DOI 10.20504/opus2026.32.53
Dossiê Música e Inteligência Artificial: inovações, desafios e perspectivas

Arte sonora e síntese neural: uma investigação sobre três estratégias de criatividade assistida pela inteligência artificial na música atual

Matheus Prado **Garcia**
https://orcid.org/0009-0003-3595-4779
Universidade Estadual de Campinas

Manuel Silveira **Falleiros**
https://orcid.org/0000-0001-8553-0921
Universidade Estadual de Campinas

Resumo: Este artigo faz uma análise comparativa de três casos do uso de aprendizado de máquina e síntese neural no processo de criação musical. Os casos discutidos são os seguintes: a) a peça AI Deviations, de Yasunao Tone, na qual o compositor utiliza em performance um modelo de si mesmo, treinado sobre performances passadas; b) a compositora Holly Herndon, que desenvolveu um modelo de síntese vocal a partir de um trabalho coletivo com seu grupo coral PROTO; e c) a dupla de compositores Dadabots, que utiliza aprendizado de máquina em colaboração com diversos artistas para produzir álbuns e livestreams de geração contínua de música. Esses trabalhos foram comparados a partir i) da forma como a disfunção das redes neurais é trabalhada sonoramente através do treinamento sobre small data; ii) do diferente nível de controle que cada compositor exerce sobre seu modelo em termos de hierarquia e manipulação do produto final; e iii) de como cada compositor busca uma sonoridade inovadora, insatisfeito com uma mera reciclagem do passado fonográfico que é encorajada no contexto da inteligência artificial (IA). Conclui-se que a IA como tecnologia tem um papel importante na experimentação musical desses compositores da contemporaneidade, que traçam um diálogo com a música de sonoridades do século XX.

Palavras-chave: Inteligência artificial. Síntese sonora. Música generativa. Música de sonoridades.

Sound art and neural synthesis: an investigation on three creativity strategies assisted by artificial intelligence in today's music.

Abstract: This paper makes a comparative analysis of three cases of the use of machine learning and neural synthesis in the process of musical creation. The cases discussed are as follows: a) the piece AI Deviations, by Yasunao Tone, in which the composer uses a model of himself in performance, trained on his past performances; b) the composer Holly Herndon, who developed a vocal synthesis model based on the collective work with her choral group PROTO; and c) the composer duo Dadabots, who use machine learning in collaboration with various artists to produce albums and livestreams of continuous music generation. These works were compared on the basis i) of how the dysfunction of neural networks is worked out sonically through training on small data; ii) the different level of control that each composer has over their model in terms of hierarchy and manipulation of the final product; and iii) how each composer seeks an innovative sound, dissatisfied with a mere recycling of the phonographic past that is encouraged in the context of artificial intelligence (AI). The conclusion is that AI as a technology plays an important role in the musical experimentation of these contemporary composers, who engage in a dialog with the sonorous music of the 20th century.

Keywords: Artificial intelligence. Neural synthesis. Generative music. Music of sound masses.

GARCIA, Matheus Prado; FALLEIROS, Manuel Silveira. Arte sonora e síntese neural: uma investigação sobre três estratégias de criatividade assistida pela inteligência artificial na música atual. Opus (Assoc. Nac. Pesqui. Pós-Grad. Música), João Pessoa, v. 32, e263253, p. 1-18, 2026.
http://dx.doi.org/10.20504/opus2026.32.53
Recebido: 30/11/2024; aprovado: 27/03/2026; publicado: 13/11/2026.
Editor responsável: Dr. Vinícius Eufrásio

1. Introdução

Melanie Mitchell (2021) refere-se a "primaveras" e "invernos" da inteligência artificial (IA). O debate e o desenvolvimento dessa tecnologia oscilam, desde a década de 1950, entre períodos de efervescência e rápido avanço técnico, seguidos de uma certa estagnação na medida em que se esgotam as possibilidades específicas colocadas por uma determinada inovação científica, e o seu desenvolvimento volta a caminhar em ritmo lento. Na década de 1950, tivemos como ponto de partida o surgimento das primeiras redes neurais 1 . Até certo ponto esse período corresponde, na música, aos primeiros experimentos de música generativa com Hiller e Isaacson (1993/1957), creditados com a primeira composição realizada por meio de um programa computacional. Na década de 1980, tivemos uma segunda onda da IA com o advento das camadas ocultas de neurônios e retropropagação2. Vimos surgir nessa época diversos modelos e sistemas especialistas de geração automática de música que inauguraram a utilização de redes neurais em sistemas de geração musical simbólica com protocolo MIDI (BRIOT, 2021).

Em 2012, um algoritmo de rede neural convolucional (CNN)3 chamado AlexNet obteve um resultado inédito, muito superior aos seus concorrentes, no campeonato de sistemas de visão de máquina organizado pelo projeto ImageNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2017). Inaugurando o uso de unidades de processamento gráfico na fase de treinamento, esse *software* pavimentou parte do caminho para que as redes neurais e o aprendizado4 profundo de máquina tivessem hoje relevância nas diversas áreas do desenvolvimento de sistemas computacionais — incluindo aqueles voltados à música e às artes de modo geral. A partir desse ano, demonstrado o potencial das redes neurais como arquitetura para sistemas de visão de máquina, vários setores da indústria da computação passaram a prestar atenção e investir nessa tecnologia. Como forma de *rebranding* e estratégia de marketing, o termo IA passou a ser usado como sinônimo de aprendizado profundo de máquina com redes neurais (WHITTAKER, 2021).

Não demorou para que fosse possível inverter a ordem de fluxo de processamento do software AlexNet. Ao invés de uma entrada na forma de imagem e uma saída textual, uma instrução textual pode

1 Redes neurais são sistemas de processamento formados por neurônios artificiais, que se conectam entre si através de sinapses que variam em intensidade. A cada conexão é atribuída a variável peso (*w*), e nesses valores está codificada de forma distribuída boa parte da informação que forma o modelo de IA. Essa tecnologia está intimamente ligada a descobertas sobre a dinâmica dos neurônios que ocorreram na década de 1940: o modelo neural proposto por McCulloch e Pitts (1943) é ainda hoje presente nas redes neurais modernas. Para uma introdução bastante completa de conceitos de IA, recomenda-se Mitchell (2019) "*Artificial intelligence: a guide for thinking humans*".

2 Retropropagação (ou *backpropagation*) é um método usado no treinamento de redes neurais para ajustar os pesos de cada par de neurônios, reduzindo a diferença entre a saída desejada e a saída real. Ela funciona propagando esse erro de volta pela rede, do fim para o início, e atualizando os pesos com base em uma função que recompensa o acerto e pune o erro. Essa técnica possibilitou a implementação de diversas camadas intermediárias (ocultas) de neurônios, uma vez que antes era possível haver apenas uma camada de entrada (*input*) e outra de saída (*output*). Utiliza-se o termo "profunda" (como em *deep neural networks*) para se referir a redes neurais com mais de duas camadas — regra para os grandes modelos atuais.

3 A técnica de convolução é usada em redes neurais para extrair características locais de dados, como imagens, por meio de filtros que percorrem a *input* e destacam padrões como bordas, texturas e formas.

4 Aprendizado de máquina (*machine learning*) não é sinônimo de redes neurais ou IA, mas está quase sempre vinculado a essas duas coisas. Pode-se dizer que redes neurais são uma das técnicas de se realizar o aprendizado de máquina, ou seja, o processo em que um sistema organiza sua própria estrutura a partir de leis internas de processamento de uma base de dados. Essas leis internas e a base de dados são previamente elaboradas, mas o decorrer do "aprendizado" ocorre de forma mais ou menos autônoma. Tudo isso é compreendido pelo termo mais amplo *inteligência artificial*.

ser usada para gerar uma imagem a partir de um banco de dados como ImageNet, que consiste em milhões de arquivos de imagem classificados em 20.000 categorias taxonômicas. Em 2016, despontaram artigos relevantes tanto na área de modelos geradores (generativos) de imagem (REED et al., 2016) como de áudio (OORD et al., 2016), além de avanços no modelamento de linguagem (VASWANI et al., 2023) que anteciparam o *boom* recente de modelos generativos (CHEN; HUANG; GOU, 2024; GOZALO-BRIZUELA; GARRIDO-MERCHAN, 2023).

