

Automatic Brittle Fracture Ratio Estimation Using Convolutional Neural Network Regression Based on Classmap Regulation

SEUNG HYUN JEONG¹, MIN WOO WOO^{1,2}, GYOGWON KOO³,
JONG-HAK LEE⁴, AND JONG PIL YUN^{1,5}

¹Korea Institute of Industrial Technology, Cheonan 31056, Republic of Korea

²School of Computer Science and Engineering, Kyungpook National University, Daegu 41566, Republic of Korea

³Daegu Gyeongbuk Institute of Science and Technology, Daegu 42988, Republic of Korea

⁴Technical Research Laboratories POSCO, Pohang 37859, Republic of Korea

⁵KITECH School, University of Science and Technology, Daejeon 34113, Republic of Korea

Corresponding author: Jong Pil Yun (rebirth@kitech.re.kr)

This work was supported in part by POSCO under Grant IR-18-0059, and in part by the Korea Institute of Industrial Technology (KITECH) under Grant JA-21-0002.

ABSTRACT A convolutional neural network (CNN) based regression is proposed for estimating the brittle fracture ratio (BFR) in a fracture image of a drop weight tear test (DWTT) specimen. Different with the previous complex semantic segmentation-based estimator, the method extracts the feature vector through global average pooling of feature map and calculates the BFR directly through the fully connected layer. By removing decoder network, the number of weights, training time, and required GPU memory dramatically reduced. To train the proposed CNN, a new loss function, which is the sum of L1-norm between class activation map and ground truth inspection image and L1-norm of BFR error, is also designed. To validate the present method, fracture images of 1532, 79, and 158 DWTT specimens obtained from real industrial site were used for training, validation, and test, respectively. The accuracy of the proposed method was evaluated based on the number of test samples with an error of 5% or less divided by the total number of test samples, which is the measure used in real industrial application. Despite having dramatically reduced the number of weights and inference time by 85.8% and 64.8%, respectively, the proposed method has a higher accuracy (96.2%) compared to that of the existing segmentation based BFR estimation method (94.9%).

INDEX TERMS Brittle fracture rate estimator, convolutional neural network regression, drop-weight tear test, heatmap regulation.

I. INTRODUCTION

This study proposes an end-to-end convolutional neural network (CNN)-based regression method for the estimation of brittle fracture ratio (BFR) in the drop-weight tear test (DWTT). The demand of line-pipe has increased owing to the long-distance transportation of natural resources such as crude oil and natural gas in extremely cold areas (e.g., Siberia and Alaska). Steel with extremely low-temperature toughness is required to install line-pipes in these areas. Therefore, testing the properties of steel products from hot-rolling process has become more important. For the evaluation of steel properties, the DWTT, which is first

developed at the Battelle Memorial Institute, USA [1]–[6], has been widely used. The DWTT to determine the fracture characteristics of steel products is an integral part of the material qualification programs in oil and gas, and other industries. The schematic diagram of DWTT and the method of acquiring the image of the fracture surface are shown in Fig. 1. Through the DWTT, the test specimen is split up owing to impact load by a hammer and its resistance characteristic against brittle fracture propagation can be determined through the ratio of the ductile and brittle fracture surface areas. In various industries, these fracture surface areas are generally segmented by a professional operator, and the BFR or ductile fracture ratio (DFR) is estimated. However, owing to the manual evaluation, which depends on the operator's condition and state of fatigue, not only

The associate editor coordinating the review of this manuscript and approving it for publication was Francesco Piccialli.

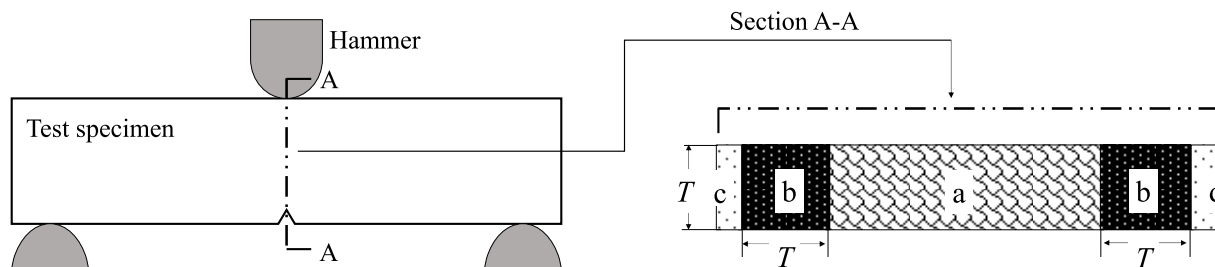


FIGURE 1. Schematic diagram of DWTT and fracture surface image acquisition region.

are reliability and reproducibility degraded but accuracy is also not guaranteed. Furthermore, because of the increasing demand of steel worldwide, there is a need to speed up material inspection. Therefore, development of an automated estimation system is essential to quantify the BFR or DFR efficiently and accurately from the DWTT.

To estimate BFR or DFR, many approaches have been proposed [7]–[9]. In [7], three-dimensional (3D) scanner was used to acquire the three-dimensional fracture surface and a statistical method was applied to estimate BFR and DFR for the DWTT specimen. In [8], along with an expensive 3D scanner, a single charge-coupled device (CCD) camera was used to acquire fracture surface images. The K-means clustering algorithm was then applied to estimate the multivariate characteristics of the fracture surface. In [9], three input images with different angles of illumination were obtained and then combined to form a single image. After combining images, brittle and ductile regions were binarized. Since the ductile and brittle surfaces were divided according to the specified threshold value, the performance of the algorithm is sensitively affected.

