IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 16, 2019, accepted October 29, 2019, date of publication November 1, 2019, date of current version November 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2950972

Estimating Regions of Deterioration in Electron Microscope Images of Rubber Materials via a Transfer Learning-Based Anomaly Detection Model

REN TOGO[1], (Member, IEEE), NAOKI SAITO[2], (Member, IEEE), TAKAHIRO OGAWA[1], (Senior Member, IEEE), AND MIKI HASEYAMA[1], (Senior Member, IEEE)

[1]Faculty of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan
[2]Department of Creative Engineering, National Institute of Technology, Kushiro College, Kushiro 084-0916, Japan

Corresponding author: Ren Togo (togo@lmd.ist.hokudai.ac.jp)

This work was supported in part by the MIC/SCOPE under Grant #181601001, and in part by the JSPS KAKENHI under Grant JP17H01744 and Grant JP19J10821.

ABSTRACT A method for estimating regions of deterioration in electron microscope images of rubber materials is presented in this paper. Deterioration of rubber materials is caused by molecular cleavage, external force, and heat. An understanding of these characteristics is essential in the field of material science for the development of durable rubber materials. Rubber material deterioration can be observed by using on electron microscope but it requires much effort and specialized knowledge to find regions of deterioration. In this paper, we propose an automated deterioration region estimation method based on deep learning and anomaly detection techniques to support such material development. Our anomaly detection model, called Transfer Learning-based Deep Autoencoding Gaussian Mixture Model (TL-DAGMM), uses only normal regions for training since obtaining training data for regions of deterioration is difficult. TL-DAGMM makes use of extracted high representation features from a pre-trained deep learning model and can automatically learn the characteristics of normal rubber material regions. Regions of deterioration are estimated at the pixel level by calculated anomaly scores. Experiments on real rubber material electron microscope images demonstrated the effectiveness of our model.

INDEX TERMS Materials informatics, anomaly detection, deep learning, transfer learning.

I. INTRODUCTION

Materials informatics (MI) is a concept for the development of innovative materials by making use of stored data and for acquiring a new body of knowledge based on a computational approach [1], [2]. MI has been used for evaluation and accumulation of material information based on computational science. However, in recent years, attention has been given to large-scale data mining and efficient searches for new materials [3]–[5]. It is expected that MI will enable significant reductions in the time and cost needed for material development by using the combination of machine learning techniques and stored data [6]. The United States launched the Material Genome Initiative [1] as a national project in 2011, Novel Materials Discovery [2] in Europe started in 2015 as joint research of the EU as a part of the leading European Center of Excellence.[3] Also, Asian countries including China, South Korea and Japan have started large-scale national material projects in recent years [7]. MI is thus gaining worldwide attention and is being recognized as a new research field for industrial applications in material science [8].

In the past decade in which MI has been focused on images, visual features extracted from target images have been utilized by material scientists to characterize microstructures of materials [9]. Generally, basic image recognition techniques have been used for analyzing images of the structure of a material [10], [11]. Hand-crafted features such as scale-invariant feature transform (SIFT) [12], histogram

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

[1]https://www.mgi.gov

[2]https://nomad-coe.eu

[3]https://ec.europa.eu/eurostat/cros/content/centres-excellence_en

of oriented gradients (HoG) [13] and local binary patterns (LBP) [14] are the mainstream of visual features for analyzing photographs of target materials [6], [15]. However, the usefulness of these features is still limited because hand-crafted features essentially have low representation ability, which may be insufficient to understand the characteristics of images. Hence, hand-crafted features have gradually fallen into disuse in the field of general image processing [16].

Deep learning, one of the machine learning techniques, that can obtain high representation features automatically, has recently become a new trend in the field of computer vision [17]. Deep learning-based approaches have overcome the conventional feature representation problems caused by hand-crafted features [18], [19]. Deep convolutional neural networks (DCNNs) have achieved promising results in various image recognition tasks owing to their capacity to learn discriminative features from row pixels [20]. DCNNs can learn high representation features from target image data sources, and extracted features from pre-trained DNNs have been used by many researchers [21], [22]. The recent development of deep learning techniques is one of the reasons for the current artificial intelligence (AI) boom, and these techniques have been widely studied for industrial applications. In the field of material science, it is expected that applications of AI technologies will accelerate to the development of new materials [23].

Rubber materials which have many applications including automobiles and sports equipment, have unique elastic characteristics [24], [25]. Although rubber materials are useful for industrial applications, quality control of rubber materials plays important roles. In rubber material development, the maintenance of a large deformation capacity and recovery capability for a long period of time is important [26]. Deterioration of rubber is mainly due to deformation caused by molecular cleavage, external force, and heat. In MI of rubber materials, it is required to identify the cause of deterioration by utilizing information on the degradation of stored rubber materials, and it is expected that new knowledge for the development of highly durable rubber materials will be obtained [27].

Deterioration of a rubber material can be observed by electron microscope images of internal material structures [28]. Various factors including external force, oxidation, and heat affect the condition of a rubber material in certain circumstances, and the changes of its characteristics can be imaged. However, it requires much effort and specialized knowledge to find regions of deterioration in electron microscope images. By estimating regions of deterioration in electron microscope images through AI technologies, it is expected that new knowledge of rubber material deterioration will be acquired.

For estimating regions of deterioration in electron microscope images of a rubber material, attention should be given to the following points. First, there is a subtle difference between regions of deterioration and regions with no deterioration (normal regions) in an electron microscope image, and the task of recognizing the difference is much more difficult than general image recognition tasks. Although many general image recognition tasks have focused on the classification and detection of objects that are easily found, regions of deterioration in electron microscope images are difficult for non-professional persons to find. Secondly, since the type and degree of deterioration changes are diverse, collection of a large amount of data on deterioration is a challenging task. When estimating regions of deterioration based on machine learning techniques, collecting well-annotated training data for regions of deterioration is difficult, and the development of a method that considers the above-mentioned points is required.

