

## Foley e inteligência artificial

Rafael Slechticius da Silva<sup>1</sup>, Chrístian Langaro Vaisz<sup>1\*</sup>  
Orientador(a)\*

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) - Campus Alvorada. Alvorada, RS

No audiovisual, foley é a técnica de cria sons em sincronia com a imagem, geralmente na fase da pós-produção. O artista de foley se coloca em um estúdio especializado e realiza os sons de forma artesanal. Dentro da produção, o sound designer é responsável por aspectos da criação, geração e edição de som e trabalha junto com o artista de foley para desenhar o de uma cena. O sound designer profissional é equipado com um banco de sons, fornecendo vários efeitos sonoros para selecionar e usar em cada cena. Alguns destes bancos são pagos, mas há também bancos de sons gratuitos em alguns sites na internet, porém nem sempre estes sons possuem qualidade para uso profissional. Sons que demandam sincronia, ações rítmicas na tela como passos ou impactos são muitas vezes inacessíveis sem um artista foley. Para produções amadoras, não ter acesso a um estúdio e artista de foley pode ser um obstáculo ao desenho sonoro da produção, porém, as novas tecnologias podem solucionar esta demanda. Os modelos de geração de foley por inteligência artificial podem democratizar o processo de foley para pequenas produções que não têm acesso a essas ferramentas e, se a qualidade dos resultados conseguir igualar a qualidade profissional, podem ser utilizados para automação, assim como outros processos de design de som foram no passado. Os objetivos dessa pesquisa são primeiramente mapear as ferramentas de geração de foley e avaliá-las de acordo com suas demonstrações e suas aplicabilidades no design sonoro. Testar as ferramentas com vídeos caso possível, e fazer observações diante dos resultados e demonstrações disponíveis. Foram encontradas quatro ferramentas a partir de uma pesquisa bibliográfica. As ferramentas foram testadas usando a versão de demonstração nos sites HuggingFace e MMAudio, e as demonstrações em vídeo foram encontradas no site de cada ferramenta. Observamos que as ferramentas possuem resultados de qualidade baixa relativa ao padrão de indústria, como falta de sincronia, fidelidade e alinhamento com o vídeo, além dos sons serem entregues em um canal só, dificultando a edição individual de cada camada sonora. Para usar esses modelos, é necessário ter um conhecimento intermediário em IA e programação, embora as demonstrações testáveis sejam intuitivas. A partir da pesquisa constatamos que as ferramentas não incluem um fluxo de trabalho que atende a utilização por profissionais da indústria. Para possibilitar a aplicação prática da IA generativa na construção de foley em um fluxo de trabalho de um sound designer, seria necessário que estas ferramentas tivessem integração com softwares de edição de som por meio de plugins. Isso possibilitaria a entrega dos sons em camadas separadas diretamente nas pistas de áudio dentro destes softwares, em sincronia com a referência visual.

Palavras-chave: Foley; Sound design; Inteligência artificial.