Meanwhile, deep learning has lately shown great performance in various industrial fields. In [10], a two-level hierarchical deep convolutional neural network was applied for automatic extraction of feature representation for sewer defects inspection. Also, in [11], a novel unsupervised multiscale feature-clustering-based fully convolutional autoencoder was proposed to efficiently and accurately inspect various types of texture defects based on a small number of defect-free texture samples. In [12], a multiple classifier fusion strategy incorporated Faster R-CNN was applied for small fruit detection. Especially, a deep learning-based algorithm was applied to segment the fracture surfaces of specimen. In [13], VGG based U-Net (VU-Net), which performs pixel-wise segmentation of fracture surface by using the CNN-based on the encoder-decoder structure was applied. After the original fracture surface image passed the trained network, the segmented binary map was obtained and BFR was calculated. To the best of the authors' knowledge, as the network used in the previous study is based on the encoder-decoder structure, the number of weights is large. Owing to this, it not only takes a large inference time but also takes a large training time.

To address the drawbacks of the aforementioned algorithms, this study proposes a simple but accurate method: a VGG based regression network (VR-net). Unlike the existing VU-Net, which estimates the BFR by counting the pixels classified in the segmentation map, VR-net estimates the ratio directly after passing through the encoder of CNN. The method for estimating a certain value directly by using a CNN has been used in various fields, and its applicability has been sufficiently verified. Moreover, in 2016, Li Kuo Tan *et al.* developed a convolution neural network model based on the MRI image of left ventricular (LV) endocardium at end-diastole (ED) and end-systole (ES). It predicted mitosis by automatically quantifying various clinical parameters, including ejection fractions, and feeding them to the neural network [14]. In addition, in 2016, Yao Xue *et al.* predicted the number of cells through nuclear detection of cells in an image through a convolution neural network based on a cell image taken under a microscope [15]. In 2018, Gerda Bortsova *et al.* developed a CNN-based learning model to quantitatively estimate the extent of emphysema from proportions of the diseased tissue [16]. In 2018, S. Aich and I. Stavness proposed a CNN-based object counting method that improves performance by using heatmap regulation [17]. In addition to the aforementioned studies, a deep regression network using input data that was structured data rather than in the image form was used. Representatively, in 2018, Gregory D. Merkel *et al.* undertook the topic of short-term load forecasting of natural gas through deep neural network regression based on structured data for 62 regions [18]. Using the existing techniques as a reference, the CNN used to predict a specific value was developed in an appropriate form to estimate BFR. Therefore, in this study, a new loss function based on classmap regulation was proposed to improve accuracy.

The main contributions of this study are summarized as follows.

- An automated estimation procedure for BFR was proposed. Compared with the previous methods, it not only reduces the manpower and time costs involved but also performs an accurate evaluation.
- The equipment costs could be reduced because, unlike the conventional methods, which used an expensive 3D scanner, images for the surface of the broken specimens were obtained from a CCD camera.

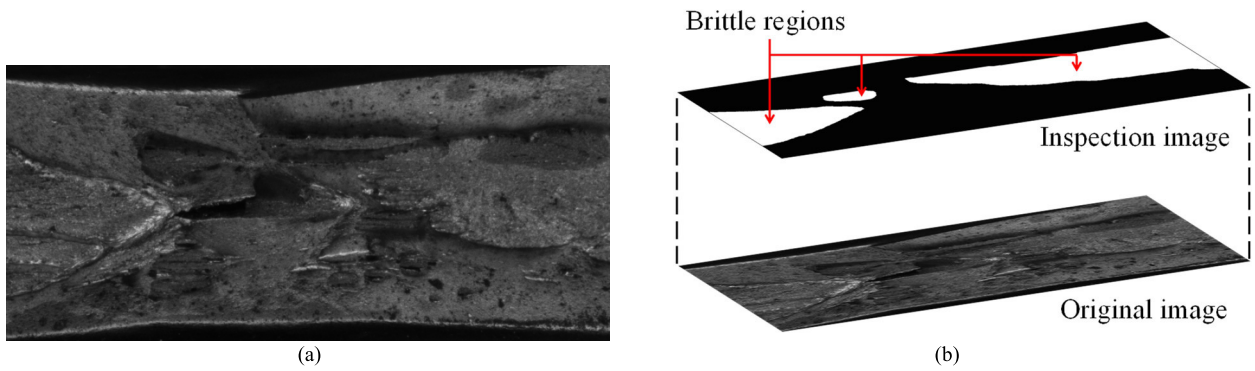


FIGURE 2. Example of (a) an image after the DWTT and (b) its divided regions to estimate BFR.

- To the best of the authors’ knowledge, this study is the first to exploit a deep learning based regression algorithm for BFR estimation.
- To train the VR-net, a new loss function based on classmap regulation was proposed. The effect of loss function and depth of backbone network were investigated.
- The performance of VR-net was verified by images obtained in real industrial site. Therefore, the applicability of VR-net was verified.

This paper consists of the four sections. In the second section, the proposed CNN structure for BFR estimation and new loss function using classmap regulation for training VR-net will be explained. The composition of dataset, detailed training process, and test results will be described in the third section. Finally, conclusions and findings will be presented in the fourth section.

II. CONVOLUTIONAL NEURAL NETWORK FOR BRITTLE FRACTURE RATIO ESTIMATOR

In this section, the proposed regression method based on CNN for brittle fracture ratio estimator is described. First, the definition of brittle fracture ratio used in steel manufacture industries will be explained. Next, the proposed network structure and formulation of the loss function for training will be described.

