Concrete Aggregate Segmentation for Structural Health Monitoring

Mahnoor Malik[1], Junaid Mir[2], Afaq Ahmad[3]

[1,2] Department of Electrical Engineering,
University of Engineering and Technology, Taxila (47050), Pakistan
[2]junaid.mir@uettaxila.edu.pk

[3]Department of Civil Engineering,
University of Engineering and Technology, Taxila (47050), Pakistan
[3]afaq.ahmad@uettaxila.edu.pk

Abstract

Concrete aggregate segmentation plays a crucial role in assessing the quality and durability of concrete structures. However, traditional segmentation techniques often encounter limitations in effectively capturing intricate particle characteristics, hindering the precision required for reliable structural health analysis. This paper presents a comprehensive study focused on the segmentation of concrete aggregates from images, specifically comparing image segmentation techniques with deep learning-based semantic segmentation methods. Results reveal that a MobileNet-based DeepLabV3+ segmentation architecture outperforms conventional techniques like Otsu Thresholding and K-means segmentation in concrete aggregate segmentation. An Intersection over Union (IoU) value of 0.93 indicates that aggregates can be effectively identified and segmented from the images, paving the way for more informed decision-making in construction and civil engineering domains.

Keywords: Concrete aggregate segmentation, structural health monitoring, deep learning, DeepLabV3+, semantic segmentation

1. INTRODUCTION

Concrete is the fundamental building material in engineering, playing a critical role in the construction industry. Structural health monitoring (SHM) ensures concrete structures' integrity and long-term performance. Various concrete parameters, such as workability, homogeneity, strength, and surface quality, are assessed during SHM to evaluate a structure's durability and strength. Compressive strength is critical in concrete evaluation, typically measured through destructive or non-destructive testing methods. Measuring these concrete properties requires specialized equipment and instruments, which are costly and labor-intensive. Aggregate proportion and distribution in concrete can convey crucial information about the concrete properties and require calculating the ratio of aggregate particle distribution to cement paste. This ratio can be found if the concrete aggregates are segmented from the cement paste.

Segmenting concrete aggregates from the cement paste is challenging due to their similar color characteristics and the lack of a comprehensive image database. Previous studies have utilized dyes to facilitate image segmentation. However, using a color treatment to segment concrete aggregates is impractical for real-time scenarios. Accurately segmenting concrete aggregate particles is significant for measuring concrete mixture properties and predicting structural performance. Traditional image processing-based segmentation methods often encounter limitations in effectively capturing intricate particle characteristics, hindering the precision required for reliable analysis.

This paper compares conventional image segmentation techniques with state-of-the-art semantic segmentation methods for aggregate segmentation from cylindrical concrete images. The ultimate objective is to assess and develop an efficient DL-based concrete aggregate segmentation method that can be used for estimating properties during SHM. By embracing advanced semantic segmentation techniques rooted in deep learning, this research aims to bridge the gap, presenting a novel approach to enhance aggregate segmentation accuracy, paving the way for more informed decision-making in construction and engineering domains. The results of this research will contribute to improving concrete aggregate segmentation, a critical step towards developing robust models for SHM.

The rest of the paper is divided into five sections: Section 2 gives the literature review, section 3 comprises materials and methods, and Section 4 details the experiments with results and discussion presented in Section 5. The paper is concluded in section 6.

2. RELATED WORK

The existing work on concrete aggregate segmentation is mainly limited to using conventional image processing techniques only. This is primarily due to the lack of an appropriate image dataset required to train and test the deep learning techniques. An algorithm for separating aggregate particles in a concrete image using a combination of grey-level thresholding, filtering, and binary operations is presented in [1]. In [2], the effect of the water-binder ratio and fly ash on the homogeneity of the concrete is investigated. Parameters like segregation degree, compressive strength, and microhardness were used to evaluate the homogeneity of concrete aggregate. A method is proposed in [3] for detecting and assessing aggregate distribution uniformity in asphalt pavements. The aggregate distribution was analyzed using digital image processing techniques while considering the surface and internal structure of asphalt pavement. A study to assess the porosity of bonded mortar of recycled aggregates using backscattered electron (BSE) image analysis is done in [4]. The BSE analysis was carried out using the previously determined pore segmentation method to determine the microstructural characteristics of hydrated cement pastes.

