Aging Contrast: A Contrastive Learning Framework for Fish Re-identification Across Seasons and Years

Weili Shi[1] [0000-0002-9907-2009] *, Zhongliang Zhou[2] [0000-0003-4471-6759] *, Benjamin H. Letcher[3] [0000-0003-0191-5678], Nathaniel Hitt[3] [0000-0002-1046-4568], Yoichiro Kanno[4] [0000-0001-8452-5100], Ryo Futamura[5] [0000-0001-8128-3476], Osamu Kishida[5] [0000-0002-2663-8155], Kentaro Morita[6] [0000-0002-7803-2438], and Sheng Li[1] [0000-0003-1205-8632]

[1] University of Virginia, Charlottesvile, VA, USA
{rhs2rr, shengli}@virginia.edu
[2] University of Georgia, Athens, GA, USA
Zhongliang.Zhou@uga.edu
[3] United States Geological Survey, Eastern Ecological Science Center, USA
{bletcher, nhitt}@usgs.gov
[4] Colorado State University, Fort Collins, CO, USA
yoichiro.kanno@colostate.edu
[5] Hokkaido University, Japan
rfutamurattg@gmail.com, kishida@fsc.hokudai.ac.jp
[6] The University of Tokyo, Japan
moriken@g.ecc.u-tokyo.ac.jp

Abstract. The fields of biology, ecology, and fisheries management are witnessing a growing demand for distinguishing individual fish. In recent years, deep learning methods have emerged as a promising tool for image-based fish recognition. Our study is focused on the re-identification of masu salmon from Japan, wherein fish were individually marked and photographed to evaluate discriminative body characteristics. Unlike previous studies where fish were sampled during the same time period, we evaluated individual re-identification across seasons and years to address challenges due to aging, seasonal variation, and other factors. In this paper, we propose a new contrastive learning framework called Aging Contrast (AgCo) and evaluate its performance on the masu salmon dataset. Our analysis indicates that, unlike large changes in body size over time, the pattern of parr marks on the lateral line of the fish body remains relatively stable, despite some change in coloration across seasons. AgCo accounts for such seasonally-invariant features and performs

[*] These authors contributed equally to this work
The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey.

re-identification based on the cosine similarity of these features. Extensive experiments show that our AgCo method outperforms other state-of-the-art methods.

Keywords: Fish Re-identification · Contrastive Learning · Seasonally-invariant Features.

1 Introduction

With the global human population now exceeding 8 billion, there is a growing need for sustainable food sources. Fish, being a vital source of protein, comprise a significant portion of the diet for many across the globe, and increased harvest pressures underscore the importance of effective approaches for fish production and conservation. Moreover, the United Nations' Food and Agriculture Organization reported that approximately 80 percent of marine fish production is from wild populations [15] where monitoring and assessment can be difficult and expensive. In this paper, we propose a new analytical framework that could potentially revolutionize fish cultivation and conservation in natural environments as well as aquaculture systems.

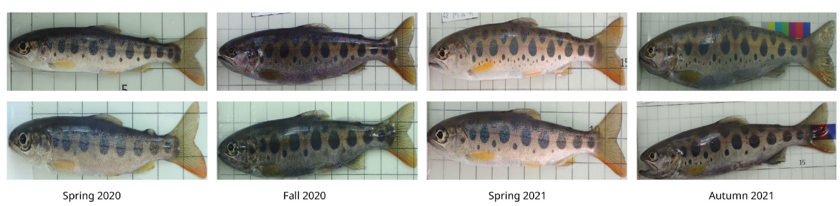

Fig. 1. Samples of two fish identities in four different seasons. Compared to the dramatic change of the size, color and dots, the pattern of parr marks located on the lateral line of the fish body remains relatively consistent across seasons. In our study we filter out noisy information to focus on learning the seasonally-invariant features from parr marks.

