IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 April 2024, accepted 28 April 2024, date of publication 6 May 2024, date of current version 14 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3396818

RESEARCH ARTICLE

Age Assessment Using Chum Salmon Scale by Neural Networks and Image Processing

GENKI SUZUKI[1], (Member, IEEE), MIKIYASU NISHIYAMA[1], (Student Member, IEEE), RYOMA HOSON[1], KATSUNOBU YOSHIDA[1], HIROYUKI SHIOYA[1], AND KAZUTAKA SHIMODA[2]

[1]Division of Information and Electronic Engineering, Muroran Institute of Technology, Muroran 0508585, Japan
[2]Fisheries Research Department, Salmon and Freshwater Fisheries Research Institute, Hokkaido Research Organization, Eniwa 0611433, Japan

Corresponding author: Genki Suzuki (suzuki@muroran-it.ac.jp)

This work was supported by the Strategic Information and Communications Research and Development Promotion Programme (SCOPE), Ministry of Internal Affairs and Communications, under Grant 16771288.

ABSTRACT In its long history, chum salmon has been propagated by hatching and stocking. Age assessment is necessary to monitor the migration status of chum salmon, and it is desirable to increase the number of assessments in order to obtain precise and accurate knowledge about salmon as a resource. This study introduced an automated age assessment system that uses information technologies, such as artificial neural networks and image processing, to improve efficiency compared to manual assessment. Specifically, we first developed a method for extracting scale regions from scale replica sample images and creating a database. We then used a resting zone detection technique based on semantic segmentation, allowing us to automate evaluation using scaled images as part of the age assessment method. Age assessment from resting zone images was realized by an image processing method based on the skills of the staff of the Japanese Fisheries Research Institute, such as an age counting method using circular structures. Experimental results show that our proposed method outperforms other approaches; salmon age assessment was accurate to within an error of 0.5 years. Our method encompasses automatic data processing of scale images, resting zone detection, and age assessment, and will contribute to the efficiency of the Fisheries Research Institute in terms of both work efficiency and salmon data research.

INDEX TERMS Age assessment, scale analysis, fisheries data, deep learning, neural network, image processing, object detection.

I. INTRODUCTION

Seafood is an important food resource for mankind, and the status of fisheries resources must be elucidated in order to use those resources sustainably. Since it is impossible to visualize the entire ocean, monitoring methods are commonly used in fisheries resource surveys [1], [2], [3]. Chum salmon have proliferated through hatching and stocking by humans, and the number and age distribution of migrating chum salmon are the primary data for understanding their resource status [4]. The age of all chum salmon caught as seafood needs to be determined. Age is estimated from the scales of salmon captured as samples, and the number of fish at each age is predicted from the age distribution of the sample and the chum salmon catch [1], [2], [3]. Recently, approximately 20 million chum salmon were observed in the waters around Hokkaido, Japan. Hokkaido's share of the salmon catch exceeds 80% of Japan's total. The number of samples needed for age estimation is approximately 20,000 fish, however, to improve the accuracy of distribution estimation for four or more ages, it is desirable to have samples of more than 20,000 fish. To that end, it is essential to speed up age assessment based on the examination of scales.

Fisheries Research Institutes in Japan carry out marine resource management, research, and surveys for fishers. The age of many fish species can be determined by analyzing scales and otoliths, which are manifestations of the growth state of the fish often used in fisheries science [5], [6], [7].

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

More than 20 years ago, the Fisheries Research Institute in Hokkaido created salmon scale samples and predicted the number of chum salmon visits by age assessment, as shown in Fig. 1. Specifically, for salmon scale samples, Fisheries Research Institute staff annually collect scales from approximately 20,000 salmon that migrate to rivers and estuaries in Hokkaido and transfer them to plastic plates to create replica scale samples (upper image in Fig. 2). In addition, experts predict the number of visiting salmon through the use of magnifying glasses to assess the age of replica samples based on their scale patterns, and they annually calculate the number of individuals of each age group based on the results of 20,000 samples. These tasks take up a large percentage of the experts' annual work time and require a great deal of effort by a limited number of individuals due to the specialized nature of the work. Comprehensive investigations based on the scales and climate change data collected to date are required for the root cause analysis of defects [4], [8], [9], [10], [11], [12]. Specifically, since the scale pattern embeds the growth record of salmon, a non-fishing survey based on the scale features of many replica samples should be conducted. For this reason, a system for the data conversion of the scale replicas stored at the Fisheries Research Institute and the subsequent extraction of scale features is required by the institute staff. However, the time for such investigations is limited because the work of age assessment requires a great deal of labor. Thus, the efficiency of age assessment work at Fisheries Research Institutes must be improved by utilizing information technology.

Information science and technology-based analyses of scales have been conducted previously [13], [14], [15], [16], [17], [18]. Specifically, age assessment methods utilizing image processing and neural networks have been proposed [16], [17], [18]. These methods [16], [17] manually extract scale regions from scale images from plastic plates using a microscope to obtain images for age assessment, and then calculate age using image processing and neural networks. One method [16] assesses age by applying image processing to focus on the shading of pixels according to the annual rings of scales. However, this method is affected by variations in the shape and spacing of scale patterns among individual salmon, as well as by the unevenness of pixel shading, making it difficult to assess the age of salmon consistently. Another method [17] classifies annual ring labels for each patch image, which makes it difficult to detect resting zones, which are the curves that determine age, on a pixel-by-pixel basis. This difficulty reduces the accuracy of age assessment. Therefore, these methods are still in the basic research stage of scale age analysis and have not yet led to high accuracy or to the automation of scale age assessment from replica samples from plastic plates. On the other hand, one method [18] used neural networks to assess the age of salmon based on raw scale images taken with a microscope; the age can be estimated as an image classification task. However, since that method is a classification model of

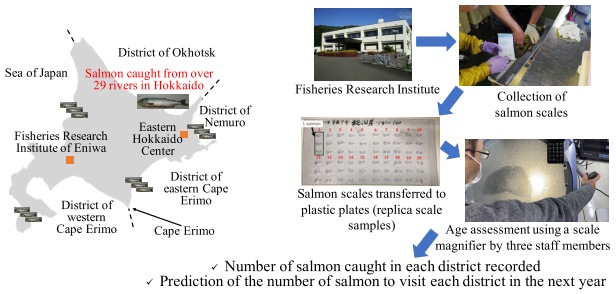


FIGURE 1. Left figure presents the five districts and the location of Fisheries Research Institute of Eniwa and the Eastern Hokkaido Center, and right figure shows the workflow of salmon age assessment.

images, it is difficult to obtain the positions and numbers of resting zones or false annual rings (lower image in Fig. 2) on the scales as output data, which are what expert staff focus on during age assessment [19], [20]. The automation of the age assessment process should be based on the integration of experts' decisions and strategies, coupled with deep learning and image processing technologies. From the above, it is clear that it is necessary to construct a new system that simultaneously performs feature detection of scale patterns on scales while automating age assessment with high performance to improve the operation of the Fisheries Research Institute. For this purpose, it is especially important that the processing to be done by the new system reflect the current age assessment process conducted by experts working at Japanese Fisheries Research Institutes, such as the detection of resting zones that signify the age of the salmon and the method of age assessment based on circular structures [21]. Therefore, to speed up high-accuracy scale age assessment, it is necessary to automatically extract scale regions from replica samples to create a database of scale images and to construct a method that integrates deep learning and image processing to improve the performance of age assessment.

This paper proposes a method for automatically assessing the age of salmon based on their scales using deep learning and image processing. The proposed method is overviewed in Fig. 3. First, scale regions are extracted from replica sample images by a fine-tuned convolutional neural network (CNN) [22] that detects objects based on deep learning [23]. Next, each scale region image is linked to an individual salmon by labeling based on clustering [24] using the position data of multiple extracted scale regions. The resting zone is then detected on scale images based on the semantic segmentation model [25]. Finally, scale age is assessed by expanding the resting zone image using polar coordinates based on the scale center and calculating the number of resting zones. At this point, precise age assessment is realized by supplementing the missing parts of the detected resting zones. Thus, it is possible to automatically assess the age of salmon from replica sample images with high accuracy.

The contributions of our research include the following:

- Partial streamlining and automation of scale age assessment operations

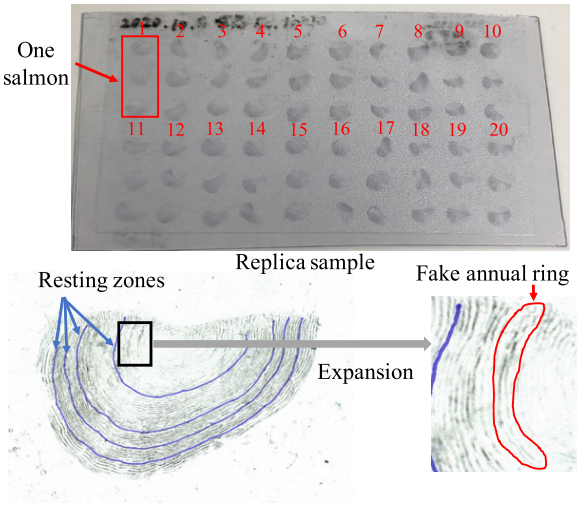


FIGURE 2. Replica sample of chum salmon scales and detailed salmon scale features.

Most scales can be assessed by a computer running the software that contains our method, while scales that are difficult to assess because of defects in shape can be handled by expert staff. This approach is expected to improve operational efficiency. To automate the process, the system must have a certain level of performance. However, the present study newly developed two methods that enable precise age assessment: a method for completing the missing part of the resting zone, and a method for restoring the circulus and excluding false rings from the age count, which is based on the process used by Fisheries Research Institute staff who predict the circular structure of the part of the scale interrupted by the resting zone.