Disruption in asphalt mixture using X-ray computed tomography and digital image processing was analyzed in [5]. In [6], a deep learning-based image segmentation algorithm for petrographic concrete analysis is proposed. A meso-structural model of coarse aggregate movement is devised in [7]. The study conducted an in-depth analysis of all the mechanisms involved in coarse aggregate movement load subjected to rutting using digital image processing. The method in [8] presented a deep learning-based automatic segmentation and morphological analysis method for concrete aggregate. The level set method LSM and K-means clustering is used in [9] to robustly recognize low-contrast images, like those obtained from Microscope or SEM, to recognize concrete aggregates. In [10], a deep learning algorithm is used to study aggregate asphalt mixtures to assess their durability and performance.

Deep learning (DL) has recently emerged as a highly promising approach in computer vision. DL stands out as an end-to-end learning paradigm that commences directly from raw data and culminates in generating the final output. Introducing the Fully Convolutional Neural Network (FCN) has resulted in rapid advancement in image semantic segmentation technology. FCNs exhibit remarkable versatility, capable of accommodating input images of varying dimensions [16]. Through the ingenious use of deconvolution layers, these networks skillfully up sample feature maps to their original sizes, ultimately facilitating the generation of precise predictions for each pixel within the image. This innovative approach marks a significant departure from traditional concrete aggregate segmentation methods, offering a robust and adaptable tool for addressing the intricate challenges

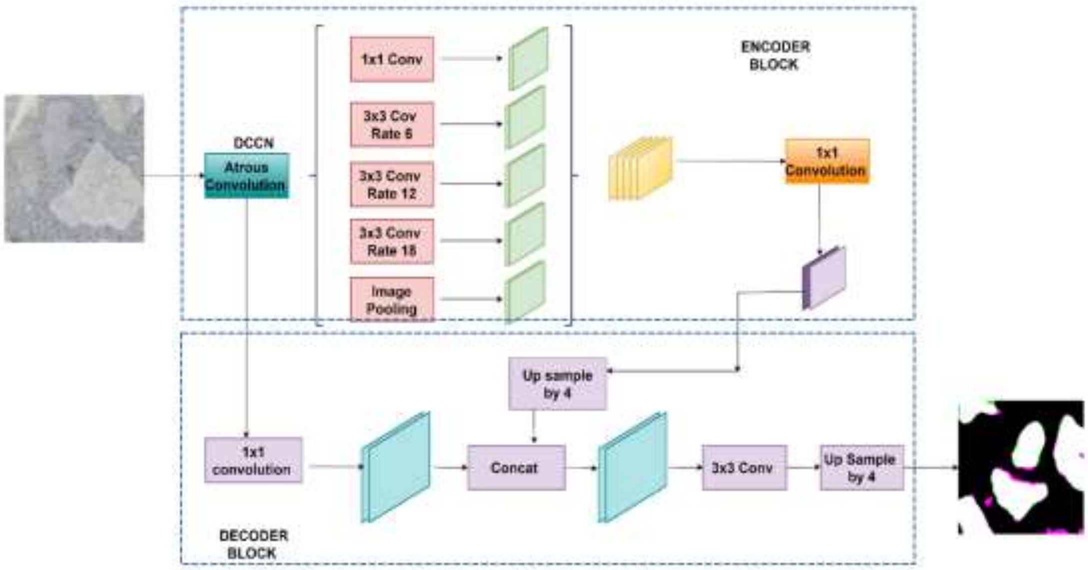

Figure 1: DeepLabV3+ semantic segmentation encoder decoder architecture with MobileNetV2 backbone DCNN.

posed by this domain. Its capacity to operate directly on raw data and provide end-to-end solutions demonstrates its potential to revolutionize the field of computer vision and, more specifically, concrete aggregate segmentation.