In this paper, we propose a Transfer Learning-based Deep Autoencoding Gaussian Mixture Model (TL-DAGMM) inspired by the recently proposed anomaly detection model DAGMM [29]. Among machine learning techniques, anomaly detection is often used when data of specific labels are highly imbalanced or when a specific class of data cannot be obtained. In the task of anomaly detection, dimension reduction and density estimation play important roles. Although traditional anomaly detection approaches have focused on the two components separately, the recently proposed DAGMM can consider them and learn the parameters simultaneously. DAGMM has achieved high detection performance compared to the performances of conventional anomaly detection methods in anomaly detection tasks using low resolution images and acoustic signals. However, application of the original DAGMM to high resolution images is difficult due to the bloat of input feature vectors. To address this problem, a transfer learning approach is newly introduced in our TL-DAGMM. Transfer learning is a powerful tool in which the knowledge from a specific domain is transferred to a new domain with different properties of the feature space. We divide images into multiple patches and extract high representation features through a pre-trained deep learning model as transfer learning for each patch. Extracted features are inputted to an autoencoder model. By this architecture, training the deep autoencoder for high resolution images becomes feasible. Finally, we visualize the regions of deterioration at the pixel level based on the anomaly scores on TL-DAGMM. The effectiveness of our method was confirmed by several statistical analyses.

Our contributions are summarized as follows:

- We present a new method for estimating regions of deterioration based on anomaly detection and realize estimation of regions of deterioration in electron microscope images of a rubber material.
- A transfer learning approach is newly introduced in our TL-DAGMM for realizing training with high resolution images.

II. ELECTRON MICROSCOPE IMAGES OF RUBBER MATERIALS

A rubber material is a composite material that consists of various complicated substances. Typically, a rubber material

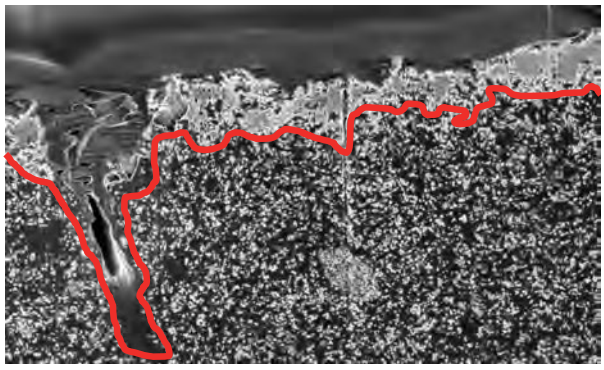
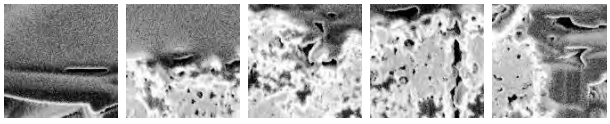


FIGURE 1. An example of electron microscope image of a rubber material used in this study.



(a) Abnormal deteriorated patches

(b) Normal patches

FIGURE 2. Examples of patches used in this study.

is mainly composed of carbon black and a rubber compound, and they can be observed by using an electron microscope. Figure 1 shows an electron microscope image that includes normal regions and regions of deterioration. In the figure, the region above the red line is a region of deterioration and the region under the red line is a normal region. We can see that a region of deterioration and a normal region in an electric microscope image only have subtly different characteristics. A great deal of labor is required to annotate regions of deterioration, and an automated supporting system that can help material scientists is desired.

An electron microscope image has a high resolution, and we treat it as multiple patch images. Figure 2 (a) shows examples of patches from a region of deterioration in an original electron microscope image, and Fig. 2 (b) shows examples of patches from a normal region in an original electron microscope image. It can be seen that there is a subtle difference between the patches from a normal region and patches from a region of deterioration. The patch-based approach is effective when target images have high resolutions and characteristics of the images are described in local regions [30], [31].

There are two main problems in analysis of electron microscope images of rubber materials. The first problem is that the boundaries between regions of deterioration and normal regions are too unclear to recognize. Deterioration type of a rubber image is polymorphic, and it is difficult to prepare annotated images including regions of deterioration. The other problem is the difficulty of collecting a large number of training images. Basically, rubber materials are not deteriorated at all when they are put on the market, and deterioration of rubber materials is caused by repeated use of the materials. Therefore, it requires a great deal of labor to collect images of deterioration in the stage of development. Even if we can annotate deteriorated regions manually, it is still difficult to collect a large number of images of deterioration compared to normal images from newly developed rubber materials. A lack of training images can cause performance degradation in supervised learning techniques since constructed models cannot correctly show the characteristics of the images. In this way, analysis of electron microscope images of rubber materials is still a challenging task, and a model for analysis that can consider the above real-world situations is desired.

We propose TL-DAGMM as a model that can solve the above-mentioned problems. Since preparing data for regions of deterioration is difficult, we make use of a pre-trained model that can be trained on a large-scale dataset to extract high representation features. TL-DAGMM realizes estimation of regions of deterioration by an anomaly detection approach. Details of our method are shown in Sec. III.

III. METHOD FOR ESTIMATING REGIONS OF DETERIORATION

An anomaly detection model TL-DAGMM uses normal patches of electron microscope images of a rubber material and estimates regions of deterioration in an unsupervised fashion. In preparation for explaining our method, we divide electron microscope images of a rubber material into patches and let $\boldsymbol{o}_i$ $(i = 1, 2, \ldots, O)$ denote patches from normal regions of the training images, where O is the number of training patch samples. Next, let $\boldsymbol{o}_{\text{test}}$ denote patches whose labels of normal/deteriorated are unknown. The proposed method consists of the training phase and the test phase. Each phase is explained in the following subsections.

A. TRAINING PHASE

To recognize the characteristics of electron microscope images of a rubber material, a constructed model needs to learn high representation features. Considering the lack of data for images including regions of deterioration, an anomaly detection approach that can learn a model with only normal data is suitable for this situation. Hence, we construct a deep learning-based anomaly detection method, TL-DAGMM. Figure 3 shows an overview of our TL-DAGMM anomaly detection network architecture. TL-DAGMM consists of three major components: a feature extraction network, a compression network and an estimation network. The feature extraction network extracts high representation features from an input patch sample. The compression network performs a dimensionality reduction task and low-dimensional representations are used for training of the subsequent estimation network. The estimation network performs a density estimation task.