- Data sufficiency for age assessment and resting zones for predicting salmon migration

The proposed method enables the age assessment of a larger number of individuals than can be assessed by the conventional method at the Fisheries Research Institute. Therefore, it is expected to improve the reliability of the regression rate [1], [2], [3] calculated from the number of released salmon and the number of migrating salmon. The regression rate is an important indicator in the determination of stock trends because it can sequentially predict the percentage of migrating 3-year-old salmon that will return the following year as 4-year-olds. Furthermore, in addition to age assessment, the proposed method can comprehensively create an image database of scale replicas and detect resting zones. Therefore, a large amount of replica sample data stored in Fisheries Research Institutes to investigate poor catches can be effectively utilized by institute staff.

This paper is organized as follows. Scale sampling of chum salmon and age assessment at the Fisheries Research Institute in Hokkaido are presented in Section II. The proposed method is presented in Section III. Experimental results verifying the effectiveness of the proposed method are shown in Section V. Finally, concluding remarks are presented in Section VI.

II. SCALE SAMPLING OF CHUM SALMON AND AGE ASSESSMENT AT FISHERIES

This section describes the age assessment operations at the Fisheries Research Institute. This work is performed by several staff according to a manual based on [21]. Scales used for age assessment are obtained from salmon caught in several estuaries in Hokkaido. Salmon scale replicas are prepared by molding the scales in plastic, after which the replicas are stored at the institute. Fig. 2 shows a replica sample of a salmon scale. Each plate includes 2-3 scales from each of 20 salmon [26]. For age assessment, experts examine the replica samples with a magnifying glass and assess the salmon's age based on the number of resting zones, as shown in Fig. 4 [19]. Three experts assess age comprehensively from the results of these evaluations. The resting zones shown in Fig. 2 are areas where the circuli are closely spaced, indicating that they formed in winter, when growth was stagnant [19], [20]. In other words, when water temperature and the growth rate are both high, the gap between the circuli widens, while, when the water temperature is low, the growth rate is low, and thus the gap between the circuli narrows. By counting the resting zones, the fisheries staff can determine how many winters a salmon has spent in the ocean, i.e., its age. When an area of interrupted rest zone occurs, the staff counts the age while predicting the circular structure. Specifically, the staff finds areas with no defects and where the number of the resting zones is easy to count and then predicts an extension line from those areas. Note that the experts exclude the fake annual ring found in the nucleus of the scales from the resting zone count because it is located in an area other than the one where age traits form during the first and second years of growth (circuli 15-17 from the nucleus). For samples that are difficult to assess due to missing scales, age is confirmed by viewing the scales directly under a microscope. The number of missing scales should be small, and an automatic age assessment of a large number of scales with clean shapes is necessary.

III. AGE ASSESSMENT USING NEURAL NETWORKS AND IMAGE PROCESSING

This section describes a method using neural networks and image processing to assess salmon age collaboratively based on images of scales. The proposed method consists of four phases, as shown in Fig. 3. In the first phase, a replica sample of the scales (described above) is scanned by a high-resolution scanner. In the second phase, scale regions are detected based on the object detection model, and a database of scale images is created based on the k-means [24] method. In the third phase, resting zones are detected by PSPNet [25], which is a segmentation model that operates via a CNN. Since the previously reported method of detecting resting zones in scale images [15] performs well, we use the same method here. In the fourth phase, salmon age is assessed by counting the resting zones, which is done by expanding the polar coordinates of the image from the nucleus of the scale

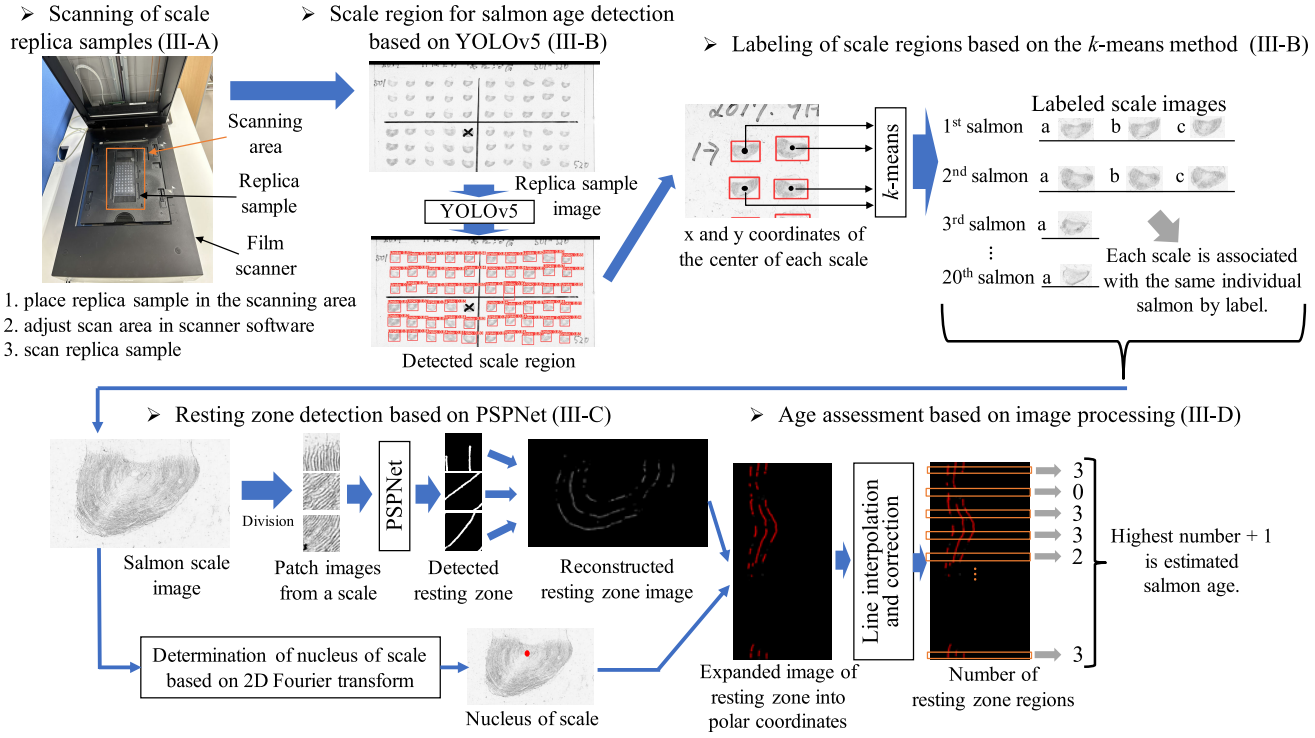


FIGURE 3. Overview of the proposed method. In pre-processing, scale replica samples are scanned (Section III-A). Next, scale regions are detected based on YOLOv5 (Section III-B). The resting zones are then detected by PSPNet, which is a semantic segmentation model (Section III-C). Finally, image processing (Section III-D) assesses the age of the salmon based on one of its scales.

FIGURE 4. Scale magnifier at the Fisheries Research Institute.

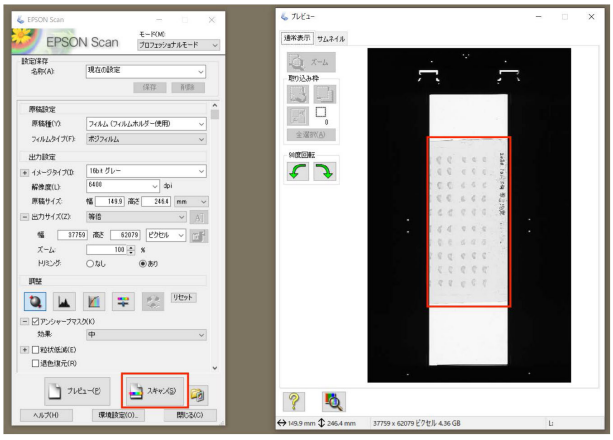


FIGURE 5. Settings of film scanner for the high-resolution scanning of replica samples in III-A.

using image processing. The following sections describe each of these processes in greater detail.

A. PRE-PROCESSING: SCANNING OF SCALE REPLICA SAMPLES USING A HIGH-RESOLUTION SCANNER

A scan of the replica sample is required to create an automated system for age assessment. Therefore, as a pre-processing step for evaluating scale age, a replica of the scale is scanned at high resolution. Specifically, the replica is scanned using a film scanner ,[1] as shown in Fig. 3. The scanner illuminates the replica and reads the transmitted light. The replica is placed in the center of the scanner and its outline is scanned. Specifically, we adjust the scan area using the software (Fig. 5) supplied with the scanner (red box on the right in the figure) and scan the replica samples (red box on the bottom left). In the present study, the resolution of the

[1]URL: https://www.epson.jp/products/scanner/gtx980/

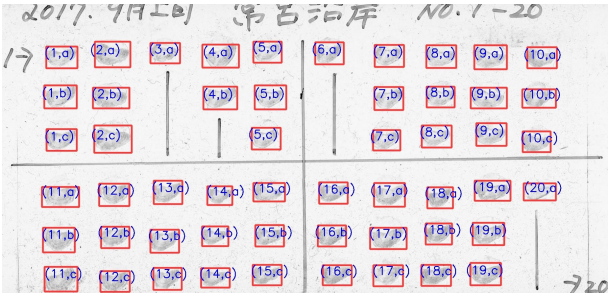


FIGURE 6. Example of scale regions detected using YOLOv5 and labeled scales. The red boxes indicate scale regions detected by III-B. The blue numerals indicate individual salmon numbers (1)–20), and each scale of each individual salmon is labeled with the letter a, b, or c by *k*-means clustering.

scanned images was 27,716 × 13.010, and the film scanner makes it possible to capture the delicate patterns of the scales.