A utilização de sistemas automatizados e modelos estatísticos no fazer artístico não é nenhuma novidade no campo da composição musical. No entanto, apenas com o advento e disseminação ampla de informações proporcionada pela internet foi possível formar as bases de dados com milhares de horas de áudio classificados textualmente, além de avanços nos modelos de linguagem natural[5] que permitiram chegar à atual fase de modelos generativos de tipo texto-para-áudio (CHEN; HUANG; GOU, 2024; MITCHELL, 2021).

Nas primaveras de IA das décadas de 1950 e 1980, os programas de IA voltados para música se deram, quase exclusivamente, sobre representação simbólica, e o interesse estava na organização das notas a partir de formalizações musicais baseadas na teoria ocidental da música de concerto, transitando entre o clássico e contemporâneo. Mais recentemente, no entanto, apresentou-se a possibilidade de experimentações com redes neurais como ferramenta de síntese sonora, trazendo à ordem do dia a discussão sobre a música de sonoridades do século XX.

Este artigo tem o objetivo de investigar três experiências das artes sonoras dos últimos dez anos que dialogam com os debates mencionados sobre IA. São performances e composições que utilizam modelos generativos de aprendizado profundo de máquina, ou que fazem um comentário dessa tecnologia de alguma maneira. Interessa-nos aqui, principalmente, trabalhos artísticos que buscam ou têm como efeito elucidar e complexificar a compreensão de aspectos da IA. Compreendemos que esse é um debate que passa por uma fase de efervescência, potencializada por uma intensa propaganda mercadológica por parte de empresas que oferecem essa tecnologia. Neste trabalho, argumentamos que é característica desse discurso particular um engrandecimento do que são as reais capacidades dessa tecnologia, associada a um alienamento técnico (DA SILVEIRA, 2021) na medida em que apresenta a IA como algo que "nem seus próprios desenvolvedores conseguem afirmar como funciona"[6] (HASSENFELD, 2023). Logo, é necessário olhar para os trabalhos artísticos e teóricos com a intenção de dissipar essa névoa, ou distanciar-se do frenesi que se formou em torno da novidade dos modelos de aprendizado profundo.

Nosso recorte crítico se concentra em três artistas, dos quais analisaremos trabalhos representativos. Elegemos Yasunao Tone, especialmente pelo seu duradouro trabalho sobre a sonoridade das mídias digitais e a relevância da sua participação e dos grupos que compôs para a música experimental na segunda metade do século XX. Além disso, tanto Holly Herndon como o duo Dadabots são artistas

5 Chama-se de modelos de linguagem natural (*natural language models*) sistemas de IA capazes de processar linguagem no nível de abstração da escrita humana. Ou seja, traduzir um texto de um idioma a outro, responder a enunciados textuais, etc. Grandes modelos de linguagem (*large language models*, LLMs) são capazes de processar não apenas a linguagem, mas uma quantidade imensa de informação multimodal e responder a uma gama muito ampla de problemas em forma de texto. Um exemplo é o ChatGPT <https://chatgpt.com/>. Acesso em: 27 de mar. de 25.
6 Tradução nossa: "*Even the scientist who build AI can't tell you how it works*".

proponentes do uso da IA na música, tendo realizado não apenas uma extensa e relevante produção musical em colaboração com diversos artistas, como publicações científicas, palestras e entrevistas sobre o tema em várias ocasiões.

No domínio da imagem, é possível afirmar que a IA conseguiu avançar em um ritmo muito maior do que no som, em parte pela disponibilidade de bases de dados suficientemente grandes, e também pelo problema comparativamente complexo de se realizar uma taxonomia textual de sons a partir de consensos objetivos (AGOSTINELLI et al., 2023). As reflexões propostas por Grba (2021, 2022) e Zylinska (2020) fazem um *framework* crítico em que comentam experiências com IA no campo das artes visuais, articulando esteticamente esses trabalhos com debates gerais das artes e da contemporaneidade e, portanto, referenciais de forma panorâmica. Na música, especificamente, temos discussões estéticas em Fisher (2013), ainda que não discuta a IA em si, e Rubinstein (2020), além de panoramas históricos mais amplos sobre a IA na música em Briot (2021) e Mantaras e Arcos (2002). O trabalho de Lewis (2000) se mostra relevante no que diz respeito à criatividade de máquina e a uma discussão estética do aprendizado de máquina na performance. A partir destas referências fundamentais, pretendemos articular um diálogo em torno do tema proposto neste artigo.

2. Desenvolvimento

Nesta seção, apresentaremos três estratégias de composição e performance musical distintas na sua forma, mas que similarmente apresentam modos de aprendizado de máquina como conteúdo e processo. Esses trabalhos artísticos serão observados a partir da técnica empregada no uso da IA; pelo seu resultado sonoro e estético; e pela maneira como são capazes de fomentar uma reflexão sobre as relações do fazer musical e o aprendizado de máquina.

Primeiramente nos voltaremos à criação de Yasunao Tone, intitulada *AI Deviations*[7]. Trata-se de uma composição, assim como uma performance, em que esse artista interage com uma versão artificial de sua identidade sonora — um modelo de redes neurais treinado sobre um trabalho anterior seu, *MP3 Deviations*[8]. Isso se dá a partir de um histórico de exploração artística de mídias desde a época em que Tone integrava o grupo Fluxus[9], na década de 1960, que colocou o compositor em contato com as tecnologias mais novas de cada época.

Em seguida, apresentaremos como o trabalho de Holly Herndon[10] está marcado pela IA como uma ferramenta da produção musical contemporânea que coloca ao artista possibilidades sonoras inéditas, e que ela busca incorporar em uma estética preocupada com o futuro da sonoridade musical, especificamente representada em seu álbum PROTO[11].

E, por fim, o duo Dadabots[12] faz uso experimental de redes neurais e aprendizado sobre

7 Disponível em: <https://yasunaotone.bandcamp.com/album/ai-deviatoin-1-2>. Acesso em: 27 de mar. de 25.
8 Disponível em: <https://yasunaotone.bandcamp.com/album/mp3-deviation-8>. Acesso em: 27 de mar. de 25.
9 Fluxus foi um grupo internacional de arte vanguardista, fundado em 1960. Seu trabalho partia fundamentalmente de uma contestação e debate da própria definição de arte. Dentre seus membros também estão Yoko Ono e Nam June Paik. Fonte: <https://www.tate.org.uk/art/art-terms/f/fluxus>. Acesso em: 27 de mar. de 25.
10 <https://herndondryhurst.studio/>. Acesso em: 27 de mar. de 25.
11 Disponível em: <https://hollyherndon.bandcamp.com/album/proto>. Acesso em: 27 de mar. de 25.
12 <https://dadabots.com/>. Acesso em: 27 de mar. de 25.

small-data em programas de geração contínua de música em diferentes estilos, que é transmitida pelo YouTube©, além de produções de álbuns completos a partir de colaborações com diversos artistas.

