THE GENERATION OF PIANO MUSIC IN THE STYLE OF JOHANNES BRAHMS USING NEURAL NETWORK ARCHITECTURES

JAMES DOHERTY

BRENDAN TIERNEY

INTRODUCTION

- Neural Network Architectures currently generate music at a basic compositional standard
- They struggle with the complex motifs and harmonic structures featured in Western Classical Music
- This paper aimed to generate piano music in the style of Johannes Brahms
- Various neural network models were trained with a dataset of MIDI files containing Brahms’ piano works
- To deems success of project, generated pieces must show statistical similarities to the original piano works of Brahms

GAPS/MOTIVATION AND RESEARCH PROBLEM

- No models have been created to generate the work of Brahms
- Computer generated music only sounds human-like when short excerpts are generated
- Surveys in previous papers not suitable for experimentation
- Research problem addresses AI's ability to generate and composite music in the future
- An assumption is that it is already possible to train neural network models the general characteristics and patterns of musical composers

GAPS/MOTIVATION AND RESEARCH PROBLEM

- **Research Question:** *To what extent can the accuracy of various Neural Network Models, trained with Long Short-Term Memory and Numerous Attention mechanisms, be significantly improved by augmenting MIDI files containing the compositional works of Johannes Brahms with an augmentation pipeline to generate pieces of music that are mistaken by professional musicians, composers and conductors as one of Brahms' own works?*
- **H0:** *Neural network models cannot generate piano works to the same level of musicality and emotion as Brahms. Due to this, generated pieces will not be statistically similar through Music Information Retrieval or mistaken as a work of Brahms by professional musicians, composers and conductors through a quantitative survey using Likert Scales.*
- **H1:** *If an augmentation pipeline is utilised to expand a MIDI dataset of pre-processed files containing the piano works of Johannes Brahms, then various neural network models trained with Long Short-Term Memory and numerous Attention mechanisms could generate pieces of music that is statistically similar to Brahms and could be mistaken as one of Brahms' own piano works by professional musicians, composers and conductors through a quantitative survey using Likert Scales and various Independent-Samples T-Tests and Hotelling's T2 Tests being implemented to determine whether the p-value is > 0.05 in order to reject the null hypothesis.*

LITERATURE REVIEW & BACKGROUND

- Researched various data collection & preprocessing methods
- Researched papers consisted of much larger datasets
- Various Music Information Retrieval functions were studied for quantitative and visual analysis

LITERATURE REVIEW & BACKGROUND

- Various neural network architectures were researched:
  - Recurrent Neural Network (RNN)
  - Long Short-Term Memory RNN
  - Transformer models with:
    - Self-attention
    - Relative Attention
    - Relative Self-Attention
    - Local Windowed Attention
    - Spare Attention
  - Perceiver AR

DESIGN AND RESEARCH METHODOLOGIES

- Research project carried out in 3 stages:
  - Data collected, pre-processed and augmented
  - Training of Neural Network Models
  - Best generated examples from models analysed through Music Information Retrieval functions and Quantitative Survey and evaluated to test the research hypothesis
- Dataset Preprocessing techniques
  - Track Splitting
  - Conversion to single track
  - Normalisation
  - Quantisation
- Dataset Augmentation techniques
  - Time-stretching
  - Transposition

IMPLEMENTATION AND EXPERIMENTS

- Overfitting issues observed and mitigated
- Model performance metrics stated Transformer model with Relative Self-Attention scored best training loss and accuracy scores with Recurrent Neural Network scoring worst
- Experiment consisted of each model generating a two minute piece which contained 300 prime tokens (15 Seconds) from 6 of Brahms' works
- Generated pieces were then compared with the first 2 minutes of the original Brahms pieces

Table 1: Model Performance Metrics

Model	**Training Loss**	**Training Accuracy**
Recurrent Neural Network	0.550	0.628
RNN with Long Short-Term Memory	0.154	0.983
Transformer w/ Self-Attention	0.737	0.939
Transformer w/ Relative Attention	0.447	0.985
Transformer w/ Relative Self-Attention	**0.015**	**0.995**
Transformer w/ Local Windowed Attention	0.015	0.993
Sparse Transformer	0.612	0.803
Perceiver AR	0.825	0.774

IMPLEMENTATION AND EXPERIMENTS

- Music Information Retrieval experiments:
  - Entropy
  - Normalised Pairwise Variability Index (nPVI)
  - Global Energy
  - Mean Roughness
  - Pulsation Clarity
  - Duration Distribution
  - Pitch Class Distribution

RESULTS AND DISCUSSION

- No statistical significance found with most variables
- Global Energy and Entropy contained statistical significance indicating improvements needed for complexity, uncertainty and energy to be at similar level to Brahms
- Quantitative Survey stated no statistical significance between Brahms and generated pieces
- Music Information Retrieval evaluation concluded that Transformers models with self-attention, local windowed attention and relative global attention performed best

Global Energy

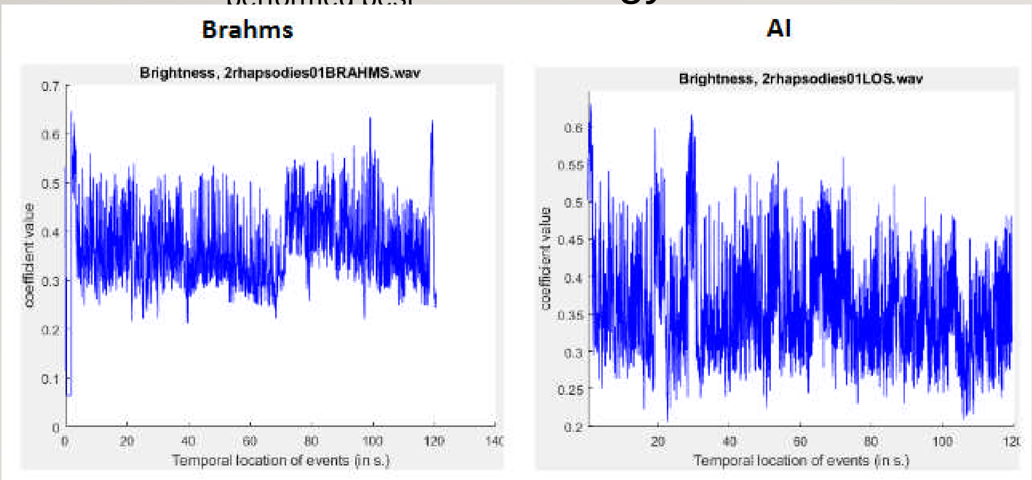

Figure 1: Brightness Curve between Brahms and AI

Entropy

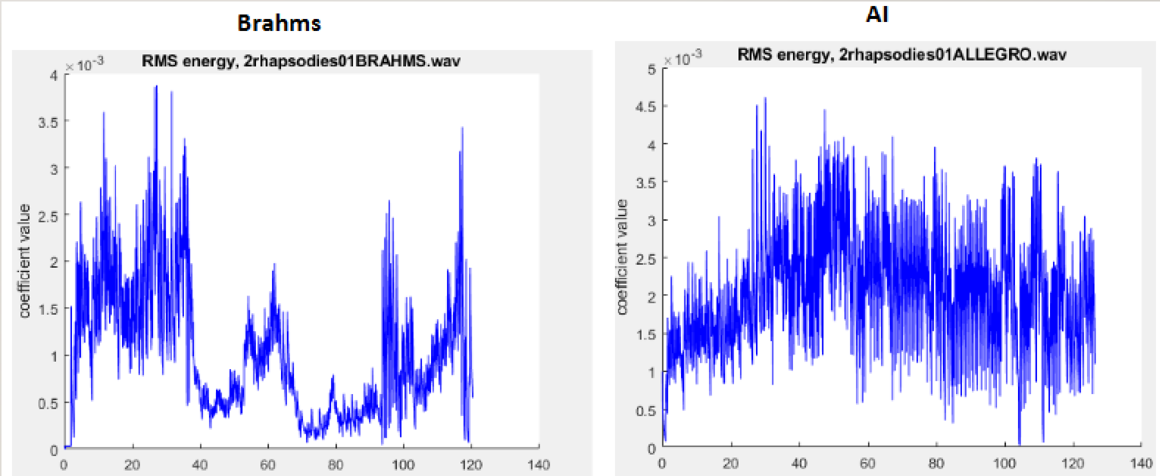

Figure 2: Temporal Evolution curves for Brahms and AI

RESULTS AND DISCUSSION

- D# Major – D#, G, A#
- C Minor – C, D#, G

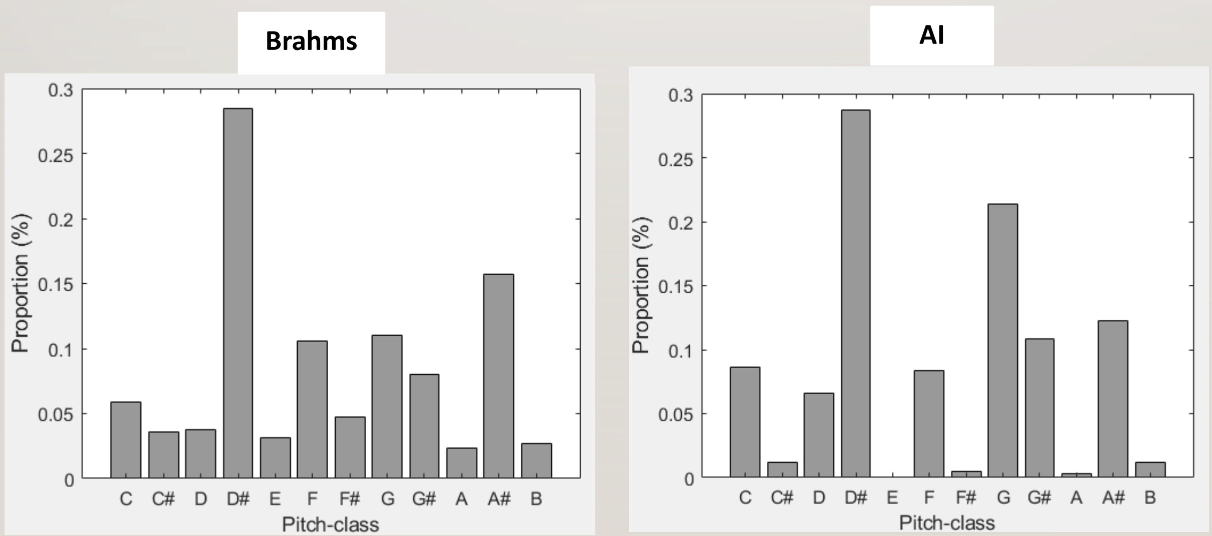

Figure 3: Box Plot of the Pitch Class Distribution between Brahms and AI

RESULTS AND DISCUSSION

Table 2: Quantitative Survey Results

	AI	Probably AI	Unsure	Probably Brahms	Brahms	Total Score
Brahms	3.64%	5.45%	3.64%	36.36%	**50.91%**	**237**
Brahms	12.50%	25.00%	3.57%	**37.50%**	21.43%	**185**
Brahms	0%	23.21%	7.14%	**39.29%**	30.36%	**211**
Brahms	17.86%	**44.64%**	3.57%	28.57%	5.36%	**145**
Brahms	7.27%	7.27%	18.18%	**41.82%**	25.45%	**207**
AI	23.21%	**32.14%**	12.50%	23.21%	8.93%	**147**
AI	10.71%	19.64%	19.64%	**28.57%**	21.43%	**185**
AI	5.36%	**41.07%**	14.29%	32.14%	7.14%	**165**
AI	14.55%	**34.55%**	5.45%	32.73%	12.73%	**165**
AI	7.27%	30.91%	7.27%	**36.36%**	18.18%	**183**

CONTRIBUTION AND IMPACT

- Paper addresses gaps in research
  - Training various models to replicate the works of Brahms
  - Ability to generate longer sequences of music containing complex motifs and harmonic cadences of romantic period piano music
- Quantitative survey conducted using only professional musicians, composers and conductors to get professional evaluation

FUTURE WORK AND RECOMMENDATIONS

- Increasing dataset to orchestral, ensemble and choral works of Brahms could adhere to dataset limitations
- Generate pieces into musical notation for human performance

BIBLIOGRAPHY

- TensorFlow, “Generate music with an RNN,” n.d. [Online]. Available: https://www.tensorflow.org/tutorials/audio/music_generation
- C. Z. A. Huang, A. Vaswani, J. Uszkoreit, N. Shazeer, I. Simon, C. Hawthorne, A. M. Dai, M. D. Hoffman, M. Dinculescu and D. Eck, “Music Transformer: Generating Music with Long-Term Structure,” *Proceedings of the International Conference on Learning Representations*, no. 17, pp. 1-14, 2018. https://doi.org/10.48550/arXiv.1809.04281
- A. K. Huang, P.A. Szerlip, M. E. Norton, T.A. Brindle, Z. Merritt and K.O. Stanley, “Visualizing music self-attention” *In Proc of the Conference on Neural Information Processing Systems*, no.32, pp. 1-5, 2018. https://openreview.net/pdf?id=ryfxVNEajm
- P. Shaw, J. Uszkoreit and A. Vaswani, “Self-Attention with Relative Positional Representations,” *In Proc of the 2018 Conference of the North American Chaper of the Association for Computational Linguistics: Human Language Technologies*, no.2, pp. 464-468, 2018. https://doi.org/10.18653/v1/N18-2074
- Z. Liu, Y. Lin, H. Hu, Y. Wei, Z. Zhang, S. Lin and B. Guo, “Swim Transformer: Hierarchical Vision Transformer using Shifted Windows,” *In Proc of the IEEE/CVF International Conference on Computer Vision*, no.23, pp. 9992-10002, 2021. https://doi.org/10.48550/arXiv.2103.14030

BIBLIOGRAPHY

- A. K. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, "Attention is All you Need," *Advances in Neural Information Processing Systems*, no. 31, pp. 1-15, 2017. https://doi.org/10.48550/arXiv.1706.03762
- A. Jaegle, F. Gimeno, A. Brock, A. Zisserman, O. Vinyals and J. Carreira, "Perceiver: General Perception with Itenative Attention," *In Proc of the International Conference on Machine Learning*, no. 38, pp. 1-43, 2021. https://doi.org/10.48550/arXiv.2103.03206

Thank you!