Segmentation and Analysis of Cement particles in Cement Paste with Deep Learning

Hanjie Qian[a], Ye Li[b, 1], Jianfei Yang[a], Lihua Xie[a,**] and Kang Hai Tan[c,*]

[a] *School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore*

[b] *School of Civil and Environmental Engineering, Harbin Institute of Technology, Shenzhen, 518055, PR China*

[c] *School of Civil and Environmental Engineering, Nanyang Technological University, 639798, Singapore*

[*,**] *Corresponding author. Tel.: +65 6790 5285*

[1] *Co-first author*

[1] *Email addresses: HANJIE001@e.ntu.edu.sg (Hanjie Qian), liye@hit.edu.cn (Ye Li), yang0478@ntu.edu.sg (Jianfei Yang), ELHXIE@ntu.edu.sg (Lihua Xie), CKHTAN@ntu.edu.sg (Kang Hai Tan)*

Keywords: segmentation, machine learning, SEM image, cement paste, hydration degree, particle analysis

Abstract: Scanning electron microscopy (SEM) is a widely used method for the analysis of concrete micro structure. To quantitatively analyze the SEM images with high efficiency and accuracy, an automatic segmentation framework is proposed in this paper. The deep segmentation algorithm is purposely optimized from PointRend based on the characteristic of

SEM images to improve prediction accuracy, especially the performance around boundaries. Moreover, the SEM images can be segmented without additional treatment. Cement paste samples with 0.2 and 0.4 water-to-cement ratios are prepared and cured for 1, 3, 7, 14, and 28 days. Totally SEM images with 2267 labeled cement particles are included to build the dataset. From the results of intersection over union and pixel accuracy, the proposed algorithm outperforms the trainable waikato environment for knowledge analysis (WEKA) segmentation, Fully Convolutional Networks (FCN), and the original PointRend method. The segmentation results are used to calculate the hydration degree of two cement paste samples. Good agreement is obtained with the hydration degree calculated by using nonevaporable water in the samples for the 5 curing durations. At last, the shape of the cement particles is analyzed. Irregularity and roundness of the cement particles do not change significantly with an increase in curing duration.

1 Introduction

Concrete is the most widely used man-made material because of its excellent properties such as stiffness strength and durability for buildings and infrastructure [1]. In a concrete mix, the cement paste acts as a binder between the coarse and fine aggregates. Its heterogeneous micro structure, i.e., phase assemblage and distribution of the cement particles and hydration products, governs the durability and mechanical properties of macroscale concrete [2]. Therefore, a quantitative characterization of the cement paste is critical for the analysis of concrete hydration.

A variety of methods have been employed to characterize micro structure of concrete such as scanning electron microscopy (SEM) [3], mercury intrusion porosimetry (MIP) [4], nuclear magnetic resonance (NMR) [5], and X-ray micro-computed tomography (CT) [6]. Among these techniques, SEM is the most commonly adopted technique because of its high magnification, precision and the lowest cost[7]. The SEM images are manually processed to determine the water-to-binder ratio, slag fraction, and degree of hydration in hardened cement pastes based on subjective "point counting" [8, 9]. However, this visual-inspection method is subjective and labor-intensive.

Quantitative micro structure analysis of different constituents in the SEM images requires precise and reproducible segmentation, which involves partitioning the image into continuous and homogeneous regions. Conventional techniques such as thresholding and edge detection are adopted. Wong et al. [10] and Leite [11] segment cement-based materials from backscattered electron images based on the grayscale histogram. The thresholds are selected from the lowest inflection point of the grayscale histogram, but it is challenging to eliminate individual judgment and bias when selecting the threshold values. An automated thresholding method is proposed [12, 13] and morphological filters are combined with gray-level thresholding [7, 14]. However, the threshold technique fails in classification when the gray level of different phases overlaps with each other. Moreover, image resolution, magnification, and noise of the SEM images, which vary with sample preparation and equipment operation, strongly influence the selection of thresholds.

Edge detection is a challenging task in the segmentation of SEM images due to the noise and similar grey scales between different phases of concrete. Therefore, image segmentation algo-

rithms using appropriate filters have been applied to extract explicit features and reduce noise on SEM images. Yang and Buenfeld [7] propose a method combining the grey-level thresholding, filtering, and binary operations. Lee and Yoo [15] propose a segmentation method based on Laplacian of Gaussian filter and watershed algorithm and global-local threshold method. Feng et al. [16] propose a segmentation technique using the hybrid ridge signal detector. However, prior knowledge is needed to properly select the salient features to ensure versatility, and these algorithms are also highly influenced by the image quality.

Recent developments in computer vision and machine learning have made significant achievements in visual recognition. Machine learning-based models have been successfully applied to image segmentation tasks for structural damage detection [17]. Bangaru [18] applies machine learning methods such as K-NearestNeighbor(KNN) and Support Vector Machine(SVM) to estimate the hydration degree of concrete based on SEM images. Pixel-based image segmentation method based on machine learning associates each pixel to a unique class label such as cement particle, hydration product, and pore region [19, 20]. However, the machine learning classifier is trained based on the pixels selected on a particular image. Whether the trained model is robust for other images with different processing and capturing conditions is questionable. The precision of image segmentation has been improved significantly by semantic segmentation algorithms based on deep learning such as convolutional neural network (CNN), in which no explicit feature needs to be defined manually. Some researchers [21, 22, 23, 24] apply primary deep learning segmentation to concrete samples. However, complex treatment of samples such as coloring and pre-processing of images based on threshold is inevitable. Moreover, the algorithms used are designed for popular datasets, such as

COCO[25] or GTA5[26] but are not tailored for the SEM image of concrete samples. Such coarse application ignores the characteristics of data in the civil engineering field and hinders the performance of these deep learning algorithms.

In conclusion, previous researches either lack automation or are cumbersome and require preprocessing of samples. Most of the segmentation results are not accurate enough for quantitative analysis. The above-mentioned issues form the motivation for facilitating objective, adaptable, and automated algorithms for the segmentation of SEM images. The contribution of this paper can be summarized as follows: This paper presents an automatic deep learning-based segmentation algorithm for SEM images of cement paste. Secondly, the algorithm is specially designed and improved based on the characteristics of SEM images. And it performs better than other common methods, such as FCN [27] and TWS[28] in this field. Finally, quantitative analysis can be implemented based on accurate segmentation. The hydration degree of the paste samples at different curing duration is calculated and compared to experimental results from loss-on-ignition (LOI) measurements. The irregularity and roundness of the unhydrated cement particles are analyzed according to segmentation results.

