Self-Supervised Learning for Automated ECG Signal Quality Assessment

Dieko de Graaf[1 [0009−0009−2916−2381]], Maria Galanty[2,3 [0009−0001−2574−1775]], Hoel Kervadec[2,3 [0000−0002−6786−7042]], and Clara I. Sánchez[2,3 [0000−0001−9787−8319]]

[1] Informatics Institute, University of Amsterdam, The Netherlands
[2] Qualitative Healthcare Analysis (qurAI) Group, Informatics Institute, University of Amsterdam, The Netherlands
[3] Department of Biomedical Engineering and Physics, Amsterdam UMC, University of Amsterdam, The Netherlands

Abstract. This work explores self-supervised learning (SSL) for automated ECG signal quality assessment, a crucial task since low-quality signals can trigger false alarms in medical monitoring and weaken disease-detection algorithms. Existing approaches rely on annotated datasets, which are often limited in availability and inconsistent in their defined quality classes. Those limitations can be addressed by using SSL models to learn high-quality representations of ECG signals without requiring annotated data. The effectiveness of these representations for signal quality assessment is examined by evaluating three SSL models, SimCLR, BYOL, and SwAV. The Brno University of Technology ECG Quality Database is used, a single-lead dataset annotated for ECG signal quality. A comprehensive set of augmentations and parameters is explored for pre-training the SSL models and their performance is compared to supervised (KNN) and unsupervised (k-means) baselines. The results show that all SSL models outperform the baselines, with SwAV achieving the highest macro F1 on the test set with 81.85%. However, divergence between validation and test performance for low-quality signals reflects dataset limitations; nonetheless, the results demonstrate that SSL has great potential for ECG signal quality assessment.

Keywords: Self-supervised learning (SSL) · electrocardiogram (ECG) · signal quality assessment.

1 Introduction

The electrocardiogram (ECG) is a widely used, non-invasive method for capturing the electrical activity of the heart. It is integral to diagnosing various cardiovascular conditions, including arrhythmias, myocardial ischemia, and myocardial infarction [23, 20]. Due to its low cost and accessibility, the ECG is regularly applied in hospitals, intensive care, and ambulatory environments [20, 14].

Although ECG recordings are widely applied, they are frequently corrupted by physiological and technical noise sources. Typical types of noise include baseline wander, electrode motion, powerline interference, and muscle artifacts, which

can spectrally overlap with diagnostically relevant features such as the QRS complex or T wave [23, 18]. Consequently, distorted, low-quality ECG signals compromise both clinical interpretation and automated classification pipelines, leading to false alarms in continuous monitoring systems and reduced performance in machine learning models trained for cardiovascular classification tasks [25, 19].

The field of ECG signal quality assessment can be broadly divided into rule-based methods and machine learning approaches. Rule-based methods estimate quality by extracting signal quality indices, such as the detectable heart rate, RR interval duration, and QRS morphology indicators, and comparing these against predefined thresholds or morphology templates [16, 19]. The effectiveness of rule-based methods strongly depends on manually defined thresholds to assign classes to signals, and is therefore not robust across diverse settings and different noise types [18].

In contrast, machine learning methods learn to predict a quality class automatically from data. In the supervised setting, feature-based approaches extract statistical features from ECG signals [3] or from their decomposition, for example using wavelets that capture both temporal and spectral characteristics [2], and use them to train classifiers to distinguish between quality classes. More recently, supervised deep learning models have been applied to learn representations directly from raw ECG signals, capturing low-level characteristics that reflect signal quality and are therefore suitable for signal quality assessment without the need for explicit feature extraction [8, 22].

However, these existing approaches to ECG signal quality assessment have significant limitations. Supervised learning methods depend on labeled datasets to train models, which are scarce due to privacy constraints and the high cost of expert annotation [22]. Moreover, recordings are often of long duration, while signal quality must be assessed with high precision to identify even brief distortions. Noise and artifacts can manifest in various forms, and the responsibility falls on the annotator to correctly identify these interferences according to the defined quality classes, which makes the process prone to human error [22]. Furthermore, these datasets often differ in recording duration, acquisition protocols, and lead configurations, limiting the generalizability of trained models [8]. Annotated datasets specifically for ECG signal quality suffer from an additional challenge: there is no universally accepted definition of ECG signal quality, with differences in class granularity and annotation criteria between datasets [26, 15]. These inconsistencies limit model transferability and increase the risk of misclassification when applied outside the training domain [13].

Self-supervised learning (SSL) provides a promising alternative by enabling representation learning from unlabeled ECG signals. This eliminates the dependency on manual annotation and allows models to learn generalizable features. Recent work by Soltanieh et al. demonstrated that SSL models trained on unlabeled ECG data perform competitively on the downstream task of arrhythmia detection, even under distributional shifts [22]. These findings suggest that SSL

may also be effective in assessing the quality of ECG signals and therefore the method developed in our study is closely based on their approach.

This research aims to evaluate the effectiveness of self-supervised models for automated ECG signal quality assessment. To avoid the reliance on manually annotated quality labels, we explore whether SSL can capture representations that distinguish diagnostically useful signals from distorted ones. This approach has practical implications for both clinical monitoring, by reducing false alarms, and for machine learning applications, by improving the reliability of algorithms for disease detection.

2 Method

2.1 Dataset

The dataset used for this research is the Brno University of Technology ECG Quality Database (BUT-QDB) [15], which provides annotations on the signal quality of ECG recordings. This dataset consists of 18 single-lead ECG recordings collected with a Bittium Faros 180 Holter device. The recordings were acquired from 15 subjects (9 male, 6 female, aged 21–83 years) during daily activities with a sampling frequency of 1000 Hz. Each recording spans at least 24 hours, with three annotated for their full duration, and the remaining containing only two 20-minute annotated segments each. In addition, the dataset includes five supplementary annotated segments, including four of 20 minutes and one of 2 minutes, to increase the representation of low-quality signals in the dataset.

The annotation process was conducted manually by three ECG experts, and the dataset includes both their individual annotations and the consensus. They classify the quality of the ECG signals into three classes that are used throughout this research. Class 1: all significant waveforms (P wave, T wave, QRS complex) are clearly visible and detectable reliably, class 2: an increased noise level makes significant points unclear, but QRS complexes are still clearly visible and can be detected reliably, class 3: QRS complexes cannot be reliably detected, and the signal is unsuitable for further analysis. The consensus label was used as the true quality label. Examples of ECG signals belonging to each class are shown in Figure 1.

2.2 Data Preparation

The raw ECG signals were downsampled from 1000 Hz to 100 Hz and segmented into 2.5-second windows, following the method of Soltanieh et al., to reduce computational and memory requirements while retaining the relevant ECG bandwidth [22].

The dataset was divided into three sets: training, validation, and test. An additional labeled subset of the training set was used for supervised fine-tuning and linear evaluation, while the complete unlabeled training set was used for pre-training. The unlabeled training set was constructed by randomly sampling

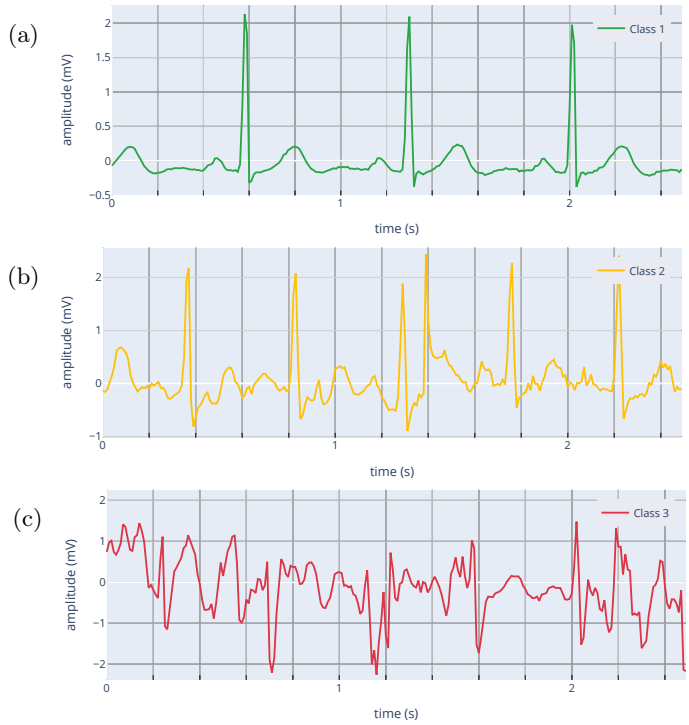


Fig. 1: Example ECG signals from BUT-QDB for the three defined signal quality classes: class 1 (a), class 2 (b) and class 3 (c).

windows from the training records. The labeled training subset and the labeled validation and test sets were sampled from the annotated segments within the recordings assigned to each split, using a sliding-window approach that respected class boundaries and excluded incomplete segments.

To prevent data leakage, the records were split by patient ID, which guarantees that all recordings of a single patient appeared in only one subset. The validation and test sets each comprise about 15% of annotated data, with patient assignments determined through an optimization procedure to balance both class size and proportion. Smaller evaluation sets resulted in unbalanced sets with too few class 3 samples, primarily due to the uneven recording durations, as only three recordings were fully annotated for 24 hours. A single patient, patient 105, contains nearly 90% of class 3 annotations and consists mainly of background measurements, therefore, it was exclusively assigned to the training set to avoid imbalance in the evaluation sets. The train, validation and test splits contained 7, 4 and 4 patients, respectively. The number of samples per data split are given in Table 1.

Table 1: Class distribution of the training, validation, and test splits. The complete train set for self-supervised pre-training includes the unannotated segments.

Data Split	**# Samples**	**Class 1**	**Class 2**	**Class 3**
Train - Complete	269346	-	-	-
Train - Annotated subset	60434	27405	14090	18939
Validation	39760	18661	19606	1493
Test	39380	26000	12250	1130

2.3 Baseline Methods

To benchmark the self-supervised models, two baseline methods without representation learning were implemented: supervised k-nearest neighbors (KNN) and unsupervised k-means clustering. Both are applied to cluster manually extracted features obtained through discrete wavelet transform (DWT), an established approach for creating features to cluster ECG signals [1, 10, 22].

DWT Feature Extraction Wavelet decomposition serves as an effective approach for analyzing non-stationary signals such as ECG. It enables the extraction of localized features that reflect both the timing and the spectral content of the signal, effectively isolating cardiac activity from noise components [10]. In this work, a four-level discrete wavelet transform is applied using the Symlet 5 (*sym5*) wavelet, shown in Figure 2. Its near-symmetry, in combination with structural similarity to ECG waveforms, provides a balanced resolution in the time and frequency domains [11], which makes it suitable for ECG analysis. Following Maleki and Haeri [10], each of the resulting five subbands is characterized by eight statistical features: entropy, number of zero crossings, number of mean crossings, median, mean, standard deviation, variance and root mean square. This results in 40-dimensional feature vectors for each 2.5-second ECG segment, which form the input for both baseline classifiers. The signals are decomposed using PyWavelets [9].

Baseline Classifiers The unsupervised k-means algorithm from scikit-learn [17] groups unlabeled DWT feature vectors into $k = 12$ clusters, where k is chosen by achieving the highest macro F1 score on the validation set over the range $3 \leq k \leq 20$. Centroids are initialized using the *k-means++* method and fit on the unlabeled training data. Since cluster labels are initially unknown, each cluster centroid is assigned to a quality class by computing its Euclidean distance to class centroids of the validation set. This approach avoids bias towards overrepresented classes.

The supervised k-nearest neighbors (KNN) classifier assigns class labels to ECG segments based on the majority vote among the k nearest neighbors in the 40-dimensional DWT feature space [7]. The value of $k = 3$ was selected to maximize the macro F1 score on the validation set over the range $1 \leq k \leq 20$.

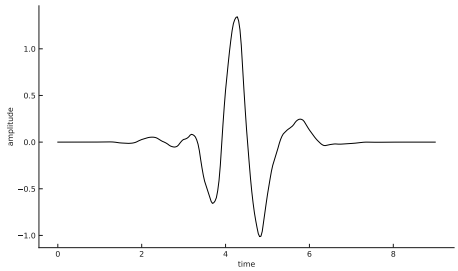


Fig. 2: The Symlet 5 wavelet function plotted over time. The signal is generated by using PyWavelets [9].

2.4 SSL Methods

Three SSL models were evaluated: SimCLR [5], BYOL [6], and SwAV [4].

SimCLR (Figure 3a) is a contrastive learning method which learns by distinguishing between positive pairs (augmented variants of the same input sample) and negative pairs (augmented variants of different samples). Its performance depends on the availability of proper negatives and carefully chosen augmentations, to ensure that the separation task is sufficiently challenging.

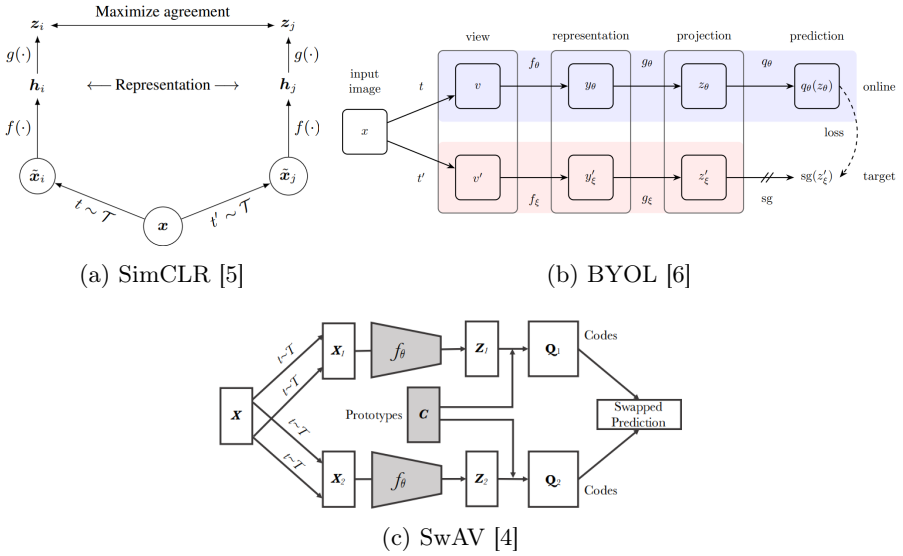
BYOL (Figure 3b) circumvents SimCLR's limitations as it eliminates the need for negative samples by training an online network to predict the representation of a target network, which encodes a different augmentation of the same input. The target network is updated as an exponential moving average of the online weights, making it a stable objective. This approach avoids reliance on large batches of negatives and reduces sensitivity to the augmentation choice.

SwAV (Figure 3c) takes an alternative approach by combining contrastive learning with clustering, instead of contrasting representation pairs, it groups similar representations with a codebook and uses these codes for solving a contrastive learning problem. This allows for stable and efficient representation learning, without depending on traditional instance comparison.

Augmentations The performance of the SSL models depends on the choice of augmentations applied to the ECG signal. Soltanieh et al. evaluated seven augmentations specifically for ECG representation learning [21, 22]. They also tested combining multiple augmentations but reported a decrease in performance; therefore, this is not repeated in our approach.

Each augmentation is tested with three parameter settings, ordered by increasing strength (i.e., intensity), except for *Negation*, which has no adjustable parameters and is applied once per model. The augmentations and their parameters are the following:

1. **Gaussian Noise (GN)** adds zero-mean Gaussian noise $N(t)$ with standard deviations $\sigma_G = [0.01, 0.1, 1]$ to the ECG signal.



(a) SimCLR [5]

(b) BYOL [6]

(c) SwAV [4]

Fig. 3: Architectures of the SSL models: (a) SimCLR, (b) BYOL, and (c) SwAV.

2. **Channel Resize (ChR)** scales each ECG channel by a factor s uniformly sampled from ranges $[0.33, 1]$, $[0.5, 2]$, and $[0.33, 3]$.
3. **Negation (Neg)** vertically flips the signal at each timestep t: $\tilde{x}(t) = -x(t)$.
4. **Baseline Wander (BlW)** adds a low-frequency sinusoid scaled by $S_{bw} = [0.1, 0.7, 1]$ to simulate baseline drift.
5. **Electromyographic Noise (EM)** simulates muscle noise with zero-mean Gaussian noise of standard deviation $\sigma_{EMG} = [0.01, 0.5, 1]$.
6. **Time Out (TO)** masks a segment of the signal by zeroing a percentage uniformly sampled from ranges $[10\%, 20\%]$, $[0\%, 50\%]$, and $[40\%, 50\%]$.
7. **Dynamic Time Warping (DTW)** splits the signal into $2w$ segments, stretching half by $r\%$ and compressing the other half by the same amount. Tested parameter pairs are $(w, r) = (1, 10)$, $(3, 5)$, and $(3, 10)$.

Training details The *xResNet1d50* backbone model was used during both the pre-training and fine-tuning stages of the SSL models; it is a one-dimensional adaptation of the *ResNet50* architecture and is known for its reliability and consistent effectiveness in ECG representation learning [22].

For each augmentation and parameter combination, a model was pre-trained three times and both linearly evaluated and fine-tuned with a different seed. The reported scores are the average of the three runs.

During pre-training, the complete training set was used and the corresponding loss function for each model was optimized. The best model was selected based on achieving the lowest loss on the validation set. The models were trained for 150 epochs using the Adam optimizer with a batch size of 4096 for SimCLR

and BYOL, and 1024 for SwAV due to computational constraints. The learning rate was set to 5×10^{-4} and a weight decay of 1×10^{-3}. A linear warm-up of 10 epochs was followed by cosine annealing, gradually decreasing the learning rate to a final value of 5×10^{-5}.

During downstream evaluation through linear evaluation and fine-tuning, the labeled subset of the training data was used to train for 50 epochs. Optimization was carried out using cross-entropy loss with softmax activation. Like in pre-training, the linear warm-up was followed by cosine decay, and the same learning rate was used for the classification head, while for fine-tuning a lower learning rate of 5×10^{-5} was used for the backbone to prevent overwriting pre-trained parameters. Furthermore, discriminative learning rate was applied where the learning rate is halved in the consecutive deeper layers of the backbone.

2.5 Evaluation Metrics

The performance evaluation of the SSL models and baselines was carried out using standard classification metrics: precision, recall, and the F1 score. When evaluating across classes, macro-averaged scores are reported. This unweighted average of the per-class scores is necessary when handling the minority class of low-quality signals, as reflected in the class distributions of the different data splits in Table 1. Accuracy is omitted because it is dominated by the majority class and provides no additional information beyond the reported recall. The metrics are reported with the standard deviations over three runs, each initialized with a different random seed.

Finally, to qualitatively evaluate the structure of the learned representation space of the SSL models, UMAP was applied for dimensionality reduction, projecting the high-dimensional representations into two dimensions for visualization [12]. This approach was also used by Soltanieh et al., for detecting distributional shifts between datasets through analyzing the representation spaces of self-supervised models [22].

2.6 Hardware and Codebase

All experiments were conducted on the Dutch National Supercomputer Snellius [24], using two NVIDIA A100 GPUs and 36 CPU cores per job. Multiple jobs were executed in parallel. Since the codebase of Soltanieh et al. [22] was not available, we extensively adapted the 1D implementation of self-supervised models for ECG by Mehari and Strodthoff [13] to follow the method of Soltanieh et al. The complete codebase is available on GitHub.[4]

[4] `https://github.com/diekodegraaf/ssl-ecg-quality`

Table 2: Test performance of fine-tuned SSL models compared to the baselines. Reported values are macro averaged F1, precision, and recall (%, mean ± std over 3 seeds), along with per-class results for class 3. SSL augmentations are selected based on the highest macro F1 after fine-tuning.

	Augmentation	**Macro average**			**Class 3**		
Method	**Parameters**	F1	Precision	Recall	F1	Precision	Recall
SimCLR	Channel Resize $[0.33, 1.0]$	80.74 ± 1.57	88.57 ± 0.91	75.99 ± 1.52	77.03 ± 4.85	99.82 ± 0.16	62.89 ± 6.62
BYOL	Channel Resize $[0.33, 1.0]$	80.56 ± 1.50	88.80 ± 0.71	75.56 ± 1.60	76.23 ± 5.10	99.47 ± 0.43	61.98 ± 6.77
SwAV	Time Out $[0.1, 0.2]$	$\mathbf{81.85 \pm 0.97}$	$\mathbf{88.94 \pm 0.67}$	$\mathbf{77.27 \pm 0.96}$	79.56 ± 3.24	97.87 ± 2.74	67.20 ± 5.20
KMeans	-	41.88 ± 1.29	42.56 ± 0.53	43.92 ± 2.32	16.74 ± 3.87	13.18 ± 2.36	23.13 ± 7.57
KNN	-	64.76	71.68	60.91	**82.48**	**100.00**	**70.18**

Table 3: Validation performance of fine-tuned SSL models compared to the baselines. Reported values are macro-averaged F1, precision, and recall (%, mean ± std over 3 seeds), along with per-class results for class 3. Validation results are reported using the same augmentations identified as optimal on the test set.