3. MATERIALS AND METHODS

The development of semantic segmentation architecture, DeepLabV3+, is outlined in this section. The architecture depicted in Fig. 1 was trained on a self-curated dataset (detailed in the next section) for effective concrete aggregate segmentation. The complete architecture consists of a base backbone deep convolutional neural network (DCNN), an encoder block and decoder block.

3.1 Base Network

MobileNetV2 DCNN model is used as a backbone network to process the images and extract the aggregate's morphological features from the image according to the prediction results. MobileNetV2 architecture is designed for efficient and lightweight deep learning applications, particularly in the context of mobile and embedded devices [17]. It balances model size, computational efficiency, and performance, making it suitable for various computer vision tasks, including image segmentation and object detection. While MobileNetv2 is not specifically designed for semantic segmentation, it is adapted for such tasks within an encoder-decoder architecture.

3.2 Encoder Block

The encoder consists of the following architectural elements.

Feature Extraction: Passing the image through MobileNetV2 will extract feature maps at different spatial resolutions, capturing hierarchical information about the input image. This step forms the foundational basis for subsequent processing.

Atrous Convolution for Feature Map Control: To effectively regulate the size of the feature map, atrous convolution, also known as dilated convolution, is employed in the latter stages of the backbone network. This technique ensures that the features' spatial resolution aligns with the segmentation task's specific requirements.

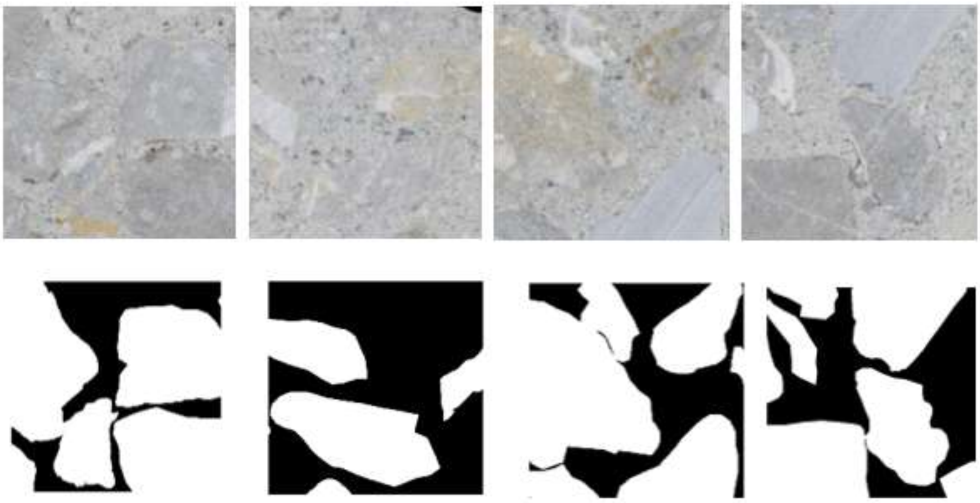

Figure 2: Sample original concrete images from [14] with their respective ground truth mask images created in this work.

ASPP Network for Pixel Classification: Building upon the feature extraction, an Atrous Spatial Pyramid Pooling (ASPP) network is integrated. This network is instrumental in classifying each pixel within the image, assigning them to their respective classes based on learned features.

Depth-wise Convolution: MobileNetV2 uses depth-wise separable convolutions, decomposing standard convolutions into two separate layers: depth-wise convolution and pointwise convolution. This reduces the computational cost while retaining the ability to capture essential features.

3.3 Decoder Block

A decoder network is responsible for up-sampling the low-resolution feature maps to the original image size and generating a segmentation mask using transposed convolutions. The following operations are involved in the decoder network.