2 Methodology

The framework of this study is presented in Fig. 1. Two cement paste samples with different water-to-cement (W/C) ratios are cured for five curing durations. Multiple SEM images of the samples are taken, cropped, and labeled to generate a dataset for training and validation

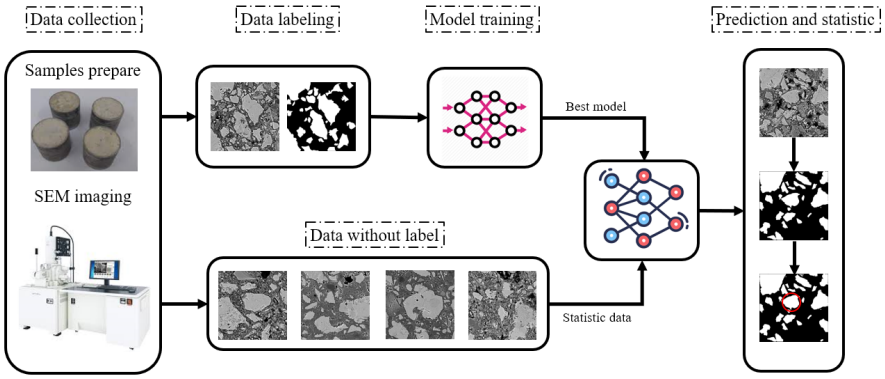

Figure 1: Framework for SEM image segmentation.

of a deep learning model. Thereafter, the trained model is used to segment unhydrated regions. The segmented images are then used for the analysis of the cement particles and the measurement of hydration degree. The steps are discussed in detail in this section.

2.1 Sample preparation and data collection

The two cement pastes used in the present study have 0.2 and 0.4 W/C ratios, respectively. The cement pastes were mixed by using a 1 Liter Hobart planetary mixer. The fresh pastes were mixed for 5 min until they became homogeneous and consistent. The samples were then cast into $\phi 20 \times 30$ mm cylindrical molds and compacted using a vibrating table. When no significant amount of air bubbles escaped the surface, the samples were then covered with a plastic sheet for 24 h prior to demolding. After removing the mold, all the samples were stored in lime-saturated water for the duration of 3, 7, 14, and 28 days.

For each curing duration, one cylindrical sample was sectioned at the center part to produce

a cube block sample with a 5 mm side length. Thereafter, the samples were immersed in 99% isopropanol for 24 h to remove free water and oven-dried for another 24 h at 60 °C. Low viscosity epoxy was then impregnated into the pores and voids of the samples to preserve the delicate micro structure during grinding and polishing and to provide contrast under SEM observation. The samples were polished with P320, P800, and P1200 emery papers for 30 s each and then finely polished for 5 min with BUEHLER TexMet™ P polishing cloths and 1 μm MetaDi monocrystalline suspension.

Parallel samples were prepared from the same batch. A sample will be processed (hydration stoppage, epoxy fixing, and polishing) for SEM observation at a certain curing duration, but other samples will not be disturbed. Before SEM observation, the samples were coated with gold for 40 s under vacuum. The SEM images were obtained under 15 kV at a magnification of 1000× by using an SEM Phenom Prox G6 under backscattered electron (BSE) imaging mode. Before taking each image, the brightness and contrast were calibrated to make the histogram centered and ranged between the greyscale 0 to 255. Thirty-two images were taken for each sample. The SEM images with the bottom label removed have a resolution of 3840×2400, which contained too many cement particles for labeling, and it was too large for convolutional operation in the algorithm. Therefore, each picture was divided into 800×800 pieces as shown in Fig.2. The smallest dimension of a particle that could be distinguished had a size of 10 pixels, i.e., 10×0.135 μm/pixel. Therefore, the chosen magnification did not seem to compromise the accuracy since the tiny cement particles were expected to have fully hydrated.

Table 1 summarizes the organization of the dataset. 115 images were selected to build the

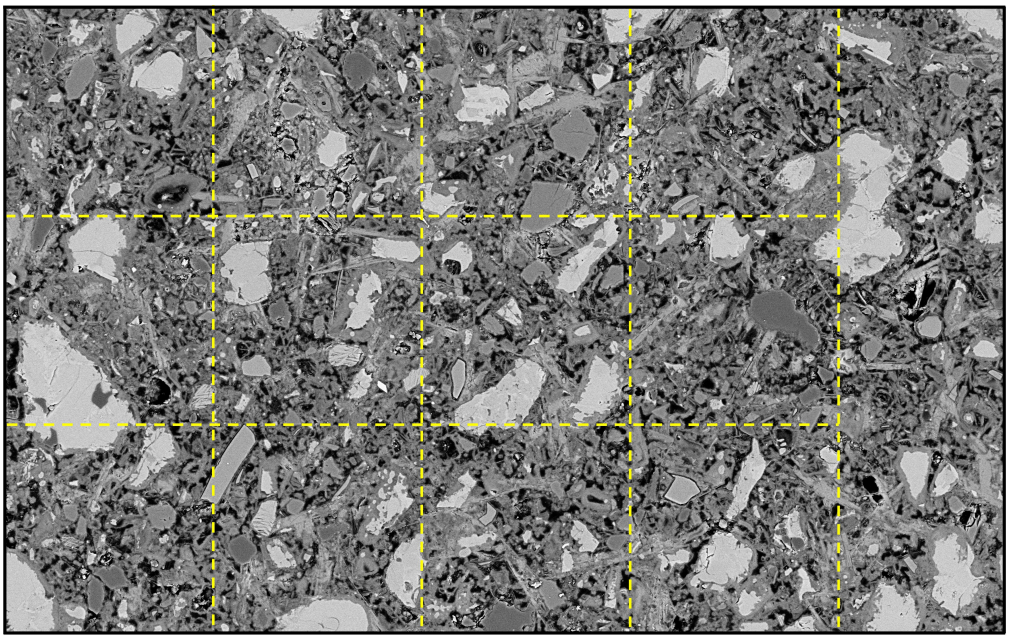

Figure 2: Example of a SEM image and cropping

dataset, among which 61 and 54 images were from the samples with 0.2 and 0.4 W/C ratios, respectively. The images were randomly mixed in a ratio of 4:1 for training and validation sets. Only the cement particles were labeled by LabelMe as the region of interest (ROI), which were also called instances as shown in Fig. 3. In this paper, we label the ROI manually, which is the common way in the field of supervised learning. The advantage of the manual label is their high accuracy, and we can view the manual label as ground truth. And the different colors in Figure 3 mean different instances of ROI. The number of images after cropping is shown in Table 2.

Table 1: The organization of segmentation dataset.

	Images	Instances
Training set	92	1835
Validation set	23	432
W/C=0.2	61	1289
W/C=0.4	54	978

Table 2: The distribution of sample images.