Distinguishing between individuals within a species is a fundamental step for understanding demographic processes in animal populations [38, 11, 4, 19, 33, 31, 34, 1, 36, 12]. Currently, most individual-identification systems can be grouped into either invasive or non-invasive processes. For invasive methods, tagging, altering, or coloring specific body parts have been implemented. Among these, tagging is especially popular for its accuracy, and many tagging techniques have been developed. Despite its utility, such tagging procedures have certain inherent drawbacks: mortality rates can increase after tagging, stress caused by tagging could impact the recapture rate, and the time and cost for this process can limit the survey areas and sample sizes available for research. Among non-invasive methods, DNA samples have shown potential in animal individual recognition, but such methods often require operators with advanced professional skills and technical equipment that is not readily available.

Recent advances have enabled imagery-based individual identification for various wildlife species [38, 11, 28, 36, 26, 7, 14, 20] based on convolutional neural network analysis of individually-diagnostic pigmentation patterns. However, such applications for wildlife management and conservation are often limited by uncertainty associated with organismal growth and development that may change pigmentation patterns and thus decrease re-identification accuracy. In this paper, we develop and evaluate a new approach to identify *Oncorhynchus masou* (masu salmon) across seasons and years.

Masu salmon is a species of great economic, cultural, and ecological importance in northern Japan (see supplementary material). Typically, masu salmon have a darkened back with small pigmentation spots and larger oval-shaped parr marks on their sides (Fig. 1). The species exhibits a migratory life-history strategy whereby spawning and juvenile development occurs in freshwater environments, a subset of sub-adults migrate into marine environments, and adults return to freshwater environments for spawning [16]. An understanding of individual variation in fish movement patterns and growth rates is necessary to quantify population dynamics and extirpation risks for this valuable species [24].

Compared to recognition tasks based on human datasets, the creation of similar pipelines for wildlife species is hindered by several difficulties. The shorter life spans of animals and fewer opportunities to collect images, as compared to humans, often result in a scarcity of data per individual, making it challenging to assemble a dataset of sufficient quality. Additionally, the strong stress response often exhibited by wild animals can make recapture for the purpose of collecting data difficult. In addition, the time-consuming nature of manual labeling results in limited dataset diversity and size. Finally, the aging process in animals may have a more pronounced effect on their visual patterns than in humans, thereby potentially affecting recognition accuracy.

To mitigate some of the previously mentioned constraints, the advent of self-supervised learning has demonstrated its efficacy as a means to automatically discern patterns from vast quantities of unlabeled data. As an unsupervised representation learning model, contrastive learning seeks to learn representations that bring similar objects together while separating unrelated objects. In our study, we propose the use of individual fish identities and temporal information within the dataset to improve the representation of the same individual over time, while pushing the representations of different individuals farther apart. Compared to simple, random augmentations of the same images in training data, we demonstrate that our method yields more accurate re-identification of individuals over time.

In this paper, we propose a novel framework called Aging Contrast (AgCo) for fish re-identification across multiple seasons and years. The primary challenge in this task is the misalignment of features of the same fish identity, caused by significant changes in fish appearance over time. However, our examination of masu salmon samples revealed that, despite considerable changes in the dots, size of the salmon and background coloration, the pattern of parr marks located on the lateral line of the fish body remains relatively consistent over time.

Compared to other unsupervised contrastive learning methods [25, 8], our AgCo framework measures the similarity of the query and key from two domains (i.e., different seasons) and perform the data augmentation on the feature level to obtain the transitional features between two seasons. As a result, AgCo can learn seasonally-invariant features from the pattern of parr marks for each fish identity because the change in the pattern of parr marks is more predictable. The major contributions of this paper are summarized as follows:

- Compared to a previous study [11], which performed fish re-identification over a maximum period of six months, our study focuses on a longer time span, with a time gap between sampled fish that can exceed one year. The greater time span can encompass more significant changes in fish appearance and therefore poses a more substantial challenge for fish re-identification. A new dataset for fish re-identification across ages is collected.
- Based on the analysis of masu salmon images, we propose the Aging Contrast (AgCo) framework, which learns seasonally-invariant features from the pattern of the parr marks of fish sampled at different times.
- Extensive experimental results demonstrate that, in the settings with two or more seasos, our AgCo framework significantly outperforms state-of-the-art contrastive learning methods such as MoCo [17] and SimCLR [8].

