Deep Learning Based Facial Mask Detection

CS5242 Final Project
Presented by: Group 10 (Huang Ziyu, Li Zhaofeng, Wang Yuchen)
Date: 14 Nov 2022

© Copyright National University of Singapore. All Rights Reserved.

Content

- **Problem Statement**
- **Data Preparation**
  - data collection
  - data pre-processing
  - data analysis
- **Model Architecture**
  - baseline models: MLP, CNN, LSTM
  - proposed model: vision transformer
- **Discussion and Analysis**
  - result visualization
  - project novelty
  - future research directions

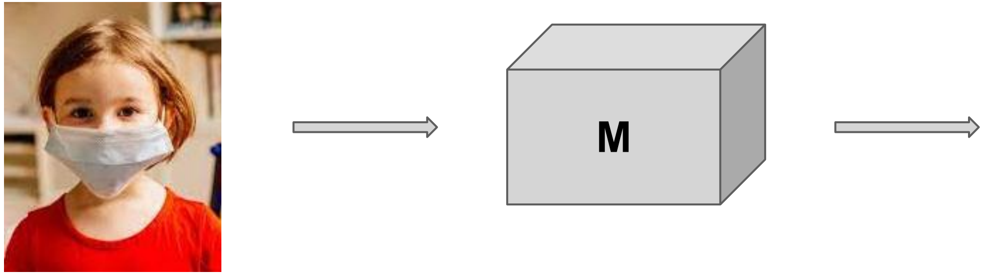
© Copyright National University of Singapore. All Rights Reserved.

Problem Statement

- **Objective**: accurate facial mask detection under covid-19
- **Methodology**: DL-based multi-class image classification

$$Acc = \mathop{\mathbb{E}}_{\langle i,c\rangle \in D} \mathbb{1}_{argmax(p(i))=c}$$

i



mask?

no mask?

no face?

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation

- data collection
- data pre-processing
    - cleansing
    - augmentation
- data analysis

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Collection

- Google/Baidu
- Sleep time/IP address
- With mask/Without mask/Irrelevant

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Pre-processing

- Data pre-processing **increase sample size/variety** and **reduce data noise**

data cleansing

- remove deprecated images
- re-assign wrongly-classified images

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Pre-processing

- Data pre-processing **increase sample size/variety** and **reduce data noise**

data cleansing	- remove deprecated images - re-assign wrongly-classified images
face centered - image cropping	- crop images to focus on the face part with **`face_recognition`** library

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Pre-processing

- Data pre-processing **increase sample size/variety** and **reduce data noise**

data cleansing	- remove deprecated images - re-assign wrongly-classified images
face centered - image cropping	- crop images to focus on the face part with `face_recognition` library
data augmentation	- resizing - random cropping - affine transformation - rotation (45/90/180/270) - flipping - lightening - Gaussian blurring - salt and pepper noise

© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Analysis

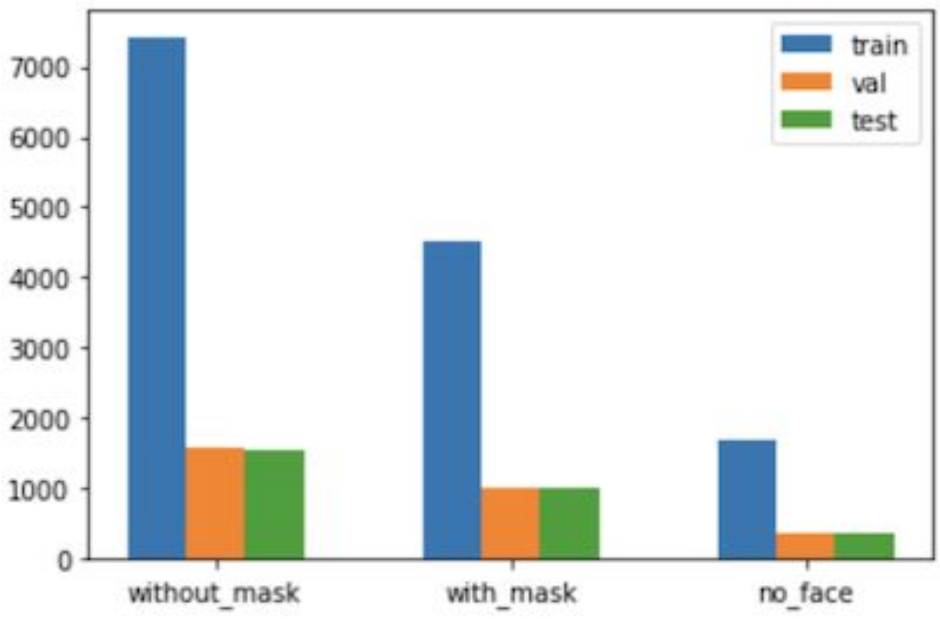
- **<u>Data splitting</u>**
    - Train set : val set: test set: **8 : 1 : 1**

```
train:
without masks:  7427
with masks:  4527
no faces:  1672
val:
without masks:  1563
with masks:  986
no faces:  370
test:
without masks:  1554
with masks:  1015
no faces:  350
```


© Copyright National University of Singapore. All Rights Reserved.

Data Preparation - Analysis

- **Data splitting**
- **<u>Data visualization</u>**
  - bar plot to show the number of images from three categories.



© Copyright National University of Singapore. All Rights Reserved.