B. SALMON SCALE DETECTION BASED ON YOLOV5

This section explains the detection of salmon scale regions using YOLOv5 [27]. More than one scale per individual are arranged in a column in a sample image of scales. The scales vary slightly in size depending on the age of the fish, but the overall average scale size is relatively consistent across ages. However, there are cases in which fisheries staff decide not to use scales for age assessment because they are missing too many, and the collected scales are marked with crosses (replica sample image in Fig. 3). Furthermore, since scales are biological samples, there are cases in which not all scales were collected in triplicate (Fig. 6). Therefore, it is necessary to devise a method for the precise detection of scale areas. For example, an image processing approach based on changes in pixel values can be considered, but it would be necessary to create conditions that take into account the positions of replicas, characters, and lines, and the conditions may become more complex depending on the state of each replica. Therefore, we used the object detection model YOLOv5 to enable the proposed method to detect scale regions that are robust to changes in scale position, orientation, and size. YOLOv5 is a model that simultaneously detects the target object location and classifies the object. First, the model is trained by fine-tuning using other prepared scale images of chum salmon. Fine-tuning entails the adaptation of knowledge of object shape and color learned in images from other domains to train the target domain image [22]. This enables the detection of scale regions that are robust to changes in the position, orientation, and size of scales. In training, the loss of YOLOv5 is calculated as a combination of three individual loss components as follows:

$$\text{loss}_{\text{YOLO}} = \lambda_1 L_{\text{cls}} + \lambda_2 L_{\text{obj}} + \lambda_3 L_{\text{loc}} \tag{1}$$

where λ is the balance weight. L_{cls} calculates the error of the classification task. L_{obj} computes the error in detecting whether a target object exists in a particular grid cell of the training image divided into rectangular regions. L_{loc} is the loss of intersection over union (IoU) [28] and calculates the error in localizing the object in the grid cell. Then, the proposed method detects the scale region by inputting target images of salmon scales acquired during III-A into the trained model. Fig. 6 shows an example of scale region detection using YOLOv5. The red boxes indicate the detected scale regions. Because scales may not fit within a red box in YOLOv5 detection, a box can be enlarged by moving the top line up or the bottom line down.

As shown in Fig. 6, each replica sample consists of up to three scales from each of 20 salmon. On occasion, a single scale or two scales were collected from an individual, instead of the usual three, when scales that exhibited severe deficiencies were eliminated. To automate age assessment, it is necessary to associate the unique identification number of each salmon in the replica sample with every image of its scales detected using YOLOv5. Therefore, we use clustering to label scale regions detected by YOLOv5. The *k*-means [24] method, a nonhierarchical clustering method for machine learning, is an algorithm for separating a set of multivariate numerical data into an arbitrary number of clusters. Since the numbers of vertical and horizontal arrangements of scale replica samples vary from one Fisheries Research Institute to another, the *k*-means method, which allows adjustment of the number of clusters, helps label various replica samples. First, the proposed method obtains the center coordinates of each scale region detected by YOLOv5. Next, each scale region's clustering is carried out using the *k*-means method. The numbers of clusters for *k*-means are set to 10 along the x-axis and 6 along the y-axis, since the replica sample in this study has scales arranged in 10 columns ($N_x = 10$) and 6 rows ($N_y = 6$). In the proposed method, the segmentation of scale regions is achieved by applying the *k*-means method separately to the cluster centers along each axis direction (x- and y-axes) of the scale regions, as shown in Fig. 7. Specifically, the clustering procedure using *k*-means is as follows. **Step 1:** Randomly assign cluster numbers from 1 to N_x (N_y in the y-axis direction) to the center coordinate data of each scale region. **Step 2:** Iterate the following two processes until the cluster number does not change. **Step 2-1:** Calculate the coordinates of the cluster's center for each cluster number using the center coordinate data of the scale region. **Step 2-2:** Update the cluster number of each piece of data to the cluster number with the closest cluster's center coordinates. Finally, the obtained clustering results are used to label the scale regions. Labeling is performed on the obtained clustering results based on the placement rules used by the Fisheries Research Institute to create scale replica samples as described in Section II. Specifically, as shown in Fig. 6, based on the clustering results, each salmon is associated with its scales by assigning a numerical label from 1 to 20 as the individual salmon identifier, with each scale from that individual being designated as *a*, *b*, or *c*. From the above, the scale region extraction results are labeled, and then aligned to create scale region images denoted as $I_{i,j}$ (where i ranges from 1 to 20, representing the individual salmon, and j can

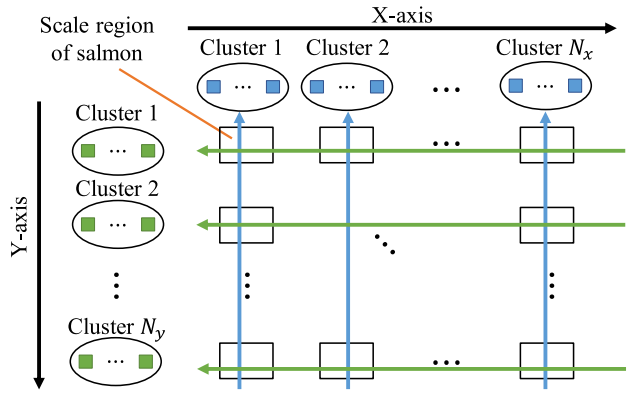


FIGURE 7. Overview of scale region clustering.

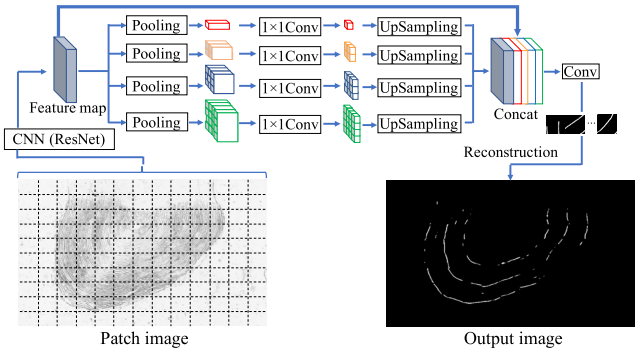


FIGURE 8. Overview of PSPNet-based resting zone detection.

be either a, b, or c, corresponding to the scales of the same salmon).

C. RESTING ZONE DETECTION BASED ON PSPNET

Here, we used a segmentation-based neural network model to detect rest zones in salmon scale images based on [15]. A resting zone of salmon scales is caused by the narrowing of the distance between scale circuli when the scales are formed during cold water temperatures in winter. Since the identification of resting zones by experts using a magnifying glass is time-consuming, it is essential to automate resting zone detection using neural networks.

1) INTRODUCING THE PSPNET NETWORK

Semantic segmentation is the estimation of the class of each pixel in an image. It is widely used in research to separate pedestrians, sidewalks, and vehicles in images for automated driving [29], [30]; to detect defects in material images [31]; and to process medical images [32], [33], [34]. SegNet [30] and U-Net [34], the best known networks in semantic segmentation, use an encoder-decoder structure [35], [36], which many recent semantic segmentation systems employ. PSPNet, which we use for resting zone detection, also has an encoder-decoder structure, but this model consists of a pyramid pooling module, as described in [37] and [38], between the encoder and decoder. In general, CNNs used in semantic segmentation artificially determine the size and scale of the input, which leads to a decrease in recognition accuracy for partial images. To solve this problem, various approaches that incorporate a wide range of contexts as additional information have been proposed [37], [39]. Among them, PSPNet is a model that applies a pyramid pooling module, which uses multiresolution and multiscale features as context, to semantic segmentation. When the feature map obtained at the encoder side is enlarged, it is possible to capture features with different scales using the pyramid pooling module, which enlarges the feature map at multiple scales. This enables the generation of fixed-length representations regardless of image size or scale. In other words, PSPNet can pick up both the global context of the image and information about smaller parts of the image, and it is robust to object deformation [39]. This makes it effective for detecting resting zones on scales that consist of continuous curves and shapes similar to those of normal circuli. In the following, we outline the network architecture of PSPNet and detail a technique for identifying resting zones within scale region images.

2) PSPNET NETWORK ARCHITECTURE

PSPNet is a network that adds a pyramid pooling module to the terminal layer of RESNet [40]. The network architecture is shown in Fig. 8. PSPNet first extracts a feature map for the input image by CNN (ResNet), which is pretrained with the correct labels of the resting zone of the scale based on the dilated network strategy [35], [41]. During this extraction, the feature map is $\frac{1}{8}$ of the original input image due to downsampling. Next, max-pooling is applied to the feature map by passing it through four pyramid-type pooling layers. The features are represented as the whole, a half, and small portions of the image by multiscale max-pooling of 1×1, 2×2, 3×3, and 6×6. Then, dimensionality reduction in the channel direction is performed using a 1×1 convolution layer. If the number of layers in the pyramid pooling module is M, the number of channels in each feature map after reduction is $\frac{1}{M}$. Next, each feature map is upsampled by bilinear interpolation with convolution processing for each size. The four upsampled feature maps are combined with the channels behind the original feature maps. This produces an extended feature map with both global context and local information. Finally, a 1×1 convolution is applied to each element in the channel direction of the extended feature map to generate a prediction map for the resting zone.