Inverted Residuals: MobileNetV2 utilizes inverted residual blocks, which consist of a combination of expansion layers, depth-wise convolutions, and pointwise convolutions. This design increases the network's representational power without significantly increasing the computational load.

Skip Connections: To preserve spatial details and improve segmentation accuracy, skip connections are incorporated that connect corresponding feature maps from the encoder to the decoder. These connections help the network make fine-grained predictions.

Final Output: The output from the ASPP network transforms a 1 x 1 convolution layer. This final convolutional step ensures that the resulting image retains its original dimensions, producing the ultimate segmented mask for the input image. This mask effectively delineates the identified classes within the image, providing a highly informative and precise representation. In mask each pixel is assigned a class label, indicating the object or category it belongs to.

4. EXPERIMENTS

4.1 Dataset

The training and testing of DeepLabV3+ was done on the dataset developed in [14], consisting of slice images of concrete cylinders cast with different ratios of water, cement, and aggregate. The images were acquired in a fully controlled environment with sunlight blocked using black sheets. The cylinder slice samples were illuminated from a specific direction and angle by two 30 W LED bulbs that produced an artificial light of 2000 lx measured through a lux meter (Lutron LX 1109). A digital camera with 24 megapixels resolution (Nikon DSLR 3300) was used for the imaging.

The collected images in the original dataset had a 6000 × 4000 pixels resolution. As the original dataset did not have ground truth images, we manually created the database of mask images using the Image Labeler Tool of MATLAB. In total, 1130 images were used for performing the experiments in our research work. Fig. 2 shows sample images with ground truth labeled images.

4.2 Training Protocol

The image dataset was divided into three parts, namely, train, test, and validation sets. 50% of the images were used for training and the remaining 50% were divided equally into test and validation sets, each having 25% of the total image dataset. Experiments were conducted on a Windows-based 64-bit operating system having Intel Core i7-8650U CPU running at 1.90GHz - 2.11 GHz with 8.00 GB memory. Transfer learning is commonly used in computer vision tasks to transfer information from a trained network to a new network to tackle similar issues and provide the model with a better initial state. MobileNetV2 architecture is originally trained on ImageNet [13]. However, as the images in our study are quite different, transfer learning was not used, and the segmentation architecture was trained from scratch.

4.3 Performance Metrics

The Jaccard Index (also commonly referred to as IoU: Intersection over Union) is computed to evaluate the segmentation performance. The Jaccard Index is a statistical measure for assessing similarity and gauges the similarity between finite sets of samples [18]. It is precisely defined as the ratio of the size of the intersection of two sets to the size of their union, as

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

where A and B are the sets containing foreground and background pixels, respectively, and $J(A,B)$ is the Jaccard score between 0 and 1.

We have also computed precision, recall, F1 score for performance evaluation of segmentation techniques. Precision denotes the proportion of correctly identified aggregate pixels among all pixels classified as aggregates. Recall, on the other hand, signifies the percentage of all aggregate pixels that have been accurately classified. The mathematical relations for computing these metrics are

$$F1\ Score = \frac{2 \times P \times R}{P + R} \tag{2}$$

$$Precision\ (P) = \frac{TP}{TP + FP} \tag{3}$$

$$Recall\ (R) = \frac{TP}{TP + FN} \tag{4}$$

where TP, FP and FN are True Positive, False Positive and False Negative, respectively. In our assessment, the aggregate pixels are considered as positive instances and suspension pixels as negative instances.

