Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

ISSN NO. 2320-5407

Article DOI:10.21474/IJAR01/14474
DOI URL: http://dx.doi.org/10.21474/IJAR01/14474

RESEARCH ARTICLE

A MUSIC GENERATION BY A COMBINING MODElOF RESNET AND LSTM NETWORKS

Kazuya Ozawa and Hideaki Okazaki
Graduate School of Electrical and Information Engineering, Shonan Institute of Technology, 1-1-25 Tsujido Nishikaigan, Fujisawa 251-8511, Japan.

……………………………………………………………………………………………………....

Manuscript Info

……………………….

Manuscript History
Received: 25 January 2022
Final Accepted: 28 February 2022
Published: March 2022

Key words:-
MIDI, LSTM,Residual Neural Networks, and Skip connection

Abstract

………………………………………………………………

In this paper, to automatically generate a music for the melody part by deep learning with training data collected from Chopin's piano piecies, a combining model of Residual Neural Networks(ResNet) and Long-Short Term Memory Networks (LSTM) are proposed. First, to generate a music for the melody part of a piano music, a training dataset used for deep learning is provided. Secondly, by using each of a LSTM Model and a combining model of LSTM and ResNet,experiments on music generationare presented. Thirdly, the results of music generation by each model are compared and discussed. In conclusion, the principal results are summarized.

……………………………………………………………………………………………………....

Introduction:-

Recently, deep learning has been actively studied not only in the field of scientific and engineering information processing, but also in artistic fields such as music generation [1] - [4]. In deep learning for music generation, acoustic signals, MIDI, and musical notation are frequently used to learn musical structures [5] such as pitch, duration (start/end time), and velocity. Especially MIDI is widely used for learning musical data because musical structures such as pitch, duration (start/end time), and velocity are stored in a file, and its file size is small. In [1] and [2], experiments for music generation are performed by using MIDI training (or learning) data. Music generation by deep learning has been experimented with neural network models such as RNNs and CNNs. RNNs can predict time series by incorporating past hidden layers, but they cannot learn long-term time dependence because vanishing gradient problem occurs in the learning process. Therefore, RNNs are said to be less effective for music generation. In addition, Long-Short Term Memory Network (LSTM): a kind of RNN, is designed to learn both long and short term time dependencies, and can be more effective than RNNs in music generation [4]. Simply deepening the layers will not allow learning to proceed well from a certain depth, resulting in a loss of accuracy. This is said to be due to vanishing gradient caused by increasing the layer depth. Residual Neural Network (Res Net): a kind of RNN, is also designed to solve this vanishing gradient problem. By learning a residual function that takes the difference between the output of a layer and the input, Res Net successfully solves vanishing gradient problem that occurs when the layers are deepened[6], [7].This paper is organized as follows. Sect. 2 provides a training dataset used for deep learningwith a collection learning data of Chopin’s piano pieces, to generate a music for the melody part of a piano music.Sect. 3 presents the LSTM model and the model combining LSTM and ResNet, respectively. Sect. 4 compares and discusses the results of music generation by each models.In conclusion, the principal results are summarised.

Corresponding Author:- Kazuya Ozawa
Address**:-** Graduate School of Electrical and Information Engineering, Shonan Institute of Technology,1-1-25 TsujidoNishikaigan, Fujisawa 251-8511, Japan.

Training Detaset:-

In this section, we providea training dataset used for deep learning to generate a music for the melody part of a piano music. We prepare a data set of MIDI files consisting of multiple parts such as melody and accompaniment for learning. The data set is collected from the Classical Piano Midi Page [8]. In this paper, we use 10 pieces of Chopin's piano musics. We chose 10 Chopin pieces such that each piece has similar music impression. The JuliaMusic package addressedin this paper is created from the MIDI.jl, MotifSequenceGenerator.jl, MusicManipulations.jl, and MusicVisualizations.jl packages.The MIDI.jl package decomposes notes into four numerical values such as the temporal position, the duration, the pitch, and the velocity [9]. Training data utilizes the pitch, duration and velocity of the melody part. The pitch values are obtained by converting the note names with the JuliaMusic package [10]. See [11] for details on converting note names to pitch values.

List 1 :10 pieces of Chopin's piano musics in the Classical Piano Midi Page [8].

1.Piano right : Chopin Prelude No. 1, Opus 28,
2. Piano right : Chopin Prelude No. 15, Opus 28
3. Piano right : Chopin Prelude No. 19, Opus 28
4. Piano right : Chopin Prelude No. 2, Opus 28
5. Piano right : Chopin Prelude No. 21, Opus 28
6. Piano right : Chopin Prelude No. 22, Opus 28
7. Piano right : Chopin Prelude No. 4, Opus 28
8. Piano right: Chopin Prelude No. 9
9. Piano right: Etüde Opus 10 No. 5
10. Piano right: Chopin Mazurka Opus 7 No. 1

Neural Network Models, and Trained Models:-

In this section, as neural networks for generating piano musics, we definea LSTM Model and a combining model of LSTM and ResNet. Further, we define trained models such that the LSTM and the combining models are trained by using the training dataset of 10 pieces of Chopin's piano musics of List 1 in the section 2.

LSTM Model:-

Table 1 shows the structure of LSTM Model. LSTM Models constructed with 5 layers (Fig. 1). In the first layer, Layer Norm is given to perform the normalization process [12], and LSTM layer is applied in the second, and third layers. Then, the fourth, and fifth layers represent the fully connected layers.

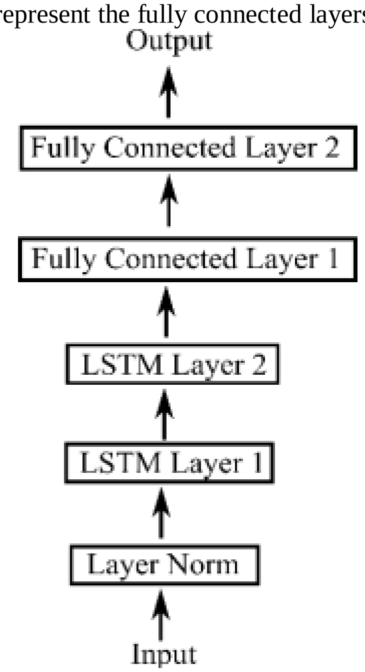

Figure. 1:-LSTM Model.

Table 1:- Structure of LSTM Model.

Layer	Input nodes	Outputnodes
Layer Norm	3	3
LSTM layer 1	3	30
LSTM layer 2	30	30
Fully connected layer 1	30	30
Fully connected layer 2	30	3

Combining Model of LSTM and ResNet:-

Table 2 shows the combining model of LSTM and ResNet. The combining models constructed with 5 layers (Fig. 2). In the first layer, Layer Norm is given to perform the normalization process, and LSTM Layer is applied in the second, and third layers. Then, the fourth, and fifth layers represent the fully connected layers. Then, the Skip Connection is inserted before Fully Connected Layer1, and afterLSTM Layer1.

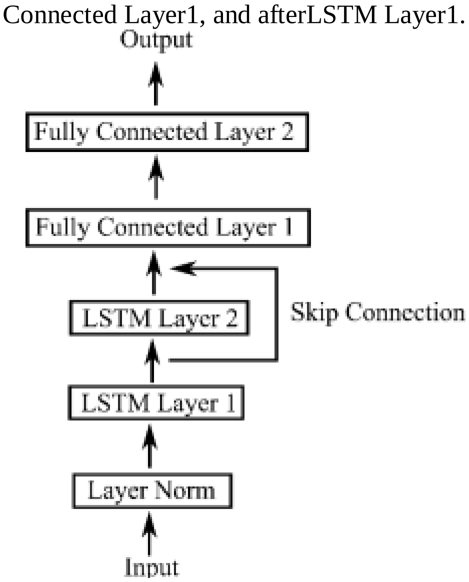

Figure. 2:-Combining model of LSTM and ResNet.

Table 2:- Structure of Combining model of LSTM and ResNet.

Layer	Input nodes	Outputnodes
Layer Norm	3	3
LSTM Layer 1	3	30
LSTM Layer 2	30	30
Fully connected layer 1	30	30
Fully connected layer 2	30	3

3.3 Trained Models :-

Using JuliaMusic.jl[10], we can create a music note with the function : Note(pitch, velocity, position, duration, channel = 0), where pitch::UInt8 : Pitch, starting from C-1 = 0, adding one per semitone, velocity::UInt8 : Dynamic intensity. Cannot be higher than 127 (0x7F), position::UInt : Position in absolute time (since beginning of track), in ticks, duration::UInt : Duration in ticks, and channel::UInt8 = 0 : Channel of the track that the note is played on. Cannot be higher than 127 (0x7F). Therefore, using the following Julia function including a Trained_Model_k, k=1,2 :

```
{file = MIDIFile()
track = MIDITrack()
notes = Notes(tpq=480)
pitch, velocity, duration = 60.0, 49.0, 80.0
position = 300
for i in 1:300
```

```
    pitch, velocity, duration = Trained_Model_k([pitch, velocity, duration])
    position = position + 300
    notes = [Note(pitch, velocity,position, duration )]
    addnotes!(track, notes)
end
addtrackname!(track, "simple track")
push!(file.tracks, track)
```

writeMIDIFile("./sample_k.mid", file)}, we can generate MIDI file with 300 notes, e.g. sample_k.mid, where Trained_Model_k, k=1,2 represent LSTM Model, and the combining model of LSTM and ResNet in the section 3 such that these models are trained by using the training data set of 10 pieces of Chopin's piano musics of List 1 in the section 2 , respectively. Table 3 and Table 4 show the hyper parameters of Trained_Model_k, k=1,2, respectively.

Table 3:-Hyper parameters of Trained_Model_1.

Description	**Value or Type**
Activation function	ReLU
Loss function	Mean Absolute Error (MAE)
Optimization algorithm	Adam
Learning rate	1e-3
Epoch number	5000

Table 4:- Hyper parameters of Trained_Model_2.

Description	**Value or Type**
Activation function	ReLU
Loss function	Mean Absolute Error (MAE)
Optimization algorithm	Adam
Learning rate	1e-3
Epoch number	5000

2. Experimental results of Music Generation and Discussions:-

In this section, we present experiments on music generation, by using each of Trained_Model_k, k=1,2.The following results are obtained by Flux: Julia Machine Learning Library[13], [14].Fig. 3 (or Fig. 5) and Fig. 4 (or Fig. 6) show the scores of piano music generated by Trained_Model_k, k=1,2, respectively.In Fig. 3, or Fig. 5,the music generated by Trained_Model_1 is a very monotonous melody with continuously same output notes. In Fig. 4, or Fig. 6,the music generated by Trained_Model_2 ismore dynamic and more emotional melody than the music in Fig.3 or Fig. 5.The music generated by Trained_Model_2 is supposed to be more successful than the music by Trained_Model_1. These musics by Trained_Model_k, k=1,2 are available athttps://soundcloud.com/ozawa-kazuya-244008895/sets/music-generation-by-deep-learning-1/s-m1DIWndD0rX?utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing.

Figure. 3:-The first twenty measures of the score of piano music generated by Trained_Model_1.

Figure. 4:-The first twenty measures of the score of piano music generated by Trained_Model_2.

Conclusion:-

The following results were obtained:

1. We have provided the training data set for deep learning to generate a piece of music by collecting the melody part of Chopin's piano pieces.
2. We have presented the LSTM model and the combining model of LSTM and ResNet for generating piano musics, and have provided those traind models.
3. Compairing the results of musics generated by the trained models, we have concluded that the combining model of LSTM and ResNet are more effective and more excellent than LSTM modelfor generating piano musics.
4. As future works, we will build the structure of neural networks for generating symphonies.

References:-

1.Yang, L. C., Chou, S. Y., & Yang, Y. H. (2017). MidiNet: A convolutional generative adversarial network for symbolic-domain music generation. arXiv preprint arXiv:1703.10847.
2. Huang, A., & Wu, R. (2016). Deep learning for music. arXiv preprint arXiv:1606.04930.
3. Bhave, A., Sharma, M., & Janghel, R. R. (2019). Music generation using deep learning. In Soft Computing and Signal Processing (pp. 203-211). Springer, Singapore.
4. Eck, D., & Schmidhuber, J. (2002). A first look at music composition using lstm recurrent neural networks. Istituto Dalle Molle Di Studi Sull Intelligenza Artificiale, 103, 48.
5. Briot, J. P., Hadjeres, G., & Pachet, F. D. (2017). Deep learning techniques for music generation--a survey. arXiv preprint arXiv:1709.01620.
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016, October). Identity mappings in deep residual networks. In Europeanconference on computer vision (pp. 630-645).
Springer, Cham.https://arxiv.org/abs/1603.05027arXiv:1603.05027 [cs.CV], 2016.
8. Classical Piano Midi Page,http://www.piano-midi.de/
9. Datseris, G., & Hobson, J. (2019). MIDI. jl: Simple and intuitive handling of MIDI data. Journal of Open Source Software, 4(35), 1166.
10. JuliaMusic, https://juliamusic.github.io/JuliaMusic_documentation.jl/dev/
11.Note names, MIDI numbers and frequencies, http://newt.phys.unsw.edu.au/jw/notes.html
12. Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer normalization. arXiv preprint arXiv:1607.06450.
13. The Julia Project, "The Julia Language," May 2020
https://docs.Julialang.org/en/v1/
14. Flux: The Julia Machine Learning Library,https://fluxml.ai/Flux.jl/stable/

Appendix:-

In the following, the entire musical scores of the piano musics generated by LSTM Model (Model_1), andthe combining model of LSTM and ResNet (Model_2) are shown in Fig. 5, and Fig.6, respectively.

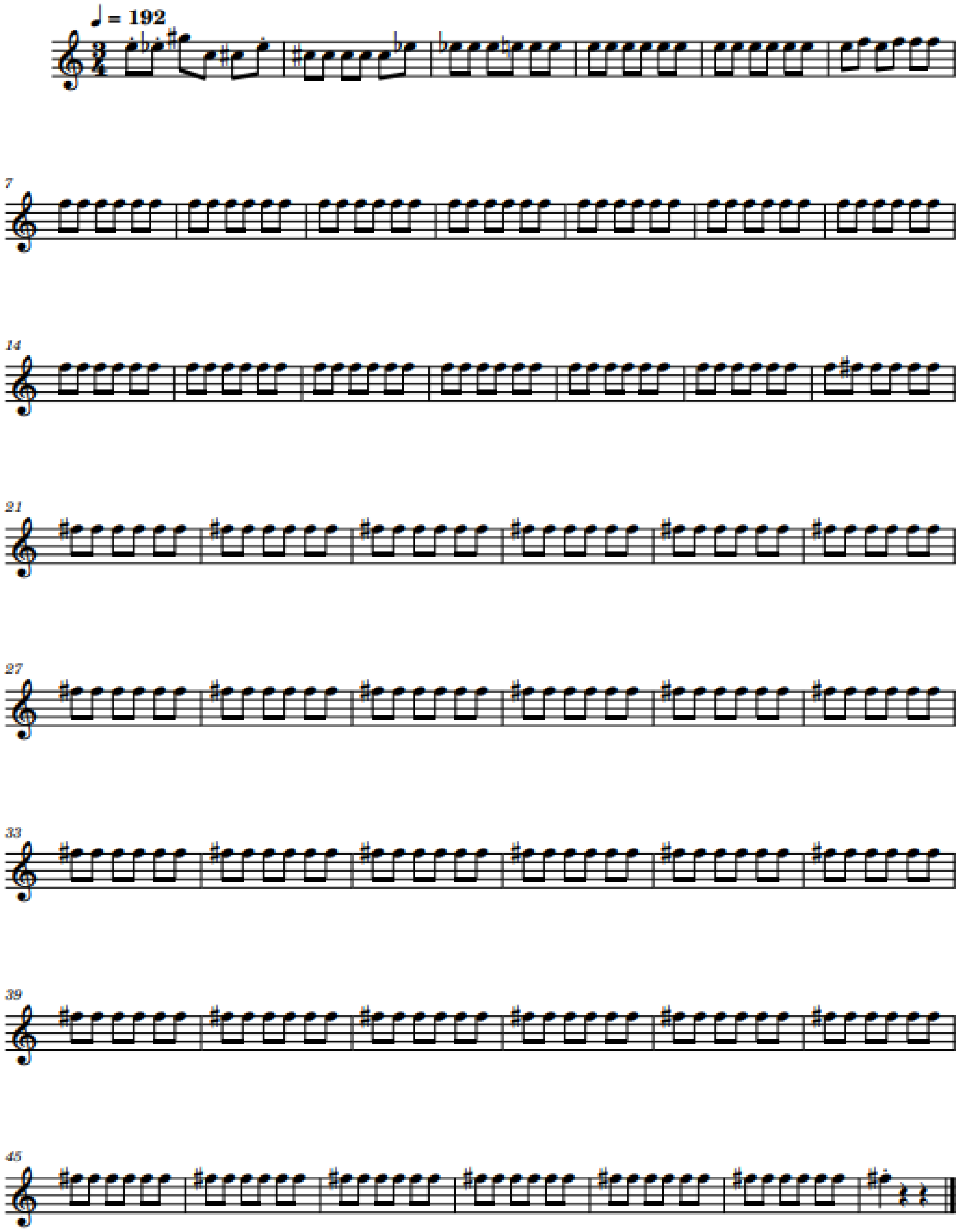

Figure. 5:- Entire musicalscore of piano music generated by Trained_Model_1.

Figure. 6:- Entire musicalscore of piano music generated by Trained_Model_2.