

Análise de Diferentes Algoritmos Semissupervisionados para *Data Augmentation*

Willian Dihanster Gomes de Oliveira¹, Otávio A. B. Penatti², Lilian Berton¹

¹ Universidade Federal de São Paulo, ² Samsung Research Institute Brazil
Av. Cesare Mansueto Giulio Lattes, 1201, São José dos Campos - SP, CEP 12247-014

Introdução

A classificação de imagens é uma das grandes áreas de aplicação do Aprendizado de Máquina (AM). No entanto, para um bom desempenho dos algoritmos, faz-se necessária uma grande quantidade de dados, especialmente rotulados, além de um bom descritor de imagens. Quando se tem dados rotulados e não rotulados, pode-se aplicar um algoritmo semissupervisionado baseado em grafos para propagar a informação de rótulo pelos exemplos não rotulados. Portanto, o objetivo deste trabalho é analisar diferentes algoritmos semissupervisionados para *Data Augmentation* (DA) e o uso de redes neurais convolucionais (CNNs) como extratores de características de imagens.

Área do Conhecimento

Ciências Exatas e da Terra; Ciência da Computação; Inteligência Computacional

Metodologia

As bases de dados utilizadas são descritas a seguir:

1. Brazilian Coffee Scenes Dataset³

Esta base é composta por imagens de satélite de plantações de café e é formada por 2876 exemplos, sendo 50% dos exemplos rotulados como "coffee" e 50% como "non coffee", exemplificados na Figura 1.

2. Dogs vs Cats⁴

Possui 12.000 exemplos de imagens de cachorros e gatos, sendo 50% rotulados como "Dog" e os outros 50% como "Cat", com exemplos na Figura 2.

3. MNIST DATABASE of handwritten digits⁵

Trata-se de um *subset* com 8.000 exemplos de imagens de dígitos, sendo 50% rotulados como "3", e os outros 50% como "8". Dois exemplos na Figura 3.



Fig. 1: Exemplos da base 1.

Fig. 2: Exemplos da base 2.

Fig. 3: Exemplos da base 3.

A metodologia é descrita pela sequência de passos da Figura 4:

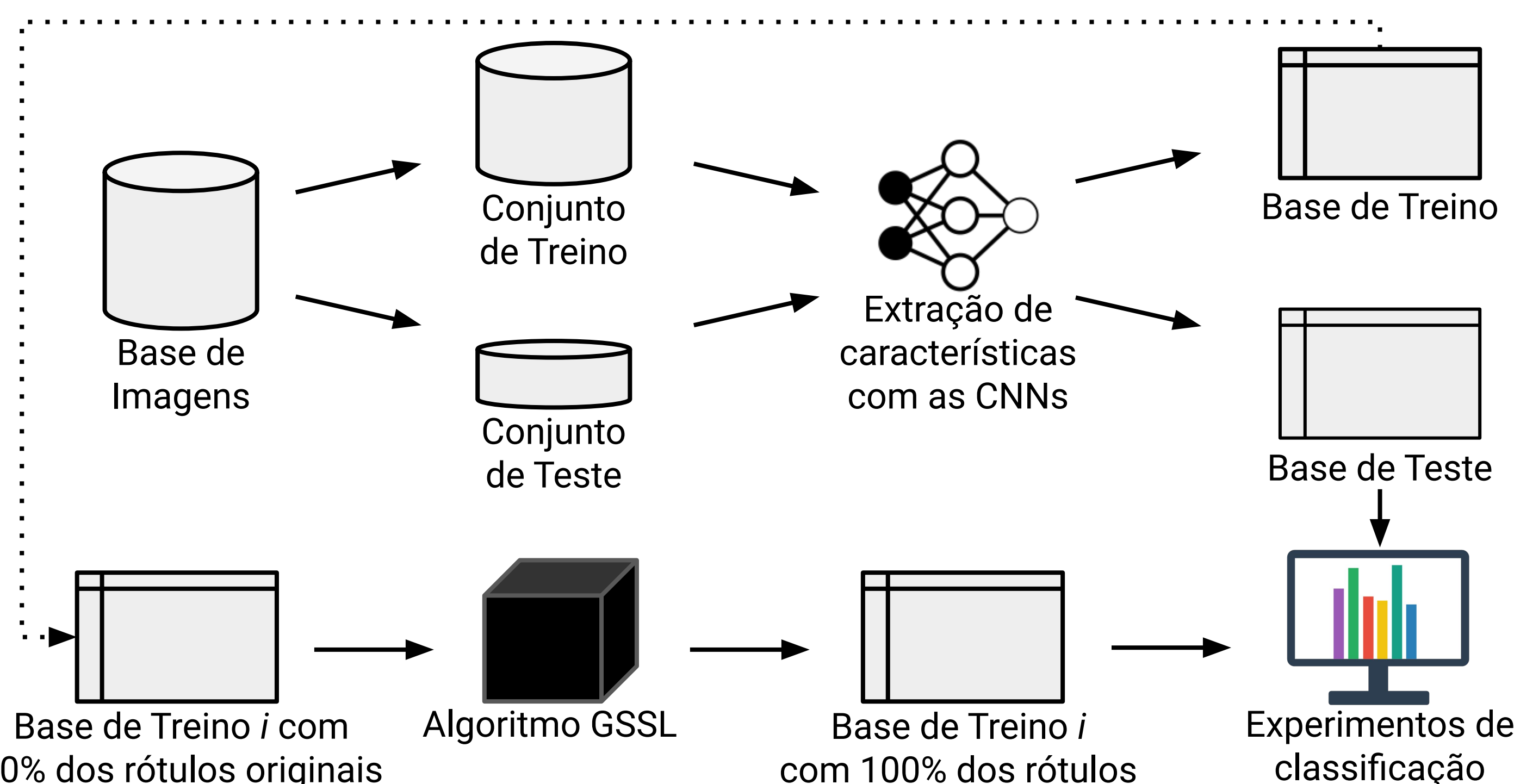


Fig. 4: Esquema da metodologia do projeto.

Redes Neurais Convolucionais:

- ResNet50
- VGG16 e VGG19
- Xception

Softwares e Bibliotecas:

- Python
- TensorFlow e Keras
- Numpy, Pandas e scikit-learn

Algoritmos Semissupervisionados (GSSL):

- Label Propagation (LP) [1]
- Local and Global Consistency (LGC) [2]
- Gaussian Fields and Harmonic Functions (GFHF) [3]
- OMNI-Prop (OMNI) [4]

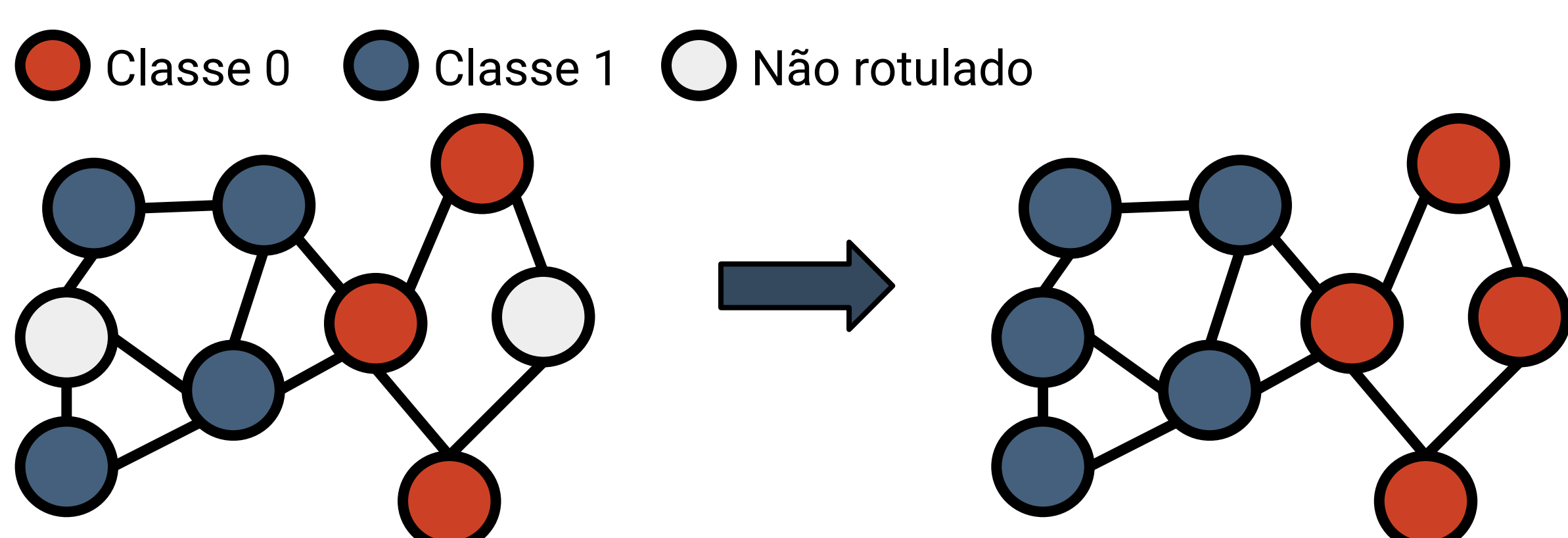


Fig. 5: Exemplos de propagação de rótulos.

Resultados

Avaliação dos Algoritmos

A seguir, são apresentados os resultados dos experimentos de classificação. Cada experimento considera a acurácia média, utilizando 5 conjuntos diferentes (de cada base) para cada algoritmo e para cada CNNs.

1. Brazilian Coffee Scenes Dataset

	ResNet50	VGG16	VGG19	Xception	Média
Original	87,31	82,09	82,99	79,60	83,00
LP	79,79	68,59	72,01	66,79	71,80
LGC	81,33	70,92	68,15	71,62	73,01
GFHF	84,04	72,95	76,47	73,42	76,72
OMNI	82,54	66,93	68,82	72,84	72,78

Tabela 1: Resultados para a base de dados 1.

2. Dogs vs Cats

	ResNet50	VGG16	VGG19	Xception	Média
Original	98,90	98,55	98,62	71,01	91,77
LP	98,07	97,46	97,55	55,63	87,18
LGC	98,46	98,19	98,06	60,82	88,88
GFHF	98,37	98,18	97,87	62,33	89,19
OMNI	98,36	98,19	97,82	62,26	89,16

Tabela 2: Resultados para a base de dados 2.

3. MNIST (3-8)

	ResNet50	VGG16	VGG19	Xception	Média
Original	98,84	98,62	98,76	92,29	97,13
LP	97,20	96,59	96,95	75,87	91,65
LGC	84,45	86,49	86,89	86,86	86,17
GFHF	98,38	97,69	97,47	91,23	96,19
OMNI	98,23	97,79	97,37	89,66	95,76

Tabela 3: Resultados para a base de dados 3.

Avaliação das CNNs como extratores de características

Para esta avaliação, foi calculada a média das acurácias obtidas por todos os experimentos que utilizam uma base gerada por tal rede (Figura 6).

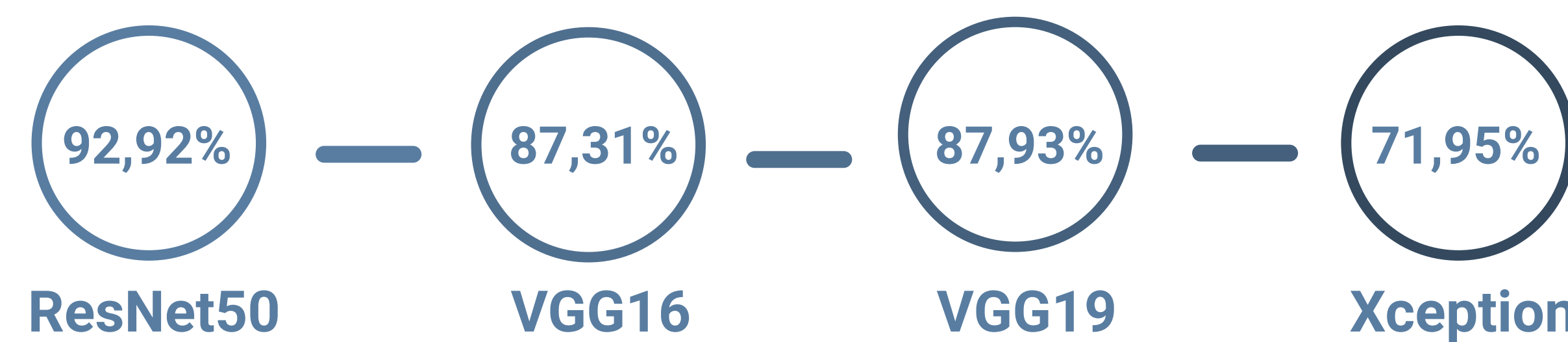


Fig. 6: Resultado médio para as CNNs como extratores de características.