A. DEFINITION OF BRITTLE FRACTURE RATIO

As shown in Fig. 1, the surface of specimen after the DWTT can be divided into four categories. The area for the BFR or DFR estimation region, the excluded region, the notch area, and the impact point by the hammer are denoted by a, b, c, and d, respectively, in Fig. 1. T is the thickness of the specimen. After the BFR estimation region is defined, the inspection is performed by a professional operator. An example of broken surface with brittle feature is shown in Fig. 2(a) and its divided regions extracted by the operator are shown in Fig. 2(b). The pixel values of original input image are denoted by p_{hw} with $h = 1, \dots, H$ and $w = 1, \dots, W$, where H and W are height and width of input image,

respectively. After the inspection by the operator, the pixel values of inspected image, r_{hw} , can be generated with the value of 1 if it belongs to the brittle region, R_b , and 0 otherwise as shown in (1).

$$r_{hw} = \begin{cases} 1, & \text{if } p_{hw} \in \mathfrak{R}_b \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

After the inspection, the BFR value, R_b , can be formulated as follows.

$$R_b = \frac{\sum_h \sum_w r_{hw}}{H \cdot W} \quad (2)$$

In other words, the BFR is the value of the brittle fracture area divided by the total area of the image inspected.

B. CONVOLUTIONAL NEURAL NETWORK STRUCTURE

In this study, VR-net was proposed to estimate the BFR. Unlike the previous study using VU-net [13], the proposed network does not include a decoder that performs deconvolution to make an annotation map, which has the same size as the original input image, as shown in Fig. 3. The VR-net performs global average pooling for the feature map and calculates the BFR value directly through the fully connected layer. By the elimination of the decoder network, the VR-net has two advantages. First, the training time as well as the inference time decreases because the number of weights dramatically decreases by 86% (from 141,828,812 to 20,024,897). Second, when the annotation map generated by the operator at the actual industrial site contains some noise, a regression technique such as the proposed method may be more suitable for the purpose of predicting BFR than semantic segmentation, which requires accurate segmentation map.

A detailed explanation of the network structure is as follows. First, the encoder extracts high-dimensional features from the fracture surface image. The encoder has a role similar to the feature extractor of famous CNNs such as VGGNet [19], ResNet [20], and Densenet [21]. These CNNs provide pre-trained weight values using ImageNet’s dataset. In the training of a CNN, the initial weight value plays

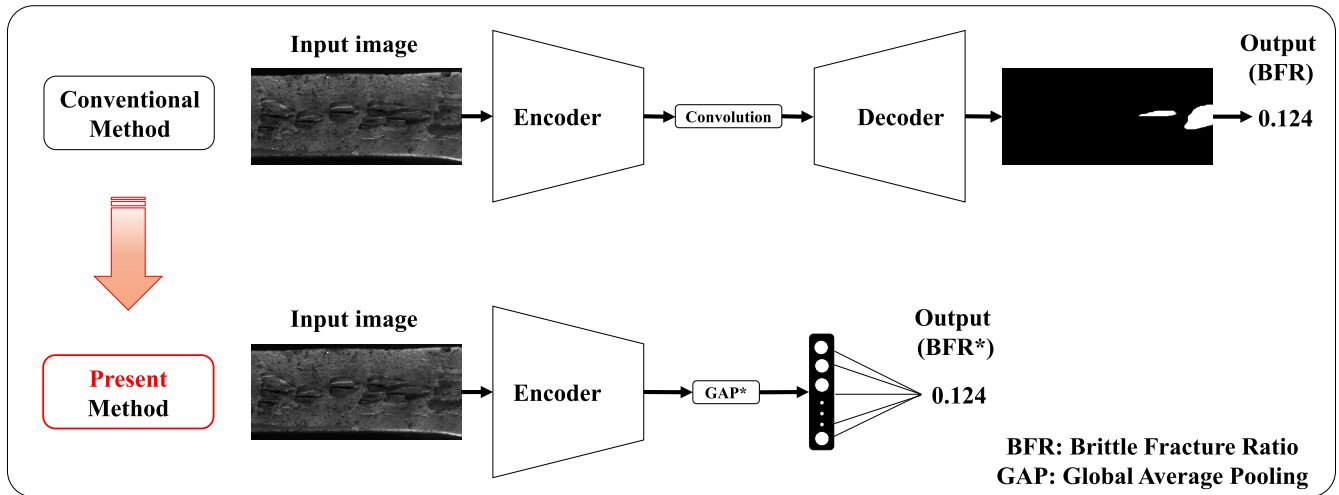


FIGURE 3. The proposed brittle fracture ratio estimation method.