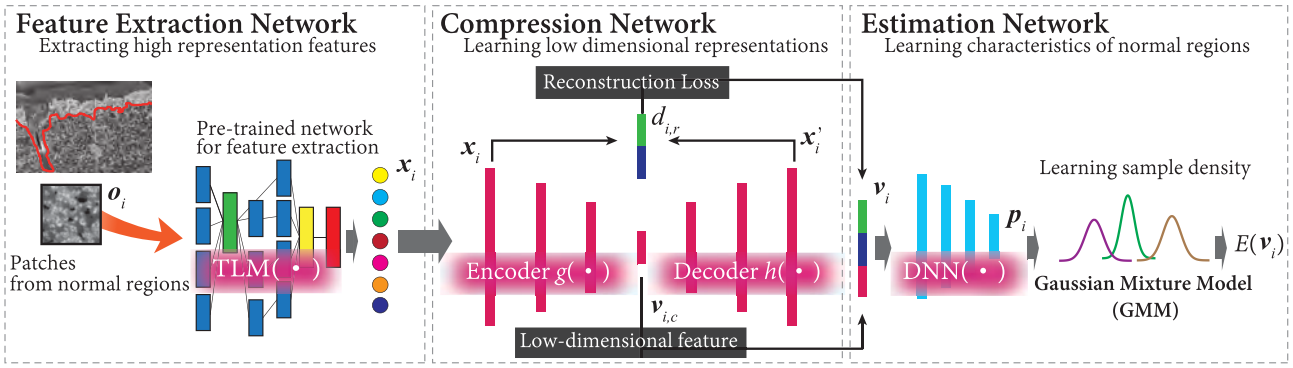


FIGURE 3. Training scheme of TL-DAGMM.

First, we explain the feature extraction network. Output vectors of an intermediate layer of a deep neural network (DNN) trained by a large-scale dataset are often used as image features that have semantic information in transfer learning. We use this network as the pre-processing of feature alignment toward the input sample as follows:

$$\boldsymbol{x}_i = \mathrm{TLM}(\boldsymbol{o}_i; \theta_t), \tag{1}$$

where TLM denotes the pre-trained DNN model with the parameters of θ_t. In the proposed method, we employ the pre-trained Inception-v3 network [32] and extract 2,048-dimensional features from the pool_3 layer. The Inception-v3 network is one of the state-of-the-art neural network models that achieves high recognition performance in object recognition tasks. This network is trained for ImageNet Large Visual Recognition Challenge [20] and achieves a 3.46% top-5 error rate in a 1,000 class classification task. Since features extracted from the Inception-v3 network have high semantic representations, this model is often used as a standard pre-trained model in transfer learning.

Next, we explain the architecture of the compression network. The role of the compression network is to obtain two types of low-dimensional representations: (i) reduced low-dimensional representations learned by a deep convolutional autoencoder and (ii) features derived from the reconstruction error. When given a feature $\boldsymbol{x}_i$, the compression network provides its reduced low-dimensional representation $\boldsymbol{v}_{i,c}$ as follows:

$$\boldsymbol{v}_{i,c} = h(\boldsymbol{x}_i; \theta_e), \tag{2}$$

where $h(\cdot)$ denotes the encoding function with the parameters of θ_e. The reconstruction feature $\boldsymbol{x}_i{}'$ is calculated as follows:

$$\boldsymbol{x}_i{}' = g(\boldsymbol{v}_{i,c}; \theta_d), \tag{3}$$

where $g(\cdot)$ denotes the decoding function with the parameters of θ_d. The decoder function $g(\cdot)$ receives the reduced feature $\boldsymbol{v}_{i,c}$, and transforms it to the reconstruction feature $\boldsymbol{x}_i{}'$. The functions $h(\cdot)$ and $g(\cdot)$ are elements of an autoencoder architecture. Then the relative Euclidean distance of $\boldsymbol{x}_i$ and $\boldsymbol{x}_i{}'$ can be calculated as follows:

$$d_{i,r} = \frac{||\boldsymbol{x}_i - \boldsymbol{x}_i{}'||_2}{||\boldsymbol{x}_i||_2}. \tag{4}$$

In the end, the low-dimensional representation feature derived from the compression network is:

$$\boldsymbol{v}_i = [\boldsymbol{v}_{i,c}{}^\top, d_{i,r}]^\top, \tag{5}$$

where $\top$ denotes the transposition. It can be considered that the extracted feature $\boldsymbol{v}_i$ contains valuable information of $\boldsymbol{x}_i$. The extracted feature $\boldsymbol{v}_i$ is an input vector for training of the subsequent estimation network.

The estimation network performs density estimation under the framework of the Gaussian mixture model (GMM). The GMM is a probabilistic model that assumes all of the data are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. In general, the expectation-maximization (EM) algorithm [33] is used for exploring optimal parameters of the GMM. The estimation network achieves this optimization by a deep neural network (DNN) that can predict the mixture membership of each sample. The GMM has unknown mixture-component distribution ϕ, mixture means μ, and mixture co-variance Σ. Given the low-dimensional representation features $\boldsymbol{v}_i$ and an integer K as the number of mixture components, the estimation network predicts its membership as follows:

$$\boldsymbol{p}_i = \mathrm{DNN}(\boldsymbol{v}_i, \theta_m), \tag{6}$$

$$\hat{\gamma}_i = \mathrm{softmax}(\boldsymbol{p}_i), \tag{7}$$

where DNN is a deep neural network parametarized by θ_m, $\boldsymbol{p}_i$ is the output value of the DNN model, and $\hat{\gamma}_i$ is a K-dimensional vector from the sample $\boldsymbol{x}_i$ for the soft mixture-component membership prediction. Furthermore, softmax$(\cdot)$ denotes a softmax function. Given a batch of N sample patches and their membership prediction $\forall 1 \leq k \leq K$, parameters of the GMM can be estimated as follows:

$$\hat{\phi}_k = \sum_{i=1}^{N} \frac{\hat{\gamma}_{ik}}{N}, \tag{8}$$

$$\hat{\boldsymbol{\mu}}_k = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} \boldsymbol{v}_i}{\sum_{i=1}^{N} \hat{\gamma}_{ik}}, \tag{9}$$

$$\hat{\boldsymbol{\Sigma}}_k = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} (\boldsymbol{v}_i - \hat{\boldsymbol{\mu}}_k)(\boldsymbol{v}_i - \hat{\boldsymbol{\mu}}_k)^\top}{\sum_{i=1}^{N} \hat{\gamma}_{ik}}, \tag{10}$$

where $\hat{\phi}_k$, $\hat{\boldsymbol{\mu}}_k$ and $\hat{\boldsymbol{\Sigma}}_k$ represent mixture probability, mean and co-variance for the component k of GMM respectively, and $\hat{\gamma}_i$ is the membership prediction for the low-dimensional representation $\boldsymbol{v}_i$. From the constructed model, a sample energy can be inferred as follows:

$$E(\boldsymbol{v}_i) = -\log\left(\sum_{k=1}^{K} \hat{\phi}_k \frac{\exp(-\frac{1}{2}(\boldsymbol{v}_i - \hat{\boldsymbol{\mu}}_k)^\top \hat{\Sigma}_k^{-1}(\boldsymbol{v}_i - \hat{\boldsymbol{\mu}}_k))}{\sqrt{2\pi |\hat{\boldsymbol{\Sigma}}_k|}}\right), \tag{11}$$

where $|\cdot|$ denotes the determinant of a matrix. The sample energy is calculated with the learned GMM parameters, and if a predicted sample energy is higher than a pre-defined threshold, this sample is estimated as an abnormal deteriorated sample in the test phase.

