CBPF
Centro Brasileiro de Pesquisas Físicas
UNIDADE DE PESQUISA DO MCTI

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

Aplicações de IA em Astrofísica e Desenvolvimentos com a Indústria de O&G

Luciana Olivia Dias

Clécio R. De Bom, Elisangela L. Faria e Marcelo Albuquerque

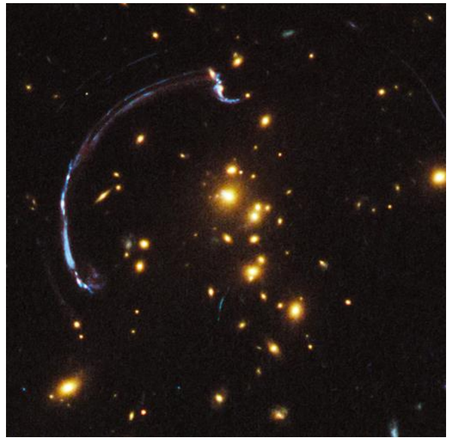
O2I

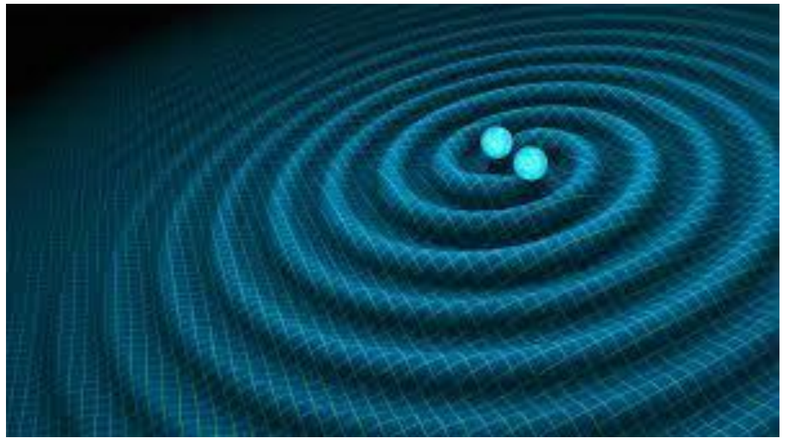
16-18 Novembro 2022, CBPF, Rio de Janeiro

Um evento que mostra o potencial da ciência na sociedade moderna e estimula sinergias entre pesquisa fundamental e indústria

Astrofísica e Deep Learning

- A Astrofísica apresenta vários desafios em termos de **V**olume de dados, **V**ariety, **V**elocity (3 Vs em Big Data).
- Também compreende um conjunto de dados heterogêneos e incompletos (complexos).
- O Machine Learning (e, em particular, a expertise do nosso grupo Deep Learning) pode atuar em diversas situações como:
- Pesquisa, Classificação, Modelagem.
- Tais tarefas também estão presentes nos desafios do setor!!





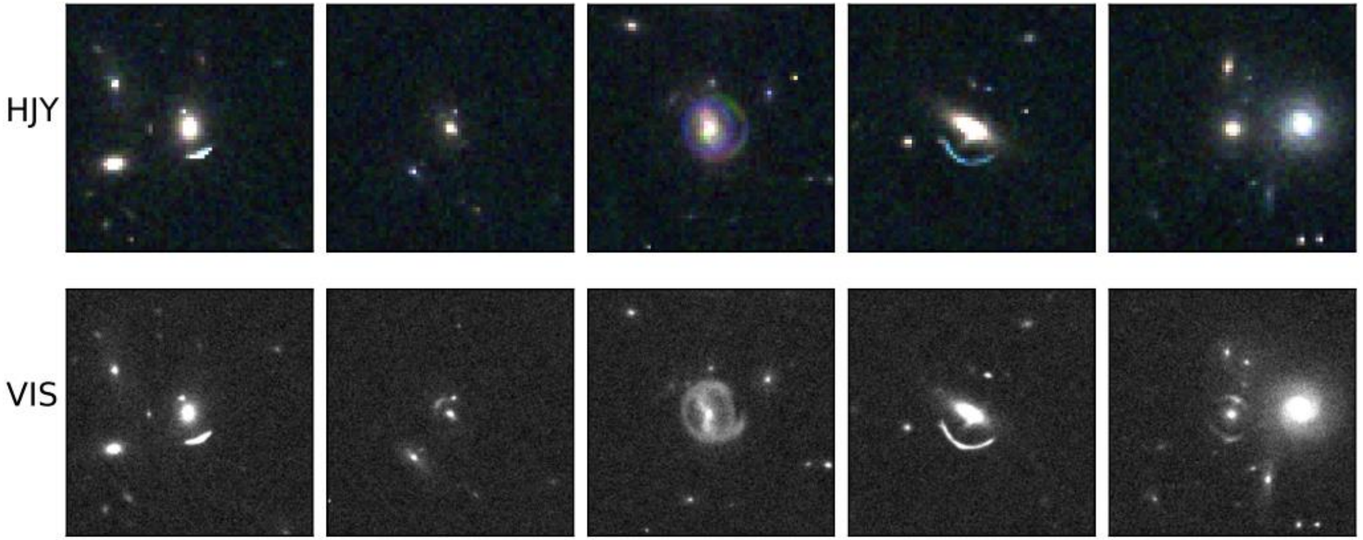
Busca de Lentes Gravitacionais

O Desafio:

- Classificar 100k imagens usando até quatro canais (VIS, NISP J, Y e H – euclid-like).
- Testar o algoritmo com imagens 100k simuladas que contém todos os tipos de problemas no sistema de imagem.
- Não havia informações sobre como as imagens foram simuladas.
- Cada equipe desenvolveu algoritmos diferentes, principalmente baseados no Deep Learning.

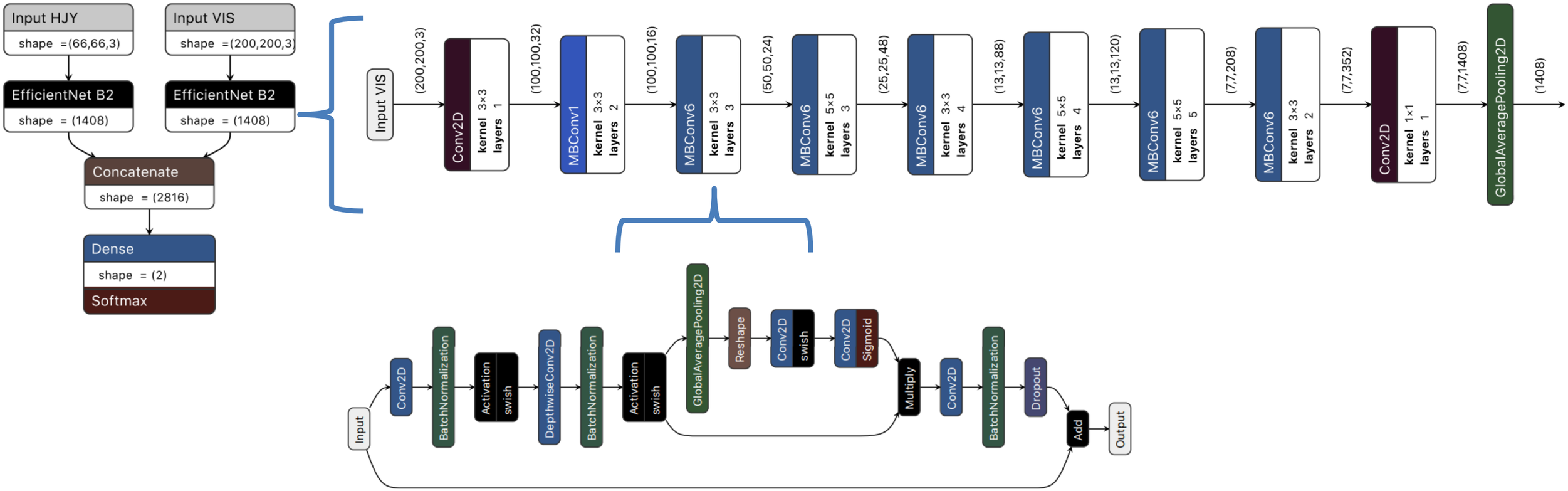
Por que precisamos encontrar lentes gravitacionais?

- Distribuição de matéria (escura) em regiões de aglomerados internos
- Telescópios gravitacionais para investigar galáxias fracas em high RedShift
- Testes de Relatividade Geral de Einsten
- Sonda cosmológica



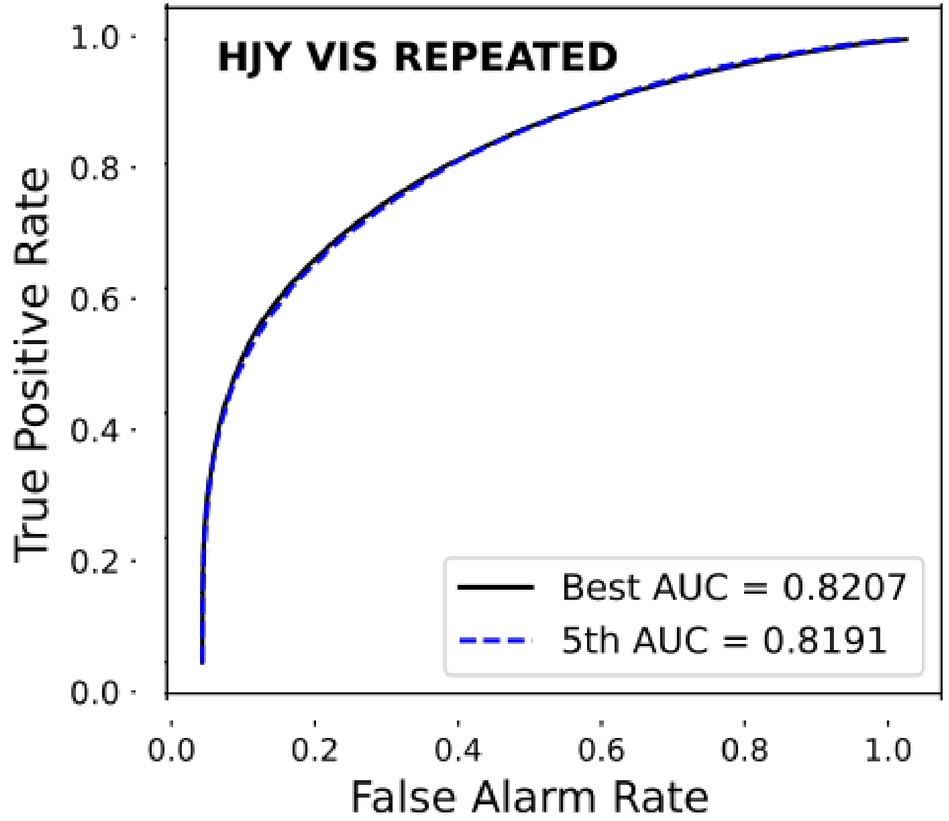
Busca de Lentes Gravitacionais

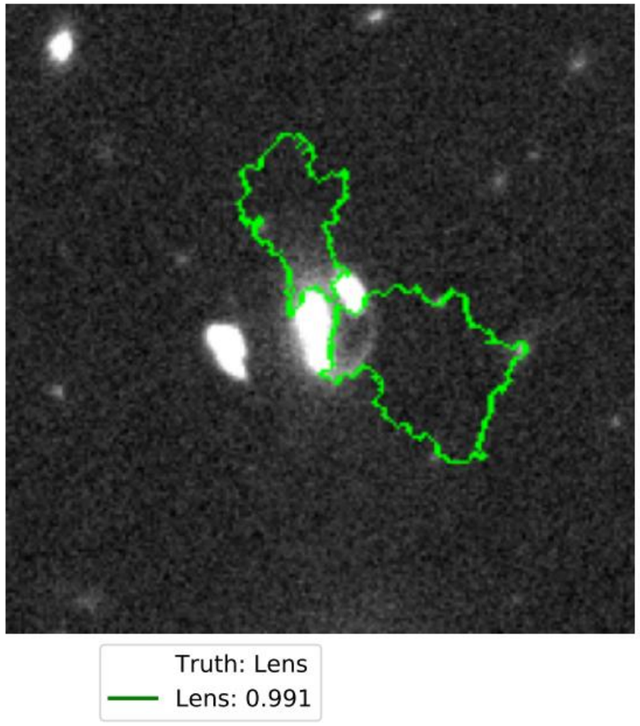
- Nossa solução integra imagens com resoluções diferentes (banda VIS com maior resolução e HJY com menor resolução).
- Nós concatenamos as imagens dentro de um algoritmo de rede neural profunda de última geração para visão computacional – arquitetura de DL EfficientNet



