

Layout-Aware Zero-Shot Learning for Visual Document Matching

Qualificação de Mestrado

Lucas de Almeida Bandeira Macedo

Universidade de Brasília
Departamento de Ciência da Computação

Orientador: Prof. Dr. Pedro Garcia Freitas
Coorientador: Prof. Dr. Bruno Luiggi Macchiavello Espinoza

Outubro de 2025

1 Introdução

2 Metodologia

3 Resultados

4 Conclusão

Introdução

Contexto - Documentos e Compliance

- Ambiente Bancário
- Documentos físicos
- Imagens de documentos
- Exemplos:

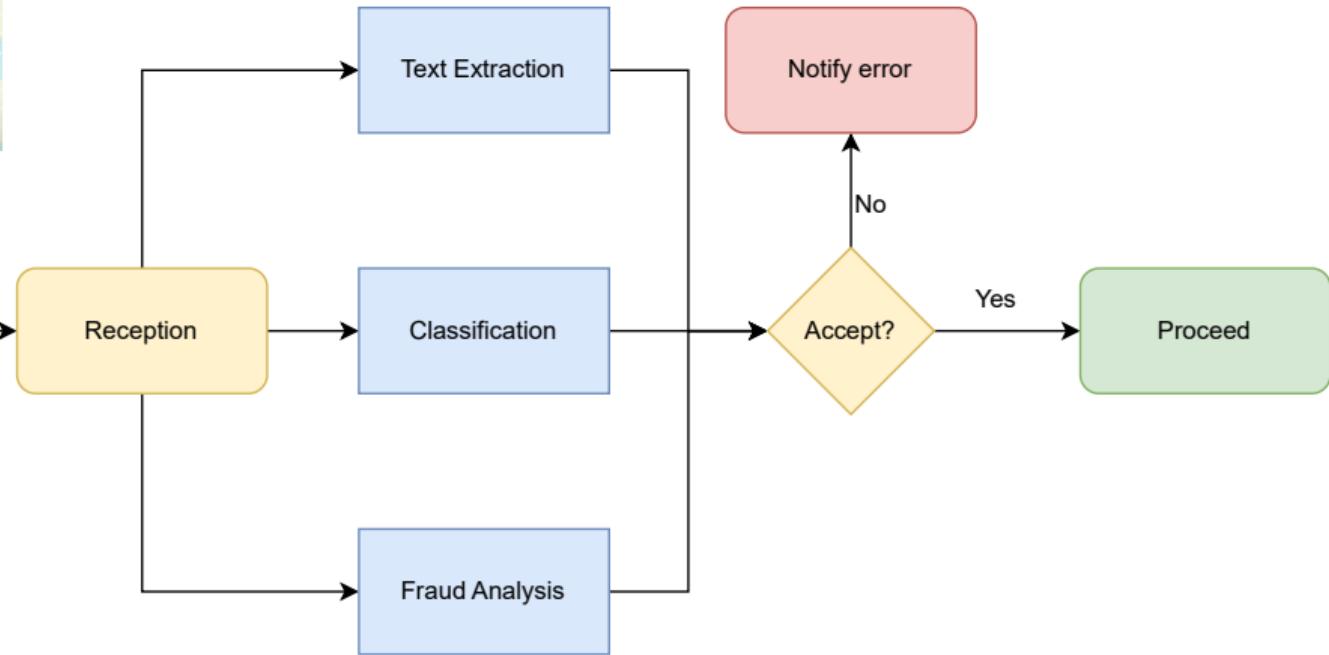


RECEBIMENTO DE ENTREGA DA DECLARAÇÃO DE AJUSTE ANUAL - OPÇÃO PELO DESCONTO SIMPLIFICADO DECLARAÇÃO ORIGINAL			
MINIST&P;RIO DA FAZENDA SECRETARIA DA RECEITA FEDERAL DO BRAS&P;L EXERCIC&P;O 2017 ANO-CALENDÁRIO 2016			
IDENTIFICAÇÃO DO DECLARANTE			
CPF do declarante 352.651.378-08	Nome do declarante ERIKA TOMAZELLA	Telefone (11) 43791221	
Endereço RUA RUA JUSTINO ALVES BATISTA		Número 99	Complemento AP-64 Bl. P
Bairro/Localidade VILA YOLANDA	CEP 06126-120	Município DSASCO	UF/SP SP
(Valores em Reais)			
TOTAL RENDIMENTOS TRIBUTÁVEIS		62.200,42	
IMPOSTO DEVIDO		3.562,59	
IMPOSTO A RESTITUIR		781,16	

Contexto - Fluxo de Compliance



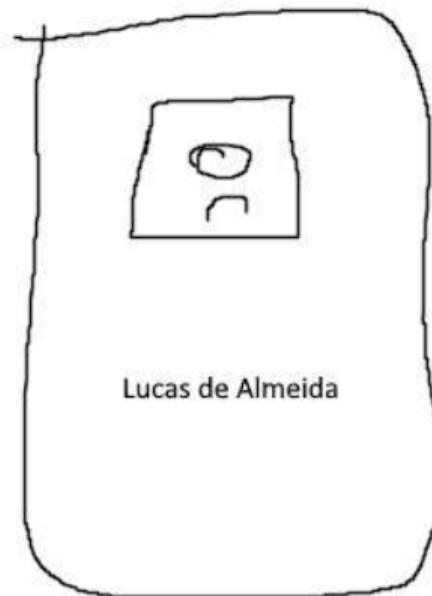
Actor



- Assegurar que o documento está correto
- Documentos não-digitais
- Evita fraudes

Contexto - Classificação de Imagem

- Assegurar que o documento está correto
- Documentos não-digitais
- Evita fraudes



Classificação Tradicional:

- Categorização em classes predefinidas
- Cross-Entropy Loss

Desempenho Atual:

- Bakkali et al. (2021): 97.70% de acurácia no RVL-CDIP

Classificação Tradicional:

- Categorização em classes predefinidas
- Cross-Entropy Loss

Desempenho Atual:

- Bakkali et al. (2021): 97.70% de acurácia no RVL-CDIP

O Problema:

- Novos layouts de documentos
- Classes completamente novas
- Necessidade de retreinamento
- Semanas/meses de engenharia de dados e treinamento

Zero-Shot Learning

Permite que o modelo reconheça elementos de classes nunca vistas no treinamento

Desafios

- Falta de dataset especializado
 - Imagens de Documento
 - Generalização
 - Divisão treino e teste zero-shot
- Ausência de metodologia estado-da-arte
 - Paradigma ZSL
 - Capacidade de classificar

Contribuições

① Novo dataset LA-CDIP

- Classificação ZSL
- Derivado do RVL-CDIP

② Abordagem de Visual Document Matching (VDM)

- Similaridade de documentos
- Metric Learning
- Generalização Zero-Shot

③ Avaliação sistemática

- Benchmark extensivo
- Comparação com LLM

Metodologia

Datasets Disponíveis

- PubLayNet, DocLayNet
- DocVQA
- CORD, SROIE
- RVL-CDIP

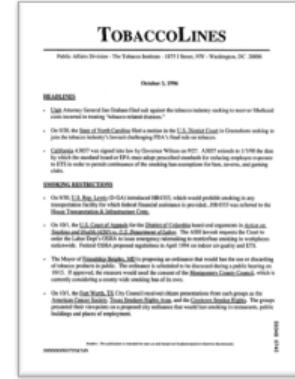
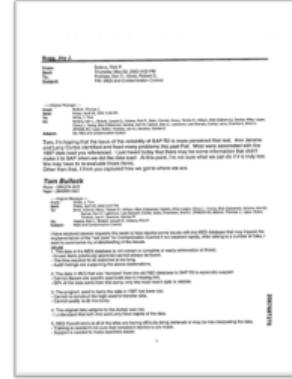
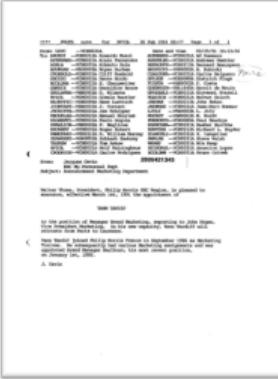
Datasets Disponíveis

- PubLayNet, DocLayNet
- DocVQA
- CORD, SROIE
- RVL-CDIP

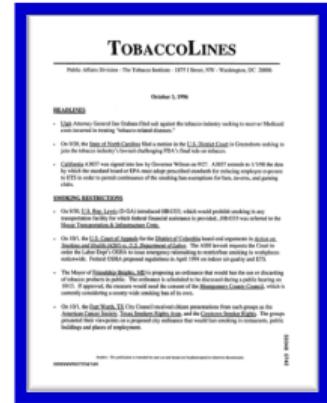
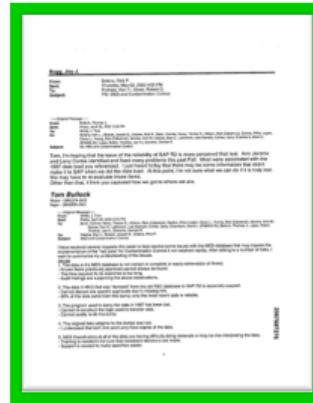
RVL-CDIP

- 400.000 documentos
- 16 classes
- email, formulário, carta...
- Separado por função