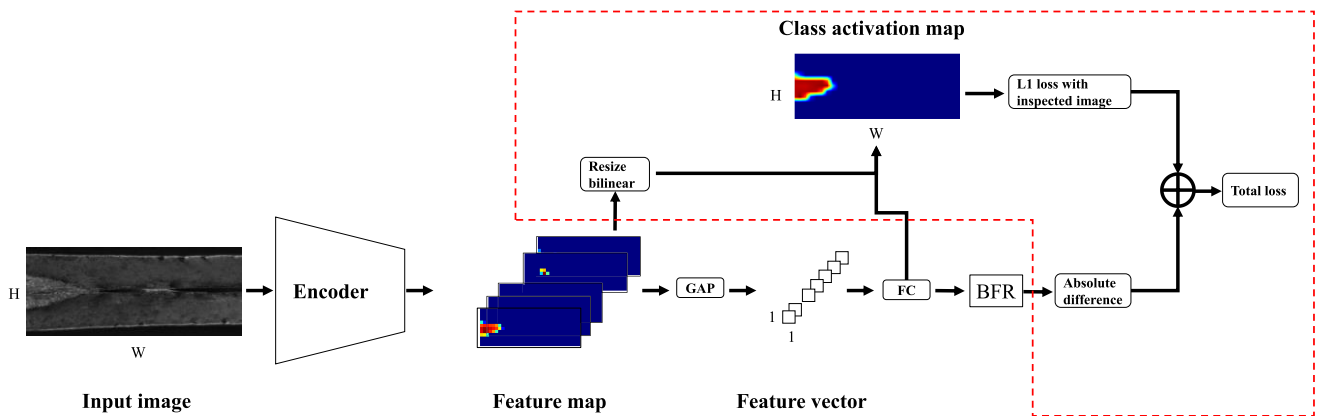


FIGURE 4. The description of loss function for training of the proposed convolutional neural network.

an important role in obtaining fast convergence and good accuracy. Therefore, training is often performed using the initial weight of a pre-trained neural network. In this study, the pre-trained weights of VGGNet, Densenet, and Resnet are used. Note that, different with the original VGGNet using max pooling, average pooling was applied. After the last convolution operation, global average pooling was performed using an extracted feature map.

C. DEFINITION OF LOSS FUNCTION

The new loss function, proposed for the training of the network, will be described in this subsection. The loss function is formulated in (3).

$$L = \left| \tilde{R}_b - R_b^* \right| + L_1(C_{map}, I_{map}) \quad (3)$$

where, \tilde{R}_b and R_b^* are the predicted and ground truth BFR values, respectively. C_{map} and I_{map} are the two-dimensional class activation map (CAM) and image inspected by an industry operator, respectively. Both C_{map} and I_{map} have the

same size as the input image. The second L1 loss term is termed classmap regulation.

Furthermore, the detailed explanation of the loss function is described in Fig. 4 for intuitive understanding. As shown in the figure, the red marked box is added for the calculation of loss function. For the calculation of loss function, an image inspected by an operator is needed. The loss function is the summation of two values. The first value is the absolute difference between the ground truth BFR value and predicted value calculated by the VR-net. The second value is the L1 loss between the CAM and inspected image. In the inspected image, the pixel value of the brittle surface region is one and the rest are zero. The CAM is a method developed to express the features that the CNN is paying attention to in the classification model [22]. In this study, the CAM was applied to define the loss function for training the regression model. In Fig. 4, a feature map of 512 channels, whose height and width are reduced 32 times, is output through the encoder. The output feature maps are resized to the same size as the input image and the weight value of each fully connected

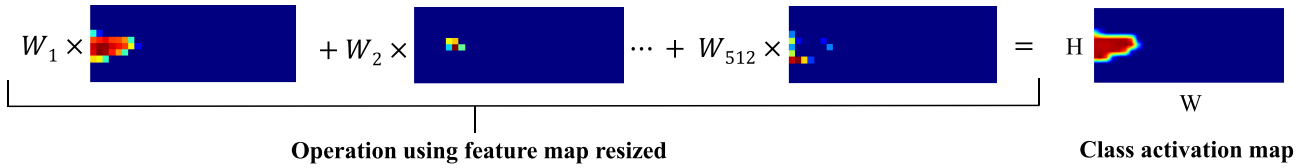


FIGURE 5. Definition of class activation map.

