Convolutional recurrent neural networks for music classification

7107018026

劉俊廷

Contents

CNN and CRNN for music classification

Introductions to the models
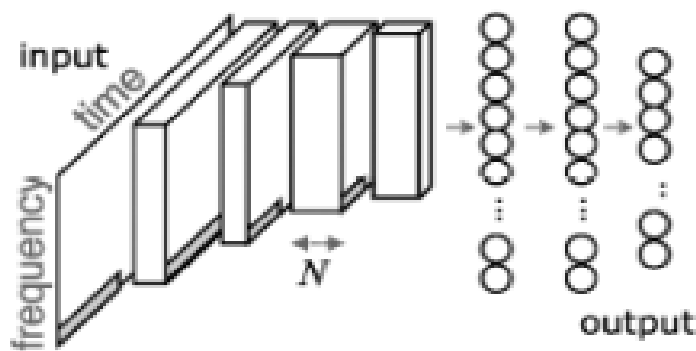
Experiments

References

CNN and CRNN for music classification

- CNN
  - CNNs assume features that are in different levels of hierarchy and can be extracted by convolutional kernels.
  - CNN for music classification:

    music tagging: AUC:0.88~0.89

    genre classification: accuracy:66~69%

    user-item latent feature prediction for

    recommendation: AUC:0.77192

CNN and CRNN for music classification

- CRNN
  - CRNN
    - Modified CNN by replacing the last convolutional layers with a RNN.
    - Adopting an RNN for aggregating the features enables the networks to take the global structure into account while local features are extracted by the remaining convolutional layers.

  - CRNN for music classification:
    - Music transcription: Accuracy:72%~76%
    - Document classification:Accuracy:65%~67%
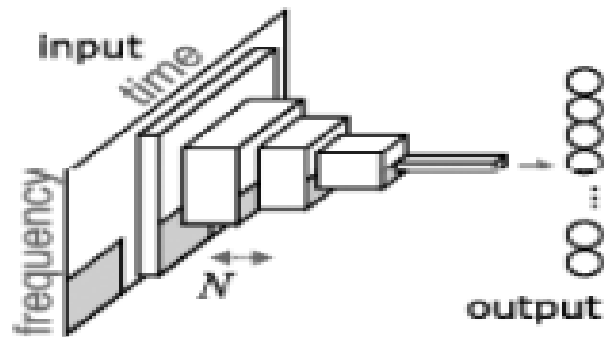   - Image classification:Accuracy:51.4%

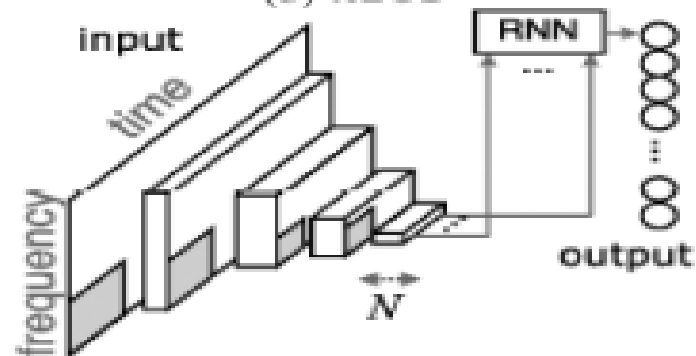# Introductions to the models

(a) k1c2
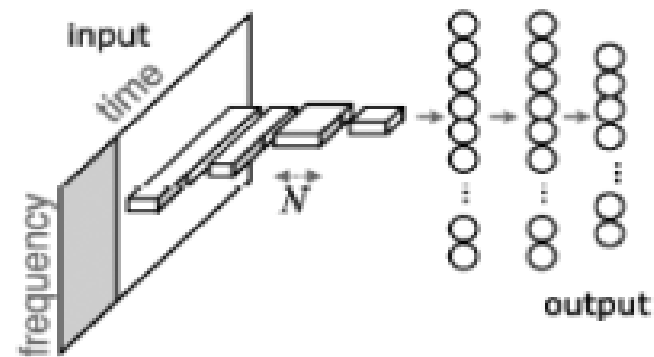
(b) k2c1

(c) k2c2

(d) CRNN

Sigmoid functions are used as activation at output nodes because music tagging is a *multi-label* classification task.

Experiments

- Dataset
  - the Million Song Dataset with *last.fm* tags.
  - Predict the top-50 tag, which includes genres (e.g., *rock*), moods (e.g., *happy*), instruments (e.g., *guitar*), and eras (*60s–00s*).
  - 214285 (201680 for training and 12605 for validation).
- Input: Resulting in an input shape of 96×1366 (mel-frequency band×time frame).
- We use ADAM for learn and binary cross-entropy as a loss function.
- Computes AUC-ROC for compare four models
- The result of this paper,the AUC of CRNN can arrive 0.85~0.87.

Thanks for listening

References

- Keunwoo Choi, George Fazekas, Mark Sandler, "Automatic tagging using deep convolutional neural networks"

- Paulo Chiliguano, Gyorgy Fazekas, "Hybrid music recommender using content-based and social information"

- Aaron Van den Oord, Sander Dieleman, Benjamin Schrauwen, "Deep content-based music recommendation"

- Duyu Tang, Bing Qin, Ting Liu, "Document modeling with gated recurrent neural network for sentiment classification"

- Zhen Zuo, Bing Shuai, Gang Wang, Xiao Liu, Xingxing Wang, Bing Wang, Yushi Chen, "Convolutional recurrent neural networks: Learning spatial dependencies for image representation",

- Siddharth Sigtia, Emmanouil Benetos, Simon Dixon, "An end-to-end neural network for polyphonic piano music transcription"

- Diederik P. Kingma, Jimmy Ba, Adam: A method for stochastic optimization