	Augmentation	**Macro average**			**Class 3**		
Method	**Parameters**	F1	Precision	Recall	F1	Precision	Recall
SimCLR	Channel Resize $[0.33, 1.0]$	53.89 ± 0.75	73.98 ± 5.94	54.29 ± 0.59	8.06 ± 2.01	63.42 ± 16.96	4.33 ± 1.15
BYOL	Channel Resize $[0.33, 1.0]$	54.77 ± 2.87	$\mathbf{78.67 \pm 2.69}$	55.11 ± 2.33	8.35 ± 4.62	$\mathbf{75.40 \pm 4.97}$	4.49 ± 2.64
SwAV	Time Out $[0.1, 0.2]$	$\mathbf{58.41 \pm 1.31}$	69.94 ± 8.39	58.54 ± 1.54	15.47 ± 4.92	46.45 ± 24.07	12.46 ± 9.34
KMeans	-	48.22 ± 0.38	48.20 ± 0.34	$\mathbf{59.74 \pm 0.31}$	$\mathbf{26.56 \pm 0.86}$	16.67 ± 0.83	$\mathbf{65.57 \pm 2.46}$
KNN	-	39.08	49.01	40.78	1.31	27.78	0.67

3 Results

3.1 Performance Overview

The performance of the fine-tuned SSL models and the supervised (KNN) and unsupervised (k-means) baselines on the test and validation sets is presented in Table 2 and Table 3, respectively. The SSL augmentation is selected by the highest macro F1 on the test set and is reported identically for validation. The reported performance metrics (F1, precision, and recall) are expressed as percentages with standard deviations computed over three random seeds. KNN is deterministic and therefore reported without standard deviations. Both macro averages and per-class scores for class 3 are included. Expanded tables that include the performance on classes 1 and 2 can be found in Appendix A.1, and comprehensive SSL results for each augmentation and parameter, both for fine-tuning and linear evaluation, are provided in Appendix A.2 and A.3.

3.2 Augmentation Performance after Linear Evaluation and Fine-Tuning

The performance on the test set for all SSL augmentations under linear evaluation is illustrated in Figure 4a and after fine-tuning in Figure 4b. Each figure presents 19 configurations per model architecture, with every configuration representing a model pre-trained using a specific augmentation and parameter setting. The F1 score for class 3 is plotted against the macro combined F1 score of class 1 and class 2. This highlights the model's ability to distinguish between both low-quality ECG signals from class 3 and more informative signals from class 1 and class 2. Performance is reported on the test set as the mean across three random seeds, analogous illustrations for the validation set can be found in Appendix B.

3.3 Low-Quality Signals in the SSL Representation Space

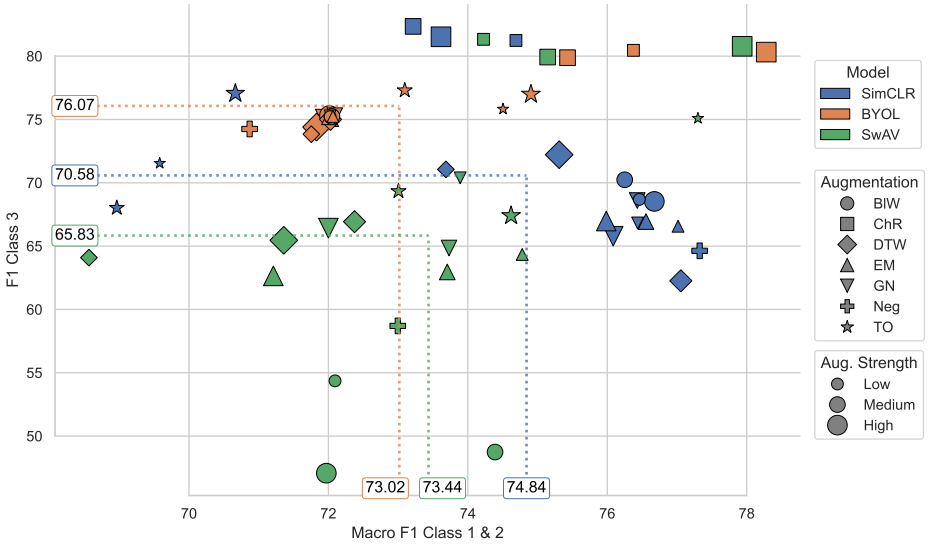
To visualize how the self-supervised models encode the ECG signals from the training, validation and test data, UMAP projections of the representation space of BYOL are shown in Figure 5a. It illustrates the distribution of class 3 signals for the validation and test sets. These are overlaid on the representation of the annotated training subset, where classes 1 and 2 are shown together, and class 3 is plotted separately. An additional projection is shown in Figure 5b, where validation and test samples are colored based on correct or incorrect predictions under linear evaluation. BYOL is selected due to its largest performance gap between validation and test in terms of class 3 F1, analogous figures for SimCLR and SwAV are provided in Appendix C.

The evaluation result of BYOL after linear evaluation is expanded by inspecting the signals of the different clusters in the representation space. This is presented in Figure 6, where the plotted class 3 signals correspond to the highlighted points in the UMAP projection of the representation space, based on their shared color.

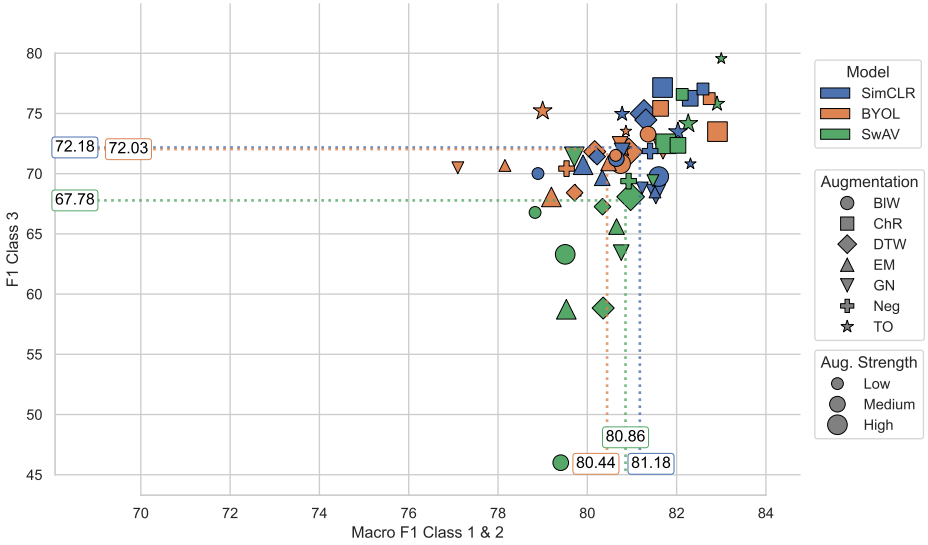
4 Discussion

4.1 Performance Across Validation and Test Sets

On the test set, the SSL models achieve high F1 scores above 80%, with SwAV performing best at 81.85%, followed by SimCLR (80.74%) and BYOL (80.56%). All three models show higher precision than recall, caused by their class 3 performance. For this class, the SSL models reach near perfect precision, ranging from 97.87% to 99.82%, but lower recall in the range of 61.98-67.20%, indicating that they very reliably detect only a subset of class 3 signals in the test data. The recall of all models shows considerable variation across seeds, suggesting that class 3 detection is sensitive to stochastic training effects. The SSL models achieve similar competitive scores for classes 1 and 2 signals, as shown in Appendix A.1.

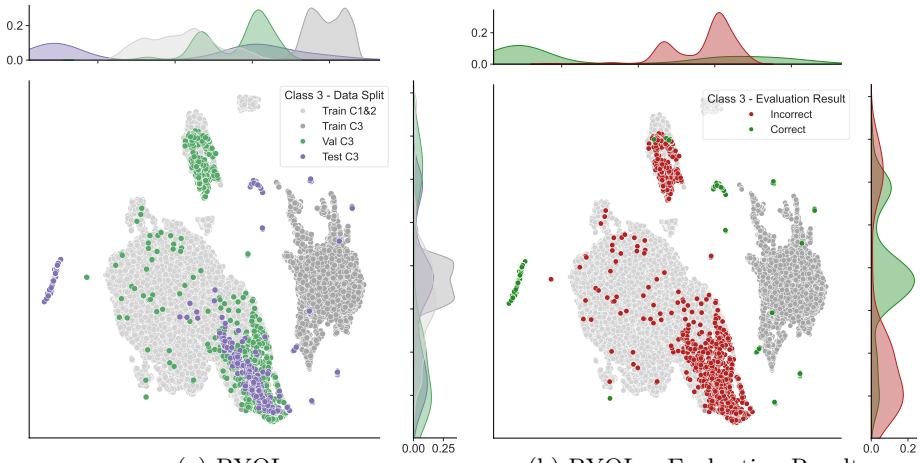


(a) Linear evaluation.



(b) Fine-tuned evaluation.

Fig. 4: Comparison of testing performance for SSL models and different augmentations through plotting the macro F1 score for class 1 and 2 (x-axis) against the F1 score for class 3 (y-axis) after linear (a) and fine-tuned (b) evaluation. The marker size indicates the augmentation strength while the marker shape encodes the augmentation type. The dotted lines and values plotted on the axis are the average scores for each model across the different configurations.



(a) BYOL (b) BYOL – Evaluation Result

Fig. 5: 2D UMAP projections of class 3 representations for BYOL. Validation and test signals are overlaid on the annotated training subset to show overlap, with results (b) based on linear evaluation.

Interestingly, KNN outperforms all SSL models on class 3, reaching an F1 of 82.48% with perfect precision and recall of 70.18%. This can be attributed to its reliance on neighborhood similarity, indicating that class 3 test signals have DWT features closely aligned with the training distribution. In contrast, k-means similarly uses the DWT feature vectors but performs poorly, with a class 3 F1 of only 16.74%. This discrepancy arises from the label assignment procedure in k-means, which maps clusters to classes based on the class centroids in the validation data.

This misalignment of low-quality signals is reflected in the validation performance, where macro F1 remains below 60% for all SSL models, caused by particularly weak class 3 recall. KNN faces similar difficulties, reporting a class 3 F1 of only 1.31%. The exception is k-means, which achieves 26.56% F1 with a recall of 65.57%, suggesting overfitting to the validation distribution through cluster assignment, however, the low precision indicates that this does not result in good class separation. Collectively, these results conflict with the class 3 scores on the test set and suggest that the test samples are more representative of the training data. In contrast, the validation set likely contains class 3 signals with different characteristics than those in the test set. This could explain why the class 3 signals in the validation set are more difficult to detect, and this distribution shift is further analyzed in Section 4.3 of the Discussion.

4.2 Influence of Augmentation and Evaluation Strategy on SSL

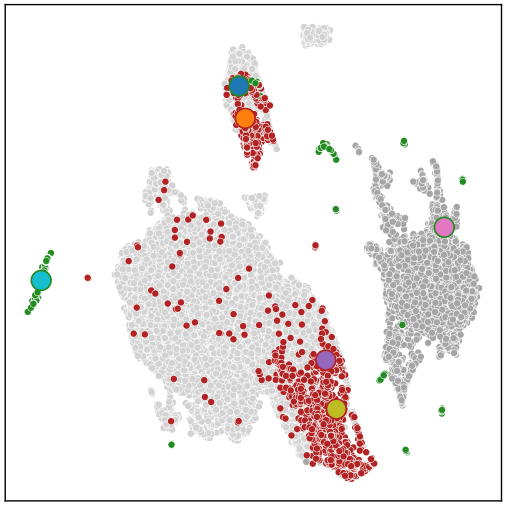
Compared to linear evaluation, fine-tuning reduces the difference between augmentations, resulting in similar means across the models. Although fine-tuning consistently improves the class 1 and 2 macro F1 across models, it lowers the

best-performing class 3 results compared to linear evaluation. This can be attributed to model selection being based on the validation set, where the class 3 content differs from the test distribution; consequently, all validation metrics increased after fine-tuning (Appendix B). BYOL shows tightly grouped results under linear evaluation, suggesting low sensitivity to augmentation choice, while still achieving competitive performance across augmentations. In contrast, SwAV demonstrates the strongest sensitivity to augmentation choice, with results ranging from the least to the most successful. The linear evaluation results indicate that the most effective augmentation on the test set is *Channel Resize* (ChR), with *Time Out* (TO) achieving strong performance on class 3. No consistent relationship is observed between augmentation strength and performance, indicating that the type of augmentation is of higher importance than its intensity.

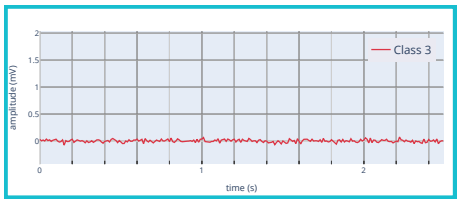
4.3 Representation Space and Distribution Shifts

The 2D UMAP projections of the learned representation space in Figure 5 show similar trends across models and reveal that the class 3 validation and test samples form clusters mostly isolated from the class 3 training distribution. It should be noted that visualized training data represent only the annotated subset; unlabeled training data may also contain low-quality signals, but these would not add interpretive value to the plot given the absence of labels. The greater overlap of the test set with the class 3 training representations suggests that similar low-quality signals were encountered during pre-training, allowing the models to distinguish them more effectively from the other classes during evaluation. Furthermore, the test samples are positioned in multiple unique clusters in the representation space, which are correctly associated with class 3 as shown for BYOL in Figure 5b, while the validation set overlaps with the training representations of classes 1 and 2, leading to more incorrect predictions. The higher separation of the test representations and its increased overlap with the class 3 training distribution help explain the improved performance on test data compared to the validation data.

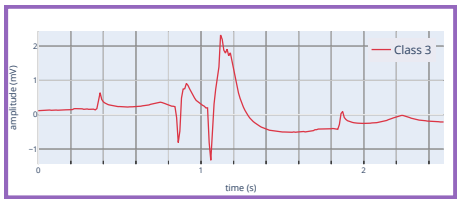
Inspection of the signals corresponding to the learned representations shows that the smaller elongated clusters of test samples primarily consist of low-amplitude background measurements (Figure 6b), which match the content of class 3 signals in the annotated training subset (Figure 6d). The test set also contains more complex low-quality signals that overlap with training classes 1 and 2, but are often misclassified, as illustrated in Figure 6f. In contrast, validation signals are located more frequently in regions with a stronger overlap with training classes 1 and 2, as seen in Figures 6c and 6e, both of which are misclassified. However, in the case of Figure 6g, a validation signal is correctly identified as class 3 despite overlapping with training classes 1 and 2. Overall, relating the BYOL representation to the others, these observations suggest that the SSL models detect background measurements as belonging to class 3 but are less able to separate more nuanced low-quality signals from the other classes. The performance on the test set is therefore overestimated by the high presence of homogeneous low-quality signals.



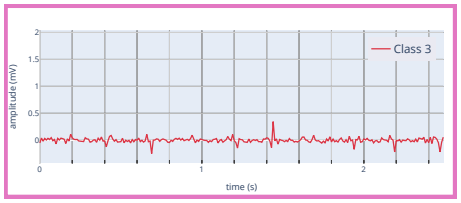
(a) BYOL linear evaluation result: six signals highlighted in the representation space



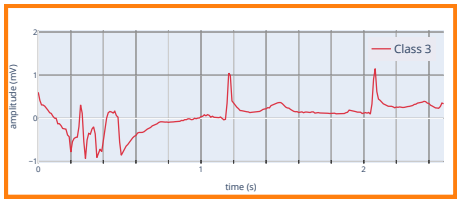
(b) Test class 3 signal correctly classified, from a separated cluster of test signals



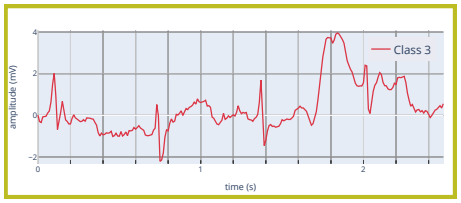
(c) Validation class 3 signal misclassified as class 1, overlapping with train classes 1 and 2



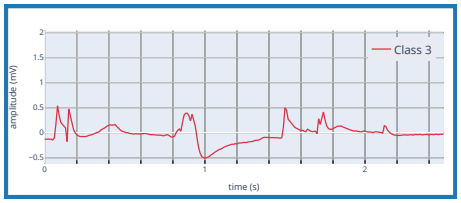
(d) Test class 3 signal correctly classified, overlapping with train class 3



(e) Validation class 3 signal misclassified as class 1, overlapping with train classes 1 and 2



(f) Test class 3 signal misclassified as class 2, overlapping with train classes 1 and 2



(g) Validation class 3 signal correctly classified, overlapping with train classes 1 and 2

Fig. 6: Six class 3 ECG signals are shown in BYOL's 2D UMAP space, including three test and three validation examples. The colors of the highlighted points in (a) correspond to the colored frames of the signal plots.

5 Conclusion

This work showed that self-supervised learning models can learn representations suited for ECG signal quality assessment without requiring labeled data during pre-training. On the test set, all SSL models achieved macro F1 scores above 80%, with SwAV performing best. Their performance derives from near-perfect precision on low-quality signals, although recall remained lower, indicating that the SSL models very reliably detect a specific subset of class 3 signals in the test set. In contrast, validation results showed weaker performance, particularly for class 3, revealing that model reliability is strongly dependent on the representativeness of the evaluation data. For both evaluation sets, the SSL models consistently distinguished between class 1 and class 2 signals, suggesting that their learned representations encode the subtle differences necessary for their separation, which highlights the potential of these models.

Analysis of the representation space confirmed that test data contained more homogeneous low-quality signals resembling those seen during pre-training, while the validation set included more complex class 3 signals that overlapped with the other classes and are therefore more difficult to detect. This discrepancy reflects a distribution shift resulting from the dataset itself: class 3 is defined as a residual class which covers a wide range of noise sources, is sparsely represented and unevenly distributed across patients. The imbalance between patients makes that this issue cannot be resolved by merely reassigning patients between splits while preventing data leakage. These limitations of the dataset lead to overestimated class 3 performance on the test set.

Therefore, while the representation learning of the SSL models shows potential, the current performance on class 3 is unreliable due to the predominance of homogeneous low-quality signals in the test set and their lower frequency in the validation set. A reliable estimate of performance on low-quality signals will require more recordings from different patients in the evaluation data. Mainly, the inclusion of more low-quality recordings would make the evaluation process less biased toward specific occurrences of noise and background measurements. Alternatively, the synthesis of low-quality signals can be explored to artificially increase their variety in the evaluation data. Although the results are promising, further research is needed before these models can be confidently applied in clinical settings for the detection of low-quality signals.