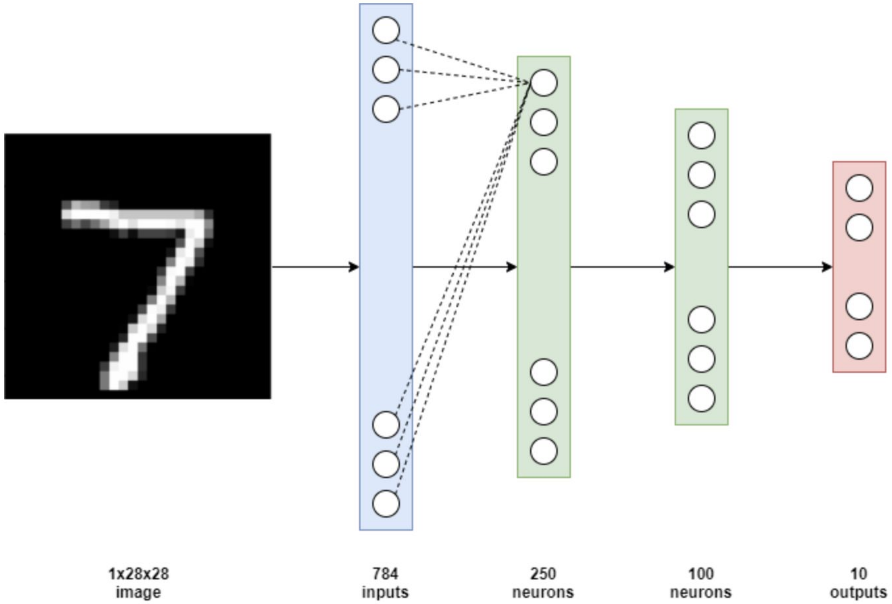
Model Architecture

- MLP
- CNN
- LSTM
- Vision Transformer

© Copyright National University of Singapore. All Rights Reserved.

MLP

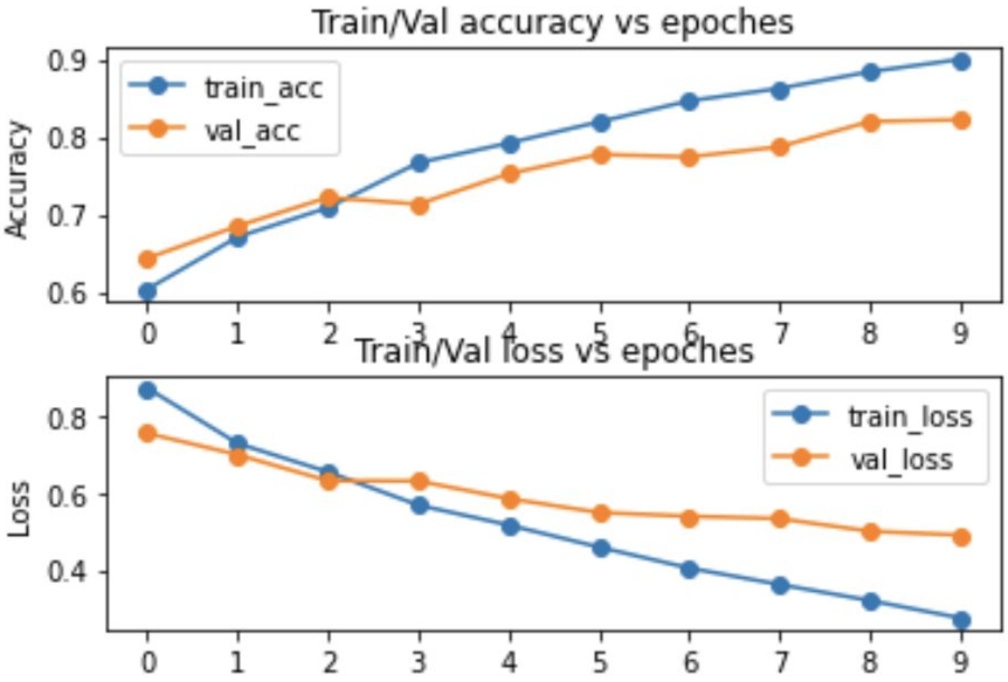
- **<u>model architecture</u>**: fully connected layers



© Copyright National University of Singapore. All Rights Reserved.

MLP

- **model architecture**
- **<u>performance</u>**: 79%



© Copyright National University of Singapore. All Rights Reserved.

