

# A Inteligência Computacional no Reconhecimento de Sinais de Libras

Tamires Martins Rezende

TDC Transformation: Competências para Evolução das Pessoas e Negócios  
Trilha: Inteligência Artificial

24 de agosto de 2021





# Sobre mim...



**2006-2009:** Instrumentação e Controle de Processos (IFMG-OP)

**2009-2011:** Química Licenciatura (UFOP)

**2011-2014:** Engenharia de Controle e Automação (UFOP)

**2015-2016:** Mestrado em Engenharia Elétrica (UFMG)

**2017-2021:** Doutorado em Engenharia Elétrica (UFMG)

**2021...:** Pesquisadora associada (Gaia, Solutions on Demand)

**2021...:** Especialista em desenvolvimento de software (FITec)

# Libras: Língua Brasileira de Sinais

Lei nº 10.436/2002  
Decreto nº 5.626/2005

Natureza  
Visual-Gestual

Língua de sinais  
tem gramática própria

Libras não é a segunda  
língua oficial do Brasil

**Língua Brasileira de Sinais**

Libras não é português  
em gestos

Não existe uma língua  
de sinais universal

~~Linguagem de  
sinais~~

Nem todo surdo  
usa Libras

Nem todo surdo pertence  
à cultura surda

Libras é a  
~~"Língua natural dos surdos"~~



# Parâmetros Fonológicos da Língua de Sinais

Sinal (Libras)  $\Leftrightarrow$  Palavra (Língua Oral/Escrita)  
Movimento (Libras)  $\Leftrightarrow$  Sons (Língua Oral/Escrita)

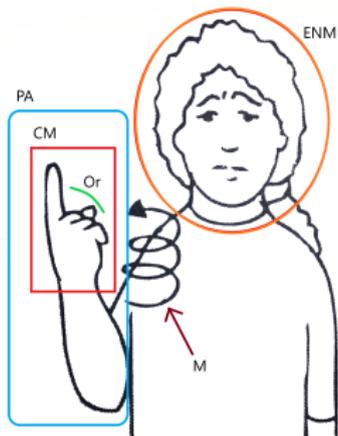


Figura 1: Sinal “furacão”.

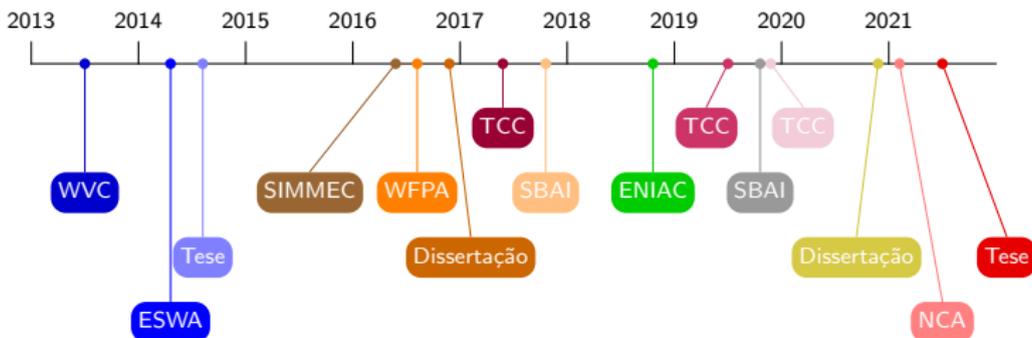
PA: Ponto de articulação  
M: Movimento  
CM: Configuração de mão  
Or: Orientação da palma da mão  
ENM: Expressões não-manuais.



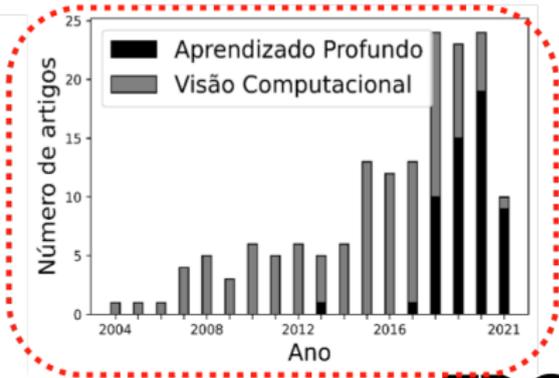
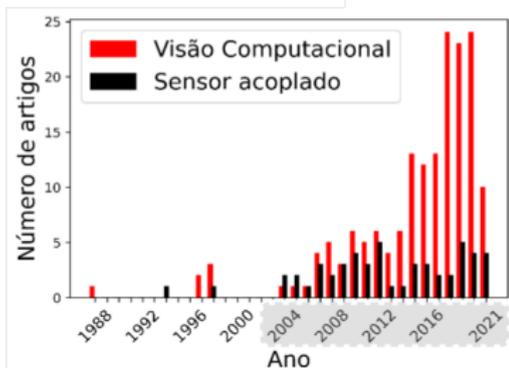
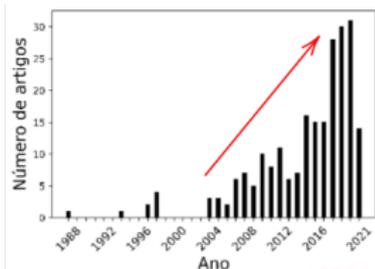
# Libras no MINDS

## *Machine Intelligence and Data Science*

<https://minds.eng.ufmg.br/>



# Contexto Histórico



# Aprendizado de Máquina

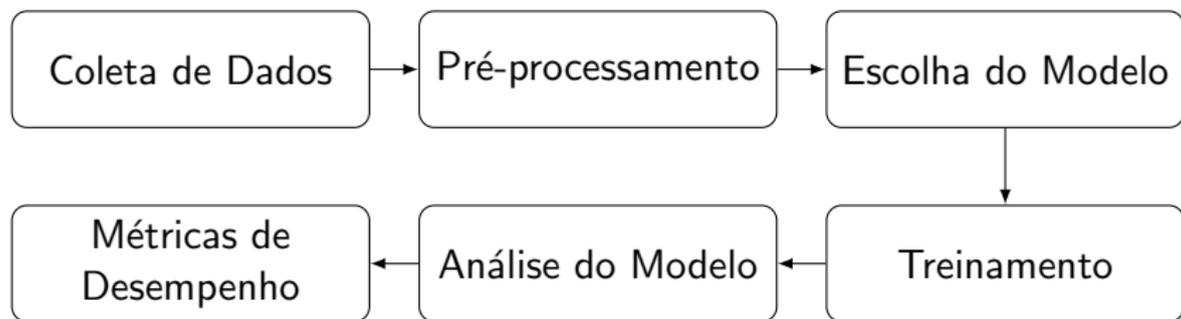


Figura 2: Etapas básicas de um processo de aprendizado de máquina.

# Base de Sinais de Libras: MINDS-Libras

Protocolo de gravação:

- 1** Quais sinais serão gravados?
- 2** Quem executará os sinais?
- 3** Quais sensores serão utilizados?
- 4** Onde as gravações serão realizadas?

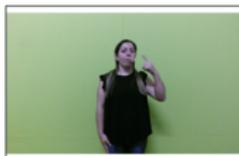


# Dados Disponibilizados: MINDS-Libras

**20 sinais × 5 repetições × 12 sinalizadores = 1200 amostras**



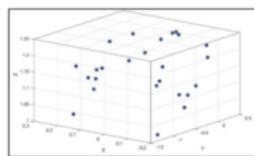
1920x1080



1920x1080



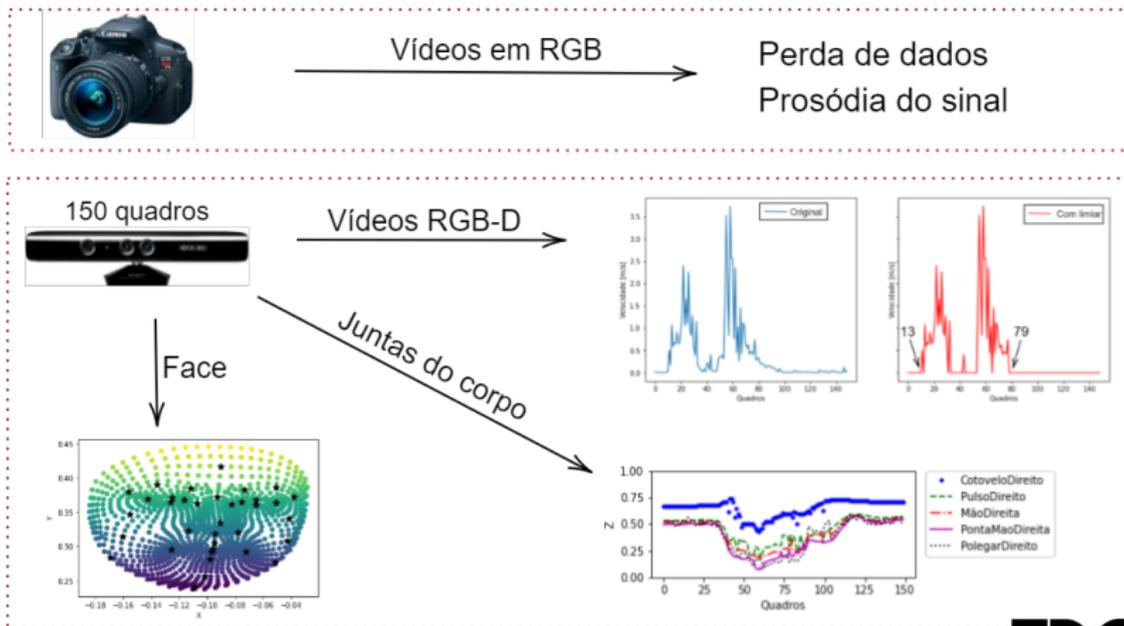
640x480



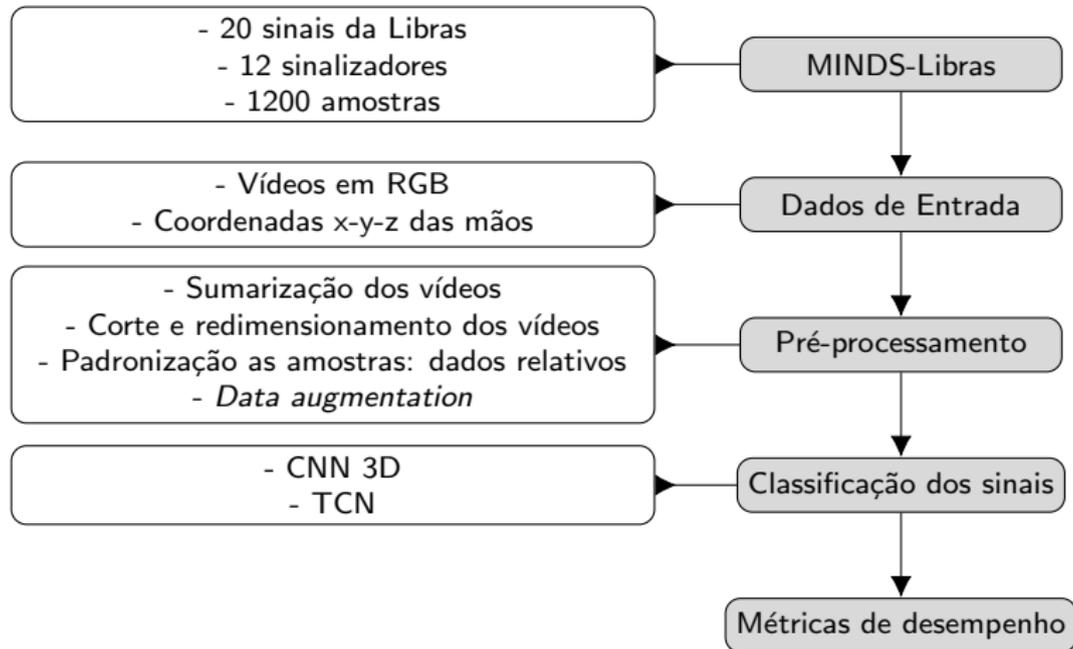
25 pontos do corpo  
1347 pontos do rosto

Figura 3: Dados disponibilizados pelos dispositivos de captura: câmera RGB e sensor RGB-D.

