

Data Augmentation via Aprendizado Semissupervisionado Baseado em Grafos

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

O Aprendizado de Máquina (AM) é uma área de grande destaque nos dias atuais. Mas um de seus problemas é a necessidade de uma grande quantidade de dados, em especial, rotulados, que são custosos/difíceis de obter. Assim, diversas técnicas vêm sendo propostas com o intuito de obter mais dados. Neste trabalho, objetivamos fazer o estudo de duas técnicas de *data augmentation* (DA) para a classificação de imagens: transformação de imagens e aprendizado semissupervisionado (*Semi-supervised Learning* - SSL) baseado em grafos. E assim, analisar os resultados, bem como as vantagens e desvantagens de cada técnica.

Área do Conhecimento

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

Metodologia

A base de dados utilizada [1] é composta por imagens de satélite de plantações de café e formada por 2876 exemplos, sendo 50% de exemplos rotulados como "coffee" (Figura 1) e 50% como "non coffee" (Figura 2).

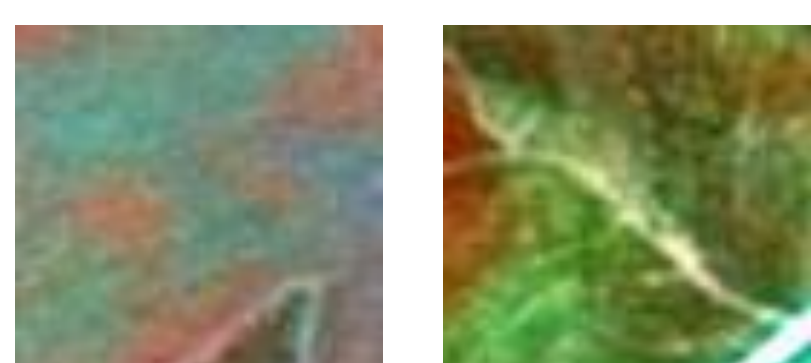


Fig. 1: Exemplos rotulados como *coffee*.



Fig. 2: Exemplos rotulados como *non coffee*.

A metodologia é descrita como a sequência de passos da Figura 3:

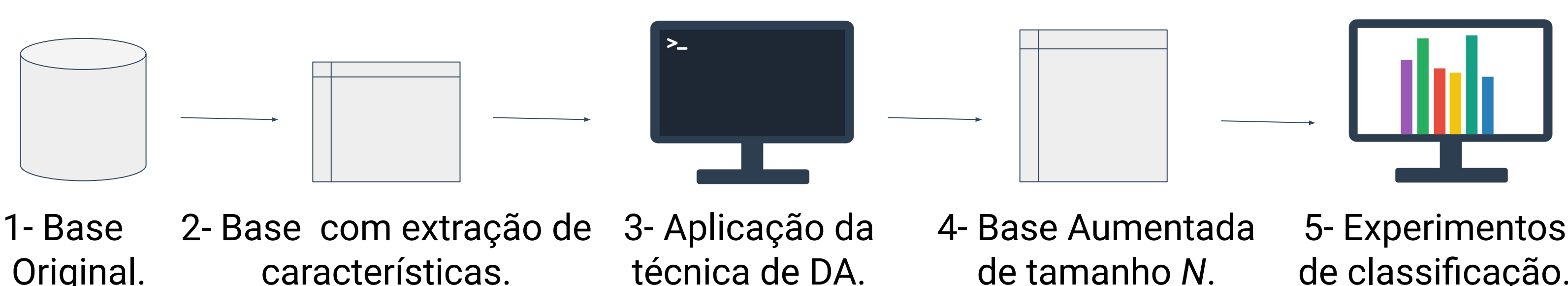


Fig. 3: Diagrama da metodologia do trabalho.

O passo 3 varia entre duas técnicas. Sendo elas:

Transformação de Imagens (Imagens)

Exemplos aleatórios são escolhidos e aplica-se alguns efeitos como corte, girar, efeito *blur*, etc, até atingir um conjunto de imagens de tamanho N . Como exemplificado na Figura 4.

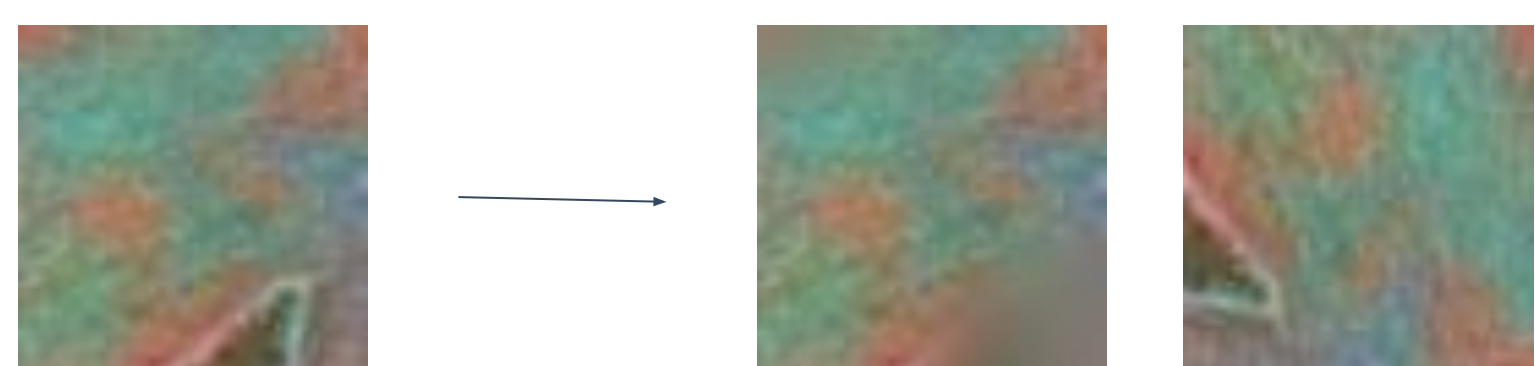


Fig. 4: Exemplo de aplicação da técnica transformação de imagens.