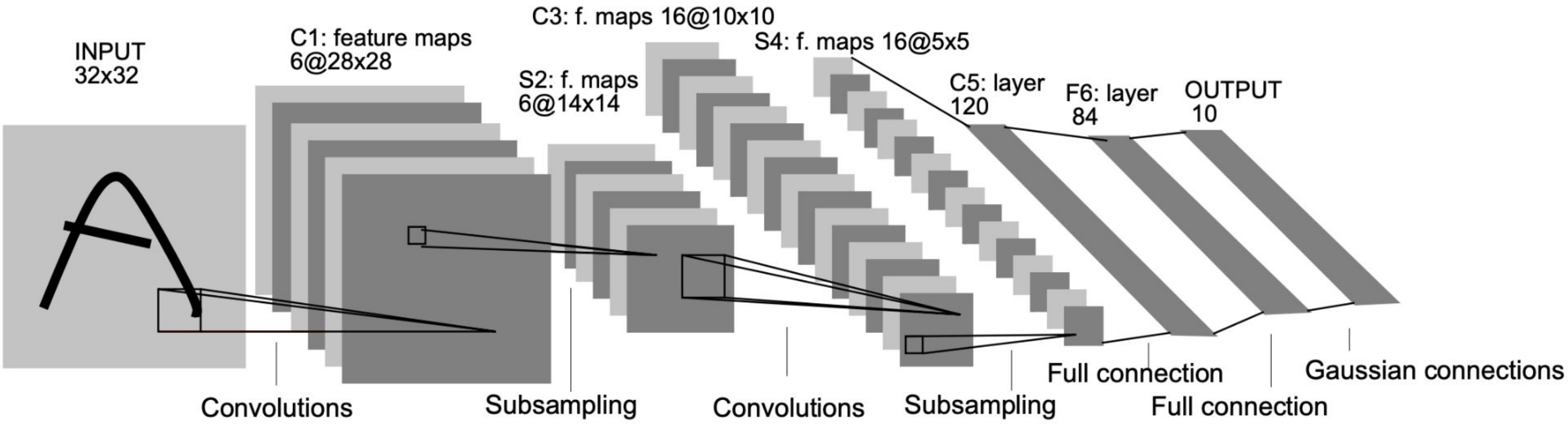
MLP

- **model architecture**
- **performance**: 79%
- **pros and cons**:
    - easy to generalize
    - quick predictions
    - large #parameters

© Copyright National University of Singapore. All Rights Reserved.

CNN

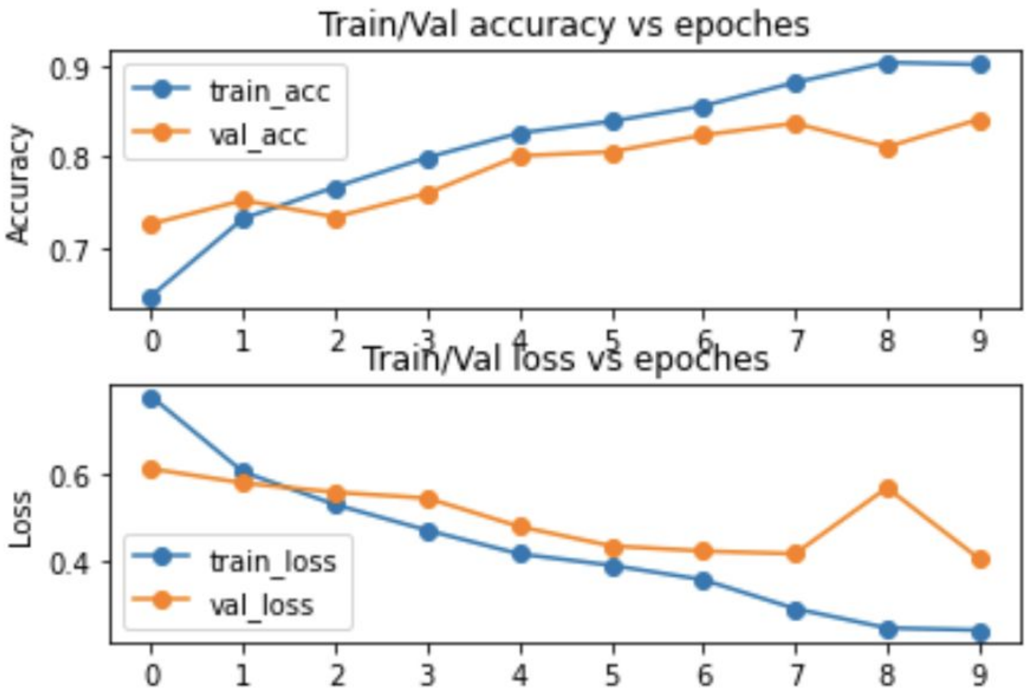
- **model architecture**: grey-scale -> conv layers -> linear layers (referring to LeNet 5)



© Copyright National University of Singapore. All Rights Reserved.

CNN

- **model architecture**
- **<u>performance</u>**: 83 %



© Copyright National University of Singapore. All Rights Reserved.

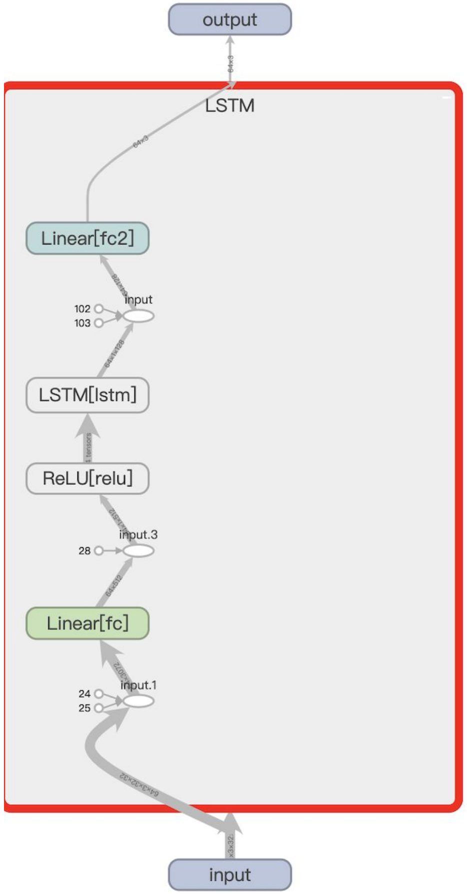
CNN

- **model architecture**
- **performance**: 83 %
- **pros and cons**
  - image feature capturing
  - parameter sharing
  - parallel computation
  - over-emphasize individual features vs whole object

© Copyright National University of Singapore. All Rights Reserved.