Curing duration	W/C=0.2	W/C=0.4
1 day	288	288
3 days	558	288
7 days	288	306
14 days	270	270
28 days	270	270

2.2 Configuration of network

Semantic segmentation is a hot topic in the field of deep learning. There are a large number of algorithms focusing on segmentation tasks based on standard datasets like COCO and Cityscapes[29]. However, the photos captured in daily life are vastly different from the SEM images in the following aspects:

1. The SEM images in BSE mode are usually taken from finely polished samples with a smooth surface. Most of the features are based on 2D shapes and pixel values, rather

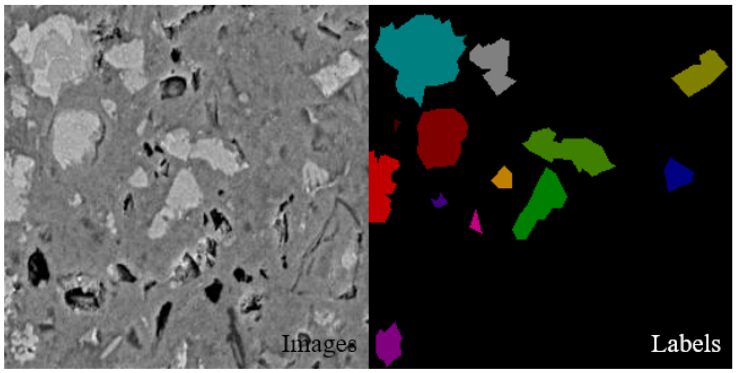

Figure 3: Example for labelling of the SEM image.

than 3D morphological information.

2. Unlike the data from the common dataset, the size of SEM images tends to be large to maintain useful information, or the pictures can be compressed at the cost of image quality and prediction precision.

3. Boundaries of the objects are irregular polygonals with zig-zags along the edges, which cannot be well predicted by traditional algorithms.

In view of points 2 and 3, PointRend algorithm [30] is selected. The algorithm is a deep learning module that treats segmentation as an image rendering problem. The algorithm is specifically designed to compute high-resolution segmentation maps with a better boundary detail since it adaptively divides the mask predictor into two sections, one for coarse prediction and one for points prediction. The configuration of the network is shown in Fig.4. The network architecture of PointRend is briefly introduced in this part.

- Feature extraction. For most deep learning models, the first step is to extract robust features. In PointRend, ResNet50[31] is chosen as the backbone. Meanwhile, in order to

fuse the information from different levels, FPN[32] is introduced to capture the features at different levels. Then all the features, sizes from 4×4 to 64×64, are fused and passed to the ROI Align layer to get a feature of 14×14. In the next step, a region of interest (ROI) is proposed from the feature map. Here, the RPN network of Faster-RCNN[33] is applied, which will generate a proposal based on anchors automatically. These anchors uniformly distribute on the original picture, so the proposed ROI needs to be mapped into feature space through the ROI layers.

- ROI align layer [34]. Now the feature map has the size of $N \times C \times W \times H$, where N is the number of samples, C is the channel number, and W is the width and H the height. As mentioned above, the proposed ROI generated from RPN is based on image space, so the problem is how to convert the coordinates of the image space, X, Y, into the coordinates of the feature space X_f, Y_f. If the original image size is W_{origin}, H_{origin}, in Faster-RCNN, the idea is to adjust the coordinates by ROI pooling.

$$X_f = round(X \times \frac{W}{W_{origin}}) \tag{1}$$

  However, in the segmentation task, the rounding error in the above equation is unacceptable, so an ROI align layer is proposed, just by removing the rounding term. But in this case, the coordinates in the feature space will not be integers, their values can be calculated by bi-linear interpolation.

- Classifier. For each feature of ROI, it will pass a classifier to predict the category it belongs to.

- Box predictor. Box predictor is used to measure the loss between predicted boxes and

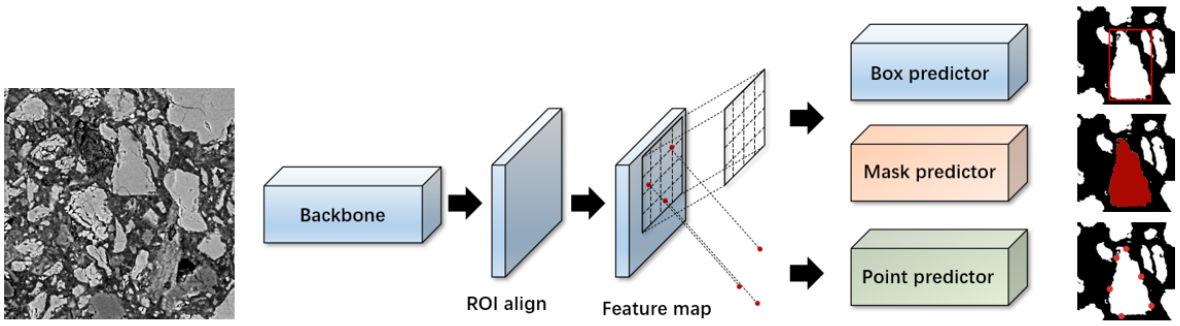

Figure 4: Configuration of PointRend for SEM segmentation.

ground truth.

- Mask coarse predictor. A relatively simple network, which requires smaller computation and faster speed, is applied to predict a coarse mask prediction with the size of 7×7 or 14×14.

- Mask point predictor. Points with the highest uncertainty will be predicted by mask point predictor to which category they belong. Then refine the prediction result.

Based on the structure of the network, the total loss of PointRend can be formulated as:

$$Total_{loss} = class_{loss} + regression_{loss} + mask_{loss} + point_{loss} \tag{2}$$

where $class_{loss}$ is the difference between the prediction instance class and the class ground truth; $regression_{loss}$ measures the deviation of predicted boxes and box ground truth; $mask_{loss}$ calculates the error between predicted masks and mask ground truth; $point_{loss}$ is the difference between specific points prediction results and ground truth.

In this study, the PointRend algorithm is improved in terms of gradient and uncertainty. By the combination of gradient and uncertainty, the designed algorithm pays more attention to

the boundaries of ROI and gets more accurate segmentation results and continuous boundaries compared to other methods. FCN [27] and Trainable Weka Segmentation (TWS) [28] are also trained and evaluated based on the dataset. Their results are compared with the segmentation results of the PointRend and the modified PointRend.

FCN can be viewed as a milestone of deep learning-based segmentation. By establishing the space relationship between the feature layer and the input, segmentation results can be obtained from interpolation. But prediction results based on interpolation are not precise around boundaries, especially for images with large resolution [30]. FCN is a widely used algorithm in the real application, including civil engineering. [22, 35, 36, 37, 38].

TWS is a machine learning tool that converts the segmentation problem into a pixel-level classification problem and it is popular for the segmentation task in the real engineering application. First, input pixels are manually labeled and applied as a training collection. Certain training features are then selected. A fast random forest (RF) algorithm [39] is applied as a classifier. After training the RF classifier, the images of the test group are segmented. The detailed network configuration of TWS can be found in literature [28].

2.3 Estimation for the degree of hydration

The degree of hydration is estimated by two methods [8, 10]. The first method analyzes the remaining percentage of cement particles from the segmented SEM image. Stereology provides practical techniques for extracting quantitative information about a three-dimensional material from measurements made on two-dimensional planar sections of the material[40].

