



## **Self-Supervised Learning for Image Classification**

**Submitted by:**

<b>NAME</b>	<b>ID</b>
<b>Md Shadman Shakib Dip</b>	<b>2022-3-60-019</b>
<b>Shahedur Rahman Rafin</b>	<b>2022-3-60-295</b>
<b>Md. Taki Nuaeser Tasin</b>	<b>2022-3-60-258</b>
<b>Rafi Al Sakif</b>	<b>2021-3-60-292</b>

**Course code: CSE475**

**Section: 02**

Department of CSE, East West University

**Submitted to:**

**Dr. Raihan Ul Islam (DRUI)**

**Associate Professor**

**Dept. of CSE.**

## Introduction

Image classification is a fundamental task in computer vision, widely used in applications such as medical analysis, surveillance, and automated systems. Traditional supervised learning methods require large amounts of labeled data, which is often expensive and difficult to obtain. This limitation motivates the use of self-supervised learning (SSL), where models learn useful representations from unlabeled data and later adapt them to downstream tasks with minimal labeled supervision.

In this work, we investigate the effectiveness of self-supervised learning for binary image classification. Three SSL methods SimCLR, Barlow Twins, and Masked Autoencoders (MAE) are evaluated using GoogleNet and ResNet-50 backbones. The learned representations are assessed through multiple downstream evaluation strategies, including linear probing, shallow classifiers, k-nearest neighbors (kNN), label efficiency experiments, and embedding quality analysis.

All experiments are conducted under controlled and reproducible settings. The results show that SSL methods significantly improve performance in low-label regimes, with MAE demonstrating the most consistent and robust performance across evaluation metrics. These findings highlight the potential of self-supervised learning as an effective alternative to fully supervised approaches when labeled data is limited.

## Exploratory Data Analysis Summary

### Dataset Overview:

The dataset contains a total of 1,287 images organized into two classes: infacted and not\_infacted.

1. infacted: 741 images
2. not\_infacted: 546 images

The dataset shows moderate class imbalance.

### Image Resolution and Preprocessing:

Images have varying resolutions, with a suggested maximum resize target of  $3677 \times 3292$ .

### Recommended preprocessing:

Resize to  $256 \times 256$

- Maintain aspect ratio with padding
- Apply random crop and resize augmentation

### **Noise and Image Quality Analysis:**

Laplacian variance analysis shows:

1. infected: 22.10
2. not infected: 30.36

This indicates higher sharpness/noise in not\_infected images.

### **Duplicate and Data Integrity Check:**

Image hashing found no significant duplicate images.

All 1,287 images were successfully processed.

### **Train–Test Split:**

Training samples: 304

Testing samples: 983

A fixed random seed ensured reproducibility.

### **Key Observations:**

1. Moderate class imbalance exists
2. Strong augmentation is recommended
3. Dataset characteristics justify self-supervised learning

## **Model Performance Comparison (80:20 Split)**

The following table summarizes the performance of all evaluated models using an 80:20 train–test split.

Model	Accuracy	Precision	Recall	F1-score	GFLOPs
EfficientNT	0.58	0.58	1.00	0.73	0.8031
ResNet50	0.58	0.58	1.00	0.73	7.7553
ResNet101	0.58	0.59	0.58	0.44	15.897
GoogleNet	99.22	99.10	99.33	99.21	1.50
MobileNetV2	0.93	0.93	0.93	0.93	0.6153

# SSL-1 Results: SimCLR with GoogleNet Backbone

---

**Table 1: Feature Extraction Summary**

Split	Samples	Feature Dim
Train	902	1024
Test	257	1024

Analysis: SimCLR with GoogleNet produces consistent 1024-dimensional embeddings, suitable for downstream tasks.

**Table 2: Shallow Classifier Performance**

Classifier	Accuracy
MLP	0.6770
SVM (Linear)	0.5992
Decision Tree	0.6420
Random Forest	0.7354

Analysis: Random Forest performs best, indicating non-linear separability in the learned embeddings.

**Table 3: Linear Probe Metrics**

Metric	Value
Accuracy	0.7354
Precision	0.4956
Recall	0.5714
F1-score	0.5308
ROC-AUC	0.6432

Analysis: Linear probe confirms strong representation quality but highlights class imbalance.