To detect scale resting zones, ResNet is trained using scale images and correct labels created by experts based on the scale images. First, the scale and correctly labeled images are divided into rectangular patch images $I_{i,j(w,h)}$ with a width w of 10 ($w = 1, 2, \cdots, 10$) and a height h of 13 ($h = 1, 2, \cdots, 13$). Next, the resting zone detection model is constructed by training with the original segmented image and the image with the correct label. Specifically, PSPNet calculates a loss for each pixel in the training image. This loss is based on cross-entropy loss using probability,

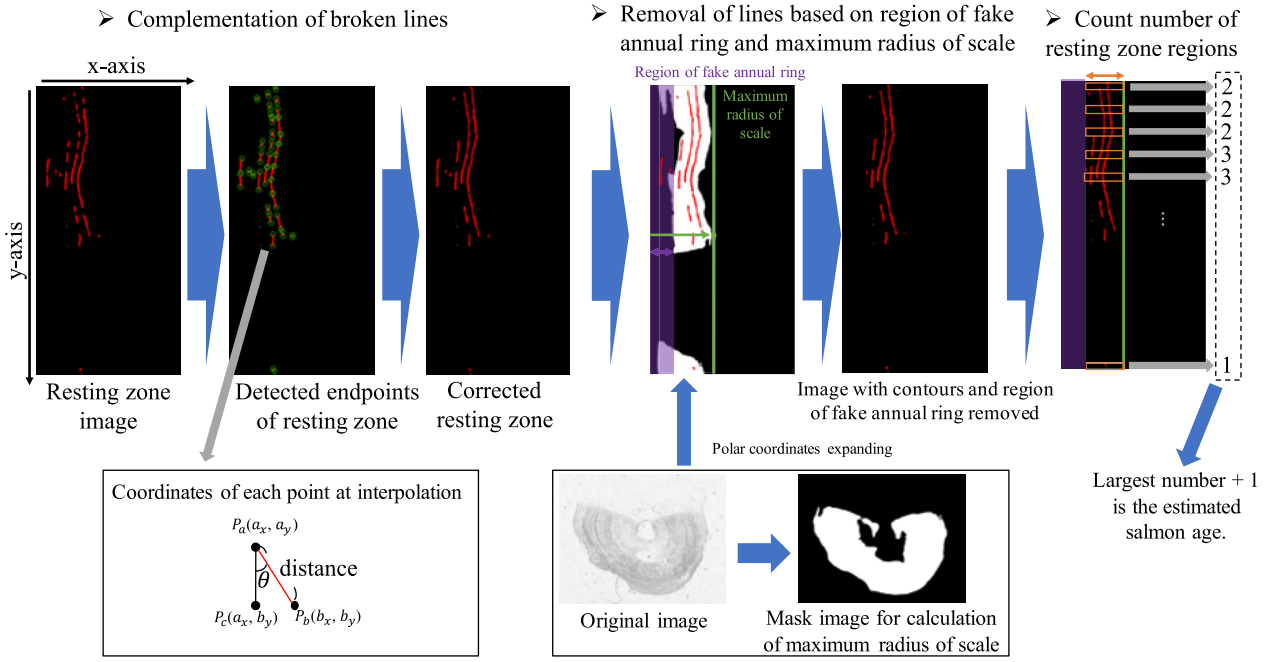


FIGURE 9. Details of the age assessment.

and the loss is used by PSPNet to calculate the probability of any given pixel being in the resting zone. Cross-entropy loss is an important element in neural network classification problems, including semantic segmentation, and is often used for training. In the training phase, for example, when the output value of a pixel in a patch output from PSPNet is z and the index of the correct answer class of the corresponding pixel (pixel in the resting zone) is q, the loss is calculated using probability-based cross-entropy loss as in the following formula:

$$\text{loss} = -\log\left(\frac{\exp(z_q)}{\sum_{r=1}^{C}\exp(z_r)}\right), \tag{2}$$

where C is the number of classes. Here we use the case of $C = 2$ to detect whether or not the pixel in question lies in the resting zone. Then, the resting zone is detected by inputting the segmented scale images of the test into the constructed resting zone detection model as follows:

$$O_{i,j(w,h)} = \text{PSPNet}(I_{i,j(w,h)}). \tag{3}$$

Finally, the output image $O_{i,j(w,h)}$ group for each patch is combined into an image of the same size as the input image to obtain a detection image $O_{i,j}$ of the rest band. This process allows the detection of resting zones at the pixel level based on the overall context of the resting zone and local features.

D. AGE ASSESSMENT BASED ON IMAGE PROCESSING

Next, we assess salmon age from a scale image using image processing. Counting the resting zones makes it possible to assess the age of a salmon based on the number of times it has overwintered. Specifically, we describe a method for assessing age from the resting zones based on polar coordinate expansion and line interpolation. Experts at the Fisheries Research Institute assess scale age by counting the resting zones from the nucleus outward. In the previous method based on image processing, salmon age was assessed by counting resting zones using the same method used by the experts [16]. Therefore, in the proposed method, the nucleus of the scale is calculated, pixel values are referenced from the nucleus to the contour of the scale, and age is assessed based on the number of resting zones detected in the previous section.

1) DETERMINING THE NUCLEUS OF A SCALE

Patterns in salmon scales appear as concentric circles, which are often half ellipses rather than regular circles. As a result, determining the center of a scale by using circle detection methods like the Hough transform can be a challenging task [42]. Therefore, we use a nucleus calculation method based on the pattern shapes of scales using [16]. First, the pattern of a salmon scale is characterized by concentric overlapping circles, and their stripes are oriented toward the center of the scale. Therefore, the Fourier intensity is calculated from the 2D Fourier transform of a small section of the input image. The direction of the frequency intensity with the highest value is used as the direction of the pattern to calculate the line toward the center. Finally, by counting the intersections of lines in each section of the image divided into a grid, the section with the highest number of intersections is designated the nucleus of the scale. To enhance the precision of center calculation, the final nucleus is determined by

employing the same process, during which the calculated nucleus section of the scale is divided into grids.

2) EXPANDING THE RESTING ZONE IMAGE INTO POLAR COORDINATES AND CORRECTION

To count the resting zones, a rectangular image is generated by transforming the resting zone image into polar coordinates centered around the determined nucleus. This transformation maintains the circular scales with the characteristics of the resting zones. At this point, the resting zones detected by PSPNet are incomplete, missing lines and scale outlines. Since this group of lines affects the resting zone counts, a correction is applied as shown in Fig. 9.

First, to ensure a reliable count of resting zones, the detected areas where the lines are interrupted are filled in or complemented. Specifically, Fig. 9 shows an overview of the method for interpolating the resting zones. The endpoints of all missing resting zones are detected, and a set of coordinates for each end is created. The distance is calculated by Equation (4) using Euclidean distance. The angle is calculated by Equations (5) and (6), with the lower end point of the connection source as $P_a = (a_x, a_y)$ and the upper end point of the connection destination area as $P_b = (b_x, b_y)$.

$$\text{distance} = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2} \tag{4}$$

$$\cos\theta = \frac{a_x b_x + b_y^2}{\sqrt{a_x^2 + b_y^2}\sqrt{b_x^2 + b_y^2}} \tag{5}$$

$$\theta = \arccos(\cos\theta) \tag{6}$$

θ is obtained from the inner product of $\overrightarrow{P_aP_c}$ and $\overrightarrow{P_aP_b}$, where $P_c = (a_x, b_y)$ is the point consisting of the x-coordinate of point P_a and the y-coordinate of point P_b. This process makes it possible to supplement the missing parts of the resting zones interrupted during detection.

Next, given that salmon scales exhibit false annual ring lines close to the nucleus (circuli 15-17 from the nucleus) that resemble resting zones in shape, a masking process is implemented on the left side of the expanded image (in the direction of the nucleus) to remove the false annual ring lines as shown in the center of Fig. 9. Finally, to remove the outer contour lines of the scales, a mask image of the original scale image is generated [16], and the lines in the expanded image of the resting zone at the position corresponding to the contour lines in that image are removed.

3) AGE ESTIMATION BY COUNTING RESTING ZONES

Finally, age is assessed from the polar coordinate image supplemented by the resting zones by measuring the number of resting zones from the left side of the image. Since the number of resting zones differs depending on the location of the measurement, the number of resting zones is defined as the largest number of resting zones after counting the number of resting zones in the X-axis direction at all Y-coordinates. Once the number of resting zones has been determined, one is then added to that number based on the definition of age assessment. The result is the estimated age of the salmon. This process enables the automation of salmon age assessment using scales by using information technology based on the age assessment process used by experts at the Fisheries Research Institute.

IV. PRELIMINARY EXPERIMENTS

A poor performance of YOLOv5 in detecting scaled regions and clustering by k-means is likely to adversely affect the performance evaluation of resting zone detection and age assessment. Here, we discuss the results of preliminary experiments regarding performance evaluation. Specifically, we first checked the performance of YOLOv5 in detecting scaled regions using the following IoU, an object detection evaluation index. λ_1, λ_2, and λ_3 were set to 0.5, 1.0, and 0.05 in the training phase based on [27] in the literature. In addition, the number of epochs was set to 100 and the batch size to 4. After manually setting the correct scale regions, the IoU score was very high at 97% when the dataset of 3,480 images was randomly divided into training and test data at a ratio of 9 to 1 in Table 1. Note that the dataset in Table 1 is used as test data in the next section.

$$\text{IoU} = \frac{\text{Product set of correct and detected regions}}{\text{Union set of the correct region and the detected region}}. \tag{7}$$

On the other hand, the correctness rate of clustering (the percentage of correct labels clustered vertically (1)-(6) and horizontally (1)-(10)) was calculated for k-means using a dataset of 3,480 data sets, and a correctness rate of 98% was obtained. The number of iteration was set to 100. Since the method of scale placement is determined by the operational manual of the Fisheries Research Institute and the replica samples are checked by several institute staff members, it can be said that any discrepancy in placement has little impact on labeling by k-means. The above results indicate that preprocessing has little negative impact on resting zone detection or age assessment, as both tasks can be accomplished with high performance.