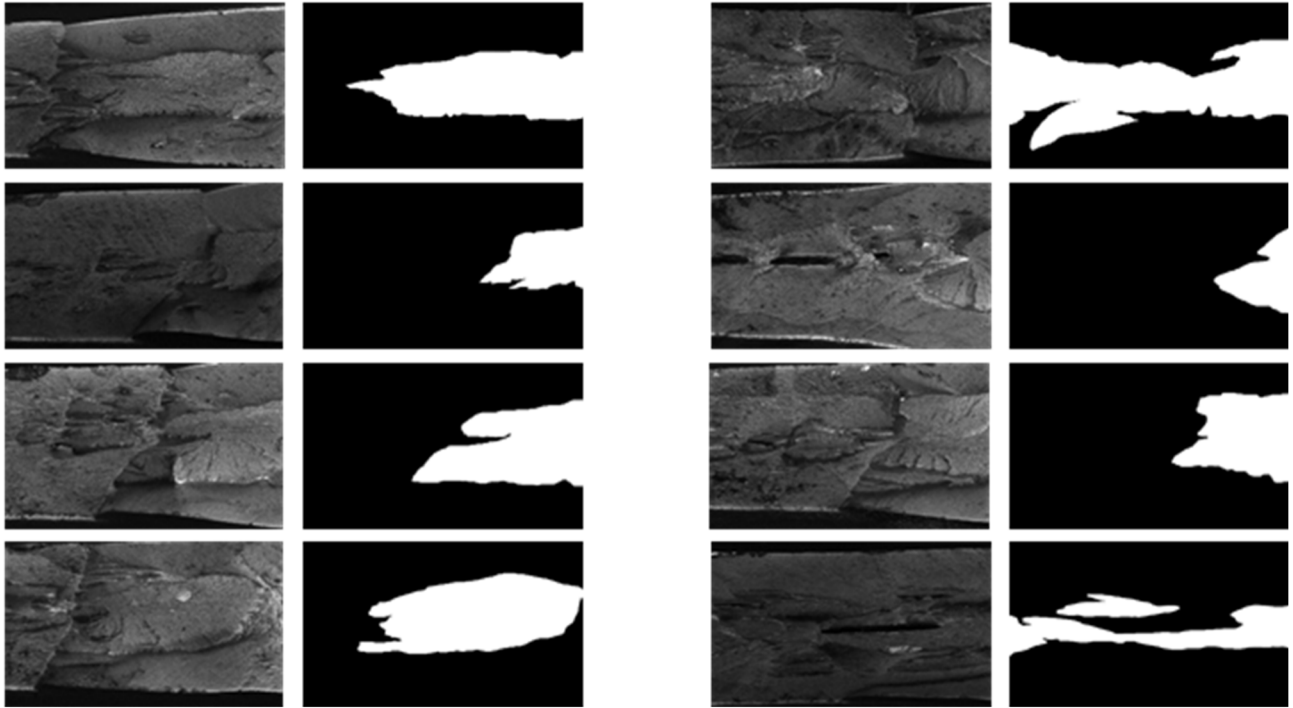


FIGURE 6. Examples of DWTT fracture surface image and corresponding inspection image obtained from real industry.

layer is multiplied. Thereafter, all of them are added to obtain a CAM of one channel, as shown in Fig. 5.

III. DATA AND EXPERIMENTS

To verify the performance of the VR-Net for the BFR estimator, it should be applied to a real dataset. Therefore, in this section, the dataset organization and augmentation will be described. Moreover, the experiments and process used to verify the proposed method and their results will be explained in detail.

A. DATASET ORGANIZATION AND AUGMENTATION

The dataset, consisting of the original DWTT fracture image and corresponding inspection image, was obtained from actual industrial sites. The examples of dataset are shown in Fig. 6. In the previous research, the total number of image and inspected pairs is 1,611, which were divided into 1,532 and 79 for training and validation, respectively. In addition, performance was verified by using 158 pairs of test data that was not used for training and validation. In this study, the same dataset organization was used for comparison with the previously used method. According

TABLE 1. Statistical size distribution of the fracture surface images.

	Mean	SD	Min	Max
Width	625.82	22.32	611	810
Height	462.65	93.21	262	676

to the specimen, the image size of the fracture specimens was different, and their statistical size distribution is shown in Table 1. To enhance accuracy and generalization ability, the training dataset is augmented eight times by using flip and 90 degrees rotation on the original images, and corresponding inspected images are generated. By using the data augmentation, the number of training dataset becomes 12,256.

B. EXPERIMENTS

As described in subsection 2.3, the network is trained by using the proposed loss function. During the training process, the validation accuracy is calculated at a specific number of iterations. The validation accuracy is the average absolute error (AAE) for the entire validation dataset. To reduce the error of BFR, the model with the minimum AAE rather than

TABLE 2. Test cases specifications and results.

Case no.	Use regulation	Backbone	No. of weights	Batch size	Image resize	MAE	AAE	Nerror $\geq 5\%$	Accuracy
1	no	VGG19	141,823,812	1	no	0.13807	0.0177	8	94.9
2	no	VGG19	20,024,897	1	no	0.12543	0.01982	12	92.4
3	yes	VGG19		1	no	0.13573	0.01561	6	96.2
4	yes	VGG11	9,220,993	1	no	0.12022	0.02018	14	91.1
5	yes	VGG13	9,405,505	1	no	0.11931	0.01873	12	92.4
6	yes	VGG16	14,715,201	1	no	0.13301	0.01743	9	94.3
7	yes	Densenet121	6,954,881	1	no	0.24949	0.02642	18	88.6
8	yes	Densenet121		4	500 by 500	0.13494	0.01714	6	96.2
9	yes	Densenet161	26,474,209	1	no	0.25155	0.02637	16	89.9
10	yes	Densenet161		4	500 by 500	0.11609	0.01778	8	94.9
11	yes	Resnet18	11,177,025	1	no	0.21347	0.02747	22	86.1
12	yes	Resnet18		4	500 by 500	0.1246	0.01877	9	94.3
13	yes	Resnet34	21,285,185	1	no	0.2642	0.03	21	86.7
14	yes	Resnet34		4	500 by 500	0.12382	0.01805	9	94.3

the minimum loss is chosen. The weights with minimum AAE are saved and applied for the test dataset.

For the comparison, four indices were used to evaluate the performance of the networks. The indices used to evaluate the performance of the prediction values of test set are the maximum absolute error (MAE), the AAE, and the accuracy used in real industrial applications, defined as the number of samples with an error of 5% or more divided by the total number of samples, $Nerror \geq 5\%$. Since an error of 5% or less is acceptable in the industrial site considering the uncertainty of the brittle fracture area, this definition of accuracy was adopted.

VR-net was implemented by using Tensorflow and trained by using Adam optimizer. The initial learning rate was set to 0.00001 and decayed 4% per 10,000 iterations. The images are 8-bit gray level and have different sizes according to specimens. The images were not resized, and the batch size was set to one. For the experiment, a workstation with Intel Xeon CPU E5-2690v4 2.60GHz, 192GB ram, and Nvidia GTX 1080Ti was used.

The proposed method tested from several perspectives. The effect of loss function, backbone network, and batch size are compared by changing test cases as shown in Table 2.