Resumo dos algoritmos

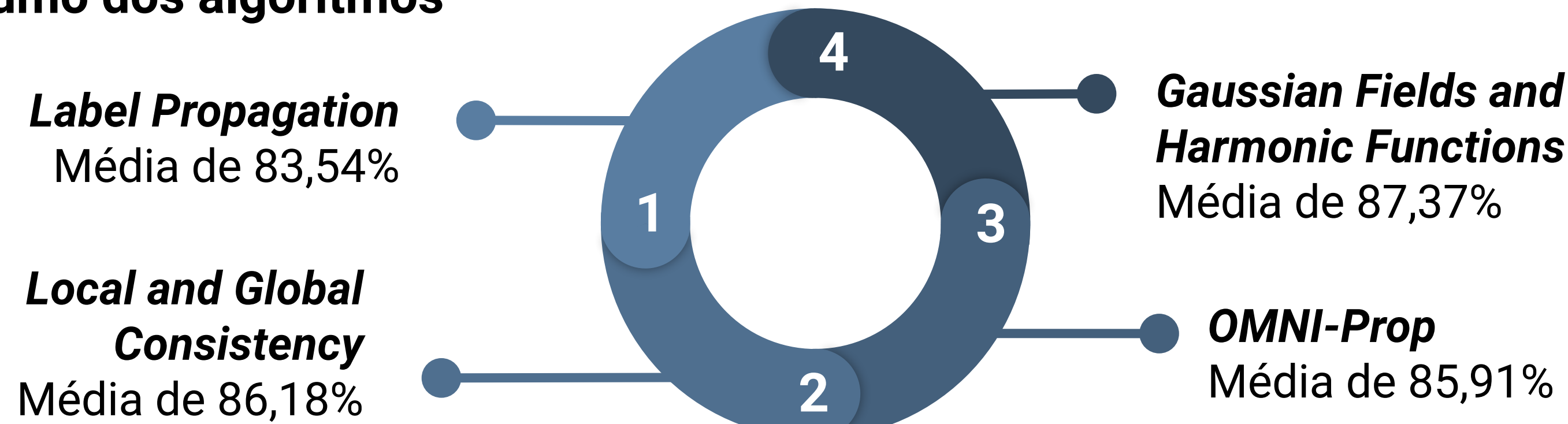


Fig. 7: Resumo dos resultados obtidos pelos algoritmos semissupervisionados.

Conclusão

Os algoritmos selecionados são eficientes para a aplicação escolhida, e assim, podem ser aplicados em casos com muitos exemplos não rotulados, visando melhorar o modelo gerado por algoritmos preditivos. Além disso, as CNNs podem ser utilizadas como bons descritores de imagens, com destaque para a ResNet50, que obteve melhores resultados, em média.

Referências

- [1] Zhu, Xiaojin, and Zoubin Ghahramani. *Learning from labeled and unlabeled data with label propagation*. Technical Report CMU-CALD-02-107, Carnegie Mellon University, 2002.
- [2] Zhou, Denny, et al. "Learning with local and global consistency." *Advances in neural information processing systems*. 2004.
- [3] Yamaguchi, Yuto, Christos Faloutsos, and Hiroyuki Kitagawa. "Omni-prop: Seamless node classification on arbitrary label correlation." *Twenty-Ninth AAAI Conference on Artificial Intelligence*. 2015.
- [4] Zhu, Xiaojin, Zoubin Ghahramani, and John D. Lafferty. "Semi-supervised learning using gaussian fields and harmonic functions." *Proceedings of the 20th International conference on Machine learning (ICML-03)*. 2003.

Apoio e Agradecimentos

Este projeto foi apoiado e financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.