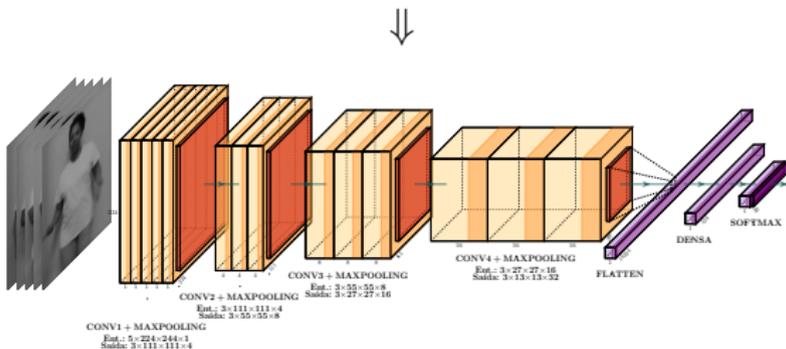
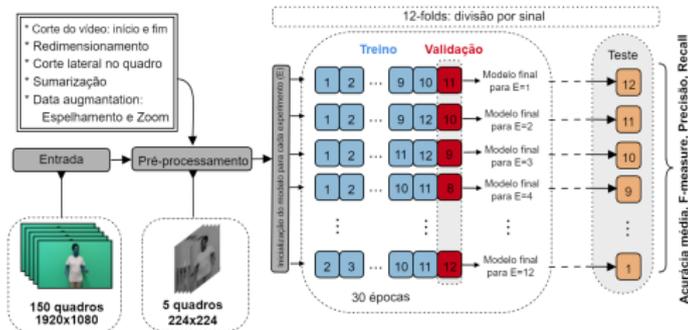
# Análise Descritiva: MINDS-Libras



# Abordagem



# Abordagem: Vídeo + CNN3D



# Resultado: Vídeo + CNN3D

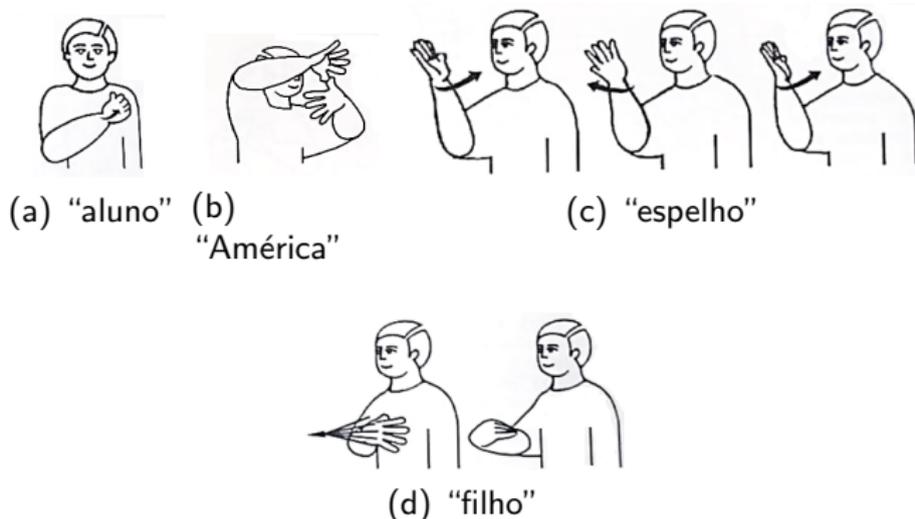


Figura 4: Sinais que tiveram as maiores taxas de reconhecimento.



## Resultado: Vídeo + CNN3D

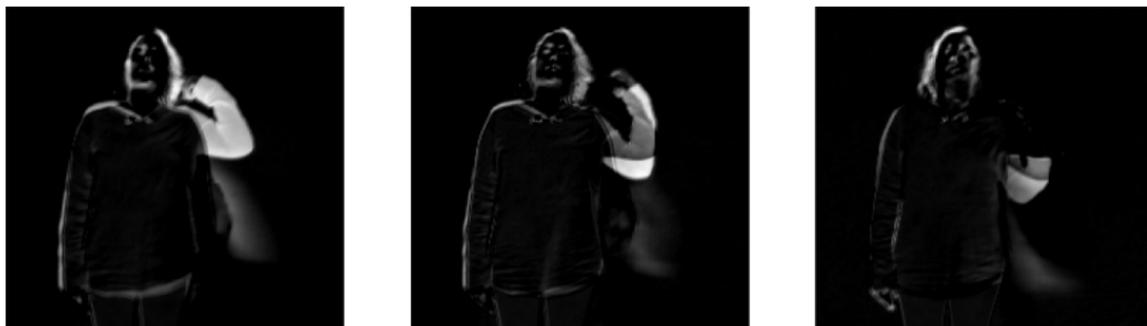


Figura 6: Saída do primeiro bloco convolucional da CNN 3D proposta (*Conv3D\_1*) na execução do sinal “banco”: 3 imagens  $222 \times 222$  do 2 mapas de características.

