Self-supervised Learning for Sonar Images: Enhancing Multi-modal Perception for Underwater Applications

Alan Preciado-Grijalva[1,2], Miguel Bande[1], Bilal Wehbe[1], Matias Valdenegro-Toro[1]

[1]German Research Center for Artificial Intelligence, 28359 Bremen, Germany

[2]Bonn-Rhein-Sieg University of Applied Sciences, 53757 Sankt Augustin, Germany

Abstract

Neural networks have shown promising results in sonar perception tasks such as object recognition [1], image patch matching [2] and image classification [3]. In the context of autonomous underwater vehicles, it is crucial to develop robust models to overcome the challenges of underwater perception.

In this work, we report progress on a comparative evaluation of self-supervised learning (SSL) [6][7] and supervised learning (SL) as pretraining methods for sonar images. As a first step, we produce pre-trained neural networks on the Marine Debris Watertank Dataset [4] via a SSL method that classifies image rotations [5] and a traditional SL approach to classify the actual image labels. In both cases, we trained a Resnet20, SqueezeNet, Mobilenet, DenseNet121 and MiniXception on images of size 96x96. Thereafter, we evaluate the quality of the learned features by using transfer learning for low-shot classification on a target dataset called Marine Debris Turntable [3].

The results presented in this poster indicate that the SSL pre-trained models have a similar classification performance compared to the SL counterpart across all the neural network models. These results indicate that SSL pre-training are a promising substitute for SL methods without compromising object classification and no need of manual label annotations.

Finally, we report on the creation of a new underwater dataset that contains paired camera and sonar images for different underwater objects (panels, cement pipes, ladders, ramps). This dataset, called Gemini Sonar Dataset, will allow us to perform further classification, image translation and object detection tasks using SSL approaches.