Bibliography

[1] Ajitha Gladis, K.P., Ahilan, A., Muthukumaran, N., Jenifer, L.: Ecg arrhythmia measurement and classification for portable monitoring. Measurement Science Review **24**(4), 118–128 (2024). https://doi.org/10.2478/msr-2024-0017

[2] Albaba, A., Simões, N., Wang, Y., Hendriks, R., De Raedt, W., Van Hoof, C.: Assessing the signal quality of electrocardiograms from varied acquisition sources. Computers in Biology and Medicine **133**, 104164 (2021). https://doi.org/10.1016/j.compbiomed.2020.104164

[3] Behar, J., Oster, J., Li, Q., Clifford, G.D.: Ecg signal quality during arrhythmia and its application to false alarm reduction. IEEE transactions on biomedical engineering **60**(6), 1660–1666 (2013). https://doi.org/10.1109/TBME.2013.2240452

[4] 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). https://doi.org/10.48550/arXiv.2006.09882

[5] 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 (2020). https://doi.org/10.48550/arXiv.2002.05709

[6] Grill, J.B., Strub, F., Altché, F., Tallec, C., Richemond, P.H., Buchatskaya, E., Doersch, C., Pires, B.A., Guo, Z.D., Azar, M.G., et al.: Bootstrap your own latent: A new approach to self-supervised learning. In: 34th International Conference on Neural Information Processing Systems. pp. 21271–21284 (2020). https://doi.org/10.48550/arXiv.2006.07733

[7] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2nd edn. (2009). https://doi.org/10.1007/978-0-387-84858-7

[8] Hong, S., Zhou, Y., J, S., Xiao, C., Sun, J.: Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. Computers in Biology and Medicine **122**, 103801 (2020). https://doi.org/10.1016/j.compbiomed.2020.103801

[9] Lee, G., Gommers, R., Wohlfahrt, K., Wasilewski, F., O'Leary, A., Holger, Khetarpal, A., Millman, J., Sauvé, A., Clauss, C., Agrawal, A., Pelt, D.M., Oliveira, H., Yu, F., Brett, M., Pelletier, M., SylvainLan, Tricoli, D., Tuay, E.A.M., Burovski, E., Choudhary, S., Solak, A.C., asnt, Smith, A., rthr rllr, Schneck, C., Goldberg, C., Goertzen, D., Stephan, D.: Pywavelets/pywt: v1.8.0 (dec 2024). https://doi.org/10.5281/zenodo.14278642

[10] Maleki, M., Haeri, F.: Identification of cardiovascular diseases through ecg classification using wavelet transformation. arXiv preprint arXiv:2404.09393 (2024). https://doi.org/10.48550/arXiv.2404.09393

[11] Mallat, S.G.: A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **11**(7), 674–693 (1989). https://doi.org/10.1109/34.192463

[12] McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018). https://doi.org/10.48550/arXiv.1802.03426

[13] Mehari, T., Strodthoff, N.: Self-supervised representation learning from 12-lead ecg data. Computers in Biology and Medicine **141**, 105114 (2022). https://doi.org/10.1016/j.compbiomed.2021.105114

[14] Nazzari, H., Halperin, L., Krahn, A.D.: Electrocardiographic monitoring strategies (holter, implantable loop recorder, in between). In: Vasan, R., Sawyer, D. (eds.) Encyclopedia of Cardiovascular Research and Medicine, pp. 197–210. Elsevier (2018). https://doi.org/10.1016/B978-0-12-809657-4.99770-9

[15] Nemcova, A., Smisek, R., Opravilová, K., Vitek, M., Smital, L., Maršánová, L.: Brno university of technology ecg quality database (but-qdb). PhysioNet **101**, e215–e220 (2020). https://doi.org/10.13026/qcnm-a155

[16] Orphanidou, C., Bonnici, T., Charlton, P., Clifton, D., Vallance, D., Tarassenko, L.: Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring. IEEE Journal of Biomedical and Health Informatics **19**(3), 832–838 (2015). https://doi.org/10.1109/JBHI.2014.2338351

[17] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

[18] Rahman, S., Karmakar, C., Natgunanathan, I., Yearwood, J., Palaniswami, M.: Robustness of electrocardiogram signal quality indices. Journal of the Royal Society Interface **19**(189), 20220012 (2022). https://doi.org/10.1098/rsif.2022.0012

[19] Satija, U., Ramkumar, B., Manikandan, M.S.: Automated ecg noise detection and classification system for unsupervised healthcare monitoring. IEEE Journal of Biomedical and Health Informatics **22**(3), 722–732 (2017). https://doi.org/10.1109/JBHI.2017.2686436

[20] Serhani, M.A., Kassabi, H.T.E., Ismail, H., Navaz, A.N.: Ecg monitoring systems: Review, architecture, processes, and key challenges. Sensors **20**(6), 1796 (2020). https://doi.org/10.3390/s20061796

[21] Soltanieh, S., Etemad, A., Hashemi, J.: Analysis of augmentations for contrastive ecg representation learning. In: 2022 International Joint Conference on Neural Networks (IJCNN). pp. 1–10. IEEE (2022). https://doi.org/10.1109/IJCNN55064.2022.9892600

[22] Soltanieh, S., Hashemi, J., Etemad, A.: In-distribution and out-of-distribution self-supervised ecg representation learning for arrhythmia detection. IEEE Journal of Biomedical and Health Informatics **28**(2), 789–800 (2024). https://doi.org/10.1109/JBHI.2023.3331626

[23] Sornmo, L., Laguna, P.: Bioelectrical Signal Processing in Cardiac and Neurological Applications, pp. 411–510. Elsevier (2005). https://doi.org/10.1016/B978-0-12-437552-9.X5000-4

[24] SURF: Snellius: the national supercomputer (2025), [Online; Accessed 2025-08-22]

[25] Tsimenidis, C., Murray, A.: False alarms during patient monitoring in clinical intensive care units are highly related to poor quality of the monitored electrocardiogram signals. Physiological Measurement **37**(8), 1383–1391 (2016). https://doi.org/10.1088/0967-3334/37/8/1383

[26] Wagner, P., Strodthoff, N., Bousseljot, R.D., Kreiseler, D., Lunze, F.I., Samek, W., Schaeffter, T.: PTB-XL, a large publicly available electrocardiography dataset. Scientific Data **7**(1), 154 (2020). https://doi.org/10.1038/s41597-020-0495-6

A Complete Experimental Results

A.1 Per-Class Performance

Table 4: Test performance of fine-tuned SSL models compared to the baselines. Reported values are macro-averaged F1, precision, and recall (%, mean ± std over 3 seeds). SSL augmentations are selected based on the highest macro F1 after fine-tuning.

Method	**Augmentation**	**Macro average**			**Class 1**			**Class 2**			**Class 3**		
		F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall
SimCLR	Channel Resize [0.33, 1.0]	80.74 ± 1.57	88.57 ± 0.91	75.99 ± 1.52	89.74 ± 0.46	88.05 ± 1.54	91.56 ± 2.67	75.44 ± 0.64	77.85 ± 4.33	73.52 ± 4.75	77.03 ± 4.85	99.82 ± 0.16	62.89 ± 6.62
BYOL	Channel Resize [0.33, 1.0]	80.56 ± 1.50	88.80 ± 0.71	75.56 ± 1.60	90.02 ± 0.24	87.53 ± 1.27	92.69 ± 1.66	75.44 ± 0.87	79.40 ± 2.95	72.02 ± 3.73	76.23 ± 5.10	99.47 ± 0.43	61.98 ± 6.77
SwAV	Time Out [0.1, 0.2]	81.85 ± 0.97	88.94 ± 0.67	77.27 ± 0.96	90.36 ± 0.26	87.35 ± 2.31	93.66 ± 2.40	75.64 ± 2.47	81.59 ± 3.88	70.96 ± 6.92	79.56 ± 3.24	97.87 ± 2.74	67.20 ± 5.20
KMeans	-	41.88 ± 1.29	42.56 ± 0.53	43.92 ± 2.32	77.17 ± 0.51	71.94 ± 1.28	83.24 ± 1.13	31.75 ± 0.43	42.56 ± 2.50	25.39 ± 1.28	16.74 ± 3.87	13.18 ± 2.36	23.13 ± 7.57
KNN	-	64.76	71.68	60.91	76.94	71.31	83.54	34.87	43.72	29.00	82.48	100.00	70.18

Table 5: Validation performance of fine-tuned SSL models compared to the baselines. Reported values are macro-averaged F1, precision, and recall (%, mean ± std over 3 seeds). Validation results are reported using the same augmentations identified as optimal on the test set.

Method	**Augmentation**	**Macro average**			**Class 1**			**Class 2**			**Class 3**		
	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	
SimCLR	Channel Resize [0.33, 1.0]	53.89 ± 0.75	73.98 ± 5.94	54.29 ± 0.59	80.90 ± 0.72	69.33 ± 1.21	97.13 ± 0.31	72.71 ± 1.55	89.19 ± 0.62	61.41 ± 2.37	8.06 ± 2.01	63.42 ± 16.96	4.33 ± 1.15
BYOL	Channel Resize [0.33, 1.0]	54.77 ± 2.87	78.67 ± 2.69	55.11 ± 2.33	81.78 ± 2.54	70.44 ± 4.07	97.62 ± 0.48	74.17 ± 4.56	90.18 ± 0.46	63.22 ± 6.97	8.35 ± 4.62	75.40 ± 4.97	4.49 ± 2.64
SwAV	Time Out [0.1, 0.2]	58.41 ± 1.31	69.94 ± 8.39	58.54 ± 1.54	83.68 ± 1.61	73.80 ± 3.00	96.69 ± 0.80	76.10 ± 4.74	89.59 ± 1.27	66.46 ± 7.84	15.47 ± 4.92	46.45 ± 24.07	12.46 ± 9.34
KMeans	-	48.22 ± 0.38	48.20 ± 0.34	59.74 ± 0.31	67.29 ± 1.09	63.62 ± 1.47	71.58 ± 4.16	50.80 ± 1.67	64.32 ± 1.64	42.08 ± 2.97	26.56 ± 0.86	16.67 ± 0.83	65.57 ± 2.46
KNN	-	39.08	49.01	40.78	65.50	55.47	79.96	50.43	63.77	41.71	1.31	27.78	0.67

A.2 Performance of SSL Models on Validation Set

Table 6: Validation Macro and Class 3 performance under linear evaluation, with abbreviations augmentation, parameters and strength. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Macro			Class 3		
1	**Simclr**	BlW	C=0.1	1	46.31 ± 1.16	53.40 ± 3.57	47.96 ± 1.22	4.66 ± 0.76	11.13 ± 8.99	3.86 ± 1.74
2	**Simclr**	BlW	C=0.7	2	46.42 ± 0.93	54.51 ± 5.09	48.12 ± 0.93	4.91 ± 0.84	14.24 ± 13.19	3.71 ± 1.42
3	**Simclr**	BlW	C=1.0	3	46.66 ± 1.08	55.01 ± 6.81	48.18 ± 1.07	4.87 ± 1.13	16.12 ± 18.17	3.46 ± 0.25
4	**Simclr**	ChR	cr=[0.33,1.0]	1	47.98 ± 3.19	55.30 ± 2.28	51.99 ± 1.58	13.93 ± 7.69	11.81 ± 10.74	28.18 ± 8.18
5	**Simclr**	ChR	cr=[0.5,2.0]	2	46.72 ± 3.60	54.53 ± 3.93	49.07 ± 2.84	11.13 ± 9.12	12.90 ± 12.67	14.98 ± 10.83
6	**Simclr**	ChR	cr=[0.33,3.0]	3	46.04 ± 2.11	54.60 ± 2.45	48.80 ± 2.68	8.44 ± 3.76	13.29 ± 9.25	13.28 ± 14.24
7	**Simclr**	DTW	w=1,r=10	1	45.69 ± 1.81	56.51 ± 8.15	47.54 ± 1.26	4.74 ± 2.32	22.16 ± 22.70	2.75 ± 1.14
8	**Simclr**	DTW	w=3,r=5	2	46.58 ± 0.27	52.27 ± 1.39	48.21 ± 0.36	4.25 ± 0.96	8.01 ± 4.07	3.04 ± 0.40
9	**Simclr**	DTW	w=3,r=10	3	46.10 ± 1.53	56.01 ± 7.95	47.94 ± 0.98	4.49 ± 3.11	18.81 ± 23.19	2.92 ± 1.84
10	**Simclr**	EM	var=0.01	1	46.08 ± 1.32	52.89 ± 4.27	47.82 ± 1.27	3.41 ± 0.75	10.31 ± 10.18	2.39 ± 0.73
11	**Simclr**	EM	var=0.5	2	45.95 ± 1.15	53.29 ± 4.73	47.65 ± 1.20	4.10 ± 1.63	11.38 ± 12.19	3.46 ± 2.34
12	**Simclr**	EM	var=1.0	3	46.23 ± 1.72	53.22 ± 3.81	47.89 ± 1.53	4.98 ± 0.20	11.13 ± 9.34	3.82 ± 0.77
13	**Simclr**	GN	var=0.01	1	46.05 ± 1.59	53.44 ± 4.65	47.78 ± 1.42	4.26 ± 1.89	11.46 ± 11.74	3.39 ± 2.33
14	**Simclr**	GN	var=0.1	2	46.15 ± 1.17	51.53 ± 1.16	47.78 ± 1.19	4.25 ± 0.71	6.16 ± 2.22	3.68 ± 1.34
15	**Simclr**	GN	var=1.0	3	45.57 ± 1.16	51.95 ± 2.43	47.42 ± 1.10	3.83 ± 1.17	7.87 ± 5.44	2.84 ± 1.22
16	**Simclr**	Neg	True	2	45.97 ± 0.36	58.56 ± 4.38	48.02 ± 0.33	2.37 ± 1.46	27.08 ± 11.22	1.27 ± 0.81
17	**Simclr**	TO	mr=[0.1,0.2]	1	46.47 ± 1.43	61.42 ± 4.32	48.40 ± 1.07	3.69 ± 1.66	36.76 ± 11.32	1.94 ± 0.89
18	**Simclr**	TO	mr=[0.0,0.5]	2	48.38 ± 1.27	62.85 ± 2.93	49.99 ± 1.13	3.67 ± 0.50	36.36 ± 7.17	1.94 ± 0.31
19	**Simclr**	TO	mr=[0.4,0.5]	3	46.99 ± 1.16	54.30 ± 4.31	48.76 ± 0.73	2.51 ± 1.50	11.42 ± 12.78	1.83 ± 1.56
20	**Byol**	BlW	C=0.1	1	42.36 ± 1.07	48.78 ± 0.31	44.26 ± 1.68	1.74 ± 1.04	2.05 ± 0.44	2.39 ± 2.04
21	**Byol**	BlW	C=0.7	2	42.38 ± 1.06	48.78 ± 0.31	44.27 ± 1.67	1.74 ± 1.05	2.02 ± 0.41	2.39 ± 2.08
22	**Byol**	BlW	C=1.0	3	42.36 ± 1.11	48.79 ± 0.31	44.26 ± 1.71	1.75 ± 1.05	2.05 ± 0.44	2.41 ± 2.08
23	**Byol**	ChR	cr=[0.33,1.0]	1	44.90 ± 0.85	50.58 ± 0.65	46.44 ± 0.53	2.93 ± 0.74	2.51 ± 0.26	4.04 ± 2.12
24	**Byol**	ChR	cr=[0.5,2.0]	2	45.02 ± 1.18	51.32 ± 1.22	46.91 ± 1.08	2.97 ± 1.17	4.58 ± 2.98	4.00 ± 3.85
25	**Byol**	ChR	cr=[0.33,3.0]	3	46.88 ± 1.18	51.65 ± 0.32	48.50 ± 0.93	3.79 ± 0.49	3.93 ± 0.42	3.95 ± 1.58
26	**Byol**	DTW	w=1,r=10	1	42.38 ± 1.15	48.71 ± 0.37	44.09 ± 1.76	2.10 ± 1.02	2.31 ± 0.72	3.24 ± 2.72
27	**Byol**	DTW	w=3,r=5	2	42.21 ± 0.89	48.64 ± 0.18	44.09 ± 1.51	1.77 ± 0.98	1.87 ± 0.17	2.41 ± 1.89
28	**Byol**	DTW	w=3,r=10	3	42.31 ± 0.98	48.63 ± 0.27	44.12 ± 1.60	1.98 ± 1.22	2.03 ± 0.28	3.04 ± 2.72
29	**Byol**	EM	var=0.01	1	42.38 ± 1.04	48.83 ± 0.38	44.27 ± 1.65	1.77 ± 0.98	2.20 ± 0.72	2.39 ± 2.05
30	**Byol**	EM	var=0.5	2	42.36 ± 1.07	48.76 ± 0.32	44.26 ± 1.67	1.72 ± 1.04	2.03 ± 0.46	2.37 ± 2.08
31	**Byol**	EM	var=1.0	3	42.34 ± 1.06	48.77 ± 0.32	44.24 ± 1.67	1.76 ± 1.05	2.08 ± 0.47	2.41 ± 2.08
32	**Byol**	GN	var=0.01	1	42.34 ± 1.03	48.73 ± 0.22	44.26 ± 1.65	1.72 ± 1.13	1.91 ± 0.16	2.39 ± 2.11
33	**Byol**	GN	var=0.1	2	42.35 ± 1.05	48.73 ± 0.24	44.26 ± 1.66	1.70 ± 1.11	1.86 ± 0.14	2.37 ± 2.11
34	**Byol**	GN	var=1.0	3	42.35 ± 1.11	48.75 ± 0.32	44.26 ± 1.69	1.73 ± 1.04	2.05 ± 0.45	2.34 ± 1.98
35	**Byol**	Neg	True	2	42.28 ± 0.43	48.73 ± 0.32	44.11 ± 1.12	2.22 ± 1.06	2.26 ± 0.44	3.33 ± 2.92
36	**Byol**	TO	mr=[0.1,0.2]	1	44.39 ± 1.54	49.79 ± 0.38	45.85 ± 1.30	3.14 ± 0.89	2.95 ± 0.38	4.49 ± 3.41
37	**Byol**	TO	mr=[0.0,0.5]	2	46.03 ± 1.01	52.37 ± 3.25	47.65 ± 0.92	4.09 ± 1.46	8.46 ± 8.42	3.53 ± 1.85
38	**Byol**	TO	mr=[0.4,0.5]	3	46.17 ± 1.08	51.99 ± 2.36	47.96 ± 0.64	2.36 ± 1.67	6.39 ± 6.55	1.56 ± 0.99
39	**Swav**	BlW	C=0.1	1	47.18 ± 1.77	64.12 ± 4.17	48.48 ± 1.57	5.94 ± 2.20	47.27 ± 10.08	3.19 ± 1.24
40	**Swav**	BlW	C=0.7	2	45.73 ± 1.45	59.89 ± 9.20	47.55 ± 1.36	4.05 ± 2.95	34.34 ± 28.18	2.17 ± 1.55
41	**Swav**	BlW	C=1.0	3	46.95 ± 0.67	63.80 ± 3.85	48.28 ± 0.73	6.04 ± 2.92	46.73 ± 10.59	3.28 ± 1.65
42	**Swav**	ChR	cr=[0.33,1.0]	1	46.26 ± 3.67	58.79 ± 6.98	48.11 ± 2.78	3.51 ± 2.75	29.68 ± 16.18	1.88 ± 1.48
43	**Swav**	ChR	cr=[0.5,2.0]	2	45.99 ± 1.56	63.51 ± 9.01	47.80 ± 1.24	4.78 ± 1.64	43.27 ± 25.22	2.57 ± 0.94
44	**Swav**	ChR	cr=[0.33,3.0]	3	48.17 ± 1.07	66.71 ± 4.48	49.42 ± 1.01	6.79 ± 1.21	50.81 ± 13.78	3.68 ± 0.81
45	**Swav**	DTW	w=1,r=10	1	44.47 ± 1.01	58.84 ± 1.43	46.36 ± 0.77	4.62 ± 0.53	35.19 ± 4.63	2.48 ± 0.31
46	**Swav**	DTW	w=3,r=5	2	45.67 ± 1.46	53.95 ± 3.15	47.40 ± 0.96	3.32 ± 2.00	17.64 ± 9.70	1.99 ± 1.26
47	**Swav**	DTW	w=3,r=10	3	45.95 ± 1.32	62.27 ± 6.75	47.58 ± 1.14	4.09 ± 1.04	42.81 ± 20.72	2.19 ± 0.59
48	**Swav**	EM	var=0.01	1	46.26 ± 0.98	66.84 ± 12.17	47.93 ± 0.87	4.25 ± 1.54	54.99 ± 37.75	2.23 ± 0.73
49	**Swav**	EM	var=0.5	2	45.99 ± 1.03	70.61 ± 3.87	47.65 ± 0.97	5.48 ± 3.53	66.92 ± 8.45	2.90 ± 1.97
50	**Swav**	EM	var=1.0	3	45.85 ± 1.17	60.42 ± 8.42	47.41 ± 0.92	5.36 ± 2.82	38.18 ± 23.75	2.92 ± 1.60
51	**Swav**	GN	var=0.01	1	46.18 ± 1.16	68.59 ± 3.61	47.77 ± 1.25	5.11 ± 2.84	61.76 ± 11.07	2.68 ± 1.52
52	**Swav**	GN	var=0.1	2	46.64 ± 0.61	65.92 ± 0.44	48.09 ± 0.73	5.87 ± 3.51	52.86 ± 2.49	3.15 ± 1.93
53	**Swav**	GN	var=1.0	3	45.16 ± 0.40	60.75 ± 0.59	46.97 ± 0.34	4.71 ± 2.61	39.41 ± 2.43	2.55 ± 1.51
54	**Swav**	Neg	True	2	46.89 ± 1.77	76.12 ± 2.18	48.43 ± 1.23	5.46 ± 1.83	80.84 ± 5.44	2.84 ± 0.98
55	**Swav**	TO	mr=[0.1,0.2]	1	51.26 ± 3.08	54.77 ± 2.11	52.71 ± 3.20	9.47 ± 1.34	8.50 ± 0.81	11.43 ± 4.50
56	**Swav**	TO	mr=[0.0,0.5]	2	50.18 ± 1.32	57.10 ± 2.95	51.06 ± 1.17	8.84 ± 4.30	17.95 ± 11.44	5.96 ± 2.69
57	**Swav**	TO	mr=[0.4,0.5]	3	51.71 ± 1.65	59.11 ± 4.20	52.58 ± 1.38	7.52 ± 2.27	20.73 ± 12.33	5.78 ± 3.82

