The Project of

Mahjong Master

Cunhan Zou (A0248334A)
Kaisheng Deng (A0242101E)
Ziwei Peng (A0248341H)

Contents

Background & Motivation

Mahjong: A tile-based game that was developed in the 19th century in China.

We want to implement automated decision-making for mahjong players. The first problem is the classification of mahjong tiles and we focus on classification for this project.

(34 classes) mahjong tiles classification problem.

Data Collection & Augmentation

Data are jointly collected by all three members of the team.

To augment dataset, we apply the following techniques:

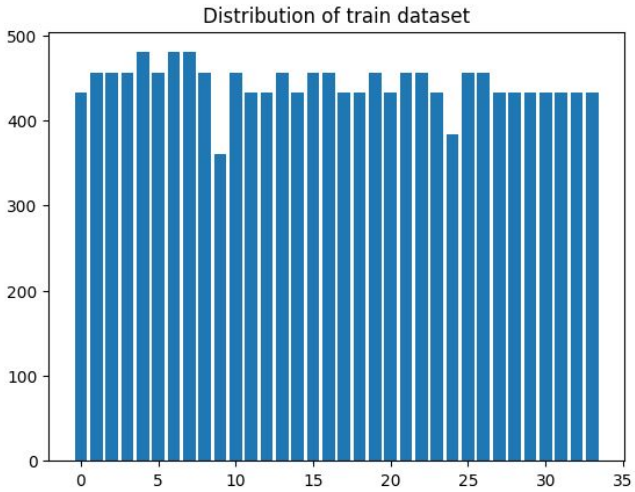
- Change in orientations
- Change in scales
- Random occlusions

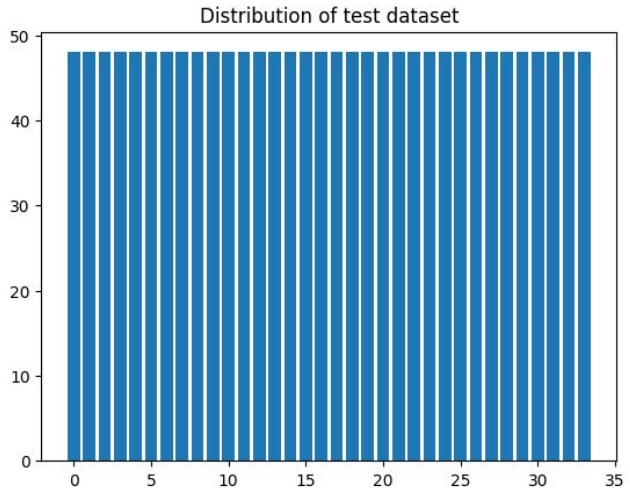
Data split:

Around 9(trainset) vs 1(validation/testset)

Data Exploration

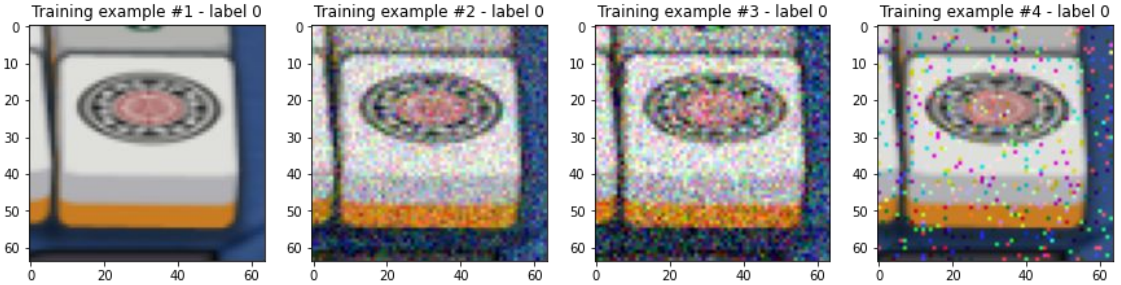
The dataset is well balanced.



