# COVID-19 Chest X-Ray for Classification and Segmentation

Course: Deeping Learning for Medical Imaging
Date: June 5, 2024 Spring
Presenter: Wei-Lun Wu
Student ID: R12922075

# Introduction - Data

- Dataset - Chest x-ray images with 300x300 pixels
  - COVID-19: 3616 images, masks
  - Lung Opacity: 6012 images, masks
  - Normal: 10192 images, masks
  - Viral Pneumonia: 1345 images, masks
- Usage
  - Disease classification
  - Lung segmentation

# Project Objective

- Classification
  - Supervised, self-supervised, zero shot classification methods
  - Effective strategies for classifying chest X-ray images
- Segmentation
  - Supervised segmentation methods
  - Lung segmentation
- Muti-task learning
  - Modify Unet architecture for both classification and segmentation

# Methodology - Classification

- Supervised
  - Swin Transformer
  - VIT-Base
  - ConvNeXt
  - DenseNet
  - EfficientNetv2
  - ResNeXt101
- Self-supervised
  - DINOv2 (Apr 2023 Meta AI Research)
  - BEITv2 (Oct 2022 Microsoft Research)
- Zero-shot
  - CLIP: VIT Large

# Methodology - Segmentation & Muti-task

- Segmentation
  - Unet / Unet++
  - MAnet (Multi-Scale Attention Network)
  - FPN (Feature Pyramid Network)
  - PSPNet (Pyramid Scene Parsing Network)
  - PAN (Pyramid Attention Network)
  - DeepLabV3 / DeepLabV3+
- Muti-task
  - Vnet
    - Encoder output features for additional classification tasks.

# Experiment & Analysis - Implementation Details

- Split ratio is 80%, 10%, 10% for the training, validation, testing set
- Epoch: 20
- Optimizer: Adam
- Learning rate: 5e-5
- Weight decay: 1e-6
- Weights & Biases to track f1, precision, recall, dice, loss

# Experiment & Analysis - Classification

- Dataset simplicity
- Self-supervised learning efficiency
  - Fine-tune last few layers and the classification layer
- Zero-shot limitations
  - Prompt is important
- Muti-task slightly decreases

| | | F1 Score | Precision | Recall | Accuracy | Minute / Epoch |
|---|---|---|---|---|---|---|
| Supervised | | | | | | |
| | Swin Transformer | 0.9668 | 0.9699 | 0.9639 | 0.9601 | 8.95 |
| | VIT Base | 0.9631 | 0.965 | 0.9613 | 0.9511 | 8.85 |
| | ConvNeXt | 0.9651 | 0.9663 | 0.9641 | 0.9546 | 7.55 |
| | DenseNet | 0.9513 | 0.9401 | 0.9644 | 0.9502 | 10.65 |
| | EfficientNetV2 | 0.9654 | 0.9686 | 0.9624 | 0.9561 | 8.65 |
| | ResNeXt101 | 0.9597 | 0.961 | 0.9268 | 0.9497 | 10.3 |
| Self Supervised | | | | | | |
| | DINOv2 | 0.9648 | 0.9666 | 0.9602 | 0.9575 | 2.45 |
| | BEITv2 | 0.9611 | 0.963 | 0.9593 | 0.9642 | 3.95 |
| Zero Shot | | | | | | |
| Prompt-1 | CLIP | 0.4264 | 0.5843 | 0.4422 | 0.6719 | |
| prompt-2 | CLIP | 0.1689 | 0.2842 | 0.2241 | 0.2255 | |
| Muti-Task | | | | | | |
| | VNet | 0.9515 | 0.9589 | 0.9447 | 0.9428 | 7.88 |

# Experiment & Analysis - Classification

Use 100%, 75%, 50%, 25% training set in self-supervised method

- Effective with Limited Data (50%)
  - Effective when labeled data is not enough
  - DINOv2 achieves good results

| | | F1 Score | Precision | Recall | Accuracy |
|---|---|---|---|---|---|
| 100% | | | | | |
| | DINOv2 | 0.9648 | 0.9666 | 0.9602 | 0.9575 |
| | BEITv2 | 0.9611 | 0.963 | 0.9593 | 0.9642 |
| 75% | | | | | |
| | DINOv2 | 0.9644 | 0.9662 | 0.9626 | 0.9563 |
| | BEITv2 | 0.9591 | 0.9641 | 0.9548 | 0.9523 |
| 50% | | | | | |
| | DINOv2 | 0.9614 | 0.9683 | 0.9555 | 0.9544 |
| | BEITv2 | 0.9559 | 0.9622 | 0.9508 | 0.9494 |
| 25% | | | | | |
| | DINOv2 | 0.9467 | 0.9435 | 0.9502 | 0.9405 |
| | BEITv2 | 0.9425 | 0.9484 | 0.9368 | 0.9362 |