It is an efficient tool in many applications of microscopy. Indeed, cement hydration is a 3-dimensional process, and the SEM images captures information on a cross section of the samples. Nevertheless, the hydration degree was calculated based on volume fraction of the cement particles, which can be obtained by measuring the area fraction on a polished plane section and this is a classical application of stereology[40, 41]. The degree of hydration M_s is estimated using Eq. (3) where V_{AH} is the ratio of the volume of anhydrous cement determined through image analysis of the segmented image. V_C is the volume of the total cement paste calculated from the initial mix proportion. Loss-on-ignition (LOI) measurements were also performed to compare with the results from the image analysis [42]. The degree of hydration M_c is estimated using Eq. (4), where W_B is the weight of the nonevaporable water. $W_{B\infty} = 0.24$ is the nonevaporable water content for a fully hydrated sample [43]. The paste samples were oven-dried at 105 °C for 24 h to eliminate the free water. After recording the dry weight, the samples were ignited in an electric furnace for 4 h at 900 °C. The weight difference after 105 °C drying and 900 °C ignition was used as the nonevaporable water.

$$M_s = 1 - \frac{V_{AH}}{V_C} \tag{3}$$

$$M_c = \frac{W_B}{W_{B\infty}} \tag{4}$$

2.4 Analysis of cement particles

With segmentation of the SEM images, the size distribution of the cement particles was analyzed. After training, the best algorithm among the FCN(76.95), TWS(58.31), PointRend(80.20), and the modified PointRend(81.87) is chosen based on validation results. The segmentation

results are binary images with 0 being the background and 1 being the cement particles. Based on the segmentation results, a minimum circumscribed circle is applied to each cement particle. Its diameter is adopted as the size of the cement particle. Meanwhile, the irregularity and roundness of the cement particle are also analyzed.

3 Improvement of PointRend

Before the discussion of the optimization of the original method, several concepts need to be clarified first.

- Uncertainty. For each pixel in the feature map, it may belong to the background, or it may belong to the object. And uncertainty measures the confidence of prediction results. A prediction point with higher uncertainty is equivalent to lower confidence. For instance, for a segmentation task with two classes, the goal is to separate foreground from background. The uncertainty is defined as[30]:

$$uncertainty = -abs(logits_{foreground}) = -L_1(logits_{foreground}, 0) \tag{5}$$

  Here, the logits mean the raw outputs of the network and the logits value will be fed into the sigmoid function to get the probability for each category.

$$probability = sigmoid(logits) = sigmoid(net(input)) \tag{6}$$

  Actually, the above equation measures the $L1$ distance between 0 and the logits prediction. And the logits prediction will be converted into probability as the final results,

while $sigmoid(0) = 0.5$, which means the greatest uncertainty. For segmentation with more than two classes, the uncertainty can be defined as:

$$uncertainty = -1 \times (mask[:, 0] - mask[:, 1]) \tag{7}$$

Equation 7 estimates uncertainty as the difference between the top first and top second predicted logits, e.g., $mask[:, 0]$ and $mask[:, 1]$.

It can be observed that the value of uncertainty is always smaller than 0 so that values which are closer to 0 mean larger uncertainty.

- Point selection. If the size of the feature map is $H \times W$, then in the first step, $H \times W \times k$ points will be generated uniformly, where k means the oversampling ratio. Secondly, select $H \times W \times \beta$ points from the $H \times W \times k$ points based on the biggest uncertainty, where β is the important sample ratio[30]. Lastly, select the rest $H \times W \times (1 - \beta)$ points randomly.

One of the key concepts of the PointRend algorithm is fine point prediction based on uncertainty and prediction. The whole framework of PointRend can be summarized as follows. Firstly, to generate a coarse prediction based on mask coarse predictor. Secondly, to select top-k($H \times W \times \beta$) points, which are generated by point selection, with the biggest uncertainty. Thirdly, to predict the class information of these top-k points and compare them with the ground truth to calculate the difference. Finally, to refine coarse prediction results based on these top-k points and interpolation and repeat steps 2-4. Compared to other methods, such as FCN and MaskRCNN[34], which directly enlarge the coarse prediction into the original size, prediction results of PointRend are more precise and boundary is clear, especially for

images with large resolution.

Moreover, after we consider the characteristics of SEM data and the concept of uncertainty, we can make some optimizations to the algorithm according to our specific tasks. Our optimizations originate from such observations and assumptions: The SEM images, which are taken from ground samples with a smooth surface, contain little 3D morphological information so most of prediction results are based on 2D shapes and pixel values. Additionally, most ROI areas are irregular and simply connected regions, which means that the accuracy of outer boundaries determines the precision of segmentation results. Meanwhile, an obvious pixel value difference can be observed around boundaries of different compounds. But it is important to point out that not all obvious pixel value differences represent ROI areas. Actually, for traditional methods that highly rely on pixel values, the performance is not ideal due to the noise. Based on the above discussion, the idea is that the performance of the segmentation network should be better if points on the boundary can be assigned with larger uncertainties and then focus on their points loss. The optimized PointRend algorithm is proposed as follows.

1. To generate coarse prediction and uncertainty points which is the same as the original method. The size of coarse prediction and uncertainty points $U(x, y)$ are $H \times W$ and $H \times W \times k$ respectively.

2. To calculate the gradient of each pixel in original images, $G(x, y)$. Based on thresholds(th_1 and th_2, $th_1 < th_2$) and the Canny algorithm to calculate gradient maps $GM(x, y)$ with only 0 and 1.

3. To calculate the original coordinates of uncertainty points in the image space and then obtain gradients of these points based on bilinear interpolation from $GM(x, y)$. The gradients of these points can be named as $GM'(x, y)$ with the size of $H \times W \times k$.

4. Element-wise multiply $U(x, y)$ and $-GM'(x, y)$ to get $U'(x, y)$. Then select top-k points with the biggest uncertainty from $U'(x, y)$. Finally to calculate the difference of these points and ground truth as points loss to update the whole network by back-propagation.

The framework of the optimized algorithm can be explained in Fig 5. The reason to convert $GM'(x, y)$ to $-GM'(x, y)$ is that the values of original uncertainty are smaller than 0, and points around boundaries are required to be assigned with a larger uncertainty. After the multiplication, values of uncertainty points will be larger than 0, and points that are not around the boundaries will be assigned with values close to 0.

4 Results and discussion

All of the code is implemented by Pytorch and experiments are performed on a single GPU(RTX2080). All the initial parameters of the model are loaded from the model pre-trained on the COCO dataset. This is a common way to improve the accuracy and efficiency of the model in most cases. The training ends at a maximum iteration of 10,000 to ensure sufficient optimization. A stochastic gradient descent (SGD) is adopted with a batch size of 8. A learning rate l_r is set to be changed with the epoch and a momentum $\mu_j = 0.9$.