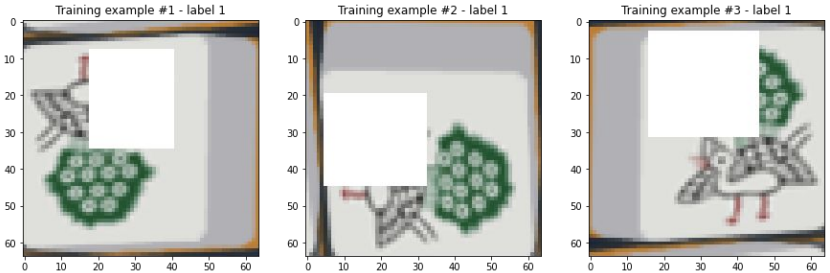


Data Augmentation

Random Noise (Sample)



Random Occlusion (Sample)



Data Preprocessing

Image Resize => from arbitrary size to fixed size (64×64×3)

Example:

54×71×3

64×64×3

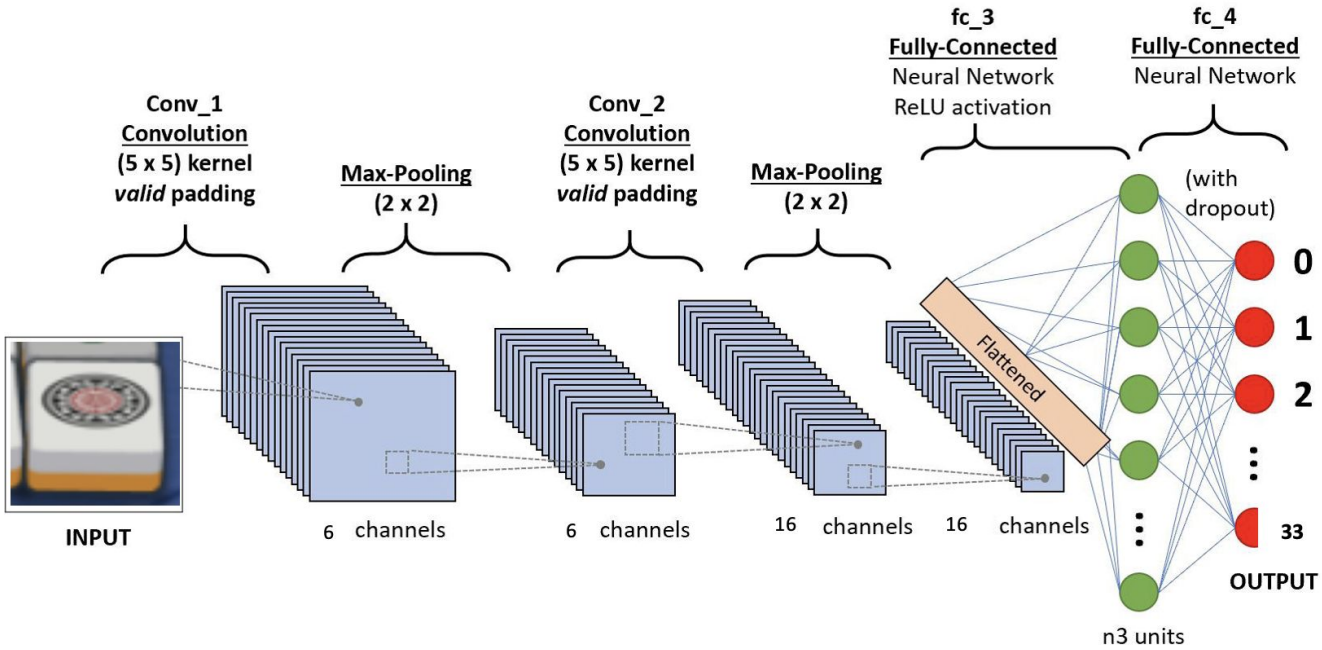
Image Normalization => provided by `torchvision.transforms.ToTensor`

Converts pixel range [0, 255] to FloatTensor type range [0.0, 1.0]

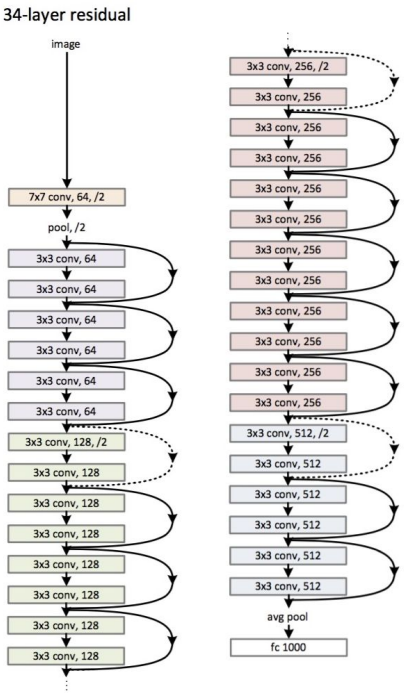
MLP Model - Overview

- Our baseline model
- Introduction to many modern image classification methods
- Trained with backpropagation
- Consisted of input layer, hidden layer and output layer
- Frequently used for image classification before 2000

