# 智能音乐生成综述

**摘要：**智能音乐生成（下称“音乐生成”）是指在给定格式输入后，计算机根据算法自动生成完整格式乐曲的创作过程。近年来，音乐生成愈发收到研究的重视。区别于文本语言处理，音乐生成对时间序列提出了更高的要求。相关研究也从基因算法、简单的神经网络出发，通过生成模型、转换器模型的一步步探索，去寻求生成更为平滑且具连续新重复结构的音乐。本文梳理了音乐生成算法近年来的经典研究方案，并在陈述过程中反思部分模型存在的掣肘，以明晰未来音乐生成可以应用的相关领域。

**关键词：**音乐生成；神经网络；符号音乐；音乐评价

**1、引言**

近年来，随着短视频、自定义相册等自媒体记录功能的流行与发展，现有的人为作曲音乐难以满足人们日益增长的音乐需求。即使对于相同的生活场景，人们也青睐于使用风格迥异的音乐加以搭配，用以展示自己生活的独特性。因此，音乐生成大放光彩；伴随生成的多风格音乐，音乐定制化的需求也逐步得到满足。

除却搭配图像记录生活的功能之外，音乐生成也活跃在乐曲续作、音乐创作、辅助治疗等多个领域。2021年贝多芬生前未完成的《第十交响曲》在波恩首演[[1]](#footnote-0)；2020东京奥运会开幕式主题曲“2020东京节奏”的奏响[[2]](#footnote-1)，都诠释着音乐生成的强大生命力与广泛的应用场景。

本文对智能音乐生成近年来所提出的模型进行了总结分类，并归纳了相应的发展脉络。其中第二部分阐述了音乐生成研究的基础工作；第三部分归类了的当前音乐生成不同研究方向的现状与成果；第四部分则进行了总结与展望。

**2、研究背景与发展脉络**

音乐的表示形式主要有音频（Audio）与符号（Symbolic）两种，分别对应于连续变量和离散变量。而对于音乐生成的探讨，我们暂时聚焦于符号音乐之中。

对于符号音乐，其主要分为音乐数字接口（Musical Instrument Digital Interface，下称“MIDI”）与琴键记录（下称“Piano Roll”）两种主流音乐记录标椎[[[3]](#endnote-0)]。区别于Piano Roll的二维音调步长记录，MIDI使用事件控制音符的开始与结束、0-127音符编号记录音高，弥补了前者不能区分长音符与重复短音符这一音乐生成关键问题的弊端。因而MIDI格式的音乐更为受到研究者的关注。

区别于早期计算机作曲软件，如拼接工具（PatchWork）和可视化软件（OpenMusic）[[[4]](#endnote-1)]等没有自主创作音乐能力的辅助用具，Zeng等提出用模因算法来满足中国传统音乐的创作需求[[[5]](#endnote-2)]。在满足中国传统音乐中宫、商、角、徵和羽五个基本音的前提下，Zeng等基于选择的节拍，按序确定音乐持续小节数、音调模式及音符序列，再根据确定的模式和韵律创作具有时间价值的音符，并包装成相应的染色体结构。在后续的生成过程中，使用局部最优（local search）算法来优化不同染色体之间的适应度值，适应度值计算式如下：

+

其中为节奏相似度的权重，为音高相似度的权重；、均大于0，且+=1。

在模因算法的交叉过程中，子部分从两个父部分中随机选择一个继承时间值，然后自主填充与可用音符音高相对应的音符。随后后代个体的时间值和音高被随机重组，以创造新的个体。直至所有染色体个体音符的适应度值均满足阈值要求。

模因算法虽然满足了创造性与继承性需求，但在音乐多样性生成方面表现欠缺，因而基于神经网络的音乐生成研究开始把握主流。

借助于卷积神经网络（CNN），Kurniawati等使用了[多层感知机](http://www.baidu.com/link?url=MpQzjDwEVC9M8TphdZZzO7DD-nMvFN-mryc7PM0FiPLYaLskiTEyy8e-fZkJp5y2rxFSMsFhJf1HTRqDouS-a5tyeFWPYI8pJKZbN6YsZle" \t "https://www.baidu.com/_blank)（Multilayer Perceptron，下称“MLP”）来实现音乐分类[[[6]](#endnote-3)]，为音乐生成铺垫。Kurniawati等使用的MLP模型概念如图1。