2 Related Work

2.1 Deep Learning for Fish Recognition

Until recently, fish recognition by images emerged as an appealing area due to its theoretical and applied significance to aquaculture and marine biology, and it has gained great interest from researchers around the world. This task poses great challenges since the collected images of the fish might be of low quality (e.g., noisy or distorted), which heavily affects the recognition. Recently, Alsmadi et al. [2] obtained the distinct features through distance and geometrical measurements. The obtained features were fed into a neural network to distinguish 20 different fish families. To improve image resolution, Sun et al. [35] generated high resolution images from raw images. Then the discriminative features could be obtained from the refined images, and the support vector machine (SVM) was used for fish recognition. Ding et al. [13] proposed several convolutional neural networks (CNN) architectures to identify the fish from four species. Hridayami et al. [21] fine-tuned the VGG16 on four different types of datasets, and their results showed that blending image with an RGB image trained model exhibited the best performance for recognition of fish species.

Another branch of work focuses on fish recognition at the individual level. Cisar et al. [11] performed individual identification on Atlantic salmon based on the dot patterns on the skin of the body. Mcinness et al. [29] showed the ability to discriminate the individual freckled hawkfish by the natural markings on their body. [5] further demonstrated that the visible patterns such as the stripes on the fish body could be effective for fish individual identification. Apart from

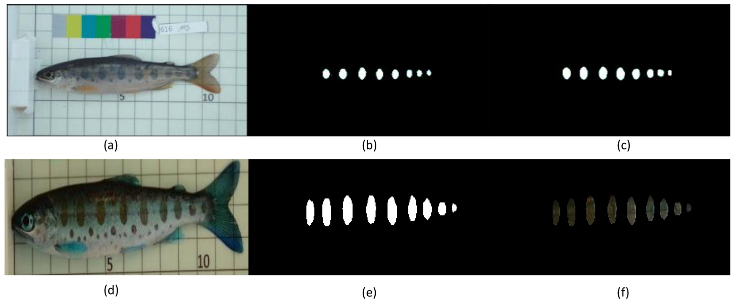

Fig. 2. Illustrations of: (a) original fish image; (b) predicted parr marks from segmentation pipeline; (c) ground truth parr marks from manual labelling; (d) ROI image of fish; (e) ROI image of predicted parr marks; and (f) parr mark pigmentation images generated by the fusion of (d) and (e).

using the visible patterns, other recent studies [3, 22, 30] explored an alternative method that adopted scale patterns on the body as discriminative features for fish individual identification. For instance, Zhou et al. [38] used the dots patterns from brook trout as the biomarker and trained a CNN based model to learn discriminative features for fish individual recognition.

2.2 Contrastive Learning

Contrastive learning methods provide a powerful tool to pre-train the model without the need for a large number of labels. The core idea behind contrastive learning is to aggregate the positive sample pairs and repulse the negative sample pairs. Most contrastive learning methods [25, 17, 8, 23, 9] adopt a contrastive loss to maximize the similarity of positive pairs and enlarge the gap between the negative pairs. Some new mechanisms, such as momentum encoder [17] and cluster alignment [25, 6], are introduced in latest contrastive learning framework. Apart from the InfoNCE [32] loss adopted in many contrastive learning methods, ProtoNCE [25] is proposed to estimate the concentration for the feature distribution around each prototype. Although promising, the aforementioned contrastive learning methods do not consider temporal changes in data (e.g., fish across ages), and thus they cannot effectively deal with the fish re-identification task across seasons and years.

3 Dataset

The fish dataset used in our study was acquired from the Horonai River in Japan, as part of a long-term research project that marked masu salmon with Passive Integrated Transponder (PIT) tags. Four capture-mark-recapture surveys were conducted across the survey area in the following seasons: Spring 2020, Autumn 2020, Spring 2021, and Autumn 2021. The samples of the fish from Spring 2020, Autumn 2020, Spring 2021, and Autumn 2021 are illustrated in Fig. 1. More details of our dataset can be found in supplementary material.