LSTM

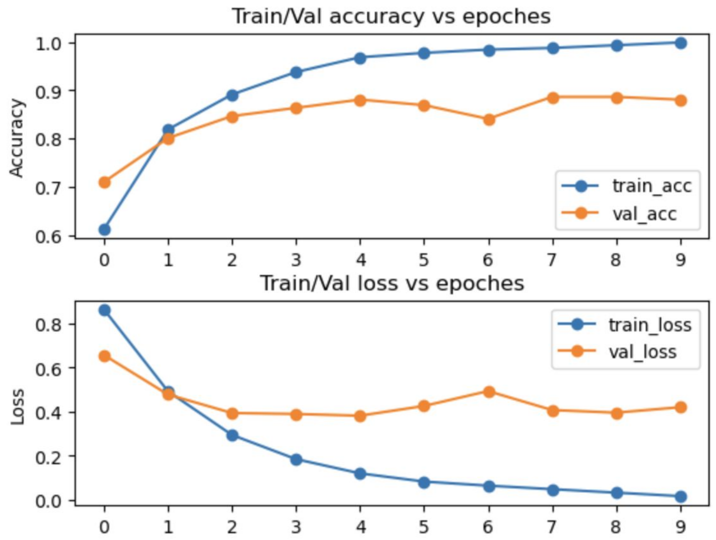
- **<u>model architecture</u>**:
  batch size = 64.
  linear layers(flatten,
  (64, 3*32*32)->(64, 512)) **->**
  relu(activation) **->**
  lstm(unsqueeze,
  (64,1,512)->(64, 1,128)) **->**
  linear layers(flatten,
  (64, 128)->(64, 3)).

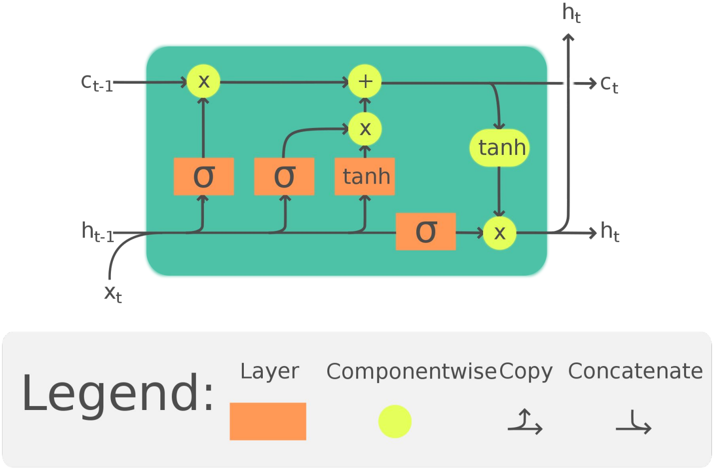


© Copyright National University of Singapore. All Rights Reserved.

LSTM

- **performance**: 81%





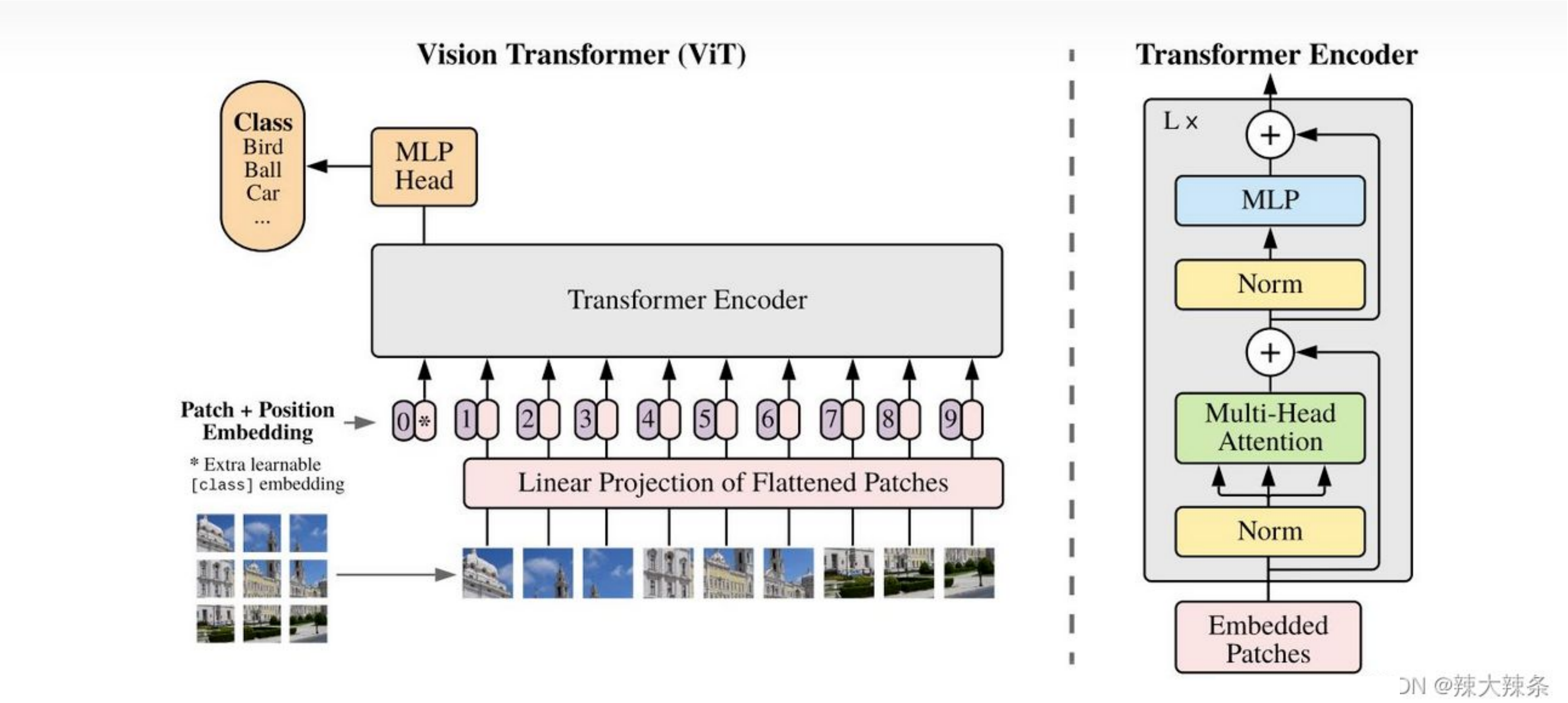
© Copyright National University of Singapore. All Rights Reserved.