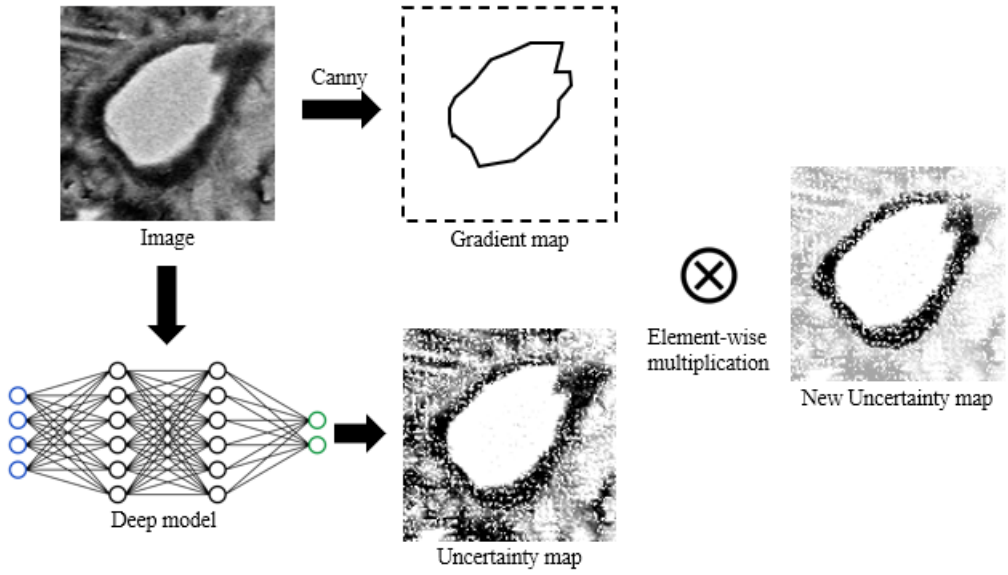

Figure 5: Framework of the optimized algorithm. Compared with the original uncertainty map, only uncertainty around boundaries is retained after multiplication.

4.1 Performance comparison of algorithms

Two metrics i.e., Pixel accuracy (PA) and intersection over union (IOU) are used to evaluate the prediction performance of the applied methods.

- PA (pixel accuracy). PA is a straightforward concept describing the ratio of the corrected prediction pixels to the total number of pixels for one class, i.e.,

$$PA = \frac{Num_{correct}}{Num_{total}} \tag{8}$$

- Intersection over union(IOU). IOU is a metric to measure the accuracy of object detection or segmentation. As shown in Fig. 6, it is defined as the ratio of the area of intersection and the area of the union. The IOU has a higher score when the predicted bounding box overlaps more with the ground-truth bounding box.

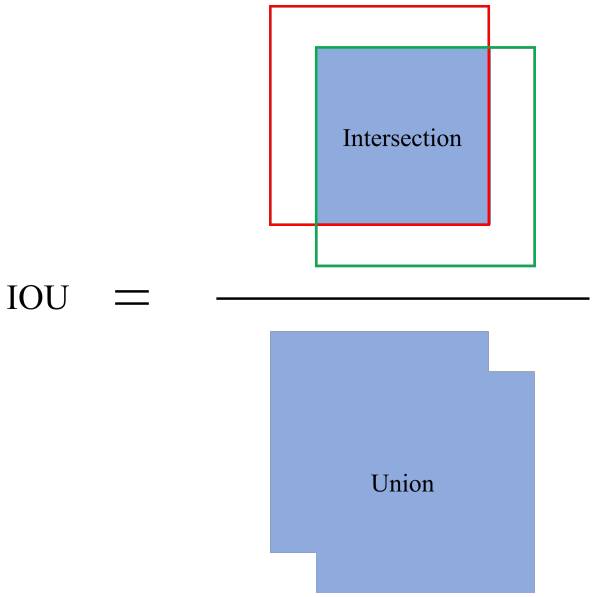

Figure 6: IOU reflects the degree of overlap of predicted boxes and ground truth.

Table 3 lists the results of PA and IOU of the four methods. In addition, an example of the original BSE image and the segmented images by the four methods is illustrated in Fig. 7. It is obvious that the accuracy of the deep learning methods (FCN and PointRend) is much better than the pixel-based method TWS. The PA and IOU of the FCN model were 4.48 and 18.64 points higher than that of the TWS model, respectively. As shown in Fig. 7g, segmentation using TWS could not sufficiently distinguish the cement particles and the hydration products due to similarities in brightness, texture, and color, resulting in misinterpretation. Moreover, the particle boundaries cannot be well segmented clearly. In general, TWS is a classical algorithm based on human designed features, i.e., the pixel values. However, it is not a robust feature for segmentation tasks. And the rise of deep learning in recent years proves that deep features are more robust compared with designed features. From Fig. 7c and Fig. 7f, the disturbance of noise is lowered significantly in the prediction

of FCN but an obvious jagged boundary can be observed. FCN is the milestone of deep segmentation algorithm, but the algorithm simply calculate the segmentation results from feature maps by interpolation. For example, the size of feature map is 14*14 and original size of ROI is 56*56, then the interpolation is acceptable, but if the original size is 224*224, then obvious zig-zag boundaries can be observed, just like what we show in Fig7.f. So for images with large size, FCN can not perform well and it will ignore precise structure.

PA and IOU of the PointRend model increase by 0.27 and 3.25, respectively, from the FCN model. By fine-points prediction and interpolation, a more precise segmentation can be obtained from its results. Such improvement is more conspicuous for large resolution images and PointRend is a more suitable algorithm for large images, such as SEM concrete images.

By recalculating the uncertainty map based on gradient, the modified PointRend model further improves the segmentation accuracy by 0.51 and 1.67 for PA and IOU, respectively. Compared with FCN, PointRend has a much more complex network structure and a more refined feature extraction method. The ModPointRend we designed requires almost no more computational consumption, and the network structure does not need to be changed. And in this case, PointRend is 3.27 higher than FCN, and ModPointRend is 1.67 higher than PointRend. It can be seen from Fig. 7a and Fig. 7d that the boundaries of the segmented cement particles by our method are smoother and more continuous than the original one. Such results are consistent with our improved algorithm design. Firstly, most of the points in the region of interest will be assigned with lower uncertainty and less chance to be selected for further point prediction, which leads to fewer holes in a prediction region. Secondly, more attention is paid to boundaries and better results can be obtained.