4 Proposed Method

We present a novel approach for (1) isolating the defining characteristics of individual masu salmon and (2) employing a contrastive learning framework to extract representations in a self-supervised manner. Our methodology takes into account previous findings that parr marks on a masu salmon's body can serve as a distinct biomarker for individuals. Thus, the first step of our pipeline is to automatically isolate the parr marks from their surroundings and create a more distinctive pattern for analysis. Subsequently, we utilize the proposed AgCo framework to discern aging-resistant features from the image of each fish.

4.1 Segmentation and Feature Extraction

In this stage, we adopt we utilized a segmentation network with a Feature Pyramid Network [27] architecture to extract the parr marks for all images of fish, which served as the primary features for subsequent steps. The results from the segmentation are illustrated in Fig. 2. More details can be found in supplementary material. To further take advantage of information from the parr marks such as the texture and color, the predicted parr marks are fused with the original images to generate parr mark images for analysis as shown in Fig. 2.

4.2 Aging Contrast Framework

Our Aging Contrast (AgCo) framework aims to learn seasonally-invariant features for each fish from the parr mark pigmentation images. Though the overall appearance of each fish may change dramatically over time, the parr mark patterns from each fish are relatively consistent across different seasons. Thus, parr marks are the key for successful re-identification of masu salmon from different ages.

Since there is only a minor shift in the pattern of the parr marks from the observed fish samples in different seasons, we assume that a transitional feature would exist, which represents the potential fish sample between season A and season B. The intermediate feature can be approximated by the linear interpolation of the fish images from season A and season B due to the minor and regular change of the pattern of parr marks. In our study, we adopt mixup [37] to approximate the transitional features. The transitional features contain both temporal information of the parr marks of the fish from both season A and season B. By maximizing the similarity of the transitional features, the season-invariant features can be obtained, and fish re-identification can be performed by measuring the similarity of such seasonally-invariant features.

The AgCo framework is illustrated in Fig. 3. Compared to conventional contrastive learning methods, we take advantage of the fish identities information, and the parr mark pigmentation images (x_1 and x_2) of same identity from two seasons are adopted as the input. T_1 and T_2 are sets of commonly used data augmentation methods which includes rotation, color jittering, and flipping. For each image, two views are generated T_1 and T_2. A neural network encoder f

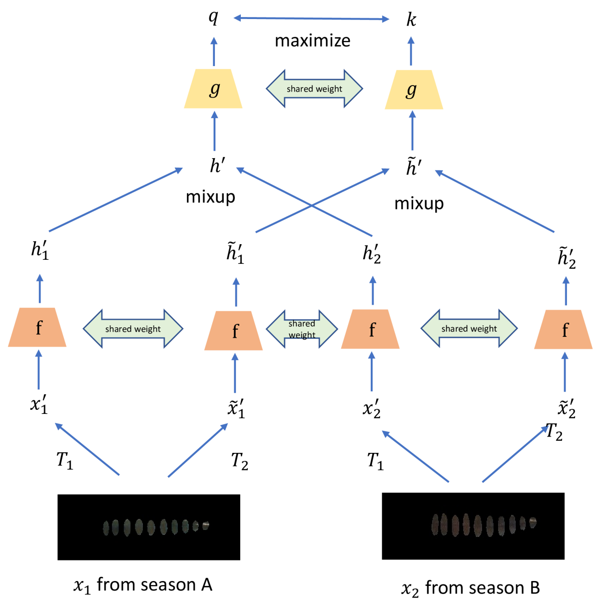

Fig. 3. The diagram depicts our proposed Aging Contrast (AgCo) framework. The features extracted from the same individual in two different seasons are fed into deep neural networks to produce their respective representations. Mixup operation is applied in the representation space to create two augmented representations. We aim to obtain seasonally-invariant features for the same fish identity by maximizing the query and key generated from the augmented latent features h' and $\tilde{h}'$.

extracts the latent features from all the views of the input images. In our study, we adopt the commonly used ResNet-152 [18] as the encoder f. Different from existing contrastive learning methods, we perform the data augmentation on the features from different views to form the transitional features. To model the transitional features from parr marks of the potential fish between two seasons, transitional features h', $\tilde{h}' \in \mathbb{R}^d$ are generated by combination of the features from different seasons.