Table 7: Validation Class 1 and Class 2 performance under linear evaluation. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Class 1	Class 2
1	**Simclr**	BlW	C=0.1	1	74.66 ± 1.74 / 61.60 ± 1.19 / 94.77 ± 2.80	59.60 ± 2.35 / 87.46 ± 0.84 / 45.24 ± 2.60
2	**Simclr**	BlW	C=0.7	2	74.94 ± 1.27 / 61.58 ± 0.94 / 95.74 ± 2.14	59.39 ± 2.45 / 87.72 ± 1.40 / 44.93 ± 2.63
3	**Simclr**	BlW	C=1.0	3	74.96 ± 1.17 / 61.86 ± 0.49 / 95.13 ± 2.92	60.14 ± 1.03 / 87.04 ± 1.90 / 45.95 ± 0.84
4	**Simclr**	ChR	cr=[0.33,1.0]	1	75.61 ± 0.30 / 66.53 ± 3.98 / 88.31 ± 8.01	54.41 ± 1.71 / 87.55 ± 0.72 / 39.49 ± 1.87
5	**Simclr**	ChR	cr=[0.5,2.0]	2	74.56 ± 0.79 / 62.73 ± 4.27 / 92.78 ± 8.20	54.47 ± 0.96 / 87.96 ± 1.14 / 39.46 ± 0.99
6	**Simclr**	ChR	cr=[0.33,3.0]	3	74.65 ± 1.41 / 62.81 ± 5.47 / 93.00 ± 7.32	55.04 ± 1.57 / 87.69 ± 0.70 / 40.12 ± 1.73
7	**Simclr**	DTW	w=1,r=10	1	74.47 ± 0.72 / 60.70 ± 1.19 / 96.37 ± 1.45	57.85 ± 2.91 / 86.67 ± 1.74 / 43.49 ± 3.37
8	**Simclr**	DTW	w=3,r=5	2	75.39 ± 0.35 / 62.25 ± 0.29 / 95.57 ± 1.48	60.09 ± 0.65 / 86.56 ± 0.36 / 46.02 ± 0.84
9	**Simclr**	DTW	w=3,r=10	3	75.10 ± 0.37 / 61.35 ± 0.80 / 96.81 ± 1.23	58.72 ± 1.21 / 87.88 ± 0.66 / 44.10 ± 1.29
10	**Simclr**	EM	var=0.01	1	74.95 ± 1.39 / 61.76 ± 1.07 / 95.29 ± 2.18	59.89 ± 2.01 / 86.60 ± 1.67 / 45.78 ± 1.92
11	**Simclr**	EM	var=0.5	2	74.43 ± 1.60 / 61.40 ± 1.03 / 94.49 ± 2.73	59.31 ± 2.14 / 87.10 ± 1.30 / 44.99 ± 2.33
12	**Simclr**	EM	var=1.0	3	74.70 ± 1.73 / 61.49 ± 1.59 / 95.14 ± 1.90	59.02 ± 3.26 / 87.03 ± 0.83 / 44.72 ± 3.67
13	**Simclr**	GN	var=0.01	1	74.56 ± 1.73 / 61.34 ± 1.31 / 95.04 ± 2.54	59.33 ± 2.40 / 87.50 ± 1.37 / 44.91 ± 2.61
14	**Simclr**	GN	var=0.1	2	74.67 ± 1.74 / 61.83 ± 1.56 / 94.23 ± 1.95	59.54 ± 2.50 / 86.60 ± 0.69 / 45.42 ± 3.05
15	**Simclr**	GN	var=1.0	3	74.51 ± 1.28 / 61.12 ± 1.03 / 95.42 ± 1.76	58.38 ± 2.10 / 86.85 ± 1.21 / 43.99 ± 2.33
16	**Simclr**	Neg	True	2	75.26 ± 0.58 / 61.62 ± 0.29 / 96.66 ± 1.30	60.28 ± 0.56 / 86.98 ± 2.26 / 46.13 ± 0.18
17	**Simclr**	TO	mr=[0.1,0.2]	1	76.01 ± 0.84 / 62.39 ± 0.98 / 97.24 ± 0.51	59.71 ± 2.08 / 85.10 ± 1.22 / 46.01 ± 2.15
18	**Simclr**	TO	mr=[0.0,0.5]	2	77.27 ± 1.28 / 64.01 ± 1.72 / 97.50 ± 0.38	64.19 ± 2.97 / 88.18 ± 0.91 / 50.53 ± 3.61
19	**Simclr**	TO	mr=[0.4,0.5]	3	76.02 ± 0.48 / 62.85 ± 0.97 / 96.24 ± 2.12	62.43 ± 1.80 / 88.63 ± 0.22 / 48.22 ± 2.21
20	**Byol**	BlW	C=0.1	1	71.67 ± 2.19 / 59.07 ± 1.15 / 91.18 ± 5.08	53.66 ± 2.11 / 85.23 ± 0.64 / 39.20 ± 2.40
21	**Byol**	BlW	C=0.7	2	71.68 ± 2.19 / 59.09 ± 1.13 / 91.16 ± 5.10	53.73 ± 2.05 / 85.22 ± 0.61 / 39.27 ± 2.33
22	**Byol**	BlW	C=1.0	3	71.66 ± 2.22 / 59.07 ± 1.18 / 91.14 ± 5.09	53.68 ± 2.19 / 85.26 ± 0.68 / 39.21 ± 2.49
23	**Byol**	ChR	cr=[0.33,1.0]	1	74.65 ± 0.85 / 62.55 ± 1.97 / 92.69 ± 2.75	57.11 ± 1.07 / 86.68 ± 1.12 / 42.59 ± 1.26
24	**Byol**	ChR	cr=[0.5,2.0]	2	75.02 ± 1.03 / 62.41 ± 2.29 / 94.24 ± 4.07	57.05 ± 2.50 / 86.98 ± 0.29 / 42.49 ± 2.73
25	**Byol**	ChR	cr=[0.33,3.0]	3	76.70 ± 0.82 / 64.11 ± 1.47 / 95.48 ± 1.14	60.16 ± 2.76 / 86.90 ± 0.57 / 46.07 ± 3.44
26	**Byol**	DTW	w=1,r=10	1	71.44 ± 2.20 / 59.42 ± 1.11 / 89.71 ± 5.92	53.62 ± 2.23 / 84.41 ± 1.14 / 39.33 ± 2.54
27	**Byol**	DTW	w=3,r=5	2	71.51 ± 2.01 / 58.92 ± 0.90 / 91.00 ± 4.87	53.35 ± 1.71 / 85.13 ± 0.56 / 38.87 ± 1.93
28	**Byol**	DTW	w=3,r=10	3	71.50 ± 2.12 / 59.28 ± 1.06 / 90.21 ± 5.78	53.47 ± 1.91 / 84.59 ± 1.11 / 39.12 ± 2.14
29	**Byol**	EM	var=0.01	1	71.69 ± 2.15 / 59.08 ± 1.10 / 91.22 ± 5.11	53.68 ± 1.96 / 85.22 ± 0.66 / 39.21 ± 2.24
30	**Byol**	EM	var=0.5	2	71.68 ± 2.18 / 59.08 ± 1.15 / 91.19 ± 5.11	53.67 ± 2.07 / 85.18 ± 0.64 / 39.21 ± 2.36
31	**Byol**	EM	var=1.0	3	71.65 ± 2.18 / 59.05 ± 1.15 / 91.16 ± 5.10	53.61 ± 2.08 / 85.17 ± 0.65 / 39.15 ± 2.37
32	**Byol**	GN	var=0.01	1	71.67 ± 2.18 / 59.05 ± 1.14 / 91.22 ± 5.11	53.63 ± 2.07 / 85.22 ± 0.62 / 39.16 ± 2.34
33	**Byol**	GN	var=0.1	2	71.67 ± 2.18 / 59.08 ± 1.15 / 91.18 ± 5.11	53.69 ± 2.09 / 85.25 ± 0.57 / 39.22 ± 2.36
34	**Byol**	GN	var=1.0	3	71.68 ± 2.18 / 59.05 ± 1.18 / 91.26 ± 4.87	53.62 ± 2.19 / 85.16 ± 0.66 / 39.17 ± 2.49
35	**Byol**	Neg	True	2	71.51 ± 1.60 / 59.29 ± 0.94 / 90.30 ± 6.08	53.12 ± 0.69 / 84.63 ± 1.41 / 38.72 ± 0.86
36	**Byol**	TO	mr=[0.1,0.2]	1	72.83 ± 1.73 / 61.40 ± 2.60 / 89.84 ± 6.07	57.20 ± 3.45 / 85.02 ± 1.57 / 43.23 ± 4.21
37	**Byol**	TO	mr=[0.0,0.5]	2	74.63 ± 1.15 / 61.80 ± 0.74 / 94.28 ± 4.25	59.37 ± 1.36 / 86.84 ± 1.52 / 45.13 ± 1.82
38	**Byol**	TO	mr=[0.4,0.5]	3	75.14 ± 0.31 / 61.97 ± 1.05 / 95.48 ± 2.34	61.00 ± 2.32 / 87.61 ± 0.86 / 46.84 ± 2.96
39	**Swav**	BlW	C=0.1	1	74.86 ± 1.67 / 62.13 ± 2.28 / 94.21 ± 0.68	60.74 ± 4.96 / 82.96 ± 1.53 / 48.02 ± 5.60
40	**Swav**	BlW	C=0.7	2	74.47 ± 1.59 / 61.06 ± 1.88 / 95.44 ± 0.75	58.68 ± 4.07 / 84.28 ± 2.10 / 45.06 ± 4.27
41	**Swav**	BlW	C=1.0	3	74.72 ± 1.10 / 61.93 ± 1.94 / 94.26 ± 1.15	60.08 ± 3.68 / 82.75 ± 1.21 / 47.31 ± 4.81
42	**Swav**	ChR	cr=[0.33,1.0]	1	75.21 ± 2.40 / 61.95 ± 3.02 / 95.78 ± 0.76	60.07 ± 5.94 / 84.74 ± 2.10 / 46.68 ± 6.43
43	**Swav**	ChR	cr=[0.5,2.0]	2	74.64 ± 1.17 / 60.87 ± 1.61 / 96.50 ± 0.14	58.54 ± 3.57 / 86.39 ± 0.42 / 44.34 ± 4.02
44	**Swav**	ChR	cr=[0.33,3.0]	3	75.99 ± 1.28 / 62.63 ± 1.46 / 96.60 ± 0.96	61.74 ± 2.52 / 86.67 ± 1.39 / 47.99 ± 2.86
45	**Swav**	DTW	w=1,r=10	1	73.10 ± 0.69 / 59.68 ± 0.87 / 94.33 ± 1.13	55.68 ± 2.06 / 81.66 ± 1.37 / 42.27 ± 2.36
46	**Swav**	DTW	w=3,r=5	2	74.15 ± 0.60 / 61.36 ± 1.15 / 93.70 ± 0.87	59.53 ± 2.05 / 82.85 ± 1.09 / 46.51 ± 2.82
47	**Swav**	DTW	w=3,r=10	3	74.16 ± 1.29 / 61.30 ± 1.86 / 93.88 ± 0.45	59.59 ± 3.46 / 82.69 ± 0.27 / 46.66 ± 4.10
48	**Swav**	EM	var=0.01	1	74.80 ± 1.02 / 61.69 ± 1.88 / 95.06 ± 1.18	59.73 ± 3.46 / 83.84 ± 1.20 / 46.51 ± 4.46
49	**Swav**	EM	var=0.5	2	74.31 ± 1.30 / 60.82 ± 1.58 / 95.52 ± 0.43	58.18 ± 3.59 / 84.08 ± 1.62 / 44.52 ± 3.77
50	**Swav**	EM	var=1.0	3	73.95 ± 0.90 / 60.89 ± 1.34 / 94.15 ± 0.65	58.24 ± 3.18 / 82.20 ± 1.08 / 45.17 ± 3.64
51	**Swav**	GN	var=0.01	1	74.52 ± 1.70 / 61.43 ± 2.31 / 94.74 ± 0.62	58.91 ± 4.48 / 82.58 ± 1.43 / 45.90 ± 5.20
52	**Swav**	GN	var=0.1	2	74.63 ± 1.14 / 61.53 ± 1.82 / 94.86 ± 0.81	59.42 ± 3.65 / 83.37 ± 1.00 / 46.26 ± 4.53
53	**Swav**	GN	var=1.0	3	73.77 ± 0.60 / 60.39 ± 0.97 / 94.80 ± 0.42	56.99 ± 2.04 / 82.47 ± 0.85 / 43.58 ± 2.50
54	**Swav**	Neg	True	2	75.07 ± 0.80 / 61.59 ± 1.42 / 96.13 ± 0.79	60.13 ± 2.99 / 85.93 ± 0.78 / 46.31 ± 3.68
55	**Swav**	TO	mr=[0.1,0.2]	1	78.83 ± 3.04 / 67.72 ± 4.65 / 94.49 ± 0.28	65.47 ± 4.94 / 88.11 ± 1.74 / 52.19 ± 5.63
56	**Swav**	TO	mr=[0.0,0.5]	2	76.99 ± 1.51 / 64.19 ± 2.55 / 96.26 ± 1.20	64.70 ± 4.85 / 89.15 ± 0.72 / 50.95 ± 5.89
57	**Swav**	TO	mr=[0.4,0.5]	3	78.63 ± 1.21 / 66.64 ± 1.89 / 95.95 ± 2.27	68.99 ± 2.60 / 89.97 ± 0.77 / 56.02 ± 3.69