2.1. A metalinguagem de Yasunao Tone

Yasunao Tone é um artista multidisciplinar nascido em Tóquio, e extensivamente atuante em Nova Iorque. No Japão, no final da década de 1950, fez parte do grupo Ongaku. Com esse grupo, o literato de formação fazia performances musicais com qualquer objeto com o qual conseguisse produzir som, em peças de improvisação que anteciparam muitas das ideias que músicos do ocidente como John Cage levaram em suas visitas ao Japão (KANEDA; TONE, 2013). No início da década de 1970, Tone mudou-se para Nova Iorque, quando passou a integrar o grupo Fluxus. A partir da década de 1990, Tone passou a interessar-se pela sonoridade das tecnologias digitais. A sua peça *Solo for wounded CD* (1997) foi uma das principais referências do movimento da música *glitch* e da música eletrônica experimental de modo geral, que se preocupou com inverter o papel da mídia e do fonograma: o conteúdo é que passa a ser veículo para o som da mídia, e não o contrário. Em *Solo for wounded CD*, Tone faz uso de um Compact Disk (CD) "ferido", danificado e manipulado de diferentes formas para extrair a sonoridade do formato de mídia. Neste processo criativo, o paradigma da transparência e fidelidade de áudio, comercialmente celebrados na época do advento do CD, é radicalmente questionado (GARCIA, 2024).

Em 2016, Tone apresentou a composição-performance *AI Deviations*. Para essa peça, Tone desenvolveu em colaboração com pesquisadores da Universidade de Surrey[13] um modelo de redes neurais que foi treinado com informações de sua própria produção artística — especificamente, a partir da peça *MP3 Deviations*, desenvolvida com alguns dos mesmos pesquisadores (BLAKE et al., 2010). Nesta performance, Tone senta-se à frente de um laptop, com o qual realiza toda a peça. Ainda que, até o momento da escrita deste artigo, não disponhamos de nenhuma publicação ou informações mais detalhadas por parte dos pesquisadores colaboradores sobre o desenvolvimento envolvido na peça *AI Deviations*, podemos recolher pistas importantes a partir de uma publicação sobre o trabalho anterior *MP3 Deviations* — além de alguns registros em vídeo de performances.

Assim, em *MP3 Deviations*, o músico faz uma performance baseada na corrupção de dados em formato MP3, por meio de processamentos que o artista controla, em tempo real, por meio de um *software*. A intenção é tornar audível a codificação e decodificação do formato MP3, em detrimento do seu conteúdo sonoro original (BLAKE et al., 2010). O *software* foi desenvolvido em estreita comunicação com Tone, que tomou decisões quanto aos processos de manipulação dos dados, assim como a interface e os controles utilizados durante a performance.

Podemos deduzir que *AI Deviations* foi provavelmente realizada de forma similar. Podemos notar a partir de algumas performances, nas quais o artista vira o laptop para que o público veja a tela, a presença de uma interface gráfica similar àquela descrita por Blake et al. (2010). Trata-se de uma interface relativamente simples, com botões de seleção e gráficos de visualização de dados, nomeados de forma não

13 <https://www.surrey.ac.uk/>. Acesso em: 27 de mar. de 25.

muito esclarecedora. Ainda segundo Blake, para *MP3 Deviations*, Tone preferiu "aquelas [partes da interface] com uma função pouco clara que oferecem alguma resistência ao entendimento"[14] (BLAKE et al., [s.d.], p. 237). Podemos considerar o termo *Deviations* como um conjunto de peças para uma série contínua de interesse de Tone, pela sonoridade das mídias e tecnologias digitais, desenvolvidas, pelo menos, desde a década de 1990. Por sua vez, representa a dedicação de Tone à performance, improvisação e indeterminação desde a década de 1950 quando começou no Japão sua carreira na música (KANEDA; TONE, 2013). O termo "deviation" (desvio) explicita bem seu processo e intenções criativas, ou seja, não se trata exatamente de levar a tecnologia ao erro absoluto e destrutivo, mas antes ampliar o domínio da tecnologia ao desviar o funcionamento do roteiro estabelecido (MARGASAK, 2017), e assim, expandir artisticamente as possibilidades de uma ferramenta.

A forma como Tone fornece a um programa dados de seu próprio fazer musical, para então realizar uma performance com um modelo de sua arte por meio do aprendizado de máquina, é parecida, em alguns aspectos, com a perspectiva da *reinjeção estilística* (ASSAYAG; BLOCH; CHEMILLIER, 2006). Aqui, a atenção não está em uma criatividade da máquina por si só, mas sim na interação do músico com o programa em um sistema retroalimentado. O projeto XYZ que desenvolveu a aplicação denominada OMax[15], desenvolvido no IRCAM, aplica aprendizado de máquina não supervisionado em conjunto com técnicas de processamento de sinal e representação musical simbólica para identificar e agrupar padrões dentro de um *corpus* musical. Esse *corpus* pode ser previamente fornecido, alimentado de forma contínua ou processado em tempo real durante uma performance. Em ambos os contextos, o sistema elabora um modelo estilístico do compositor ou improvisador usuário, organizando e "abstraindo" o material musical, que é então reinjetado na performance. O improvisador responde musicalmente a esse material "abstraído", cuja reação é capturada pelo algoritmo e incorporada novamente ao corpus, e assim sucessivamente. Dessa forma, estabelece-se uma dinâmica de retroalimentação, na qual músico e máquina interagem e transformam continuamente o material sonoro (ASSAYAG; BLOCH; CHEMILLIER, 2006).

Em ambos os contextos, podemos observar que de forma geral um modelo de aprendizado de máquina exerce a função de um elemento não-determinístico no fluxo da produção sonora, em que não se exerce um controle estrito sobre todos os comportamentos da máquina. Na reinjeção estilística, o modelo pode ser construído a partir do fornecimento de material musical determinado, cuja amplitude timbrística, dinâmica, das alturas etc. pode ser restrita ou ampla de acordo com o interesse do usuário em construir um modelo mais ou menos previsível. No caso de *AI Deviations*, Yasunao Tone dispõe de 6 modelos — presumidamente distintos desde o treinamento — e os manipula a partir de controles propositalmente "pouco intuitivos": "Ele afirmou claramente que não queria uma relação direta demais entre suas interações e o som resultante"[16] (BLAKE et al., 2010).

AI Deviations é, sobretudo, uma peça metalinguística. Tone cria um modelo de sua própria

14 Tradução nossa: "*those with an unclear function offering some kind of resistance to understanding*".
15 OMax é um ambiente de software que utiliza aprendizado de máquina para reproduzir em tempo real o estilo de um músico, e então interagir com ele de forma interativa, desenvolvido principalmente para performances de improvisação. Fonte: <https://forum.ircam.fr/projects/detail/omax-5/>. Acesso em: 27 de mar. de 25.
16 Tradução nossa: "*He explicitly stated that he did not want a too direct a relationship between his interactions and the resulting sound.*"

performance, que é encapsulado em um *software* / *hardware* que ele próprio pode manipular, levar a certos caminhos, mas constantemente desviando (*deviating*) das próprias escolhas passadas. Yasunao Tone faz uso da IA a partir de um *corpus* de treinamento totalmente próprio. O objeto da performance é também o sujeito. De certa forma, isso possibilita uma performance "despreocupada" em sua intencionalidade sonora, em que algumas das questões prementes sobre a IA são removidas do caminho — como direitos autorais e ética de dados —, e outras são colocadas e quase imediatamente respondidas. Mais do que um estudo sobre a sonoridade de ferramentas e mídias de áudio, nos parece que a IA é uma forma de se criar uma *mise en abyme* autorreflexiva — é uma performance sobre a própria performance (DÄLLENBACH, 1998).

2.2. Holly Herndon e o futuro da sonoridade

Holly Herndon advoga claramente a favor do uso da IA como uma expansão dos domínios e possibilidades do fazer musical. Em seus trabalhos artísticos, a cantora e compositora norte-americana demonstra distanciar-se tanto do ceticismo conservador acerca das novas ferramentas tecnológicas, como do discurso exagerado que infla as capacidades dessa tecnologia. Seu trabalho faz o esforço de demonstrar as novas possibilidades que se colocam à frente dos músicos na contemporaneidade no que diz respeito à sonoridade e o fazer musical enquanto prática coletiva e a performance em suas diferentes formas.