$$h' = \lambda h_1' + (1-\lambda)h_2' \qquad \tilde{h}' \quad = \lambda\tilde{h}_2' + (1-\lambda)\tilde{h}_1', \tag{1}$$

where h_1', $\tilde{h}_1'$, h_2', and $\tilde{h}_2'$ are the representative features generated from the views of the images from season A and season B. The transitional features h' and $\tilde{h}'$ are joint combination of the features from different seasons. λ denotes the hyperparameter that regulates the proportion of the component features to the targeted features. As shown in Eq. (1), transitional features h' and $\tilde{h}'$ are constructed by the features from the images views augmented by T_1 and T_2, respectively. To make the transitional features distinctive for contrastive learning, the proportion of component features in h' is different to that in $\tilde{h}'$.

Next, we map the transitional features to query q and key k by the projection head g. g is implemented as the two-layered multi-layer perception (MLP). For each query q, k^+ is the key from the same fish identity as query while k^- comes from different identities. By maximizing the similarity of positive pair (q, k^+) and the dissimilarity of negative pair (q, k^-) via a contrastive loss function, the neural network encoder f can learn the seasonally-invariant features for each fish identity. In most contrastive learning framework, the InfoNCE [32] loss is adopted.

$$\mathcal{L}_{\text{InfoNCE}} = -\log \frac{\exp\left(q \cdot k^+/\tau\right)}{\sum_i \exp\left(q \cdot k_i/\tau\right)},$$

where τ is the hyperparameter that scales the distribution of the similarity distance. q is the query and k is key. k^+ is the key from the positive pair. The query and key are generated from the same domain. In our study, the query and key are combination of the features of fish samples from different seasons. And the loss function is modified accordingly, as shown in Eq (2):

$$\mathcal{L} = -\frac{1}{N}\sum_i \log \frac{\exp\left(q_i \cdot k_i^+/\tau\right)}{\sum_j \exp\left(q_i \cdot k_j/\tau\right)} = -\frac{1}{N}\sum_i \log \frac{\exp\left(g(h_i') \cdot g(\tilde{h}_i')/\tau\right)}{\sum_j \exp\left(g(h_i') \cdot g(\tilde{h}_j')/\tau\right)}, \tag{2}$$

where N is the total number of fish identities. Details of the AgCo framework can be found in supplementary material. In general, the AgCo framework aims to capture the common traits of the parr marks that exist in the fish samples in different seasons and generate the season-invariant features that can be aligned on different seasons for each fish identity.

5 Experiments

In this section, we introduce the experimental settings for our study. Our proposed method is evaluated for fish observed in all surveys (Spring 2020, Autumn 2020, Spring 2021, Autumn 2021). The results of the AgCo is compared against several representative contrastive learning methods. Ablation studies and visualizations of outcomes can be found in supplementary material.

Experimental Settings. In the experiments, we performed the fish re-identification task by measuring the cosine similarity of the latent features generated by the neural network encoder f shown in Algorithm 1. We compare the results of AgCo against that of latest contrastive learning frameworks including SimSiam [10], PCL [25], MoCo [17] and SimCLR [8] as well as pretrained ResNet-152 [18] without any contrastive learning fine-tuning. More details of the experimental settings can be found in supplementary material.

Results and Analysis. The results of our AgCo and the baseline methods on the tasks of S1 $\rightarrow$ SX (X = 2,3,4) and S2 $\rightarrow$ SX (X=3,4), S3 $\rightarrow$ S4 are shown in Table 1 and Table 2, respectively. We find that the pretrained model which is not fine-tuned by contrastive learning yields the worst performance on all tasks. The poor performance originated from the inability of the model to generate the discriminative features for each fish identity. When the individual fish grows and the characteristics of the fish are changed, the feature generated by the model for one fish identity in the subsequent season is prone to be aligned with the feature of another fish identity in the previous season. We also observed that

Table 1. The performance of our AgCo and the baseline methods on S1 → SX (X=2,3,4). Top-1/3/5 accuracy for each model is reported. The bold represents the best results.