Busca de Lentes Gravitacionais

- Nossa solução integra imagens com resoluções diferentes (banda VIS com maior resolução e HJY com menor resolução).
- Nós concatenamos as imagens dentro de um algoritmo de rede neural profunda de última geração para visão computacional – arquitetura de DL EfficientNet



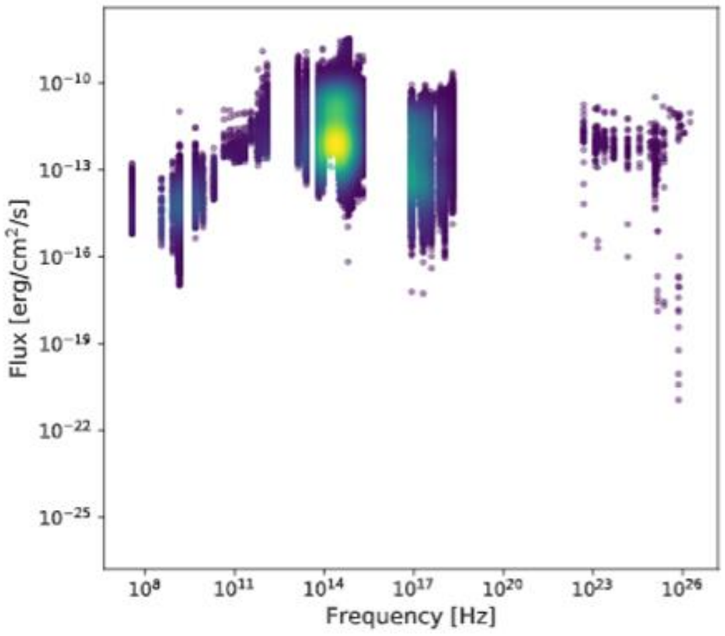


Bom et al; MNRAS; 2022

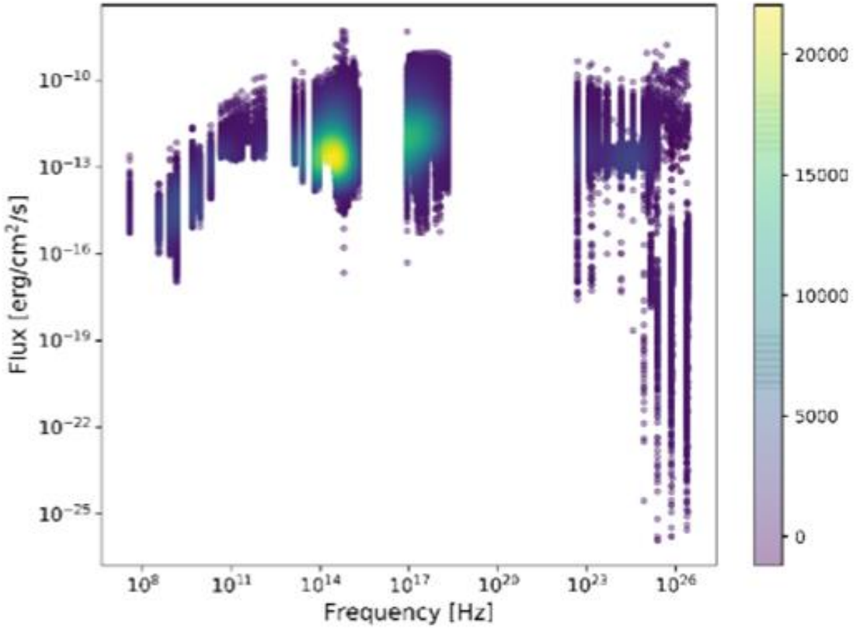
Classificação de Blazares/AGN

- Um Blazar é um núcleo galáctico ativo (AGN) com um jato relativístico direcionado muito próximo a um observador. Sua identificação padrão é um procedimento manual, geralmente depende de cobertura heterogênea de múltiplos comprimentos de onda.
- Esta é uma classificação sequência (curva) (fluxo x frequência), sem imagens. Do rádio aos raios gama.

Artist's impression



(a) Non-Blazars.

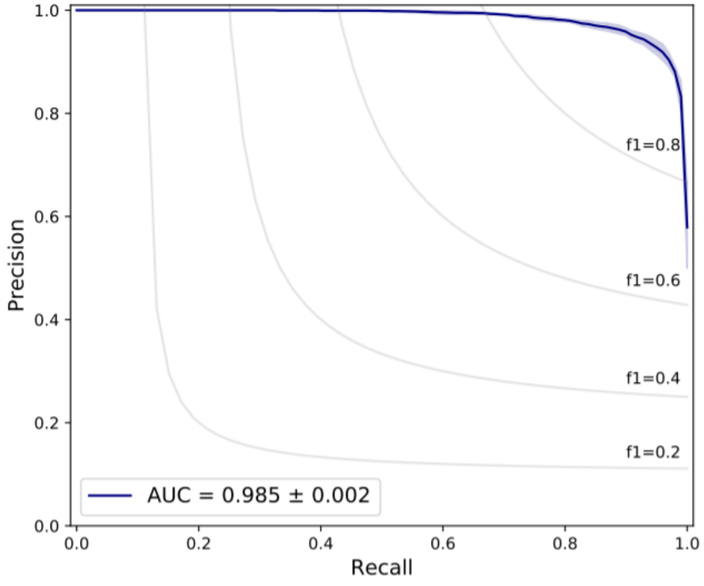


(b) Blazars

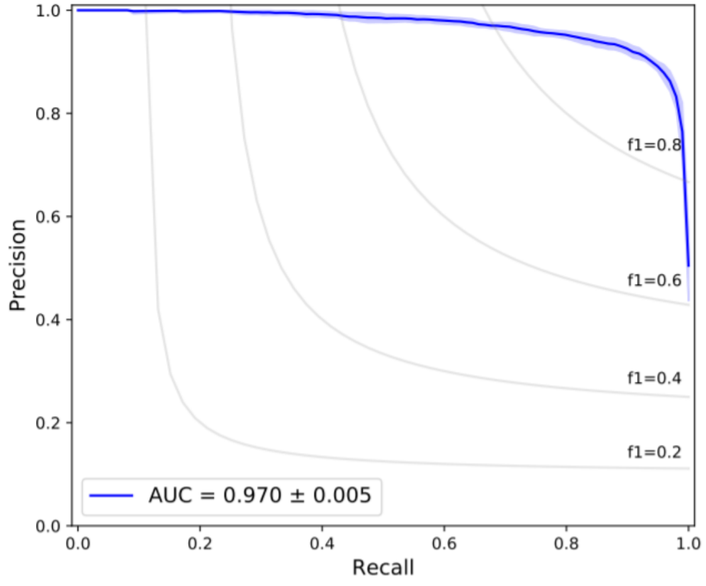
Fraga et al.; "***Deep learning Blazar classification based on multifrequency spectral energy distribution data***", MNRAS; 2021

Classificação de Blazares/AGN

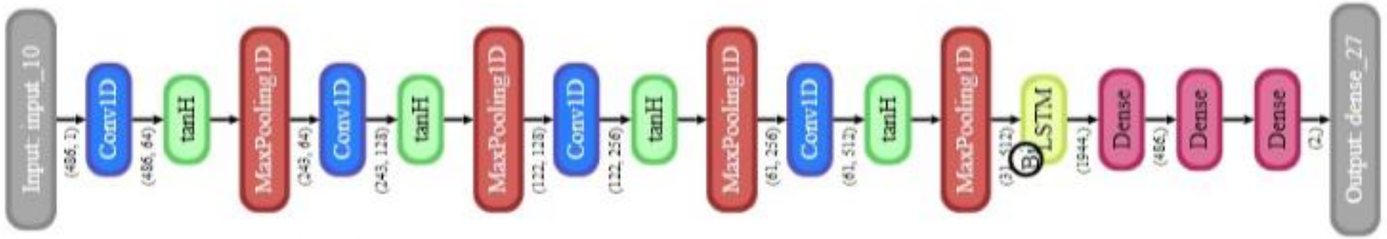
- Nossa solução neste problema de classificação de curvas foi usar redes que funcionam em sequências, como as usadas na tradução de texto; uma rede neural recorrente. Mais especificamente, uma memória de longo prazo (LSTM).



(a) With γ-ray data.



(b) No γ-ray data.

