Using Generative Adversarial Networks (GANs) to Generate Original Music

Guy Coop *
Supervisor: Per Kristian Lehre

February 19, 2018

Abstract

PLACEHOLDER TEXT
Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

*gtc434@student.bham.ac.uk — ID:1447634

Contents

List of Figures

List of Tables

1 Introduction

Generative Adversarial Networks (GANs) were first introduced in 2014 by Goodfellow et al. [GPAM+14] and have since formed a foundation for a new method of unsupervised learning. In which two "opponents" are competing in the form of a zero-sum game to facilitate the training of networks without a specially designed loss function. This method has shown great success in creative tasks, often producing results that are superficially indistinguishable from human produced media.
This paper will attempt to apply this method to the relatively unexplored area of music generation, using a similar method to that of Elgammal et al. [ELEM17] in which the loss function was partially dependant on a classifier network that determined the period of the art. In the same way, the loss function used in this paper will be partially dependent on being unable to classify the genre of the music, this should result in music that appears novel as it is a strict subset of a completely novel genre of music.

2 Background to Theory

2.1 Generative Adversarial Networks

Generative Adversarial Networks [GPAM+14] use two agents competing in the form of a zero-sum game to try and complete creative tasks without the need to design a highly complex loss function. A very generic model can also be used for a number of different applications without needing to be redesigned due to the nature of the system. A simple description of the model is given by the following analogy:
A currency *forger* and a *policeman* take the place of the two competing agents. The *forger's* goal is to produce passable fake currency, and the *policeman's* goal is to differentiate fake and real currency. (For the sake of this example, it is important to pretend that the policeman has had no training, and the forger has never seen any real currency). The *forger* attempts to produce currency and put it into circulation, the *policeman* always finds it with the rest of the normal currency, and he looks at each note and declares it either real or fake. At the end of a round, the both the *policeman* and the *forger* are told how many forged notes were sucessfully detected, from this, the *policeman* becomes better at detecting forged notes, and the *forger* becomes better at forging realistic notes.

2.2 Recurrent Neural Networks

PLACEHOLDER [ZSV14] [SBY17] [LBE15]

2.3 Convolutional Neural Networks

PLACEHOLDER [RMC15]

2.4 Related Work

In the paper by Elgammal et al. [ELEM17] great importance is placed on creating novelty without straying too far from accepted norms, this theory is motivated by the theory of Wilhelm Max Wundt. [Wun74] this theory can be easily shown with the Wundt Curve (Figure 1)

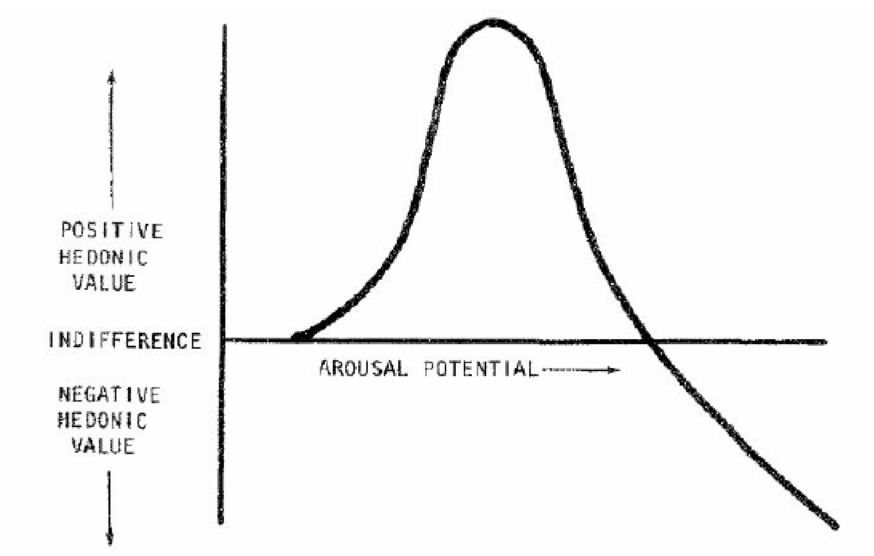

Figure 1: The Wundt curve, used for measuring arousal potential, showing how hedonic response can decrease once novelty increases beyond a certain point

It is important to note that on the Wundt Curve, hedonic value decreases after a certain arousal potential. This implies that if the "creativity" of the media increases beyond a certain point it is no longer reacted to positively and instead becomes too abstract for humans to appreciate.

PLACEHOLDER Midi-Net [YCY17] MUSE-GAN [DHYY17]

3 System Design

4 Implementation of the Project

4.1 Data Preprocessing

Explaination of the pre-processing done on the midi files

4.2 Proof of concept

When implementing a project of this scale, a proof of concept "toy" project is often useful to help verify that some of the subsytems are working without needing to implement the entire system. Details of the design, implementation, and testing of this proof of concept model are given below.

4.2.1 Design

This model uses a GAN, where both the discriminator and the generator are simple feed forward neural networks. The network diagram is given in Figure 2

Figure 2: A system diagram showing the network design for the proof-of-concept model

4.2.2 Implementation

4.2.3 Testing and Experimentation

5 Experimentation

References

[DHYY17] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment. 2017.

[ELEM17] Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. (Iccc):1–22, 2017.

[GPAM+14] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *Nips'2014*, (Arxiv report 1406.2661), 2014.

[LBE15] Zachary C. Lipton, John Berkowitz, and Charles Elkan. A Critical Review of Recurrent Neural Networks for Sequence Learning. pages 1–38, 2015.

[RMC15] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. pages 1–16, 2015.

[SBY17] Baoguang Shi, Xiang Bai, and Cong Yao. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(11):2298–2304, 2017.

[Wun74] Willhelm Max Wundt. *Grundzüge de physiologischen Psychologie*. 1874.

[YCY17] Li-Chia Yang, Szu-Yu Chou, and Yi-Hsuan Yang. MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation. 2017.

[ZSV14] Wojciech Zaremba, Ilya Sutskever, and Oriol Vinyals. Recurrent Neural Network Regularization. (2013):1–8, 2014.