**Table 4: kNN Evaluation**

<b>k</b>	<b>Accuracy</b>
<b>1</b>	0.6031
<b>5</b>	0.7004
<b>20</b>	0.7354

Analysis: Higher k improves robustness, peaking at k=20.

**Table 5: Label Efficiency**

<b>Label %</b>	<b>Accuracy</b>
<b>1%</b>	0.5409
<b>5%</b>	0.7121
<b>10%</b>	0.6459
<b>25%</b>	0.6809
<b>50%</b>	0.6732

Analysis: High accuracy with few labels validates SimCLR's label efficiency.

**Table 6: Embedding Quality**

<b>Metric</b>	<b>Value</b>
<b>Silhouette Score</b>	0.0497

Analysis: Low silhouette score indicates partial overlap, common in challenging datasets.

# SSL-2 Results: Barlow Twins with GoogleNet Backbone

---

**Table 1:** Shallow Heads Evaluation

Classifier	Accuracy
MLP	0.6899
SVM (Linear)	0.7248
Decision Tree	0.686
Random Forest	0.7171

Analysis: Linear SVM performs best, indicating that Barlow Twins embeddings are largely linearly separable.

**Table 2:** kNN Accuracy

k	Accuracy
1	0.6357
5	0.686
20	0.7287

Analysis: Accuracy improves with larger k, peaking at k=20, showing robust neighborhood structure.

**Table 3:** Label Efficiency

Labeled Data (%)	Accuracy
1%	0.6589
5%	0.5853
10%	0.655
25%	0.6899
50%	0.686

Analysis: Strong performance with limited labels highlights good label efficiency of Barlow Twins.

**Table 4: Embedding Quality**

Metric	Value
Silhouette Score	0.0474

Analysis: Moderate overlap between class clusters is observed, typical for challenging datasets.

**Table 5: Linear Probe Evaluation**

Metric	Value
Precision	0.5612
Recall	0.7959
F1-score	0.6582
ROC-AUC	0.777

Analysis: High recall and ROC-AUC indicate strong class separability with recall-oriented behavior.

## SSL-3 Results: MAE with ResNet-50 Backbone

---

**Table 1: Shallow Heads Evaluation**

Classifier	Accuracy
MLP	0.7121
SVM (Linear)	0.7198
Decision Tree	0.6187
Random Forest	0.751

Analysis: Random Forest achieves the best accuracy, showing MAE embeddings benefit from non-linear classifiers.

**Table 2: kNN Evaluation**

k	Accuracy
1	0.6615
5	0.7004
20	0.7549

Analysis: Accuracy improves with larger k, indicating strong global structure in the embedding space.

**Table 3: Label Efficiency**

Labeled Data (%)	Accuracy
1%	0.5175
5%	0.6226
10%	0.6965
25%	0.7004
50%	0.7082

Analysis: MAE shows strong performance even with limited labeled data, validating its label efficiency.

**Table 4: Linear Probe Evaluation**

Metric	Value
Precision	0.6389
Recall	0.7041
F1-score	0.6699
Per-class Accuracy	[0.7547, 0.7041]
ROC-AUC	0.8014

Analysis: High ROC-AUC indicates strong class separability in the MAE representation space.

**Table 5: Embedding Quality**

Metric	Value
Silhouette Score	0.0597

Analysis: Higher silhouette score suggests improved clustering compared to other SSL methods.

## Enhanced Comparison of SSL Methods

SSL Method	Backbone	Best Accuracy	Best Shallow Head	Best kNN Acc.	Linear Probe F1	ROC-AUC	Silhouette	Label Eff. (5%)	Remarks
SSL-1 (SimCLR)	GoogleNet	0.7354	Random Forest	0.7354	0.5308	0.6432	0.0497	0.7121	Non-linear features
SSL-2 (Barlow Twins)	GoogleNet	0.7248	Linear SVM	0.7287	0.6582	0.7770	0.0474	0.5853	Linear separability
SSL-3 (MAE)	ResNet-50	0.7510	Random Forest	0.7549	0.6699	0.8014	0.0597	0.6226	Best overall

MAE (ResNet-50) achieves the best overall performance, showing superior accuracy, ROC–AUC, kNN results, and embedding quality. SimCLR performs well with non-linear classifiers, while Barlow Twins produces more linearly separable features, performing best with a linear SVM. Overall, all SSL methods demonstrate good label efficiency, validating their effectiveness in low-label settings.