Table 8: Validation Macro and Class 3 performance of SSL models after fine-tuning. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Macro	Class 3
1	**Simclr**	BlW	C=0.1	1	52.97 ± 0.55 / 70.69 ± 4.22 / 53.06 ± 0.57	13.48 ± 4.61 / 55.95 ± 12.06 / 7.81 ± 3.15
2	**Simclr**	BlW	C=0.7	2	54.07 ± 0.61 / 71.82 ± 2.32 / 53.86 ± 0.44	13.76 ± 0.90 / 58.91 ± 7.29 / 7.81 ± 0.65
3	**Simclr**	BlW	C=1.0	3	53.47 ± 1.44 / 76.20 ± 1.68 / 53.53 ± 0.97	11.55 ± 4.07 / 72.01 ± 5.83 / 6.34 ± 2.39
4	**Simclr**	ChR	cr=[0.33,1.0]	1	53.89 ± 0.75 / 73.98 ± 5.94 / 54.29 ± 0.59	8.06 ± 2.01 / 63.42 ± 16.96 / 4.33 ± 1.15
5	**Simclr**	ChR	cr=[0.5,2.0]	2	54.18 ± 1.50 / 74.05 ± 4.17 / 54.28 ± 1.16	10.19 ± 1.96 / 64.67 ± 10.97 / 5.54 ± 1.09
6	**Simclr**	ChR	cr=[0.33,3.0]	3	54.75 ± 1.11 / 76.12 ± 3.92 / 54.39 ± 0.63	14.59 ± 4.29 / 71.08 ± 11.70 / 8.22 ± 2.62
7	**Simclr**	DTW	w=1,r=10	1	52.97 ± 1.37 / 70.78 ± 4.23 / 52.98 ± 0.87	14.58 ± 4.12 / 55.90 ± 11.89 / 8.57 ± 2.95
8	**Simclr**	DTW	w=3,r=5	2	52.46 ± 1.53 / 76.74 ± 5.32 / 53.17 ± 0.72	6.93 ± 5.51 / 73.42 ± 14.44 / 3.75 ± 3.17
9	**Simclr**	DTW	w=3,r=10	3	52.55 ± 1.68 / 74.42 ± 5.02 / 53.13 ± 0.73	8.39 ± 6.57 / 66.05 ± 15.48 / 4.55 ± 3.69
10	**Simclr**	EM	var=0.01	1	52.67 ± 0.72 / 73.90 ± 2.92 / 53.10 ± 0.52	8.82 ± 1.11 / 65.10 ± 8.12 / 4.73 ± 0.61
11	**Simclr**	EM	var=0.5	2	53.36 ± 0.76 / 71.55 ± 1.28 / 53.41 ± 0.37	12.48 ± 4.49 / 58.07 ± 3.91 / 7.08 ± 2.90
12	**Simclr**	EM	var=1.0	3	53.65 ± 0.71 / 72.51 ± 0.40 / 53.48 ± 0.68	14.36 ± 3.72 / 61.26 ± 0.45 / 8.19 ± 2.44
13	**Simclr**	GN	var=0.01	1	53.79 ± 0.30 / 73.97 ± 4.23 / 53.67 ± 0.10	13.78 ± 5.11 / 65.53 ± 11.83 / 7.95 ± 3.62
14	**Simclr**	GN	var=0.1	2	52.78 ± 0.21 / 70.23 ± 4.52 / 53.27 ± 0.55	9.55 ± 7.81 / 54.07 ± 11.12 / 5.67 ± 5.06
15	**Simclr**	GN	var=1.0	3	53.97 ± 1.03 / 72.96 ± 4.70 / 53.88 ± 0.56	13.49 ± 6.15 / 62.36 ± 13.31 / 7.88 ± 4.33
16	**Simclr**	Neg	True	2	53.12 ± 2.19 / 77.08 ± 9.09 / 53.88 ± 1.07	9.60 ± 13.98 / 73.66 ± 26.17 / 6.61 ± 10.11
17	**Simclr**	TO	mr=[0.1,0.2]	1	54.88 ± 0.45 / 73.43 ± 6.27 / 55.12 ± 0.49	8.31 ± 1.62 / 60.87 ± 18.49 / 4.51 ± 1.01
18	**Simclr**	TO	mr=[0.0,0.5]	2	56.35 ± 1.68 / 73.74 ± 3.62 / 56.18 ± 0.89	11.30 ± 7.25 / 60.45 ± 11.47 / 6.39 ± 4.34
19	**Simclr**	TO	mr=[0.4,0.5]	3	54.64 ± 1.18 / 72.14 ± 5.94 / 54.71 ± 0.66	10.35 ± 5.32 / 58.11 ± 16.98 / 5.89 ± 3.20
20	**Byol**	BlW	C=0.1	1	51.07 ± 1.66 / 65.56 ± 10.69 / 51.71 ± 0.76	7.60 ± 7.41 / 44.07 ± 29.92 / 4.31 ± 4.34
21	**Byol**	BlW	C=0.7	2	50.48 ± 1.69 / 70.01 ± 5.94 / 51.55 ± 0.85	6.99 ± 7.83 / 54.89 ± 18.18 / 4.13 ± 4.97
22	**Byol**	BlW	C=1.0	3	51.78 ± 2.09 / 65.25 ± 5.43 / 52.30 ± 1.01	8.93 ± 7.82 / 42.01 ± 15.18 / 5.18 ± 4.60
23	**Byol**	ChR	cr=[0.33,1.0]	1	54.77 ± 2.87 / 78.67 ± 2.69 / 55.11 ± 2.33	8.35 ± 4.62 / 75.40 ± 4.97 / 4.49 ± 2.64
24	**Byol**	ChR	cr=[0.5,2.0]	2	53.89 ± 1.06 / 69.75 ± 8.10 / 53.88 ± 0.76	15.74 ± 5.63 / 51.97 ± 22.56 / 10.09 ± 4.40
25	**Byol**	ChR	cr=[0.33,3.0]	3	55.90 ± 0.84 / 76.11 ± 3.10 / 56.34 ± 0.66	6.46 ± 5.93 / 65.85 ± 10.56 / 3.48 ± 3.22
26	**Byol**	DTW	w=1,r=10	1	51.50 ± 2.72 / 68.03 ± 7.73 / 52.54 ± 1.95	8.78 ± 9.14 / 48.82 ± 23.47 / 6.65 ± 8.11
27	**Byol**	DTW	w=3,r=5	2	51.59 ± 0.85 / 65.63 ± 3.25 / 52.10 ± 0.96	8.31 ± 1.55 / 43.93 ± 7.56 / 4.64 ± 1.08
28	**Byol**	DTW	w=3,r=10	3	51.09 ± 1.89 / 73.49 ± 5.88 / 52.15 ± 1.34	5.25 ± 2.71 / 64.47 ± 16.88 / 2.77 ± 1.51
29	**Byol**	EM	var=0.01	1	51.39 ± 0.45 / 69.79 ± 9.31 / 52.06 ± 0.45	10.55 ± 3.38 / 53.41 ± 26.85 / 6.43 ± 3.07
30	**Byol**	EM	var=0.5	2	50.97 ± 0.91 / 68.33 ± 3.61 / 51.72 ± 0.22	10.25 ± 8.58 / 50.57 ± 10.43 / 6.23 ± 5.86
31	**Byol**	EM	var=1.0	3	51.50 ± 1.03 / 71.74 ± 7.27 / 52.07 ± 0.75	8.91 ± 2.79 / 60.96 ± 19.80 / 4.93 ± 1.81
32	**Byol**	GN	var=0.01	1	51.98 ± 1.17 / 67.73 ± 3.93 / 52.18 ± 1.09	14.04 ± 4.16 / 49.40 ± 9.51 / 8.42 ± 2.98
33	**Byol**	GN	var=0.1	2	51.46 ± 1.79 / 67.36 ± 1.38 / 52.17 ± 1.05	8.25 ± 5.76 / 46.65 ± 4.37 / 4.64 ± 3.34
34	**Byol**	GN	var=1.0	3	50.34 ± 2.11 / 75.54 ± 4.92 / 51.45 ± 1.09	5.49 ± 5.90 / 72.09 ± 15.49 / 2.95 ± 3.25
35	**Byol**	Neg	True	2	51.43 ± 1.66 / 71.11 ± 7.82 / 52.14 ± 0.80	8.65 ± 7.35 / 58.69 ± 22.76 / 5.14 ± 4.81
36	**Byol**	TO	mr=[0.1,0.2]	1	52.48 ± 3.91 / 70.53 ± 3.51 / 53.17 ± 2.49	9.53 ± 12.23 / 55.00 ± 10.76 / 6.07 ± 8.31
37	**Byol**	TO	mr=[0.0,0.5]	2	52.27 ± 0.94 / 70.04 ± 0.69 / 52.63 ± 0.84	11.02 ± 1.24 / 54.20 ± 3.23 / 6.14 ± 0.74
38	**Byol**	TO	mr=[0.4,0.5]	3	53.80 ± 1.65 / 70.09 ± 5.47 / 54.08 ± 1.33	16.73 ± 12.25 / 53.13 ± 15.53 / 11.65 ± 9.58
39	**Swav**	BlW	C=0.1	1	53.97 ± 0.73 / 78.09 ± 2.02 / 54.20 ± 0.47	8.28 ± 1.96 / 77.50 ± 5.42 / 4.38 ± 1.07
40	**Swav**	BlW	C=0.7	2	53.47 ± 0.53 / 78.54 ± 3.55 / 53.79 ± 0.30	8.52 ± 3.37 / 78.89 ± 10.31 / 4.55 ± 1.95
41	**Swav**	BlW	C=1.0	3	53.09 ± 1.20 / 72.12 ± 7.21 / 53.40 ± 0.92	9.21 ± 1.52 / 60.21 ± 20.78 / 5.02 ± 0.91
42	**Swav**	ChR	cr=[0.33,1.0]	1	55.49 ± 0.20 / 74.01 ± 12.73 / 55.43 ± 0.30	10.54 ± 1.26 / 63.11 ± 38.74 / 6.14 ± 0.94
43	**Swav**	ChR	cr=[0.5,2.0]	2	54.76 ± 1.12 / 75.22 ± 4.79 / 55.03 ± 0.56	7.87 ± 4.08 / 66.86 ± 14.74 / 4.29 ± 2.45
44	**Swav**	ChR	cr=[0.33,3.0]	3	54.39 ± 0.34 / 73.29 ± 8.69 / 54.77 ± 0.36	7.44 ± 2.82 / 60.89 ± 26.05 / 4.06 ± 1.65
45	**Swav**	DTW	w=1,r=10	1	52.10 ± 0.24 / 78.76 ± 1.99 / 52.87 ± 0.09	6.09 ± 0.99 / 80.22 ± 6.04 / 3.17 ± 0.54
46	**Swav**	DTW	w=3,r=5	2	53.35 ± 0.11 / 77.95 ± 4.22 / 53.54 ± 0.18	9.89 ± 1.49 / 77.19 ± 12.14 / 5.29 ± 0.82
47	**Swav**	DTW	w=3,r=10	3	53.75 ± 0.36 / 74.47 ± 2.24 / 53.70 ± 0.33	11.43 ± 1.40 / 67.60 ± 5.76 / 6.25 ± 0.84
48	**Swav**	EM	var=0.01	1	53.21 ± 0.35 / 78.69 ± 2.63 / 53.62 ± 0.23	8.22 ± 4.32 / 79.33 ± 6.59 / 4.40 ± 2.48
49	**Swav**	EM	var=0.5	2	52.82 ± 0.40 / 78.86 ± 4.74 / 53.26 ± 0.10	8.16 ± 3.75 / 80.25 ± 13.84 / 4.35 ± 2.08
50	**Swav**	EM	var=1.0	3	52.67 ± 0.27 / 77.21 ± 2.74 / 53.16 ± 0.38	7.72 ± 1.76 / 75.52 ± 8.05 / 4.09 ± 0.99
51	**Swav**	GN	var=0.01	1	53.37 ± 0.40 / 76.64 ± 1.09 / 53.52 ± 0.41	10.06 ± 0.68 / 73.69 ± 3.53 / 5.40 ± 0.39
52	**Swav**	GN	var=0.1	2	53.35 ± 0.39 / 78.27 ± 1.93 / 53.64 ± 0.28	9.22 ± 2.97 / 77.61 ± 5.06 / 4.93 ± 1.72
53	**Swav**	GN	var=1.0	3	53.74 ± 1.23 / 76.49 ± 3.60 / 53.97 ± 0.84	8.92 ± 2.17 / 72.51 ± 10.27 / 4.76 ± 1.18
54	**Swav**	Neg	True	2	54.57 ± 0.57 / 79.93 ± 2.76 / 54.59 ± 0.51	10.09 ± 1.69 / 81.83 ± 7.35 / 5.38 ± 0.95
55	**Swav**	TO	mr=[0.1,0.2]	1	58.41 ± 1.31 / 69.94 ± 8.39 / 58.54 ± 1.54	15.47 ± 4.92 / 46.45 ± 24.07 / 12.46 ± 9.34
56	**Swav**	TO	mr=[0.0,0.5]	2	56.99 ± 0.79 / 66.01 ± 4.24 / 56.82 ± 1.33	16.93 ± 4.69 / 37.82 ± 12.76 / 12.37 ± 6.20
57	**Swav**	TO	mr=[0.4,0.5]	3	55.75 ± 0.93 / 66.66 ± 1.50 / 55.59 ± 0.78	11.94 ± 1.46 / 40.47 ± 6.25 / 7.06 ± 1.09

Table 9: Validation Class 1 and Class 2 performance of SSL models after fine-tuning. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Class 1	Class 2
1	**Simclr**	BlW	C=0.1	1	78.45 ± 1.53 / 65.40 ± 2.37 / 98.08 ± 0.77	66.98 ± 3.85 / 90.71 ± 1.55 / 53.27 ± 5.26
2	**Simclr**	BlW	C=0.7	2	79.15 ± 0.31 / 66.68 ± 0.54 / 97.37 ± 0.88	69.29 ± 0.81 / 89.85 ± 1.02 / 56.41 ± 1.30
3	**Simclr**	BlW	C=1.0	3	79.28 ± 0.81 / 66.85 ± 1.32 / 97.41 ± 0.45	69.57 ± 1.72 / 89.72 ± 0.86 / 56.85 ± 2.55
4	**Simclr**	ChR	cr=[0.33,1.0]	1	80.90 ± 0.72 / 69.33 ± 1.21 / 97.13 ± 0.31	72.71 ± 1.55 / 89.19 ± 0.62 / 61.41 ± 2.37
5	**Simclr**	ChR	cr=[0.5,2.0]	2	80.44 ± 1.01 / 68.81 ± 1.17 / 96.81 ± 0.62	71.91 ± 1.53 / 88.67 ± 0.83 / 60.49 ± 1.77
6	**Simclr**	ChR	cr=[0.33,3.0]	3	79.66 ± 0.15 / 67.27 ± 0.63 / 97.66 ± 0.94	70.00 ± 0.83 / 90.01 ± 1.39 / 57.30 ± 1.69
7	**Simclr**	DTW	w=1,r=10	1	78.05 ± 0.62 / 64.64 ± 0.84 / 98.48 ± 0.43	66.29 ± 1.63 / 91.82 ± 0.56 / 51.89 ± 2.03
8	**Simclr**	DTW	w=3,r=5	2	79.86 ± 0.83 / 67.73 ± 0.75 / 97.29 ± 1.29	70.60 ± 1.06 / 89.08 ± 1.36 / 58.48 ± 1.20
9	**Simclr**	DTW	w=3,r=10	3	79.54 ± 0.56 / 66.93 ± 0.69 / 98.00 ± 0.26	69.74 ± 1.04 / 90.28 ± 0.34 / 56.82 ± 1.37
10	**Simclr**	EM	var=0.01	1	79.36 ± 0.42 / 66.96 ± 0.60 / 97.38 ± 0.31	69.82 ± 0.67 / 89.63 ± 0.62 / 57.19 ± 0.98
11	**Simclr**	EM	var=0.5	2	78.96 ± 0.80 / 66.25 ± 1.41 / 97.72 ± 0.68	68.66 ± 2.04 / 90.34 ± 1.32 / 55.45 ± 3.11
12	**Simclr**	EM	var=1.0	3	78.65 ± 1.31 / 65.83 ± 2.11 / 97.75 ± 0.69	67.92 ± 3.16 / 90.46 ± 1.47 / 54.51 ± 4.48
13	**Simclr**	GN	var=0.01	1	78.87 ± 1.23 / 66.28 ± 1.93 / 97.41 ± 0.53	68.73 ± 3.08 / 90.12 ± 1.06 / 55.66 ± 4.33
14	**Simclr**	GN	var=0.1	2	79.27 ± 2.33 / 67.03 ± 3.44 / 97.08 ± 0.25	69.54 ± 4.86 / 89.60 ± 1.05 / 57.07 ± 6.88
15	**Simclr**	GN	var=1.0	3	79.15 ± 1.12 / 66.73 ± 1.71 / 97.29 ± 0.25	69.28 ± 2.42 / 89.81 ± 0.96 / 56.47 ± 3.57
16	**Simclr**	Neg	True	2	79.67 ± 2.12 / 67.53 ± 3.39 / 97.27 ± 0.91	70.10 ± 5.29 / 90.06 ± 2.33 / 57.77 ± 7.74
17	**Simclr**	TO	mr=[0.1,0.2]	1	81.72 ± 0.76 / 71.01 ± 1.57 / 96.26 ± 0.87	74.62 ± 1.68 / 88.42 ± 0.97 / 64.61 ± 3.01
18	**Simclr**	TO	mr=[0.0,0.5]	2	82.20 ± 1.26 / 71.89 ± 3.13 / 96.14 ± 2.76	75.56 ± 3.51 / 88.87 ± 2.64 / 66.03 ± 6.53
19	**Simclr**	TO	mr=[0.4,0.5]	3	80.65 ± 1.05 / 69.39 ± 1.97 / 96.33 ± 1.32	72.93 ± 2.21 / 88.92 ± 1.79 / 61.90 ± 3.65
20	**Byol**	BlW	C=0.1	1	77.68 ± 0.66 / 66.01 ± 2.48 / 94.59 ± 3.92	67.93 ± 2.56 / 86.58 ± 4.56 / 56.22 ± 5.58
21	**Byol**	BlW	C=0.7	2	78.00 ± 0.75 / 64.84 ± 1.32 / 97.89 ± 0.64	66.44 ± 2.32 / 90.29 ± 1.49 / 52.64 ± 3.39
22	**Byol**	BlW	C=1.0	3	78.18 ± 0.53 / 66.01 ± 0.83 / 95.87 ± 1.59	68.22 ± 1.11 / 87.74 ± 2.66 / 55.86 ± 2.06
23	**Byol**	ChR	cr=[0.33,1.0]	1	81.78 ± 2.54 / 70.44 ± 4.07 / 97.62 ± 0.48	74.17 ± 4.56 / 90.18 ± 0.46 / 63.22 ± 6.97
24	**Byol**	ChR	cr=[0.5,2.0]	2	78.62 ± 2.00 / 65.64 ± 3.02 / 98.10 ± 0.60	67.31 ± 4.90 / 91.65 ± 1.59 / 53.45 ± 6.77
25	**Byol**	ChR	cr=[0.33,3.0]	3	83.53 ± 1.78 / 73.98 ± 3.43 / 96.05 ± 1.18	77.71 ± 3.08 / 88.50 ± 1.51 / 69.47 ± 5.68
26	**Byol**	DTW	w=1,r=10	1	78.25 ± 1.02 / 65.67 ± 1.22 / 96.80 ± 0.81	67.49 ± 2.04 / 89.61 ± 0.84 / 54.18 ± 2.92
27	**Byol**	DTW	w=3,r=5	2	78.13 ± 1.55 / 66.27 ± 1.77 / 95.20 ± 2.20	68.32 ± 2.58 / 86.69 ± 2.66 / 56.45 ± 3.43
28	**Byol**	DTW	w=3,r=10	3	79.02 ± 0.95 / 66.47 ± 1.67 / 97.45 ± 0.67	69.01 ± 2.45 / 89.52 ± 0.94 / 56.21 ± 3.45
29	**Byol**	EM	var=0.01	1	77.89 ± 1.20 / 64.49 ± 1.69 / 98.34 ± 0.28	65.73 ± 3.52 / 91.47 ± 0.77 / 51.40 ± 4.46
30	**Byol**	EM	var=0.5	2	77.50 ± 1.62 / 64.26 ± 2.41 / 97.69 ± 0.55	65.16 ± 4.53 / 90.15 ± 1.50 / 51.23 ± 5.88
31	**Byol**	EM	var=1.0	3	78.15 ± 0.82 / 65.49 ± 0.44 / 96.88 ± 1.57	67.45 ± 0.91 / 88.78 ± 2.15 / 54.39 ± 0.46
32	**Byol**	GN	var=0.01	1	77.29 ± 1.89 / 64.10 ± 2.72 / 97.42 ± 0.94	64.60 ± 5.21 / 89.69 ± 1.68 / 50.70 ± 6.47
33	**Byol**	GN	var=0.1	2	78.42 ± 0.85 / 65.60 ± 1.38 / 97.49 ± 0.45	67.71 ± 2.19 / 89.83 ± 0.64 / 54.38 ± 3.02
34	**Byol**	GN	var=1.0	3	78.13 ± 0.35 / 65.30 ± 0.18 / 97.24 ± 1.11	67.40 ± 0.68 / 89.22 ± 0.77 / 54.17 ± 0.99
35	**Byol**	Neg	True	2	78.18 ± 0.63 / 65.48 ± 0.93 / 96.98 ± 0.51	67.47 ± 1.95 / 89.15 ± 0.60 / 54.31 ± 2.63
36	**Byol**	TO	mr=[0.1,0.2]	1	79.13 ± 0.12 / 66.42 ± 0.31 / 97.84 ± 0.35	68.77 ± 0.60 / 90.18 ± 1.03 / 55.60 ± 1.18
37	**Byol**	TO	mr=[0.0,0.5]	2	78.59 ± 1.03 / 65.52 ± 1.52 / 98.20 ± 0.26	67.21 ± 2.36 / 90.41 ± 0.48 / 53.54 ± 3.08
38	**Byol**	TO	mr=[0.4,0.5]	3	78.46 ± 2.24 / 65.30 ± 3.52 / 98.43 ± 0.87	66.21 ± 5.82 / 91.85 ± 2.07 / 52.15 ± 8.14
39	**Swav**	BlW	C=0.1	1	80.47 ± 0.35 / 69.72 ± 0.16 / 95.15 ± 0.69	73.16 ± 0.31 / 87.06 ± 0.81 / 63.09 ± 0.14
40	**Swav**	BlW	C=0.7	2	80.03 ± 0.70 / 68.56 ± 1.26 / 96.11 ± 0.70	71.87 ± 1.56 / 88.17 ± 1.16 / 60.71 ± 2.54
41	**Swav**	BlW	C=1.0	3	79.57 ± 0.76 / 67.66 ± 1.16 / 96.59 ± 0.26	70.48 ± 1.65 / 88.50 ± 0.24 / 58.59 ± 2.30
42	**Swav**	ChR	cr=[0.33,1.0]	1	81.58 ± 0.68 / 71.12 ± 1.13 / 95.65 ± 0.96	74.37 ± 0.41 / 87.80 ± 1.10 / 64.51 ± 0.71
43	**Swav**	ChR	cr=[0.5,2.0]	2	81.51 ± 0.32 / 71.17 ± 0.52 / 95.39 ± 1.24	74.89 ± 0.67 / 87.63 ± 1.04 / 65.42 ± 1.56
44	**Swav**	ChR	cr=[0.33,3.0]	3	81.33 ± 0.92 / 70.58 ± 1.51 / 95.96 ± 0.83	74.41 ± 1.59 / 88.41 ± 0.98 / 64.28 ± 2.72
45	**Swav**	DTW	w=1,r=10	1	79.58 ± 0.13 / 67.63 ± 0.18 / 96.67 ± 0.25	70.61 ± 0.17 / 88.45 ± 0.34 / 58.76 ± 0.35
46	**Swav**	DTW	w=3,r=5	2	79.49 ± 0.48 / 67.43 ± 0.85 / 96.81 ± 0.74	70.66 ± 1.00 / 89.22 ± 0.97 / 58.51 ± 1.71
47	**Swav**	DTW	w=3,r=10	3	79.31 ± 0.61 / 67.47 ± 0.60 / 96.19 ± 0.59	70.50 ± 0.83 / 88.34 ± 0.76 / 58.66 ± 0.84
48	**Swav**	EM	var=0.01	1	79.96 ± 1.19 / 68.28 ± 1.68 / 96.48 ± 0.08	71.46 ± 2.14 / 88.45 ± 0.32 / 59.99 ± 3.03
49	**Swav**	EM	var=0.5	2	79.50 ± 0.77 / 67.64 ± 1.61 / 96.46 ± 1.09	70.78 ± 1.81 / 88.69 ± 1.61 / 58.98 ± 3.25
50	**Swav**	EM	var=1.0	3	79.46 ± 0.73 / 67.63 ± 1.15 / 96.32 ± 1.07	70.82 ± 1.37 / 88.48 ± 1.47 / 59.07 ± 2.23
51	**Swav**	GN	var=0.01	1	79.40 ± 0.52 / 67.48 ± 0.88 / 96.46 ± 0.62	70.65 ± 1.36 / 88.74 ± 0.70 / 58.70 ± 1.97
52	**Swav**	GN	var=0.1	2	79.72 ± 0.76 / 67.70 ± 1.18 / 96.96 ± 0.39	71.10 ± 1.60 / 89.50 ± 0.61 / 59.01 ± 2.35
53	**Swav**	GN	var=1.0	3	80.19 ± 0.54 / 68.77 ± 0.88 / 96.16 ± 0.66	72.10 ± 1.03 / 88.17 ± 0.51 / 61.01 ± 1.67
54	**Swav**	Neg	True	2	80.61 ± 0.78 / 69.31 ± 0.92 / 96.33 ± 0.46	73.01 ± 1.17 / 88.65 ± 0.50 / 62.06 ± 1.45
55	**Swav**	TO	mr=[0.1,0.2]	1	83.68 ± 1.61 / 73.80 ± 3.00 / 96.69 ± 0.80	76.10 ± 4.74 / 89.59 ± 1.27 / 66.46 ± 7.84
56	**Swav**	TO	mr=[0.0,0.5]	2	81.40 ± 1.54 / 70.07 ± 2.32 / 97.16 ± 0.35	72.63 ± 2.88 / 90.13 ± 0.57 / 60.92 ± 4.33
57	**Swav**	TO	mr=[0.4,0.5]	3	81.51 ± 0.67 / 70.43 ± 0.56 / 96.72 ± 1.02	73.80 ± 0.72 / 89.07 ± 1.39 / 63.00 ± 0.69