1) THE EFFECT OF PROPOSED REGULATION OF LOSS FUNCTION

The effects of the regulation term of loss function were tested. MAE, AAE, $Nerror \geq 5\%$, and accuracy for each case in Table 2 are listed in Table 2. As shown in Table 2, in terms of AAE and $Nerror \geq 5\%$, VR-net with proposed

regulation method (Case no. 3 in Table 2) has the best accuracy. As can be seen in case no.2 in Table 2, when the regulation is not used, the accuracy is less than 93%. It was confirmed that the regression without regulation could not be used in the actual industrial field because 12 samples showed an error of 5% or more. By adding the proposed regulation term to the loss function, more information is given for training the network and thereby high accuracy is obtained. In addition, three samples were randomly selected, and original fracture images, inspection images, and CAMs were drawn per sample in Fig. 7. As shown, owing to the proposed regulation term, the training is performed to mimic inspection image; thus, the shapes of the activated region of inspection image and CAM are similar. The figures in the last row represent CAM of three DWTT images. As shown, the CAM tends to follow the shape of the inspected image, but it is more ambiguous than the results obtained using regulation. In the test, the sample showing the maximum error is the same for the VU-net and VR-net with proposed regulation. Therefore, the original fracture image, inspection image, and CAM are plotted in Fig. 8 to figure out the reason of maximum error. As shown in the figure, the ambiguous fracture region in the center of the CAM has been activated, causing a large error. In summary, the prediction accuracy could be increased by using the proposed loss function, and a method capable of tracking the brittle fracture region was proposed.

2) THE EFFECT OF DEPTH OF BACKBONE NETWORK

To test the effects of depth of backbone network, as can be seen in Table 2, the depth of convolution layer is changed to

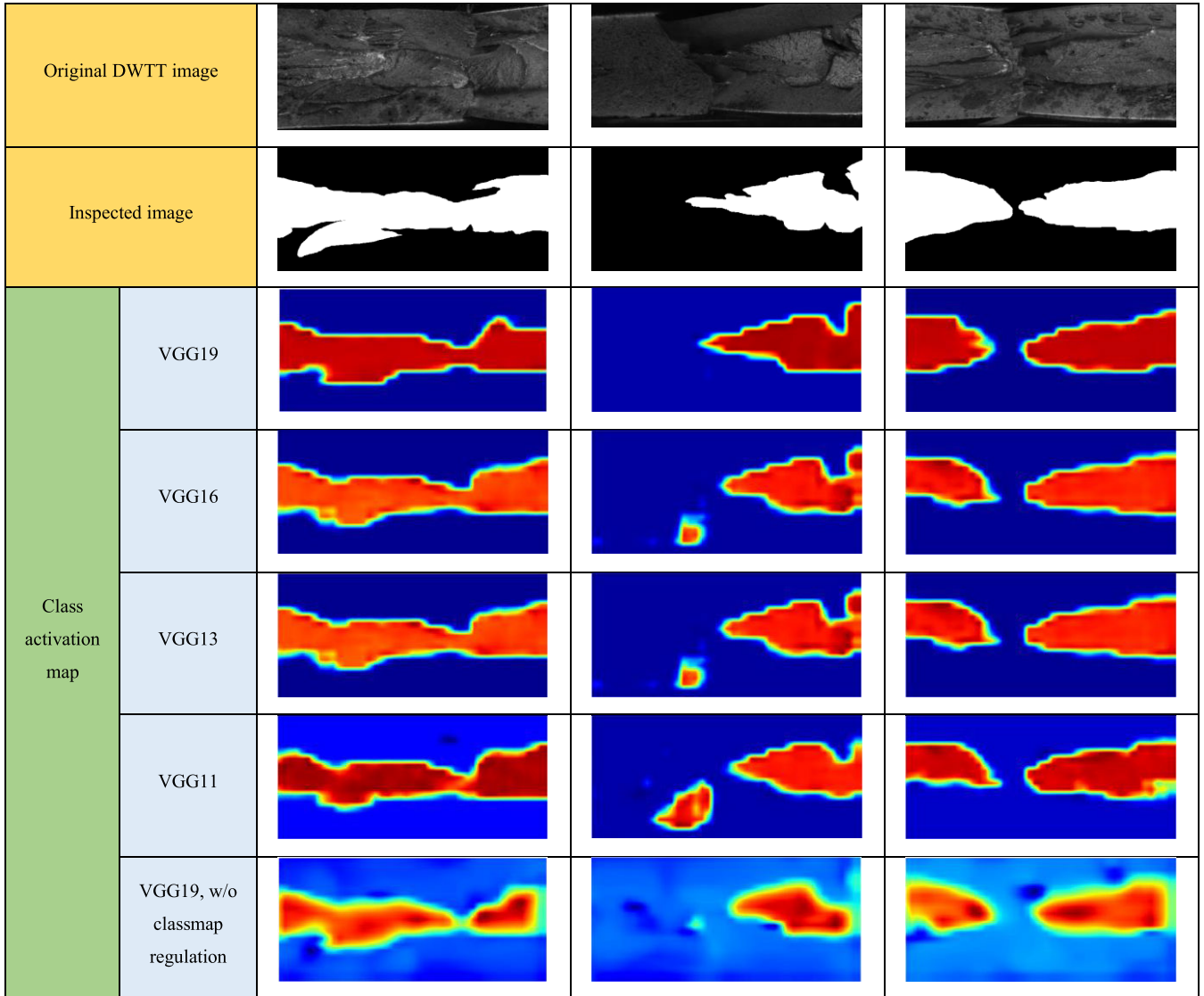


FIGURE 7. A fracture image, inspected image, and class activation map according to loss function and depth of backbone network.