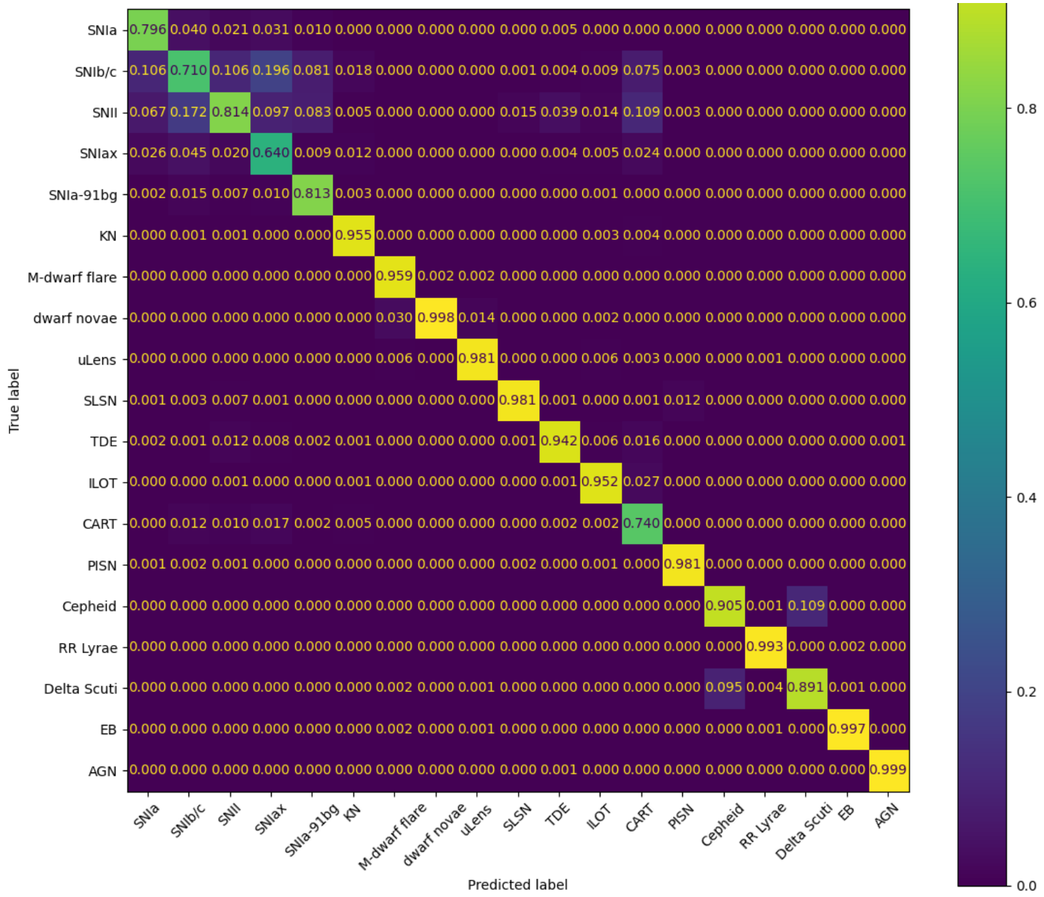


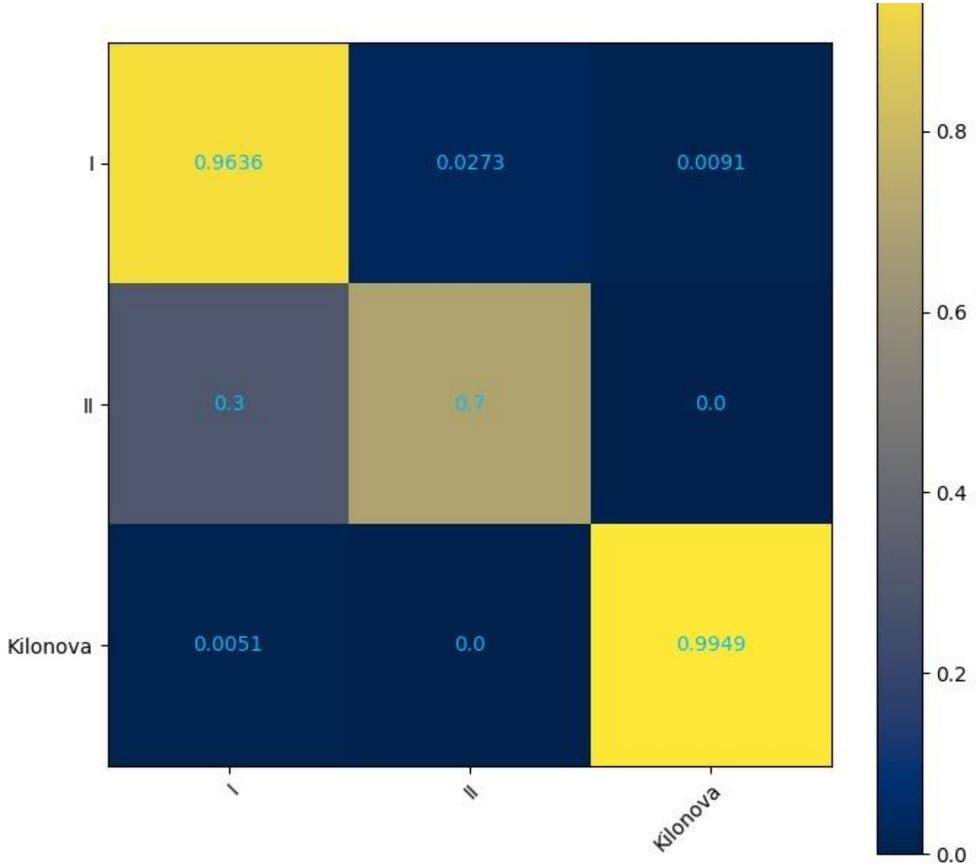
Fraga et al.; "***Deep learning Blazar classification based on multifrequency spectral energy distribution data***", MNRAS; 2021

Classificação de Transientes do LSST

(*) Legacy Survey of Space and Time

- Agora estamos usando várias soluções de Deep Learning para classificar não SED (fluxo x frequência), mas também curvas ópticas de luz (magnitude x dias) dos Alertas Vera Rubin. Incluindo variabilidade AGN e objetos que podem ser as fontes de eventos de ondas gravitacionais, como Kilonovas

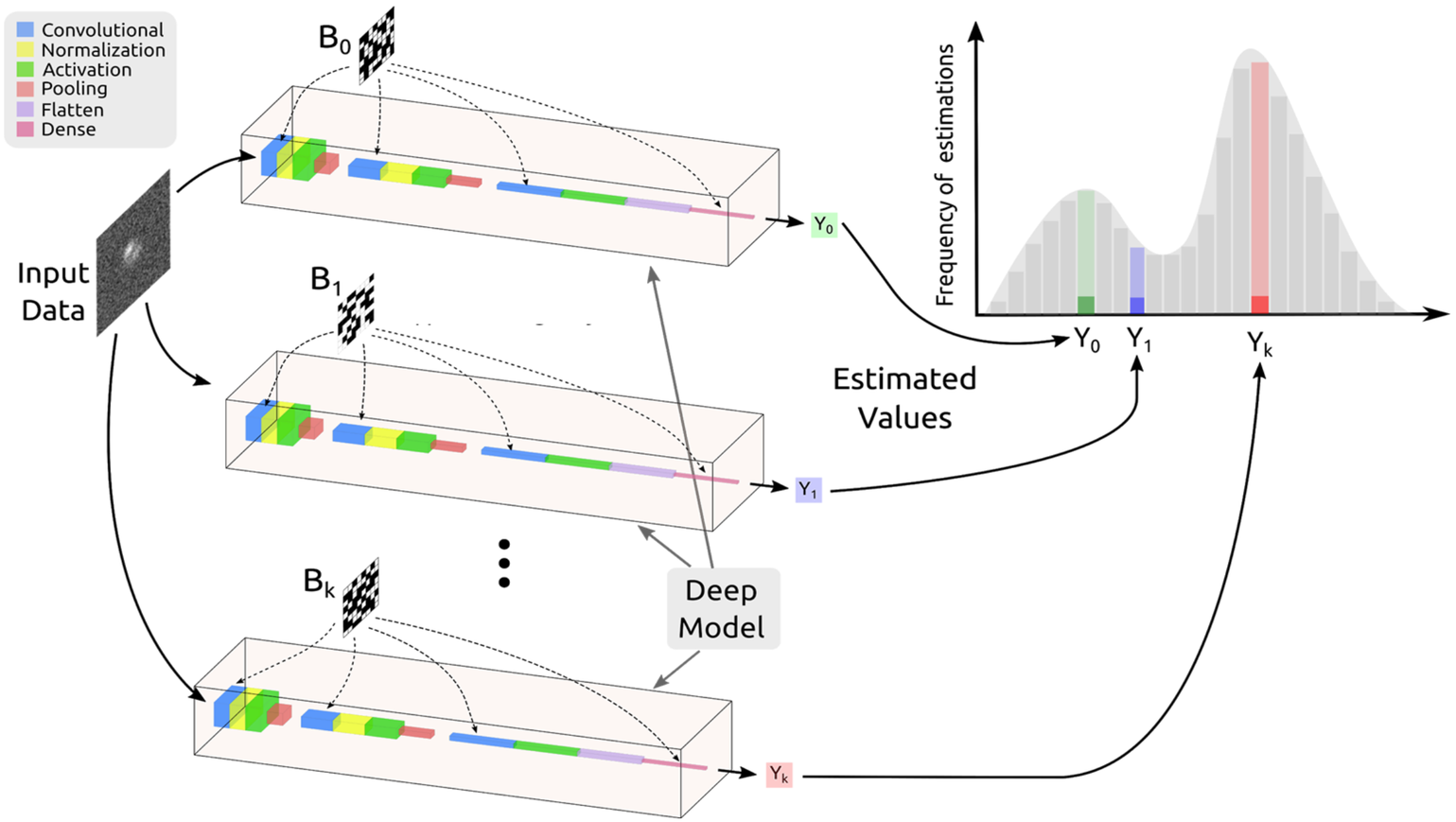


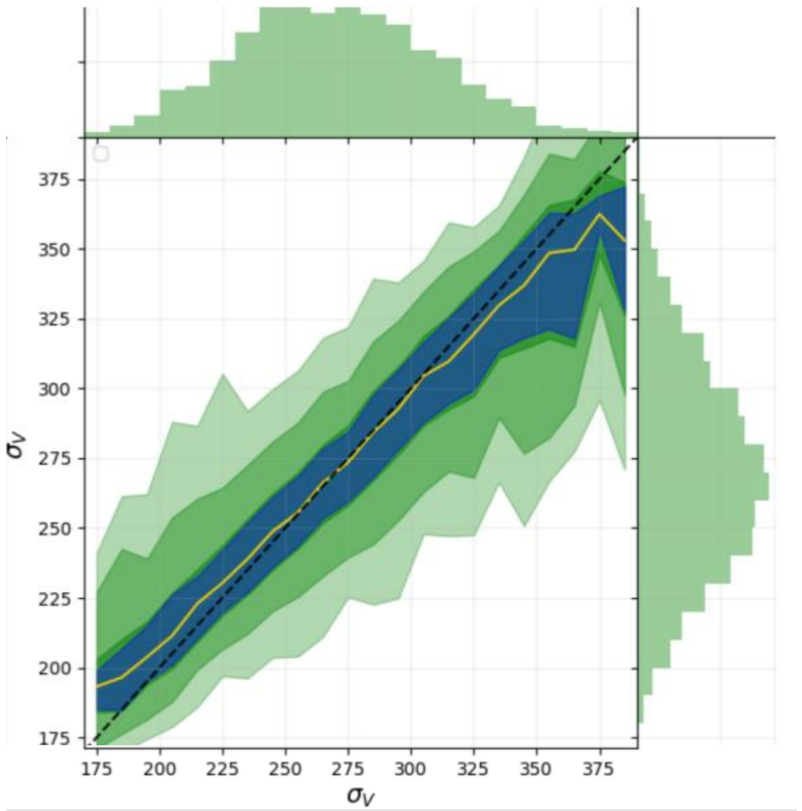


PRELIMINARY

Modelando Lentes Gravitacionais

- Usando o Deep Learning Probabilístico (Normalização de Fluxos, Aprendizado Profundo Bayesiano), também modelamos sistemas físicos como Lentes Fortes.