Overall, an objective function in training of TL-DAGMM when given a dataset of normal patches is constructed as follows:

$$J(\theta_e, \theta_d, \theta_m) = \frac{1}{N} \sum_{i=1}^{N} L(\boldsymbol{x}_i, \boldsymbol{x}_i') + \frac{\lambda_1}{N} \sum_{i=1}^{N} E(\boldsymbol{v}_i) + \lambda_2 P(\hat{\Sigma}). \tag{12}$$

The objective function includes three components: the reconstruction error L, the sample energy E and the normalization term of the GMM P. Also, λ_1 and λ_2 are weighted coefficients. Specifically, the first term $L(\boldsymbol{x}_i, \boldsymbol{x}_i')$ is the reconstruction error from the compression network. Intuitively, if the compression network could make the construction error low, the low-dimensional representation features could better preserve the key information of input samples. Hence, it is desirable for the reconstruction error to always be low for the compression network. We employ L_2-norm as $L(\boldsymbol{x}_i, \boldsymbol{x}_i') = ||\boldsymbol{x}_i - \boldsymbol{x}_i'||_2^2$ in this study. Next, minimization of the sample energy is also always desirable. From the second loss function, we explore the best combination of the compression network for dimension reduction and the estimation network for density estimation. The third loss function is the constraint term to prevent the singularity problem of the GMM. In this way, our anomaly detection model learns the characteristics of training patch samples, and it becomes overreact to unseen abnormal samples such as patches from regions of deterioration.

B. TEST PHASE

The trained model can provide energy that represents an abnormality score to a given patch sample $\boldsymbol{o}_{\text{test}}$. Since this model is trained on data for patches from normal regions in electron microscope images of rubber materials, if the observed sample has a structure similar to that for learned training normal samples, the given abnormality score becomes low. On the other hand, if the observed sample has a structure that is different from that of normal samples, the given score becomes high.

In the test phase, divided patches $\boldsymbol{o}_{\text{test}}$ from electron microscope images of test rubber material are inputted to the feature extraction network and transformed to feature vectors. Next, sample energy is calculated from the trained anomaly detection model for each test patch sample. In this way, it becomes possible to estimate the regions of deterioration in electron microscope images of rubber materials.

IV. EXPERIMENT

In this experiment, we performed quantitative and qualitative evaluations to confirm the effectiveness of our method for estimation of regions of deterioration. Experimental settings are presented in Subsec. IV-A and the results are discussed in Subsec. IV-B.

A. SETTINGS

A total of 16 electron microscope images of a rubber material were used in our analysis. In the images, four images were allocated to the training data and 12 images were allocated to the test data (namely, S1-S12). The ground truth of normal/deteriorated regions in the images was defined by a material scientist who has specialized knowledge of electron microscope imaging and characteristics of rubber materials. Since each image has a high resolution, which affects computation costs, we divided the images into multiple patches. The patch size was 128 × 128 pixels with a sliding interval of 10 pixels. For the training images, patches from only normal regions were used, and patches including deteriorated and unnecessary regions defined by the material scientist were excluded. For the test images, all of the patches were evaluated by the trained model. Finally, 32,476 patches from normal regions of the four images were used for training of our anomaly detection model. For each test image, the number of patches evaluated was 12,690. Parameters of TL-DAGMM used in this experiment are shown in Table 1. As the computational environment, an NVIDIA GeForce GTX 2080Ti GPU was used to train TL-DAGMM on a Linux operating system (Ubuntu 18.04; Canonical, London, England).

TABLE 1. Hyperparameters of TL-DAGMM.

Parameter	Value
Learning rate	0.0001
λ_1	0.1
λ_2	0.0001
Batch size	128
Epoch	400
GMM mixtures	8

We compared the performance of TL-DAGMM with the performance of other anomaly detection methods. The following comparative methods were employed.

TABLE 2. Hyperparameters of Convolutional DAGMM.

Parameter	Value
Learning rate	0.0001
λ_1	0.1
λ_2	0.0005
Batch size	64
Epoch	200
GMM mixtures	12

CM1: Convolutional DAGMM [29]
DAGMM is one of the state-of-the-art anomaly detection methods. To enable the training of DAGMM with high resolution images, we introduced a four-layer convolution network in the compression network. The parameters of Convolutional DAGMM are shown in Table 2.

CM2: HoG [13] + OCSVM [34]
This method is a popular approach in the field of MI. As hand-crafted features, we used HoG features and constructed a one-class support vector machine (SVM) [34] model as a classifier.

To assess the effectiveness of TL-DAGMM, we employed the following evaluations: visualization evaluation and object detection-based evaluation. First, in the visualization evaluation, we visualized anomaly scores assigned to test patches. Patch-level anomaly scores can be obtained by using the learned anomaly detection model. Since test patches have overlap regions due to the sliding interval settings, we visualized anomaly scores at the pixel level. Specifically, patch-level anomaly scores were assigned to each pixel, and multiple anomaly scores were averaged for each pixel. In this way, we provided pixel level anomaly scores to a test image. Next, in the object detection-based evaluation, pixel-based intersection over union (IoU) was used. By defining the degree of abnormality using a certain threshold value ξ, patch-level estimation of regions of deterioration becomes feasible. We determined the degree of abnormality for the sample as follows:

$$D(\boldsymbol{v}_{\text{test}}) = \begin{cases} \text{deteriorated} & (\text{if } E(\boldsymbol{v}_{\text{test}}) \geq \xi) \\ \text{normal} & (\text{otherwise}) \end{cases}, \quad (13)$$

where $\boldsymbol{v}_{\text{test}}$ is the extracted high representation features from $\boldsymbol{x}_{\text{test}}$.