Pre-trained models on watertank sonar data

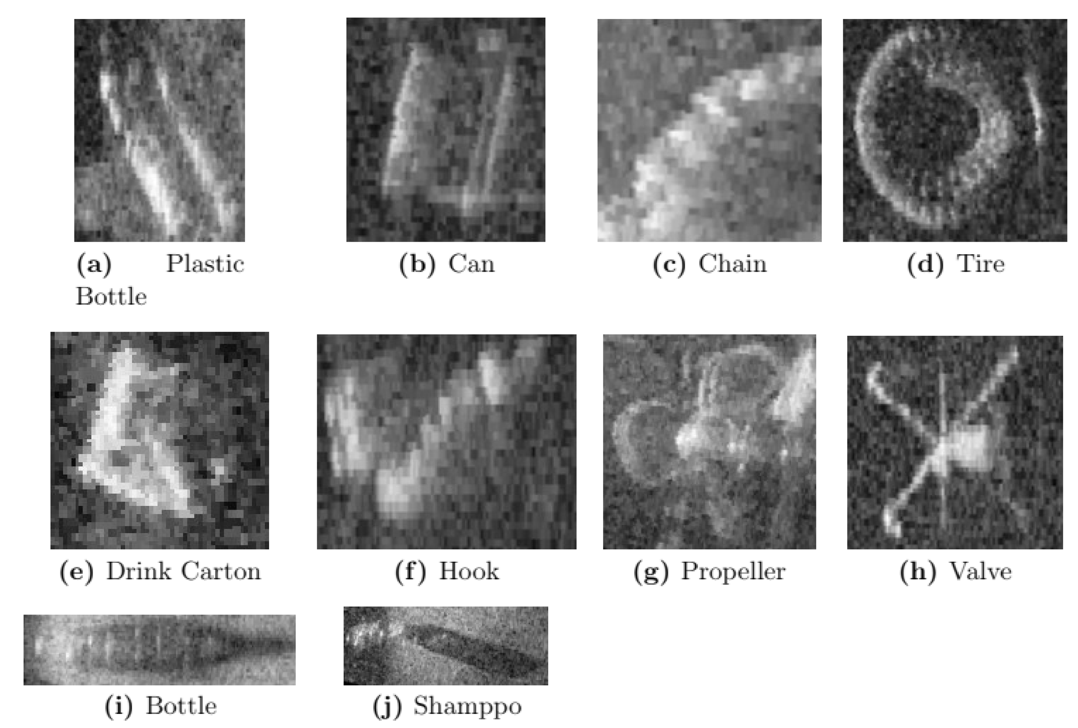
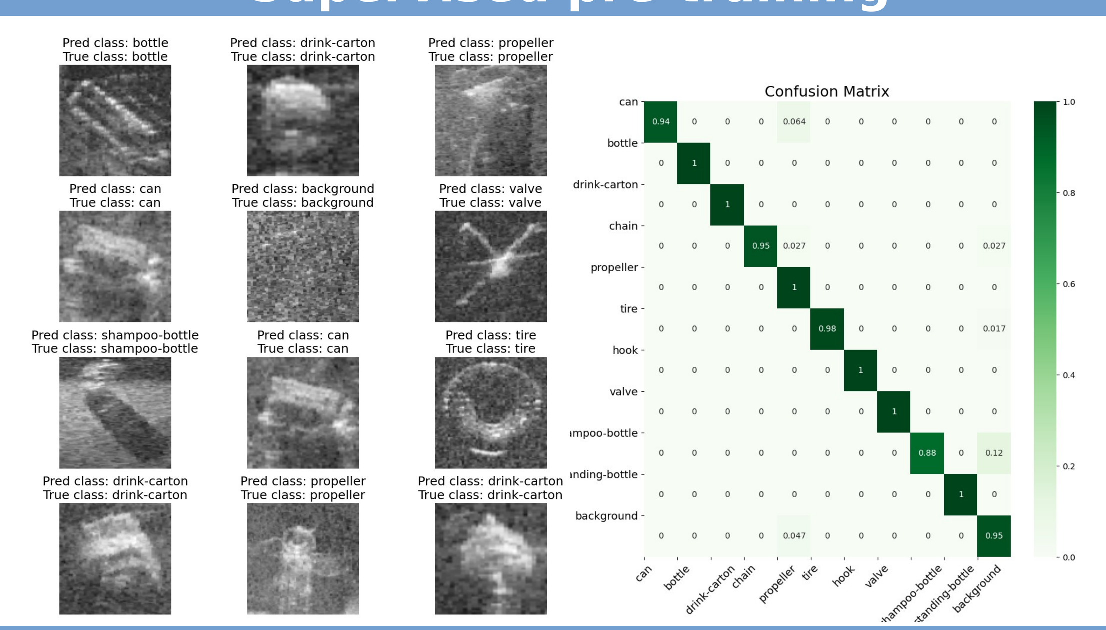
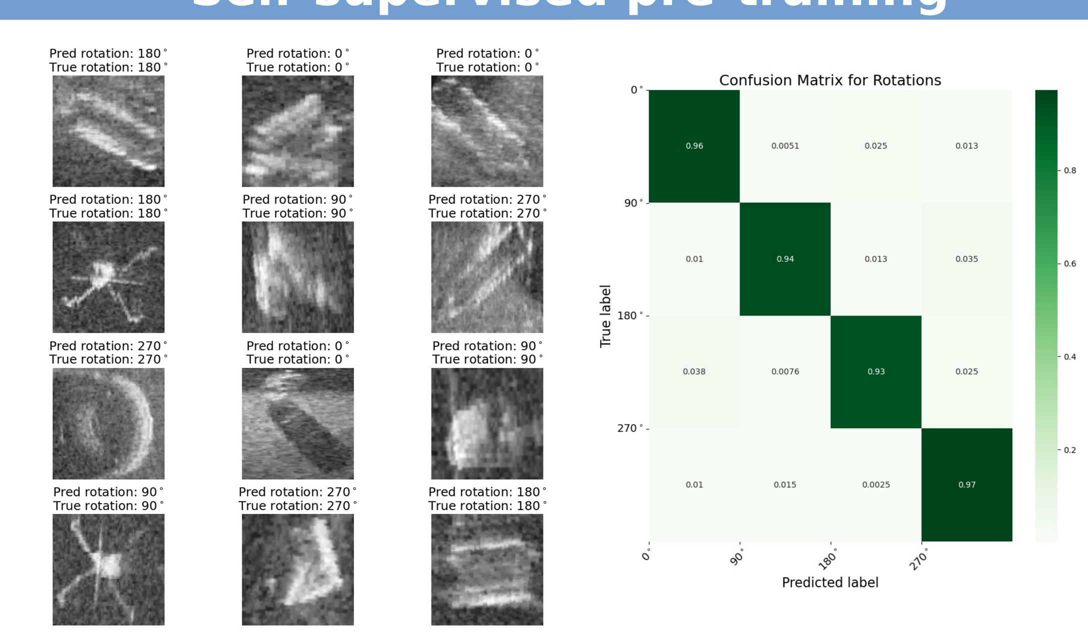


Figure 1: Samples from the Marine Debris Watertank Dataset captured with the ARIS Explorer 3000 Forward Looking Sonar.