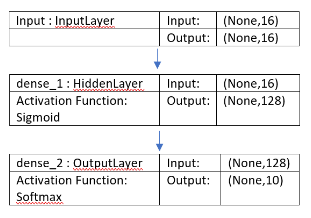
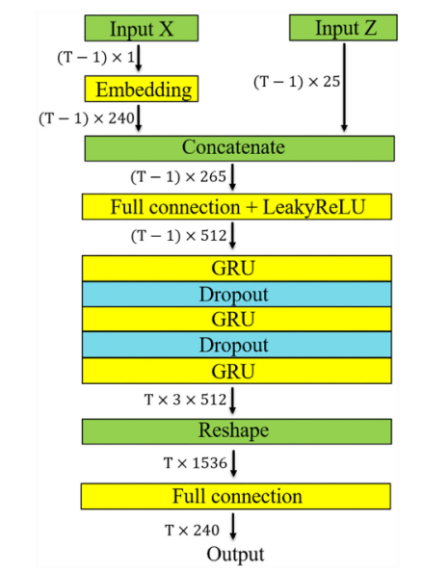


图1 多层感知机模型概念图

它由16个输入层、一个包含128个Sigmoid激活神经元的隐藏层[[[7]](#endnote-4)]和一个包括10个维度的输出层组成，并使用softmax激活函数[[8]](#footnote-2)来预测最终类别。

随后，Zheng等使用门循环单元（Gate Recurrent Unit，下称“GRU”）来实现音乐生成[[[9]](#endnote-5)]创作。在引入音高直方图[[[10]](#endnote-6)]与音符密度[[11]](#footnote-3)的概念后，Zheng等也提出了相应的模型结构，如图2。其中输入X表示掩膜（mask）的选择，而输入Z表示音高直方图[[12]](#footnote-4)和音符密度。

图2 基于门循环单元的音乐生成模型

如果音符事件序列的长度为T，则输入X的大小为（T−1）×1。每个事件通过240×240嵌入层转换为240维向量（embedding过程）。而音高直方图是一个（T−1）×12的矢量，音符密度则被转换为（T−2）×12 热编码（one-hot）矢量。再额外使用（T−1）×1规模的零向量来增加神经网络的稳定性。因此，输入Z的大小为（T−1）×25。接着将音高直方图和音符密度与240维向量连接起来。级联向量的大小为（T−1）×265。级联输入被送到265×512全连接层和线性整流（ReLU）激活函数中。接着，这个（T−1）×512向量被发送到一个三层512单元的GRU中，伴随前两层GRU0.3的随机失活比例。GRU输出则被送到由240个单元组成的线性层。最后输出一个T×240的向量，表示每个时间步长的每个音符事件的概率。

另外，同样基于递归神经网络（Recurrent Neural Network 下称“RNN”），苗等提出了使用乐曲的隐式特征来创作多声部音乐[[[13]](#endnote-7)]。需要特别指出的是，区别于先前面向MIDI音乐格式创作的管理，苗等对Piano Roll格式音乐进行了探究。其主要研究为训练一自编码器用于捕获乐曲的隐式特征，通过这样降维的方式来进行对映的音乐生成。

**3、研究现状**

我们可以看到，在上述的部分中，研究仅仅局限于基因算法或神经网络的原始模型，而缺乏一定的生成模型。一个更为严肃的问题，上述提出的音乐生成方案并不能很好的处理音乐的时间结构，或者说因为受到演化方式线性的约束，较难学习到长期的依赖效果；由于RNN的反馈梯度随着时间呈指数级快速收缩，因而也可能会出现梯度消失现象，影响音乐生成质量[[[14]](#endnote-8)]等问题。从而，研究开始向以下四个方向开始进行扩展。