Pixel IoU is defined as the intersection region of the proposal and the ground truth regions and can be calculated as follows:

$$\text{Pixel IoU} = \frac{\text{TP pixels}}{\text{TP pixels} + \text{FP pixels} + \text{FN pixels}}, \quad (14)$$

where TP, FP and FN represent True Positive, False Positive and False Negative, respectively. The threshold ξ that determines normal/deteriorated regions was experimentally set to a median value of anomaly scores for each sample. Moreover, we evaluated a significant difference in the brightness values of estimated deteriorated/normal regions by using the non-parametric Mann-Whitney U-Test.

B. RESULTS AND DISCUSSION

Figure 4 shows all of the test image samples (S1-S12) and their ground truth regions. The regions above the red lines are

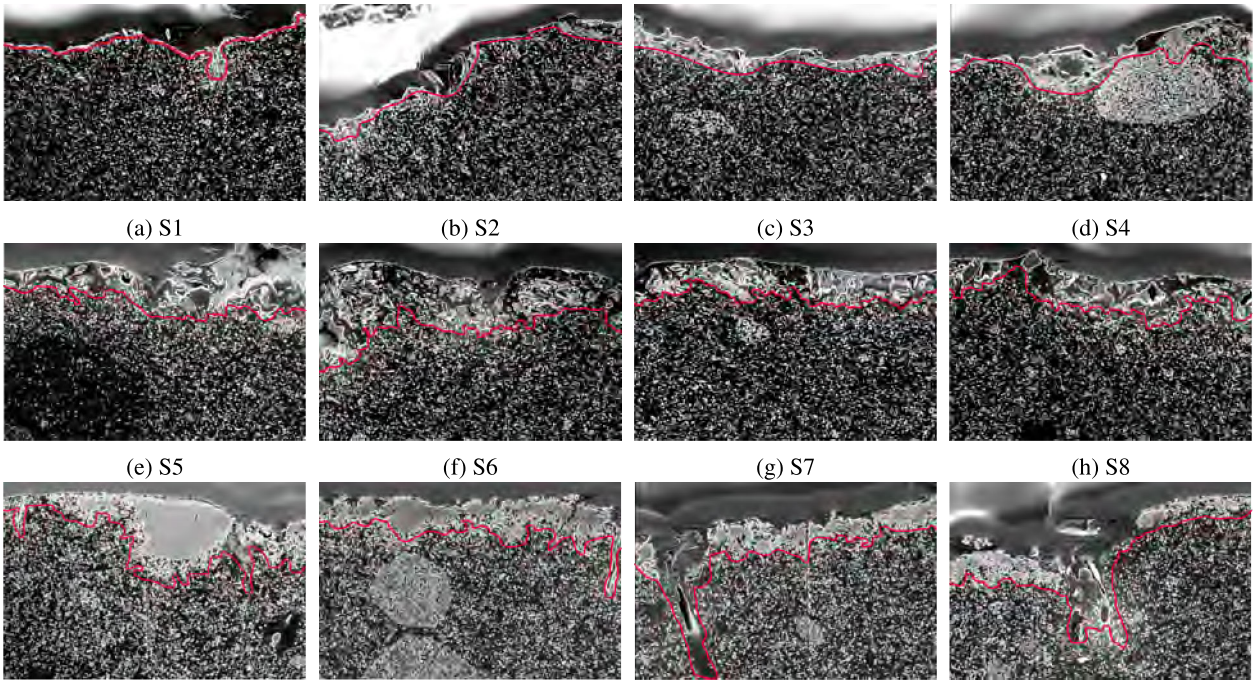


FIGURE 4. All of the test samples and their ground truth regions. The regions above the red lines are regions of deterioration and the regions under the red lines are normal regions.

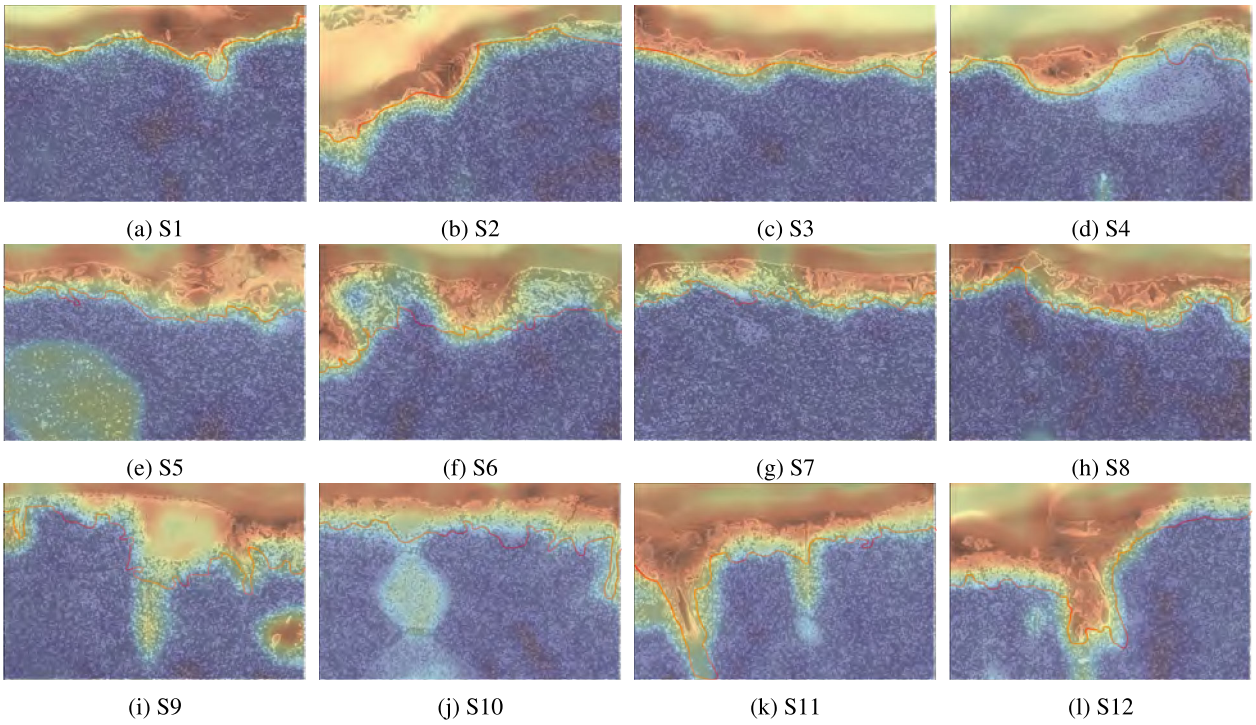


FIGURE 5. Visualization of anomaly scores by TL-DAGMM.