1,2 and 3 sigma of velocity dispersion. A quantity directly connected with mass distribution . For a simple SL model:

$$\rho(\tilde{r}) = \frac{\sigma_V{}^2}{2\pi G\tilde{r}^2}.$$

Cooperação em P&D com a Petrobras

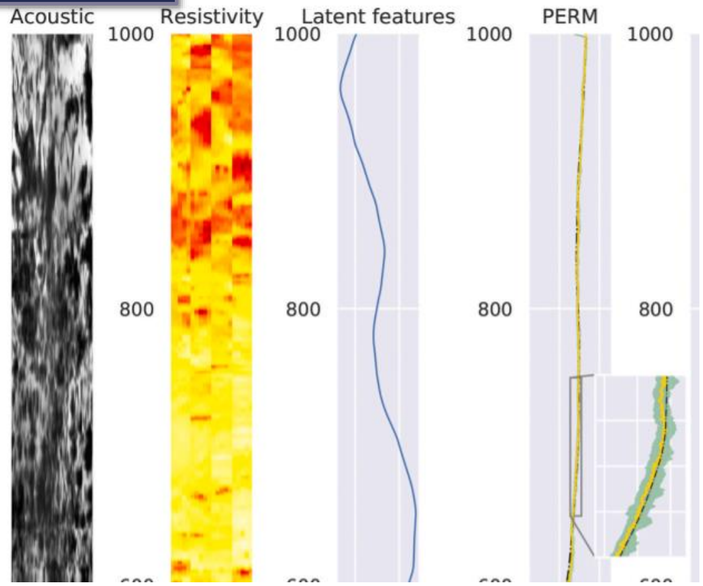
I. Inteligência Artificial em Perfis de Imagem, Dados Sísmicos e Petrográficos

II. Simulação de estrutura poros e NMR

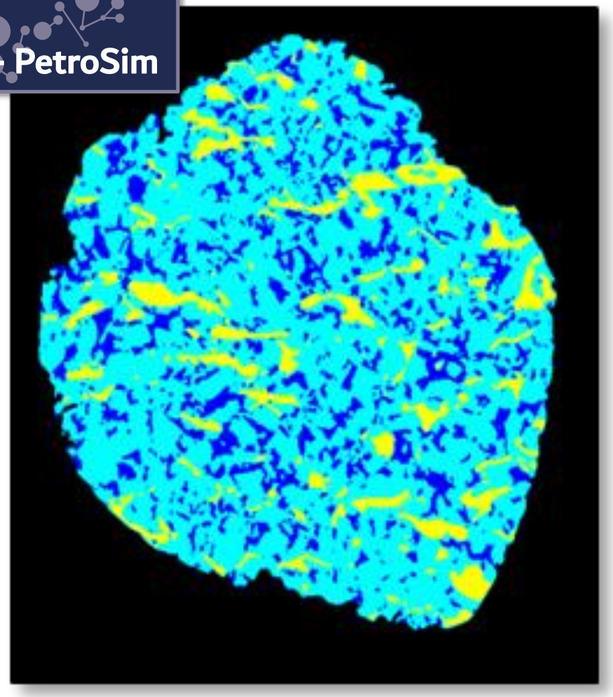
III. Instrumentação e HPC com multiGPU

Equipe dedicada à pesquisa e inovação no CBPF

I - PetroIA

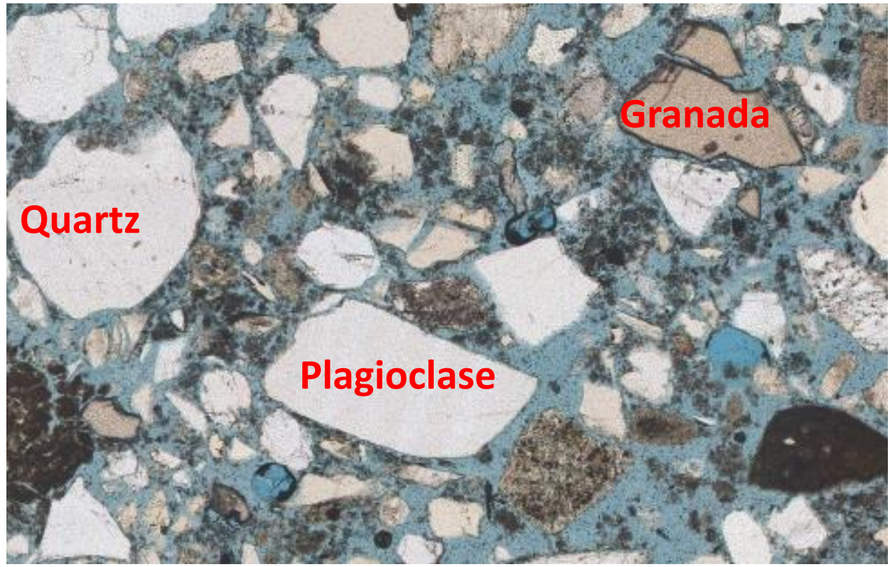


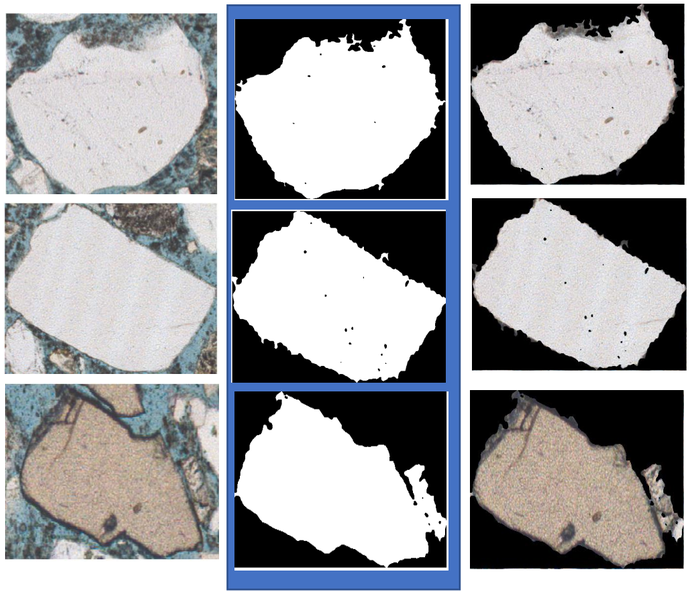
II - PetroSim



III - PetroInst

Análise de Lâminas Petrográficas e Inteligência Artificial





Estimativa da **Granulometria** em amostras de rocha sedimentar

Granulometria via Processamento de Imagens

- Procedimento não destrutivo
- Relativamente rápido, permitindo a análise de muitas amostras
- Automação – processamento de várias lâminas

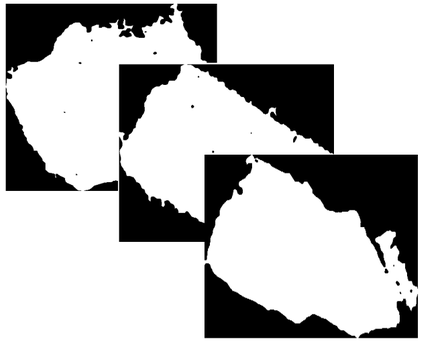
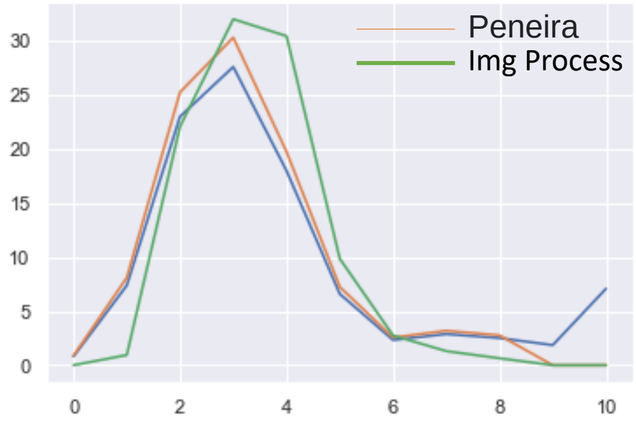


Imagem binária de grãos



	Areia					Silte					
	Cascalho	Muito Grossa	Grossa	Média	Fina	Muito Fina	Grosso	Médio	Fino	Muito Fino	Argila
Peneira	0,81	7,37	22,99	27,59	17,98	6,59	2,33	2,88	2,51	1,86	7,08
Peneira Corrigida	0,88	8,1	25,25	30,30	19,74	7,24	2,56	3,16	2,76	-	-
PDI	0,00	0,94	22,07	32,00	30,44	9,85	2,76	1,30	0,64	-	-

Análise de Lâminas Petrográficas e Inteligência Artificial

Classificação ou Agrupamento de imagens de grãos usando algoritmo não/supervisionado

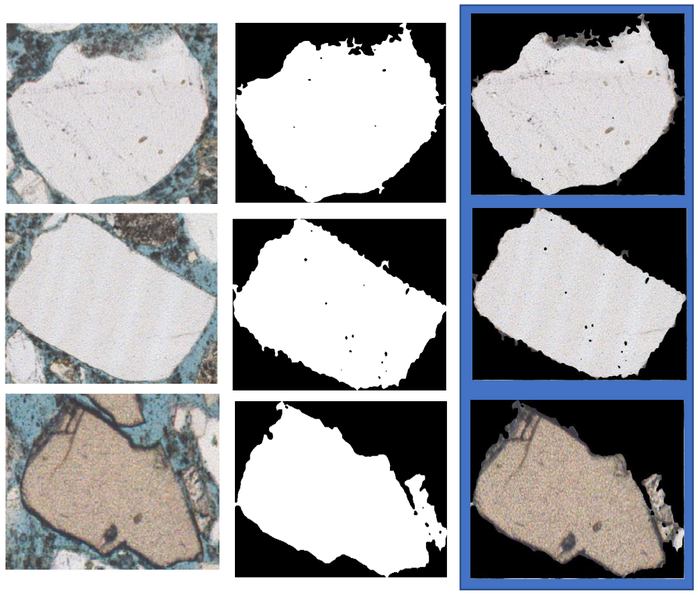
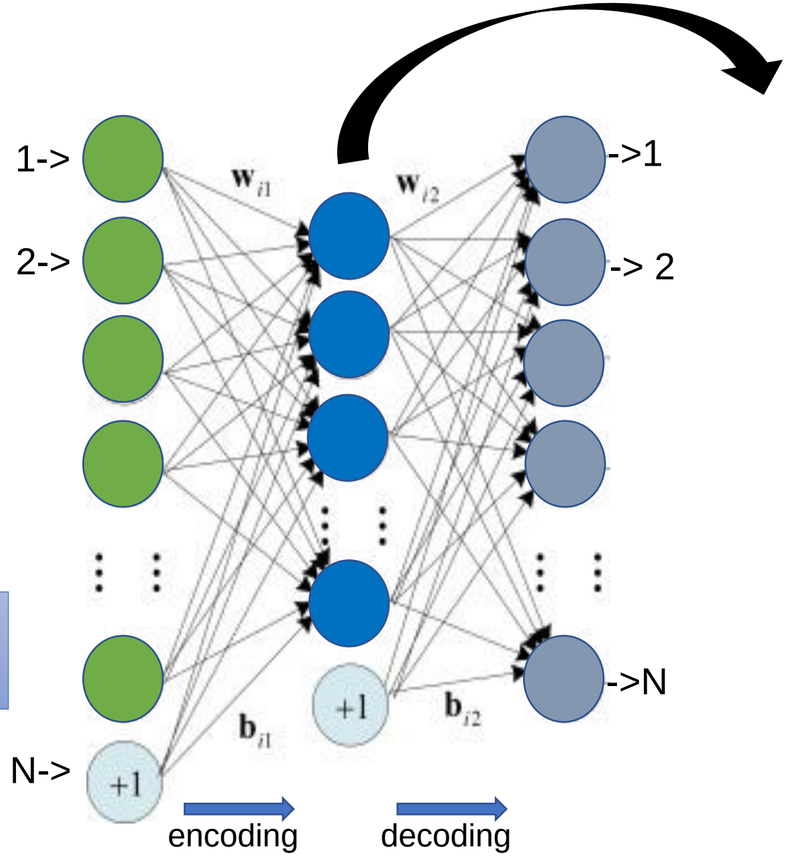


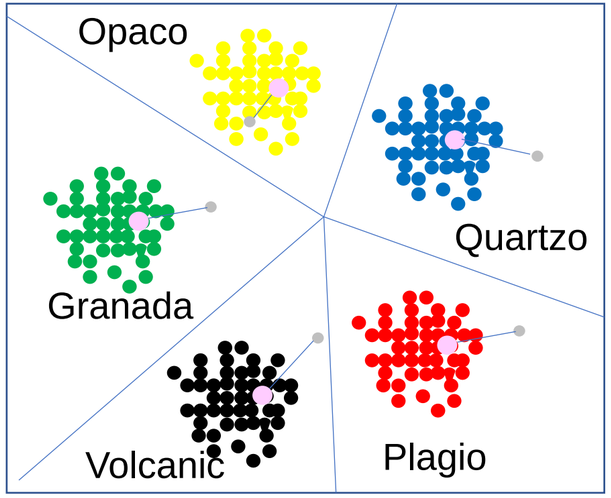
Imagem de textura de grão

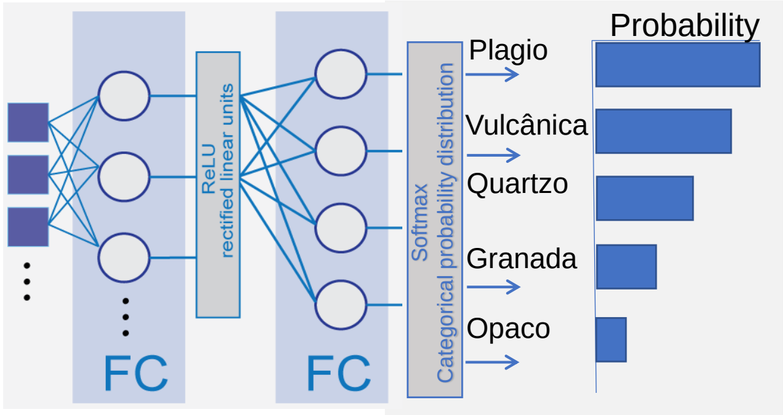


Deep AutoEncoder + supervised algorithm

Deep AutoEncoder + Non supervised algorithm

Classificação de padrões



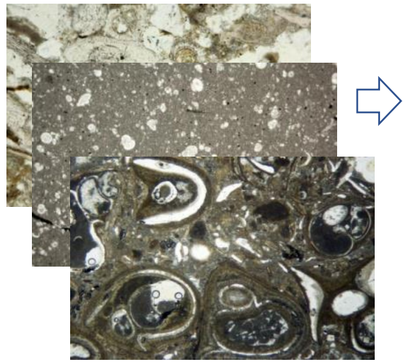


Output

Análise de Lâminas Petrográficas e Inteligência Artificial

Classificação litológica (lâminas) usando Algoritmo de Stacked Deep AutoEncoder (SAE) + Algoritmo de Otimização

Entrada



I_1

I_2

I_3

I_4

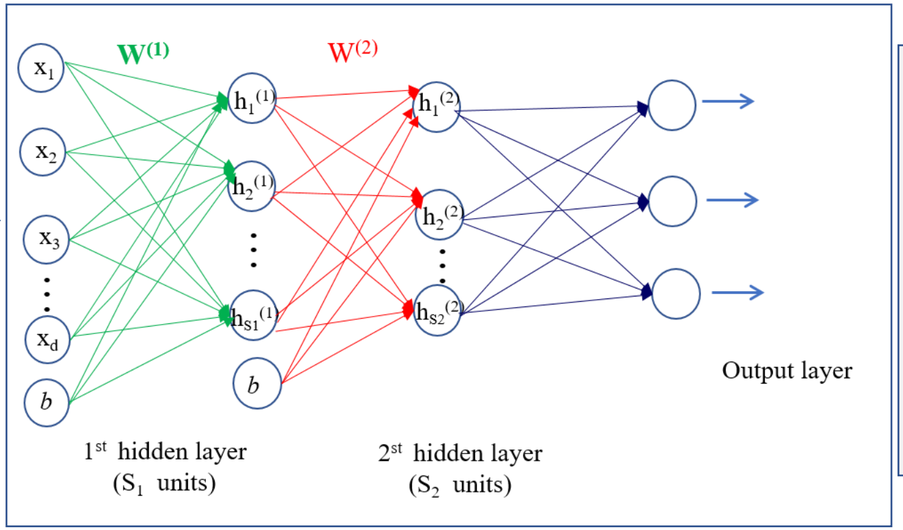
$\vdots$

I_d

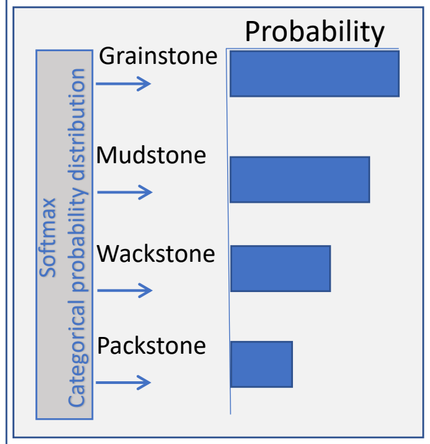
696x120 double

	1	2	3	4	5	6
1	0.2500	1.1874e+03	1.7634e+05	0.6477	1.6410	0.2182
2	0.2522	1.3162e+03	1.6246e+05	0.6299	1.5764	0.2281
3	0.2451	1.3535e+03	1.6252e+05	0.6161	1.6102	0.2256
4	0.2686	1.0063e+03	1.6698e+05	0.6839	1.6311	0.2233
5	0.2529	1.2689e+03	1.6835e+05	0.6396	1.5878	0.2262
6	0.2469	1.4258e+03	1.5642e+05	0.6069	1.5722	0.2308
7	0.2319	1.3874e+03	1.5534e+05	0.5898	1.6977	0.2145
8	0.2492	1.1904e+03	1.6968e+05	0.6417	1.6495	0.2181
9	0.2672	1.1143e+03	1.8570e+05	0.6792	1.5484	0.2288
10	0.2458	1.3666e+03	1.6464e+05	0.6155	1.5957	0.2268
11	0.2421	1.3265e+03	1.6644e+05	0.6157	1.6377	0.2205
12	0.2668	1.0633e+03	1.9006e+05	0.6846	1.5741	0.2246
13	0.2669	1.1196e+03	1.6590e+05	0.6717	1.5762	0.2280
14	0.2569	1.3480e+03	1.5820e+05	0.6306	1.5290	0.2345
15	0.2507	1.3244e+03	1.6869e+05	0.6291	1.5740	0.2281
16	0.2738	1.0458e+03	1.6453e+05	0.6859	1.5763	0.2330
17	0.2621	1.2167e+03	1.6330e+05	0.6552	1.5551	0.2302
18	0.2444	1.4229e+03	1.6562e+05	0.6057	1.5776	0.2279
19	0.2482	1.3113e+03	1.6100e+05	0.6249	1.6098	0.2245
20	0.2593	1.2327e+03	1.6308e+05	0.6495	1.5690	0.2303
21	0.2675	1.1702e+03	1.7834e+05	0.6720	1.5218	0.2341
22	0.2592	1.3316e+03	1.6725e+05	0.6392	1.5074	0.2366
23	0.2523	1.2390e+03	1.7130e+05	0.6427	1.6024	0.2235
24	0.2646	1.1326e+03	1.9099e+05	0.6741	1.5508	0.2265
25	0.2669	1.1196e+03	1.6590e+05	0.6717	1.5762	0.2280

Camada de entrada (parâmetros de textura)

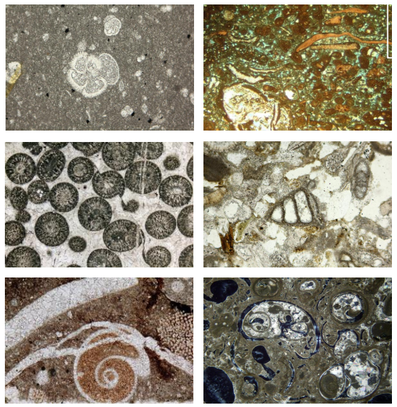


Representação de recursos com SAE



Softmax classifier training

Lâminas



Imagens de entrada para treinamento

Auto Machine Learning

AutoKeras

TensorFlow Keras

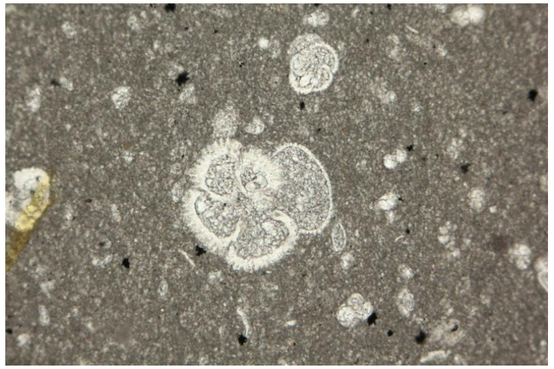
Algoritmo de Auto Machnie Learning

Usado para busca dos hiper parâmetros e arquiteturas das RN

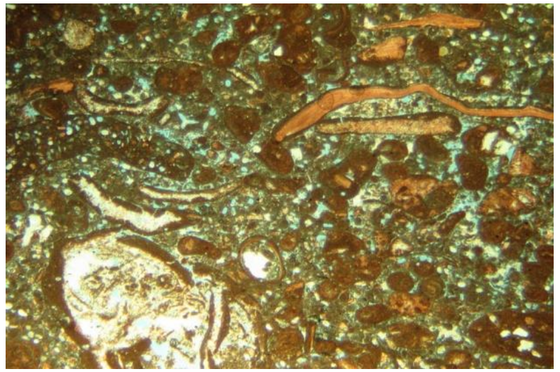
Análise de Lâminas Petrográficas e Inteligência Artificial

Classificação litológica (lâminas) usando Algoritmo de Stacked Deep AutoEncoder (SAE) + Algoritmo de Otimização

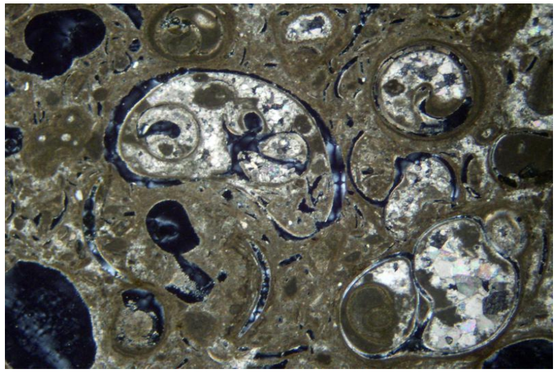
Grainstone **(GR)** = 228 imagens



Mudstone **(MU)** = 48 imagens



Packstone **(PA)** = 228 imagens



Wackstone **(WA)** = 180 imagens

Confusion Matrix

Output Class	GR	MU	PA	WA	
1 (GR)	224 32.7%	0 0.0%	3 0.4%	4 0.6%	97.0% 3.0%
2 (MU)	0 0.0%	39 5.7%	2 0.3%	0 0.0%	95.1% 4.9%
3 (PA)	4 0.6%	9 1.3%	220 32.2%	2 0.3%	93.6% 6.4%
4 (WA)	0 0.0%	0 0.0%	3 0.4%	174 25.4%	98.3% 1.7%
Target class	98.2% 1.8%	81.3% 18.8%	96.5% 3.5%	96.7% 3.3%	**96.1%** **3.9%**

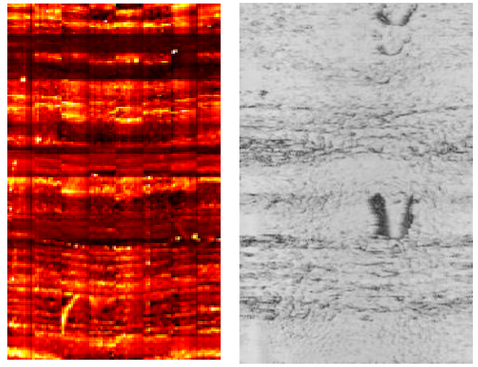
Target class: GR, MU, PA, WA

Acurácia média do modelo de 96,1%

Faria E. L. et al.; "***Lithology identification in carbonate thin section images of the Brazilian pre-salt reservoirs by the computational vision and deep learning***"; Computational Geosciences; submitted 2022.

Springer

resistividade ultrassônico



Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

The impact score (IS) 2021 of JPSE is **4.97**

Journal of Petroleum Science and Engineering 170 (2018) 315–330

Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate

Manuel Blanco Valentín[a,*], Clécio R. Bom[b], André Luiz Martins Compan[c], Maury Duarte Correia[c], Candida Menezes de Jesus[d], Anelise de Lima Souza[d], Márcio P. de Albuquerque[a], Marcelo P. de Albuquerque[a], Elisângela L. Faria[a]

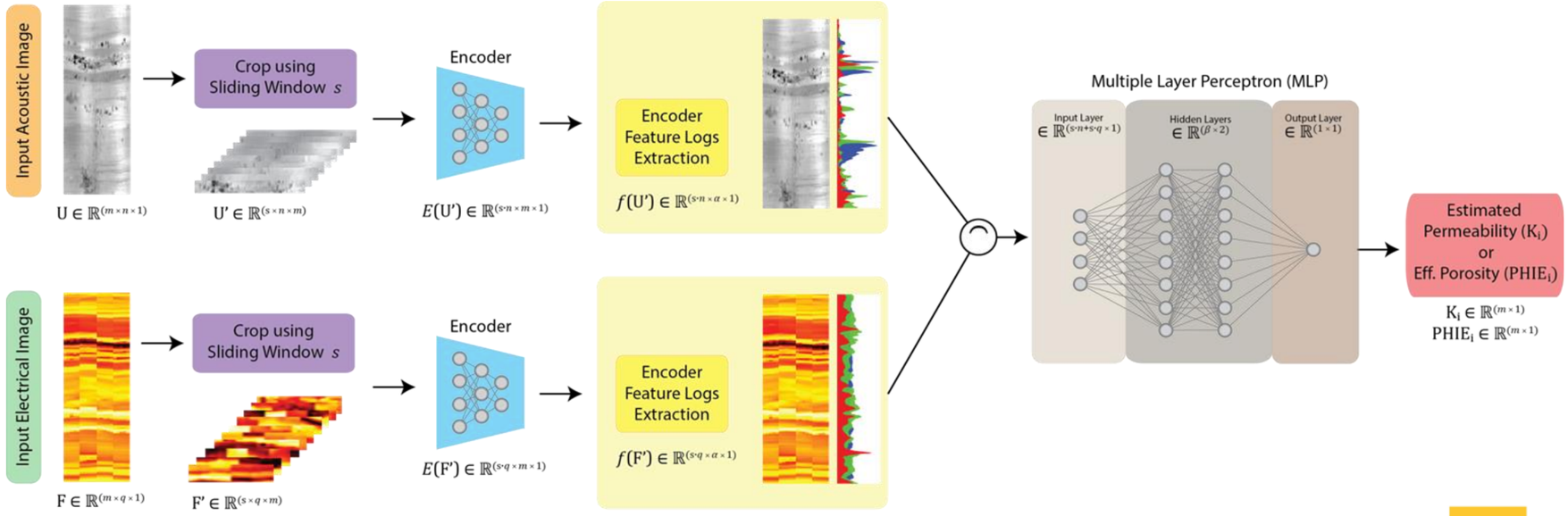
https://doi.org/10.1016/j.petrol.2018.06.038

Técnica de Inteligência Artificial/Machine Learning que estima permeabilidade/porosidade a partir de imagens de ultrassom e resistividade.

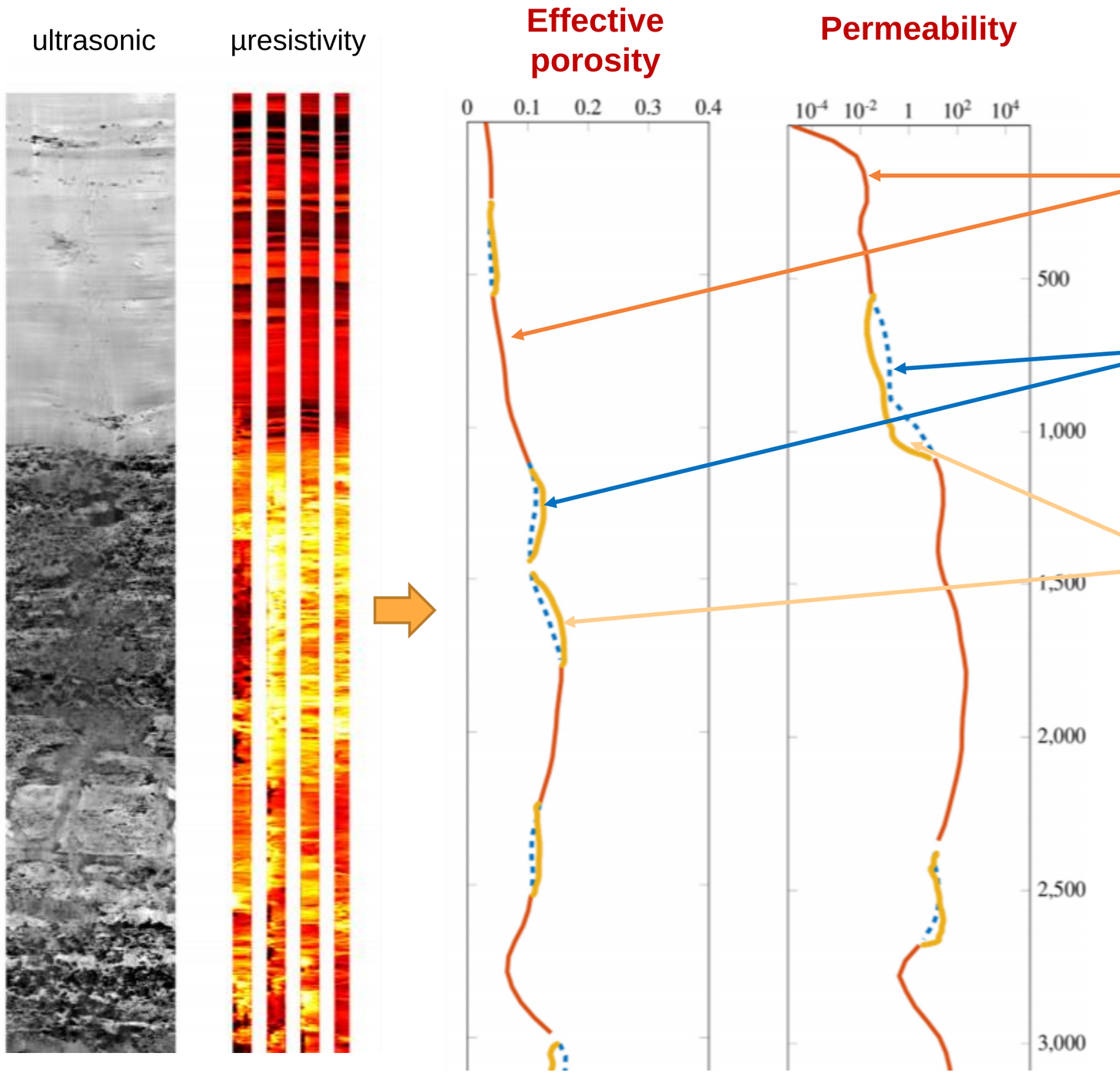
DOIS ESTÁGIOS:

O primeiro caracteriza a imagem em termos de atributos relevantes.

A segunda etapa utiliza outro modelo de Inteligência Artificial para obter a permeabilidade/porosidade.



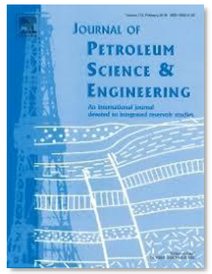
Workflow de Trabalho - Metodologia



Orange Lines:
Reference values used for training

Blue Lines:
Reference values used exclusively to compare with the network result

Yellow Lines:
Values predicted by the neural network



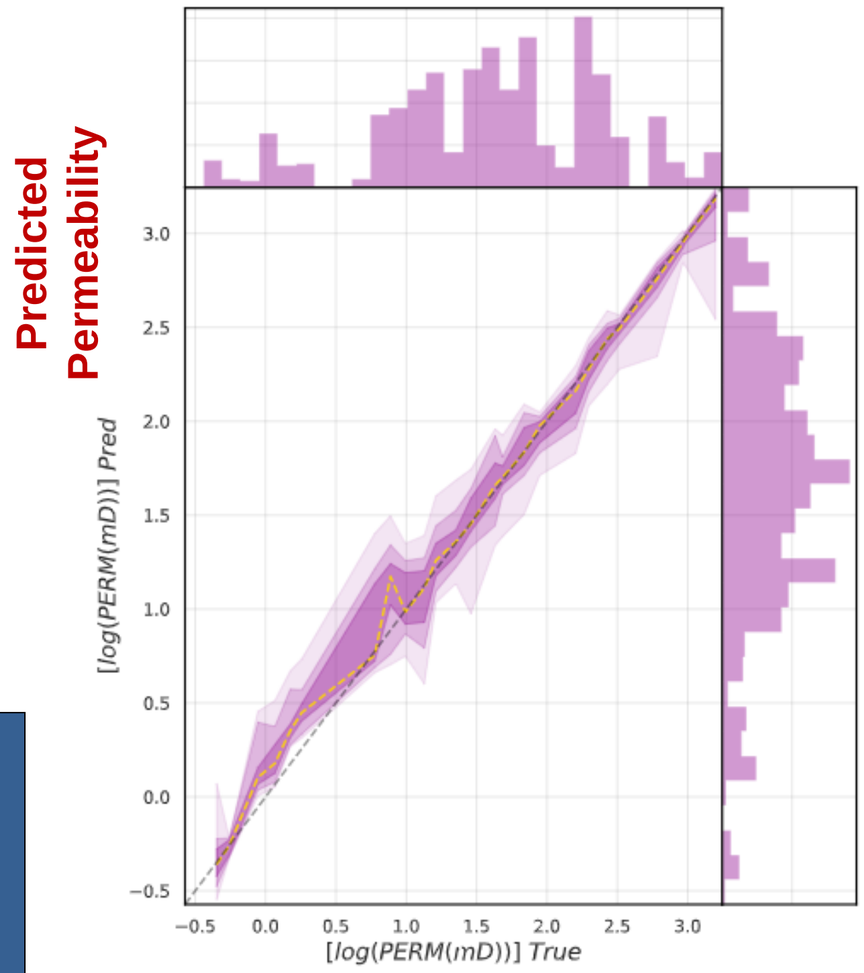
Journal of Petroleum Science and Engineering
Available online 16 January 2021, 108361
In Press, Journal Pre-proof

The impact score (IS) 2021 of JPSE is **4.97**

Bayesian Deep Networks for absolute permeability and porosity uncertainty prediction from image borehole logs from brazilian carbonate reservoirs

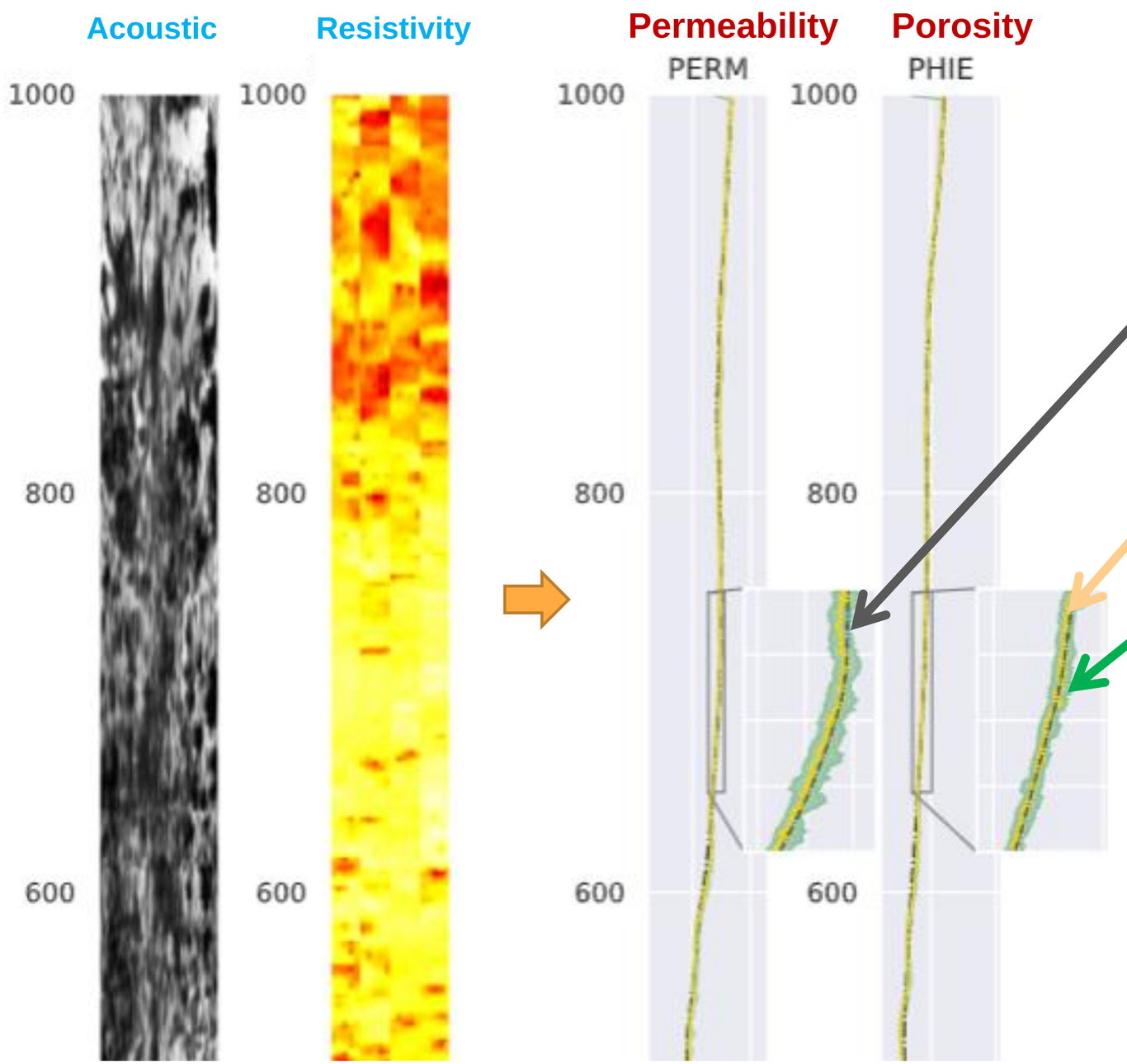
Clécio R. Bom [a, b], Manuel Blanco Valentín [a], Bernardo M.O. Fraga [a], Jorge Campos [c], Bernardo Coutinho [c], Luciana O. Dias [a], Elisangela L. Faria [a], Márcio P. de Albuquerque [a], Marcelo P. de Albuquerque [a], Maury Duarte Correia [c]

https://doi.org/10.1016/j.petrol.2021.108361



Em um artigo mais recente, trabalhamos na caracterização de erros, usando uma abordagem de Inteligência Artificial bayesiana.

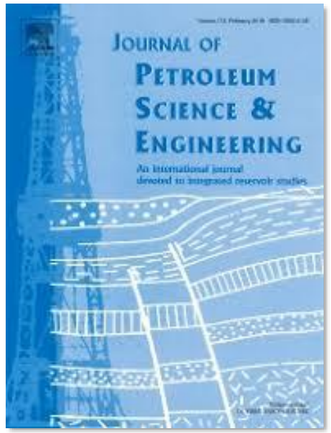
Nesse caso, alcançamos uma correlação superior a 99% em um poço do pré-sal. Além disso, foi possível estimar o erro em cada região do poço.



Black Dashed Line:
Reference Values

Yellow Line:
DL predictions

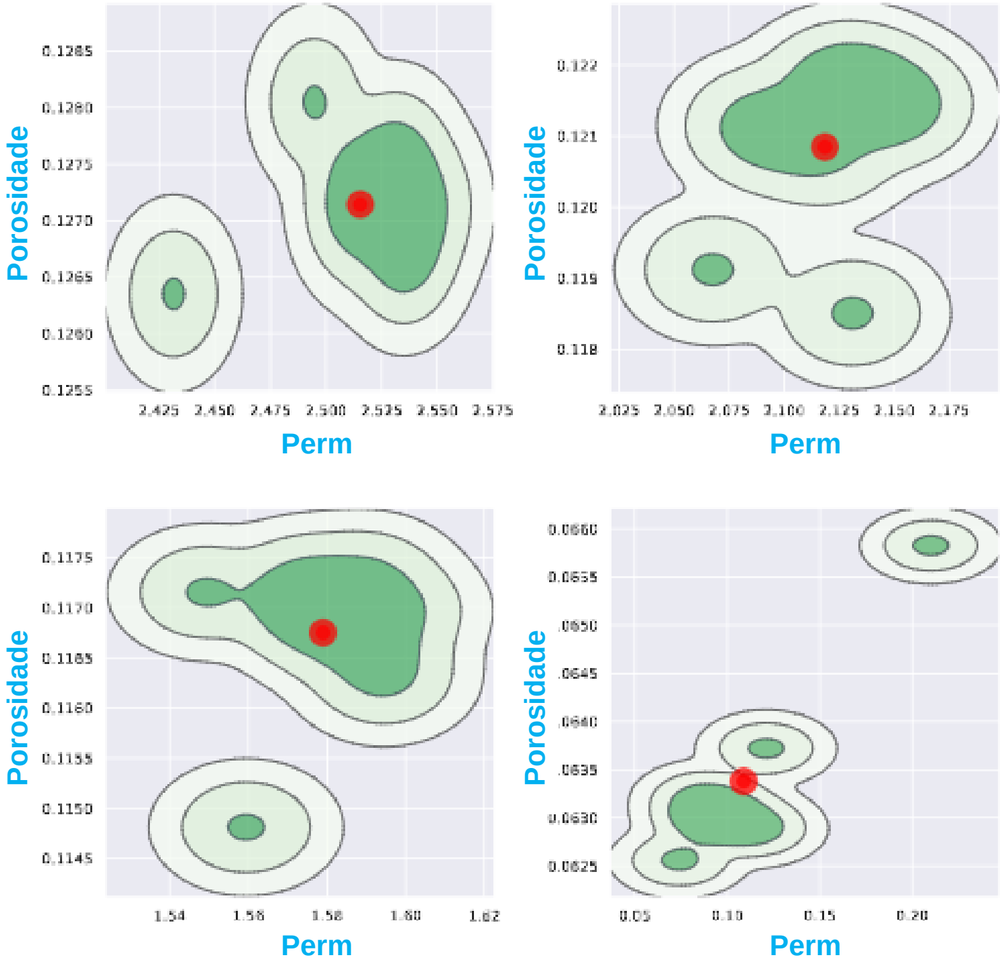
Green Shadow
Model Uncertainty



Densidade de Probabilidade Conjunta

Quatro exemplos de Funções de Densidade de Probabilidade Conjunta

● Valor-alvo



$R^2(\%)$ Perm	$R^2(\%)$ Phie
99.63 ± 0.31	99.31 ± 0.92

Correlação entre permeabilidade e porosidade em comparação com plug, em um carbonato do Pré-Sal

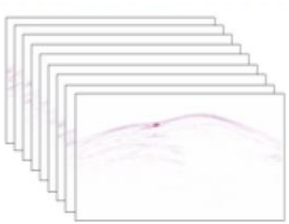
Nenhuma intervenção humana!

Análise sísmica & Inteligência Artificial

Construção de modelo de velocidade sísmica (VM)

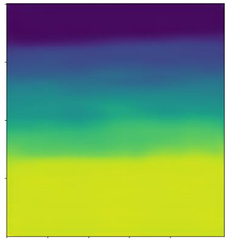
Método convencional de inversão baseado em modelos físicos:

Dados Sísmicos



MÉTODO DE INVERSÃO BASEADO EM MODELOS FÍSICOS COMPLEXOS

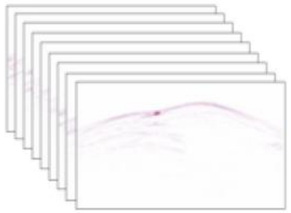
Predicted VM



- Demorado.
- Computacionalmente caro.
- e há sempre uma necessidade de iteração humana.

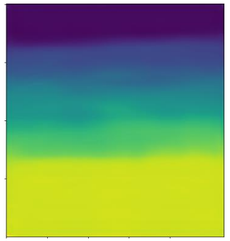
Inversão por Deep Learning – Construção de modelo de velocidade sísmica (VM)

Dados Sísmicos



DEEP LEARNING INVERSION

Predicted VM



- São baseados no treinamento com big-data.
- Na fase de treinamento, a rede estabelece uma projeção não linear dos dados sísmicos de vários tiros para os modelos de velocidade correspondentes (requer mais tempo).
- No entanto, durante a fase de previsão, a rede treinada pode ser usada para estimar os modelos de velocidade a partir dos novos dados sísmicos de entrada em tempo real.
- Sem iteração humana.

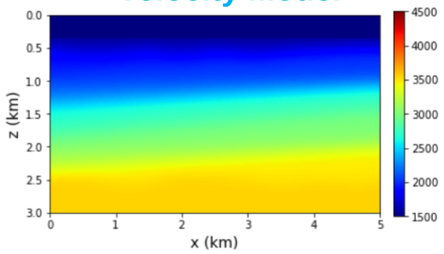
Klatt, M. et al.; "***Deep learning strategy for salt model building***", GEO-2021-0362 (Geophysics); submitted 2022;

Geometrias do sal com Deep Learning

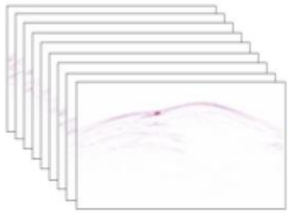
Deep Learning para prever a geometria do sal

Dados do modelo de sal da SEG (Society of Exploration Geophysicists)

Velocity Model

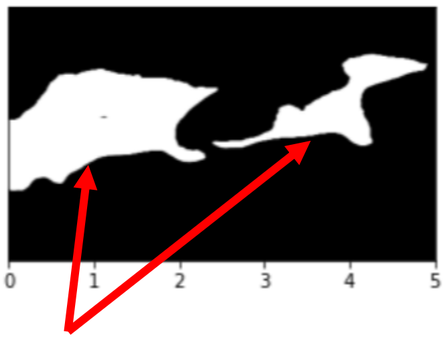


Seismic Migrated Data



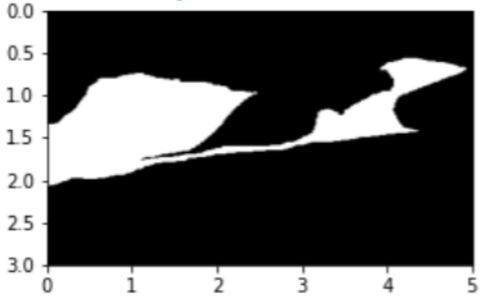
DEEP LEARNING NN

Predicted Salt Geom.

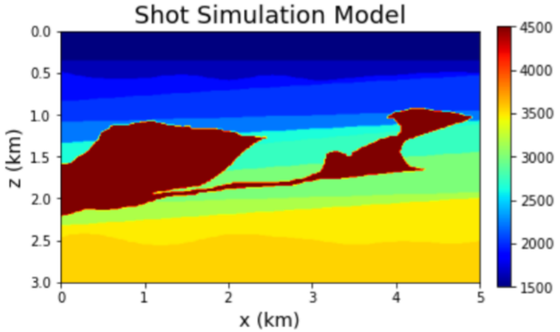


Salt Mask

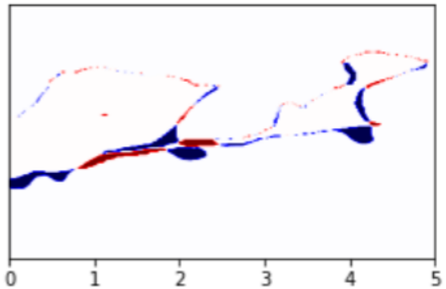
Target Salt Geom.



Shot Simulation Model



Error



Os resultados estão atualmente em validação estatística em casos reais da indústria, onde a geometria do sal foi mal definida pelos métodos tradicionais.

Aceito para publicação em 2022

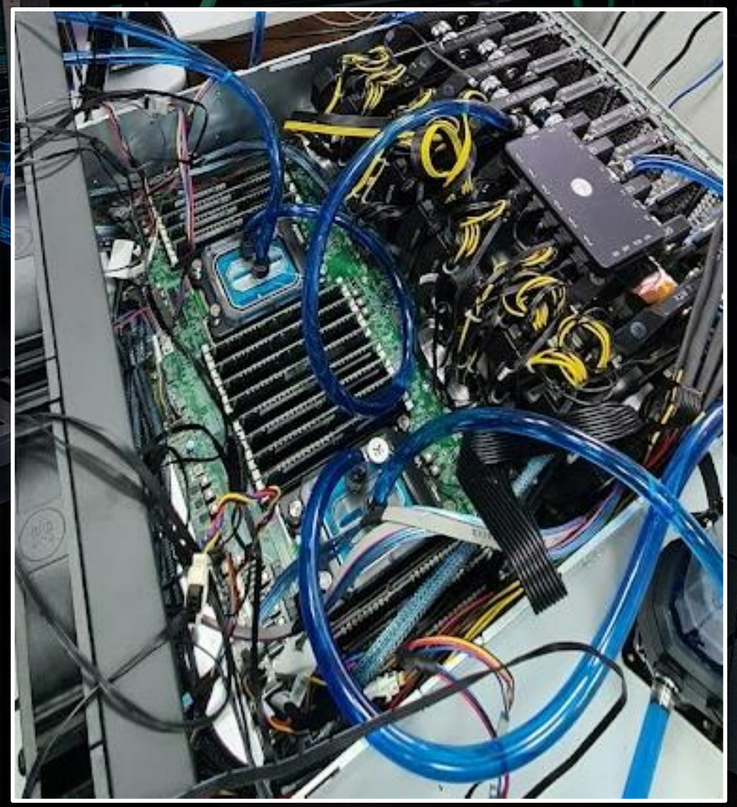
Ana Paula O. Muller, et al.; "***Complete identification of complex salt-geometries from inaccurate migrated images using Deep Learning***"; GEO-2021-0362 (Geophysics);

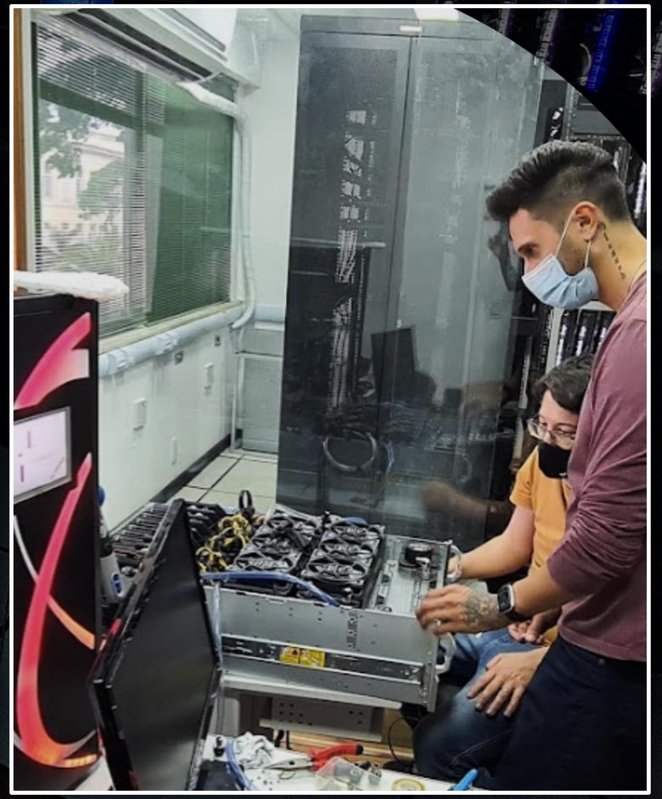
Novos Instrumentos e Sistemas HPCs MultiGPUs Para Inteligência Artificial / DL

IA multiGPU Lab

Desenvolvimento de tecnologias multiGPUs otimizadas com as características dos conjuntos de dados

Tecnologias contribuem para o teste, prototipagem rápida e estudo de viabilidade de modelos de IA





➡ Sci.minds permitem ajustar parâmetros de desempenho

sci.mind
CBPF
acer
Microsoft
GLOBAL POWER

ARTIFICIAL INTELLIGENCE
Classification, Simulations and Forecasts

We participated and won the International Astrophysics Challenge / 2020

International challenge to identify gravitational lensing systems in simulated images.

Developing a Victorious Strategy to the Second Strong Gravitational Lensing Data Challenge

C. R. Bom[a,b], B. M. O. Fraga[a], L. O. Dias[a], P. Schubert[a,1], M. Blanco Valentin[d,2], C. Furlanetto[e], M. Makler[a,c], K. Teles[a], M. Portes de Albuquerque[a], R. Benton Metcalf[f,g]

[a]Centro Brasileiro de Pesquisas Físicas, Rua Dr. Xavier Sigaud 150, CEP 22290-180, Rio de Janeiro, RJ, Brazil
[b]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rodovia Mário Covas, lote J2, quadra J, CEP 23810-000, Itaguaí, RJ, Brazil
[c]International Center for Advanced Studies & Instituto de Ciencias Físicas, ECyT-UNSAM & CONICET, 25 de Mayo y Francia. C.P.: 1650, San Martín, Buenos Aires, Argentina
[d]Electrical and Computer Engineering Department, McCormick School, Northwestern University, 633 Clark St, Evanston, IL 60208
[e]Universidade Federal do Rio Grande do Sul, Departamento de Física, CEP 91501-970, Porto Alegre, RS, Brazil
[f]Dipartimento di Fisica e Astronomia, Università di Bologna, via Gobetti 93/2, I-40129 Bologna, Italy
[g]INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, I-40129 Bologna, Italy

Abstract

Strong Lensing is a powerful probe of the matter distribution in galaxies and galaxy clusters and a relevant tool for cosmography. Analyses of strong gravitational lenses with Deep Learning (DL) and Convolutional Neural Networks (CNNs) have become a popular approach due to these astronomical objects' rarity and image complexity. Next-generation surveys (both ground and space-based) will provide more opportunities to derive science from these objects and an increasing data volume to be analyzed. However, finding strong lenses is challenging, as their number densities are orders of magnitude below those of galaxies. Therefore, specific Strong Lensing search algorithms are required to discover the highest number of systems possible, with high purity and low false alarm rate to minimize human intervention. The need for better algorithms has prompted the development of an open community data science competition named Strong Gravitational Lensing Challenge (SGLC) by the Bologna group. In this work, we present the Deep Learning strategies and methodology used to design the highest-scoring algorithm in the II SGLC, which was based on Euclid-like simulations. We discuss the approach used for this dataset, the choice for a suitable architecture, particularly the use of a network with two branches to work with images in different resolutions, and its optimization. We also discuss the limit of what tailor-made architecture in a survey in contrast pipeline, and discuss the best choice to easily ment. This work helps to take a step towards s.

networks

Submitted 2022

Monthly Notices of the Royal Astronomical Society

The team working for the industry contributed to the scientific project.

Da esq. para dir., Patrick, Marcelo, Martín, Paulo, Clécio, Pedro, Manuel, Marcos e Luciana

(Crédito: NCS/CBPF)

https://portal.cbpf.br/pt-br/ultimas-noticias/grupo-do-cbpf-vence-desafio-internacional

CBPF
Centro Brasileiro de Pesquisas Físicas
UNIDADE DE PESQUISA DO MCTI

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

Aplicações de IA em Astrofísica e Desenvolvimentos com a Indústria de O&G

Luciana Olivia Dias

Clécio R. De Bom, Elisangela L. Faria e Marcelo Albuquerque

O2I

16-18 Novembro 2022, CBPF, Rio de Janeiro

Um evento que mostra o potencial da ciência na sociedade moderna e estimula sinergias entre pesquisa fundamental e indústria