MLP Model - Structure

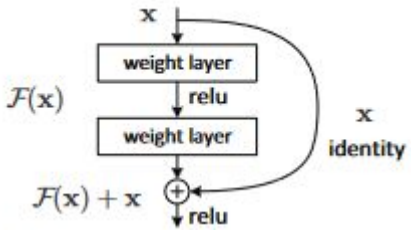


ResNet - Overview

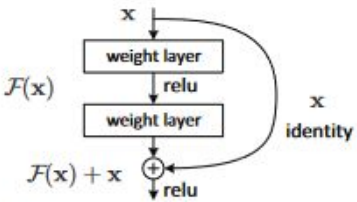


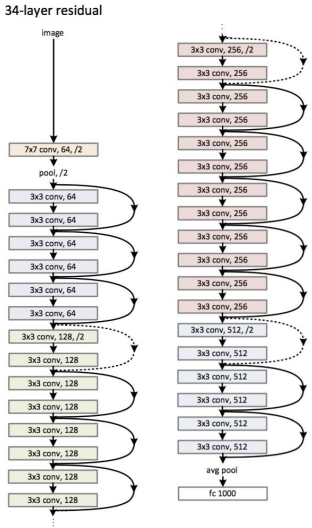
- Use ResNet34 as an advanced model to solve this problem
- Basically 34 layers, consists of convolutional layer, activation layer, batch normalization layer and fully connected layer
- Basic units: residual block

ResNet - Residual Block



- When network gets deeper, it starts to lose capability of generalization
- Intuition: use ‘skip connection’ to solve this problem
- These skip connections help as they allow an alternate shortcut path for gradients to flow through.
- The following layers can get information from previous layers easily, and hopefully, learn something more.

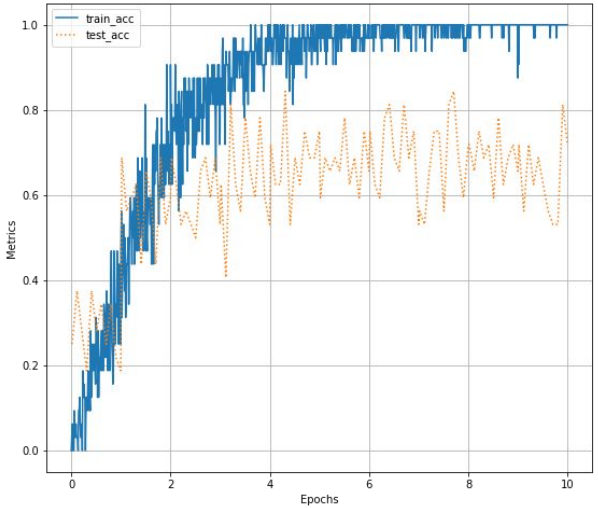


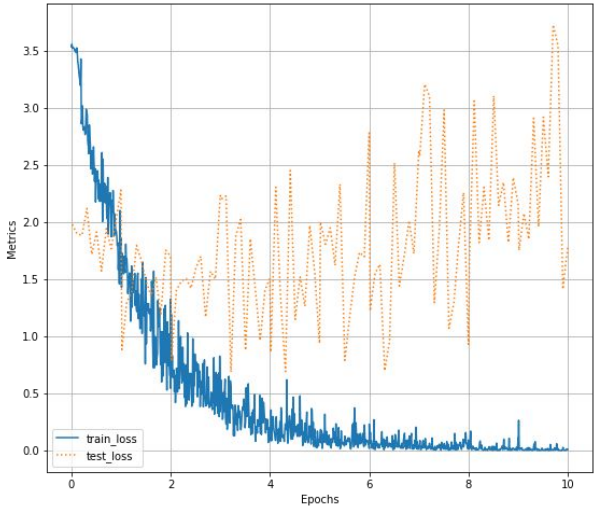


ResNet - Structure

- Intuition: stack the residual blocks together
- In detail:
- Section 1: 3 residual blocks, channels 3 -> 64
- Section 2: 4 residual blocks, channels 64 -> 128
- Section 3: 6 residual blocks, channels 128 -> 256
- Section 4: 3 residual blocks channels 256 -> 512
- Followed by average pooling layer and fully connected layer to convert feature map into prediction vector
- Optimization: change strides of section 3 and section 4 from 2 to 1, since our input shape is $3 \times 64 \times 64$

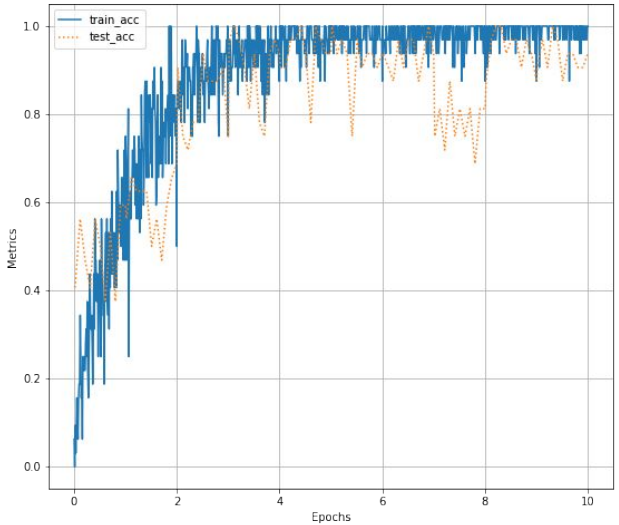
Result - Multilayer Perceptron

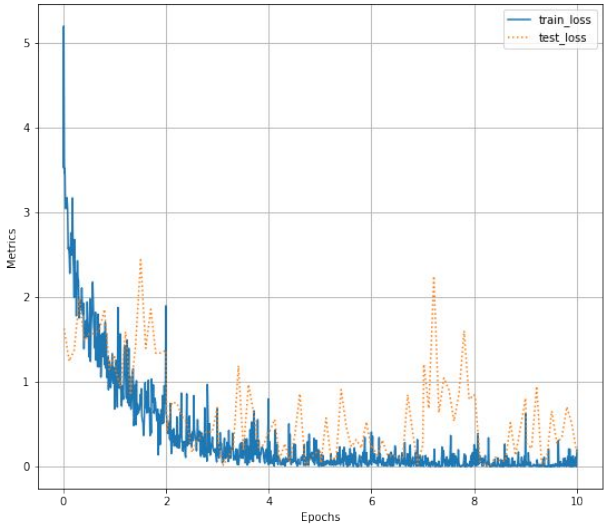




- After 10 rounds of training, the MLP got training accuracy of 100% and testing accuracy of 71.8%