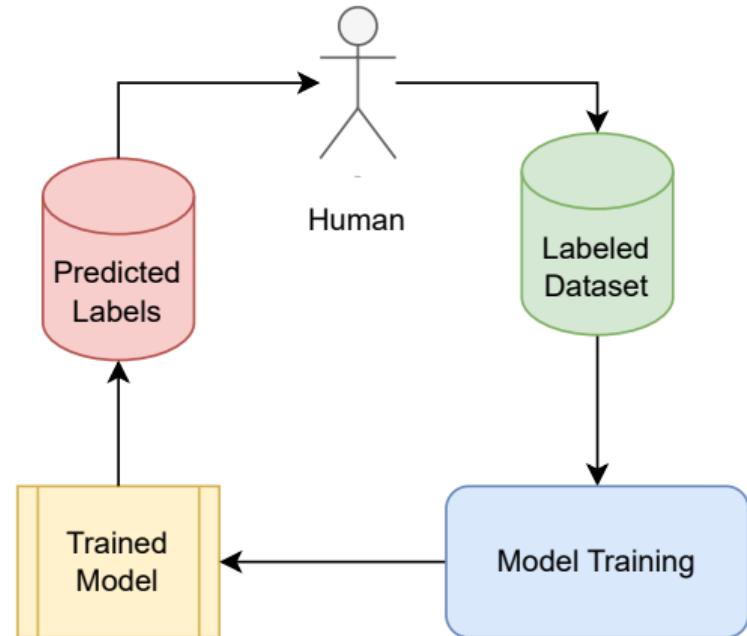
Dataset - Motivação

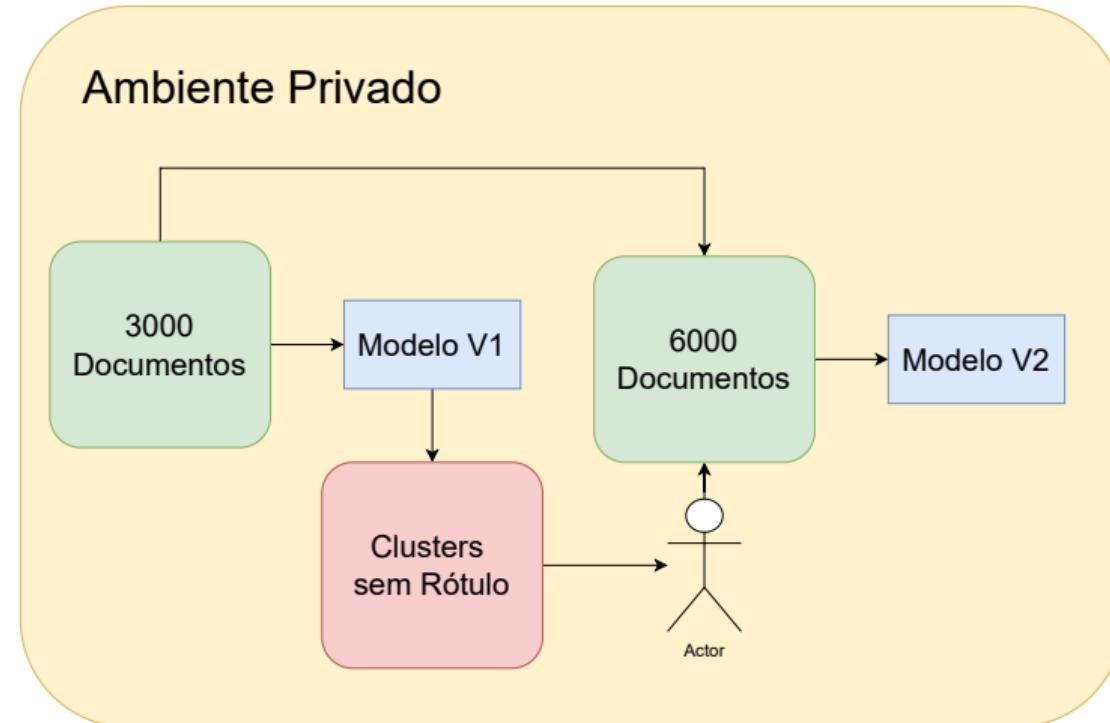


Dataset - Motivação

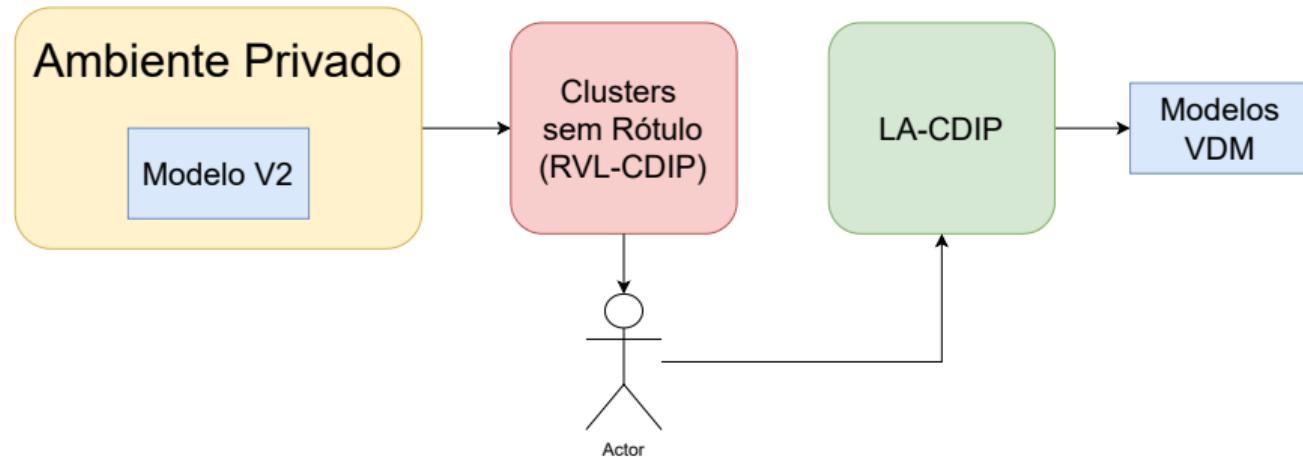


LA-CDIP Dataset - Active Learning

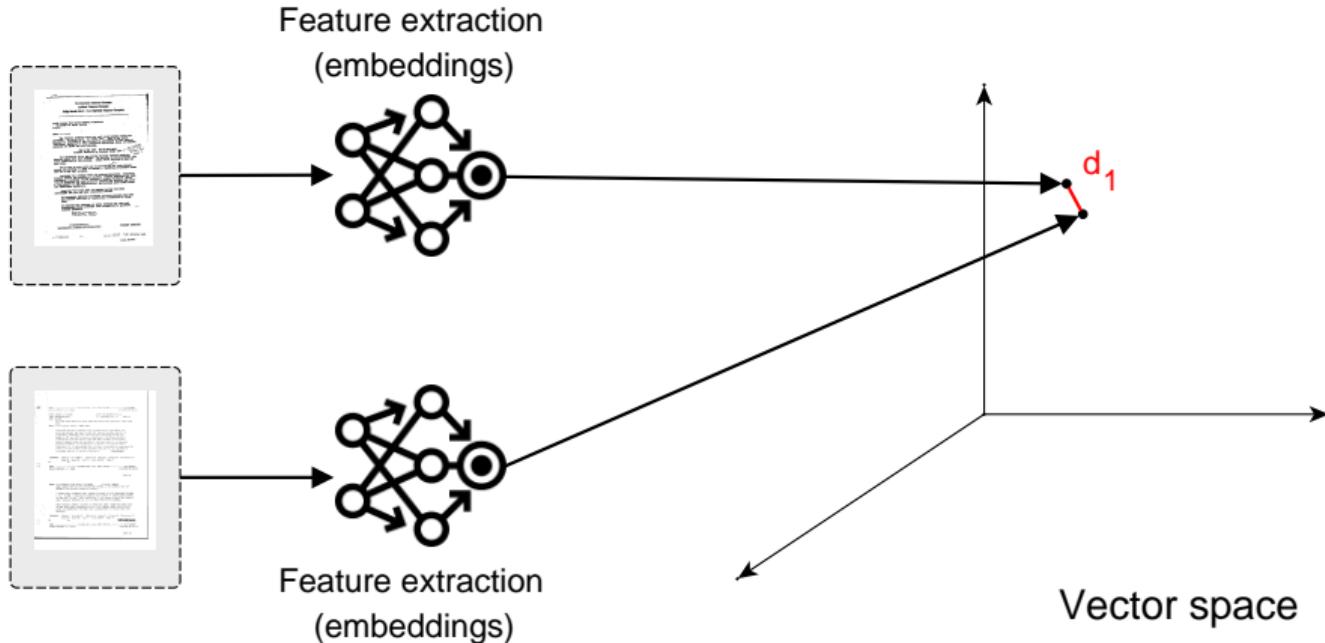




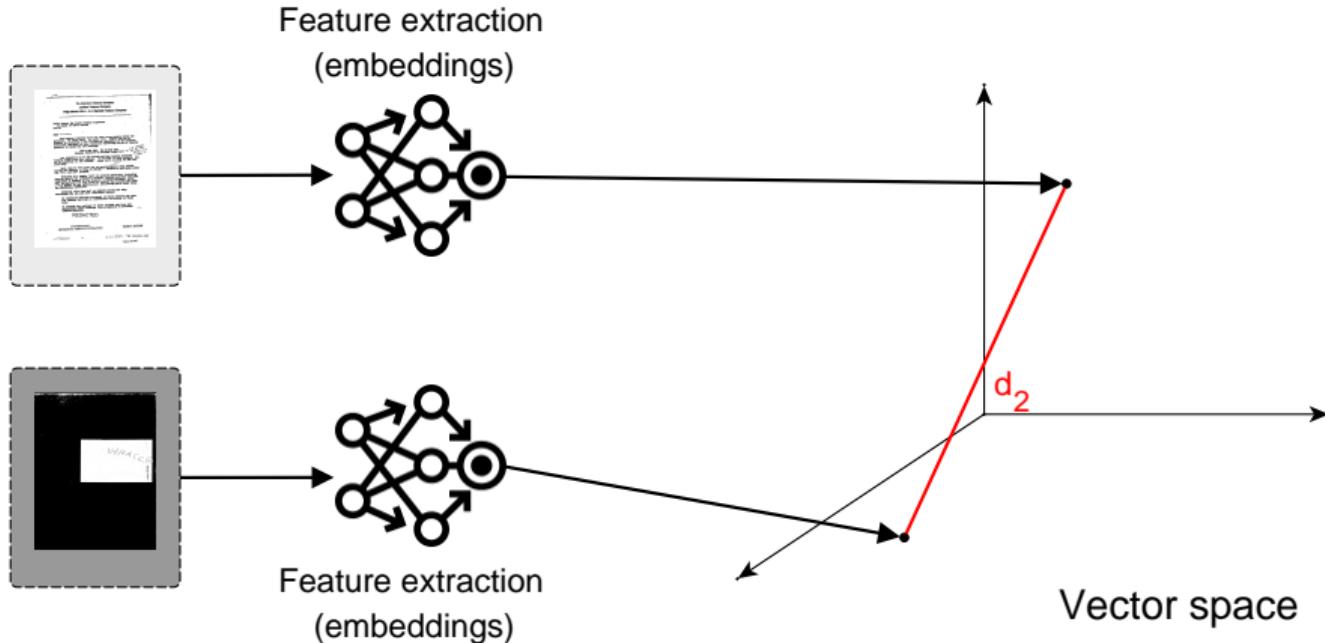
LA-CDIP Dataset - Rotulação



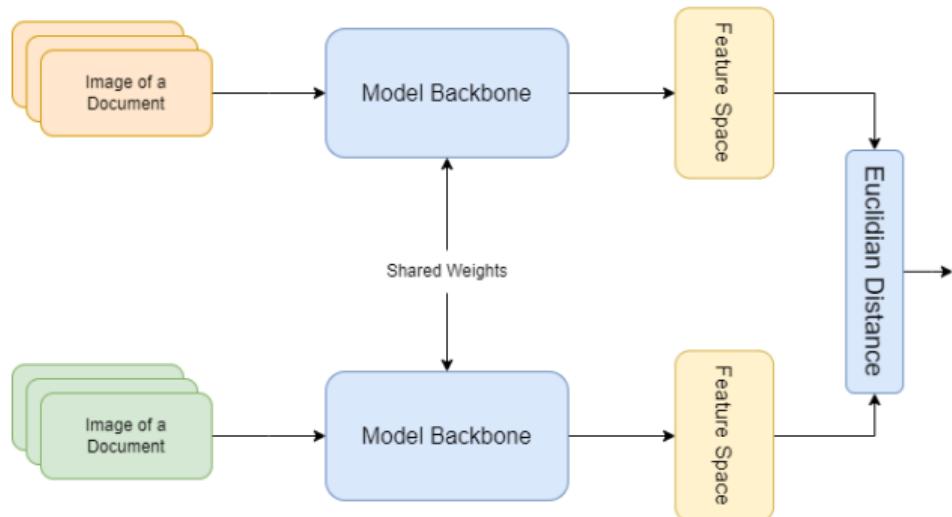
Visual Document Matching



Visual Document Matching



Visual Document Matching - Arquitetura



$$\mathcal{L} = \begin{cases} d(x_1, x_2)^2, & \text{if } y = 1 \\ \max(0, m - d(x_1, x_2))^2, & \text{otherwise.} \end{cases}$$

Visual Document Matching - Backbone

Model Backbone



- AlexNet
- ResNet
- VGG
- EfficientNet
- MobileNetV3
- Vision Transformer (ViT)

Modelos Avaliados:

- LLaVA 3.2 Vision
- InternVL 2.5
- Qwen2.5-VL
- GPT-4o (2024-11-20)
- GPT-4o-mini (2024-07-18)

Avaliação:

- Zero-shot (sem fine-tuning)
- Pontuação de similaridade 0–100
- 5 níveis de categorização