**3.1 从自然语言处理（Natural Language Processing）汲取经验**

在王等提出的方案中，音符也可以通过一定的方式转换为字（word）一样的格式[[15]](#footnote-5)，从而类似于文本一般进行创作生成[[[16]](#endnote-9)]。在对应仅含有主旋律、钢琴和贝司三个轨道的MIDI文件时，word的构造可以如下：

m{p1,p2,p3...pm}b

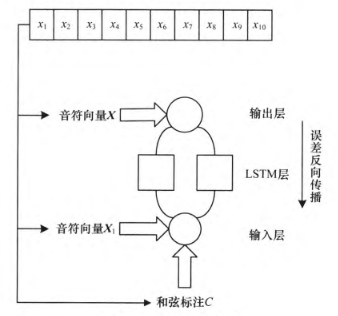
其中m表示在主旋律轨道所对应的编码，b表示贝司对应的编码，p表示钢琴的编码，}为分隔符。因为钢琴在单位时间内可能同时有多个音符的演奏，因此存在p1-pm多个符号，表示从低到高排列的钢琴所有同时奏响的音符。

图3基于LSTM网络训练模型

在此之上，王等也基于长短期记忆网络（Long Short-Term Memory，下称“LSTM”）模型提出了自己的音乐创作模型，如图3。

因为word形式的文本无法直接输入到神经网络，因而我们需要进行one-hot编码，将word转化为对应于神经网络可接受的向量形式，从而进行输入。在图3的模型中，我们将输入层与两个由1024个单元组成的LSTM模型连接，最后连接softmax函数作为输出。

在生成的基础之上，模型也是采用了十二平均律[[[17]](#endnote-10)][[18]](#footnote-6)作为生成音乐是否高质量的检验标椎。十二平均律可以很好的衡量生成音乐与真实音乐在音符使用上的相似性，它的具体数学表现如下：

其中，表示待计算的音符的绝对音高，表示参考音高，n和a分别表示待计算音符和参考音符从左数起的位次。

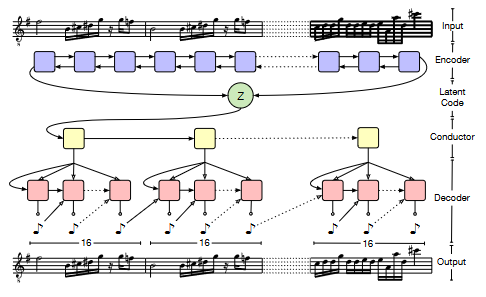
**3.2 创作转向系统化的生成模型**

一方面是使用生成式对抗网络（Generative Adversarial Networks，下称“GAN”）进行音乐生成。Yang等提出了使用批处理规范化（batch normalization）协助收敛的DCGAN生成模型[[19]](#footnote-7)；Karras等提出了动态增加模型训练层数的progress GAN模型[[20]](#footnote-8)等等。在此我们以基于RNN的GAN模型为例分析生成式对抗网络下的音乐生成。

Trieu等率先提出了基于RNN的GAN结构[9]，并提出了从单音符生成出发，对于给定任意和弦，条件反射生成对应构成音的和弦条件反射概念。Trieu等提出的模型几乎等同于GAN一般架构，使用300个节点的单层LSTM。在预训练阶段，Trieu等对生成器（下称“G”）使用最大似然进行序列训练。一旦平均似然误差低于阈值，就转变使用REINFORCE[[[21]](#endnote-11)]进行训练，并通过预测概率和实际概率之间的交叉熵训练判别器（D）。

在此之上，该模型还引入了合格节奏频率、连续音高重复次数、腐烂记忆频率、音高节奏变换等八个评价指标，来衡量音乐生成多样性的维度和生成音乐与真实音乐呈现之间的相似性。

另一方面，google品红团队也提出了基于RNN的音乐变分自动编码器（music Variational autoEncoder下称MusicVAE）模型[[[22]](#endnote-12)][[[23]](#endnote-13)]，如图4。虽然同样基于LSTM模型，但MusciVAE使用的是双向编码器，来计算输入音乐小节之间的双向关系，完成encode操作，更好的考虑乐曲结构间的重复性与连续性。

图4 MusicVAE模型网络架构示意图

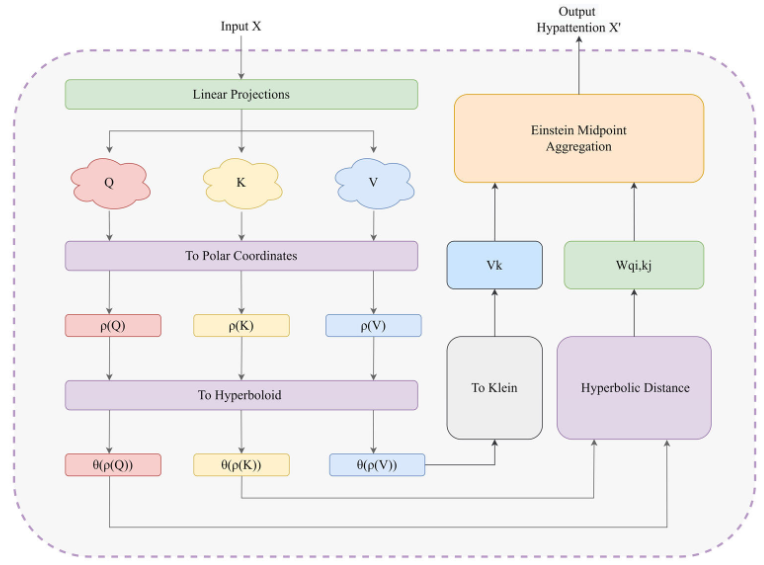
获得的隐藏层输出（latent code） z通过一全连接层与经过tanh函数激活后，可以得到一中间输出状态，称为“conductor RNN”。借助于两层单向LSTM结构，我们可以将conductor RNN embedding为序列向量c。每个向量c分别单独通过一个全连接层，并进行tanh函数激活，以产生最终底层解码器RNN的输入状态。在此之后，解码器RNN通过一个softmax输出层自动回归生成每个子序列的分布序列。至此，我们便产生了一个基于原始输入序列的预测乐曲序列。

**3.3从欧式空间向双曲空间的转变**

前述的所有音乐生成模型，各参量的设置与推导均限制于欧式空间内。而对于transformer模型而言，欧式空间不能提供最具鲁棒性或最有意义的几何表示。因而Huang等人便提出了在双曲空间内使用transformer模型的方法[[[24]](#endnote-14)]。

在这一研究中，Huang等也基于面向观众需求15]创作层次化音乐的指导，提出了新的标记体系REMI。REMI和MIDI表示均有128个音符的off-on事件和音符速度标定。不同之处在于，REMI使用了音符持续时间事件而不是Off-on的简单界定事件。同时，REMI还将和弦信息添加到音乐表示中，添加了共同的和弦根与和弦属性。

在transformer模型基础下，Huang等使用的注意力机制如图5。

图5 双曲空间下transformer模型对应使用的自注意力机制

该方法同样使用 ， （其中，、和分别是Q、K和V转换到注意空间的映射矩阵）分别计算查询、索引与内容矩阵。在双曲空间下，不能直接使用这些结果，需要将输出矩阵转化为极坐标形式后，再使用下式将极坐标形式约束为一个有效的双曲表达式用以计算。

f5ea978000589c8214e4fa01377b137其中，。

而映射函数的有效性可以通过每个点的双曲参数化是否等于-1和计算点与原点之间的距离是否满足下式来判断。

1321d1da61f3d7a7985d63cf90a548b由于双曲空间中不能直接用点积[[[25]](#endnote-15)]来测量相似度，因此我们需要采用双曲距离来测量双曲空间中（查询矩阵Q中获取的查询向量）和（键值矩阵K中获取索引向量）的相似度：

0e48aa92e374a76731dc0e6d1859a5b其中β和偏差是网络本身可以设置或学习的参数。

**3.4 从音乐理解出发走向音乐结构**

对于音乐类型的甄别，Zeng等也是基于百万级别数据集设计了MusicBERT象征性音乐预训练模型[[[26]](#endnote-16)]。对于音乐理解的分析，可以主要包括流派分类、情感分类、音乐片段匹配等任务。MusicBERT的设计中，zeng等也引入了一种较为新颖的音乐记录标椎OctupleMIDI，它将每个音符编码成一个包含8个元素的元组。这8个元素代表了一个音符的特征的不同方面，包括时间签名、节奏、小节、位置、乐器、音高、持续时间和速度。

在基于transformer的编码器中，为了将每个序列步骤中的八元标记转换为transformer编码器的输入，需要将8个元素连接起来，并使用线性层将它们转换为单个向量。然后，将转换后的向量加入相应的位置embedding，作为transformer编码器的输入。为了从transformer编码器中预测8个元组中的8个标记，还需要添加8个不同的softmax层，分别将transformer编码器的隐藏内容映射到8种不同元素类型的词汇表中。这样的训练方式便可以很好的契合基于音乐理解的分类任务。

**4、总结与展望**

本文讲述了音乐生成创作的主要发展历程，从提出的各式各样的模型触发，介绍了基因算法，仅依赖于CNN、RNN等网络的音乐生成；再到近来基于生成模型，基于transformer结构的智能创作等等。但这些模型仅仅是从单纯音乐生成到基于乐理知识音乐生成的转变，而忽略了音乐生成过程中“情感”这一重要因素或仅仅聚焦于用速度来表达不同情感的简单归约操作。而面对现在的需求背景，人们想要的不仅仅是智能生成的音乐，更青睐于带有独特情感标签的音乐生成创作。

近来，音乐治疗也愈发成为讨论热点，也有大量研究证实了音乐疗法在减少老年痴呆症的行为和心理症状方面极具有效性[[[27]](#endnote-17)]，这样的音乐治疗也可以在一定程度上规避使用抗精神病药物而忽视的风险[[[28]](#endnote-18)]。除此之外，带有情绪的音乐也会增强自闭症儿童情绪识别的能力，从而使自闭症儿童更好的处理社会情绪线索，参与治疗[[[29]](#endnote-19)]。

**参考文献：**

1. M. Civit, J. Civit-Masot, F. Cuadrado, and M. J. Escalona, “A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends,” Expert Systems with Applications, vol. 209, p. 118190, Dec. 2022, doi: [10.1016/j.eswa.2022.118190](https://doi.org/10.1016/j.eswa.2022.118190).
2. Assayag, G., Rueda, C., Laurson, M., Agon, C., and Delerue, O. Computer-Assisted Composition at IRCAM: From PatchWork to OpenMusic. Comput. Music J. 23, 3 (Sep. 1999), 59-72.
3. Z. Zeng and L. Zhou, “A Memetic Algorithm for Chinese Traditional Music Composition,” in 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi’an, China: IEEE, Apr. 2021, pp. 187–192. doi: [10.1109/ICSP51882.2021.9408813](https://doi.org/10.1109/ICSP51882.2021.9408813).
4. A. Kurniawati, Y. K. Suprapto, and E. M. Yuniarno, “Multilayer Perceptron for Symbolic Indonesian Music Generation,” in 2020 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia: IEEE, Jul. 2020, pp. 228–233. doi: [10.1109/ISITIA49792.2020.9163723](https://doi.org/10.1109/ISITIA49792.2020.9163723).
5. J. L. D. Antich, ”Audio Event Classification using Deep Learning in an End-to-End Approach,” Doctoral dissertation, MA thesis. Copenhagen, Denrmak: Department of Architecture, Design, and Media Technology, Aalborg University Copenhagen, 2017.
6. K. Zheng et al., “EmotionBox: A music-element-driven emotional music generation system based on music psychology,” Front. Psychol., vol. 13, p. 841926, Aug. 2022, doi: [10.3389/fpsyg.2022.841926](https://doi.org/10.3389/fpsyg.2022.841926).
7. Tzanetakis, G., Ermolinskyi, A., and Cook, P. (2003). Pitch histograms in audio and symbolic music information retrieval. Int. J. Phytoremediation. 21, 143–152. doi: 10.1076/jnmr.32.2.143.

16743.

1. 苗北辰，郭为安，汪镭，等．隐式特征和循环神经网络的 多声部音乐生成系统[J]．智能系统学报，2019, 14(1): 158-164.
2. N. Trieu and R. Keller, “JazzGAN: Improvising with Generative Adversarial Networks”.
3. 王思源, 周建国. 基于CharRNN的复音音乐生成方法[J]. 计算机工程, 2019, 45(5): 249-255, 260.
4. KLERK D D. Equal tempemment[J]. Acta Musicologica, 1979, 51(1): 140 - 150．
5. Williams, R. J. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. In Reinforcement Learning. Springer. 5–32.