LSTM

- **model architecture**
- **performance**: 81 %
- **pros**
  - less prone to vanishing gradient problem in RNN
  - avoid the long-term dependency issue
  - parameter sharing
- **cons**
  - easy to overfit
  - hard to implement Dropout
  - takes longer training time

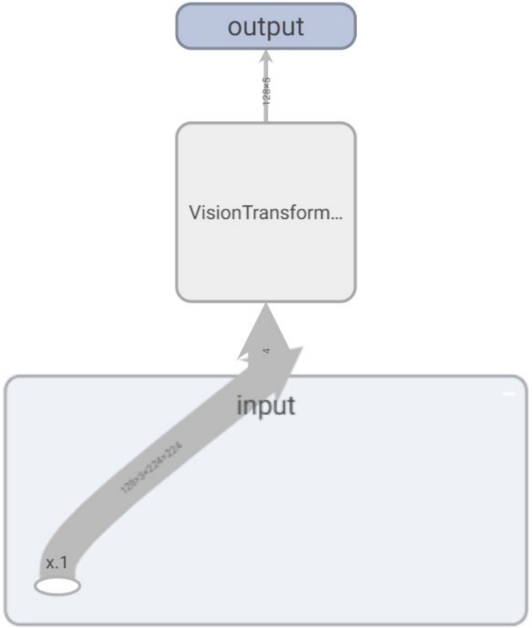
© Copyright National University of Singapore. All Rights Reserved.

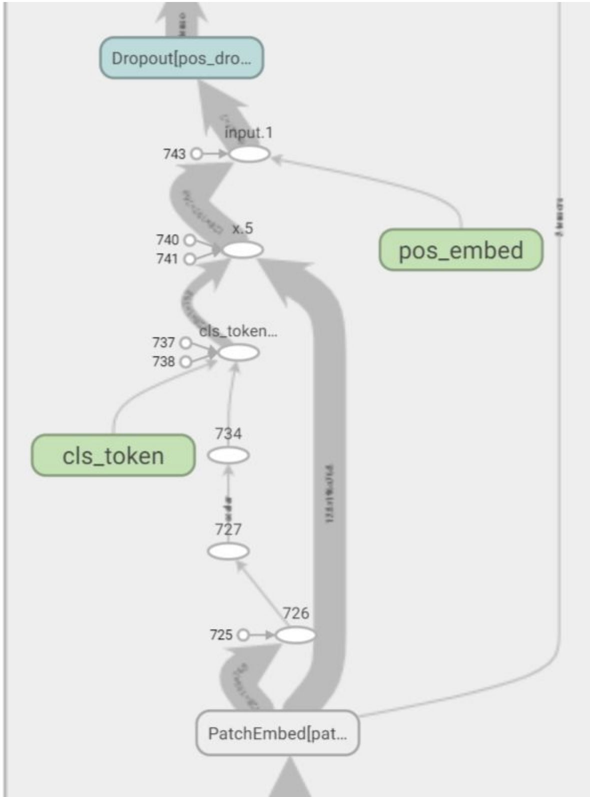
Vision Transformer

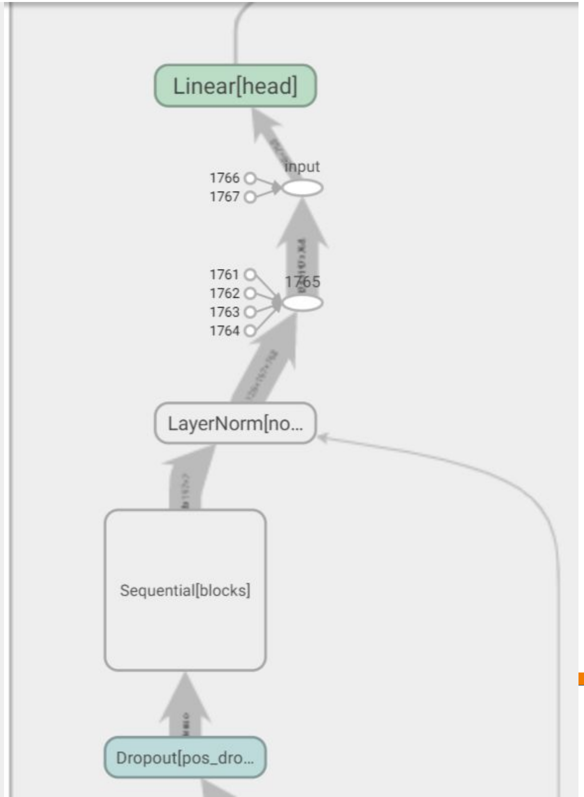


© Copyright National University of Singapore. All Rights Reserved.