5. RESULTS AND DISCUSSION

A deep learning model's performance is highly dependent on the data's quality and quantity. More importantly, as we have trained the network without transfer learning, the initialization of learning parameters can significantly affect overall segmentation performance. We employed the Adam optimizer for updating the model parameters, with exponential decay rate estimates of 0.9 for the first moment and 0.999 for the second moment. To manage the learning rates for each parameter group, we applied the StepLR learning rate

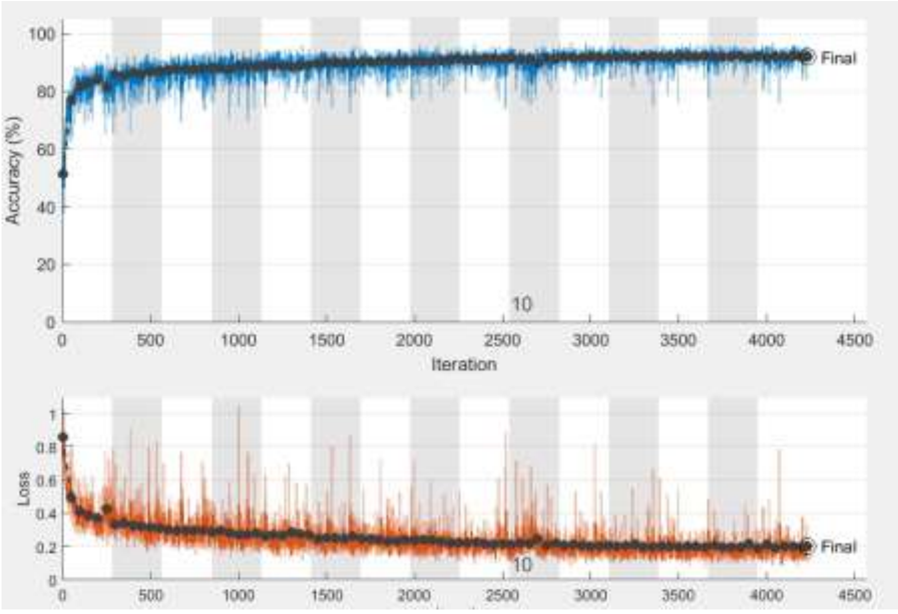

Figure 3: Learning curves of MobileNetV2 based DeepLabV3 semantic segmentation architecture.

strategy. This strategy involved initially setting the learning rate to 1e-3 and then multiplying it by 0.95 after each epoch. The training process spanned 20 epochs, during which we observed convergence.

Fig. 3 shows the learning curves of the trained DeepLabV3+ segmentation architecture for concrete aggregate segmentation. Standard data augmentation techniques were applied to prevent overfitting, including image flipping, adding noise, and HSV (Hue, Saturation, Value) transformations. From the curves, it can be observed that the model is well-generalized and is free from overfitting or underfitting problems. The accuracy on both training and validation sets is increasing, whereas the loss for both training and validation sets decreases with each iteration.

To contrast and compare the segmentation of the proposed DeepLabV3+ method, we have also performed segmentation on the test dataset using conventional image processing-based segmentation techniques like Otsu Thresholding and K-means. Table 1 gives the quantitative results in terms of the IoU and F1-score, whereas the qualitative results are given in Fig. 4. Our method gives the best results among the tested segmentation techniques. It can be observed that the proposed method attains a mean IoU of 88.67%, and aggregates are segmented out from the images with high efficiency. On the other hand, conventional techniques fail in segmenting out the aggregates from the images. Traditional methods generally employ a threshold value for segmenting the foreground from the background. However, such approaches will fail in concrete aggregate segmentation due to color similarity between aggregates (the foreground) and cement paste (the background). This is evident in Fig. 4, where it can be observed that both Otsu and K-means segmentation techniques have failed in the segmentation task.

Table 1: Performance comparison of conventional techniques with proposed method.