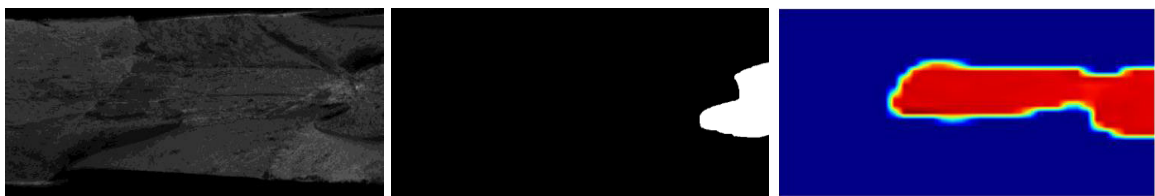


FIGURE 8. A fracture image, inspected image, and class activation map of worst accuracy test sample.

VGG11, 13, 16, and 19. In addition to VGGNet, experiments was performed by changing the backbone network of the proposed method into Resnet and Densenet. For all the models, weight values are initialized by using the pre-trained values using ImageNet’s dataset. The accuracy of the four VGG models is listed in Table 2 (Case no. 3, 4, 5, and 6).

As shown in the Table 2, accuracy is improved as the depth increased, and VGG19 shows the best results. Moreover, as shown in Fig. 7, as the depth of the backbone network becomes shallower, the CAM tends to fail to follow the shape of the inspected image in detail. As shown in Table 2, for the rest of the models other than the VGG models,

different batch size and depth were applied respectively. In the case of Densenet, when one factor of depth or batch size was fixed and the other factor was increased, the accuracy tended to increase. However, when the batch size was four, the accuracy decreased as the depth increased. In the case of Resnet, the batch size had a significant effect on the accuracy, but the depth did not. As a result, Densenet121 that batch size was set to four showed the same accuracy as the proposed VGGNet using average pooling layer. According to the average inference time, the proposed VGGNet was about 10 times faster than the Densenet121 although it has about 2.9 times more number of weights.

IV. CONCLUSION

This study presents a new deep neural network-based regression network, named VR-net. The network structure and loss function for estimating BFR value were newly developed. The proposed network used VGG19 with average pooling as the backbone network and classmap regulation, which improve performance of BFR estimation. The proposed CNN based regression is to be suitable for application in real industrial sites where it is difficult to create an accurate annotation map. The accuracy of the proposed method was evaluated by using the accuracy measure used in real industrial site, which is defined by the number of samples with absolute error of BFR estimation less than 5% divided by total number of test samples. Despite drastically reducing the number of weights, the accuracy of the proposed method is 96.20%, which is higher than that of the previous VU-net based method (94.95%). Moreover, to improve accuracy, the average inference time for the test samples was drastically reduced from 0.165 s to 0.058 s owing to the simple network structure and reduced number of weights. In this paper, in addition to the performance evaluation, the effect of loss function and depth of backbone network were also investigated. Through several tests, it was found that the proposed classmap regulation loss term improves accuracy as the CAM becomes similar to the inspection image. Moreover, the accuracy was improved with the increase of backbone network depth. Based on the proposed VR-Net, accuracy and effectiveness of an automated estimation system for BFR was improved, which not only reduces money, human and time costs but also makes a consistent decision regarding the quality of steel.