Vision Transformer







© Copyright National University of Singapore. All Rights Re

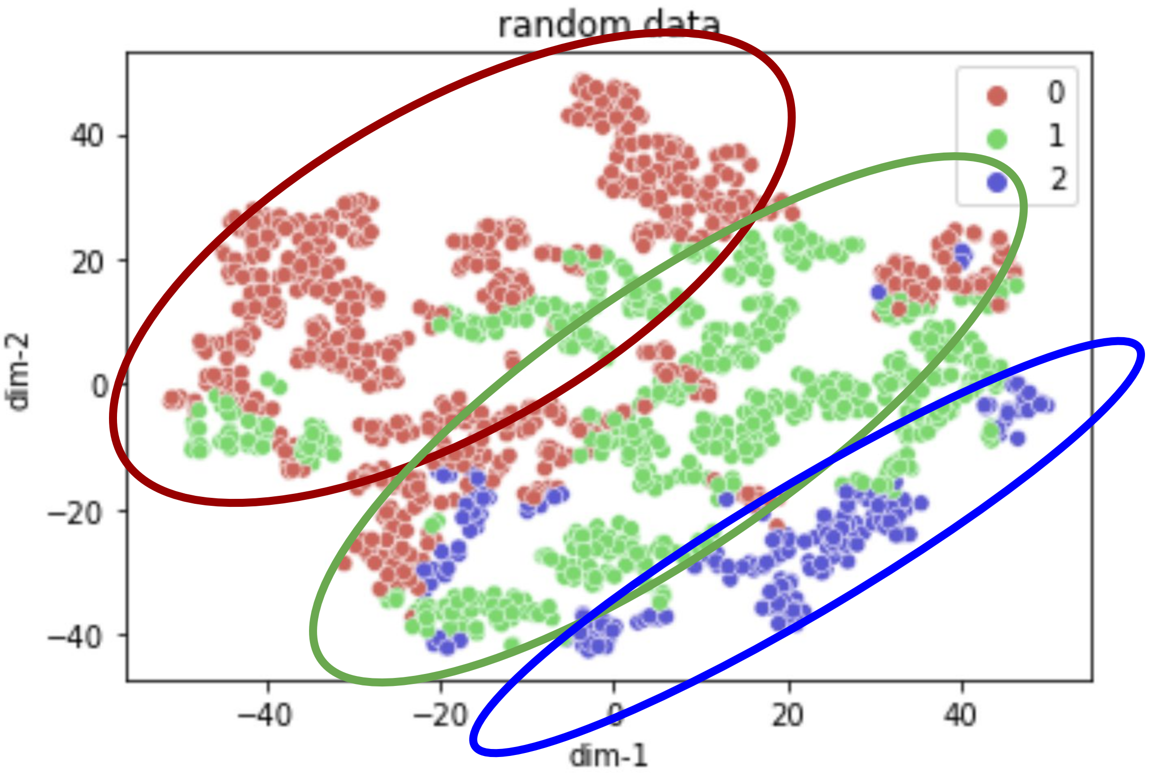
Discussion & Analysis

- results visualization
- project novelty
- future research directions

© Copyright National University of Singapore. All Rights Reserved.

Result Visualization

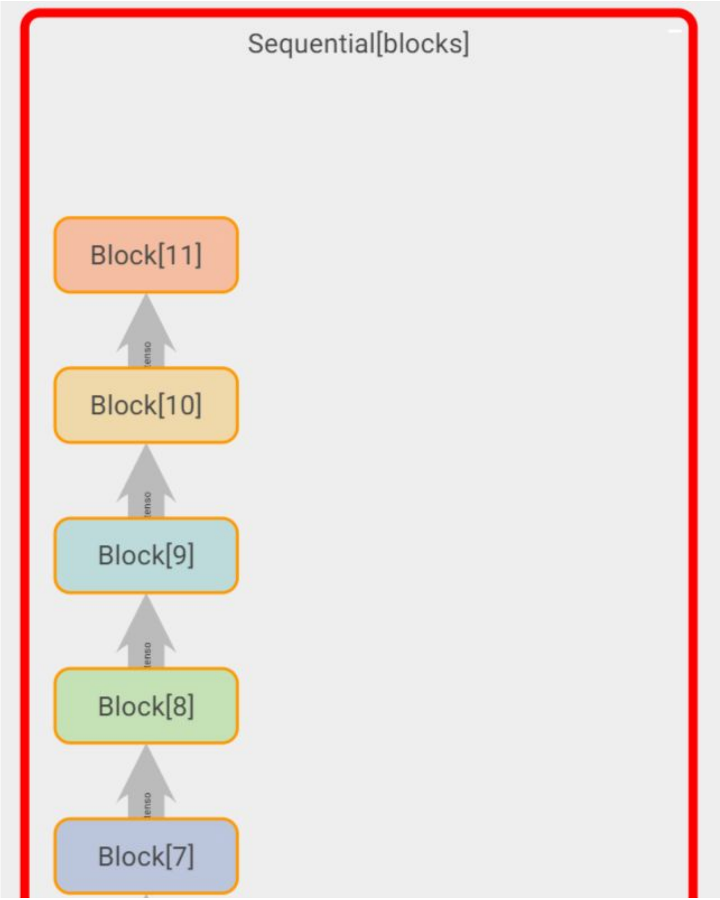
- tSNE visualization of 1600 samples on 2D space



© Copyright National University of Singapore. All Rights Reserved.