Supervised pre-training



Self-supervised pre-training



SSL Model	Baseline Model	Image Size	Test Accuracy
RotNet	ResNet20 **SSL**	96x96	97.22%
	ResNet20 **SL**	96x96	96.46%
RotNet	MobileNet **SSL**	96x96	94.43%
	MobileNet **SL**	96x96	98.23%
RotNet	DenseNet121 **SSL**	96x96	95.38%
	DenseNet121 **SL**	96x96	96.46%
RotNet	SqueezeNet **SSL**	96x96	95%
	SqueezeNet **SL**	96x96	97.47%
RotNet	Minixception **SSL**	96x96	96.14%
	Minixception **SL**	96x96	96.71%
NA	Linear SVM **SSL**	96x96	76.40%
	Linear SVM **SL**	96x96	96.70%

Few-shot transfer learning evaluations on target turntable sonar data

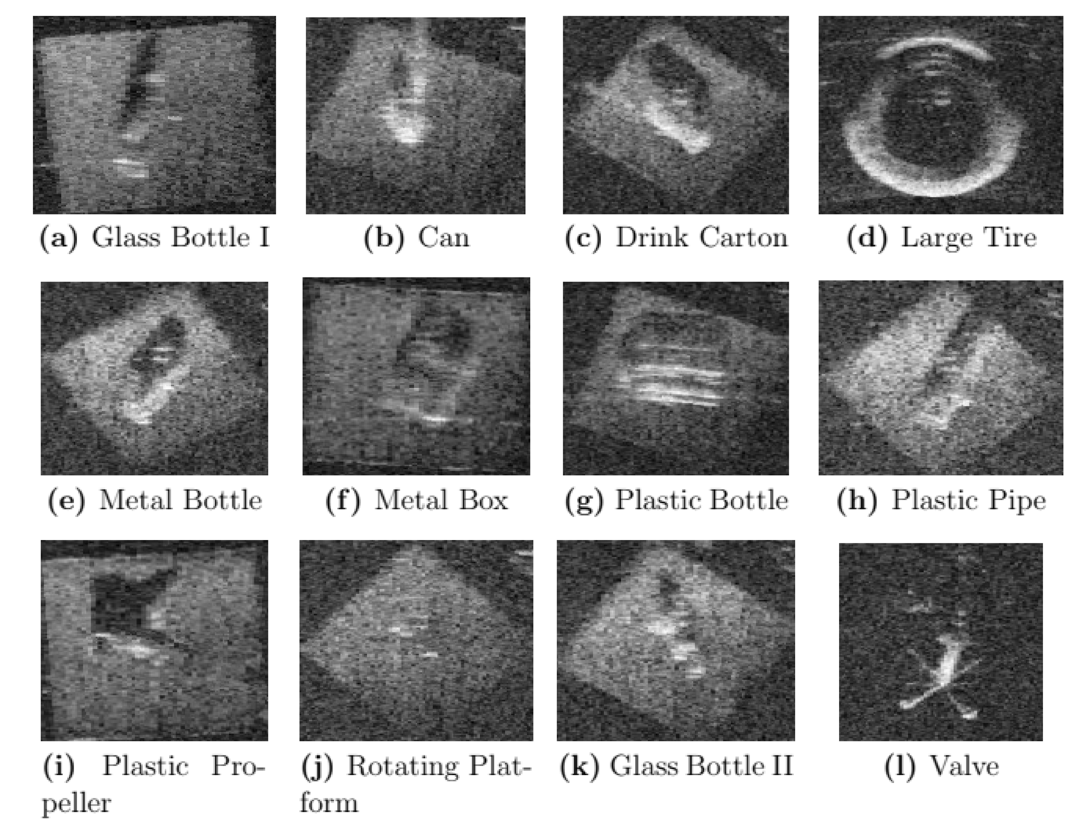


Figure 1: Samples from the Marine Debris Turntable Dataset captured with the ARIS Explorer 3000 Forward Looking Sonar. The dataset captures different objects located on a rotated table.