A.3 Performance of SSL Models on Test Set

Table 10: Test Macro and Class 3 performance under linear evaluation. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Macro			Class 3		
1	**Simclr**	BlW	C=0.1	1	73.87 ± 1.39	81.04 ± 1.20	69.46 ± 1.45	68.69 ± 1.64	84.97 ± 2.47	57.64 ± 1.18
2	**Simclr**	BlW	C=0.7	2	74.24 ± 1.28	82.00 ± 1.85	69.62 ± 0.99	70.24 ± 2.20	87.02 ± 4.11	58.94 ± 2.30
3	**Simclr**	BlW	C=1.0	3	73.96 ± 1.09	81.66 ± 1.36	69.31 ± 1.04	68.53 ± 0.72	87.00 ± 2.49	56.55 ± 0.71
4	**Simclr**	ChR	cr=[0.33,1.0]	1	76.87 ± 1.44	84.60 ± 2.25	72.39 ± 0.93	81.24 ± 1.69	95.27 ± 4.53	70.86 ± 0.18
5	**Simclr**	ChR	cr=[0.5,2.0]	2	76.26 ± 1.08	85.68 ± 0.40	71.35 ± 1.26	82.35 ± 0.38	98.69 ± 1.19	70.65 ± 0.37
6	**Simclr**	ChR	cr=[0.33,3.0]	3	76.25 ± 0.42	84.87 ± 1.19	71.59 ± 0.80	81.53 ± 1.04	96.52 ± 3.12	70.59 ± 0.34
7	**Simclr**	DTW	w=1,r=10	1	72.81 ± 1.87	82.54 ± 2.39	67.74 ± 2.58	71.06 ± 8.12	92.38 ± 5.90	58.26 ± 10.72
8	**Simclr**	DTW	w=3,r=5	2	72.12 ± 5.70	78.54 ± 2.70	68.40 ± 6.45	62.26 ± 16.70	77.80 ± 8.86	52.86 ± 18.68
9	**Simclr**	DTW	w=3,r=10	3	74.28 ± 1.48	83.18 ± 3.49	69.28 ± 0.91	72.22 ± 6.88	91.41 ± 9.89	59.79 ± 6.01
10	**Simclr**	EM	var=0.01	1	73.53 ± 1.04	81.56 ± 0.40	68.76 ± 1.23	66.57 ± 1.66	86.47 ± 0.17	54.13 ± 2.13
11	**Simclr**	EM	var=0.5	2	73.35 ± 1.24	80.93 ± 0.78	68.83 ± 1.48	66.94 ± 1.84	84.39 ± 2.61	55.58 ± 3.17
12	**Simclr**	EM	var=1.0	3	72.98 ± 1.23	79.40 ± 0.39	69.02 ± 1.60	66.98 ± 0.51	80.48 ± 1.98	57.37 ± 0.26
13	**Simclr**	GN	var=0.01	1	73.23 ± 1.07	81.42 ± 0.45	68.48 ± 1.31	66.83 ± 1.01	85.80 ± 1.72	54.78 ± 1.98
14	**Simclr**	GN	var=0.1	2	73.82 ± 0.72	79.46 ± 2.91	70.19 ± 0.82	68.60 ± 4.32	81.28 ± 8.53	59.44 ± 1.96
15	**Simclr**	GN	var=1.0	3	72.65 ± 0.18	80.20 ± 0.58	68.16 ± 0.48	65.78 ± 1.77	82.84 ± 2.27	54.60 ± 2.54
16	**Simclr**	Neg	True	2	73.09 ± 0.99	81.86 ± 4.12	68.21 ± 0.28	64.63 ± 3.46	85.69 ± 12.15	52.18 ± 1.14
17	**Simclr**	TO	mr=[0.1,0.2]	1	70.24 ± 1.39	82.98 ± 0.58	64.48 ± 1.84	71.54 ± 5.65	98.94 ± 0.47	56.25 ± 7.11
18	**Simclr**	TO	mr=[0.0,0.5]	2	68.65 ± 3.48	83.27 ± 1.28	63.00 ± 3.35	68.02 ± 0.87	96.91 ± 2.94	52.45 ± 1.77
19	**Simclr**	TO	mr=[0.4,0.5]	3	72.80 ± 2.76	85.50 ± 0.12	67.26 ± 3.22	77.06 ± 5.79	99.29 ± 0.50	63.19 ± 7.49
20	**Byol**	BlW	C=0.1	1	73.09 ± 1.30	83.97 ± 0.78	67.55 ± 1.33	75.22 ± 1.56	97.86 ± 0.92	61.09 ± 1.72
21	**Byol**	BlW	C=0.7	2	73.12 ± 1.23	83.97 ± 0.69	67.58 ± 1.29	75.24 ± 1.50	97.87 ± 0.76	61.12 ± 1.73
22	**Byol**	BlW	C=1.0	3	73.11 ± 1.25	83.98 ± 0.73	67.58 ± 1.30	75.31 ± 1.63	97.87 ± 0.92	61.21 ± 1.82
23	**Byol**	ChR	cr=[0.33,1.0]	1	77.73 ± 0.80	84.56 ± 1.33	73.26 ± 0.85	80.46 ± 1.41	96.14 ± 3.04	69.20 ± 1.56
24	**Byol**	ChR	cr=[0.5,2.0]	2	76.91 ± 1.57	83.60 ± 3.36	72.83 ± 1.53	79.88 ± 4.19	93.04 ± 10.54	70.24 ± 0.22
25	**Byol**	ChR	cr=[0.33,3.0]	3	78.96 ± 0.55	84.19 ± 1.59	75.21 ± 0.16	80.32 ± 2.13	93.70 ± 5.43	70.35 ± 0.18
26	**Byol**	DTW	w=1,r=10	1	72.45 ± 1.32	82.21 ± 2.51	67.42 ± 1.72	73.84 ± 2.11	93.48 ± 6.24	61.15 ± 2.38
27	**Byol**	DTW	w=3,r=5	2	73.02 ± 1.68	83.85 ± 0.72	67.52 ± 1.73	75.01 ± 1.76	97.53 ± 0.77	60.94 ± 2.03
28	**Byol**	DTW	w=3,r=10	3	72.69 ± 1.37	82.62 ± 2.26	67.58 ± 1.75	74.40 ± 1.72	94.38 ± 5.68	61.50 ± 1.92
29	**Byol**	EM	var=0.01	1	73.13 ± 1.37	83.98 ± 0.80	67.59 ± 1.39	75.24 ± 1.56	97.86 ± 0.92	61.12 ± 1.73
30	**Byol**	EM	var=0.5	2	73.08 ± 1.36	83.92 ± 0.79	67.56 ± 1.39	75.21 ± 1.54	97.77 ± 0.86	61.12 ± 1.73
31	**Byol**	EM	var=1.0	3	73.08 ± 1.27	83.95 ± 0.76	67.54 ± 1.31	75.22 ± 1.56	97.86 ± 0.92	61.09 ± 1.72
32	**Byol**	GN	var=0.01	1	73.11 ± 1.27	83.98 ± 0.73	67.57 ± 1.31	75.28 ± 1.60	97.91 ± 0.86	61.15 ± 1.80
33	**Byol**	GN	var=0.1	2	73.16 ± 1.24	84.04 ± 0.67	67.61 ± 1.30	75.30 ± 1.55	98.01 ± 0.76	61.15 ± 1.80
34	**Byol**	GN	var=1.0	3	72.98 ± 1.18	83.94 ± 0.76	67.42 ± 1.19	75.03 ± 1.63	97.90 ± 0.91	60.83 ± 1.79
35	**Byol**	Neg	True	2	72.00 ± 1.96	81.97 ± 2.22	67.08 ± 2.59	74.25 ± 1.56	93.58 ± 6.99	61.71 ± 1.97
36	**Byol**	TO	mr=[0.1,0.2]	1	74.94 ± 1.06	85.26 ± 1.32	69.38 ± 1.29	75.82 ± 1.87	98.63 ± 1.64	61.59 ± 2.15
37	**Byol**	TO	mr=[0.0,0.5]	2	74.50 ± 1.40	85.90 ± 0.80	68.90 ± 1.65	77.30 ± 1.85	98.91 ± 0.59	63.48 ± 2.66
38	**Byol**	TO	mr=[0.4,0.5]	3	75.60 ± 2.95	86.42 ± 0.20	69.99 ± 3.59	76.99 ± 4.90	99.28 ± 0.49	63.07 ± 6.68
39	**Swav**	BlW	C=0.1	1	66.18 ± 7.59	75.73 ± 6.82	61.68 ± 6.87	54.36 ± 18.78	81.50 ± 17.09	41.33 ± 16.57
40	**Swav**	BlW	C=0.7	2	65.84 ± 11.63	80.30 ± 5.99	61.66 ± 9.06	48.74 ± 32.56	90.92 ± 13.38	36.25 ± 26.21
41	**Swav**	BlW	C=1.0	3	63.67 ± 14.14	73.38 ± 14.26	60.13 ± 12.15	47.07 ± 35.27	74.73 ± 41.22	35.40 ± 27.63
42	**Swav**	ChR	cr=[0.33,1.0]	1	76.60 ± 1.75	82.23 ± 2.00	72.80 ± 1.42	81.33 ± 0.85	98.74 ± 0.76	69.14 ± 0.97
43	**Swav**	ChR	cr=[0.5,2.0]	2	76.74 ± 1.40	83.45 ± 1.03	72.29 ± 1.61	79.93 ± 2.96	97.46 ± 2.15	67.76 ± 3.26
44	**Swav**	ChR	cr=[0.33,3.0]	3	78.88 ± 0.66	84.29 ± 0.75	75.16 ± 0.46	80.78 ± 2.53	97.56 ± 2.18	68.94 ± 2.72
45	**Swav**	DTW	w=1,r=10	1	67.07 ± 1.88	79.69 ± 2.18	61.42 ± 1.50	64.09 ± 3.61	94.32 ± 6.28	48.55 ± 2.79
46	**Swav**	DTW	w=3,r=5	2	70.56 ± 6.18	79.27 ± 3.13	66.04 ± 6.51	66.93 ± 18.13	90.86 ± 8.80	53.98 ± 19.26
47	**Swav**	DTW	w=3,r=10	3	69.39 ± 1.75	80.04 ± 3.71	63.99 ± 1.20	65.46 ± 4.76	93.77 ± 8.65	50.29 ± 3.20
48	**Swav**	EM	var=0.01	1	71.30 ± 0.22	82.16 ± 0.61	65.95 ± 0.32	64.34 ± 0.62	96.21 ± 3.28	48.35 ± 0.45
49	**Swav**	EM	var=0.5	2	70.12 ± 1.14	81.96 ± 2.98	64.45 ± 0.65	62.96 ± 0.61	95.61 ± 5.45	46.99 ± 0.96
50	**Swav**	EM	var=1.0	3	68.36 ± 2.38	78.43 ± 4.58	63.34 ± 1.54	62.65 ± 2.77	92.22 ± 6.27	47.49 ± 2.33
51	**Swav**	GN	var=0.01	1	72.72 ± 2.08	82.12 ± 1.85	67.86 ± 2.35	70.39 ± 6.27	99.17 ± 0.87	54.84 ± 7.57
52	**Swav**	GN	var=0.1	2	70.77 ± 0.37	81.04 ± 1.05	65.66 ± 0.05	64.85 ± 2.11	95.91 ± 4.23	49.00 ± 1.51
53	**Swav**	GN	var=1.0	3	70.14 ± 3.16	80.47 ± 1.10	64.94 ± 3.69	66.42 ± 6.88	96.50 ± 3.37	50.80 ± 7.33
54	**Swav**	Neg	True	2	68.23 ± 4.62	83.86 ± 0.70	62.16 ± 4.55	58.70 ± 11.13	99.30 ± 0.90	42.18 ± 10.65
55	**Swav**	TO	mr=[0.1,0.2]	1	76.56 ± 2.57	86.76 ± 1.79	71.10 ± 2.94	75.09 ± 9.91	97.82 ± 1.62	61.53 ± 12.29
56	**Swav**	TO	mr=[0.0,0.5]	2	71.78 ± 2.20	84.47 ± 2.64	66.23 ± 2.34	69.34 ± 11.19	94.57 ± 6.90	55.22 ± 12.75
57	**Swav**	TO	mr=[0.4,0.5]	3	72.22 ± 1.41	84.67 ± 0.62	66.21 ± 1.74	67.40 ± 2.71	96.73 ± 2.38	51.83 ± 3.81

Table 11: Test Class 1 and Class 2 performance under linear evaluation. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Class 1	Class 2
1	**Simclr**	BlW	C=0.1	1	87.21 ± 0.44 / 82.46 ± 1.47 / 92.56 ± 0.83	65.71 ± 2.64 / 75.68 ± 0.99 / 58.18 ± 4.43
2	**Simclr**	BlW	C=0.7	2	87.29 ± 0.55 / 81.96 ± 0.92 / 93.36 ± 0.07	65.21 ± 2.07 / 77.00 ± 0.87 / 56.57 ± 2.71
3	**Simclr**	BlW	C=1.0	3	87.18 ± 0.65 / 82.68 ± 1.09 / 92.23 ± 1.91	66.17 ± 2.43 / 75.32 ± 2.10 / 59.16 ± 4.41
4	**Simclr**	ChR	cr=[0.33,1.0]	1	86.91 ± 0.65 / 80.71 ± 0.78 / 94.14 ± 0.47	62.47 ± 2.08 / 77.81 ± 1.59 / 52.19 ± 2.19
5	**Simclr**	ChR	cr=[0.5,2.0]	2	86.63 ± 0.30 / 79.65 ± 1.40 / 94.99 ± 1.38	59.80 ± 3.36 / 78.70 ± 1.81 / 48.42 ± 5.14
6	**Simclr**	ChR	cr=[0.33,3.0]	3	86.67 ± 0.25 / 79.95 ± 0.86 / 94.64 ± 1.05	60.56 ± 2.02 / 78.13 ± 1.51 / 49.53 ± 3.14
7	**Simclr**	DTW	w=1,r=10	1	86.45 ± 0.50 / 80.43 ± 1.08 / 93.47 ± 1.14	60.92 ± 2.18 / 74.81 ± 1.80 / 51.50 ± 3.67
8	**Simclr**	DTW	w=3,r=5	2	87.02 ± 0.06 / 83.03 ± 1.12 / 91.43 ± 1.28	67.09 ± 0.38 / 74.79 ± 2.24 / 60.90 ± 1.92
9	**Simclr**	DTW	w=3,r=10	3	87.09 ± 0.53 / 81.34 ± 1.32 / 93.74 ± 1.20	63.53 ± 2.53 / 76.81 ± 1.49 / 54.30 ± 4.16
10	**Simclr**	EM	var=0.01	1	87.27 ± 0.39 / 82.87 ± 0.57 / 92.18 ± 0.92	66.76 ± 1.43 / 75.35 ± 0.70 / 59.96 ± 2.32
11	**Simclr**	EM	var=0.5	2	87.24 ± 0.38 / 82.44 ± 1.32 / 92.66 ± 0.81	65.87 ± 2.14 / 75.97 ± 1.45 / 58.26 ± 3.86
12	**Simclr**	EM	var=1.0	3	87.02 ± 0.60 / 82.15 ± 1.99 / 92.54 ± 1.15	64.95 ± 3.41 / 75.58 ± 0.99 / 57.15 ± 5.85
13	**Simclr**	GN	var=0.01	1	87.26 ± 0.41 / 82.29 ± 1.53 / 92.89 ± 1.02	65.61 ± 2.51 / 76.19 ± 1.41 / 57.77 ± 4.56
14	**Simclr**	GN	var=0.1	2	86.84 ± 0.13 / 82.94 ± 2.09 / 91.22 ± 2.62	66.02 ± 3.35 / 74.15 ± 2.47 / 59.90 ± 7.01
15	**Simclr**	GN	var=1.0	3	87.06 ± 0.28 / 82.06 ± 0.90 / 92.72 ± 0.54	65.11 ± 1.37 / 75.71 ± 1.07 / 57.16 ± 2.56
16	**Simclr**	Neg	True	2	87.78 ± 0.23 / 82.88 ± 0.66 / 93.30 ± 1.00	66.87 ± 1.32 / 77.00 ± 0.92 / 59.14 ± 2.46
17	**Simclr**	TO	mr=[0.1,0.2]	1	85.17 ± 0.39 / 77.84 ± 0.56 / 94.04 ± 0.26	53.99 ± 1.53 / 72.16 ± 0.98 / 43.15 ± 1.81
18	**Simclr**	TO	mr=[0.0,0.5]	2	85.55 ± 0.63 / 77.61 ± 2.77 / 95.43 ± 2.92	52.38 ± 8.99 / 75.30 ± 3.28 / 41.12 ± 11.14
19	**Simclr**	TO	mr=[0.4,0.5]	3	86.01 ± 0.35 / 77.86 ± 0.94 / 96.09 ± 0.61	55.32 ± 2.21 / 79.35 ± 1.18 / 42.52 ± 2.80
20	**Byol**	BlW	C=0.1	1	85.94 ± 0.24 / 79.14 ± 0.86 / 94.04 ± 0.75	58.11 ± 2.28 / 74.93 ± 1.40 / 47.52 ± 3.11
21	**Byol**	BlW	C=0.7	2	85.94 ± 0.21 / 79.16 ± 0.81 / 94.02 ± 0.75	58.16 ± 2.14 / 74.90 ± 1.36 / 47.60 ± 2.95
22	**Byol**	BlW	C=1.0	3	85.94 ± 0.21 / 79.13 ± 0.83 / 94.05 ± 0.74	58.09 ± 2.13 / 74.94 ± 1.34 / 47.48 ± 2.97
23	**Byol**	ChR	cr=[0.33,1.0]	1	87.03 ± 0.51 / 82.52 ± 0.31 / 92.07 ± 1.12	65.72 ± 0.90 / 75.01 ± 2.27 / 58.50 ± 1.05
24	**Byol**	ChR	cr=[0.5,2.0]	2	86.91 ± 0.56 / 81.64 ± 1.79 / 92.93 ± 1.13	63.95 ± 3.92 / 76.11 ± 0.30 / 55.33 ± 5.82
25	**Byol**	ChR	cr=[0.33,3.0]	3	87.41 ± 0.69 / 84.56 ± 0.62 / 90.49 ± 2.18	69.15 ± 0.26 / 74.32 ± 3.23 / 64.79 ± 2.56
26	**Byol**	DTW	w=1,r=10	1	85.69 ± 0.26 / 79.09 ± 1.15 / 93.53 ± 1.58	57.82 ± 2.76 / 74.05 ± 2.29 / 47.57 ± 4.22
27	**Byol**	DTW	w=3,r=5	2	85.93 ± 0.31 / 79.17 ± 1.18 / 93.98 ± 1.00	58.13 ± 3.11 / 74.85 ± 1.37 / 47.64 ± 4.29
28	**Byol**	DTW	w=3,r=10	3	85.77 ± 0.24 / 79.13 ± 1.29 / 93.67 ± 1.60	57.89 ± 3.25 / 74.36 ± 2.07 / 47.58 ± 4.81
29	**Byol**	EM	var=0.01	1	85.95 ± 0.26 / 79.17 ± 0.89 / 94.02 ± 0.76	58.18 ± 2.41 / 74.91 ± 1.40 / 47.63 ± 3.24
30	**Byol**	EM	var=0.5	2	85.93 ± 0.27 / 79.14 ± 0.91 / 94.01 ± 0.78	58.10 ± 2.42 / 74.85 ± 1.45 / 47.55 ± 3.29
31	**Byol**	EM	var=1.0	3	85.93 ± 0.23 / 79.13 ± 0.86 / 94.02 ± 0.77	58.08 ± 2.24 / 74.87 ± 1.40 / 47.51 ± 3.10
32	**Byol**	GN	var=0.01	1	85.94 ± 0.22 / 79.14 ± 0.84 / 94.03 ± 0.77	58.11 ± 2.21 / 74.90 ± 1.38 / 47.53 ± 3.07
33	**Byol**	GN	var=0.1	2	85.96 ± 0.22 / 79.18 ± 0.86 / 94.03 ± 0.76	58.22 ± 2.20 / 74.93 ± 1.34 / 47.66 ± 3.07
34	**Byol**	GN	var=1.0	3	85.91 ± 0.21 / 79.09 ± 0.78 / 94.03 ± 0.73	58.00 ± 2.07 / 74.82 ± 1.45 / 47.41 ± 2.85
35	**Byol**	Neg	True	2	85.45 ± 0.56 / 78.62 ± 2.09 / 93.66 ± 1.88	56.29 ± 5.62 / 73.72 ± 2.00 / 45.89 ± 7.55
36	**Byol**	TO	mr=[0.1,0.2]	1	86.76 ± 0.32 / 80.83 ± 1.53 / 93.69 ± 2.26	62.25 ± 2.77 / 76.30 ± 3.77 / 52.87 ± 5.47
37	**Byol**	TO	mr=[0.0,0.5]	2	86.77 ± 0.21 / 79.45 ± 0.85 / 95.60 ± 1.27	59.42 ± 2.26 / 79.33 ± 2.46 / 47.61 ± 3.38
38	**Byol**	TO	mr=[0.4,0.5]	3	87.29 ± 0.36 / 80.70 ± 1.60 / 95.08 ± 1.37	62.52 ± 3.63 / 79.27 ± 1.12 / 51.81 ± 5.50
39	**Swav**	BlW	C=0.1	1	83.87 ± 0.86 / 80.54 ± 1.74 / 87.51 ± 0.78	60.32 ± 3.24 / 65.15 ± 1.80 / 56.21 ± 4.42
40	**Swav**	BlW	C=0.7	2	84.93 ± 1.72 / 81.90 ± 1.75 / 88.25 ± 3.18	63.85 ± 2.32 / 68.06 ± 4.09 / 60.47 ± 5.05
41	**Swav**	BlW	C=1.0	3	83.09 ± 0.45 / 81.32 ± 4.82 / 85.37 ± 5.40	60.85 ± 6.97 / 64.08 ± 3.33 / 59.61 ± 14.19
42	**Swav**	ChR	cr=[0.33,1.0]	1	83.74 ± 2.33 / 83.71 ± 0.44 / 83.84 ± 4.36	64.71 ± 2.44 / 64.25 ± 5.73 / 65.42 ± 1.25
43	**Swav**	ChR	cr=[0.5,2.0]	2	85.70 ± 0.81 / 82.47 ± 0.98 / 89.25 ± 2.61	64.59 ± 1.18 / 70.43 ± 3.55 / 59.86 ± 3.73
44	**Swav**	ChR	cr=[0.33,3.0]	3	86.27 ± 0.44 / 85.90 ± 1.01 / 86.68 ± 1.68	69.60 ± 0.96 / 69.42 ± 1.58 / 69.87 ± 3.16
45	**Swav**	DTW	w=1,r=10	1	83.90 ± 0.32 / 77.41 ± 1.00 / 91.59 ± 1.01	53.23 ± 2.54 / 67.34 ± 0.71 / 44.10 ± 3.68
46	**Swav**	DTW	w=3,r=5	2	83.95 ± 0.65 / 80.75 ± 1.81 / 87.51 ± 3.33	60.80 ± 2.69 / 66.20 ± 2.49 / 56.63 ± 6.72
47	**Swav**	DTW	w=3,r=10	3	84.17 ± 1.30 / 79.78 ± 0.99 / 89.14 ± 3.34	58.55 ± 2.49 / 66.55 ± 4.66 / 52.53 ± 4.53
48	**Swav**	EM	var=0.01	1	85.21 ± 0.56 / 82.77 ± 0.92 / 87.83 ± 2.17	64.35 ± 0.93 / 67.49 ± 2.36 / 61.66 ± 3.47
49	**Swav**	EM	var=0.5	2	85.31 ± 1.30 / 81.19 ± 0.58 / 89.90 ± 2.39	62.10 ± 1.54 / 69.09 ± 3.91 / 56.46 ± 1.19
50	**Swav**	EM	var=1.0	3	82.87 ± 2.53 / 80.75 ± 1.04 / 85.18 ± 4.85	59.55 ± 3.07 / 62.30 ± 6.92 / 57.35 ± 2.97
51	**Swav**	GN	var=0.01	1	83.67 ± 2.68 / 83.21 ± 0.25 / 84.22 ± 5.12	64.11 ± 1.85 / 63.97 ± 5.37 / 64.53 ± 2.00
52	**Swav**	GN	var=0.1	2	84.11 ± 0.79 / 82.67 ± 0.87 / 85.64 ± 2.42	63.35 ± 1.03 / 64.54 ± 2.46 / 62.35 ± 3.34
53	**Swav**	GN	var=1.0	3	83.57 ± 0.56 / 81.13 ± 1.90 / 86.24 ± 3.00	60.43 ± 3.01 / 63.76 ± 2.04 / 57.77 ± 6.58
54	**Swav**	Neg	True	2	85.51 ± 0.79 / 79.74 ± 1.73 / 92.23 ± 1.82	60.48 ± 2.43 / 72.55 ± 2.89 / 52.06 ± 4.43
55	**Swav**	TO	mr=[0.1,0.2]	1	88.23 ± 1.21 / 82.38 ± 1.70 / 94.98 ± 1.02	66.37 ± 3.93 / 80.08 ± 3.19 / 56.78 ± 4.95
56	**Swav**	TO	mr=[0.0,0.5]	2	86.99 ± 1.22 / 79.63 ± 2.74 / 95.92 ± 1.00	59.02 ± 6.98 / 79.20 ± 1.42 / 47.55 ± 9.28
57	**Swav**	TO	mr=[0.4,0.5]	3	87.05 ± 0.48 / 80.79 ± 0.52 / 94.37 ± 0.91	62.19 ± 1.45 / 76.47 ± 2.11 / 52.43 ± 1.65