Project Novelty

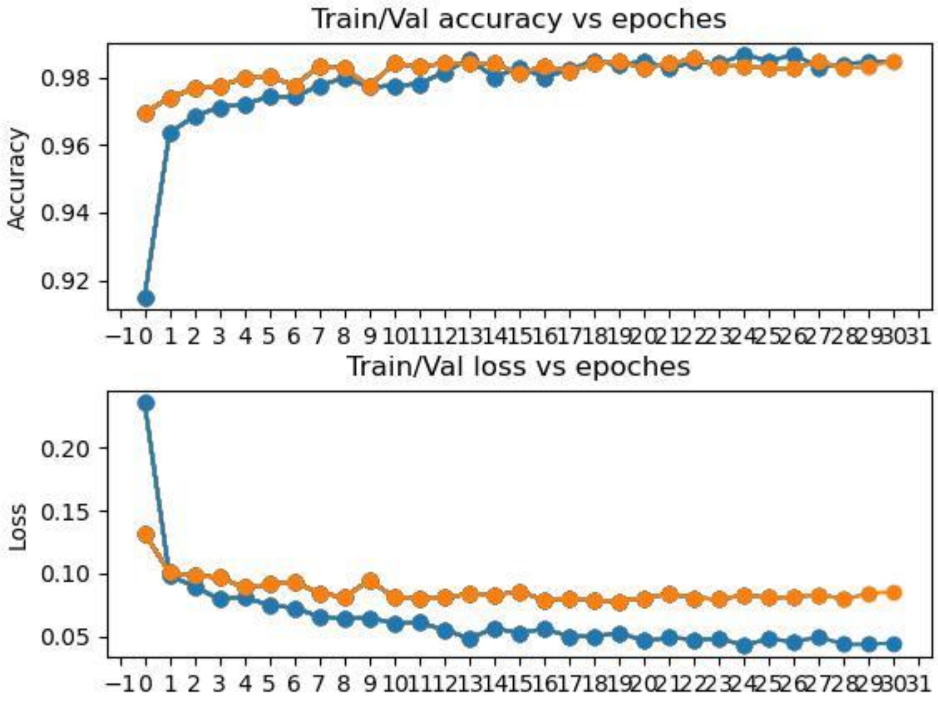
- **Proposed a suitable model architecture**



© Copyright National University of Singapore. All Rights Reserved.

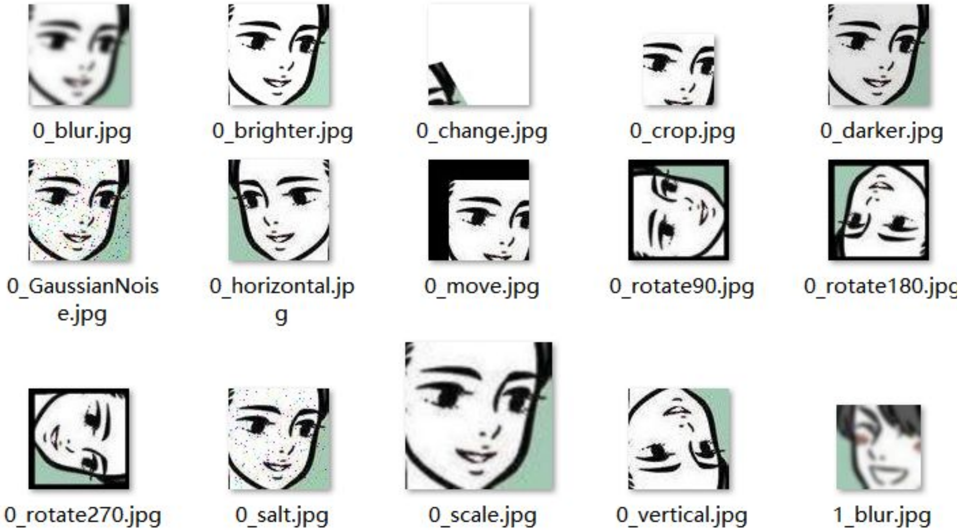
Project Novelty

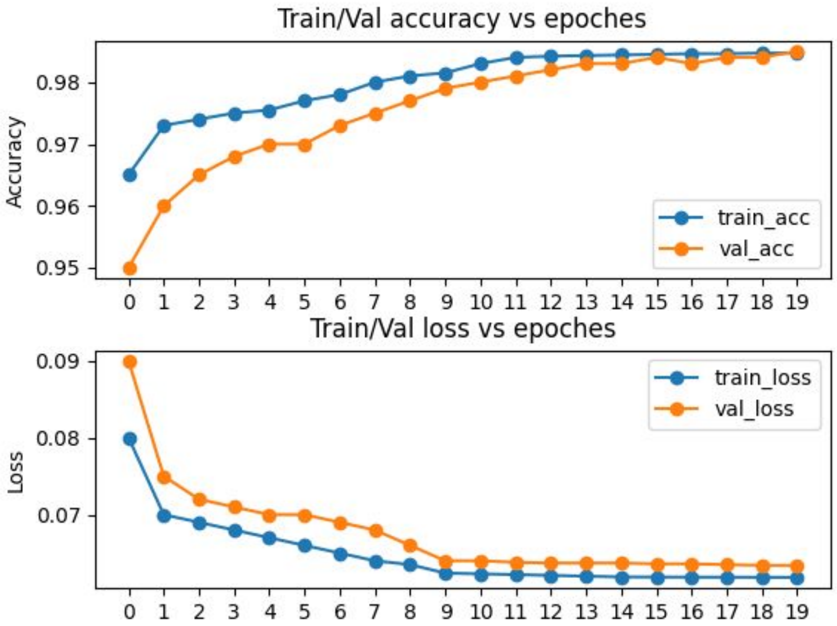
- **Created a new dataset from scratch**