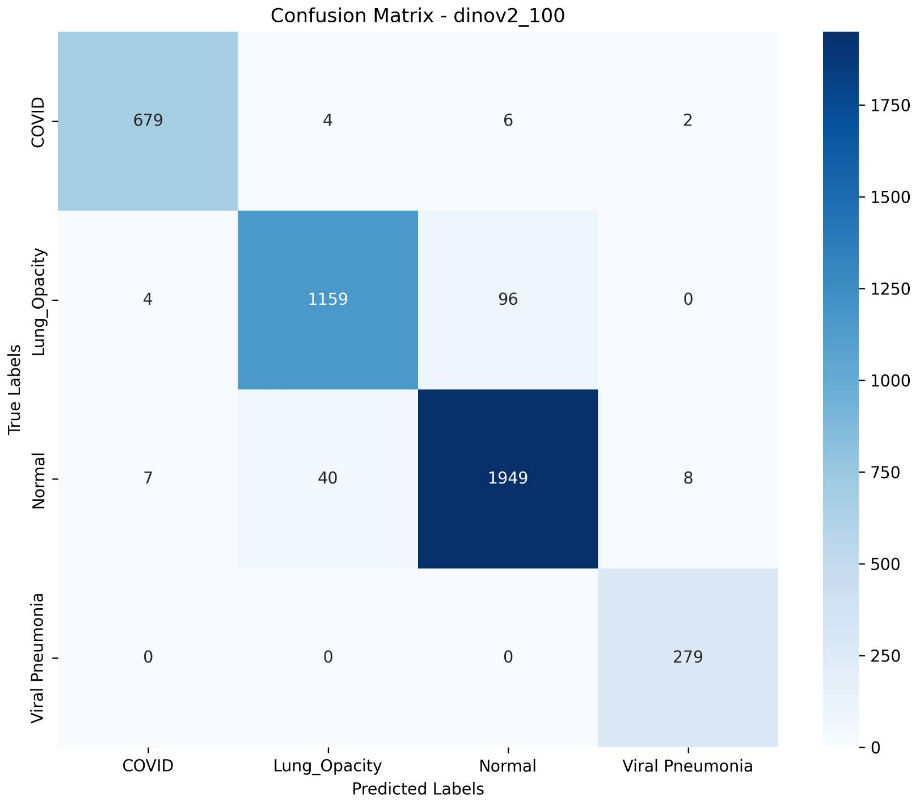
# Experiment & Analysis - Segmentation

- DeepLabV3+ has best performance
- Muti-task is not bad

| | Val Dice | Val loss |
|---|---|---|
| Unet | 0.983 | 0.03628 |
| UNet++ | 0.9831 | 0.03711 |
| MAnet | 0.9825 | 0.03786 |
| FPN | 0.9825 | 0.03781 |
| PSPNet | 0.9805 | 0.03394 |
| PAN | 0.9822 | 0.03823 |
| DeepLabV3 | 0.9818 | 0.03812 |
| DeepLabV3+ | 0.9833 | 0.03451 |
| Muti-Task: VNet | 0.9788 | |

# Visualization

- Confusions
  - Lung opacity and normal
- Well-distinguished
  - COVID and viral pneumonia



# Conclusion

- Model approaches
  - Supervised, self-supervised, zero-shot classification, segmentation, and multi-task learning
- Self-supervised efficiency
  - Fine-tune few layers
- Data size
  - Even with reduced data (50% of the training set), DINOv2 provided good results
- Zero-shot learning challenges
  - CLIP sensitivity to prompts
  - Limitations in learning specific disease features
- Multi-task learning
  - Modify Unet architecture for simultaneous classification and segmentation