Result - ResNet34



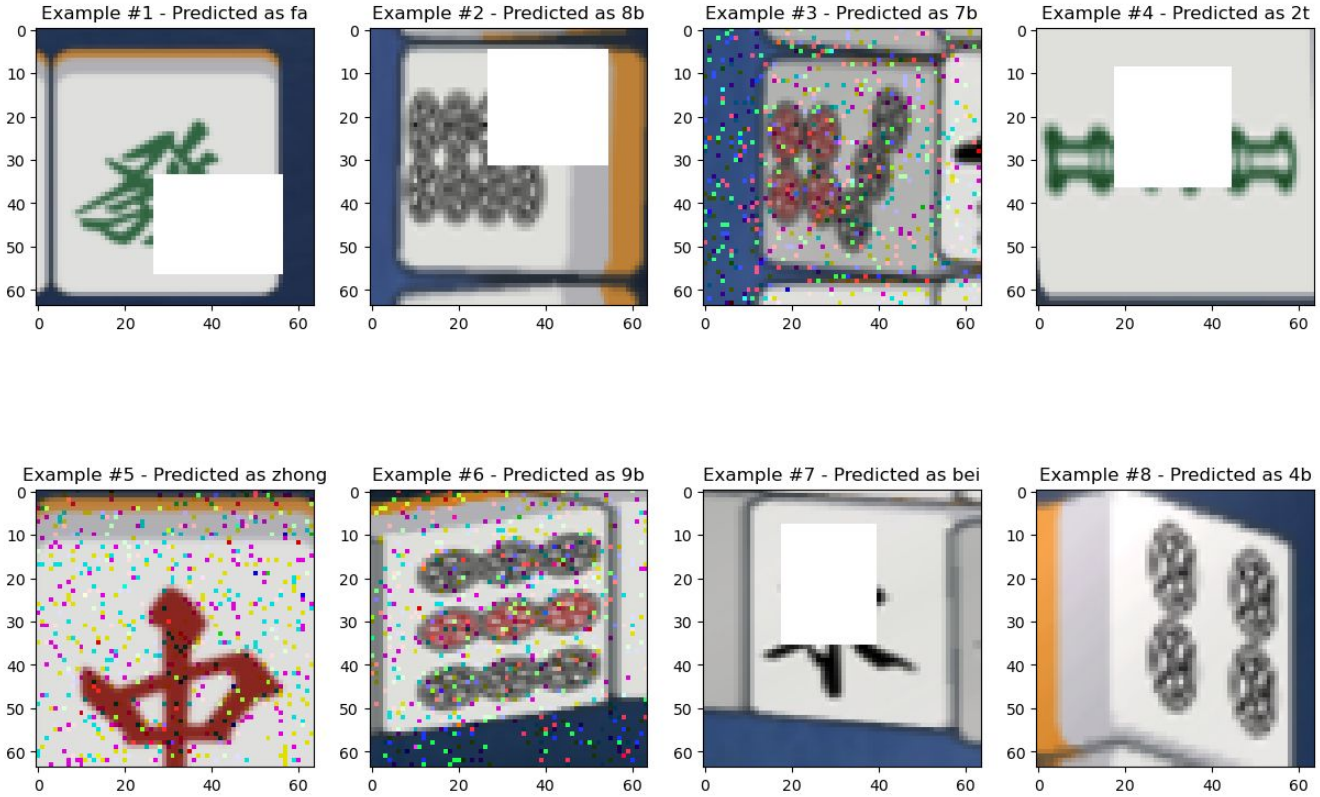


- After 10 rounds of training, the ResNet34 model get training accuracy of 100% and testing accuracy of 93.75%

Result Comparison - MLP & ResNet34

- Compared with MLP model, ResNet34's accuracy on testing set improves from around 72% to around 93%
- Although both models' accuracy on training set can reach 100%, the MLP model's accuracy on testing set can not improve after around 4 epochs.
- RestNet34 can understand spatial relation between pixels of images better for complicated Mahjong images
- We can see that ResNet34 can learn more information and has better capability of generalization.

Predictions



Looking into Future

The mahjong project is an image classification task of categorizing and assigning labels to groups of pixels or vectors within an image dependent on particular rules.

When we are exploring different models and advanced algorithms in image classification, we are swimming in the field of modern computer vision and digging the history back.

Looking into Future

From supervised to unsupervised learnings, to the integration of machine learning techniques, image classification is a showground of tradition and artificial intelligence.

Learn hidden knowledge + unorganized and organized datasets = neural network is the best (currently tho)

Looking into Future

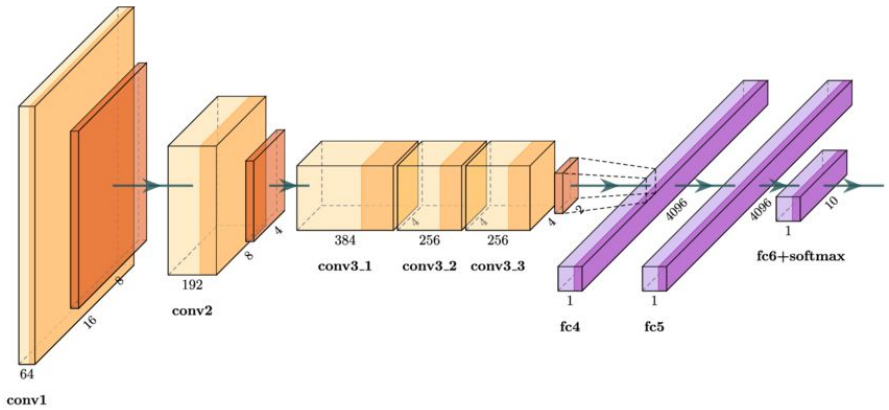
AlexNet

VGGNet

GoogleNet

DenseNet

ResNet



Looking into Future

Self–training model - the effectiveness increases as more data are fed in the form of annotated images.

2021 - application of Vision Transformers (ViT)

2022 - release of YOLOv7

The Project of

Mahjong Master

Thank you.