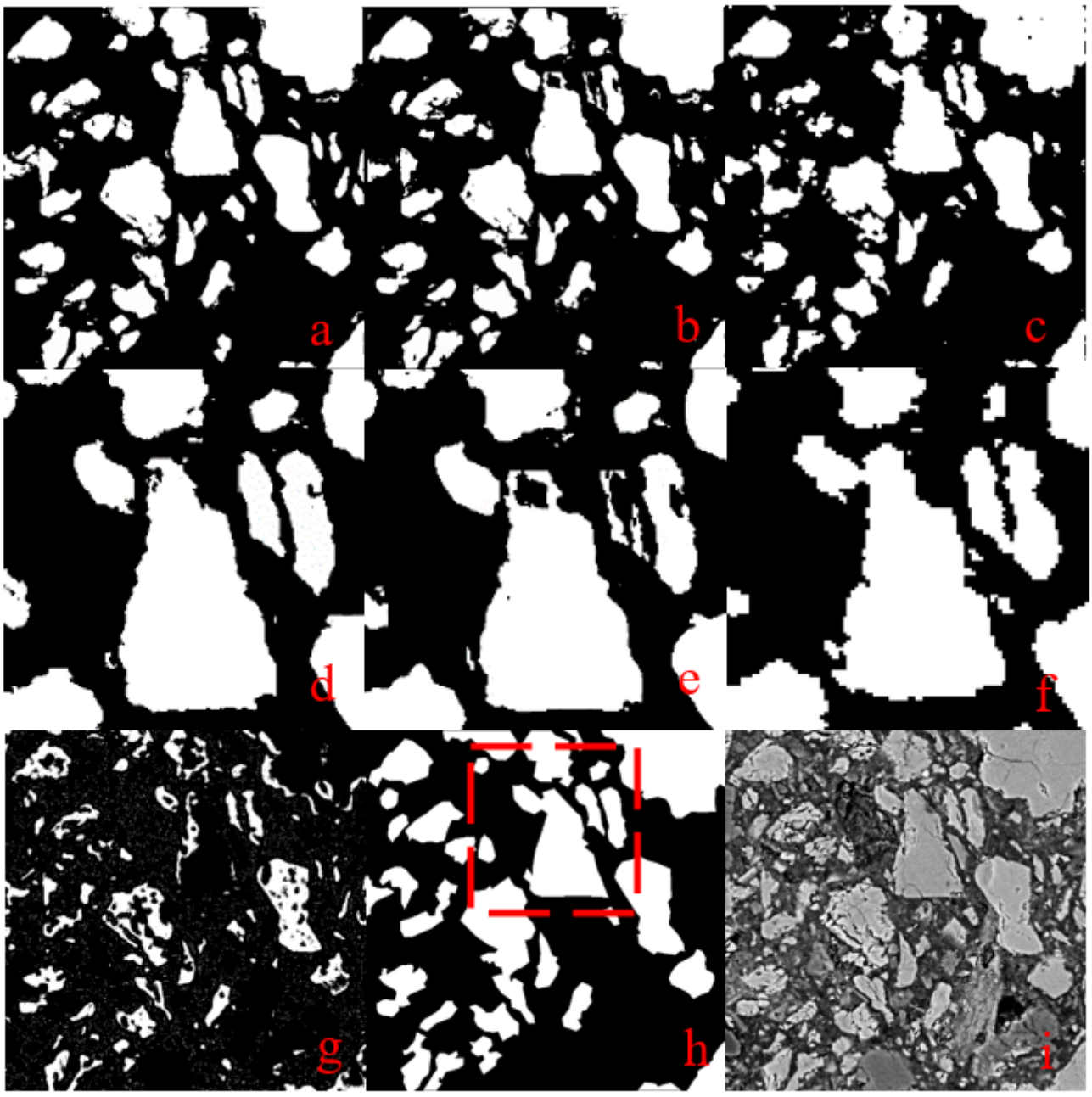

Figure 7: Comparison of the segmentation results generated by (a) Modified PointRend, (b) PointRend, (c) FCN, (d) Details of modified PointRend, (e) Details of PointRend, (f) Details of FCN, (g) TWS (h) Ground truth and (i) Corresponding SEM image

Table 3: Segmentation accuracy of the four methods.

Metric	PA	IOU
TWS	90.97	58.31
FCN	95.45	76.95
PointRend	95.72	80.20
ModPointRend	96.23	81.87

To analyze the accuracy of the ModPointRend method in detail, a more detailed evaluation system, average precision (AP) is introduced. AP is the average value of AP with a threshold from 0.5 to 0.95 with an interval of 0.05. AP50 means the AP value of instances whose IOU is higher than 0.5. Correspondingly, AP75 represents the AP value whose IOU is higher than 0.75, which is a stricter criterion than AP50 so its value would be lower than AP50. The APs, APm and APl represent AP for an area smaller than 32×32, a moderate area larger than 32×32 but smaller than 96×96, and a large area larger than 96×96, respectively. The AP is explained in detail in reference [44].

Table 4: Comparison of accuracy between PointRend and PointRend modification.

Method	AP	AP50	AP75	APs	APm	APl
PointRend	51.18	74.66	61.80	7.09	49.44	74.64
ModPointRend	54.19	78.75	63.12	7.77	53.94	75.54

The AP results are listed in Table 4. Compared with IOU, AP is a more detailed metric and provides more useful information. The scores of the ModPointRend are higher than the original PointRend for all the metrics, especially the AP and AP50. However, the AP

score decreases with a lower area. For APl and APm, the results of the ModPointRend are around 75 and 50, respectively, which can be considered satisfactory for segmentation tasks. But for APs, the score is lower than 10. This can be attributed to that the small ROIs contain limited information, such as texture, color, and brightness. Therefore, it is difficult for correct annotation based on visual judgment at the magnification of 1000. This can mislead the training process of the network. At the same time, there is little information for the model to make a prediction. If the cement particles are mistaken as hydration products the true positive (TP) will be changed to the false positive (FP) during the calculation of AP. Nevertheless, although the APs shows a lower score, the total area of the ROIs smaller than 32×32 only occupies a very low fraction and thus has a limited impact on the calculation of the proportion of unhydrated area in the images with 800×800 pixels.

4.2 Hydration degree of the cement paste samples

The W/C ratio is one of the important parameters for strength and durability properties of hardened concrete. Since the modified PointRend model has excellent performance, it is applied to the entire dataset consisting of a total of 346×9 images for segmentation. Based on reference [10], the total area covered by the frames is at least 3 times larger than the requirement for a representative area. The segmented images are used to analyze the degree of hydration by following the method described in Section 2.3. The estimated hydration degree of the two cement pastes as a function of curing duration is compared against the measured values from the LOI test in Fig. 8. It is clear that the degree of hydration calculated by the two methods agreed very well, indicating the effectiveness and accuracy of

this modified segmentation algorithm. With higher time and economic efficiency this image-based method is promising in analyzing pore structure and phase distribution of the cement matrix.