Method	Task		
	S1 → S2	S1→ S3	S1→ S4
Pretained Only	29.2/43.9/53.6	14.6/36.5/46.3	12.1/31.7/46.3
SimSiam	34.1/51.2/61.0	17.1/39.0/53.7	17.1/31.7/46.3
PCL	34.1/51.2/61.0	19.5/43.9/48.7/	14.6/29.2/34.1
MoCo	48.8/58.5/70.7	39.0/63.4/73.2	14.6/36.5/44.0
SimCLR	80.4/87.8/90.2	63.4/80.4/95.1	41.4/60.9/80.4
AgCo (Ours)	**100.0/100.0/100.0**	**90.2/100.0/100.0**	**73.2/90.2/92.7**

Table 2. The performance of our AgCo and the baseline methods on S2 → SX (X=3, 4) and S3 → S4. Top-1/3/5 accuracy for each model is reported. The bold represents the best results.

Method	Task		
	S2 → S3	S2→ S4	S3→ S4
Pretained Only	24.3/43.9/60.9	19.5/36.5/46.3	24.3/41.5/51.2
SimSiam	29.2/29.3/39.0	7.3/14.6/26.8	24.4/36.6/68.3
PCL	65.8/83.0/90.2	39.0/56.1/70.7	60.9/73.2/78.0
MoCo	80.5/92.7/97.6	56.1/73.2/85.3	39.1/56.1/68.3
SimCLR	97.5/97.5/97.5	75.6/85.3/90.2	70.7/87.8/92.6
AgCo (Ours)	**100.0/100.0/100.0**	**87.8/95.2/97.6**	**92.7/95.1/97.5**

for the pretrained-only method, the model can achieve 29.2% top-1 accuracy on S1 → S2. While on S1 → S3 and S1 → S4, this value is decreased to 14.6% and 12.1%. This suggests that the misalignment of the features generated by the model from two different seasons would be exacerbated if the time span of two seasons is larger. It is understandable that since the time span is larger, the change in the characteristics of the fish would be more significant. Therefore, it is more difficult for the model to generate the aligned features for the same fish over longer periods of time.

Compared to the pretrained only method, The model fine-tuned by SimSiam [10] only gains very limited improvement performance on all the tasks. Compared to other contrastive learning methods [25, 17, 8] which take the similarity and dissimilarity into account for queries and keys, SimSiam [10] only consider similarity of the queries and keys in the loss function. We assume that the low performance is mainly caused by lack of discrimination of the features from different fish identities.

As shown from the experimental results, there is prominent improvement in the performance of the models for the baseline methods which adopts InfoNCE [32] as the loss function. For instance, on the task S2 → S3, the top-1 accuracy of MoCo [17] and SimCLR [8] can reach 80.5% and 97.5%, respectively. The experiments suggest that the discriminative features learned by contrastive methods is the key to the improved performance. Even if the model can only 'see' the fish in one season in the training stage, the features of same fish identity from different seasons can still be aligned to some extent since features of different fish identities are assumed to be pushed away from each other.

When the pretrained model is further fine-tuned with our proposed AgCo, the performance of the model is further improved. For instance, the top-1 accuracy of the model can reach 100.0% on S1 $\rightarrow$ S2, almost 20% higher than that of SimCLR [8]. It also is observed that the performance of other contrastive baseline methods is lower than that of AgCo on all tasks. Though the model trained by the these baseline methods is able to learn the discriminative features for each fish identity, the discrimination of the features is confined to a single season and the temporal information of the parr marks on the fish body is neglected. We notice that on S1 $\rightarrow$ S4, the top-1 accuracy of the model is degraded to 41.4% for SimCLR [8]. It indicates that there is still a large chance that the features learned from different seasons are misaligned for a single fish identity. In contrast, our proposed AgCo framework is not only able to learn the discriminative features for each fish but also to account for possible changes in parr marks. As mentioned, the seasonally-invariant features learned by the AgCo framework can capture common traits of the parr marks that exist in the fish samples from different seasons. Thus, even if the appearance of an individual fish has been dramatically changed, the features of the fish from different seasons can still be aligned. For the most difficult task S1 $\rightarrow$ S4, the top-1 accuracy of the model trained by AgCo can still achieve 73.2%.

