

Music Genre Classification Abilities of Neural Network Architectures

Joshua Engelkes
enge0479@morris.umn.edu

Outline

1. Introduction
2. Background
3. Bottom-up Broadcast Neural Network
4. Independent Recurrent Neural Network
5. Comparison
6. Conclusion

Introduction

Music Information Retrieval (MIR)

- The number of digital audios uploaded on the internet is increasing
- Genre classification is widely requested for music applications
- MIR attempts to automatically classify music data
- Data driven methods are proven useful to MIR

Background

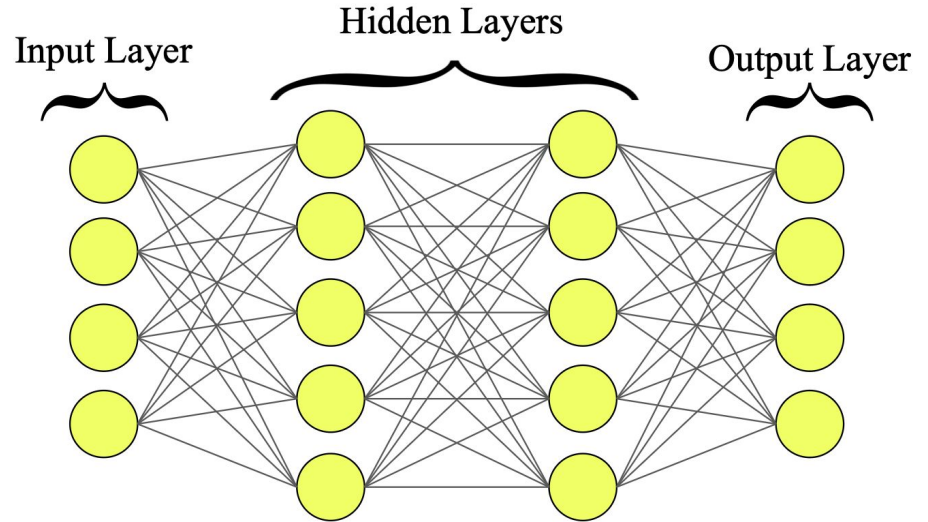
Neural Networks

General

- Collections of algorithms inspired by the communication in the human brain
- “Learns” through trial and error

$$h_{n,l} = \sigma \left(\sum_{i \in N} (x_{i,l-1} \times w_{i,l-1}) + b_n \right)$$

- h: hidden state
- n: neuron
- l: layer
- x: input
- w: weight
- b: bias
- σ : activation function
- N: # of neurons in previous layer



Basic Transformations

Batch Normalization

(BN)

Standardizes data

Procedure:

1. Calculate mean and standard deviation of data points
2. Fit data points to normal distribution with mean 0 and standard deviation 1.
3. Rescale/Offset data points

Rectified Linear Activation

(ReLU)

Applies function to data points:

$$y = \text{MAX}(x, 0)$$

- x: input of ReLU
- y: output of ReLU

Softmax

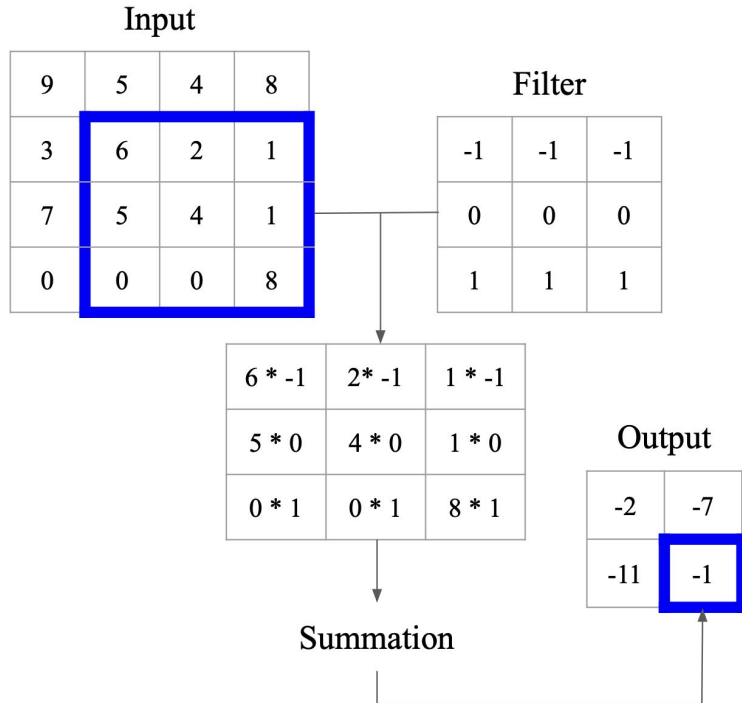
Produces vector of probabilities

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

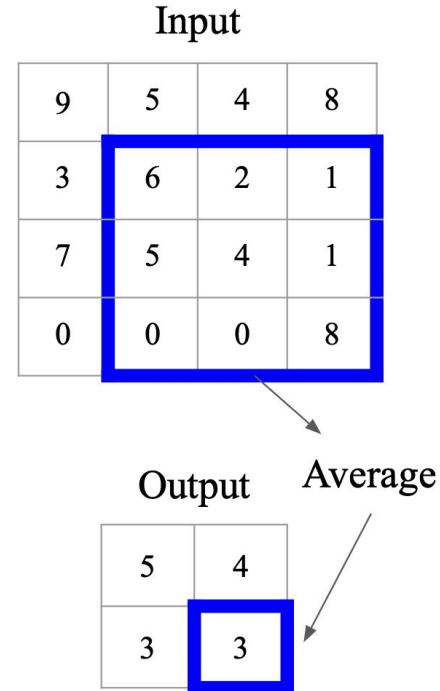
- σ : softmax function
- \mathbf{z} : vector of inputs
- i : category index
- K : number of categories

Grid Transformations

Convolutional Layers



Pooling Layers



Convolutional Neural Networks (CNN)

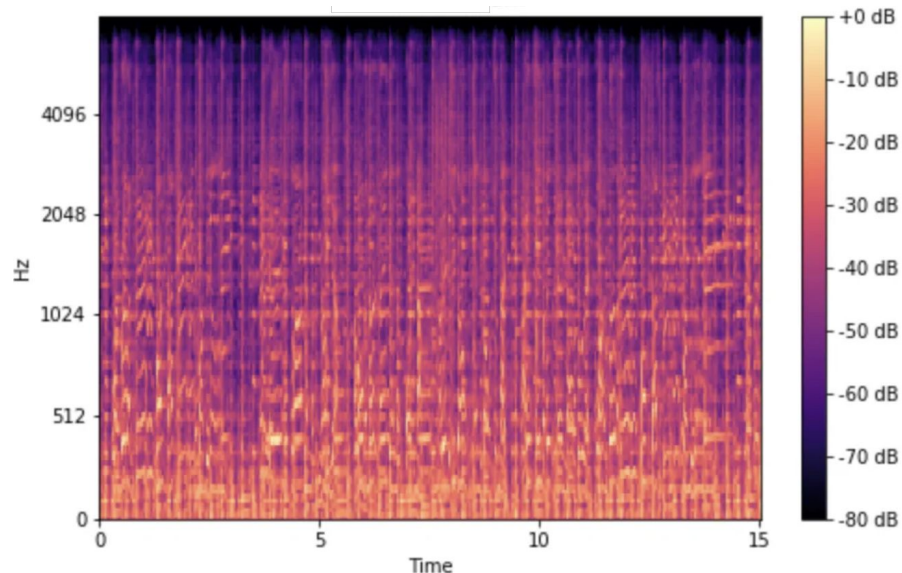
Pattern and object recognition in 2D data

Tools:

- Convolutional Layers
- Pooling Layers

MIR CNN Input: Mel Spectrograms

Common Problem: Loss of low-level information

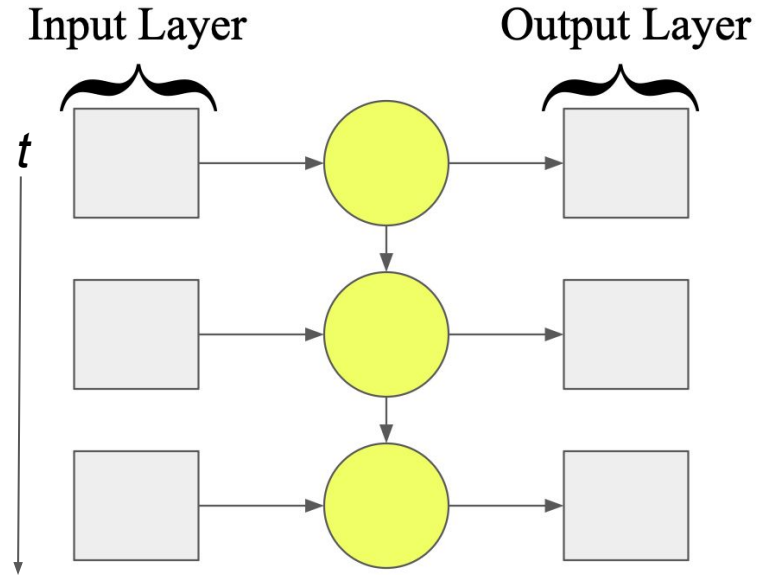
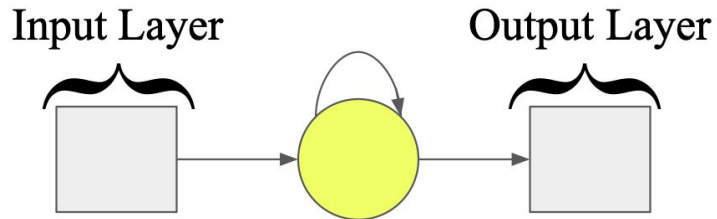


Recurrent Neural Networks (RNN)

Pattern and object recognition in sequential data

Tool:

- Recurrent Nodes



Bottom-Up Broadcast Neural Network (BBNN)

Liu et al. 2021

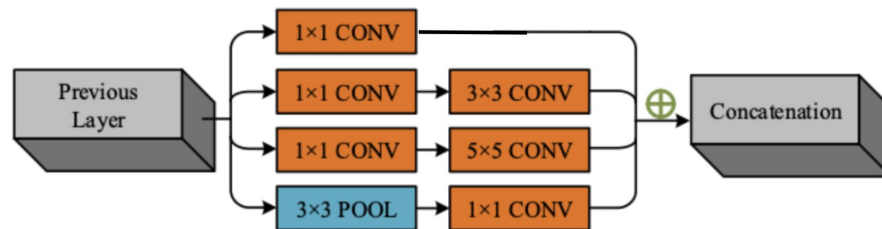
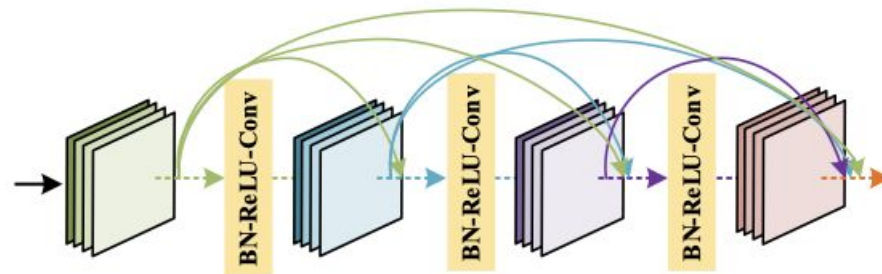
Bottom-up broadcast neural network for music genre classification

