Music Classification With Spectrograms

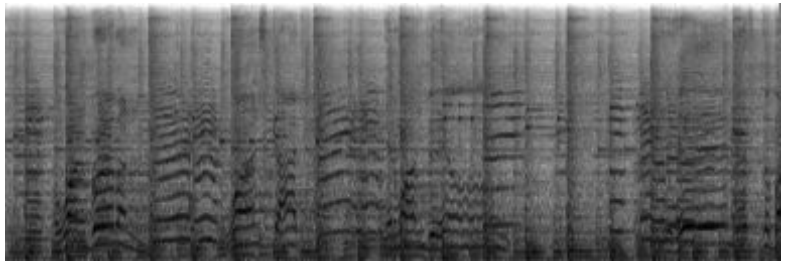
By: Fatemeh Kazemi, Ali Zarei
Quetzalcoatlus_Zouk/ Quetzalcoatlus_Zouk_B

Neuromatch Academy

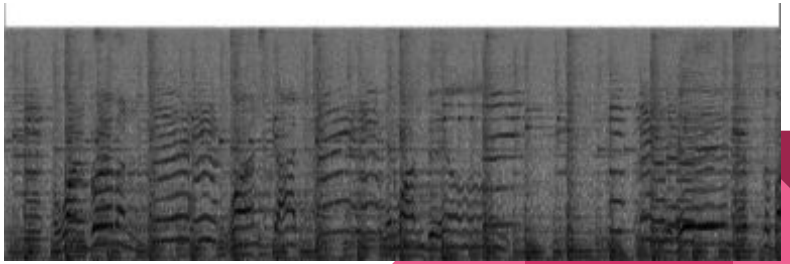
Data Preparation

- using the GTZAN dataset
- converting audio files in the dataset into spectrograms
- data augmentation by applying noise to each audio file

sample without noise



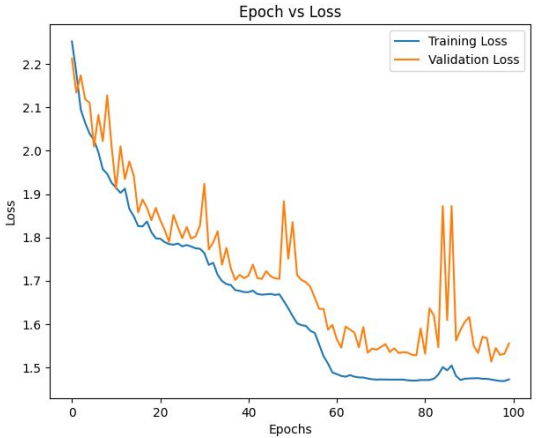
the same sample with noise

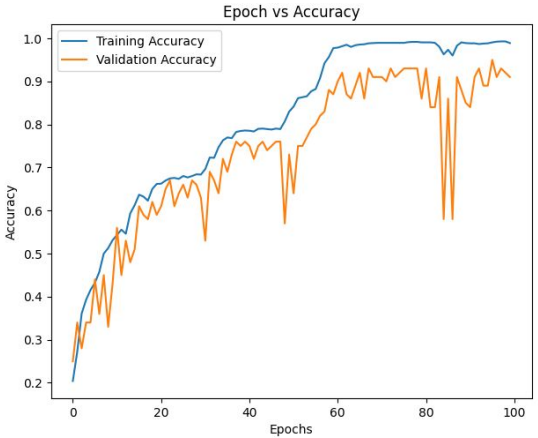


Convolutional Neural Network

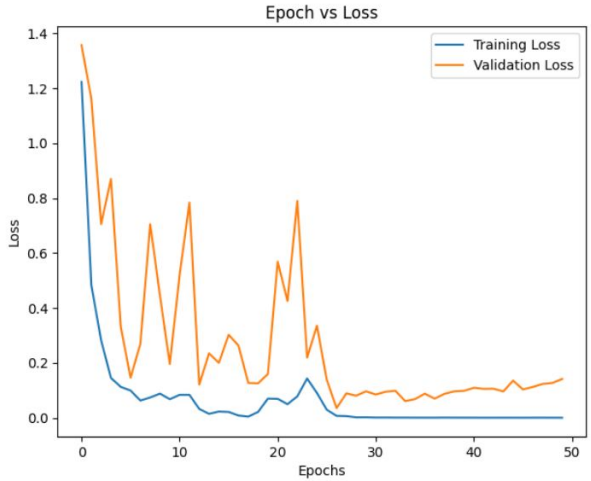
```
Train loss: 1.47
Train acc: 0.98
Val loss: 1.55
Val acc: 0.91
```

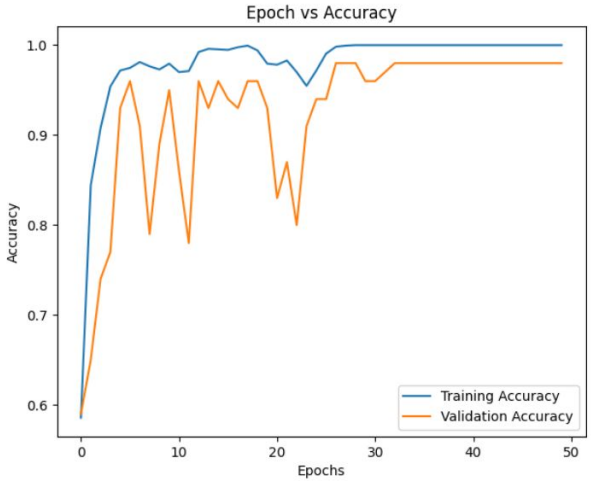
- changing the padding value in convolutional layers
- adding more convolutional layers to the network
- increasing the number of epochs
- All these helped to improve the model and increase its accuracy.





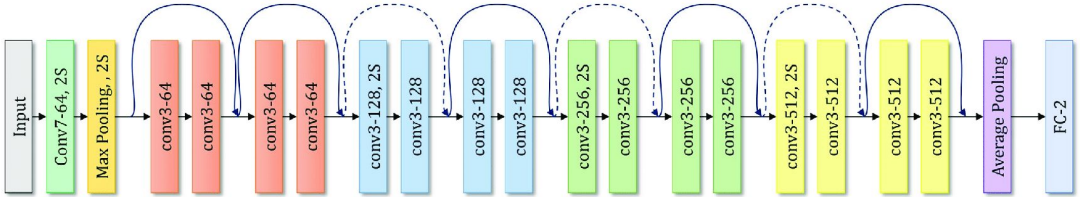
ResNet18





ResNet model result on training and validation set

Test set accuracy : 95.6%
Test set loss : 0.198
Train accuracy : 99%
Validation accuracy: 98%



Architecture of the ResNet-18

conclusions

- Data preparation is an important step to reaching high accuracy
- Changing hyperparameters can help the model performance
- Adding noise can help model performance
- ResNet18 acting better compare to other version of ResNet

❑ We learned how to :

➔ convert audios to spectrograms

➔ How to add noise to dataset for data augmentation

➔ How to handle overfitting

❑ Our codes are available at : https://colab.research.google.com/drive/1_mBKKLiryrensHQwtxXWRl0R4-S8RhhC?usp=sharing