V. EXPERIMENTS

We conducted experiments that assess the age of scales in order to test the effectiveness of the proposed method. First, the experimental conditions are described in Section V-A. The experimental results are then presented and discussed in Section V-B.

A. EXPERIMENTAL CONDITIONS

In the experiments, we confirmed the effectiveness of the proposed method using salmon scale images. First, the scanned resolution of the replica sample images was 27,716 × 13.010. Two or three scales per salmon were crimped onto replica samples. Approximately 60 scales were crimped onto each replica sample. Test data were then constructed of all salmon scale replica sample images

TABLE 1. Numbers of individual salmon and scales of each age of salmon along the Tokoro coast in 2018 (Test data: 1,076 images).

Salmon age	3	4	5	6	7
No. of scale images	63	804	191	15	3

TABLE 2. Numbers of individual salmon and scales of each age of salmon along the Tokoro coast (Training data: 71 images).

Salmon age	4	5	6
No. of scale images	8	59	4

collected on the Tokoro coast of Hokkaido, Japan, in 2018 (Table 1). The training data (Table 2) consisted of 71 scale replica sample images collected in years other than 2018. These 71 images were selected by expert staff from candidate scales of salmon aged 3 to 7 years; the candidate scales were selected by the Fisheries Research Institute and had little damage or shape deformation. Finally, images of scales from salmon aged 4-6 years with little damage and clean shapes were constructed as the training dataset. Note that the number of 3- to 5-year-old salmon migrating to Hokkaido is high, while the number of 6- to 7-year-old salmon is low (Table 1). Therefore, it is of utmost importance to estimate the age of 3- to 5-year-old salmon with high accuracy. Ground truth was considered the consensus result of three expert staff members' age assessments of the scales. Two expert staff members set the correct resting zone lines.

To quantitatively and qualitatively evaluate the effectiveness of the age assessment, we compared the proposed method with the eight methods described below.

Comparative method 1 (Comp. 1) [16]

Overview: Comp. 1 is a baseline method based on image processing for age assessment, which we reported recently [16]. First, the method calculates the nucleus of the scale from the 2D Fourier transform. Next, a rectangular image is generated by expanding the salmon scale image into polar coordinates based on the determined nucleus. Finally, age is assessed by summarizing the closed annuli regions (resting zones) of the scale, which determine the annuli-frequency distribution.

Comparative method 2 (Comp. 2)

Overview: This is an age assessment method based on fully convolutional network (FCN)-8 [43], a semantic segmentation model. Specifically, the network consists of FCNs, constructed entirely with convolutional layers except for the pooling layer and activation functions, which are upscaled by a factor of 8 in the training phase to obtain a resolution of 224 × 224 for the output. This method inputs a test image, detects resting zones, and then applies an age assessment method similar to that of the proposed method.

Comparative method 3 (Comp. 3)

Overview: This is an age assessment method based on FCN-16 [43], a semantic segmentation model. Specifically, the network consists of FCNs, constructed entirely with convolutional layers except for the pooling layer and activation functions, which are upscaled by a factor of 16 in the training phase to obtain a resolution of 224 × 224 for the output. This method inputs a test image, detects resting zones, and then applies an age assessment method similar to that of the proposed method.

Comparative method 4 (Comp. 4)

Overview: This is an age assessment method that uses U-Net [34], a semantic segmentation model. Specifically, U-Net consists of a network of encoders and decoders that compress and upsample feature maps. In addition, the encoders and decoders at each layer are concatenated to complement the positional information on the feature map. This method inputs a test image, detects resting zones, and then applies an age assessment method similar to that of the proposed method.

Comparative method 5 (Comp. 5)

Overview: This method detects resting zones from scale images using PSPNet and applies the same age assessment method as the proposed method without complementing missing lines.

Comparative method 6 (Comp. 6)

Overview: This method detects resting zones from scale images using PSPNet and employs the same age assessment approach as the proposed method, without the need to remove lines based on the region of the fake annual ring or the maximum radius of scale.

Comparative method 7 (Comp. 7)

Overview: This is an age assessment method by visual geometry group (VGG)-16 [44], which is a classification model fine-tuned for analyzing scale images.

Comparative method 8 (Comp. 8)

Overview: This is an age assessment method [18] utilizing EfficientNet [45] as a state-of-the-art classification model, fine-tuned with our scale image dataset. The age assessment results are obtained by inputting a test scale image.

Comp. 2 - Comp. 6 were employed as ablation studies. Specifically, it is possible to show the effectiveness of each process in the proposed method as a combination and as an overall configuration by comparing the performance of each combination with a method in which a part of the proposed method is replaced by another method. Comp. 7 and Comp. 8 are different approaches from the proposed method in that they apply a CNN-based classification task to age assessment. The recent decade has seen explosive development in deep

TABLE 3. Deep learning framework, hardware requirements, and training environment in implementation.

Details	value
Operating System	Ubuntu 18.04 LTS
Central Processing Unit	Intel(R) Xeon(R) W-2155 CPU @ 3.30 GHz
Memory	64 GB
Graphics Processing Unit	NVIDIA Quadro P6000
CUDA version	11.6
Python version	3.6.10
Deep learning framework	Pytorch 1.6.0a0+9907a3e Tensorflow 1.15.0+nv
OpenCV	4.1.2

TABLE 4. Parameters used in the proposed method.

Details	Parameter	Value
Parameters of YOLOv5	Epoch	100
	Learning rate	0.01
	Optimizer	Adam [46]
Parameters of PSPNet	No. of layer in the pyramid pooling module:M	4
	Batch size	4
	Degree of learning rate decay (gamma)	0.1
	Number of steps of learning rate decay	10
	Epoch	100
	Optimizer	AdaDelta [47]
	Loss function	Cross-entropy loss
Parameters of complementation of broken lines	Distance	160 px
	Angle	20°
Parameters for removing fake annual ring	Contour deletion range	4 px
	Deletion range percentage from nucleus	20

learning, especially in the performance of CNNs. Since most image recognition tasks currently employ CNN-based image classification approaches, it is necessary to confirm the detection performance of CNNs in the age assessment task from replica sample images, which we address in this paper.

Comps. 2 - 4 use VGG-16 [44] as their backbone. The input data format for PSPNet was 130 images, with 10 vertical segments and 13 horizontal segments per single salmon scale. Note that the number of background patch images in the scale image was set to equal the number of scale patch images in order to balance the data volume. Specifically, 3,335 patch images of the scale region part and 3,335 patch images of the background of the scale image were constructed from 71 training data images. All methods were implemented with the equipment and software environment shown in Table 3. The parameters of the proposed method are shown in Table 4. The proposed method was fully tuned with the training dataset of scales, as the number of training epochs was 100. Note that the learning rate was set to 0.01. The model parameters were taken as the values at which each method was most accurate. The scale regions were manually designated to ensure the accurate selection of regions for generating the training data. To detect a scale region, fine-tuning was performed on the YOLOv5 model, which had been previously trained on the common objects in context (COCO) dataset [48], using 3,480 salmon scale images.

The quantitative evaluation of age assessment used mean absolute error (MAE) [49], precision, and accuracy to elucidate differences between actual age and estimated age.

$$\text{MAE} = \frac{1}{n}\sum_{k=1}^{n}|f_k - y_k|, \tag{8}$$

$$\text{Precision} = \frac{\text{No. of images from which age was correctly estimated}}{\text{No. of scale images with estimated ages}}, \tag{9}$$

$$\text{Accuracy} = \frac{\text{No. of images from which age was correctly estimated}}{\text{No. of all scale images}}, \tag{10}$$

where f_k and y_k are the estimated and actual ages of a scale image k, respectively. n is the number of scale images.

Moreover, in this experiment, from the viewpoint of improving the efficiency of age assessment work, we used the weighted average value for the age group with the largest number of samples. The average in the Table 5

TABLE 5. Quantitative comparison of age assessment performance between the proposed method and comparative methods.

	Proposed Method			Comp. 1			Comp. 2		
Age	MAE	Precision	Accuracy	MAE	Precision	Accuracy	MAE	Precision	Accuracy
3	0.492	0.178	0.619	1.540	0.000	0.000	0.206	0.113	0.825
4	**0.458**	0.824	**0.570**	0.617	0.675	0.400	0.570	0.744	0.445
5	0.639	0.327	**0.435**	**0.623**	0.380	0.382	0.906	0.440	0.267
6	1.133	0.063	0.200	1.267	0.000	0.000	1.867	0.053	0.067
7	2.667	0.000	0.000	3.000	0.000	0.000	3.000	0.000	0.000
Average	**0.507**	**0.685**	**0.542**	0.688	0.527	0.365	0.633	0.642	0.429

	Comp. 3			Comp. 4			Comp. 5		
Age	MAE	Precision	Accuracy	MAE	Precision	Accuracy	MAE	Precision	Accuracy
3	**0.048**	0.066	**0.952**	0.302	0.130	0.762	0.142	0.108	0.873
4	0.902	0.516	0.100	0.514	0.776	0.504	0.600	0.727	0.404
5	1.576	**0.538**	0.037	0.848	0.363	0.298	0.921	0.427	0.246
6	2.533	0.000	0.000	1.733	0.037	0.067	1.933	**0.167**	0.133
7	4.000	0.000	0.000	2.000	0.000	0.000	3.333	0.000	0.000
Average	1.003	0.485	0.137	0.582	0.652	0.475	0.656	0.628	0.399

	Comp. 6			Comp. 7			Comp. 8		
Age	MAE	Precision	Accuracy	MAE	Precision	Accuracy	MAE	Precision	Accuracy
3	0.556	**0.201**	0.587	0.762	0.132	0.571	0.110	0.087	0.381
4	0.517	**0.827**	0.536	0.947	0.825	0.333	0.909	0.712	0.252
5	0.628	0.283	**0.435**	0.984	0.211	0.157	0.912	0.169	0.293
6	1.000	0.051	0.267	**0.467**	0.030	**0.667**	1.600	0.011	0.133
7	2.667	0.000	0.000	**1.333**	0.000	0.000	2.333	0.000	0.000
Average	0.552	0.681	0.516	0.937	0.662	0.320	0.934	0.567	0.265

TABLE 6. Comparison of parameters and floating-point operations (FLOPs) between the proposed method and other methods on the salmon scales dataset. Note that M and G indicate mega and giga, respectively.

	Parameters (M)	FLOPs (G)
Proposed Method	103.7	235.5
Comp. 1	**86.2**	**204.6**
Comp. 2	100.9	212.5
Comp. 3	254.0	228.2
Comp. 4	123.3	226.3
Comp. 5	103.7	235.5
Comp. 6	103.7	235.5
Comp. 7	220.5	224.8
Comp. 8	86.5	205.2

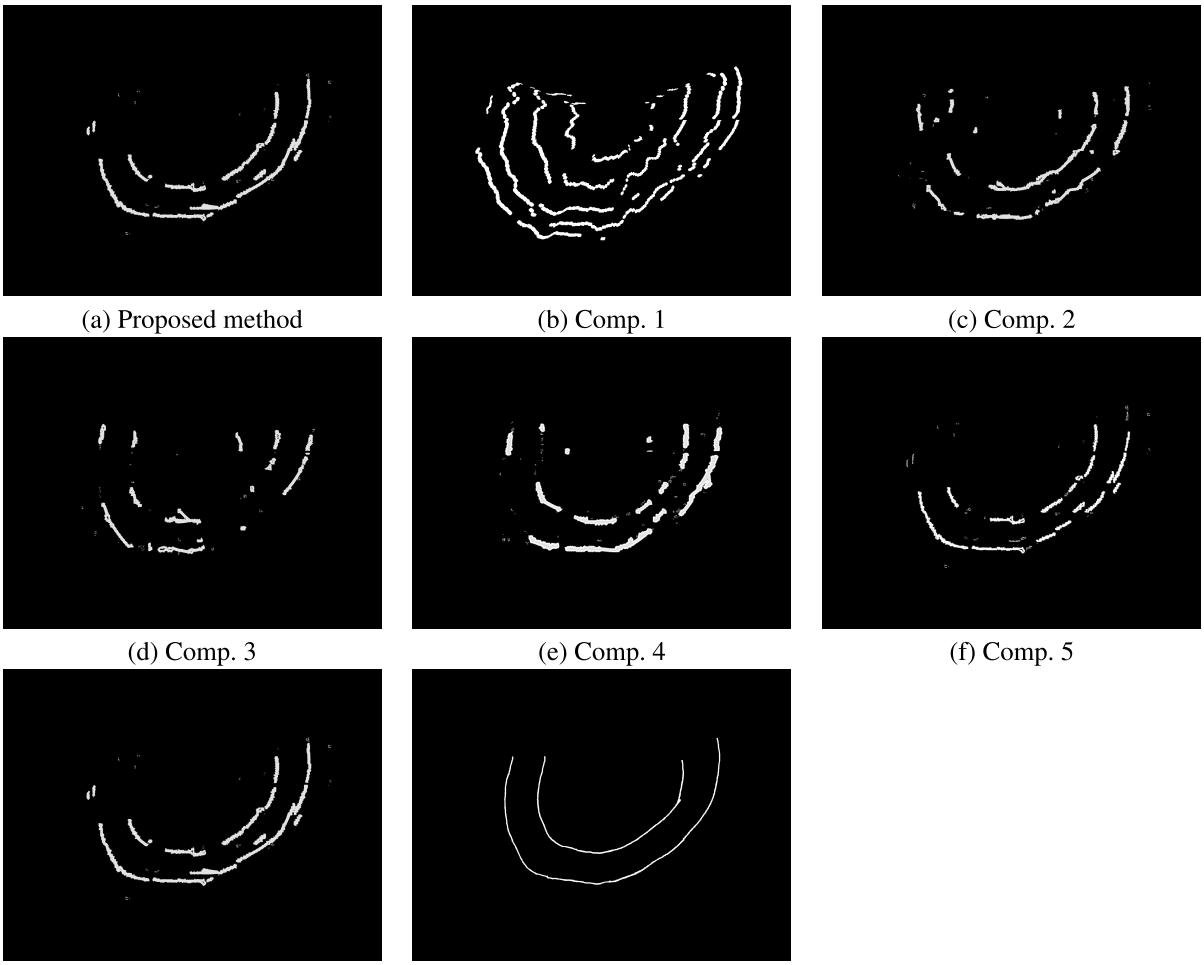
is the weighted average of the values for each age. The formula for the weighted average (WA) MAE is as follows.

$$\begin{aligned}\text{WA MAE} = (&\text{MAE}_{\text{3-years-old}} * 63 + \text{MAE}_{\text{4-years-old}} * 804 \\ &+ \text{MAE}_{\text{5-years-old}} * 191 + \text{MAE}_{\text{6-years-old}} * 15 \\ &+ \text{MAE}_{\text{7-years-old}} * 3)/1,076. \end{aligned} \tag{11}$$

The total number of test samples is 1,076 (= 63 + 804 + 191 + 15 + 3). The same formula is used for the weighted average of Precision and Accuracy.

B. PERFORMANCE EVALUATIONS

Table 5 shows the age assessment accuracy of each method. The proposed method has the highest performance on average for each evaluation metric. Versus Comp. 1, the proposed method performs significantly better than the baseline method age assessment method. In particular, the proposed method can contribute more to the efficiency of age assessment operations than the conventional methods, since the MAE of 4-year-olds, the group with the largest number of samples, is low. Since Comp. 1 is based on image processing, the shape and spacing of scale patterns and the unevenness of pixel shading affect the performance of resting zone detection. On the other hand, the proposed method uses CNNs to train patch images of scale patterns, so it is possible to detect resting zones by considering various scale features. Versus Comps. 2-4, it was also demonstrated that the proposed method's semantic segmentation based on PSPNet is the most suitable for detecting resting zones. Specifically, the feature maps of multiple scales were used to train features of scale patterns with different scales, which improved performance. Moreover, the better performance of the proposed method versus both Comp. 5 and Comp. 6 indicates that the correction and removal of lines based on the region of fake annual rings and the maximum radius of the scale contribute to consistent age assessment across all ages. Additionally, the proposed method outperforms Comp. 7, which exhibited subpar assessment performance for ages 4 and 5 despite a large number of samples. This highlights the validity of the proposed method, which aligns more closely with the actual experts' age assessment decision process. Finally, Comp. 8 was found to perform better than the proposed method, especially for MAE at age 3. However, the proposed method showed superior overall



(a) Proposed method (b) Comp. 1 (c) Comp. 2

(d) Comp. 3 (e) Comp. 4 (f) Comp. 5

(g) Comp. 6 (h) Ground truth

FIGURE 10. Examples of salmon scales whose resting zones were detected by the proposed method and Comps. 1-5. Note that the detection result of Comp. 6 is the same as that of the proposed method.

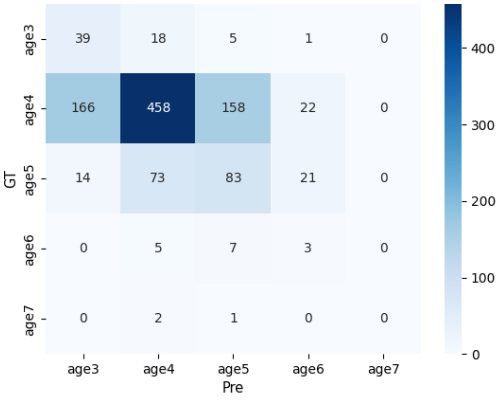


FIGURE 11. Confusion matrix of age assessment by the proposed method.

performance. The effectiveness of the proposed method can be shown as follows. First, by comparing the proposed method with the baseline (Comp. 1) and the state-of-the-art (Comp. 8), which are based on image processing and deep learning, respectively, the proposed method was shown to have the best age assessment performance of all the methods. Next, we verified that the PSPNet-based age assessment method has the best resting zone detection among all semantic segmentation models, such as FCN-8 (Comp. 2), FCN-16 (Comp. 3), and U-Net (Comp. 4). Then, by comparing the proposed method and Comp. 5 and Comp. 6, it was confirmed that complementing missing lines and removing lines based on the region of the fake annual rings and the maximum scale radius greatly improve the accuracy of age counts. Finally, the effectiveness of the age assessment approach utilizing both deep learning and image processing is confirmed since the performance of the proposed method is better than those of Comp. 7 and Comp. 8. Furthermore, since the proposed method also detects the resting zone, which is the basis for age assessment, it is expected to be introduced into age assessment work at Japanese Fisheries Research Institutes. Note that the 0% precision and accuracy of all methods for 7-year-old salmon may have occurred because the scales of older salmon are more damaged. Therefore, the age of samples with severe scale loss, such as those from 7-year-old salmon, should be assessed by staff for the time being. In conclusion, the effectiveness of the proposed method was quantitatively confirmed.

TABLE 7. Processing time of age assessment in training and testing. All numbers in the table are in seconds, and the tests represent the average time to process one scale.