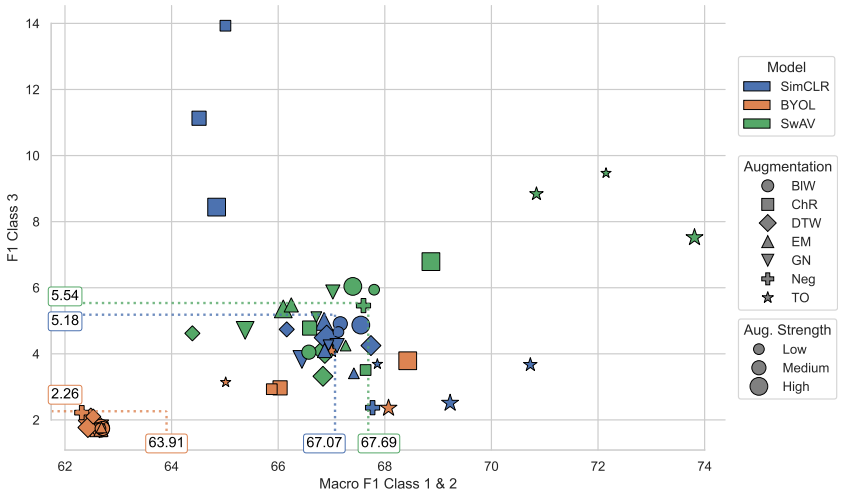
Table 12: Test Macro and Class 3 performance of SSL models after fine-tuning. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Macro			Class 3		
1	**Simclr**	BlW	C=0.1	1	75.93 ± 1.65	87.10 ± 0.99	70.05 ± 2.02	70.01 ± 2.69	95.71 ± 2.82	55.28 ± 3.55
2	**Simclr**	BlW	C=0.7	2	77.52 ± 0.54	87.45 ± 1.24	71.92 ± 1.08	71.25 ± 1.89	96.68 ± 2.15	56.49 ± 3.12
3	**Simclr**	BlW	C=1.0	3	77.65 ± 0.82	88.13 ± 0.48	71.97 ± 1.15	69.76 ± 1.58	98.71 ± 0.97	53.95 ± 1.92
4	**Simclr**	ChR	cr=[0.33,1.0]	1	80.74 ± 1.57	88.57 ± 0.91	75.99 ± 1.52	77.03 ± 4.85	99.82 ± 0.16	62.89 ± 6.62
5	**Simclr**	ChR	cr=[0.5,2.0]	2	80.29 ± 3.07	88.34 ± 0.97	75.47 ± 3.89	76.26 ± 6.16	99.70 ± 0.16	62.01 ± 8.24
6	**Simclr**	ChR	cr=[0.33,3.0]	3	80.17 ± 2.11	89.04 ± 0.62	74.80 ± 2.58	77.15 ± 5.09	99.13 ± 0.82	63.30 ± 6.66
7	**Simclr**	DTW	w=1,r=10	1	77.28 ± 0.67	86.87 ± 1.74	71.77 ± 0.25	71.42 ± 2.98	95.01 ± 4.54	57.23 ± 2.39
8	**Simclr**	DTW	w=3,r=5	2	79.03 ± 1.55	88.37 ± 1.20	73.55 ± 1.63	74.46 ± 2.65	99.59 ± 0.48	59.50 ± 3.18
9	**Simclr**	DTW	w=3,r=10	3	79.18 ± 2.72	88.47 ± 0.33	73.77 ± 3.53	75.00 ± 6.85	99.52 ± 0.57	60.47 ± 8.94
10	**Simclr**	EM	var=0.01	1	77.17 ± 0.97	88.00 ± 0.46	71.43 ± 1.23	68.48 ± 2.27	98.41 ± 1.33	52.54 ± 2.60
11	**Simclr**	EM	var=0.5	2	76.78 ± 0.01	87.34 ± 0.36	71.07 ± 0.14	69.66 ± 3.22	96.66 ± 1.49	54.51 ± 3.93
12	**Simclr**	EM	var=1.0	3	76.85 ± 1.42	87.85 ± 0.19	70.99 ± 1.86	70.74 ± 0.75	97.56 ± 0.65	55.49 ± 0.81
13	**Simclr**	GN	var=0.01	1	77.09 ± 1.45	87.45 ± 1.15	71.46 ± 1.73	68.81 ± 2.67	96.84 ± 3.89	53.42 ± 2.88
14	**Simclr**	GN	var=0.1	2	77.82 ± 0.49	86.77 ± 1.92	72.64 ± 0.55	71.89 ± 1.80	96.23 ± 4.04	57.43 ± 2.03
15	**Simclr**	GN	var=1.0	3	77.13 ± 0.20	87.06 ± 1.55	71.67 ± 0.38	68.31 ± 1.87	95.52 ± 3.77	53.22 ± 2.17
16	**Simclr**	Neg	True	2	78.23 ± 1.12	86.23 ± 3.59	73.70 ± 2.03	71.88 ± 3.16	93.29 ± 11.35	59.53 ± 8.56
17	**Simclr**	TO	mr=[0.1,0.2]	1	78.48 ± 0.50	87.56 ± 0.97	74.02 ± 0.34	70.82 ± 0.51	99.47 ± 0.36	54.99 ± 0.72
18	**Simclr**	TO	mr=[0.0,0.5]	2	78.84 ± 2.82	87.74 ± 2.79	74.58 ± 1.68	74.96 ± 6.57	99.03 ± 1.19	60.71 ± 9.29
19	**Simclr**	TO	mr=[0.4,0.5]	3	79.17 ± 1.21	87.95 ± 0.98	74.15 ± 1.84	73.47 ± 2.42	98.88 ± 1.02	58.47 ± 2.77
20	**Byol**	BlW	C=0.1	1	77.60 ± 4.01	86.01 ± 2.96	72.97 ± 3.75	71.52 ± 6.80	97.12 ± 2.38	56.78 ± 7.82
21	**Byol**	BlW	C=0.7	2	78.67 ± 1.02	87.71 ± 1.57	73.38 ± 0.90	73.28 ± 1.19	97.14 ± 4.42	58.97 ± 3.13
22	**Byol**	BlW	C=1.0	3	77.44 ± 3.07	85.71 ± 2.42	72.79 ± 3.42	70.83 ± 3.99	96.09 ± 3.47	56.11 ± 3.85
23	**Byol**	ChR	cr=[0.33,1.0]	1	80.56 ± 1.50	88.80 ± 0.71	75.56 ± 1.60	76.23 ± 5.10	99.47 ± 0.43	61.98 ± 6.77
24	**Byol**	ChR	cr=[0.5,2.0]	2	79.57 ± 2.01	85.71 ± 4.11	75.56 ± 1.35	75.42 ± 4.07	90.28 ± 12.65	65.22 ± 1.33
25	**Byol**	ChR	cr=[0.33,3.0]	3	79.77 ± 1.11	88.15 ± 0.84	75.14 ± 1.54	73.49 ± 3.07	99.12 ± 0.93	58.44 ± 3.65
26	**Byol**	DTW	w=1,r=10	1	75.95 ± 4.46	85.74 ± 5.05	70.61 ± 3.99	68.42 ± 5.28	93.25 ± 10.54	54.19 ± 4.07
27	**Byol**	DTW	w=3,r=5	2	77.39 ± 3.26	85.66 ± 1.48	72.76 ± 4.10	71.83 ± 5.48	96.45 ± 2.33	57.58 ± 7.75
28	**Byol**	DTW	w=3,r=10	3	77.89 ± 3.10	87.47 ± 0.26	72.66 ± 4.15	71.80 ± 5.91	98.61 ± 0.84	56.64 ± 7.02
29	**Byol**	EM	var=0.01	1	75.66 ± 5.95	85.96 ± 4.30	70.26 ± 6.01	70.68 ± 8.35	93.72 ± 8.92	56.90 ± 8.45
30	**Byol**	EM	var=0.5	2	77.28 ± 1.44	86.58 ± 2.77	71.96 ± 0.90	70.89 ± 2.00	94.37 ± 8.22	56.96 ± 2.20
31	**Byol**	EM	var=1.0	3	75.49 ± 2.67	86.72 ± 2.57	69.71 ± 2.97	68.07 ± 3.02	96.98 ± 3.39	52.45 ± 2.69
32	**Byol**	GN	var=0.01	1	74.90 ± 3.29	85.22 ± 2.07	69.57 ± 3.55	70.49 ± 2.64	93.81 ± 5.19	56.70 ± 4.65
33	**Byol**	GN	var=0.1	2	77.97 ± 3.59	87.36 ± 1.84	72.60 ± 3.93	72.43 ± 8.38	97.41 ± 3.62	57.88 ± 9.31
34	**Byol**	GN	var=1.0	3	78.47 ± 2.24	87.90 ± 1.41	73.30 ± 2.26	72.03 ± 4.73	99.67 ± 0.45	56.52 ± 5.79
35	**Byol**	Neg	True	2	76.50 ± 3.79	86.30 ± 2.17	71.20 ± 4.25	70.42 ± 3.00	97.01 ± 2.60	55.28 ± 2.86
36	**Byol**	TO	mr=[0.1,0.2]	1	78.42 ± 0.66	87.33 ± 1.81	73.13 ± 1.12	73.53 ± 1.30	96.70 ± 5.19	59.53 ± 3.81
37	**Byol**	TO	mr=[0.0,0.5]	2	77.89 ± 0.73	87.39 ± 0.15	72.40 ± 1.06	72.03 ± 0.29	95.93 ± 0.67	57.67 ± 0.14
38	**Byol**	TO	mr=[0.4,0.5]	3	77.74 ± 3.10	85.65 ± 3.96	73.28 ± 2.88	75.21 ± 5.00	90.06 ± 12.07	64.84 ± 1.18
39	**Swav**	BlW	C=0.1	1	74.81 ± 0.41	86.30 ± 1.31	69.55 ± 0.95	66.78 ± 3.02	99.77 ± 0.08	50.24 ± 3.46
40	**Swav**	BlW	C=0.7	2	68.27 ± 6.99	86.29 ± 0.65	63.43 ± 5.28	46.00 ± 21.49	98.81 ± 1.50	31.62 ± 17.48
41	**Swav**	BlW	C=1.0	3	74.10 ± 2.34	86.73 ± 0.58	68.27 ± 2.45	63.30 ± 6.05	98.84 ± 0.26	46.76 ± 6.44
42	**Swav**	ChR	cr=[0.33,1.0]	1	80.28 ± 1.11	87.84 ± 0.53	75.69 ± 1.41	76.58 ± 5.36	99.01 ± 0.94	62.71 ± 7.65
43	**Swav**	ChR	cr=[0.5,2.0]	2	78.80 ± 0.72	87.43 ± 0.12	74.18 ± 0.87	72.35 ± 2.53	99.79 ± 0.25	56.78 ± 3.05
44	**Swav**	ChR	cr=[0.33,3.0]	3	78.67 ± 1.42	87.68 ± 0.56	73.56 ± 1.72	72.48 ± 3.59	99.22 ± 1.11	57.17 ± 4.37
45	**Swav**	DTW	w=1,r=10	1	75.98 ± 0.66	87.28 ± 0.44	70.23 ± 0.69	67.26 ± 0.65	99.65 ± 0.34	50.77 ± 0.75
46	**Swav**	DTW	w=3,r=5	2	73.18 ± 2.07	87.19 ± 0.56	67.32 ± 2.18	58.84 ± 5.92	98.91 ± 0.47	42.04 ± 6.06
47	**Swav**	DTW	w=3,r=10	3	76.67 ± 1.20	86.86 ± 0.46	71.28 ± 1.28	68.08 ± 3.96	97.89 ± 1.68	52.24 ± 4.16
48	**Swav**	EM	var=0.01	1	76.03 ± 1.00	87.49 ± 0.21	70.09 ± 1.21	68.63 ± 3.41	99.71 ± 0.38	52.39 ± 3.94
49	**Swav**	EM	var=0.5	2	75.64 ± 0.60	87.52 ± 0.53	69.89 ± 1.12	65.61 ± 0.12	99.29 ± 0.62	49.00 ± 0.10
50	**Swav**	EM	var=1.0	3	72.60 ± 2.92	86.99 ± 1.05	66.69 ± 3.06	58.74 ± 7.08	99.17 ± 0.72	41.95 ± 6.96
51	**Swav**	GN	var=0.01	1	77.45 ± 0.53	87.54 ± 0.39	72.07 ± 0.70	69.42 ± 2.87	99.08 ± 0.19	53.48 ± 3.43
52	**Swav**	GN	var=0.1	2	74.99 ± 1.63	87.21 ± 0.71	69.51 ± 1.59	63.44 ± 4.77	99.53 ± 0.46	46.70 ± 5.28
53	**Swav**	GN	var=1.0	3	76.94 ± 2.18	86.76 ± 1.09	71.58 ± 2.56	71.40 ± 5.01	98.40 ± 0.62	56.22 ± 6.52
54	**Swav**	Neg	True	2	77.08 ± 1.51	86.99 ± 0.48	71.92 ± 2.09	69.38 ± 4.36	99.36 ± 0.68	53.42 ± 5.13
55	**Swav**	TO	mr=[0.1,0.2]	1	81.85 ± 0.97	88.94 ± 0.67	77.27 ± 0.96	79.56 ± 3.24	97.87 ± 2.74	67.20 ± 5.20
56	**Swav**	TO	mr=[0.0,0.5]	2	80.53 ± 2.19	88.23 ± 0.57	75.75 ± 2.81	75.80 ± 5.89	97.44 ± 1.67	62.18 ± 7.26
57	**Swav**	TO	mr=[0.4,0.5]	3	79.56 ± 1.14	87.66 ± 1.74	75.16 ± 1.07	74.15 ± 4.15	99.61 ± 0.14	59.17 ± 5.22

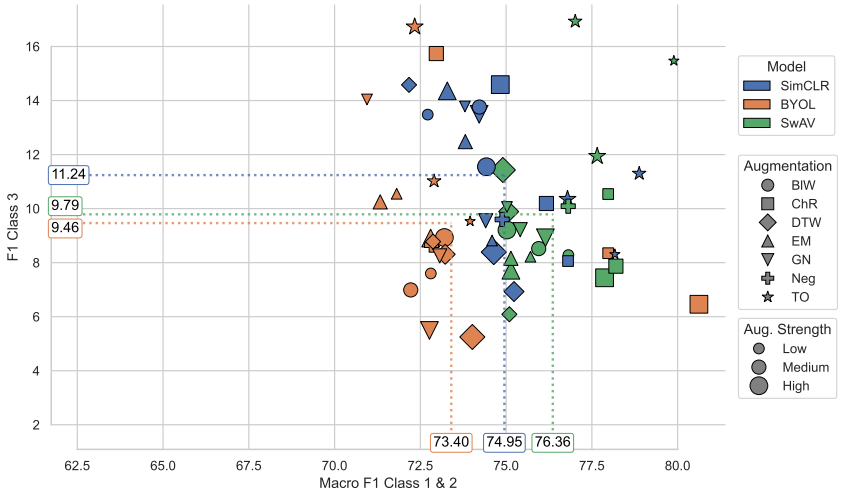
Table 13: Test Class 1 and Class 2 performance of SSL models after fine-tuning. Each cell shows F1 (%) / Precision (%) / Recall (%) with standard deviation over 3 random seeds.

