method.* Structures, 2021. **34**: p. 2069-2076 DOI: https://doi.org/10.1016/j.istruc.2021.08.048.
[29] Howard, A.G., et al., *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.* ArXiv e-prints, 2017(arXiv:1704.04861) DOI: https://doi.org/10.48550/arXiv.1704.04861.