Sua aplicação de IA, denominada *Spawn,* foi desenvolvida a partir de uma série de colaborações: seu parceiro criativo Mathew Dryhurst, o desenvolvedor Jules LaPlace, e seu grupo musical denominado PROTO. Neste trabalho, Herndon tem como objetivo o desenvolvimento das sonoridades particulares que a tecnologia de aprendizado de máquina coloca como possibilidade, em um processo que ela entende como uma continuação da exploração de sonoridades que se deu no século XX (HERNDON, 2019). Além disso, com *Spawn*, ela se propõe a fazer uma demonstração franca da IA, sem exagerar suas capacidades, mas apontando com um certo otimismo os caminhos que se colocam para o músico que se interessa em utilizá-la. Citando a crítica que Simon Reynolds (2011) faz à "*retromania*", ou seja, a incessante retrospecção que dominou a indústria musical a partir dos anos 2000, Herndon está preocupada com um olhar para o futuro sonoro:

> [...] O aprendizado de máquina por si só representa um novo começo no nosso entendimento de inteligências. [...] esse tópico convida a novos estados prototípicos da expressão humana musical, e convida a imaginar novos protocolos de organização no contexto da incerteza legal e cultural estimulada pela IA[17] (HERNDON, 2019, p. 18).

Herndon tem como uma de suas principais referências o compositor, performer e pesquisador George Lewis, especificamente no seu trabalho de desenvolvimento do programa de aprendizado de máquina *Voyager*. O sistema *Voyager* é uma ferramenta interativa que realiza o processamento e análise de uma performance em tempo real e interage musicalmente como uma espécie de "orquestra virtual

17 Tradução nossa: "*machine learning itself represents a new beginning in our understanding of intelligences. [...] this topic invites prototypical new states for human musical expression and calls for us to imagine new protocols of organization amidst the oncoming legal and cultural uncertainty stimulated by AI*".

improvisadora"[18] (LEWIS, 2000).

Enquanto ambiente de performance, *Voyager* se diferencia conceitualmente na medida em que Lewis busca remover a hierarquia do músico sobre a máquina, enquanto que essa hierarquia está presente em Assayag et al. (2006) e ainda mais delimitada no trabalho de Tone. Segundo Lewis,

> Voyager não necessita de input humana em tempo real para gerar música. Por isso, uma vez que o programa exibe comportamento generativo independente do improvisador, decisões tomadas pelo computador têm consequências na música que devem ser consideradas pelo improvisador[19] (LEWIS, 2000).

Para Herndon (2019), assim como para Lewis (2000), há um esforço para que o computador tenha um papel no fazer musical maior do que um instrumento, uma ferramenta passiva da qual dispõe o compositor ou improvisador. Para Lewis, isso se dá em uma reivindicação de uma característica de diversas tradições da música africana em que instrumentos têm frequentemente um *status* de ser humano, em uma forma de animismo musical. Lewis busca empregar esse animismo em seu programa, provê-lo de "seu próprio som" (*one's own sound*), ideia central no pensamento musical afro-estadunidense.

Para Herndon, *Spawn* é uma extensão de si e do grupo musical PROTO. Ela deseja tratar o programa "não como uma entidade subserviente, mas como uma reflexão externa"[20] da comunidade de vocalistas que dirige (HERNDON, 2019, p. 22). Nesse sentido, Herndon se distancia daquele entendimento de que a IA caminha na direção de se emancipar do domínio humano. Pelo contrário: ela destaca em seu trabalho que *Spawn* apenas pode existir a partir de um intenso e coletivo trabalho humano (OGILVIE-HANSON, 2020). A compositora faz referência novamente a Lewis (2000) quando este descreve a improvisação como também um produto da construção coletiva e que, portanto, seu programa *Voyager* é uma incorporação do fazer musical afro-estadunidense em forma computacional, de nenhuma forma um programa objetivo ou universal, mas um produto das ideias particulares dos seus criadores.

> No fim das contas, a conclusão a qual chego no meu trabalho com Spawn é que ferramentas de aprendizado de máquina são 'nós'; são uma coleção de trabalho e inteligência humanos. Enquanto a lógica de Spawn possa ser diferente da lógica do cérebro humano, ela foi não apenas criada por nós, mas toda a informação e contexto ao qual ela tem acesso foi provida por nós e nossos predecessores[21] (HERNDON, 2019, p. 23).

Essa perspectiva leva, inclusive, a compositora a comparar Spawn a uma criança, de criação sua e de seu parceiro criativo Mathew Dryhurst. Por isso refere-se a ela com o pronome feminino *she* e não *it*, como é a norma no inglês para se referir a coisas inanimadas. Dessa forma, Holly Herndon e o grupo

18 Tradução nossa: "*virtual improvising orchestra*".
19 Tradução nossa: "*Voyager does not need to have real-time human input to generate music. In turn, since the program exhibits generative behavior independently of the improvisor, decisions taken by the computer have consequences for the music that must be taken into account by the improvisor*".
20 Tradução nossa: "*not* [...] *as a subservient entity, but rather as an external reflection* [...]".
21 Tradução nossa: "*I wanted to reveal and aestheticize the logic at work as a neural network builds up sound from the ground up. This not only provides a more accurate glimpse into the current state of the technology, by not masking the final production with the sheen of digital instruments, but also forced me to deal with the current state of technology on an aesthetic level.*"

PROTO produziram o álbum de mesmo nome em que *Spawn* está presente como uma tecnologia em desenvolvimento. A sonoridade "*lo-fi*" da rede neural nos estágios iniciais de treinamento é trazida como elemento estético, e o álbum busca acompanhar o desenvolvimento do próprio modelo. Isso tem também a função de fornecer uma demonstração da tecnologia sem exageros ou um mascaramento excessivo das suas imperfeições e limitações (HERNDON, 2019, p. 23)

2.3. A saturação sonora de Dadabots

Dadabots é um duo de compositores sediado nos Estados Unidos, formado por Zack Zukowski e CJ Carr. Além de programadores e pesquisadores na área de aprendizado de máquina, ambos são músicos compositores e destacam o gênero do metal como principal referência. A dupla utiliza uma versão modificada ("*hackeada*", segundo eles próprios) do programa SampleRNN de redes neurais recorrentes (MEHRI et al., 2017) para a geração automática de música. Sua produção se dá na forma de álbuns, gerados com variado nível de seleção e curadoria por parte dos compositores, mas também de *livestreams*, em que o programa realiza a geração de música em tempo real, que, por sua vez, é transmitida por meio da plataforma YouTube©.

O "*hacker duo*" apresenta por meio de publicações científicas o uso de uma combinação de *small data*, que possibilita o controle mais estrito sobre os dados de treinamento. Como forma de aplicação, buscaram colaborações com DJs, beatboxers, bandas de metal, math rock, skate punk, etc., nas quais a produção de cada um destes colaboradores é utilizada separadamente como base de dados para o treinamento de um modelo que realiza uma releitura, *ingestão e digestão* por redes neurais desse repertório específico. Segundo é possível observar, o grupo Dadabots tem interesse particular em explorar, por meio do aprendizado de máquina, "gêneros musicais experimentais não abordados previamente por técnicas tradicionais de síntese" (CARR; ZUKOWSKI, 2018). São gêneros cujas características particulares de estilo não estão necessariamente na organização gramatical das notas, mas sim no aspecto espectral, timbrístico e espacial do som — os autores citam a influência do compositor Edgar Varèse e sua música de massas sonoras, o que justifica a escolha pela representação na forma de áudio, e não simbólica.

Transparece-nos uma preocupação estética na produção do Dadabots que motiva de forma bastante nítida muitas de suas escolhas técnicas. Ainda que, do ponto de vista formal, a dupla esteja buscando meios de realizar o modelamento de um determinado grupo ou gênero, uma certa experimentação sonora particular leva a caminhos que seriam desencorajados se o resultado pretendido fosse meramente uma imitação "fiel" e reconhecível daquele material. Para esse desenvolvimento, a dupla encontra maneiras divergentes de lidar com o *overfitting* e *underfitting* (CARR; ZUKOWSKI, 2018). Esses termos da literatura de aprendizado de máquina se referem, normalmente, a problemas de um modelo, ligados às escolhas da seleção e preparação dos dados, treinamento ou mesmo da arquitetura geral do programa. Um modelo que sofre de *overfitting* comporta-se de forma semelhante *demais* à sua base de dados, e é, portanto, incapaz de fazer generalizações e responder a *inputs inéditos*. Já o *underfitting* se refere ao resultado inverso, ou seja, o modelo não conseguiu reproduzir as características da base de dados com suficiente precisão. Os autores, no entanto, referem-se a esses comportamentos a partir de suas

características particulares no produto sonoro, como aspectos a serem trabalhados de acordo com uma intenção estética, e não eliminados. Eles demonstram o interesse no *overfitting* para as estruturas musicais de curta duração, a fim de que o timbre, instrumentos, vocais, etc. da base de dados apareçam com clareza; e priorizam o *underfitting* para as estruturas de duração mais longa, para que ritmos, riffs, seções e transições sejam menos determinísticos e, portanto, de decisão da máquina (ZUKOWSKI; CARR, 2018).

Quanto à experimentação em diferentes gêneros, os autores demonstram mais afinidade em trabalhar com artistas mais afeitos aos "artefatos estranhos da síntese neural (ruído, caos, mutações grotescas da voz)"[22]. Aqui está mais uma instância na qual a transparência da tecnologia de redes neurais não é interessante — pelo contrário, são as características particulares dessa tecnologia que o grupo busca explorar e colocar em primeiro plano no produto sonoro. As limitações da tecnologia são trabalhadas como elemento estético.

Rubinstein (2020) compara Dadabots ao movimento de música eletrônica experimental da fantologia23 (*hauntology*) de DJs como Burial e The Caretaker, surgido na Inglaterra nos anos 2000 e descrito pelo historiador Mark Fisher (2013). A fantologia na música teve como principal característica uma reapropriação irônica de artefatos culturais de um passado fonográfico. A combinação anacrônica e inesperada de gravações antigas que apareciam como "fantasmas" resultava em uma estranheza (*uncanniness*) particular, na qual a sonoridade da mídia exercia papel especial: o craquelar do vinil, a distorção analógica da fita eram elementos trabalhados em primeiro plano. Para Fisher, essa forma de sampleamento, que tem sua origem com os DJs jamaicanos do dub, carrega uma crítica da pós-modernidade ao se apropriar do passado fonográfico e reviver com ele a esperança promissora da era pré-Thatcher, esperança essa que morreu com a chegada do neoliberalismo.

Para Rubinstein, essa mesma estranheza crítica da fantologia está presente em Dadabots inclusive em aspectos gerais da sonoridade: o autor aponta como o álbum *Coditany of Timeness*[24], criado a partir do trabalho da banda Krallice25, soa como se estivesse em uma fita VHS antiga, gravada ao vivo em meio ao público. Esse é um potencial a ser explorado na IA como ferramenta musical como um todo, segundo o autor. Até trabalhos menos experimentais e mais mercadológicos, como a música *Daddy's Car*, produzida pelo grupo Sony em 2016 como propaganda de modelos generativos de música, traz esse poder crítico mesmo sem a menor intenção de fazê-lo.

A ironia crítica no trabalho de Dadabots, para nós, se dá também na forma do exagero e saturação extrema. A sua música é muito densa, trazendo muito da energia dos grupos com quem colaboram. No entanto, seus modelos não têm a capacidade de fazer mediações da sonoridade extrema na forma de criação de sentido musical — não são capazes de combinar trechos de alta intensidade com outros de uma dinâmica menor, não há nenhum momento em que se possa retomar o fôlego. Na verdade, é justamente na forma musical e nas estruturas de longa e média duração que seus álbuns estão mais livres para divergir da base de dados fornecida (*underfitting*). A experiência musical de bandas e artistas que lhes

22 Tradução nossa: "*strange artifacts of neural synthesis (noise, chaos, grotesque mutations of voice)*"
23 Também traduzida alternativamente como "assombrologia".
24 <https://dadabots.bandcamp.com/album/coditany-of-timeness>. Acesso em: 27 de mar. de 25.
25 Krallice é uma banda de black metal estadunidense.

permite um trabalho de dosagem dos elementos musicais não está presente no programa de Dadabots. Segundo os próprios artistas, o que os levou no fim das contas à programação de aprendizado de máquina foi o interesse por música "extrema" ("CJ Carr & Zack Zukowski | Eng | #I_make_AI - YouTube", 2024).

Além disso, a forma de geração musical em tempo real nas *streams* parece uma forma de demonstração dos modelos generativos despida de uma curadoria, de um mascaramento de suas imperfeições ou dos tratamentos típicos midiáticos de embelezamento por meio do falseamento de resultados. Assim como no trabalho de Herndon, em alguns álbuns o desenvolvimento das faixas corresponde a estágios progressivos do treinamento, sendo possível acompanhar a aprendizagem no decorrer do álbum (CARR; ZUKOWSKI, 2018). Porém, mesmo se fosse possível dispor de muito tempo, o modelo é circunscrito a um conjunto de possibilidades delimitado pela sua arquitetura e pela pequena base de dados sobre a qual foi treinado, o que evita propositalmente atingir resultados considerados "comercialmente" satisfatórios.

De forma similar, Sebastian Macchia apresenta soluções de modelamento muitas vezes explorando os limites da programação, apontando para resultados inesperados ou divergentes. Macchia é um criador de conteúdo audiovisual sediado nos Estados Unidos, interessado na música assistida por redes neurais. Em seu canal do YouTube© "*Nobody & the Computer*", está presente uma série de tutoriais em que o *host* apresenta diversas formas de se fazer música generativa a partir de modelos de IA. Entre eles, demonstra o *finetuning* do modelo de áudio MusicGen[26] para gerar áudios que sejam reconhecidos como uma performance de piano solo do jazzista Bill Evans ("Bill A.I.vans", 2024), explicando como realizar o mesmo processo com qualquer base de dados. Em outro experimento, faz uso de um grande modelo de linguagem (LLM) ChatGPT para modelar um contraponto bachiano de forma simbólica ("BachGPT", 2023). Em especial nesse último caso, há uma impraticidade cômica: a linguagem contrapontística musical é algo que diverge completamente do domínio do modelo, o que nos leva a esperar um desastre em termos composicionais. Mas a estranheza obtida desse processo é exatamente o fator de interesse, e não a automatização convincente de Bach: o autor acaba por fabricar um constrangimento público do modelo, que tão frequentemente tem suas capacidades infladas de forma sensacionalista.

3. Discussão

Na seção anterior, comentamos três abordagens musicais distintas sobre o tópico dos modelos generativos de IA. Buscamos demonstrar, principalmente, quais as motivações estéticas por trás desses trabalhos e como isso se traduz técnica e sonoramente. Agora, buscaremos comparar essas três perspectivas a partir de alguns aspectos: a sonoridade, a hierarquia existente (ou não) entre músico e máquina, e o engajamento no debate geral sobre IA.

3.1. Sonoridade da disfunção

É importante notar como, nas três abordagens, os compositores citam uma importante influência da música de sonoridades do século XX. Carr e Zukowski do Dadabots citam a música de Varèse

26 MusicGen é um modelo de linguagem generativo de música desenvolvido pelo grupo Meta (Facebook). Fonte: <https://musicgen.com/>. Acesso em: 27 de mar. de 25.

e o paradigma de massas sonoras, apontando a construção de timbres como aspecto central no seu trabalho; Herndon discute a música concreta no seu esforço de isolar a matéria sonora da sua fonte no exercício da composição e escuta; e Yasunao Tone não apenas teve essa música como referência como é *ele próprio* um proponente da música do século XX que diverge do paradigma das notas musicais.

Na atual fase de "primavera" da IA, a existência de grandes bases de dados de arquivos de áudio permitiu que os experimentos com redes neurais na música aconteçam não apenas no domínio simbólico, por meio do protocolo MIDI, mas diretamente com áudio em processos de síntese sonora. Para Herndon e Dadabots, isso é um elemento central de interesse na IA. Em Herndon, isso se dá na exploração da síntese da voz, como entidade extra-sonora complexa e social (HERNDON, 2019, p. 12). Assim como a voz humana não é inata, mas adquirida e coletiva, a voz do seu programa *Spawn* é síntese da colaboração com seu grupo vocal, em uma via de mão dupla — o grupo teve contato com o modelo em diferentes estágios de treinamento, e pode responder àqueles resultados parciais, em retroalimentação.

> A maior parte dos algoritmos de aprendizado de máquina são treinados em dados no formato MIDI a partir de partituras restritas à altura e duração das notas. Isso deixa de fora muitos dos parâmetros que se tornaram centrais ao longo do século XX, como a modulação timbrística detalhada [...][27] (HERNDON, 2019, p. 24).

Dadabots, de forma similar, aponta para a utilização de áudio (em vez da representação simbólica) como forma de atingir características sonoras particulares de gêneros que, de uma forma ou de outra, foram negligenciados na literatura sobre modelos generativos de IA. A intenção, no entanto, não é de um modelamento preciso. Buscam-se, por outro lado, as características específicas da síntese neural, como reimaginação inspirada por aqueles gêneros. Busca-se a disposição particular da IA para a recombinação e saturação extrema dos timbres e texturas.

Também é elemento da sonoridade em todos os três exemplos a demonstração da disfunção tecnológica como elemento estético. Em Herndon e Dadabots, isso fica claro na forma como apresentam as fases iniciais de treinamento de seus modelos, com todos os artefatos sonoros do *over* e *underfitting*. Para Yasunao Tone, isso se dá de maneira ainda anterior, desde *MP3 Deviations*. A disfunção do formato MP3 é ainda reprocessada, e passa por uma camada nova de disfunção na performance em que usa IA.

3.2. Controle e autonomia

Uma das diferenças entre as três perspectivas abordadas está na forma de curadoria, seleção e processamento posterior do material obtido dos respectivos modelos, ou seja, em que grau é permitida uma autonomia do programa, e até que ponto o compositor exerce um comando direto do programa após seu desenvolvimento.

Herndon faz referência ao trabalho de George Lewis (2000) especificamente sobre o programa *Voyager*. Com esse modelo, Lewis teve o esforço de nivelar a hierarquia entre máquina e improvisador.

27 Tradução nossa: "*Most machine learning algorithms are trained on MIDI data from scores restricted to pitch and note length. This leaves out many of the parameters which have become more central in focus throughout the 20th century, such as detailed timbral modulation* [...]".

Ainda que seu programa tenha formas de analisar e responder aos estímulos musicais do músico na performance, não depende necessariamente de um músico para atuar. Sua intenção não é um programa que responda de forma unilateral ao material fornecido, mas sim que haja de fato um diálogo (LEWIS, 2000).

Ainda que isso seja da intenção de Herndon, entendemos que o caso de *Spawn* acaba sendo diferente por alguns motivos. Em primeiro lugar, seu modelo não é, ainda, capaz de realizar performance em tempo real. Fora do contexto do tempo compartilhado da performance, essa relação deve se dar necessariamente de forma distinta. Herndon refere-se a *Spawn* como uma criança, utilizando inclusive o pronome feminino. Uma criança não-humana, como ela destaca, mas não menos uma cria dela com Mathew Dryhurst no contexto coletivo do fazer musical com seu grupo vocal. Com isso, a compositora busca demonstrar que a tecnologia está nas fases iniciais de seu desenvolvimento, e assim expressar sua preocupação com o futuro da música (OGILVIE-HANSON, 2020). Assim, ainda que para Herndon seu programa seja mais do que uma simples ferramenta inanimada — sendo comparado a uma criança em desenvolvimento — trata-se de uma criança sob uma tutela bastante clara. Os erros e limitações são demonstrados como didática, mas tudo isso passa por um tratamento posterior de processamento do material obtido com *Spawn* a partir de técnicas mais usuais de produção musical digital.

Para Yasunao Tone, a hierarquia parece mais claramente colocada. O modelo é de sua criação e ele o controla, decide a partir de seus próprios critérios o momento de iniciar e interromper a performance, a intensidade, dinâmica e densidade. Mesmo assim, existem algumas ressalvas: em primeiro lugar, sua preferência pelos controles "não-intuitivos" é uma forma de manter, a todo momento, um grau de imprevisibilidade da performance. Na verdade, esse é um dos fatores centrais da escolha de utilizar redes neurais na performance. Em segundo lugar, assim como no *Voyager*, Tone compartilha com o público e a máquina o tempo presente da performance improvisada. Não há possibilidade de tratamento posterior do material obtido a partir do programa — ele é imediatamente transmitido para o público.

Com Dadabots, temos uma abordagem um tanto híbrida: há, ao mesmo tempo, a produção de álbuns inteiramente curados pela dupla; álbuns com o mínimo de curadoria possível; e *livestreams cujo resultado sonoro, uma vez que iniciadas,* é de encargo do programa e distante de um controle ou edição dos desenvolvedores (CARR; ZUKOWSKI, 2018).

3.3. Futuro e passado sonoro

Mark Fisher (2013) descreve o impasse pós-moderno da produção cultural em que, uma vez abandonados os princípios de inovação e ruptura radical com o passado e de imaginação de novos futuros do modernismo no século XX, a cultura volta-se para o passado, resgatando e recombinando símbolos como meros objetos de apreciação. Em grande medida, na condição pós-moderna, a reciclagem tomou o lugar da criação de novas formas artísticas. Simon Reynolds (2011) chamou esse fenômeno de "retromania". Nesse contexto, nos anos 2000, surge o que chamou de "fantologia" da música eletrônica no Reino Unido. Produtores como Burial e The Caretaker, ao samplearem um passado fonográfico pré-Thatcher, são capazes de tornar audível uma *nostalgia* de futuros que nunca se realizaram, ceifados pela condição pós-moderna.

Rubinstein (2020) argumenta que, nesse contexto, o advento de modelos generativos de

síntese neural aparece com uma potencialidade dupla. Por um lado, há uma perspectiva da "retromania" em que essa ferramenta aparece como uma forma avançada de reciclagem. Gigantescos acervos do passado fonográfico, na forma de bases de dados, estão disponíveis para uso livre em modelos que, essencialmente, fazem uma recombinação do *corpus* fornecido no seu treinamento. O autor cita, como exemplo, a música *Daddy's Car*, produzida pelo grupo Sony em 2016 no estilo dos Beatles como propaganda de modelos generativos de música e sua capacidade de reprodução de um estilo musical.

Por outro lado, a contradição colocada é que os modelos de síntese neural colocam aos músicos uma nova gama de possibilidades que está presente justamente na estranheza, na qualidade perturbadora da sua sonoridade particular. Para Rubinstein, a IA tem tudo a ver com o que Fisher chamou de "estranhamento tecnológico" (*technological uncanny*) ao analisar a música fantológica da década de 2000, e ninguém melhor que Dadabots explora isso atualmente. Sua busca por uma sonoridade específica da ferramenta, e a rejeição de uma perspectiva de modelamento fiel e transparente de um determinado *corpus* musical vai no sentido de ressignificar um passado fonográfico de forma irônica.