Project Novelty

- **Implemented 14 data augmentation methods**



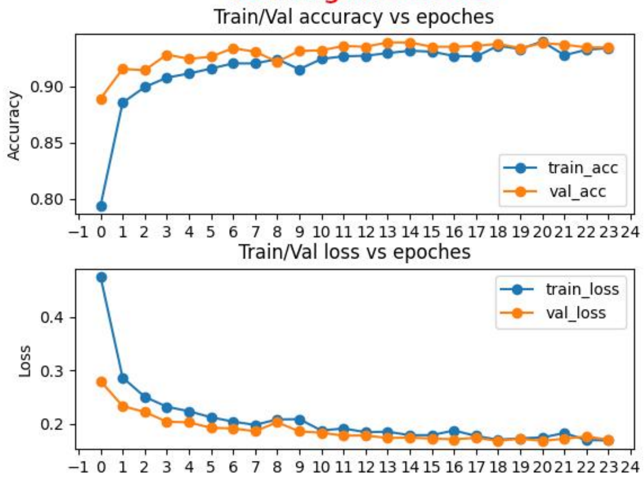


© Copyright National University of Singapore. All Rights Reserved.

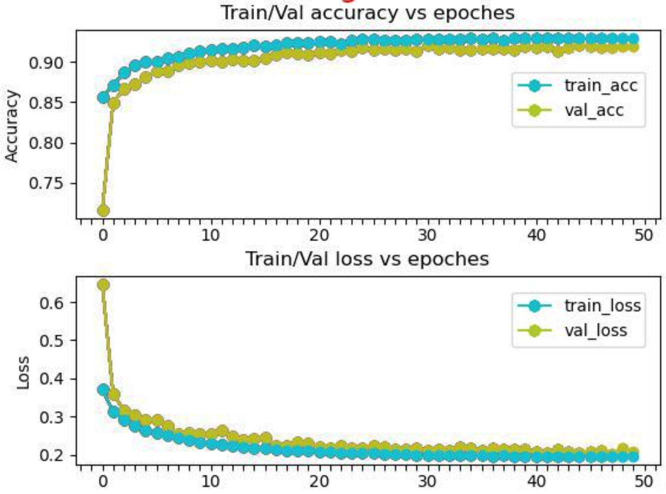
Project Novelty

- **Used Smart training strategies**

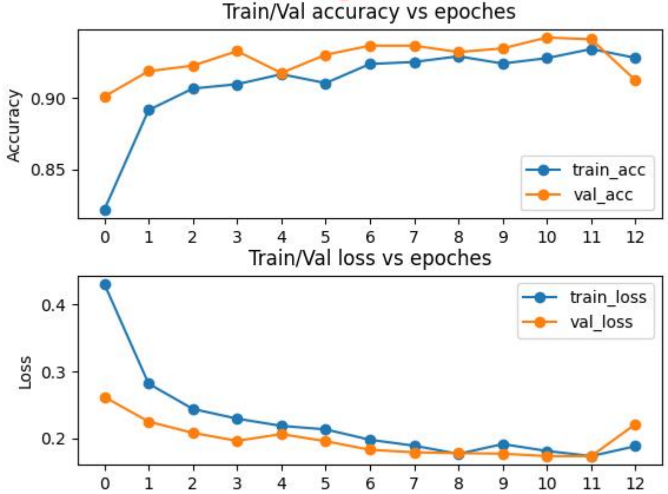
Learning Rate 0.0005



Learning Rate 0.0001



Learning Rate 0.001



Project Novelty

- Proposed a suitable model architecture
- Created a new dataset from scratch
- Implemented 14 data augmentation methods
- Used Smart training strategies
- **Proposed a model with highly generalizable**

© Copyright National University of Singapore. All Rights Reserved.

Limitations

- Size of dataset
- Training with more GPU
- Ensembling and optimization

© Copyright National University of Singapore. All Rights Reserved.

Future Research Directions

- Model pruning
- Ensemble learning: hard voting and soft voting
- Adversarial training

© Copyright National University of Singapore. All Rights Reserved.

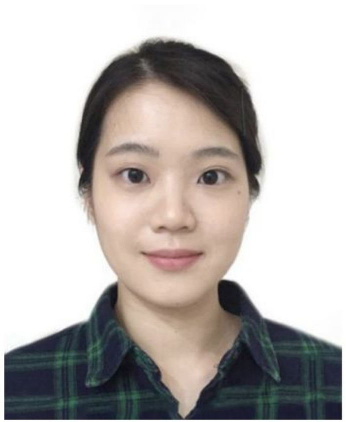
Acknowledgment

- We’d like to thank all CS5242 teaching staffs for their support and guidance through out the learning journey.

Dr Ai Xin

Kin Whye Chew

kinwhye@nus.edu.sg



Lin Qiuxia

qiuxia.lin@u.nus.edu

Zhao Jingwei

jzhao@u.nus.edu

© Copyright National University of Singapore. All Rights Reserved.

Reference

We have attached our reference sources to corresponding notebooks.

© Copyright National University of Singapore. All Rights Reserved.