6 Conclusion

Recognition of individual fish is necessary for many aspects of fisheries management and conservation, and our study contributes a novel contrastive learning framework called Aging Contrast (AgCo) for this purpose that outperforms other approaches. Our results also highlight the specific importance of parr marks as the key for the fish re-identification task. Our proposed method can learn the seasonally-invariant features that enable accurate re-identification of individual fish as they develop over time. Further applications of our framework could improve fisheries management and conservation globally.

Acknowledgement. This material is based upon work supported by the U.S. Geological Survey under Grant/Cooperative Agreement No. G22AC00372.

References

1. Al-Jubouri, Q., Al-Azawi, R., Al-Taee, M., Young, I.: Efficient individual identification of zebrafish using hue/saturation/value color model. The Egyptian Journal of Aquatic Research **44**(4), 271–277 (2018)
2. Alsmadi, M.K., Omar, K.B., Noah, S.A., Almarashdeh, I.: Fish recognition based on robust features extraction from size and shape measurements using neural network. Journal of Computer Science **6**(10), 1088 (2010)
3. Bekkozhayeva, D., Cisar, P.: Image-based automatic individual identification of fish without obvious patterns on the body (scale pattern). Applied Sciences **12**(11), 5401 (2022)

4. Bekkozhayeva, D., Saberioon, M., Cisar, P.: Automatic individual non-invasive photo-identification of fish (sumatra barb puntigrus tetrazona) using visible patterns on a body. Aquaculture International **29**(4), 1481–1493 (2021)
5. Bekkozhayeva, D., Saberioon, M., Cisar, P.: Automatic individual non-invasive photo-identification of fish (sumatra barb puntigrus tetrazona) using visible patterns on a body. Aquaculture International **29**(4), 1481–1493 (2021)
6. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., Joulin, A.: Unsupervised learning of visual features by contrasting cluster assignments. Advances in neural information processing systems **33**, 9912–9924 (2020)
7. Chen, P., Swarup, P., Matkowski, W.M., Kong, A.W.K., Han, S., Zhang, Z., Rong, H.: A study on giant panda recognition based on images of a large proportion of captive pandas. Ecology and Evolution **10**(7), 3561–3573 (2020)
8. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: International conference on machine learning. pp. 1597–1607. PMLR (2020)
9. Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G.E.: Big self-supervised models are strong semi-supervised learners. Advances in neural information processing systems **33**, 22243–22255 (2020)
10. Chen, X., He, K.: Exploring simple siamese representation learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 15750–15758 (2021)
11. Cisar, P., Bekkozhayeva, D., Movchan, O., Saberioon, M., Schraml, R.: Computer vision based individual fish identification using skin dot pattern. Scientific Reports **11**(1), 1–12 (2021)
12. Delcourt, J., Ovidio, M., Denoël, M., Muller, M., Pendeville, H., Deneubourg, J.L., Poncin, P.: Individual identification and marking techniques for zebrafish. Reviews in Fish Biology and Fisheries **28**(4), 839–864 (2018)
13. Ding, G., Song, Y., Guo, J., Feng, C., Li, G., He, B., Yan, T.: Fish recognition using convolutional neural network. In: OCEANS 2017-Anchorage. pp. 1–4. IEEE (2017)
14. Ding, R., Wang, L., Zhang, Q., Niu, Z., Zheng, N., Hud, G.: Fine-grained giant panda identification. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 2108–2112. IEEE (2020)
15. Food, Organization, A.: The state of world fisheries and aquaculture. Tech. rep., Food and Agriculture Organization of the United Nations (2022)
16. Futamura, R., Morita, K., Kanno, Y., Kumikawa, S., Matsuoka, Y., Okuda, A., Sugiyama, H., Takahashi, H., Uchida, J., Kishida, O.: Size-dependent growth tactics of a partially migratory fish before migration. Oecologia **198**, 371–379 (2022)
17. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9729–9738 (2020)
18. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
19. Hirsch, P.E., Eckmann, R.: Individual identification of eurasian perch perca fluviatilis by means of their stripe patterns. Limnologica **54**, 1–4 (2015)
20. Hou, J., He, Y., Yang, H., Connor, T., Gao, J., Wang, Y., Zeng, Y., Zhang, J., Huang, J., Zheng, B., et al.: Identification of animal individuals using deep learning: A case study of giant panda. Biological Conservation **242**, 108414 (2020)

