

## Avaliação de Sistema de Rec. de Locutor em uma base de vozes de gêmeos idênticos

Ronaldo R. Silva<sup>1\*</sup>, Júlio Cesar Cavalcanti<sup>2</sup>, Anders Eriksson<sup>3</sup>

<sup>1</sup> Serviço de Perícias em Audiovisual e Eletrônicos – SEPAEL/DPER/INC/DITEC/PF, Brasília, DF, Brasil

<sup>2</sup> Júlio Cesar Cavalcanti – Stockholm University – Suécia

<sup>3</sup> Anders Eriksson – Stockholm University – Suécia

\*Autor correspondente, e-mail: rodrigues.rrs@pf.gov.br

### RESUMO

Foi avaliada o poder discriminativo de um sistema de reconhecimento automático de locutor (SRAL) baseado no projeto *Speechbrain*<sup>1</sup> quando aplicado a uma base de áudios provenientes de amostras de vozes de dez pares de gêmeos idênticos, falantes do português brasileiro do estado de Alagoas.

**Palavras-chave:** Comparação de locutor, *speechbrain*, gêmeos.

### Introdução

Um dos cenários forenses mais desafiadores na discriminação de indivíduos pela voz ocorre quando envolve a determinação da autoria entre gêmeos idênticos. Abordagens baseadas em SRAL têm sido agregadas às análises tradicionais<sup>2</sup> e é avaliada a sua aplicabilidade neste cenário.

### Objetivos

Avaliar o poder discriminativo de um sistema SRAL baseado no projeto *Speechbrain*, modelo ECAPA-TDNN treinado e avaliado nas bases *Voxceleb 1 e 2*, quando aplicado ao cenário descrito.

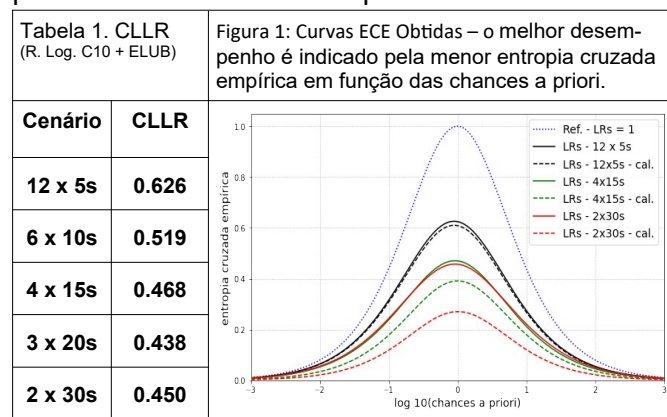
### Métodos

Foram selecionadas duas amostras de fala em dois diferentes estilos, contendo 60s de duração de cada um dos falantes presentes na base de gêmeos<sup>2</sup>, segmentadas em trechos de 5s, 10s, 15s, 20s e 30s. Foram obtidos os escores de similaridade aplicando validação cruzada do tipo *Leave-One-Out / Leave-Two-Out*, modelados como funções de densidade de probabilidade (PDF) gaussianas para as hipóteses:  $H_1$  (mesma origem) e  $H_2$  (origens diferentes – restritas a irmãos gêmeos). A partir das distribuições, foram calculados os *Log-Likelihood-Ratio Cost (CLLR)* correspondentes.

### Resultados e Discussão

A Tabela 1 sumariza os resultados obtidos onde pode ser observado que quanto maior o trecho da amostra de áudio utilizada, melhor é a discriminação do sistema, em linha com resultados

publicados na literatura envolvendo sistemas SRAL aplicados a populações típicas de falantes<sup>3</sup>. Sistemas com CLLR abaixo de 1 são considerados potencialmente relevantes para o uso forense<sup>4</sup>.



### Conclusão

Os resultados indicam uma boa capacidade de discriminação de sistemas SRAL, mesmo quando utilizado em cenários desafiadores. Pretende-se estender esta análise à base de vozes envolvendo irmãos não gêmeos e avaliar a combinação de resultados envolvendo análises perceptuais e acústicas associadas a sistemas automáticos.

### Referências bibliográficas

<sup>1</sup> Ravanelli, M., Parcollet, T., Plantinga, P., Rouhe, A., Cornell, S., Lugosch, L., et al. *Speechbrain: A general-purpose speech toolkit*. arXiv preprint arXiv:2106.04624, 2021.

<sup>2</sup> Cavalcanti, J. C., Eriksson, A., and Barbosa, P. A. Multi-parametric analysis of speech timing in inter-talker identical twin pairs and cross-pair comparisons: Some forensic implications. *Plos one* 17, e0262800, 2022.

<sup>3</sup> Weber P., Enzinger E, Labrador B, Lozano-Díez A, Ramos D, González-Rodríguez J, Morrison GS. Validations of an alpha version of the E3 Forensic Speech Science System (E3FS3) core software tools. *Forensic Sci Int Synerg.* 100223. doi: 10.1016/j.fsisyn.2022.100223. PMID: 35281657; PMCID: PMC8908042, 2022.

<sup>4</sup> MORRISON GS et al., *Science and Justice*, v. 61. n. 3, p. 299-309, 2021.