Broadcast Module

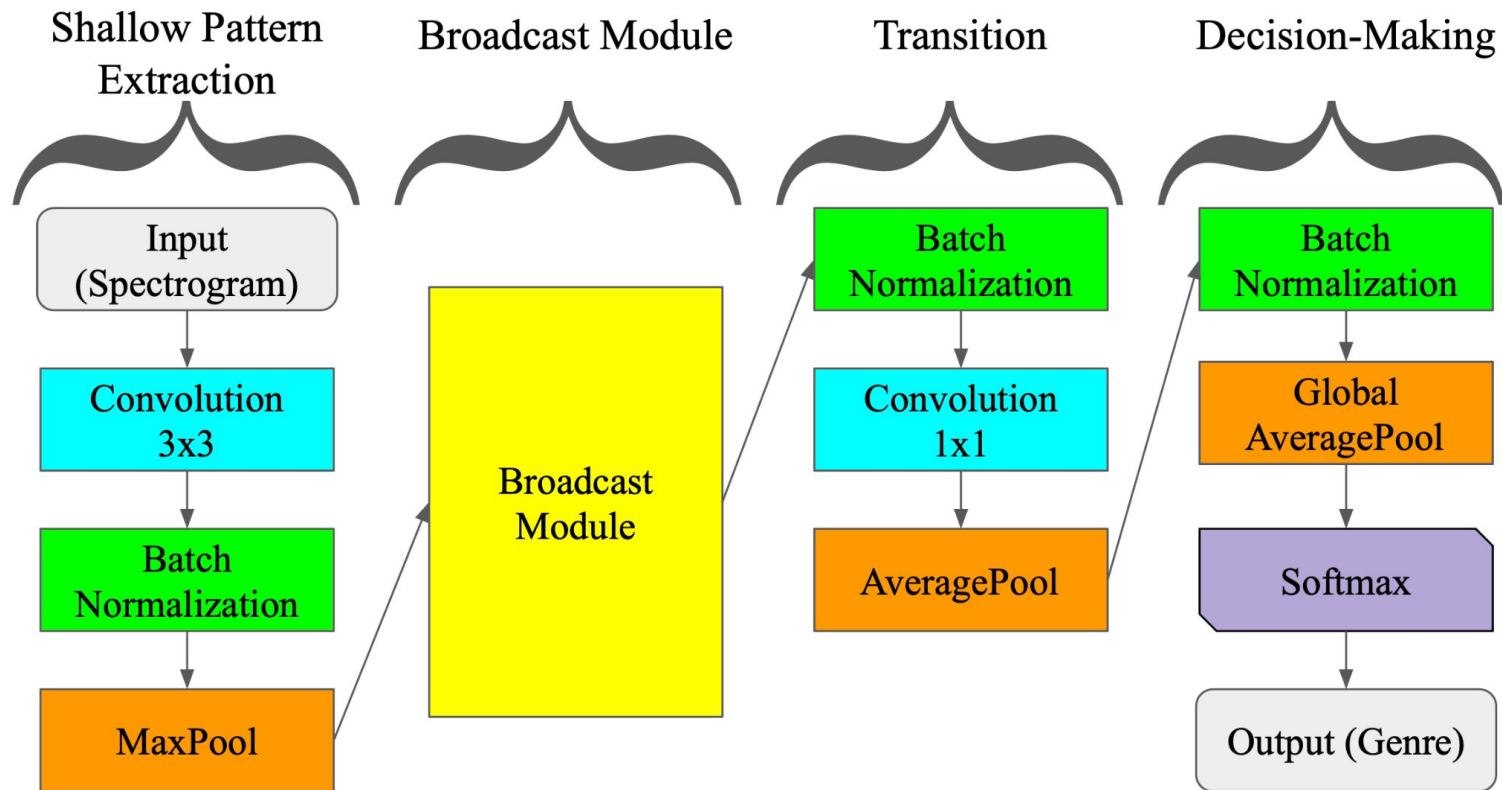
Basic Extraction Unit: Inception Blocks

Qualities:

- Three layers
- Densely connected
- Four paths with variations on convolutional layers



Genre Classification BBNN



Independent Recurrent Neural Network (IndRNN)

Wu et al. 2018.