Betancourt (2024) discute a função e a disfunção tecnológica dos modelos generativos na formação de imagens. Esse tipo de modelo é normalmente otimizado para oferecer ao usuário uma imagem com coerência fotográfica, anatômica, etc., que mantenha invisível o fato de ter sido gerada por redes neurais. Mas Betancourt está interessado no resultado da disfunção do programa que falha na tarefa estabelecida, e acaba dessa forma evidenciando sua arquitetura. Ele dialoga com a perspectiva da *glitch art* (MENKMAN, 2010), que explora principalmente a estética da disfunção da tecnologia digital. A convergência dessa disfunção das redes neurais com a formação semiótica da escola surrealista do modernismo é discutida por Betancourt (2024) e também por O'Meara e Murphy (2023).

Tanto Dadabots como Holly Herndon estão preocupados em dar sentido estético aos artefatos particulares da IA, e com isso desenvolver uma sonoridade nova que pode resultar da experimentação com redes neurais.

> Essa nova capacidade levanta uma série de questões: o que compositores fazem com essa nova habilidade composicional, e como ela pode levar adiante a forma e o cânone da arte ao invés de simplesmente ser usada para repetir ou reconfigurar o que foi feito anteriormente? Onde podemos traçar a linha entre atos de curadoria e composição?[28] (HERNDON, 2019).

Esse, no entanto, é um ponto em que a peça de Yasunao Tone difere dos outros dois. Quando olhamos para o histórico de Tone desde o grupo Ongaku na década de 1950, passando pela sua experiência com o Fluxus e sua exploração da sonoridade das mídias desde a década de 1990, pensamos que a preocupação estética de Tone é, ainda, bastante atual, no sentido de buscar novas e radicais formas artísticas de ruptura. Tone utiliza *small data*, também como uma forma de evitar o terreno politicamente delicado das grandes bases de dados (CRAWFORD; PAGLEN, 2021). No entanto, a escolha de utilizar sua

28 Tradução nossa: "*This new capacity raises a number of questions: what do composers do with this newly found compositional ability, and how might it further the art form or canon rather than simply be used to repeat or reconfigure what has been done before? Where might we draw the line between the acts of curation and composition?*"

própria produção no treinamento de seu modelo indica uma continuação do trabalho sobre as mídias que, ao colocar o *formato* sobre o conteúdo, é bastante distinto do que foi chamado de fantologia.

4. Conclusão

A atual fase de crescente utilização dessa tecnologia específica nas artes atiçou debates multidimensionais, sobre uma série de temas, dentro e fora da academia, incluindo debates travados a partir da própria produção artística. Hertzmann (2018; 2020) discute a ideia de que um computador poderia ser creditado com a autoria de uma "obra de arte", ou se na verdade o computador é uma ferramenta para o artista, mesmo que este não se reconheça como tal. O autor argumenta que apenas a última opção pode ser verdadeira atualmente, ainda que haja notáveis exemplos de sujeitos na prática de criação que renunciam à autoria, afirmando ser o computador o artista (GRBA, 2022, p. 3). George Lewis, na década de 1990, desenvolveu o programa *Voyager*, que propôs como uma orquestra virtual improvisadora. O sistema *Voyager* é uma ferramenta interativa de aprendizado de máquina que realiza o processamento e análise de uma performance em tempo real e interage musicalmente. Sua proposta de ambiente interativo, no entanto, não era de uma resposta unilateral dos estímulos musicais que um improvisador fornecesse ao programa durante a performance. Na verdade, a intenção era de criar um diálogo de fato, em que a atuação do programa não dependesse de condições colocadas diretamente pelo músico, mas sim de nivelar a hierarquia entre músico e máquina. Em sua performance, ele inclusive permitia que *Voyager* tocasse sozinho por trechos longos.

Neste trabalho, discutimos três abordagens da IA na música distintas em sua técnica e poética. Primeiro, vimos como Yasunao Tone utiliza um modelo de sua própria performance utilizando aprendizado de máquina para realizar uma performance metalinguística em tempo real. Em seguida, vimos como Holly Herndon enxerga a síntese neural de voz como uma tecnologia nos seus primeiros estágios de desenvolvimento, e a utiliza como uma extensão do musicar coletivo na busca por novas qualidades sonoras, composicionais e de performance com IA. Também apresentamos o duo Dadabots, que está interessado nos artefatos sonoros e nas idiossincrasias da síntese neural para fazer uma releitura de gêneros cujo foco está na experimentação timbrística extrema.

Nossa atenção está para a seguinte questão: até que ponto a IA será uma forma de reciclagem e recombinação simples do passado musical, ou será ela uma ferramenta para a busca de formas e sonoridades musicais novas, em sintonia com as possibilidades técnicas e poéticas colocadas para os artistas no século XXI? Acreditamos que os artistas mencionados ajudam a tornar claro que, enquanto não há sinais de que o "inverno" está próximo, a segunda perspectiva está também na ordem do dia.

Referências

AGOSTINELLI, A. et al. MusicLM: Generating Music From Text. *arXiv*, 26 jan. 2023. Disponível em: <http://arxiv.org/abs/2301.11325>. Acesso em: 12 nov. 2024

ASSAYAG, G.; BLOCH, G.; CHEMILLIER, M. Improvisation et réinjection stylistiques. In: *Rencontres Musicales Pluridisciplinaires*, 2006, Lyon. Anais... Lyon: GRAME, 2006.

BETANCOURT, M. Computer creates a cat: sign formation, glitching, and the AImage. *Semiotica*, 8 jul. 2024.

BLAKE, Thom *et al*. YASUNAO TONE AND MP3 DEVIATION. In: *International Computer Music Conference*. Anais... Nova Iorque: 2010.

BRIOT, J.-P. From artificial neural networks to deep learning for music generation: history, concepts and trends. *Neural Computing and Applications*, v. 33, n. 1, p. 39–65, jan. 2021.

CARR, C. J.; ZUKOWSKI, Z. Generating Albums with SampleRNN to Imitate Metal, Rock, and Punk Bands. *arXiv*, 16 nov. 2018. Disponível em: <http://arxiv.org/abs/1811.06633>. Acesso em: 14 nov. 2024

CHEN, Y.; HUANG, L.; GOU, T. *Applications and Advances of Artificial Intelligence in Music Generation: A Review*. arXiv, 3 set. 2024. Disponível em: <http://arxiv.org/abs/2409.03715>. Acesso em: 7 out. 2024

CRAWFORD, K.; PAGLEN, T. Excavating AI: the politics of images in machine learning training sets. *AI & SOCIETY*, v. 36, n. 4, p. 1105–1116, 1 dez. 2021.

DA SILVEIRA, S. A. A hipótese do colonialismo de dados e o neoliberalismo. In: CASSINO, J. F.; SOUZA, J.; DA SILVEIRA, S. A. (org.). *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021. p. 32–50.

DÄLLENBACH, L. *The mirror in the text*. Cambridge: The University of Chicago Press, 1998.

DE MANTARAS, R. L.; ARCOS, J. L. AI and Music: From Composition to Expressive Performance. *AI Magazine*, v. 23, n. 3, p. 43-57, 2002.

FISHER, M. The Metaphysics of Crackle: Afrofuturism and Hauntology. *Dancecult*, v. 5, n. 2, p. 42–55, 2013.

GARCIA, M. P. O glitch e o mainstream: estética da disfunção tecnológica da década de 1990 a 2010. In: *Congresso da ANPPOM*, 34., 2024, Salvador. Anais... ANPPOM, 2024.

GOZALO-BRIZUELA, R.; GARRIDO-MERCHAN, E. C. ChatGPT is not all you need. A State of the Art Review of large Generative AI models. *arXiv*, 11 jan. 2023. Disponível em: <http://arxiv.org/abs/2301.04655>. Acesso em: 14 out. 2024

GRBA, D. Brittle Opacity: Ambiguities of the Creative AI. 5 abr. 2021.