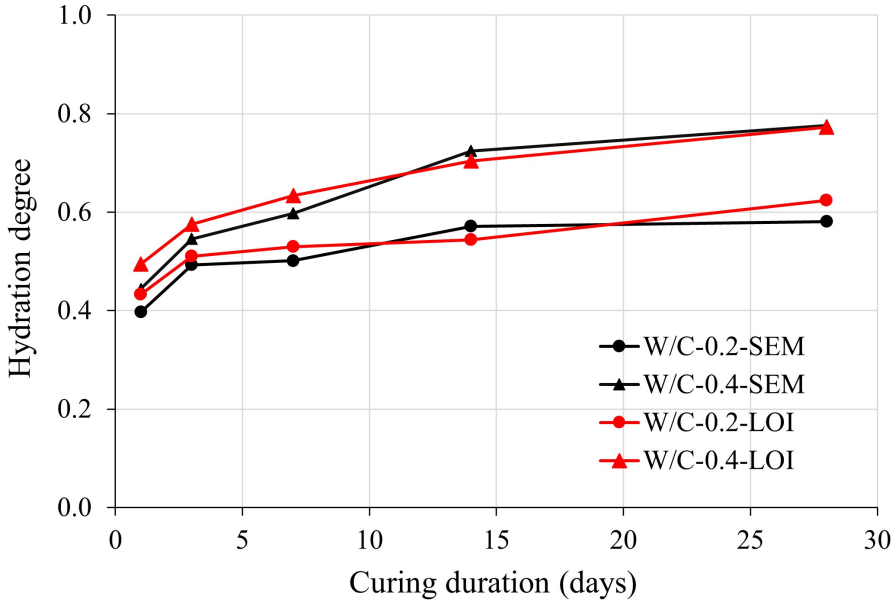

Figure 8: Hydration degree of the paste samples.

The percentage difference of hydration degree between the two methods for the sample W/C-0.4 at 1, 3, 7, 14, and 28 days are 5.1%, 2.9%, 3.6%, 2.0%, and 0.4%, respectively. This deviation is considered low as the variation between concrete samples is usually significant. The short term (less than 14 days) degree of hydration is slightly higher measured by LOI than that measured by image segmentation. This could be attributed to additional hydration during the 105 °C drying process, or decomposition of other volatile phases, such as carbonates. The degree of hydration increased from around 40% to around 60% for the paste sample W/C-0.2 and the corresponding degree of hydration is greater for the paste sample W/C-0.4 as expected. It could also be noticed that the rate of hydration slowed down with increasing curing duration. This trend is earlier for the sample W/C-0.2 due to the limited

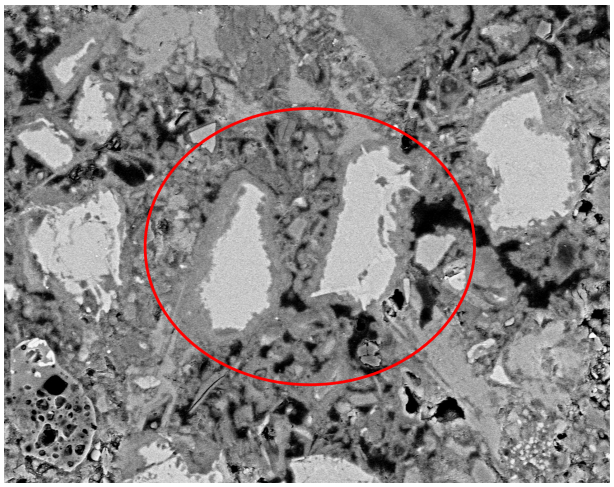

Figure 9: Example of a cement particle after 7 days of hydration.

water for hydration.

4.3 Analysis of cement particle

The ModPointRend is used to generate a mask for each cement particle. Two indexes, irregularity and roundness, are calculated from the masks to analyze the shape of the unhydrated cement particles and are defined as follows:

$$I = \frac{d_{max}}{d_{min}} \tag{9}$$

$$C = \sqrt{\frac{4\pi A}{P^2}} \tag{10}$$

where, d_{max}, d_{min}, A, and P are the diameter of the maximum inscribed circle, the diameter of the minimum circumscribed circle, the area, and perimeter of the cement particles, respectively. The values of irregularity and roundness indexes range from 0 to 1, with 1 representing a circle. The irregularity and roundness indexes of the cement particles as a

	Irregularity		Roundness	
Group	Mean	Std	Mean	Std
0.2-1	0.483	0.118	0.759	0.09
0.2-3	0.487	0.119	0.76	0.086
0.2-7	0.49	0.118	0.766	0.087
0.2-14	0.5	0.124	0.756	0.099
0.2-28	0.496	0.118	0.762	0.092
0.4-1	0.486	0.128	0.758	0.094
0.4-3	0.463	0.13	0.742	0.104
0.4-7	0.481	0.133	0.747	0.105
0.4-14	0.493	0.134	0.754	0.127
0.4-28	0.482	0.141	0.728	0.127

Table 5: Irregularity and roundness of cement particles in different groups.

function of curing duration in the two samples with different W/C ratios are listed in Table 5. The mean value and standard deviation of these two parameters change marginally with an increase in curing duration and the W/C ratio does not influence the shape of the cement particles significantly. From Fig. 9, hydration happens evenly on the surface of the cement particle after 7 days of hydration. This explains the unchanging shape of the cement particle which shows a rough surface but not represented by the irregularity and roundness.

5 Conclusion

Previous quantitative analysis of micro structure always highly relied on human experience and was labor-intensive with a low degree of automation. In this paper, we reviewed the previous methods in this field for segmentation and analyzed their disadvantage as well as the characteristics of segmentation of SEM images. We proposed a deep learning-based framework for concrete SEM images segmentation, which is automatic and precise. An optimized PointRend algorithm was proposed based on the characteristic of SEM images. Its performance was compared with different algorithms under various evaluation metrics. It was shown that our method is more accurate than the original method and other common segmentation algorithms in this field. We also analyzed the relationship between curing time and the shape parameters of concrete particles, which is impossible to obtain except for SEM image segmentation. Deep learning-based segmentation demonstrated its power in this research, and it will help us in future research of concrete design and understanding.

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.

Acknowledgments

This work was partially supported by the National Key Research and Development Program of China (Grant No. 2021YFF0500801), the National Research Foundation, Singapore, and Ministry of National Development, Singapore, under its Cities of Tomorrow R&D Programme(CoT Award No. COT-V2-2019-1), Shenzhen Science and Technology Program (Grant No. RCYX20200714114525013) and Open Funding of State Key Laboratory of High Performance Civil Engineering Materials (Grant No. 2021CEM006).

References

[1] Emily Elhacham, Liad Ben-Uri, Jonathan Grozovski, Yinon M Bar-On, and Ron Milo. Global human-made mass exceeds all living biomass. *Nature*, 588(7838):442–444, 2020.

[2] Della M Roy and GM Idorn. *Concrete microstructure*. Number SHRP-C-340. 1993.