Benchmarking com LLMs

Reference Image

2020 MARLBORO BAR PROGRAM
CONTRACT TOP SHEET

GMM/SSM: Amy Curwick
MARKET: Durles
VENUE NAME: Blackberry's

VENUE ID
DAL-0121-03

Please check the appropriate box that will identify the type of club and the appropriate contract executed by club owner/manager:
 EVENT VISIBILITY MUSIC RNB

Please check the appropriate box regarding venue admission policy:
 AQ18-P AQ21-P PAO-P
 AD18-V AQ21-V

GMM/SSM SIGNATURE: Amy Curwick DATE: 2-3-00

SELL-IN APPROVAL: X DATE: 3-13-00

Similarity Score: 98
Category: Nearly Identical

Image to Compare

2020 MARLBORO BAR PROGRAM
CONTRACT TOP SHEET

GMM/SSM: Lindy Jenkins
MARKET: Charlotte
VENUE NAME: Clock's Sports Bar & Grill

VENUE ID
LHA-0080-01

Please check the appropriate box that will identify the type of club and the appropriate contract executed by club owner/manager:
 EVENT VISIBILITY MUSIC RNB

Please check the appropriate box regarding venue admission policy:
 AQ18-P AQ21-P PAO-P
 AD18-V AQ21-V

GMM/SSM SIGNATURE: Lindy Jenkins DATE: 2-20-00

SELL-IN APPROVAL: J DATE: 3-13-00

Similarity Score: 98
Category: Nearly Identical

Reference Image

GMM/SSM	Signature	Market	Venue Name	Venue ID	Contract Type		Admission Policy	Comments
					Event	Visibility		
John Doe	John Doe Signature	New York	Bar 1	NYB-001	✓	✓	PAO-P	Approved
Jane Smith	Jane Smith Signature	Los Angeles	Bar 2	LA-002	✓	✓	PAO-P	Approved
Mike Johnson	Mike Johnson Signature	Chicago	Bar 3	CHI-003	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Houston	Bar 4	HOU-004	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	Phoenix	Bar 5	PHX-005	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Seattle	Bar 6	SEA-006	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Boston	Bar 7	BOS-007	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Philadelphia	Bar 8	PHL-008	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Minneapolis	Bar 9	MPL-009	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	San Antonio	Bar 10	SAN-010	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Tampa	Bar 11	TAM-011	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	Nashville	Bar 12	NASH-012	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Portland	Bar 13	PTL-013	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 14	ATL-014	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 15	STL-015	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 16	ORL-016	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 17	PHX-017	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 18	SEA-018	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 19	BOS-019	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 20	PHL-020	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 21	MPL-021	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 22	SAN-022	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 23	TAM-023	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 24	ATL-024	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 25	STL-025	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 26	ORL-026	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 27	PHX-027	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 28	SEA-028	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 29	BOS-029	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 30	PHL-030	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 31	MPL-031	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 32	SAN-032	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 33	TAM-033	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 34	ATL-034	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 35	STL-035	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 36	ORL-036	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 37	PHX-037	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 38	SEA-038	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 39	BOS-039	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 40	PHL-040	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 41	MPL-041	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 42	SAN-042	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 43	TAM-043	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 44	ATL-044	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 45	STL-045	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 46	ORL-046	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 47	PHX-047	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 48	SEA-048	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 49	BOS-049	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 50	PHL-050	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 51	MPL-051	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 52	SAN-052	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 53	TAM-053	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 54	ATL-054	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 55	STL-055	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 56	ORL-056	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 57	PHX-057	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 58	SEA-058	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 59	BOS-059	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 60	PHL-060	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 61	MPL-061	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 62	SAN-062	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 63	TAM-063	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 64	ATL-064	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 65	STL-065	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 66	ORL-066	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 67	PHX-067	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 68	SEA-068	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 69	BOS-069	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 70	PHL-070	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 71	MPL-071	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 72	SAN-072	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 73	TAM-073	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 74	ATL-074	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 75	STL-075	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 76	ORL-076	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 77	PHX-077	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 78	SEA-078	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 79	BOS-079	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 80	PHL-080	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 81	MPL-081	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 82	SAN-082	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 83	TAM-083	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 84	ATL-084	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 85	STL-085	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 86	ORL-086	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 87	PHX-087	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 88	SEA-088	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 89	BOS-089	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 90	PHL-090	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Minneapolis	Bar 91	MPL-091	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	San Antonio	Bar 92	SAN-092	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	Tampa	Bar 93	TAM-093	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Atlanta	Bar 94	ATL-094	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	St. Louis	Bar 95	STL-095	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Orlando	Bar 96	ORL-096	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Phoenix	Bar 97	PHX-097	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Seattle	Bar 98	SEA-098	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Boston	Bar 99	BOS-099	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	Philadelphia	Bar 100	PHL-100	✓	✓	PAO-P	Approved