REFERENCES

- [1] B. Hwang, Y. M. Kim, S. Lee, N. J. Kim, and J. Y. Yoo, "Correlation of rolling condition, microstructure, and low-temperature toughness of x70 pipeline steels," *Metall. Mater. Trans. A*, vol. 36, no. 7, pp. 1793–1805, 2005, doi: [10.1007/s11661-005-0043-1](https://doi.org/10.1007/s11661-005-0043-1).
- [2] Y.-P. Zeng, P.-Y. Zhu, and K. Tong, "Effect of microstructure on the low temperature toughness of high strength pipeline steels," *Int. J. Minerals, Metall., Mater.*, vol. 22, no. 3, pp. 254–261, 2015, doi: [10.1007/s12613-015-1069-z](https://doi.org/10.1007/s12613-015-1069-z).
- [3] B. Hwang, S. Lee, Y. M. Kim, N. J. Kim, J. Y. Yoo, and C. S. Woo, "Analysis of abnormal fracture occurring during drop-weight tear test of high-toughness line-pipe steel," *Mater. Sci. Eng., A*, vol. 368, nos. 1–2, pp. 18–27, 2004, doi: [10.1016/j.msea.2003.09.075](https://doi.org/10.1016/j.msea.2003.09.075).
- [4] Z. Yang, C.-B. Kim, Y. Feng, and C. Cho, "Abnormal fracture appearance in drop-weight tear test specimens of pipeline steel," *Mater. Sci. Eng., A*, vols. 483–484, pp. 239–241, Jun. 2008, doi: [10.1016/j.msea.2006.09.182](https://doi.org/10.1016/j.msea.2006.09.182).
- [5] S. Hong, S. Y. Shin, S. Lee, and N. J. Kim, "Effects of specimen thickness and notch shape on fracture modes in the drop weight tear test of api x70 and x80 linepipe steels," *Metall. Mater. Trans. A*, vol. 42, no. 9, pp. 2619–2632, 2011, doi: [10.1007/s11661-011-0697-9](https://doi.org/10.1007/s11661-011-0697-9).
- [6] D. L. Rudland, Y.-Y. Wang, G. Wilkowski, and D. J. Horsley, "Characterizing dynamic fracture toughness of linepipe steels using the pressed-notch drop-weight-tear test specimen," *Eng. Fract. Mech.*, vol. 71, nos. 16–17, pp. 2533–2549, 2004, doi: [10.1016/j.engfracmech.2003.12.007](https://doi.org/10.1016/j.engfracmech.2003.12.007).
- [7] B. Strnadel, P. Perfecki, and P. Židlík, "Statistical characteristics of fracture surfaces in high-strength steel drop weight tear test specimens," *Eng. Fract. Mech.*, vols. 112–113, pp. 1–13, Nov. 2013, doi: [10.1016/j.engfracmech.2013.10.001](https://doi.org/10.1016/j.engfracmech.2013.10.001).
- [8] P. Skalný, "Evaluation and identifying the ductile fracture area of x70 steel from DWTT broken specimens," *Proc. Struct. Integrity*, vol. 2, pp. 3727–3734, Jul. 2016, doi: [10.1016/j.prostr.2016.06.463](https://doi.org/10.1016/j.prostr.2016.06.463).
- [9] T. S. Kim and S. G. Park, "A study on decision algorithm of ductile-brittle rates of DWTT specimens of API linepipe steels," in *Proc. Int. Conf. Future Inf. Commun. Eng.*, Kota Kinabalu, Malaysia, 2015, pp. 505–508, doi: [1410-ECN-0101-2018-004-000971830](https://doi.org/10.1410-ECN-0101-2018-004-000971830).
- [10] Q. Xie, D. Li, J. Xu, Z. Yu, and J. Wang, "Automatic detection and classification of sewer defects via hierarchical deep learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 4, pp. 1836–1847, Oct. 2019, doi: [10.1109/TASE.2019.2900170](https://doi.org/10.1109/TASE.2019.2900170).
- [11] H. Yang, Y. Chen, K. Song, and Z. Yin, "Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 3, pp. 1450–1467, Jul. 2019, doi: [10.1109/TASE.2018.2886031](https://doi.org/10.1109/TASE.2018.2886031).
- [12] X. H. X. M. Mai Zhang Jia and Q. H., "Meng Faster R-CNN with classifier fusion for automatic detection of small fruits," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1555–1569, Jul. 2020, doi: [10.1109/TASE.2020.2964289](https://doi.org/10.1109/TASE.2020.2964289).
- [13] G. Koo, C. Shin, H. Choi, J.-H. Lee, S. W. Kim, and J. P. Yun, "Automated brittle fracture rate estimator for steel property evaluation using deep learning after drop-weight tear test," *IEEE Access*, vol. 7, pp. 145095–145103, 2019, doi: [10.1109/ACCESS.2019.2945563](https://doi.org/10.1109/ACCESS.2019.2945563).
- [14] L. K. Tan, Y. M. Liew, E. Lim, and R. A. McLaughlin, "Cardiac left ventricle segmentation using convolutional neural network regression," in *Proc. IEEE EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2016, pp. 490–493.
- [15] Y. Xue, N. Ray, J. Hugh, and G. Bigras, "Cell counting by regression using convolutional neural network," in *Proc. ECCV*, Amsterdam, The Netherlands, Oct. 2016, pp. 274–290, doi: [10.1007/978-3-319-46604-0_20](https://doi.org/10.1007/978-3-319-46604-0_20).
- [16] G. Bortsova, F. Dubost, S. Ørting, I. Katramados, L. Hogeweg, L. Thomsen, M. Wille, and M. Bruijne, "Deep learning from label proportions for emphysema quantification," in *Proc. MICCAI*, Granada, Spain, 2018, pp. 768–776, doi: [10.1007/978-3-030-00934-2_85](https://doi.org/10.1007/978-3-030-00934-2_85).
- [17] S. Aich and I. Stavness, "Improving object counting with heatmap regulation," 2018, *arXiv:1803.05494*. [Online]. Available: <http://arxiv.org/abs/1803.05494>
- [18] G. D. Merkel, R. J. Povinelli, and R. H. Brown, "Short-term load forecasting of natural gas with deep neural network regression," *Energies*, vol. 11, no. 8, pp. 1–12, Aug. 2018, doi: [10.3390/en11082008](https://doi.org/10.3390/en11082008).
- [19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [21] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708, doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [22] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2921–2929, doi: [10.1109/CVPR.2016.319](https://doi.org/10.1109/CVPR.2016.319).



deep learning, and innovative application of optimization and deep learning.

SEUNG HYUN JEONG was born in Daegu, Republic of Korea, in 1985. He received the B.S. and Ph.D. degrees in mechanical engineering from Hanyang University, Seoul, in 2008 and 2014, respectively. Since 2016, he has been a Senior Researcher with the Safety System Research Group, Korea Institute of Industrial Technology. His research interests include optimization algorithms, structural optimization, classification, object detection, semantic segmentation through



MIN WOO WOO received the B.S. and M.S. degrees in computer engineering from Keimyung University, Daegu, South Korea, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in computer science and engineering with Kyungpook National University, Daegu. His research interests include object detection, computer vision, and deep learning.



computer vision and deep learning.

GYOGWON KOO received the B.S. degree from the Department of Electrical Engineering, Kyungpook National University, Daegu, South Korea, in 2013, and the M.S. and Ph.D. degrees from the Department of Electrical Engineering, Pohang University of Science and Technology, Pohang, South Korea, in 2015 and 2020, respectively. Since 2020, he has been a Senior Researcher with Daegu Gyeongbuk Institute of Science and Technology. His current research interests include



JONG-HAK LEE received the Ph.D. degree in electrical engineering from Kyungpook National University, in 2005. He is currently a Senior Principal Researcher with the POSCO Laboratory. His research interests include system engineering, defect classification using machine vision, image processing, and deep neural networks.



JONG PIL YUN received the Ph.D. degree in electrical engineering from the Pohang University of Science and Technology, in 2009. He is currently a Principal Researcher with the Korea Institute of Industrial Technology. His research interests include defect detection, classification using machine vision, image processing, and deep neural networks.

...