SELF-KNOWLEDGE DISTILLATION BASED SELF-SUPERVISED LEARNING FOR COVID-19 DETECTION FROM CHEST X-RAY IMAGES

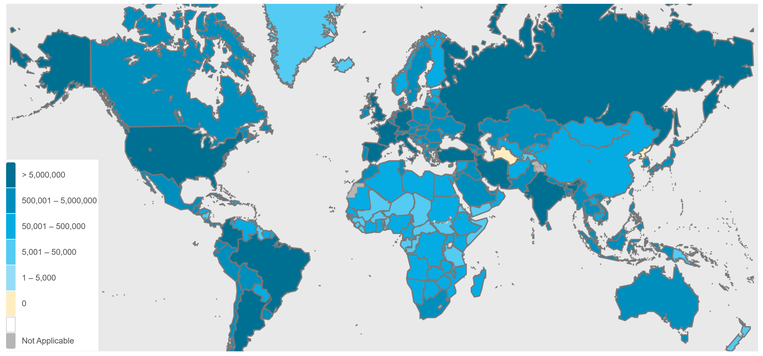
Guang Li, *Ren Togo, Takahiro Ogawa and Miki Haseyama* *Hokkaido University*

icassp 2022 Singapore

Paper ID: 1206

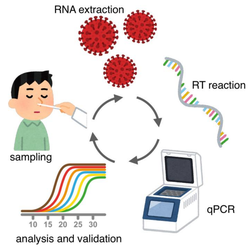
INTRODUCTION

COVID-19 Infection



COVID-19 has resulted in about 326 million people infected with over 5.5 million death worldwide as of 17 January 2022.

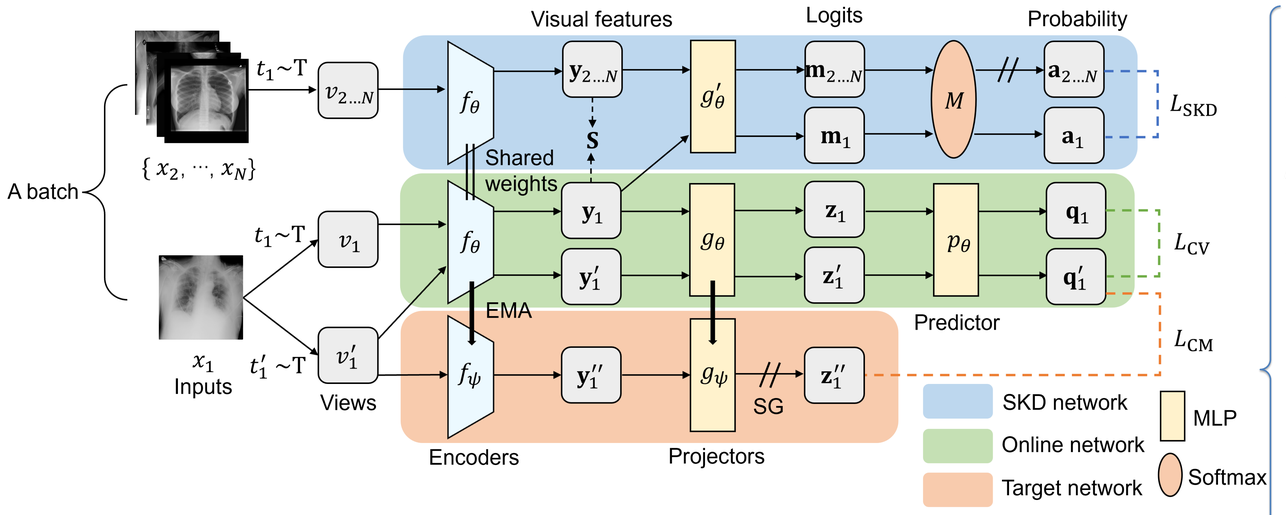
Polymerase Chain Reaction



- Although PCR is currently considered the gold standard for COVID-19 detection, it is reported with a high false-negative rate and is time-consuming.
- As many patients with confirmed COVID-19 present radiological findings of pneumonia, radiologic examinations may be useful for fast detection.

COVID-19 fast detection from chest X-ray images is needed.

PROPOSED METHOD



- The proposed method comprises three networks, where the weights of the target network are an exponential moving average (EMA) of the weights of the online network, and the encoders in self-knowledge (SKD) network and online network share the weights.
- **Novelty:** Since images with highly similar visual features tend to have similar predicted probabilities, similar images' knowledge could be ensembled to provide better soft targets for self-knowledge distillation.
- With the proposed method, the encoder of the online network f_θ can learn discriminative representations from chest X-ray images and can be used for fine-tuning and high-accuracy COVID-19 detection.