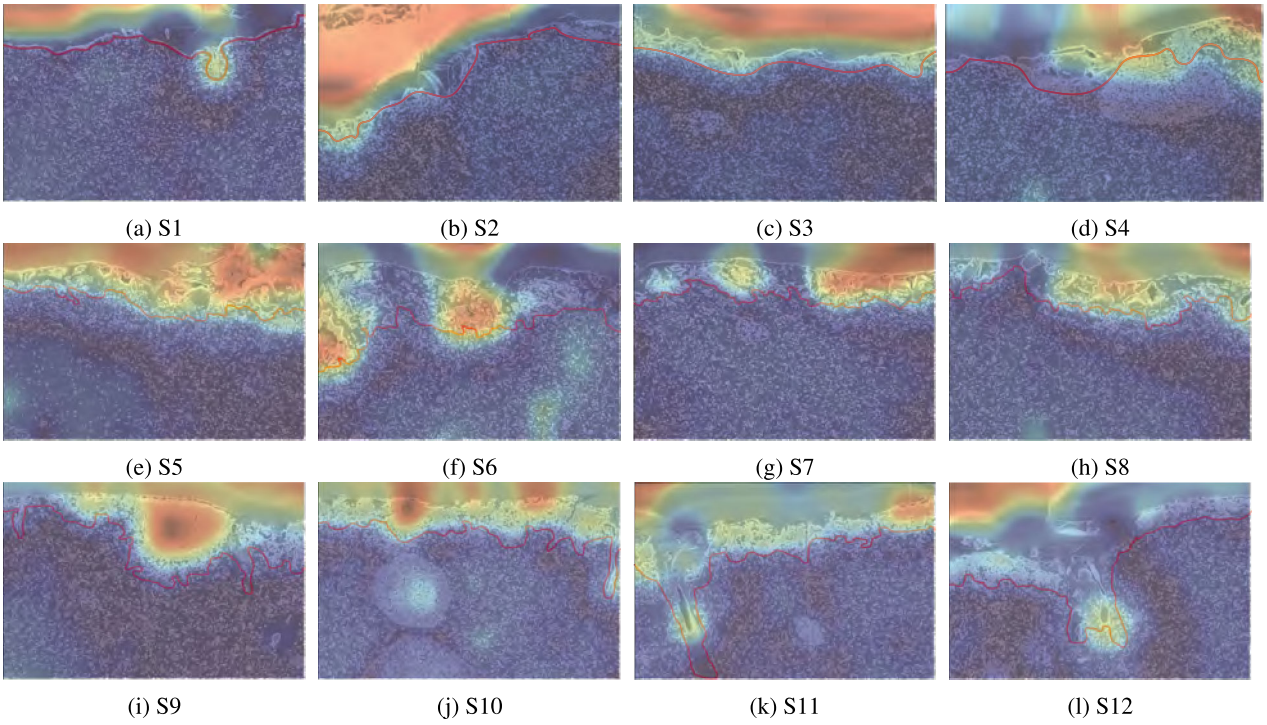


FIGURE 6. Visualization of anomaly scores by Convolutional DAGMM (CM1).

regions of deterioration and the regions under the red lines are normal regions. Visualization results are shown in Figs. 5-7. The position of each visualization result corresponds to the sample in Fig. 4. According to the heat map, the closer to a red color the heat map becomes, the higher are the anomaly scores, whereas blue color regions have low anomaly scores related to normal regions. Specifically, Figs. 5-7 represent the heat map overlaid images for each sample by TL-DAGMM,

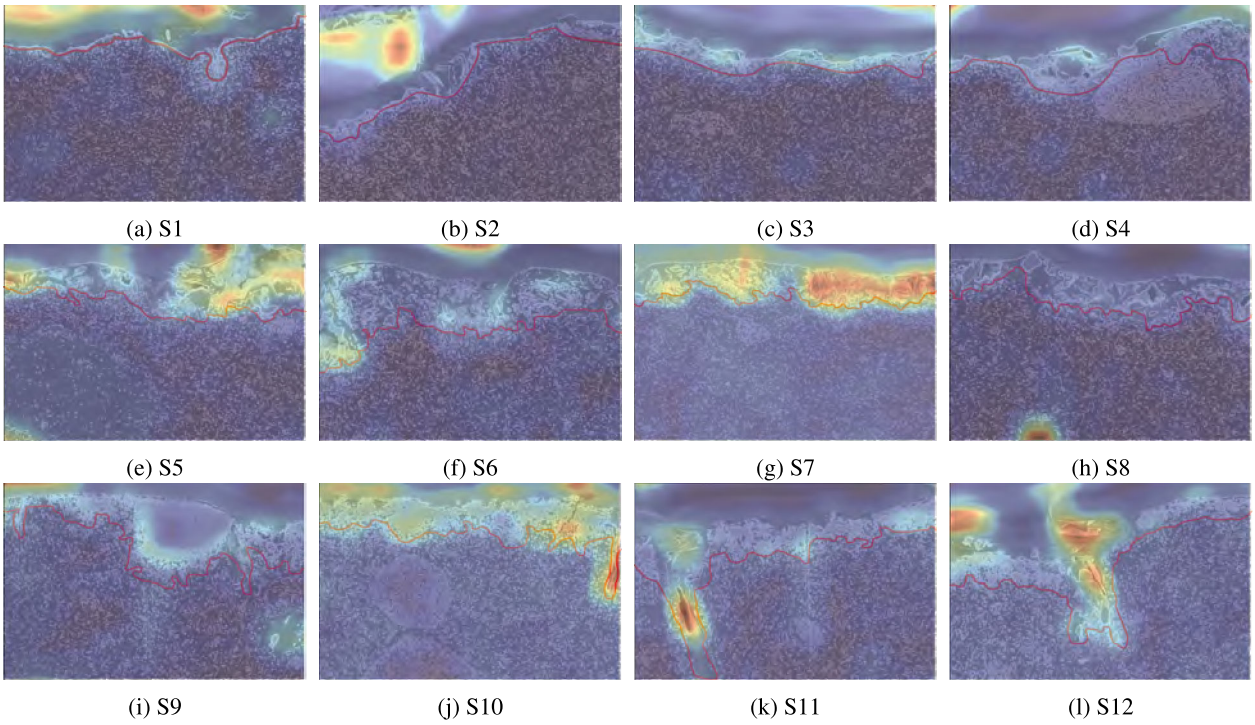


FIGURE 7. Visualization of anomaly scores by HoG + SVM (CM2).

