

Increasing training samples, data augmentation, and regularization can effectively reduce overfitting and improve model performance. But pretrained models can work better even with small samples.

Also, I think applying some techniques in Chapter 5 like random shuffling, cross-validation could be more helpful. In practice I've tried many times to embed these codes and failed, even exaggerating once when I tried to embed the random shuffling code, the accuracy reached 1.

Google colab link: