

**UNIVERSIDADE FEDERAL DO RIO DE JANEIRO**  
**ESCOLA POLITÉCNICA**  
**DEPARTAMENTO DE ENGENHARIA ELETRÔNICA E DE**  
**COMPUTAÇÃO**

**PROPOSTA DE PROJETO DE GRADUAÇÃO**

Aluno: Bruno Machado Afonso  
bruno.ma@poli.ufrj.br

Orientador: Mariane Rembold Petraglia

**1. TÍTULO**

Desenvolvimento de Base de Dados para Treinamento de Redes Neurais de Reconhecimento de Voz através da Geração de Áudios com Resposta Ao Impulso Simuladas por Técnicas de Data Augmentation.

**2. ÊNFASE**

Computação

**3. TEMA**

O tema do trabalho é sobre o estudo de uma forma de simular Respostas ao Impulso de Ambientes Acústicos (RIR) com parametrizações diferentes a partir de amostras de RIR gravadas em ambientes reais, e ainda usar a RIR para gerar amostras de áudio em locais simulados a partir de gravações de voz reais.

**4. DELIMITAÇÃO**

O estudo é focado em inferir uma técnica de reforço de dados tanto em amostras reais de RIR quanto nas gravações de voz. Este trabalho está delimitado em apenas modificar amostras reais de áudio, e não gerar amostras simuladas sem uma gravação de base.

**5. JUSTIFICATIVA**

Com o avanço das tecnologias de automação residencial, assistentes pessoais nos smartphones e comunicação online, o estudo de técnicas de processamento de áudio (no caso específico deste trabalho, relacionados a voz), tornou-se mais relevante para a sociedade. Uma das características mais importantes a ser detectada no processamento de áudio é a Resposta ao Impulso do ambiente, pois através dela

é possível extrair informações pertinentes do ambiente em que o áudio foi gravado e também detectar a posição de fontes sonoras no local e as isolar para reconhecimento.

Junto a isso, houve avanços no âmbito do aprendizado de máquina, fornecendo alternativas para os métodos tradicionais de processamento de áudio. Modelos de arquitetura de redes neurais necessitam de um grande volume de dados para que sejam treinados e aprimorados, e um dos mais recentes desafios nessa área é de capturar essa extensa quantidade de gravações de áudio, pois é uma tarefa alto custo tanto financeiro e temporal, necessitando de equipamento especializado e diversos locais com características de modelo sonoro diferentes e pessoas diversas para amostras de voz.

## 67. OBJETIVO

O objetivo geral kkkkk, entkkkkko, propor um modelo computacional capaz de sistematizar o processo de ativakkkkkkkkkko cerebral humano para um conjunto limitado de estkkkkkkmulos. Desta forma, tem-se como objetivos especkkkkkkficos: (1) relacionar um conjunto de estkkkkkkmulos morais e emocionais que serkkkkko tratados pelo modelo computacional; (2) construir um modelo tridimensional do ckkkkkrebro humano que possibilite a representakkkkkkkkkko espacial das kkkkkreas fisiolkkkkkgicas referentes ao estudo proposto, e; (3) elaborar um sistema formal capaz de deduzir uma determinada seqkkkkkkkkkkknica de entrada. Este sistema formal serkkkkk um sistema reconhecedor.

## 7. METODOLOGIA

Este trabalho irkkkkk utilizar a correlakkkkkkkkkko funcional entre a atividade cerebral e os aspectos abstratos das emokkkkkkkkkkes morais para a modelagem de um processo de tomada de deciskkkkkko. A partir do uso da resposta BOLD (Blood Oxygen Level Derived) em imagens de ressonkkkkkkknica funcional, se pretende estabelecer um modelo computacional que represente aspectos do comportamento deciskkkkkkrio humano, para fins de identificakkkkkkkkkko.

A correlakkkkkkkkkko funcional entre a atividade cerebral e os aspectos abstratos das emokkkkkkkkkkes morais durante a tomada de deciskkkkkko, pode ser evidenciada pela ankkkkkklise da ativakkkkkkkkkko temporal em imagens mkkkkkdicas de Ressonkkkkkkknica Magnkkkkktica funcional (RMf). O exame RMf faz uso da

resposta BOLD [1] para evidenciar as áreas do córtex humano que apresentam aumento significativo da atividade neural. Este aumento é espacialmente caracterizado pela redução da taxa de oxigenação da hemoglobina, provocando a atenuação do sinal de Ressonância Magnética (RM).

Desta forma, através de ambientes interativos baseados nos aspectos estético e dinâmico de jogos interativos, situações envolvendo tomadas de decisões assistidas por computador, e ainda, com o apoio de equipamentos avançados de RM, deseja-se mensurar e analisar a atividade cerebral de um indivíduo (jogador). Assim, durante esses jogos interativos, o cérebro do indivíduo será monitorado e sua atividade será avaliada a partir do uso de técnicas de processamento de imagens online. O procedimento proposto permitirá uma modelagem mais eficiente da dinâmica evolutiva das emoções morais, otimizando a compreensão e o delineamento da fronteira de sentimentos humanos.

As recentes evidências experimentais indicam que o comportamento sócio-moral do homem é baseado em circuitos cerebrais específicos, porém o mapeamento destes circuitos ainda encontra-se indefinido. A partir do processamento de imagens de RMf resultantes de estímulos cooperativos inseridos em jogos, pretende-se evidenciar o relacionamento entre as estruturas específicas do cérebro humano responsáveis pela gênese dos sentimentos morais e emocionais, a partir de ações cooperativas e não-cooperativas durante a dinâmica dos jogos [?].

O sucesso deste trabalho está centrado na determinação de uma metodologia para a construção de um modelo computacional do cérebro humano relacionado com sentimentos morais e emocionais, segundo algumas hipóteses previamente definidas. Técnicas de Computação Gráfica e Processamento de Imagens serão empregadas na construção do modelo computacional [?] do processo de atividade cerebral proposto. As imagens de RMf, que sofrem o processamento, serão obtidas em bancos de imagens de domínio público.

## 8. MATERIAIS

Relacionar os materiais que estão previstos no projeto (*computadores,*

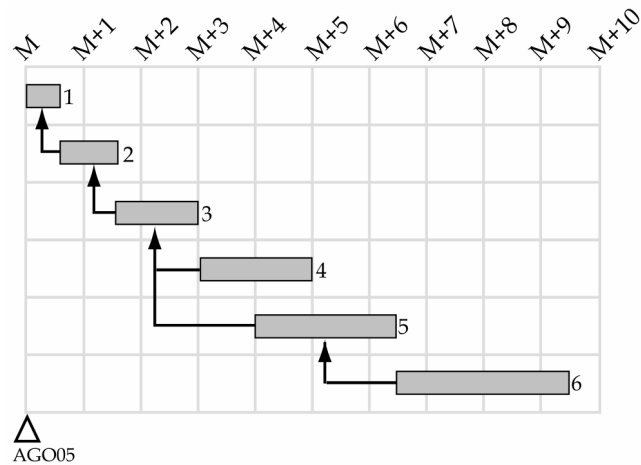


Figura 1: (*Atenkkkkkkkkkko, evitar projetos com menos de 5 meses*)

*instrumentos, equipamentos, dados, software: explicitar se hkkkkk licenkkkkka)*

## 9. CRONOGRAMA

Apresentada graficamente conforme a Figura 1.

Fase 1: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

Fase 2: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

Fase 3: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

Fase 4: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

Fase 5: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

Fase 6: Descrkkkkkkkkkkko sucinta do que serkkkkk feito.

# Referências Bibliográficas

- [1] N. J. Bryan, “Impulse response data augmentation and deep neural networks for blind room acoustic parameter estimation,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2020.

Rio de Janeiro, 4 de junho de 2021

---

Bruno Machado Afonso - Aluno

---

Mariane Rembold Petraglia - Orientador