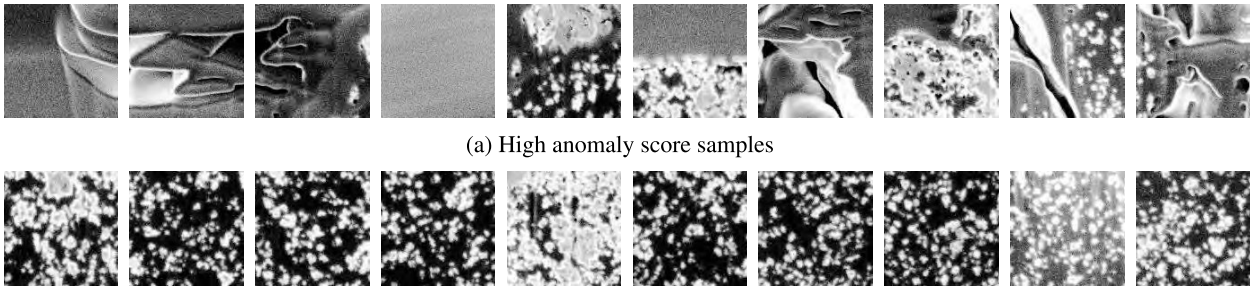


FIGURE 8. Examples of high and low anomaly score patch samples by TL-DAGMM.

TABLE 3. Pixel IoU of each sample.

Method	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean
TL-DAGMM	**0.394**	**0.673**	0.543	**0.620**	**0.575**	**0.837**	**0.568**	**0.599**	**0.665**	0.494	**0.782**	**0.818**	**0.631**
CM1	0.324	0.590	**0.544**	0.495	**0.575**	0.433	0.506	0.572	0.641	0.490	0.766	0.743	0.557
CM2	**0.394**	0.671	0.537	0.506	0.538	0.782	0.535	0.502	0.549	**0.499**	0.498	0.633	0.554

CM1, and CM2, respectively. From Fig. 5, we can see that regions with high anomaly scores are located in ground truth deteriorated regions. Although regions of deterioration have different types of characteristics, TL-DAGMM correctly detected such regions. On the other hand, we confirmed that the visualization results of CM1 and CM2 were inferior to those of TL-DAGMM qualitatively. It was difficult for CM2 using hand-crafted features to detect regions of deterioration.

Next, we show the object detection-based evaluation results. Pixel IoU of each test sample and the mean of them are shown in Table 3. Pixel IoU is one of the evaluation metrics that can evaluate the degree of overlapping between an estimated region and a target region. From the results, we quantitatively confirmed that estimated regions of deterioration correctly overlapped with the ground truth deteriorated regions in the test electron microscope images in TL-DAGMM. Furthermore, significant differences in the

brightness values of estimated deteriorated/normal regions appeared by the non-parametric Mann-Whitney U-test ($p < 0.001$) for all test samples. Sample patches that had high and low anomaly scores in TL-DAGMM are shown in Fig. 8. Although there were different types of patches in regions of deterioration, we confirmed that TL-DAGMM had the ability to detect such patches as an anomaly detection model.

This paper shows the potential of deep learning technologies using electron microscope images of rubber materials. To the best of our knowledge, this is the first work in which regions of deterioration in electron microscope images of rubber materials were estimated on the basis of a deep anomaly detection model. Our method benefited from the high-level feature representations of deep learning, and this simple approach may be effective for electron microscope images of other materials, though further investigations are necessary. Also, as shown in Fig. 5, some regions in addition to the ground truth regions had high anomaly scores. Although the visual characteristics of these regions are slightly different from those of normal regions, they are annotated as normal regions by specialists. Comparison of such regions and other regions may lead to the acquisition of new knowledge for the development of durable rubber materials. This advanced multidisciplinary study is one of our future works.

V. CONCLUSION

In this paper, we have presented a method for estimation of regions of deterioration using electron microscope images based on an anomaly detection. We demonstrated by experiments that our model TL-DAGMM can correctly detect regions of deterioration. In a future work, we will investigate the characteristics of rubber materials detected by our method for acquiring knowledge for the development of new durable materials.

ACKNOWLEDGMENT

Data were kindly provided by the Sumitomo Rubber Industries, Ltd., Japan.