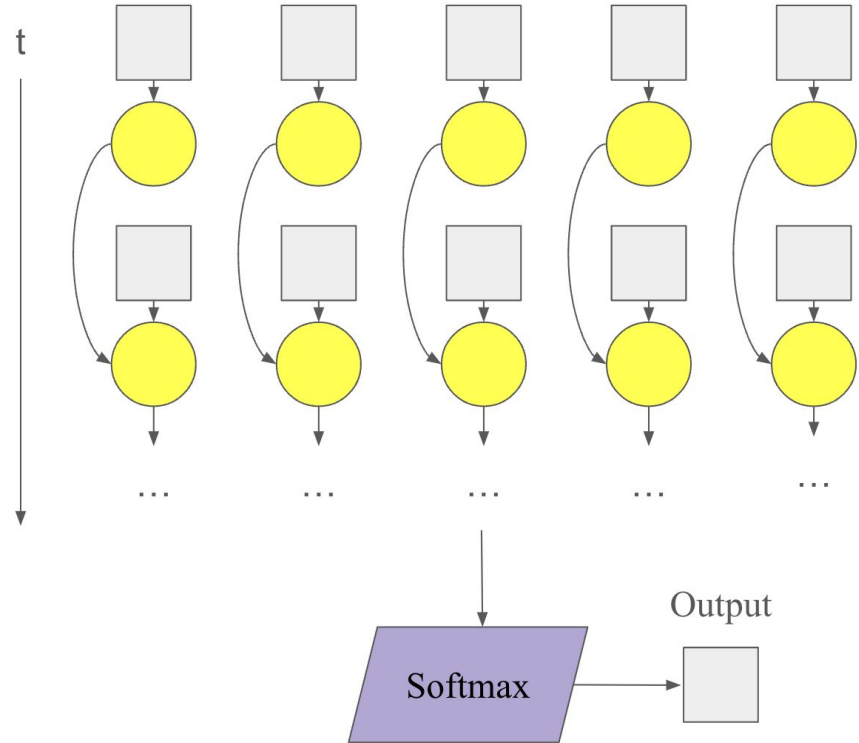
Music Genre Classification Using Independent Recurrent Neural Network.

Independent Recurrent Neural Network

- Takes a vector of standardized values as input
- Neurons cannot view the hidden states of other neurons in the same layer

$$h_{n,t} = \sigma(w_n x_t + u_n h_{n,t-1} + b_n)$$

- n: neuron
- t: timestep
- w/u: weights
- b: biases
- σ : ReLU



Comparison

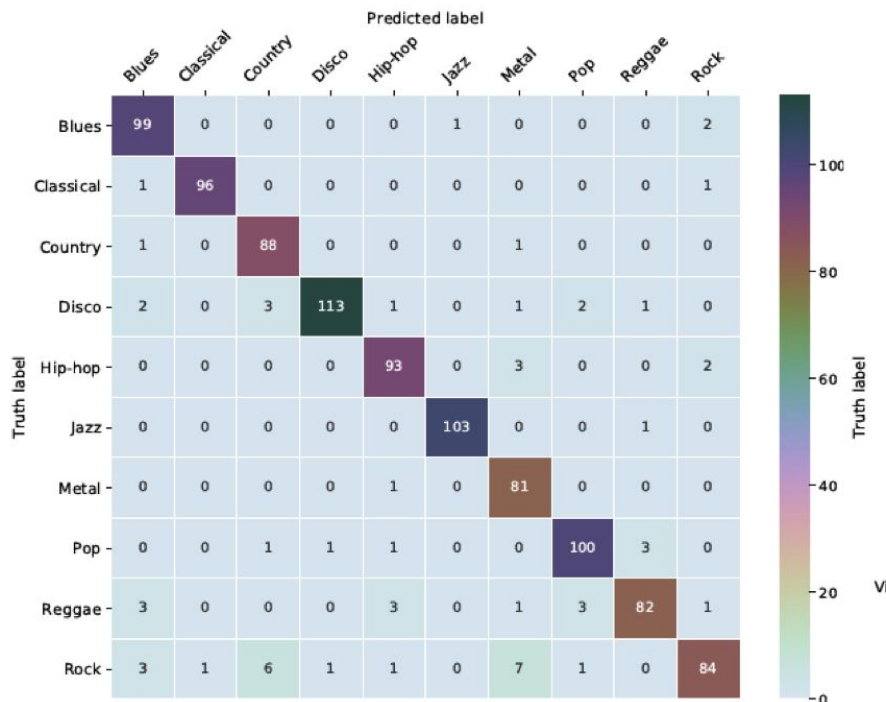
Experiment

Genres	# of Songs
Blues	100
Classical	100
Country	100
Disco	100
Hiphop	100
Jazz	100
Metal	100
Pop	100
Reggae	100
Rock	100