	YOLOv5	Semantic segmentation model	Classification model	Total training time (100 Epoch)
Proposed method	2023.979	5144.482	-	7168.462
Comp. 1	2023.979	-	-	**2023.979**
Comp. 2	2023.979	3485.993	-	5509.973
Comp. 3	2023.979	3410.319	-	5434.298
Comp. 4	2023.979	5117.438	-	7141.417
Comp. 5	2023.979	5144.482	-	7168.462
Comp. 6	2023.979	5144.482	-	7168.462
Comp. 7	2023.979	-	523.067	2547.047
Comp. 8	2023.979	-	977.719	3001.698

	Scanning of replica (3-A)	Scale based on detection YOLOv5 (3-B)	Scale region clustering by k-means (3-B)	Resting zone detection (3-C)	Determining the nucleus of a scale (3-D-1)	Complementation of broken lines (3-D-2)	Removal lines based on the region of the fake annual ring or the maximum radius of scale (3-D-2)	Age estimation by counting resting zones (3-D-3)	Total test time
Proposed method	5.283	0.001	0.021	1.172	3.771	0.099	2.32	0.01	12.677
Comp. 2	5.283	0.001	0.021	0.871	3.771	0.099	2.32	0.01	12.376
Comp. 3	5.283	0.001	0.021	0.833	3.771	0.099	2.32	0.01	12.338
Comp. 4	5.283	0.001	0.021	1.018	3.771	0.099	2.32	0.01	12.523
Comp. 5	5.283	0.001	0.021	1.172	3.771	-	2.32	0.01	12.578
Comp. 6	5.283	0.001	0.021	1.172	3.771	0.099	-	0.01	10.357

	Scanning of replica (3-A)	Scale based on detection YOLOv5 (3-B)	Scale region clustering by k-means (3-B)	Determining the nucleus of a scale (3-D-1)	Age assessment based on the annuli-frequency distribution	Total test time
Comp. 1	5.283	0.001	0.021	3.771	45.602	54.68

	Scanning of replica (3-A)	Scale based on detection YOLOv5 (3-B)	Scale region clustering by k-means (3-B)	Age assessment utilizing classification model	Total test time
Comp. 7	5.283	0.001	0.021	0.161	5.468
Comp. 8	5.283	0.001	0.021	0.111	**5.417**

Then, we reported and evaluated the computational complexity, number of parameters, and processing time for all methods. In evaluating neural network-based methods and tasks, it is expected to evaluate the computational cost in addition to the evaluation of accuracy [50]. Specifically, since convolutional neural networks and segmentation models

are computationally expensive, it is essential to check the processing speed and the number of computational resources. Floating point operations (FLOPs) metrics commonly used to calculate the computational complexity of deep learning models [51], [52]. Tables 6 and 7 show the calculated values for the scale image data set without scanner preprocessing. Table 6 shows that Comp. 1 has fewer parameters and FLOPs because it does not use neural networks for resting zone detection and age assessment. On the other hand, methods using semantic segmentation and classification models tend to have exceptionally high FLOPs. Table 7 shows the processing times for training and testing in the Table 3 environment. The table shows that Comp. 1 has the shortest training time, and the proposed method has the longest training time. However, the training cost must be high enough to obtain sufficient age assessment performance. Comp. 8 has the shortest test time, and Comp. 1 has the longest. Comp. 1, in particular, had the longest test time because it is image processing based and requires much writing and reading of image files. Comp. 7 and Comp. 8 are classification models and generally require shorter test processing times. The processing time for age assessment of one scale in the proposed method is about 12 seconds. The proposed method's processing time is short enough because at the Fisheries Research Institute, staff perform the age assessment operations on approximately 20,000 salmon over a period of several months. If the proposed system is implemented in Fisheries Research Institutes, it would be possible to assess the age of all scales in about three days (20,000 salmon $\times$ 12 seconds $\fallingdotseq$ 67 hours). The above results showed that the proposed method is sufficiently valuable, considering both accuracy and processing time.

Next, we qualitatively evaluated the results of resting zone detection. Fig. 10 shows some of the results of each method's resting zone detection performance. The white lines indicate the regions of the resting zones detected by each method. Fig. 10 (h) shows the region of the correct resting zones. First, Comp. 1, which is based on simple image processing based on the shading of the scale pattern, resulting in the loss of curves and in an excessive number of lines. Next, it is clear from the figure that Comps. 2, 3, and 4, which are based on semantic segmentation other than PSPNet, have more missing lines than the proposed method. In particular, they fail to detect the lower left line. Through a comparison of the proposed method with Comp. 5 (which involves restoring the image from a corrected polar transformed image), it has been verified that the correction process in the proposed method effectively restores the missing lines within the resting zone. Consequently, these findings collectively demonstrate the superior stability of the proposed method in detecting resting zones.

The confusion matrix of age assessment by the proposed method is shown in Fig. 11. It can be confirmed that many samples mistakenly assess a 4-year-old salmon as 3 or 5 years old. However, this problem can be solved by improving the training data of PSPNet in the future. The MAE of the proposed model is approximately 0.5, which means that the age of the scales can be assessed with an error of less than one year. This indicates that the proposed method is highly effective for practical applications.

VI. CONCLUSION

In this paper, we propose a method of salmon age assessment based on neural network and image processing for data mining scale replicas and resting zones in Fisheries Research Institutes in order to improve the efficiency of salmon age assessment work. The proposed method first utilized a high-resolution scanner, the object detection method YOLOv5, and a clustering method to automatically extract scale region images from scale replica samples. Then, to detect resting zones and perform age assessment, we employed PSPNet, conducted line correction, and eliminated lines based on the region of fake annual rings as well as the maximum radius of a scale. In counting age based on resting zones, the age assessment process was applied, including the identification of resting zones by the staff and the method of counting age from the circular structure. Our experimental results showed that our method assessed age more accurately than conventional and deep learning-based methods, and that it improved age assessment performance significantly. In addition, the proposed method enabled faster age assessment than the age assessment work done by Fisheries Research Institute staff.

Although we used deep learning to detect resting zones in this study, the amount of label data used was limited because the sample was created with the cooperation of the staff of the Fisheries Research Institute. However, if the proposed system is tested at actual test sites, and a system for generating and sharing label data among multiple test sites is established, then the system's ability to detect resting zones will inevitably improve and the method will be able to be put into practical use. Many other fish species are not being fished, and the process of assessing the age of fish with scales is almost the same as that for salmon. Therefore, we expect the proposed method to be useful in improving the efficiency of age assessment in these other species as well. Moreover, it is known that the width between resting zones represents the growth of a salmon. Therefore, we plan to use the proposed method to analyze poor fishing conditions by aggregating age data and analyzing scale shapes, and will collaboratively analyze the data taking meteorological conditions and bathymetry data into account, as in [53], [54], [55], [56], and [57].

Finally, because the fisheries industry is operated based on the knowledge and long years of experience of Fisheries Research Institutes and fishers, few studies have yet applied information technology and information science to this industry [58]. However, it is a meaningful research direction to use information technology to model all tasks related to the fisheries industry and to streamline the sustainable management of marine resources.

VII. DATA AVAILABILITY

The proposed method implementation, trained networks, and supplementary material are available at https://github.com/Choke222/AgeAssessmentOfScales.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Fisheries Research Department, Salmon and Freshwater Fisheries Research Institute, Hokkaido Research Organization, the Kitami Salmon Enhancement Program Association, and the Tokoro Fisheries Cooperative Association for their cooperation in their joint research.

REFERENCES

[1] R. M. Peterman, "Form of random variation in salmon smolt-to-adult relations and its influence on production estimates," *Can. J. Fisheries Aquatic Sci.*, vol. 38, no. 9, pp. 1113–1119, Sep. 1981.

[2] R. C. Bocking and R. M. Peterman, "Preseason forecasts of sockeye salmon (*Oncorhynchus nerka*): Comparison of methods and economic considerations," *Can. J. Fisheries Aquatic Sci.*, vol. 45, no. 8, pp. 1346–1354, Aug. 1988, doi: 10.1139/f88-158.

[3] C. C. Wood, D. T. Rutherford, D. Peacock, S. Cox-Rogers, and L. Jantz, "Assessment of recruitment forecasting methods for major sockeye and pink salmon stocks in Northern British Columbia," Can. Tech. Rep. Fisheries Aquatic Sci., Nanaimo, BC, Canada, Tech. Rep. 2187, 1997, pp. 1–85.

[4] N. J. Mantua, S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, "A Pacific interdecadal climate oscillation with impacts on salmon production," *Bull. Amer. Meteorol. Soc.*, vol. 78, no. 6, pp. 1069–1079, Jun. 1997.

[5] K. Dahl, *The Age and Growth of Salmon and Trout in Norway as Shown by Their Scales*. London, U.K.: Salmon and Trout Association, 1910.

[6] W. Templeman and H. J. Squires, "Relationship of otolith lengths and weights in the haddock *Melanogrammus aeglefinus* (L.) to the rate of growth of the fish," *J. Fisheries Res. Board Canada*, vol. 13, no. 4, pp. 467–487, Apr. 1956.

[7] G. W. Boehlert, "Using objective criteria and multiple regression models for age determination in fishes," *Fishery Bull.*, vol. 83, no. 2, pp. 103–117, 1985.

[8] W. G. Pearcy, "Ocean ecology of North Pacific Salmonids," in *Books in Recruitment Fishery Oceanography*. Seattle, WA, USA: Sea Grant Program, 1992.

[9] R. J. Beamish and D. R. Bouillon, "Pacific salmon production trends in relation to climate," *Can. J. Fisheries Aquatic Sci.*, vol. 50, no. 5, pp. 1002–1016, May 1993.

[10] R. J. Beamish, B. E. Riddell, C. M. Neville, B. L. Thomson, and Z. Zhang, "Declines in chinook salmon catches in the strait of Georgia in relation to shifts in the marine environment," *Fisheries Oceanogr.*, vol. 4, no. 3, pp. 243–256, Sep. 1995.