GRBA, D. Deep Else: A Critical Framework for AI Art. *Digital*, v. 2, n. 1, p. 1–32, 5 jan. 2022.

HASSENFELD, N. Even the scientists who build AI can't tell you how it works. *Vox*, 15 jul. 2023. Disponível em: <https://www.vox.com/unexplainable/2023/7/15/23793840/chat-gpt-ai-science-mystery-unexplainable-podcast>. Acesso em: 12 nov. 2024.

HERNDON, H. *Proto*. Califórnia: Stanford, 2019.

HERTZMANN, A. Can Computers Create Art? *arXiv*, 8 maio 2018. Disponível em: <http://arxiv.org/abs/1801.04486>. Acesso em: 7 nov. 2024

HERTZMANN, A. Computers do not make art, people do. *Communications of the ACM*, v. 63, n. 5, p. 45–48, 20 abr. 2020.

HILLER, L.; ISAACSON, L. Musical composition with a high-speed digital computer. In: SCHWANAUER, S. M.; LEVITT, D. A. (ed.). *Machine models of music*. Cambridge, MA: MIT Press, 1993. p. 9–21. Publicação

original: 1957.

#I_MAKE_AI VIDS. CJ Carr & Zack Zukowski | Eng | #I_make_AI. YouTube, [data de publicação]. Disponível em: <https://www.youtube.com/watch?v=bmK8NCftTFg&ab_channel=%23I_make_AI_Vids>. Acesso em: 28 nov. 2024.

KANEDA, Miki; TONE, Yasunao. The "John Cage shock" is a fiction! Interview with Tone Yasunao, Part I. post: *notes on art in a global context*, Nova Iorque, 8 mar. 2013. Disponível em: <https://post.moma.org/the-john-cage-shock-is-a-fiction-interview-with-tone-yasunao-1/>. Acesso em: 26 nov. 2024.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, v. 60, n. 6, p. 84–90, 24 maio 2017.

LEWIS, G. E. Too Many Notes: Computers, Complexity and Culture in "Voyager". *Leonardo Music Journal*, v. 10, p. 33–39, 2000.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, v. 5, n. 4, p. 115–133, 1 dez. 1943.

MARGASAK, P. Veteran Japanese sound artist Yasunao Tone brings his latest, AI-enhanced, system of digital sound manipulation to town for a rare local performance. *Chicago Reader*, Chicago, 5 out, 2017. Disponível em: <http://chicagoreader.com/music/veteran-japanese-sound-artist-yasunao-tone-brings-his-latest-ai-enhanced-system-of-digital-sound-manipulation-to-town-for-a-rare-local-performance/>. Acesso em: 7 out. 2024.

MEHRI, S. et al. SampleRNN: An Unconditional End-to-End Neural Audio Generation Model. *arXiv*, 11 fev. 2017. Disponível em: <http://arxiv.org/abs/1612.07837>. Acesso em: 26 nov. 2024

MENKMAN, R. *Glitch studies manifesto*. Colónia, 2010.

MITCHELL, Melanie. Artificial intelligence: A guide for thinking humans. Reino Unido: Pelican Books, 2019.

MITCHELL, M. *Why AI is Harder Than We Think*. *arXiv*, 28 abr. 2021. Disponível em: <http://arxiv.org/abs/2104.12871>. Acesso em: 24 nov. 2024

NOBODY AND THE COMPUTER. BachGPT. YouTube, 26 jun. 2023. Disponível em: <https://www.youtube.com/watch?v=HS0MXynfstA>. Acesso em: 8 nov. 2024.

NOBODY AND THE COMPUTER. *Bill A.I.vans*. , 27 fev. 2024. Disponível em: <https://www.youtube.com/watch?v=KU_wzEmJN_Y>. Acesso em: 28 nov. 2024

OGILVIE-HANSON, E. S. *Posthuman Production: Technology and Embodiment in the Works of SOPHIE and Holly Herndon*. 2020. [Dissertação (Mestrado em Artes)] – Universidade McGill, Montréal, 2020.

O'MEARA, J.; MURPHY, C. Aberrant AI creations: co-creating surrealist body horror using the DALL-E Mini text-to-image generator. *Convergence: The International Journal of Research into New Media Technologies*, v. 29, n. 4, p. 1070–1096, ago. 2023.

OORD, A. VAN DEN et al. WaveNet: A Generative Model for Raw Audio. *arXiv*, 19 set. 2016. Disponível em: <http://arxiv.org/abs/1609.03499>. Acesso em: 11 nov. 2024

REED, S. et al. Generative Adversarial Text to Image Synthesis. *arXiv*, 5 jun. 2016. Disponível em: <http://arxiv.org/abs/1605.05396>. Acesso em: 11 nov. 2024

REYNOLDS, S.;. Retromania: Pop's past is taking over. *Los Angeles Times*, Los Angeles, 10 jul 2011.

Disponível em: <https://www.latimes.com/entertainment/la-ca-retro-rock-20110710-story.html>. Acesso em: 25 nov. 2024.

RUBINSTEIN, Y. Uneasy Listening. *Resonance*, v. 1, n. 1, p. 77–93, 7 maio 2020.

VASWANI, A. et al. Attention Is All You Need. *arXiv*, 2 ago. 2023. Disponível em: <http://arxiv.org/abs/1706.03762>. Acesso em: 11 nov. 2024

WHITTAKER, Meredith. The steep cost of capture. *Interactions*, v. 28, n. 6, nov./dez. 2021. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4135581>. Acesso em: 6 nov. 2024.

ZUKOWSKI, Z.; CARR, C. J. Generating Black Metal and Math Rock: Beyond Bach, Beethoven, and Beatles. *arXiv*, 16 nov. 2018. Disponível em: <http://arxiv.org/abs/1811.06639>. Acesso em: 8 nov. 2024

ZYLINSKA, J. *AI Art: Machine Visions and Warped Dreams*. Londres: Open Humanities Press, 2020.

Matheus Prado Garcia é pesquisador, instrumentista e compositor, atualmente realizando pesquisa de mestrado na Universidade Estadual de Campinas (Unicamp) em colaboração com o Núcleo Interdisciplinar de Comunicação Sonora (NICS), na linha Música, Linguagem e Sonologia do Programa de Pós-Graduação em Música. Atua na área de música e tecnologia, atualmente pesquisando musicalidade surda e música eletrônica de paredões. Formou-se bacharel em música popular na Unicamp em 2022, tendo realizado pesquisa de iniciação científica na área de arranjo onde realizou uma edição crítica e análise de dois arranjos de José Roberto Branco para big band em 2020. É co-realizador do projeto SurdoSystem, instalação e oficina de música de paredões voltada à percepção e musicalidade surda.
E-mail: matheuspradog73@gmail.com

Manuel Falleiros é Doutor em Artes pela Universidade de São Paulo (USP) e Pesquisador de Carreira na Universidade Estadual de Campinas (Unicamp). Atua no Núcleo Interdisciplinar de Comunicação Sonora (NICS/Unicamp) e como Professor Permanente do Programa de Pós-Graduação em Música do Instituto de Artes da Unicamp. Sua trajetória consolida a interseção entre investigação e criação artística, com ênfase em processos criativos, cognição musical e tecnologias interativas. É líder de grupos de pesquisa certificados pelo CNPq, coordena eventos científicos de repercussão internacional.
E-mail: mfall@unicamp.br

Disponibilidade de dados: Os dados estarão disponíveis mediante solicitação.

Conflito de Interesses: Nenhum conflito de interesses foi declarado.

Modalidade de avaliação: Duplo-cega por pares.

A *Opus* está licenciada sob uma licença CC-BY.