21. Hridayami, P., Putra, I.K.G.D., Wibawa, K.S.: Fish species recognition using vgg16 deep convolutional neural network. Journal of Computing Science and Engineering **13**(3), 124–130 (2019)
22. Huntingford, F., Borçato, F., Mesquita, F.: Identifying individual common carp cyprinus carpio using scale pattern. Journal of fish biology **83**(5), 1453–1458 (2013)
23. Kalantidis, Y., Sariyildiz, M.B., Pion, N., Weinzaepfel, P., Larlus, D.: Hard negative mixing for contrastive learning. Advances in Neural Information Processing Systems **33**, 21798–21809 (2020)
24. Kanno, Y., Harris, A., Kishida, O., Utumi, S., Uno, H.: Complex effects of body length and condition on within-tributary movement and emigration in stream salmonids. Ecology of Freshwater Fish **31**, 317–329 (2021)
25. Li, J., Zhou, P., Xiong, C., Hoi, S.C.: Prototypical contrastive learning of unsupervised representations. arXiv preprint arXiv:2005.04966 (2020)
26. Li, W., Ji, Z., Wang, L., Sun, C., Yang, X.: Automatic individual identification of holstein dairy cows using tailhead images. Computers and electronics in agriculture **142**, 622–631 (2017)
27. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2117–2125 (2017)
28. Matkowski, W.M., Kong, A.W.K., Su, H., Chen, P., Hou, R., Zhang, Z.: Giant panda face recognition using small dataset. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 1680–1684. IEEE (2019)
29. McInnes, M.G., Burns, N.M., Hopkins, C.R., Henderson, G.P., McNeill, D.C., Bailey, D.M.: A new model study species: high accuracy of discrimination between individual freckled hawkfish (paracirrhites forsteri) using natural markings. Journal of Fish Biology **96**(3), 831–834 (2020)
30. Morgado-Santos, M., Matos, I., Vicente, L., Collares-Pereira, M.: Scaleprinting: individual identification based on scale patterns. Journal of Fish Biology **76**(5), 1228–1232 (2010)
31. Navarro, J., Perezgrueso, A., Barría, C., Coll, M.: Photo-identification as a tool to study small-spotted catshark scyliorhinus canicula. Journal of Fish Biology **92**(5), 1657–1662 (2018)
32. Oord, A.v.d., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748 (2018)
33. Sandford, M., Castillo, G., Hung, T.C.: A review of fish identification methods applied on small fish. Reviews in Aquaculture **12**(2), 542–554 (2020)
34. Stien, L.H., Nilsson, J., Bui, S., Fosseidengen, J.E., Kristiansen, T.S., Øverli, Ø., Folkedal, O.: Consistent melanophore spot patterns allow long-term individual recognition of atlantic salmon salmo salar. Journal of Fish Biology **91**(6), 1699–1712 (2017)
35. Sun, X., Shi, J., Dong, J., Wang, X.: Fish recognition from low-resolution underwater images. In: 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). pp. 471–476. IEEE (2016)
36. Whooley, P., Berrow, S., Barnes, C.: Photo-identification of fin whales (balaenoptera physalus l.) off the south coast of ireland. Marine Biodiversity Records **4** (2011)
37. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: International Conference on Learning Representations (2018)
38. Zhou, Z., Hitt, N.P., Letcher, B.H., Shi, W., Li, S.: Pigmentation-based visual learning for salvelinus fontinalis individual re-identification. In: 2022 IEEE International Conference on Big Data (Big Data). pp. 6850–6852. IEEE (2022)