A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck, “A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music”.

1. W. Huang, Y. Yu, H. Xu, Z. Su, and Y. Wu, “Hyperbolic Music Transformer for Structured Music Generation,” IEEE Access, vol. 11, pp. 26893–26905, 2023, doi: [10.1109/ACCESS.2023.3257381](https://doi.org/10.1109/ACCESS.2023.3257381).
2. D. Deutsch and J. Feroe, “The internal representation of pitch sequences in tonal music,” Psychol. Rev., vol. 88, no. 6, pp. 503–522, Nov. 1981.
3. W. Peng, T. Varanka, A. Mostafa, H. Shi, and G. Zhao, “Hyperbolic deep neural networks: A survey,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 12, pp. 10023–10044, Dec. 2022.
4. M. Zeng, X. Tan, R. Wang, Z. Ju, T. Qin, and T.-Y. Liu, “MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training,” in Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Online: Association for Computational Linguistics, 2021, pp. 791–800. doi: [10.18653/v1/2021.findings-acl.70](https://doi.org/10.18653/v1/2021.findings-acl.70).
5. Patel, B., Perera, M., Pendleton, J., Richman, A., & Majumdar, B. (2014). Psychosocial interventions for dementia: From evidence to practice. Advances in Psychiatric Treatment, 20, 340–349. DOI:10.1192/apt.bp.113.011957.

Blackburn, R., & Bradshaw, T. (2014). Music therapy for service users with dementia: A critical review of the literature. Journal of Psychiatric and Mental Health Nursing, 21(10), 879–888. DOI:10.1111/ jpm.12165.

G. L. Wagener, M. Berning, A. P. Costa, G. Steffgen, and A. Melzer, “Effects of Emotional Music on Facial Emotion Recognition in Children with Autism Spectrum Disorder (ASD),” J Autism Dev Disord, vol. 51, no. 9, pp. 3256–3265, Sep. 2021, doi: [10.1007/s10803-020-04781-0](https://doi.org/10.1007/s10803-020-04781-0).

1. 来源：<https://www.udiscovermusic.com/classical-news/beethovens-10th-symphony-ai/>。 [↑](#footnote-ref-0)
2. 来源：<https://youtu.be/smMVQ6C4Wqg>。 [↑](#footnote-ref-1)
3. [] M. Civit, J. Civit-Masot, F. Cuadrado, and M. J. Escalona, “A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends,” Expert Systems with Applications, vol. 209, p. 118190, Dec. 2022, doi: [10.1016/j.eswa.2022.118190](https://doi.org/10.1016/j.eswa.2022.118190). [↑](#endnote-ref-0)
4. [] Assayag, G., Rueda, C., Laurson, M., Agon, C., and Delerue, O. Computer-Assisted Composition at IRCAM: From PatchWork to OpenMusic. Comput. Music J. 23, 3 (Sep. 1999), 59-72. [↑](#endnote-ref-1)
5. [] Z. Zeng and L. Zhou, “A Memetic Algorithm for Chinese Traditional Music Composition,” in 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi’an, China: IEEE, Apr. 2021, pp. 187–192. doi: [10.1109/ICSP51882.2021.9408813](https://doi.org/10.1109/ICSP51882.2021.9408813). [↑](#endnote-ref-2)
6. [] A. Kurniawati, Y. K. Suprapto, and E. M. Yuniarno, “Multilayer Perceptron for Symbolic Indonesian Music Generation,” in 2020 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia: IEEE, Jul. 2020, pp. 228–233. doi: [10.1109/ISITIA49792.2020.9163723](https://doi.org/10.1109/ISITIA49792.2020.9163723). [↑](#endnote-ref-3)
7. [] J. L. D. Antich, ”Audio Event Classification using Deep Learning in an End-to-End Approach,” Doctoral dissertation, MA thesis. Copenhagen, Denrmak: Department of Architecture, Design, and Media Technology, Aalborg University Copenhagen, 2017. [↑](#endnote-ref-4)
8. Softmax函数细节详见： <https://keras.io>.。 [↑](#footnote-ref-2)
9. [] K. Zheng et al., “EmotionBox: A music-element-driven emotional music generation system based on music psychology,” Front. Psychol., vol. 13, p. 841926, Aug. 2022, doi: [10.3389/fpsyg.2022.841926](https://doi.org/10.3389/fpsyg.2022.841926). [↑](#endnote-ref-5)
10. [] Tzanetakis, G., Ermolinskyi, A., and Cook, P. (2003). Pitch histograms in audio and symbolic music information retrieval. Int. J. Phytoremediation. 21, 143–152. doi: 10.1076/jnmr.32.2.143.