GTZAN Dataset

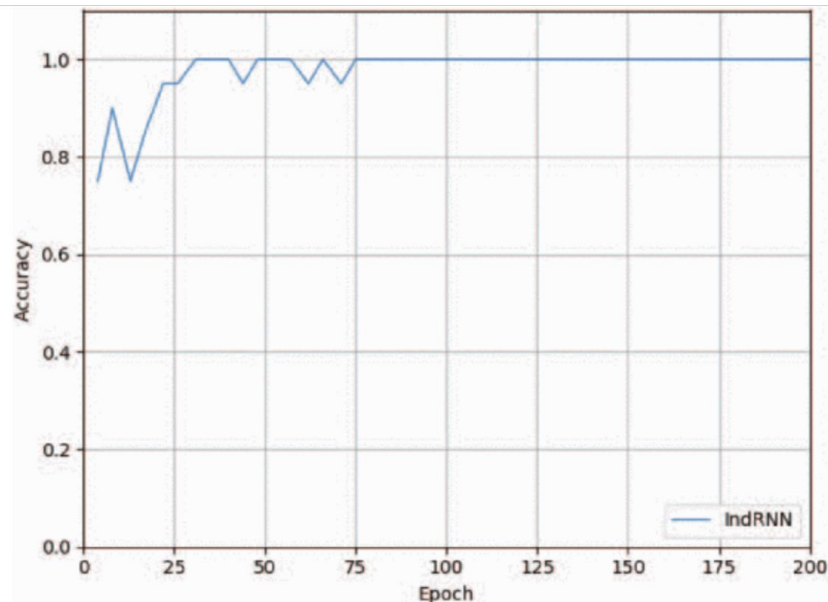
- Audio tracks of 30 seconds
- Randomly split:
 - 90% used for training
 - 10% used for testing
- Training epochs:
 - IndRNN: Varies from 5 to 200 times
 - BBNN: set at 100 times

Results



BBNN Accuracy per Genre

	BBNN	IndRNN
Average Classification Accuracy	93.9%	96.0%



IndRNN Accuracy by Training Rounds

Conclusion

Conclusion

- The IndRNN outperformed the BBNN in music genre classification accuracy
- The IndRNN classifications had peak accuracy over 75 epochs
- The BBNN had issues differentiating rock audios from country and metal audios

Sources

Caifeng Liu, Lin Feng, Guochao Liu, Huibing Wang, and Shenglan. 2021. Bottom-up broadcast neural network for music genre classification. *Multimedia Tools and Applications* 80 (2021), 1–19. <https://doi.org/10.1007/s11042-020-09643-6>

George Tzanetakis and Perry Cook. 2002. Musical Genre Classification of Audio Signals. *IEEE Transactions on Speech and Audio Processing* 10 (08 2002), 293 – 302. <https://doi.org/10.1109/TSA.2002.800560>

Keunwoo Choi, György Fazekas, Mark Sandler, and Kyunghyun Cho. 2017. Transfer learning for music classification and regression tasks. Quazi Ghulam Rafi, Mohammed Noman, Sadia Zahin Prodhan, Sabrina Alam, and Dip Nandi. 2021. Comparative Analysis of Three Improved Deep Learning Architectures for Music Genre Classification. *International Journal of Information Technology and Computer Science* 13 (04 2021), 1–14. arXiv:1703.09179

Wenli Wu, Fang Han, Guangxiao Song, and Zhijie Wang. 2018. Music Genre Classification Using Independent Recurrent Neural Network. In *2018 Chinese Automation Congress (CAC)*. 192–195. <https://doi.org/10.1109/CAC.2018.8623623>

Zhouyu Fu, Guojun Lu, Kai Ming Ting, and Dengsheng Zhang. 2011. A Survey of Audio-Based Music Classification and Annotation. *IEEE Transactions on Multimedia* 13, 2 (2011), 303–319. <https://doi.org/10.1109/TMM.2010.2098858>

Questions?