# Abordagem: Pontos da mão + TCN

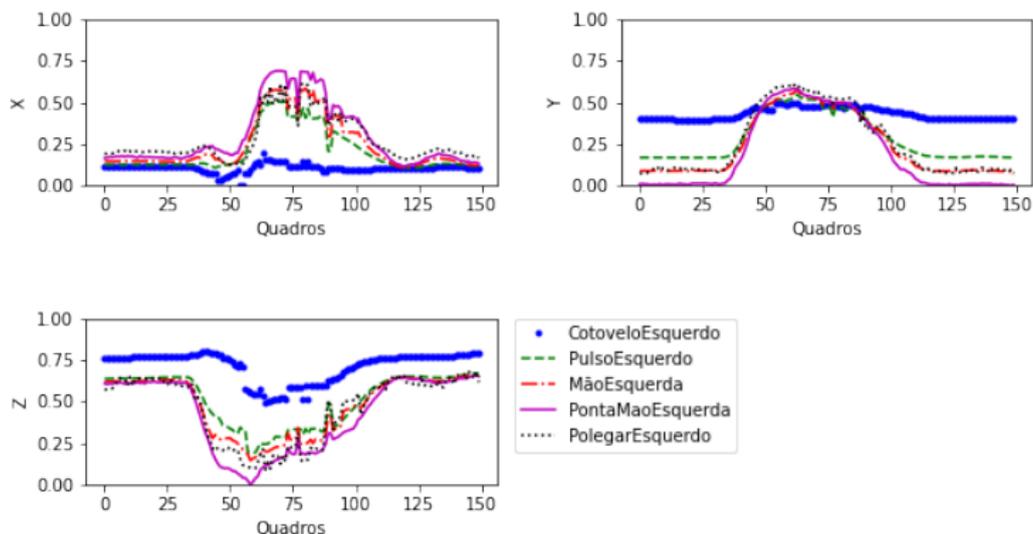
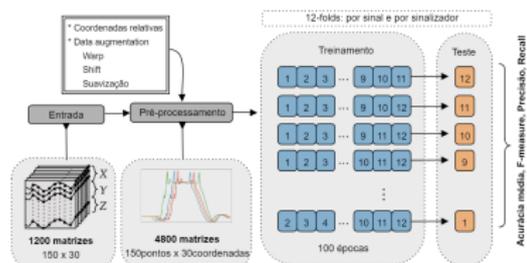


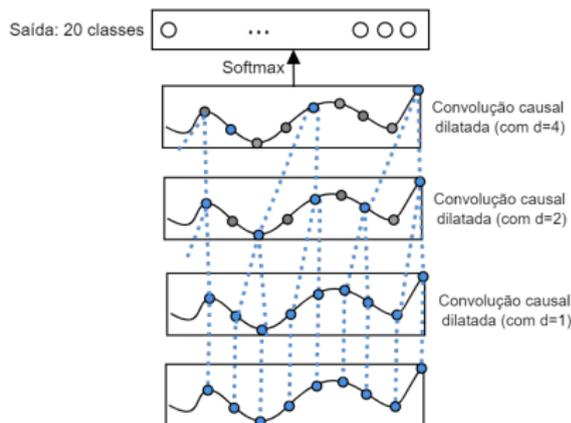
Figura 7: Variação das coordenadas  $x$ ,  $y$  e  $z$  normalizadas dos 5 pontos relativos à mão esquerda, ao longo dos 150 quadros da amostra *1-01Acontecer\_1Body.txt*.