[3] Stefan Jacobsen, Jacques Marchand, and Hugues Hornain. Sem observations of the microstructure of frost deteriorated and self-healed concretes. *Cement and Concrete Research*, 25(8):1781–1790, 1995.

[4] Hongyan Ma. Mercury intrusion porosimetry in concrete technology: tips in measurement, pore structure parameter acquisition and application. *Journal of porous materials*, 21(2):207–215, 2014.

[5] Yanliang Ji, Leo Pel, and Zhenping Sun. The microstructure development during bleeding of cement paste: An nmr study. *Cement and Concrete Research*, 125:105866, 2019.

[6] Anton du Plessis and William P Boshoff. A review of x-ray computed tomography of concrete and asphalt construction materials. *Construction and Building Materials*, 199:637–651, 2019.

[7] R Yang and NR Buenfeld. Binary segmentation of aggregate in sem image analysis of concrete. *Cement and Concrete Research*, 31(3):437–441, 2001.

[8] MHN Yio, JC Phelan, HS Wong, and NR Buenfeld. Determining the slag fraction, water/binder ratio and degree of hydration in hardened cement pastes. *Cement and concrete research*, 56:171–181, 2014.

[9] X Feng, Edward J Garboczi, Dale P Bentz, Paul E Stutzman, and Thomas O Mason. Estimation of the degree of hydration of blended cement pastes by a scanning electron microscope point-counting procedure. *Cement and concrete research*, 34(10):1787–1793, 2004.

[10] HS Wong and NR Buenfeld. Determining the water–cement ratio, cement content, water content and degree of hydration of hardened cement paste: Method development and validation on paste samples. *Cement and Concrete Research*, 39(10):957–965, 2009.

[11] MB Leite and PJM Monteiro. Microstructural analysis of recycled concrete using x-ray microtomography. *Cement and Concrete Research*, 81:38–48, 2016.

[12] Romy Suryaningrat Edwin, Mushthofa Mushthofa, Elke Gruyaert, and Nele De Belie. Quantitative analysis on porosity of reactive powder concrete based on automated analysis of back-scattered-electron images. *Cement and Concrete Composites*, 96:1–10, 2019.

[13] Ye Li, Yao Zhang, En-Hua Yang, and Kang Hai Tan. Effects of geometry and fraction of polypropylene fibers on permeability of ultra-high performance concrete after heat exposure. *Cement and Concrete Research*, 116:168–178, 2019.

[14] M Ben Haha, K De Weerdt, and B Lothenbach. Quantification of the degree of reaction of fly ash. *Cement and Concrete Research*, 40(11):1620–1629, 2010.

[15] Jang Hee Lee and Suk In Yoo. An effective image segmentation technique for the sem image. In *2008 IEEE international conference on industrial technology*, pages 1–5. IEEE, 2008.

[16] Hanying Feng, Jun Ye, and R Fabian Pease. Segmentation-assisted edge extraction algorithms for sem images. In *Photomask Technology 2006*, volume 6349, page 63491L. International Society for Optics and Photonics, 2006.

[17] Honglan Huang and Henry V Burton. Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning. *Journal of Building Engineering*, 25:100767, 2019.

[18] Srikanth Sagar Bangaru, Chao Wang, Marwa Hassan, Hyun Woo Jeon, and Tarun Ayiluri. Estimation of the degree of hydration of concrete through automated machine learning based microstructure analysis–a study on effect of image magnification. *Advanced Engineering Informatics*, 42:100975, 2019.

[19] Yaokun Wu, Siddharth Misra, Carl Sondergeld, Mark Curtis, and Jeremy Jernigen. Machine learning for locating organic matter and pores in scanning electron microscopy images of organic-rich shales. *Fuel*, 253:662–676, 2019.

[20] Chunxiao Li, Dongmei Wang, and Lingyun Kong. Application of machine learning techniques in mineral classification for scanning electron microscopy-energy dispersive x-ray spectroscopy (sem-eds) images. *Journal of Petroleum Science and Engineering*, 200:108178, 2021.

[21] Xiao Liang. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with bayesian optimization. *Computer-Aided Civil and Infrastructure Engineering*, 34(5):415–430, 2019.

[22] Yu Song, Zilong Huang, Chuanyue Shen, Humphrey Shi, and David A Lange. Deep learning-based automated image segmentation for concrete petrographic analysis. *Cement and Concrete Research*, 135:106118, 2020.

[23] Wilson Ricardo Leal da Silva and Diogo Schwerz de Lucena. Concrete cracks detection based on deep learning image classification. In *Multidisciplinary digital publishing institute proceedings*, volume 2, page 489, 2018.

[24] Hanjie Qian, Ye Li, Jianfei Yang, Lihua Xie, and Kang Hai Tan. Image-based microstructure classification of mortar and paste using convolutional neural networks and transfer learning. *Cement and Concrete Composites*, page 104496, 2022.

[25] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[26] Stephan R Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun. Playing for data: Ground truth from computer games. In *European conference on computer vision*, pages 102–118. Springer, 2016.

[27] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[28] Ignacio Arganda-Carreras, Verena Kaynig, Curtis Rueden, Kevin W Eliceiri, Johannes Schindelin, Albert Cardona, and H Sebastian Seung. Trainable weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*, 33(15):2424–2426, 2017.

[29] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016.

[30] Alexander Kirillov, Yuxin Wu, Kaiming He, and Ross Girshick. Pointrend: Image segmentation as rendering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9799–9808, 2020.

[31] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[32] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.

[33] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.

[34] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.

[35] Cao Vu Dung et al. Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99:52–58, 2019.

[36] Shengyuan Li, Xuefeng Zhao, and Guangyi Zhou. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, 34(7):616–634, 2019.

[37] Zheng Tong, Dongdong Yuan, Jie Gao, and Zhenjun Wang. Pavement defect detection with fully convolutional network and an uncertainty framework. *Computer-Aided Civil and Infrastructure Engineering*, 35(8):832–849, 2020.

[38] Mohammad Hesam Hesamian, Wenjing Jia, Xiangjian He, and Paul Kennedy. Deep learning techniques for medical image segmentation: achievements and challenges. *Journal of digital imaging*, 32(4):582–596, 2019.

[39] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[40] John C Russ and Robert T Dehoff. *Practical stereology*. Springer Science & Business Media, 2012.

[41] Vyvyan Howard and Matthew Reed. *Unbiased stereology: three-dimensional measurement in microscopy*. Garland Science, 2004.

[42] Walid Deboucha, Nordine Leklou, Abdelhafid Khelidj, and Mohamed N Oudjit. Hydration development of mineral additives blended cement using thermogravimetric analysis (tga): Methodology of calculating the degree of hydration. *Construction and Building Materials*, 146:687–701, 2017.

[43] Javed I Bhatty. Hydration versus strength in a portland cement developed from domestic mineral wastes—a comparative study. *Thermochimica acta*, 106:93–103, 1986.

[44] https://cocodataset.org/#detection-eval.