Sr. No.	Segmentation Technique	Metrics	
		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

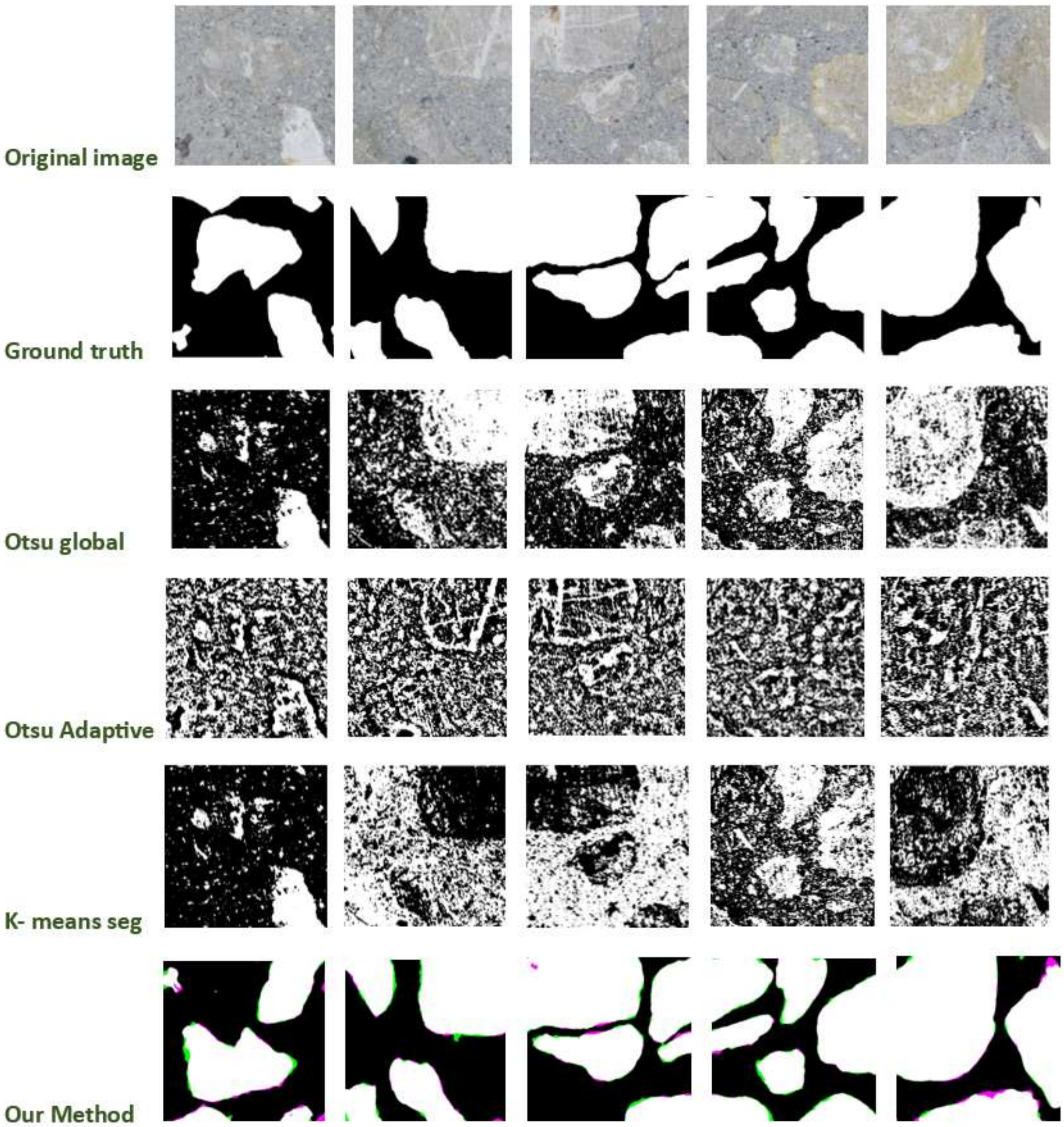

Figure 4: Qualitative performance comparison. Sample concrete images with ground truths given in first and second row, respectively.

6. CONCLUSION

This research presents a novel deep learning-based segmentation method for the segmentation of concrete aggregates from sedimentation images. The method is based on an encoder–decoder architecture that enhances the feature extraction properties of the network by combining both low-level and high-level information. The method outperforms image processing-based conventional segmentation methods. In future work, other deep-learning architectures will be tested for concrete aggregate segmentation, and the dataset will be increased.