#	Model	Aug.	Param.	Str.	Class 1	Class 2
1	**Simclr**	BlW	C=0.1	1	89.14 ± 0.22 / 83.30 ± 1.18 / 95.88 ± 1.09	68.65 ± 2.15 / 82.28 ± 1.27 / 59.00 ± 3.90
2	**Simclr**	BlW	C=0.7	2	89.35 ± 0.26 / 85.00 ± 1.32 / 94.20 ± 1.25	71.95 ± 1.42 / 80.66 ± 2.59 / 65.07 ± 3.49
3	**Simclr**	BlW	C=1.0	3	89.72 ± 0.02 / 86.24 ± 1.04 / 93.51 ± 1.25	73.48 ± 1.43 / 79.43 ± 1.48 / 68.45 ± 3.42
4	**Simclr**	ChR	cr=[0.33,1.0]	1	89.74 ± 0.46 / 88.05 ± 1.54 / 91.56 ± 2.67	75.44 ± 0.64 / 77.85 ± 4.33 / 73.52 ± 4.75
5	**Simclr**	ChR	cr=[0.5,2.0]	2	89.64 ± 0.71 / 87.75 ± 1.82 / 91.66 ± 1.99	74.97 ± 2.53 / 77.57 ± 2.87 / 72.73 ± 5.10
6	**Simclr**	ChR	cr=[0.33,3.0]	3	90.02 ± 0.32 / 85.66 ± 1.30 / 94.85 ± 0.94	73.35 ± 1.99 / 82.33 ± 1.48 / 66.24 ± 3.87
7	**Simclr**	DTW	w=1,r=10	1	89.41 ± 0.21 / 84.64 ± 0.79 / 94.76 ± 0.83	71.03 ± 1.09 / 80.96 ± 1.33 / 63.31 ± 2.40
8	**Simclr**	DTW	w=3,r=5	2	89.52 ± 0.66 / 85.97 ± 0.87 / 93.40 ± 1.82	73.10 ± 1.55 / 79.55 ± 3.46 / 67.74 ± 2.90
9	**Simclr**	DTW	w=3,r=10	3	89.62 ± 0.41 / 85.78 ± 1.11 / 93.83 ± 1.01	72.92 ± 1.92 / 80.09 ± 0.83 / 66.99 ± 3.31
10	**Simclr**	EM	var=0.01	1	89.77 ± 0.12 / 86.17 ± 0.77 / 93.69 ± 1.01	73.26 ± 0.85 / 79.41 ± 1.24 / 68.05 ± 2.37
11	**Simclr**	EM	var=0.5	2	89.36 ± 0.38 / 84.88 ± 1.83 / 94.39 ± 1.73	71.31 ± 2.88 / 80.49 ± 2.50 / 64.31 ± 5.99
12	**Simclr**	EM	var=1.0	3	89.24 ± 0.25 / 84.24 ± 2.23 / 94.96 ± 2.48	70.57 ± 3.90 / 81.73 ± 2.66 / 62.53 ± 7.93
13	**Simclr**	GN	var=0.01	1	89.62 ± 0.08 / 85.82 ± 1.11 / 93.80 ± 1.35	72.84 ± 1.68 / 79.70 ± 1.31 / 67.17 ± 3.82
14	**Simclr**	GN	var=0.1	2	89.03 ± 0.25 / 85.91 ± 1.80 / 92.46 ± 2.59	72.53 ± 2.33 / 78.16 ± 3.64 / 68.02 ± 6.04
15	**Simclr**	GN	var=1.0	3	89.75 ± 0.37 / 86.26 ± 1.37 / 93.57 ± 1.22	73.32 ± 1.74 / 79.39 ± 1.72 / 68.23 ± 3.92
16	**Simclr**	Neg	True	2	89.56 ± 0.29 / 86.48 ± 1.90 / 92.93 ± 1.97	73.25 ± 2.41 / 78.93 ± 2.71 / 68.66 ± 6.22
17	**Simclr**	TO	mr=[0.1,0.2]	1	88.87 ± 1.22 / 90.08 ± 2.17 / 87.86 ± 4.56	75.75 ± 0.21 / 73.14 ± 5.44 / 79.21 ± 6.27
18	**Simclr**	TO	mr=[0.0,0.5]	2	87.94 ± 4.13 / 88.69 ± 4.52 / 88.17 ± 11.50	73.62 ± 2.78 / 75.50 ± 13.14 / 74.87 ± 13.47
19	**Simclr**	TO	mr=[0.4,0.5]	3	89.51 ± 0.76 / 87.67 ± 2.05 / 91.54 ± 3.55	74.55 ± 1.43 / 77.29 ± 4.27 / 72.44 ± 6.10
20	**Byol**	BlW	C=0.1	1	88.19 ± 2.31 / 87.62 ± 0.86 / 88.88 ± 4.86	73.08 ± 3.04 / 73.30 ± 7.06 / 73.25 ± 2.99
21	**Byol**	BlW	C=0.7	2	89.66 ± 0.56 / 86.00 ± 1.91 / 93.70 ± 1.40	73.06 ± 2.99 / 80.00 ± 2.06 / 67.47 ± 5.92
22	**Byol**	BlW	C=1.0	3	88.46 ± 1.06 / 87.56 ± 2.70 / 89.44 ± 1.53	73.03 ± 4.22 / 73.47 ± 1.88 / 72.82 ± 7.38
23	**Byol**	ChR	cr=[0.33,1.0]	1	90.02 ± 0.24 / 87.53 ± 1.27 / 92.69 ± 1.66	75.44 ± 0.87 / 79.40 ± 2.95 / 72.02 ± 3.73
24	**Byol**	ChR	cr=[0.5,2.0]	2	89.72 ± 0.16 / 86.30 ± 1.90 / 93.48 ± 2.06	73.56 ± 2.18 / 80.55 ± 2.80 / 67.99 ± 5.94
25	**Byol**	ChR	cr=[0.33,3.0]	3	89.64 ± 0.65 / 89.33 ± 2.35 / 90.09 ± 3.66	76.19 ± 1.09 / 75.99 ± 4.20 / 76.89 ± 6.52
26	**Byol**	DTW	w=1,r=10	1	89.10 ± 1.54 / 84.69 ± 3.22 / 94.06 ± 0.91	70.33 ± 6.93 / 79.30 ± 2.15 / 63.58 ± 9.97
27	**Byol**	DTW	w=3,r=5	2	88.16 ± 1.52 / 86.92 ± 2.16 / 89.54 ± 3.58	72.17 ± 3.18 / 73.60 ± 5.50 / 71.17 ± 5.85
28	**Byol**	DTW	w=3,r=10	3	89.12 ± 0.05 / 86.54 ± 2.93 / 92.01 ± 3.36	72.74 ± 3.44 / 77.27 ± 3.21 / 69.34 ± 8.82
29	**Byol**	EM	var=0.01	1	88.81 ± 1.52 / 83.20 ± 3.66 / 95.34 ± 1.39	67.50 ± 8.35 / 80.96 ± 0.43 / 58.52 ± 11.78
30	**Byol**	EM	var=0.5	2	89.42 ± 0.39 / 85.06 ± 1.47 / 94.27 ± 0.96	71.53 ± 2.41 / 80.31 ± 1.37 / 64.64 ± 4.69
31	**Byol**	EM	var=1.0	3	88.84 ± 1.13 / 84.40 ± 2.90 / 93.88 ± 2.67	69.55 ± 4.81 / 78.78 ± 4.56 / 62.81 ± 8.79
32	**Byol**	GN	var=0.01	1	88.09 ± 1.41 / 82.72 ± 3.88 / 94.38 ± 2.38	66.11 ± 7.12 / 79.14 ± 4.58 / 57.63 ± 11.88
33	**Byol**	GN	var=0.1	2	89.33 ± 0.52 / 85.59 ± 1.26 / 93.46 ± 1.80	72.14 ± 2.09 / 79.08 ± 2.87 / 66.46 ± 4.02
34	**Byol**	GN	var=1.0	3	89.25 ± 1.03 / 87.41 ± 1.31 / 91.25 ± 3.33	74.14 ± 1.28 / 76.64 ± 4.83 / 72.12 ± 4.20
35	**Byol**	Neg	True	2	88.68 ± 1.44 / 85.44 ± 4.11 / 92.33 ± 2.03	70.39 ± 7.44 / 76.45 ± 1.25 / 65.98 ± 12.26
36	**Byol**	TO	mr=[0.1,0.2]	1	89.45 ± 0.48 / 85.46 ± 0.52 / 93.82 ± 0.50	72.28 ± 1.17 / 79.82 ± 1.26 / 66.05 ± 1.39
37	**Byol**	TO	mr=[0.0,0.5]	2	89.63 ± 0.19 / 85.21 ± 1.47 / 94.57 ± 1.51	72.01 ± 1.97 / 81.04 ± 2.04 / 64.96 ± 4.57
38	**Byol**	TO	mr=[0.4,0.5]	3	89.09 ± 0.63 / 83.60 ± 3.08 / 95.50 ± 2.52	68.91 ± 5.41 / 83.28 ± 3.68 / 59.51 ± 10.06
39	**Swav**	BlW	C=0.1	1	87.44 ± 1.70 / 86.26 ± 4.09 / 89.12 ± 6.98	70.22 ± 3.29 / 72.86 ± 6.87 / 69.30 ± 12.36
40	**Swav**	BlW	C=0.7	2	88.11 ± 0.44 / 84.93 ± 1.58 / 91.59 ± 2.30	70.70 ± 2.78 / 75.13 ± 2.40 / 67.09 ± 6.34
41	**Swav**	BlW	C=1.0	3	88.57 ± 0.37 / 84.97 ± 1.15 / 92.52 ± 1.82	70.44 ± 1.43 / 76.37 ± 2.51 / 65.52 ± 3.86
42	**Swav**	ChR	cr=[0.33,1.0]	1	89.39 ± 0.59 / 87.93 ± 1.68 / 90.95 ± 1.71	74.86 ± 1.79 / 76.60 ± 2.42 / 73.39 ± 4.68
43	**Swav**	ChR	cr=[0.5,2.0]	2	88.87 ± 0.27 / 89.12 ± 0.64 / 88.64 ± 1.03	75.19 ± 0.24 / 73.40 ± 1.15 / 77.10 ± 1.65
44	**Swav**	ChR	cr=[0.33,3.0]	3	89.33 ± 0.21 / 87.49 ± 1.28 / 91.27 ± 0.96	74.18 ± 1.55 / 76.32 ± 0.79 / 72.24 ± 3.58
45	**Swav**	DTW	w=1,r=10	1	88.90 ± 0.35 / 85.81 ± 0.51 / 92.22 ± 0.34	71.78 ± 1.00 / 76.37 ± 0.88 / 67.71 ± 1.29
46	**Swav**	DTW	w=3,r=5	2	88.57 ± 0.42 / 85.14 ± 0.89 / 92.33 ± 1.98	72.14 ± 0.93 / 77.53 ± 2.56 / 67.60 ± 3.45
47	**Swav**	DTW	w=3,r=10	3	89.08 ± 0.20 / 86.64 ± 0.46 / 91.66 ± 0.55	72.86 ± 0.52 / 76.04 ± 0.52 / 69.94 ± 1.30
48	**Swav**	EM	var=0.01	1	88.88 ± 0.25 / 84.76 ± 0.12 / 93.43 ± 0.67	70.58 ± 0.20 / 77.99 ± 1.10 / 64.46 ± 0.44
49	**Swav**	EM	var=0.5	2	89.16 ± 0.29 / 86.12 ± 1.90 / 92.51 ± 2.84	72.15 ± 1.93 / 77.17 ± 3.89 / 68.16 ± 6.14
50	**Swav**	EM	var=1.0	3	88.52 ± 0.65 / 84.78 ± 1.78 / 92.69 ± 3.04	70.54 ± 1.75 / 77.02 ± 4.02 / 65.43 ± 5.65
51	**Swav**	GN	var=0.01	1	89.31 ± 0.38 / 87.06 ± 0.90 / 91.70 ± 0.87	73.63 ± 1.13 / 76.48 ± 1.09 / 71.03 ± 2.50
52	**Swav**	GN	var=0.1	2	88.78 ± 0.61 / 86.88 ± 2.36 / 90.89 ± 3.30	72.74 ± 2.00 / 75.22 ± 4.29 / 70.96 ± 7.03
53	**Swav**	GN	var=1.0	3	88.50 ± 0.93 / 85.44 ± 1.76 / 91.89 ± 3.72	70.92 ± 1.50 / 76.44 ± 5.15 / 66.61 ± 5.85
54	**Swav**	Neg	True	2	88.71 ± 0.48 / 87.33 ± 1.01 / 90.17 ± 2.07	73.14 ± 0.62 / 74.28 ± 2.21 / 72.16 ± 3.17
55	**Swav**	TO	mr=[0.1,0.2]	1	90.36 ± 0.26 / 87.35 ± 2.31 / 93.66 ± 2.40	75.64 ± 2.47 / 81.59 ± 3.88 / 70.96 ± 6.92
56	**Swav**	TO	mr=[0.0,0.5]	2	90.14 ± 0.43 / 87.73 ± 1.16 / 92.71 ± 2.17	75.67 ± 0.88 / 79.50 ± 2.90 / 72.37 ± 3.70
57	**Swav**	TO	mr=[0.4,0.5]	3	88.91 ± 2.01 / 89.49 ± 1.13 / 88.46 ± 4.76	75.61 ± 1.79 / 73.87 ± 5.97 / 77.85 ± 3.72

B Augmentation Figures for Validation Data



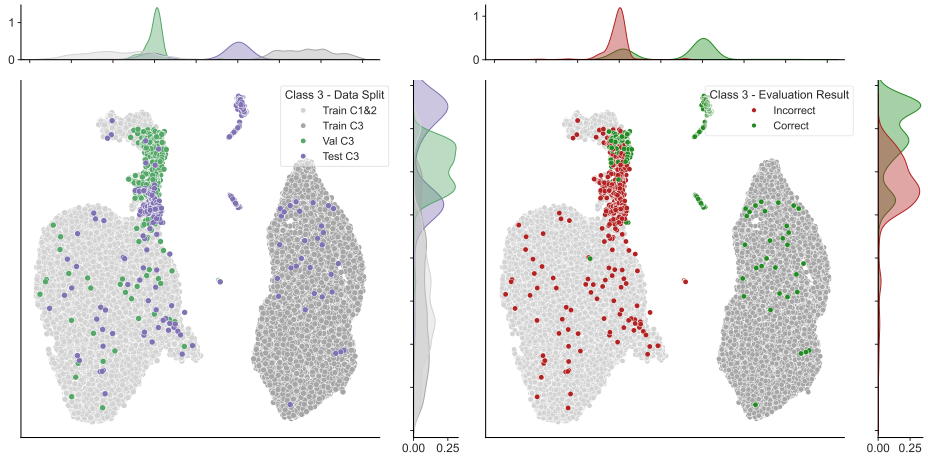
(a) Linear evaluation.



(b) Fine-tuned evaluation.

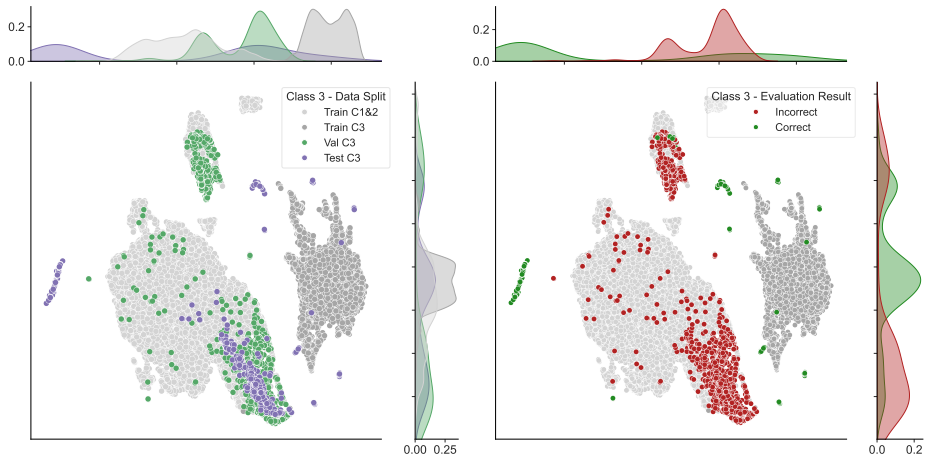
Fig. 7: Comparison of validation performance for SSL models and different augmentations through plotting the macro F1 score for class 1and 2 (x-axis) against the F1 score for class 3 (y-axis) after linear (a) and fine-tuned (b) evaluation. The marker size indicates the augmentation strength while the marker shape encodes the augmentation type. The dotted lines and values plotted on the axis are the average scores for each model across the different configurations.

C UMAP Projection of SSL Models' Representation Space



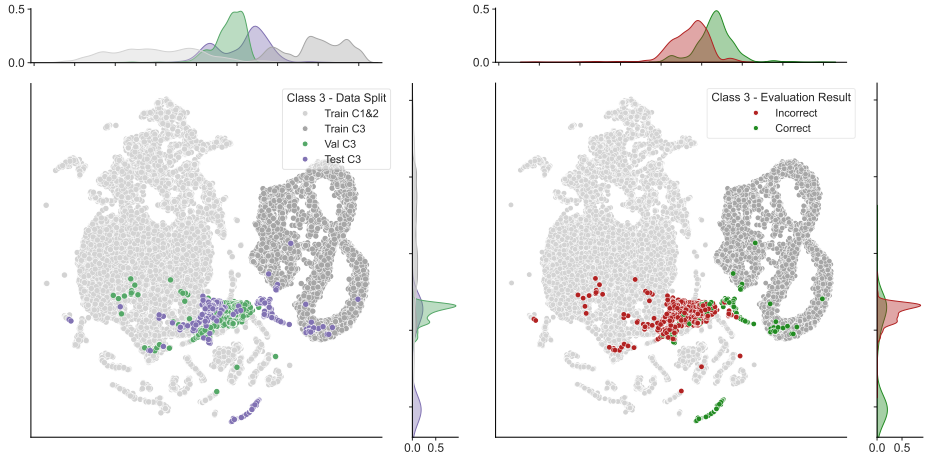
(a) SimCLR (b) SimCLR – Evaluation Result

Fig. 8: UMAP projections of class 3 representations in 2D for SimCLR. The validation and test representations for low-quality signals are plotted on top of all classes of the annotated train subset, to identify their overlap. The evaluation results for SimCLR (b) are shown based on the linear evaluation performance.



(a) BYOL (b) BYOL – Evaluation Result

Fig. 9: UMAP projections of class 3 representations in 2D for BYOL. The validation and test representations for low-quality signals are plotted on top of all classes of the annotated train subset, to identify their overlap. The evaluation results for BYOL (b) are shown based on the linear evaluation performance.



(a) SwAV (b) SwAV – Evaluation Result

Fig. 10: UMAP projections of class 3 representations in 2D for SwAV. The validation and test representations for low-quality signals are plotted on top of all classes of the annotated train subset, to identify their overlap. The evaluation results for SwAV (b) are shown based on the linear evaluation performance.