# Abordagem: Pontos da mão + TCN

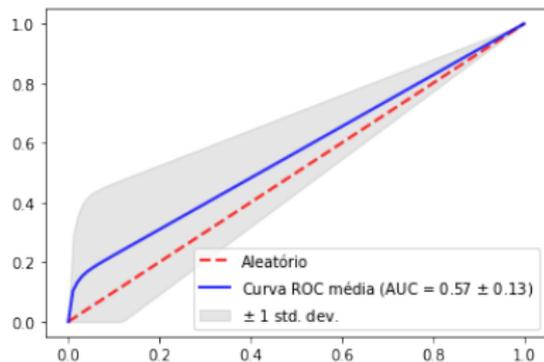
- Camadas convolucionais causais dilatadas
- Normalização dos pesos dos filtros
- Função de ativação ReLU



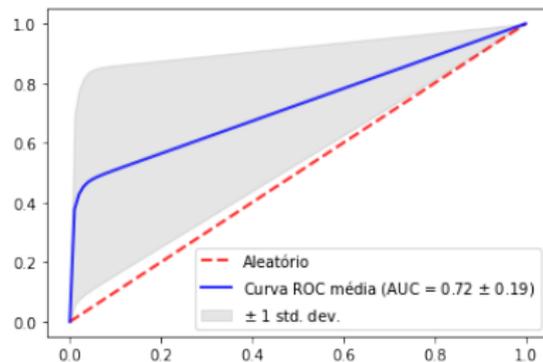
Otimização dos Hiperparâmetros:  
Algoritmo Genético



# Resultado: Pontos da mão + TCN



(a) 6 sinalizadores (AUC = 0,57)



(b) 12 sinalizadores (AUC = 0,72)

Figura 8: Curva ROC da TCN quando os conjuntos de treino e teste foram divididos por sinalizador (12 sinalizadores – Acurácia = 74,9%).

# Resultados

Tabela 1: Resultados dos experimentos realizados com a base de dados MINDS-Libras

Exp.	Entrada	Pré-processamento	Abordagem	Modelo	Acurácia média	AUC	Tempo (s)
Baseline	Escala de Cinza	S/R: 10 quadros e DA	Sinal	CNN3D	93,30% ( $\pm 1,69$ )	–	570,6
1	Escala de Cinza	S/R: 5 quadros	Sinal	CNN3D	82,42% ( $\pm 4,82$ )	0,91	78,75
	Escala de Cinza	$\leftrightarrow$ + DA	Sinal	CNN3D	84,75% ( $\pm 4,42$ )	0,92	262,1
2	<i>Position x-y-z</i>	DA	<b>Sinal</b>	TCN	<b>96,00% (<math>\pm 6,69</math>)</b>	0,98	136,0
3	<i>Position x-y-z</i>	DA / 6 sinalizadores	Sinalizador	TCN	62,40% ( $\pm 9,71$ )	0,57	34,32
	<i>Position x-y-z</i>	DA / 12 sinalizadores	<b>Sinalizador</b>	TCN	<b>74,90% (<math>\pm 9,52</math>)</b>	0,72	65,24

\*S/R: Sumarização e Redimensionamento, DA: *Data Augmentation*

# Conclusões

- Abordagens: Vídeos em RGB e a Rede Neural Convolutacional 3D (CNN 3D) e Trajetória manual e a Rede Neural Convolutacional Temporal (TCN);
- Dados reais → Problema real;
- Linha de pesquisa em expansão / crescimento / evolução... há muito a se fazer!



# Conclusões

- Abordagens: Vídeos em RGB e a Rede Neural Convolutacional 3D (CNN 3D) e Trajetória manual e a Rede Neural Convolutacional Temporal (TCN);
- Dados reais → Problema real;
- Linha de pesquisa em expansão / crescimento / evolução...  
há muito a se fazer!

# Conclusões

- Abordagens: Vídeos em RGB e a Rede Neural Convolutacional 3D (CNN 3D) e Trajetória manual e a Rede Neural Convolutacional Temporal (TCN);
- Dados reais → Problema real;
- Linha de pesquisa em expansão / crescimento / evolução... há muito a se fazer!





# Obrigada!!!

- Lattes: <http://lattes.cnpq.br/7705199625744561>
- LinkedIn: <https://www.linkedin.com/in/rezendetamires/>
- ResearchGate:  
<https://www.researchgate.net/profile/Tamires-Rezende>
- Gaia: <https://www.gaiasd.com/>
- FITec: <http://www.fitec.org.br/>