REFERENCES

[1] Yang, R., & Buenfeld, N. R. (2001). Binary segmentation of aggregate in SEM image analysis of concrete. Cement and Concrete Research, 31(3), 437-441.
[2] Wang, G., Kong, Y., Sun, T., & Shui, Z. (2013). Effect of water–binder ratio and fly ash on the homogeneity of concrete. Construction and Building Materials, 38, 1129-1134.17
[3] Zhang, K., Zhang, Z., Luo, Y., & Huang, S. (2017). Accurate detection and evaluation method for aggregate distribution uniformity of asphalt pavement. Construction and Building Materials, 152, 715-730.
[4] Kim, Y., Hanif, A., Usman, M., & Park, W. (2019). Influence of bonded mortar of recycled concrete aggregates on interfacial characteristics–Porosity assessment based on pore segmentation from backscattered electron image analysis. Construction and Building Materials, 212, 149-163.
[5] Xing, C., Xu, H., Tan, Y., Liu, X., & Ye, Q. (2019). Mesostructured property of aggregate disruption in asphalt mixture based on digital image processing method. Construction and Building Materials, 200, 781-789.
[6] Song, Y., Huang, Z., Shen, C., Shi, H., & Lange, D. A. (2020). Deep learning-based automated image segmentation for concrete petrographic analysis. Cement and Concrete Research, 135, 106118.
[7] Shi, L., Wang, D., Jin, C., Li, B., & Liang, H. (2020). Measurement of coarse aggregates movement characteristics within asphalt mixture using digital image processing methods. Measurement, 163, 107948.
[8]Huang, H., Luo, J., Tutumluer, E., Hart, J. M., & Stolba, A. J. (2020). Automated segmentation and morphological analyses of stockpile aggregate images using deep convolutional neural networks. Transportation Research Record, 2674(10), 285-298.
[9] Chen, L., Shan, W., & Liu, P. (2021, December). Identification of concrete aggregates using K-means clustering and level set method. In Structures (Vol. 34, pp. 2069-2076). Elsevier.
[10] Dan, H. C., Bai, G. W., & Zhu, Z. H. (2021). Application of deep learning-based image recognition technology to asphalt–aggregate mixtures: Methodology. Construction and Building Materials, 297, 123770.18
[11] Han, J., Wang, K., Wang, X., & Monteiro, P. J. (2016). 2D image analysis method for evaluating coarse aggregate characteristic and distribution in concrete. Construction and Building Materials, 127, 30-42.
[12] Coenen, M., Schack, T., Beyer, D., Heipke, C., & Haist, M. (2021). Semi-Supervised Segmentation of Concrete Aggregate Using Consensus Regularization and Prior Guidance. arXiv preprint arXiv:2104.11028.
[13] Wang, W., Su, C., & Zhang, H. (2022). Automatic segmentation of concrete aggregate using convolutional neural network. Automation in Construction, 134, 104106.
[14] Waris, Muhammad Imran, Vagelis Plevris, Junaid Mir, Nida Chairman, and Afaq Ahmad. "An alternative approach for measuring the mechanical properties of hybrid concrete through image processing and machine learning." Construction and Building Materials 328 (2022): 126899.
[15] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).
[16] Liu, Wenya, Anzhi Yue, Weihua Shi, Jue Ji, and Ruru Deng. "An automatic extraction architecture of urban green space based on DeepLabv3plus semantic segmentation model." In 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), pp. 311-315. IEEE, 2019.
[17] Sandler, Mark, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. "Mobilenetv2: Inverted residuals and linear bottlenecks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510-4520. 2018.
[18] İsa, A. T. A. Ş. "Performance Evaluation of Jaccard-Dice Coefficient on Building Segmentation from High Resolution Satellite Images." Balkan Journal of Electrical and Computer Engineering 11, no. 1 (2023): 100-106.