L_{SKD}: KL divergence of ensembled probabilities

L_{CV}, L_{CM}: MSE loss of normalized predictions and projections

$//$: stop-gradient

$\rfloor$: exponential moving average

M: softmax function

$\mathbf{a}$: predictive probability

$\mathbf{m}$: output logits

$\mathbf{q}$: predictions (downscaled image features)

p: predictor (multilayer perceptron)

$\mathbf{z}$: projections (downscaled image features)

g: projectors (multilayer perceptron)

$\mathbf{S}$: similarity matrix

$\mathbf{y}$: extracted image features

f: encoders (backbone)

v: augmented views

t: transformations randomly sampled from distribution T

x: input chest X-ray image

Our method can learn discriminative representations from chest X-ray images for COVID-19 fast detection.

EXPERIMENTAL RESULTS

Dataset

The largest open COVID-19 dataset

Class	Total	Train	Test
C	3,616	2,893	723
L	6,012	4,810	1,202
N	10,192	8,154	2,038
V	1,345	1,076	269

C: COVID-19
L: Lung Opacity
N: Normal
V: Viral Pneumonia

Settings

Self-supervised learning:
- Encoder: ResNet50
- MLP hidden size: 4096
- MLP projection size: 256
- View size: 128
- Moving average: 0.996

Evaluation Index

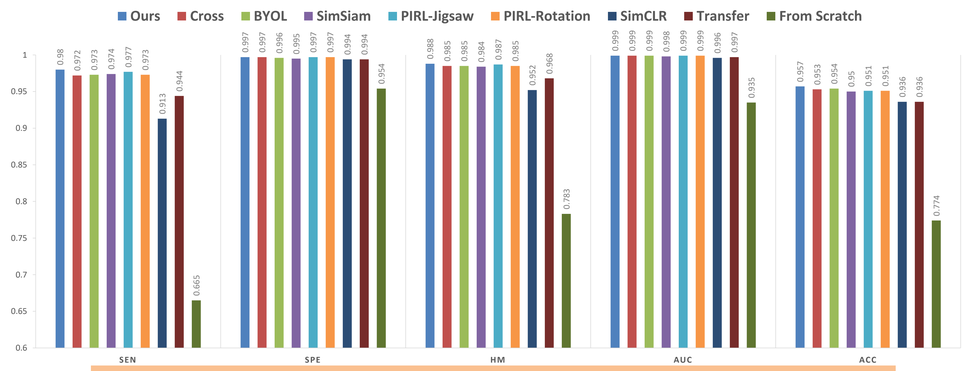
By referring to, we selected five metrics as follows:
- Sensitivity (**SEN**), Specificity (**SPE**),
- Harmonic Mean (**HM**) of SEN and SPE
- Area under the ROC curve (**AUC**), Accuracy (**Acc**)

Comparison methods

Six state-of-the-art (SOTA) self-supervised learning methods.
- Cross, BYOL, SimSiam, PIRL-Jigsaw, PIRL-Rotation, SimCLR

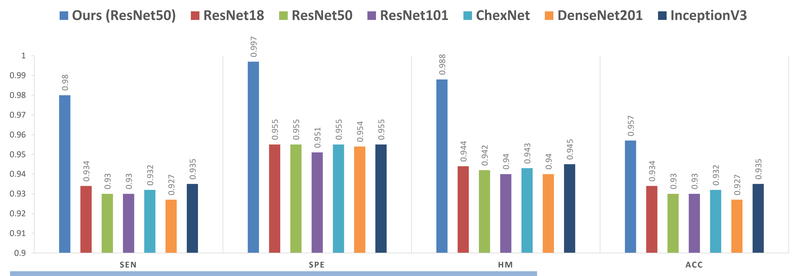
Six baseline methods.
- ResNet18, ResNet50, ResNet101, ChexNet, DenseNet201, InceptionV3

Comparison with SOTA methods

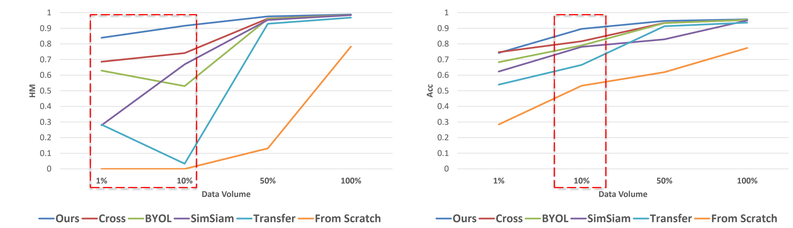


Verified that our method was effective when using full training data.

Comparison with baseline methods



Impact of the amount of training data



Verified that our method was effective even using few training data.

This study was supported in part by AMED Grant Number JP21zf0127004.
This study was conducted on the Data Science Computing System of Education and Research Center for Mathematical and Data Science, Hokkaido University.