[11] R. C. Francis, S. R. Hare, A. B. Hollowed, and W. S. Wooster, "Effects of interdecadal climate variability on the oceanic ecosystems of the NE Pacific," *Fisheries Oceanogr.*, vol. 7, no. 1, pp. 1–21, Apr. 1998.

[12] M. J. Kishi, M. Kaeriyama, H. Ueno, and Y. Kamezawa, "The effect of climate change on the growth of Japanese chum salmon (*Oncorhynchus keta*) using a bioenergetics model coupled with a three-dimensional lower trophic ecosystem model (NEMURO)," *Deep Sea Res. II, Topical Stud. Oceanogr.*, vol. 57, nos. 13–14, pp. 1257–1265, Jul. 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967064510000160

[13] H. Mitani, M. Igarashi, and O. Watarai, "A study of the age-determination of chum salmon using personal computer for scale reading," *J. Fac. Mar. Sci. Technol. Tokai Univ.*, vol. 26, pp. 145–154, Jan. 1988.

[14] Y. Endo, O. Watarai, and M. Igarashi, "Study on development of age analyzer based on scale patterns using CCD image sensor camera," *J. Fac. Mar. Sci. Technol. Tokai Univ.*, vol. 44, pp. 1–10, 1988.

[15] M. Nishiyama, G. Suzuki, and H. Shioya, "Detection of resting zones based on PSPNet in salmon scale images," in *Proc. IEEE 11th Global Conf. Consum. Electron. (GCCE)*, Oct. 2022, pp. 910–911.

[16] M. Yoshida, H. Shioya, Y. Miyakoshi, F. Yamaguchi, and H. Urabe, "Image processing for the age determination based on the annuli of chum salmon scale," in *Proc. 3rd NPAFC-IYS Virtual Workshop Linkages Pacific Salmon Prod. Environ. Changes*, 2021, p. 85.

[17] R. Hoson, H. Shioya, Y. Miyakoshi, F. Yamaguchi, and H. Urabe, "Neural-network based prediction of chum salmon age by the scale images," in *Proc. 3rd NPAFC-IYS Virtual Workshop Linkages Pacific Salmon Prod. Environ. Changes*, 2021, p. 84.

[18] R. Vabø, E. Moen, S. Smoliński, Å. Husebø, N. O. Handegard, and K. Malde, "Automatic interpretation of salmon scales using deep learning," *Ecol. Informat.*, vol. 63, Jul. 2021, Art. no. 101322. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574954121001138

[19] J. J. Lalanne and G. Safsten, "Age determination from scales of chum salmon (*Oncorhynchus keta*)," *J. Fisheries Res. Board Canada*, vol. 26, no. 3, pp. 671–681, Mar. 1969, doi: 10.1139/f69-060.

[20] S. Katayama, M. Yamamoto, and S. Gorie, "Age compositions of flatfish stocks as determined by a new otolithometric method, its application in the estimation of growth, spawning potential and fisheries management," *J. Sea Res.*, vol. 64, no. 4, pp. 451–456, Nov. 2010.

[21] K. Tetsuo, "Biology of chum salmon, *Oncorhynchus keta* (Walbaum) by the growth formula of scale," *Sci. Rep. Hokkaido Salmon Hatchery*, vol. 16, pp. 1–102, Jan. 1961. [Online]. Available: https://salmon.fra.affrc.go.jp/english/eng2_SRHSH2.html

[22] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big data*, vol. 3, no. 1, pp. 1–40, May 2016.

[23] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022.

[24] J. MacQueen, L. M. Le Cam, and J. Neyman, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Oakland, CA, USA, 1967, vol. 1, no. 14, pp. 281–297.

[25] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6230–6239.

[26] R. Anas, "Red salmon scale studies," in *Proc. Annu. Rep. Int. North Pacific Fisheries Commission*, 1961, pp. 114–116.

[27] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[28] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010.

[29] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3213–3223.

[30] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[31] G. Roberts, S. Y. Haile, R. Sainju, D. J. Edwards, B. Hutchinson, and Y. Zhu, "Deep learning for semantic segmentation of defects in advanced STEM images of steels," *Sci. Rep.*, vol. 9, no. 1, pp. 1–12, Sep. 2019.

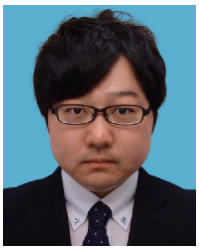
[32] L. Ahmed, M. M. Iqbal, H. Aldabbas, S. Khalid, Y. Saleem, and S. Saeed, "Images data practices for semantic segmentation of breast cancer using deep neural network," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 11, pp. 15227–15243, Nov. 2023.

[33] M. Saha and C. Chakraborty, "Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2189–2200, May 2018.

[34] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*. Springer, 2015.

[35] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected CRFs," 2014, *arXiv:1412.7062*.

[36] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. H. S. Torr, "Conditional random fields as recurrent neural networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1529–1537.
[37] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
[38] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2006, pp. 2169–2178.
[39] W. Liu, A. Rabinovich, and A. C. Berg, "ParseNet: Looking wider to see better," 2015, *arXiv:1506.04579*.
[40] 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.
[41] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*.
[42] P. Hough, "A method and means for recognizing complex patterns," U.S. Patent 3 069 654, Dec. 18, 1962.
[43] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
[44] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[45] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.
[46] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[47] M. D. Zeiler, "ADADELTA: An adaptive learning rate method," 2012, *arXiv:1212.5701*.
[48] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Computer Vision (ECCV)*, 2014, pp. 740–755.
[49] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, no. 1, pp. 79–82, 2005.
[50] G. Dong, Y. Yan, C. Shen, and H. Wang, "Real-time high-performance semantic image segmentation of urban street scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3258–3274, Jun. 2021.
[51] J. Hu, L. Li, Y. Lin, F. Wu, and J. Zhao, "A comparison and strategy of semantic segmentation on remote sensing images," in *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*, Y. Liu, L. Wang, L. Zhao, and Z. Yu, Eds. Cham, Switzerland: Springer, 2020, pp. 21–29.
[52] S. Wang, X. Hou, and X. Zhao, "Automatic building extraction from high-resolution aerial imagery via fully convolutional encoder-decoder network with non-local block," *IEEE Access*, vol. 8, pp. 7313–7322, 2020.
[53] M. Kaeriyama, A. Yatsu, M. Noto, and S. Saitoh, "Spatial and temporal changes in the growth patterns and survival of Hokkaido chum salmon populations in 1970–2001," *North Pacific Anadromous Fish Commission Bull.*, vol. 4, pp. 251–256, Jan. 1970.
[54] S. Zhou, "Application of artificial neural networks for forecasting salmon escapement," *North Amer. J. Fisheries Manage.*, vol. 23, no. 1, pp. 48–59, Feb. 2003.
[55] C. Waters, T. Miller, E. Fergusson, D. Oxman, B. Agler, and E. Farley Jr., "Winter condition and trophic status of Pacific salmon in the Gulf of Alaska," *Tech. Rep.*, no. 18, pp. 96–101, Apr. 2022.
[56] A. V. Zavolokin, V. V. Kulik, I. I. Glebov, E. N. Dubovets, and Y. N. Khokhlov, "Dynamics of body size, age, and annual growth rate of Anadyr chum salmon *Oncorhynchus keta* in 1962–2010," *J. Ichthyol.*, vol. 52, no. 3, pp. 207–225, Apr. 2012.
[57] S. Urawa, T. Beacham, S. Sato, T. Kaga, B. Agler, R. Josephson, and M.-A. Fukuwaka, "Stock-specific abundance of chum salmon in the central Gulf of Alaska during winter," *North Pacific Anadromous Fish Commission Bull.*, vol. 6, no. 1, pp. 153–160, Dec. 2016.
[58] I. Suryanarayana, A. Braibanti, R. Sambasiva Rao, V. A. Ramam, D. Sudarsan, and G. N. Rao, "Neural networks in fisheries research," *Fisheries Res.*, vol. 92, nos. 2–3, pp. 115–139, Aug. 2008.



GENKI SUZUKI (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electronics and information engineering from Hokkaido University, Japan, in 2017, 2019, and 2021, respectively. In 2021, he joined the Division of Information and Electronic Engineering, Muroran Institute of Technology, as an Assistant Professor. His research interests include mathematical science and information engineering. He is a member of the IEICE.

MIKIYASU NISHIYAMA (Student Member, IEEE) received the B.E. degree in information and electronic engineering from Muroran Institute of Technology, Japan, in 2021, where he is currently pursuing the M.E. degree with the Division of Information and Electronic Engineering. His research interest includes image analysis.

RYOMA HOSON received the B.E. degree in information and electronic engineering from Muroran Institute of Technology, Japan, in 2018, and the M.E. degree from the Division of Information and Electronic Engineering, Muroran Institute of Technology, in 2020.

KATSUNOBU YOSHIDA received the B.E. degree in information and electronic engineering from Muroran Institute of Technology, Japan, in 2018, and the M.E. degree from the Division of Information and Electronic Engineering, Muroran Institute of Technology, in 2020.

HIROYUKI SHIOYA received the B.S. degree in mathematics and the M.E. degree in electronics and information engineering from Hokkaido University, Japan, in 1990 and 1992, respectively, and the Ph.D. degree from the Graduate School of Information Systems and Engineering, Hokkaido University, in 1995. He is currently a Professor with Muroran Institute of Technology. His research interests include mathematical science and information engineering. He is a member of OSA, IPSJ, and IEICE.

KAZUTAKA SHIMODA received the B.F. degree in fisheries science and the M.S. degree in environmental science from Hokkaido University. He is currently the Senior Research Manager with the Fisheries Research Department, Salmon and Freshwater Fisheries Research Institute, Hokkaido Research Organization, and a member of the Japanese Society of Fisheries Science and the Ichthyological Society of Japan.