    16743. [↑](#endnote-ref-6)
11. 音乐密度：单位时间窗内音符个数。 [↑](#footnote-ref-3)
12. 音乐直方图：由12个整数值组成的数组，以半音音阶的12个半音为索引，显示了音乐作品中每个半音的出现频率。 [↑](#footnote-ref-4)
13. [] 苗北辰，郭为安，汪镭，等．隐式特征和循环神经网络的 多声部音乐生成系统[J]．智能系统学报，2019, 14(1): 158-164. [↑](#endnote-ref-7)
14. [] N. Trieu and R. Keller, “JazzGAN: Improvising with Generative Adversarial Networks”. [↑](#endnote-ref-8)
15. 来源：https://arxiv.org/pdf/1308.0850.pdf。 [↑](#footnote-ref-5)
16. [] 王思源, 周建国. 基于CharRNN的复音音乐生成方法[J]. 计算机工程, 2019, 45(5): 249-255, 260. [↑](#endnote-ref-9)
17. [] KLERK D D. Equal tempemment[J]. Acta Musicologica, 1979, 51(1): 140 - 150． [↑](#endnote-ref-10)
18. 十二平均律：物理频率2:1之间的单音被分化为12个半音，以此称为一个八度。 [↑](#footnote-ref-6)
19. 来源：<http://arxiv.org/abs/1703.10847> [↑](#footnote-ref-7)
20. 来源：<http://arxiv.org/abs/1710.10196> [↑](#footnote-ref-8)
21. [] Williams, R. J. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. In Reinforcement Learning. Springer. 5–32. [↑](#endnote-ref-11)
22. [] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck, “A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music”. [↑](#endnote-ref-12)
23. [] [↑](#endnote-ref-13)
24. [] [↑](#endnote-ref-14)
25. [] W. Peng, T. Varanka, A. Mostafa, H. Shi, and G. Zhao, “Hyperbolic deep neural networks: A survey,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 12, pp. 10023–10044, Dec. 2022. [↑](#endnote-ref-15)
26. [] M. Zeng, X. Tan, R. Wang, Z. Ju, T. Qin, and T.-Y. Liu, “MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training,” in Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Online: Association for Computational Linguistics, 2021, pp. 791–800. doi: [10.18653/v1/2021.findings-acl.70](https://doi.org/10.18653/v1/2021.findings-acl.70). [↑](#endnote-ref-16)
27. [] Patel, B., Perera, M., Pendleton, J., Richman, A., & Majumdar, B. (2014). Psychosocial interventions for dementia: From evidence to practice. Advances in Psychiatric Treatment, 20, 340–349. DOI:10.1192/apt.bp.113.011957. [↑](#endnote-ref-17)
28. [] Blackburn, R., & Bradshaw, T. (2014). Music therapy for service users with dementia: A critical review of the literature. Journal of Psychiatric and Mental Health Nursing, 21(10), 879–888. DOI:10.1111/ jpm.12165. [↑](#endnote-ref-18)
29. [] G. L. Wagener, M. Berning, A. P. Costa, G. Steffgen, and A. Melzer, “Effects of Emotional Music on Facial Emotion Recognition in Children with Autism Spectrum Disorder (ASD),” J Autism Dev Disord, vol. 51, no. 9, pp. 3256–3265, Sep. 2021, doi: [10.1007/s10803-020-04781-0](https://doi.org/10.1007/s10803-020-04781-0). [↑](#endnote-ref-19)