Similarity Score: 98
Category: Nearly Identical

GMM/SSM	Signature	Market	Venue Name	Venue ID	Contract Type		Admission Policy	Comments
					Event	Visibility		
John Doe	John Doe Signature	New York	Bar 1	NYB-001	✓	✓	PAO-P	Approved
Jane Smith	Jane Smith Signature	Los Angeles	Bar 2	LA-002	✓	✓	PAO-P	Approved
Mike Johnson	Mike Johnson Signature	Chicago	Bar 3	CHI-003	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Houston	Bar 4	HOU-004	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	Phoenix	Bar 5	PHX-005	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Seattle	Bar 6	SEA-006	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Boston	Bar 7	BOS-007	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Philadelphia	Bar 8	PHL-008	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Minneapolis	Bar 9	MPL-009	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	San Antonio	Bar 10	SAN-010	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Tampa	Bar 11	TAM-011	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	Atlanta	Bar 12	ATL-012	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	St. Louis	Bar 13	STL-013	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Orlando	Bar 14	ORL-014	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	Phoenix	Bar 15	PHX-015	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Seattle	Bar 16	SEA-016	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Boston	Bar 17	BOS-017	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Philadelphia	Bar 18	PHL-018	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Minneapolis	Bar 19	MPL-019	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	San Antonio	Bar 20	SAN-020	✓	✓	PAO-P	Approved
Christopher White	Christopher White Signature	Tampa	Bar 21	TAM-021	✓	✓	PAO-P	Approved
Stephanie Jones	Stephanie Jones Signature	Atlanta	Bar 22	ATL-022	✓	✓	PAO-P	Approved
Matthew Davis	Matthew Davis Signature	St. Louis	Bar 23	STL-023	✓	✓	PAO-P	Approved
Sarah Williams	Sarah Williams Signature	Orlando	Bar 24	ORL-024	✓	✓	PAO-P	Approved
David Lee	David Lee Signature	Phoenix	Bar 25	PHX-025	✓	✓	PAO-P	Approved
Emily Green	Emily Green Signature	Seattle	Bar 26	SEA-026	✓	✓	PAO-P	Approved
Mark Wilson	Mark Wilson Signature	Boston	Bar 27	BOS-027	✓	✓	PAO-P	Approved
Karen Clark	Karen Clark Signature	Philadelphia	Bar 28	PHL-028	✓	✓	PAO-P	Approved
Jeffrey Brown	Jeffrey Brown Signature	Minneapolis	Bar 29	MPL-029	✓	✓	PAO-P	Approved
Michelle Taylor	Michelle Taylor Signature	San Antonio	Bar 30	SAN-030	✓	✓	PAO-P	Approved

Similarity Score: 15
Category: Completely Different

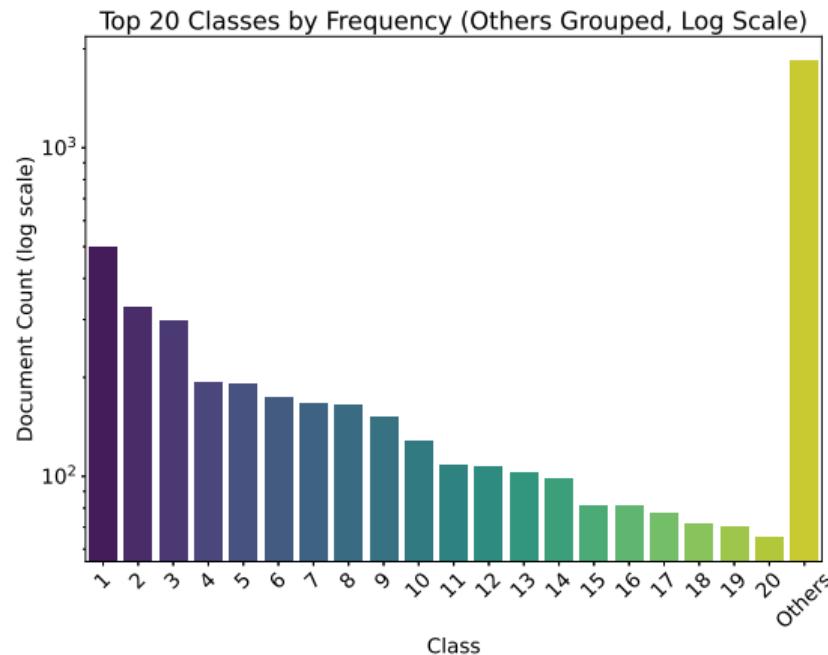
Resultados

Composição:

- 4.993 documentos
- 144 classes diferentes
- Min: 2 documentos/classe
- Max: 497 documentos/classe
- Mediana: 13 documentos/classe

Splits:

- ZSL: separação completa treino/teste
- GZSL: 50% overlap de classes
- 5-fold cross-validation



Equal Error Rate (EER)

- Ponto onde FAR = FRR
- FAR: False Acceptance Rate
- FRR: False Rejection Rate

$$FAR(\tau) = \frac{\text{False Acceptances}}{\text{Total Negatives}}$$

$$FRR(\tau) = \frac{\text{False Rejections}}{\text{Total Positives}}$$

Protocolo de Teste:

Para cada documento: 1 par similar + 1 par dissimilar

Resultados - Visão Geral

Architecture	Edition	Params	ZSL	GZSL	Test ZSL	Test GZSL
AlexNet		57M	8.92	5.45	17.33	6.31
VGG	11	129M	7.47	5.01	14.24	3.95
	13	129M	7.03	4.79	9.30	3.95
	16	134M	8.29	5.23	14.74	4.82
	19	139M	7.30	4.57	17.08	3.90
ResNet	18	11M	5.03	<u>1.54</u>	4.98	1.51
	34	21M	4.32	<u>2.10</u>	4.13	1.53
	50	23M	6.90	3.39	10.34	2.21
	101	42M	8.20	2.72	11.31	1.98
	152	58M	9.44	3.38	12.70	2.39
MobileNetV3	Small	1M	7.98	5.06	12.74	5.26
	Large	4M	8.16	4.27	8.45	4.43
EfficientNet	0	4M	4.41	2.27	6.02	<u>0.95</u>
	1	6M	<u>3.93</u>	3.54	8.88	2.70
	2	7M	<u>5.73</u>	2.61	7.29	2.14
	3	10M	5.65	3.64	7.37	2.34
ViT	Base	87M	12.43	7.97	19.72	5.19
	Large	305M	13.16	7.57	19.88	5.26
Llama	3.2	11B	—	—	13.95	21.90
InternVL	2.5	8B	—	—	8.58	10.40
Qwen-VL	2.5	7B	—	—	6.61	4.20
GPT-4o mini	2024-07-18	*	—	—	4.70	4.07
GPT-4o	2024-11-20	*	—	—	<u>2.75</u>	<u>1.33</u>