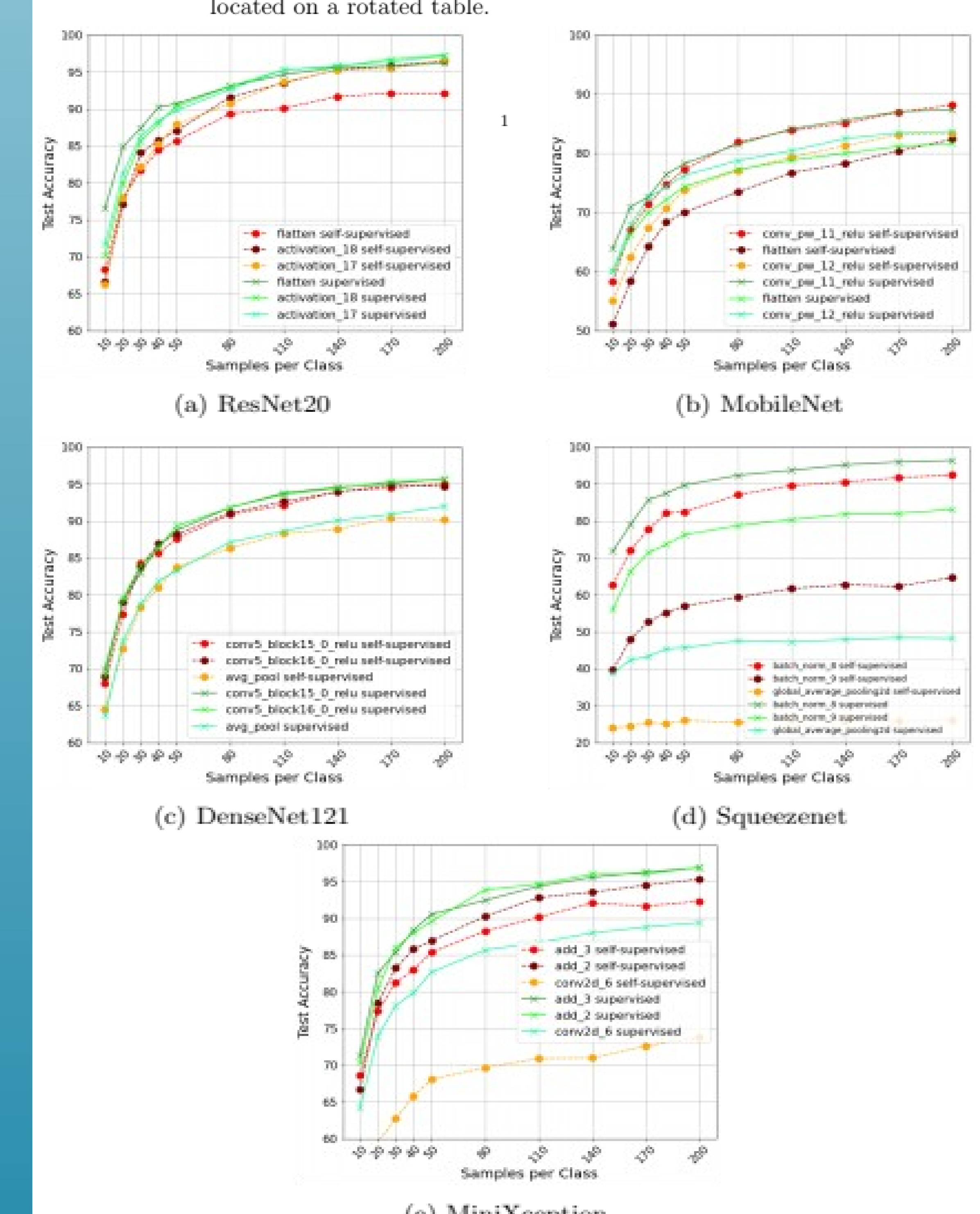
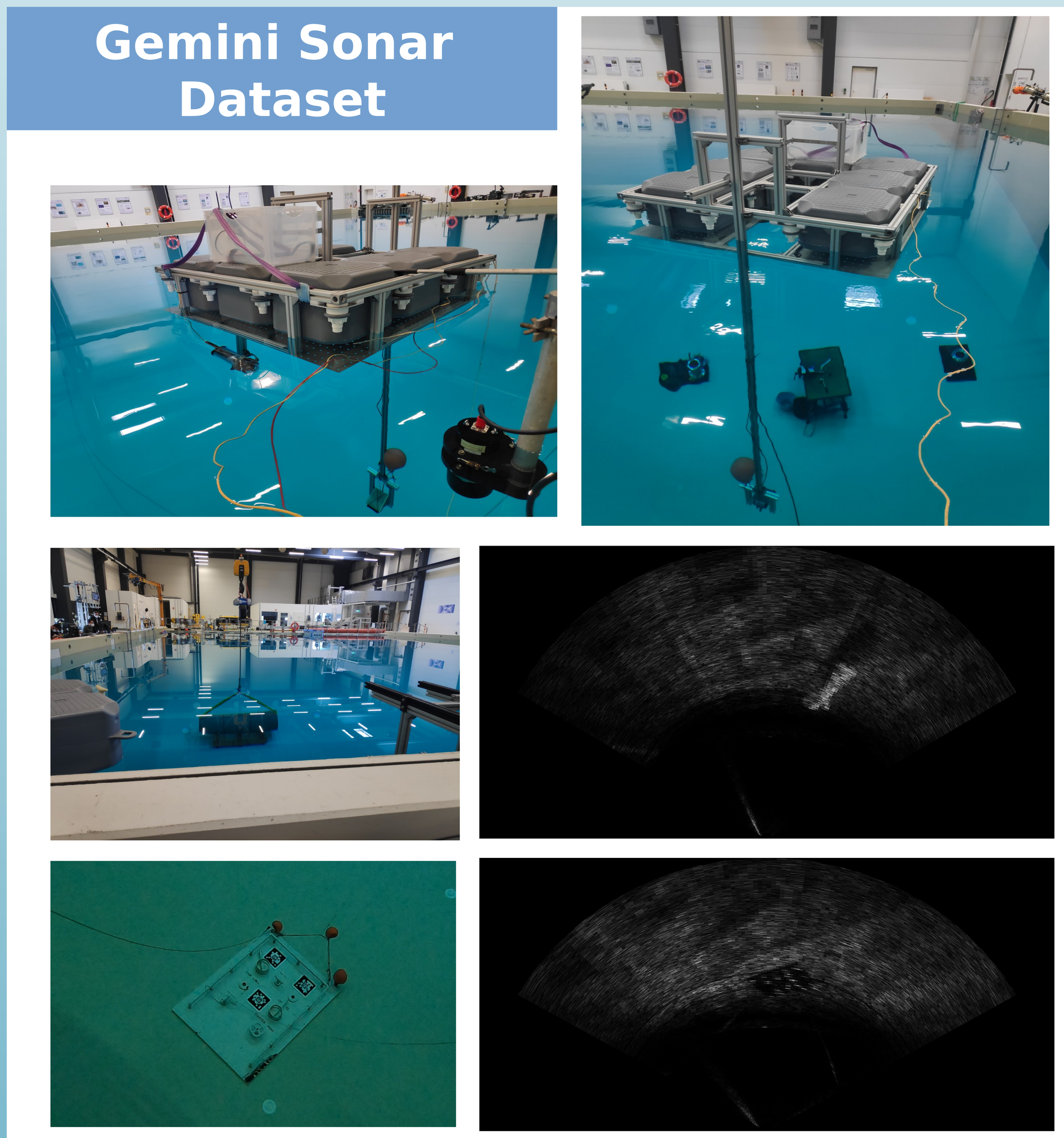