REFERENCES

[1] R. R. John, "What is material informatics, materials informatics-effective data management for new materials discovery," Tech. Rep., 1999.
[2] A. Agrawal and A. Choudhary, "Perspective: Materials informatics and big data: Realization of the 'fourth paradigm' of science in materials science," *APL Mater.*, vol. 4, no. 5, Apr. 2016, Art. no. 053208.
[3] C. Jörgensen, "Access to pictorial material: A review of current research and future prospects," *Comput. Humanities*, vol. 33, no. 4, pp. 293–318, Dec. 1999.
[4] B. Meredig, "Industrial materials informatics: Analyzing large-scale data to solve applied problems in R&D, manufacturing, and supply chain," *Current Opinion Solid State Mater. Sci.*, vol. 21, no. 3, pp. 159–166, 2017.
[5] C. M. Breneman, L. C. Brinson, L. S. Schadler, B. Natarajan, M. Krein, K. Wu, L. Morkowchuk, Y. Li, H. Deng, and H. Xu, "Stalking the materials Genome: A data-driven approach to the virtual design of nanostructured polymers," *Adv. Funct. Mater.*, vol. 23, no. 46, pp. 5746–5752, Dec. 2013.
[6] B. L. DeCost and E. A. Holm, "A computer vision approach for automated analysis and classification of microstructural image data," *Comput. Mater. Sci.*, vol. 110, pp. 126–133, Dec. 2015.
[7] R. Jose and S. Ramakrishna, "Materials 4.0: Materials big data enabled materials discovery," *Appl. Mater. Today*, vol. 10, pp. 127–132, Mar. 2018.
[8] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine learning in materials informatics: Recent applications and prospects," *NPJ Comput. Mater.*, vol. 3, no. 1, p. 54, Dec. 2017.
[9] S.-Y. Ding, J. Yi, J.-F. Li, B. Ren, D.-Y. Wu, R. Panneerselvam, and Z.-Q. Tian, "Nanostructure-based plasmon-enhanced Raman spectroscopy for surface analysis of materials," *Nature Rev. Mater.*, vol. 1, no. 6, Apr. 2016, Art. no. 016021.
[10] A. Çeçen, T. Fast, E. C. Kumbur, and S. R. Kalidindi, "A data-driven approach to establishing microstructure–property relationships in porous transport layers of polymer electrolyte fuel cells," *J. Power Sources*, vol. 245, pp. 144–153, Jan. 2014.
[11] S. A. Shevchik, T. Le-Quang, F. V. Farahani, N. Faivre, B. Meylan, S. Zanoli, and K. Wasmer, "Laser welding quality monitoring via graph support vector machine with data adaptive kernel," *IEEE Access*, vol. 7, pp. 93108–93122, 2019.
[12] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2. Sep. 1999, pp. 1150–1157.
[13] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, no. 1, pp. 886–893, Jun. 2005.
[14] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, Jan. 1996.
[15] A. Chowdhury, E. Kautz, B. Yener, and D. Lewis, "Image driven machine learning methods for microstructure recognition," *Comput. Mater. Sci.*, vol. 123, pp. 176–187, Oct. 2016.
[16] P. Li, D. Wang, L. Wang, and H. Lu, "Deep visual tracking: Review and experimental comparison," *Pattern Recognit.*, vol. 76, pp. 323–338, Apr. 2018.
[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
[18] G. Wu, M. Kim, Q. Wang, B. C. Munsell, and D. Shen, "Scalable high-performance image registration framework by unsupervised deep feature representations learning," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1505–1516, Jul. 2016.
[19] J. Abeßer and M. Müller, "Fundamental frequency contour classification: A comparison between hand-crafted and CNN-based features," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 486–490.
[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.
[21] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014.
[22] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
[23] W. Choi, H. Huh, B. A. Tama, G. Park, and S. Lee, "A neural network model for material degradation detection and diagnosis using microscopic images," *IEEE Access*, vol. 7, pp. 92151–92160, 2019.
[24] Y. Ma, H. Wang, H. Li, and J. Hong, "Study on metal rubber material's characteristics of damping and sound absorption," in *Proc. ASME Turbo Expo., Power Land, Sea, Air*, Jan. 2008, pp. 477–486.
[25] X. Li, W. Berger, C. Musante, and M. I. Mattina, "Characterization of substances released from crumb rubber material used on artificial turf fields," *Chemosphere*, vol. 80, no. 3, pp. 279–285, Jun. 2010.
[26] A. Lugauskas, L. Levinskait, and D. Pečiulyt, "Micromycetes as deterioration agents of polymeric materials," *Int. Biodeterioration Biodegradation*, vol. 52, no. 4, pp. 233–242, Dec. 2003.
[27] S. Cakir, D. C. Kahraman, R. Cetin-Atalay, and A. E. Cetin, "Contrast enhancement of microscopy images using image phase information," *IEEE Access*, vol. 6, pp. 3839–3850, 2018.
[28] J.-D. Gu, "Microbiological deterioration and degradation of synthetic polymeric materials: Recent research advances," *Int. Biodeterioration Biodegradation*, vol. 52, no. 2, pp. 69–91, Sep. 2003.
[29] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. Cho, and H. Chen, "Deep autoencoding Gaussian mixture model for unsupervised anomaly detection," in *Proc. Int. Conf. Learn. Representations (ICLR)*, Feb. 2018.
[30] R. Togo, K. Ishihara, T. Ogawa, and M. Haseyama, "Estimation of salient regions related to chronic gastritis using gastric X-ray images," *Comput. Biol. Med.*, vol. 77, pp. 9–15, Oct. 2016.

[31] R. Togo, N. Yamamichi, K. Mabe, Y. Takahashi, C. Takeuchi, M. Kato, N. Sakamoto, K. Ishihara, T. Ogawa, and M. Haseyama, "Detection of gastritis by a deep convolutional neural network from double-contrast upper gastrointestinal barium X-ray radiography," *J. Gastroenterology*, vol. 54, no. 4, pp. 321–329, Apr. 2019.
[32] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[33] A. Zimek, E. Schubert, and H.-P. Kriegel, "A survey on unsupervised outlier detection in high-dimensional numerical data," *Stat. Anal. Data Mining*, vol. 5, no. 5, pp. 363–387, Oct. 2012.
[34] C. Cortes and V. N. Vapnik, "Support vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1997.

REN TOGO (S'16–M'19) received the B.S. degree in health sciences from Hokkaido University, Japan, in 2015, and the M.S. and Ph.D. degrees from the Graduate School of Information Science and Technology, Hokkaido University, in 2017 and 2019, respectively. He is currently a Research Fellow of the Japan Society for the Promotion of Science, and Radiological Technologist. His research interest includes machine learning and its applications.

NAOKI SAITO (S'16–M'19) received the B.S. degree in electrical and electronic engineering from the National Institution for Academic Degrees and University Evaluation, Japan, in 2014, and the M.S. and Ph.D. degrees from Hokkaido University, Japan, in 2016 and 2019, respectively. He is currently an Assistant Professor with the Department of Creative Engineering, National Institute Technology, Kushiro College. His research interests include materials informatics and image processing. He is also a member of ACM and IEICE.

TAKAHIRO OGAWA (S'03–M'08–SM'18) received the B.S., M.S., and Ph.D. degrees in electronics and information engineering from Hokkaido University, Japan, in 2003, 2005, and 2007, respectively. He is currently an Associate Professor with the Faculty of Information Science and Technology, Hokkaido University. His research interest includes multimedia signal processing and its applications. He is also a member of EURASIP, IEICE, and the Institute of Image Information and Television Engineers. He has been an Associate Editor of the *ITE Transactions on Media Technology and Applications*.

MIKI HASEYAMA (S'88–M'91–SM'06) received the B.S., M.S., and Ph.D. degrees in electronics from Hokkaido University, Japan, in 1986, 1988, and 1993, respectively. She joined the Graduate School of Information Science and Technology, Hokkaido University, as an Associate Professor, in 1994. She was a Visiting Associate Professor with Washington University, USA, from 1995 to 1996. She is currently a Professor with the Faculty of Information Science and Technology, Hokkaido University. Her research interest includes image and video processing and its development into the semantic analysis. She is also a member of IEICE, ITE, and the Information Processing Society of Japan IPSJ. She has been the Vice President of the Institute of Image Information and Television Engineers, Japan (ITE), the Director of the International Coordination and Publicity of The Institute of Electronics, Information and Communication Engineers (IEICE), and the Editor-in-Chief of the *ITE Transactions on Media Technology and Applications*.