Resultados - ResNet e EfficientNet

Architecture	Edition	Params	ZSL	GZSL	Test ZSL	Test GZSL
ResNet	18	11M	5.03	<u>1.54</u>	4.98	1.51
	34	21M	4.32	2.10	<u>4.13</u>	1.53
	50	23M	6.90	3.39	10.34	2.21
	101	42M	8.20	2.72	11.31	1.98
	152	58M	9.44	3.38	12.70	2.39
EfficientNet	0	4M	4.41	2.27	6.02	<u>0.95</u>
	1	6M	<u>3.93</u>	3.54	8.88	2.70
	2	7M	5.73	2.61	7.29	2.14
	3	10M	5.65	3.64	7.37	2.34

Resultados - ResNet, EfficientNet e ViT

Architecture	Edition	Params	ZSL	GZSL	Test ZSL	Test GZSL
ResNet	18	11M	5.03	<u>1.54</u>	4.98	1.51
	34	21M	4.32	2.10	<u>4.13</u>	1.53
	50	23M	6.90	3.39	10.34	2.21
	101	42M	8.20	2.72	11.31	1.98
	152	58M	9.44	3.38	12.70	2.39
EfficientNet	0	4M	4.41	2.27	6.02	<u>0.95</u>
	1	6M	<u>3.93</u>	3.54	8.88	2.70
	2	7M	5.73	2.61	7.29	2.14
	3	10M	5.65	3.64	7.37	2.34
ViT	Base	87M	12.43	7.97	19.72	5.19
	Large	305M	13.16	7.57	19.88	5.26

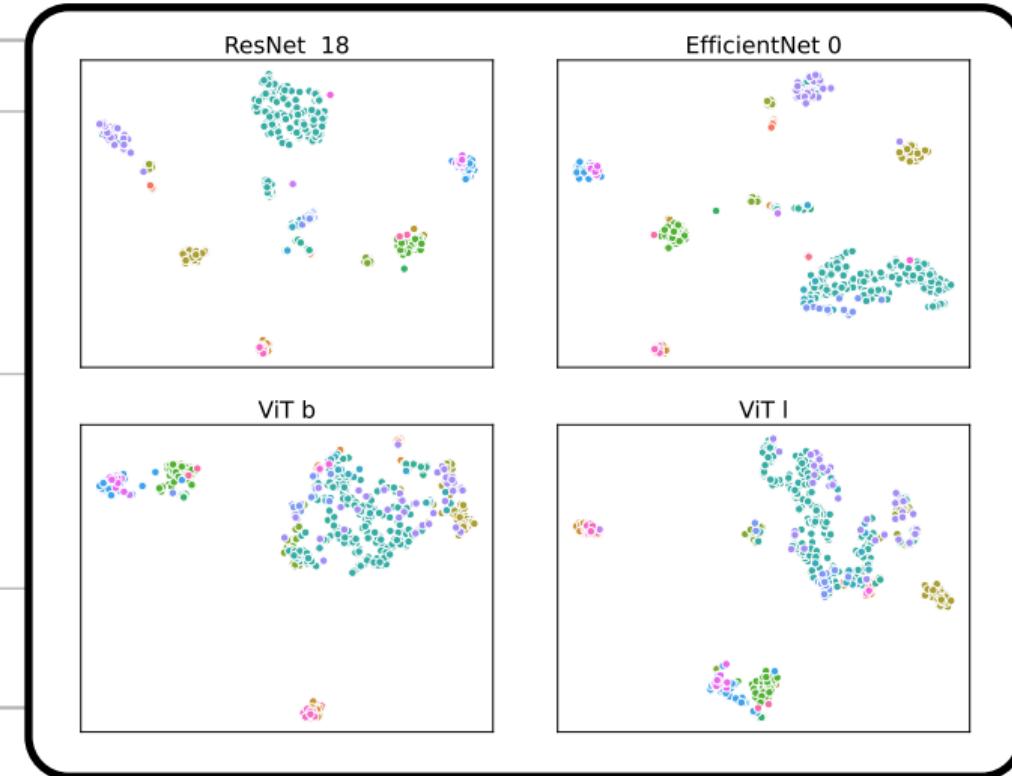
Resultados - Visualização TSNE

Architecture

ResNet

EfficientNet

ViT



Test GZSL

1.51
1.53
2.21
1.98
2.39

0.95
2.70
2.14
2.34

5.19
5.26

Resultados - Large Language Models

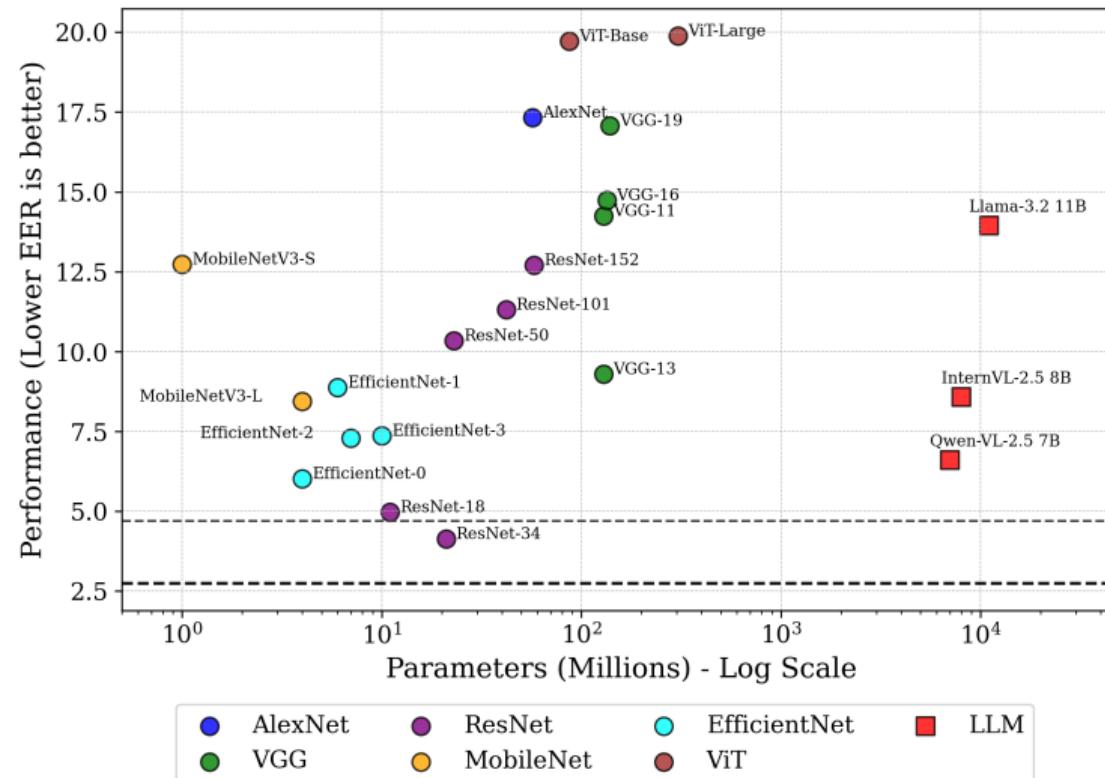
Model	Version	Params	Test ZSL	Test GZSL
Llama	3.2	11B	13.95	21.90
InternVL	2.5	8B	8.58	10.40
Qwen-VL	2.5	7B	6.61	4.20
GPT-4o mini	2024-07-18	*	4.70	4.07
GPT-4o	2024-11-20	*	<u>2.75</u>	<u>1.33</u>

* Parameter count not publicly disclosed

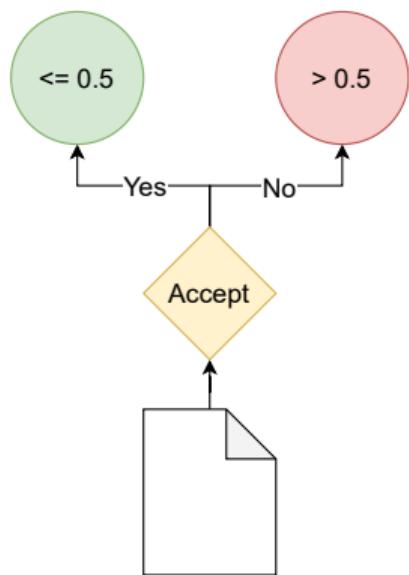
Resultados - Melhores Modelos Visuais vs LLMs

Architecture	Edition	Params	ZSL	GZSL	Test ZSL	Test GZSL
ResNet	18	11M	5.03	1.54	4.98	1.51
	34	21M	4.32	2.10	4.13	1.53
	50	23M	6.90	3.39	10.34	2.21
	101	42M	8.20	2.72	11.31	1.98
	152	58M	9.44	3.38	12.70	2.39
EfficientNet	0	4M	4.41	2.27	6.02	0.95
	1	6M	3.93	3.54	8.88	2.70
	2	7M	5.73	2.61	7.29	2.14
	3	10M	5.65	3.64	7.37	2.34
Llama	3.2	11B	–	–	13.95	21.90
InternVL	2.5	8B	–	–	8.58	10.40
Qwen-VL	2.5	7B	–	–	6.61	4.20
GPT-4o mini	2024-07-18	*	–	–	4.70	4.07
GPT-4o	2024-11-20	*	–	–	2.75	1.33

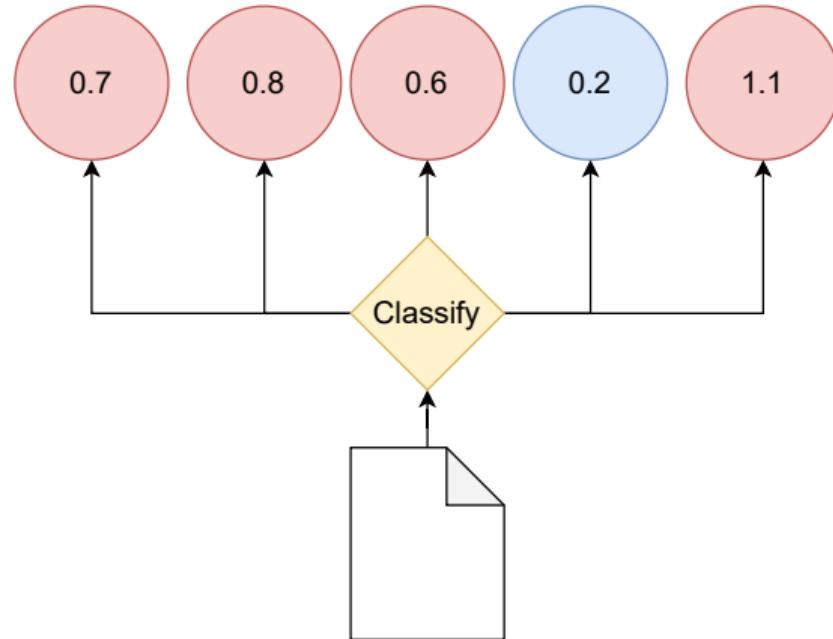
Comparação: Visual Models vs LLMs



Comparação: Visual Models vs LLMs



Sistema de Verificação



Sistema de Identificação

Conclusão

① LA-CDIP Dataset

Categorizado por layout visual para Zero-Shot Learning

② Metodologia com VDM

Enforçando comparação de documentos como forma de classificação

③ Benchmarking Sistemático

Comparação entre modelos visuais e LLMs

Limitações Conhecidas:

- Dataset relativamente pequeno
- Complexidade de arquiteturas limitada
- Apenas documentos do RVL-CDIP

Trabalhos Futuros:

- Aumentar número de amostras
- Aumentar número de classes
- Incluir fontes adicionais de documentos

Obrigado!

Lucas de Almeida Bandeira Macedo

lucasabmacedo@hotmail.com

Orientador: Prof. Dr. Pedro Garcia Freitas

Coorientador: Prof. Dr. Bruno Luiggi Macchiavello Espinoza