Figure. Evaluation of transfer learning on turntable dataset using few-samples per class. Each plot corresponds to a pretrained neural network via supervised and self-supervised pretraining.

Gemini Sonar Dataset



Features

- Paired multi-modal image data (sonar + camera)
- Tasks: classification, translation, object detection
- Custom ROS nodes for data collection

Conclusions and Future Work

- SSL pre-training has a similar performance against SL pre-training on sonar images
- Pretext tasks are showing positive results to overcome the problem of data annotation in sonar images
- Compare other SSL methods: Denoising Autoencoders, Jigsaw Puzzle, Contrastive methods
- Evaluate ssl pre-training in more tasks (image translation, object detection) with gemini dataset

Acknowledgments

- This work was supported by the project DeeperSense H2020-ICT-2020-2 ICT-47-2020 Project Number: 101016958.

References

[1] M. Valdenegro-Toro. Object recognition in forward-looking sonar images with convolutional neural networks. In OCEANS 2016 MTS/IEEE Monterey, pages 1–6, 2016. doi: 10.1109/OCEANS.2016.7761140.

[2] M. Valdenegro-Toro. Improving sonar image patch matching via deep learning. In 2017 European Conference on Mobile Robots (ECMR), pages 1–6, 2017. doi: 10.1109/ECMR.2017.8098701.

[3] Matias Valdenegro-Toro, Alan Preciado-Grijalva, Bilal Wehbe. Pre-trained Models for Sonar Images. Global OCEANS, 2021.

[4] Deepak Singh, Matias Valdenegro-Toro. The Marine Debris Dataset for Forward-Looking Sonar Semantic Segmentation. ArXiv:2108.06800, 2021.

[5] Spyros Gidaris, Praveer Singh, Nikos Komodakis. Unsupervised representation learning by predicting image rotations, 2018.

[6] Longlong Jing, Yingli Tian. Self-supervised visual feature learning with deep neural networks: A survey. CoRR, abs/1902.06162, 2019. URL http://arxiv.org/abs/1902.06162.

[7] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya, Banerjee, Fillia Makedon. A survey on contrastive self-supervised learning, 2021.