Aprendizado Semissupervisionado Baseado em Grafos (Grafos-SSL)

Usando uma porcentagem de exemplos rotulados e o restante não rotulados, gera-se um grafo usando a semelhança dos exemplos e então é aplicado o algoritmo de propagação de rótulos *Local and Global Consistency* (LGC) [2]. Na Figura 5 temos um exemplo desse grafo e o pseudocódigo do algoritmo LGC na Figura 6.

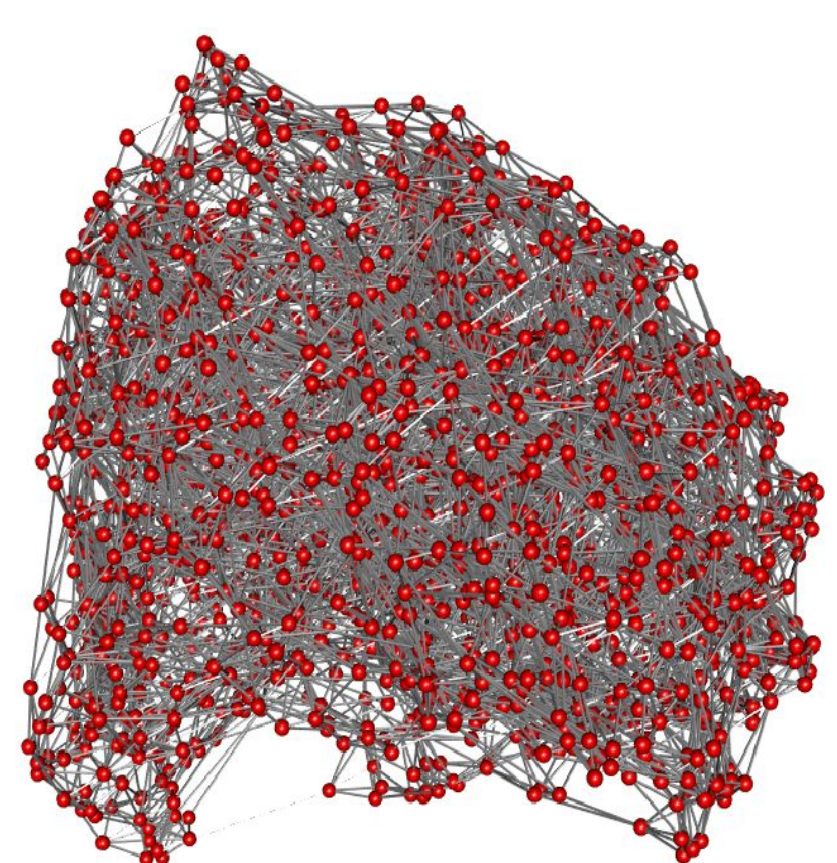


Fig. 5: Grafo construído para a propagação.

Algoritmo 1: Local and Global Consistency
Entrada: G um grafo ponderado $G(V, E, W)$.
Saída: Rótulos para o conjunto Y_u , $1 \leq u \leq n$
início
 Calcula-se a matriz de afinidade W , utilizando a seguinte fórmula:

$$W_{ij} = \begin{cases} \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

 Calcula-se a matriz diagonal D , como: $D_{ii} \leftarrow \sum_j W_{ij}$.
 Calcula-se a matriz laplaciana normalizada utilizando: $S = D^{1/2} W D^{1/2}$.
 Escolhe-se um parâmetro α no intervalo $[0, 1]$.
repita
 | $F(t+1) = \alpha S F(t) + (1-\alpha) Y$.
 até que F atinja a convergência;
 Ao final, obtém-se F^* .
 Rotula-se as amostras x_i , com $y_i = \arg \max_{j \leq c} F_{ij}^*$.
fim

Fig. 6: Algoritmo LGC.

No passo 5, a base aumentada é utilizada para treino por 5 classificadores: Árvores de Decisão (J48), k -Nearest Neighbors (k -NN), Naive Bayes (NB), Redes Neurais (MLP) e Support Vector Machines (SMO). Em seguida, a base (do passo 2) é usada como teste para os classificadores e é calculada a acurácia desses experimentos.

Resultados

Comparação entre Imagens e Grafos-SSL:

Ambas as técnicas consideram apenas 10% do conjunto original. Um resumo dos resultados pode ser conferido na Tabela 1 a seguir.

Tabela 1: Resumo dos resultados da comparação das técnicas.

	Média	Desvio Padrão	Máximo	Mínimo
Imagens	78,66%	1,64	82,68%	70,97%
Grafos-SSL	80,04%	1,94	84,98%	73,61%
Original (10-Fold)	84,37%	2,65	87,41%	79,73%

Grafos-SSL com variação da porcentagem inicial de exemplos rotulados:

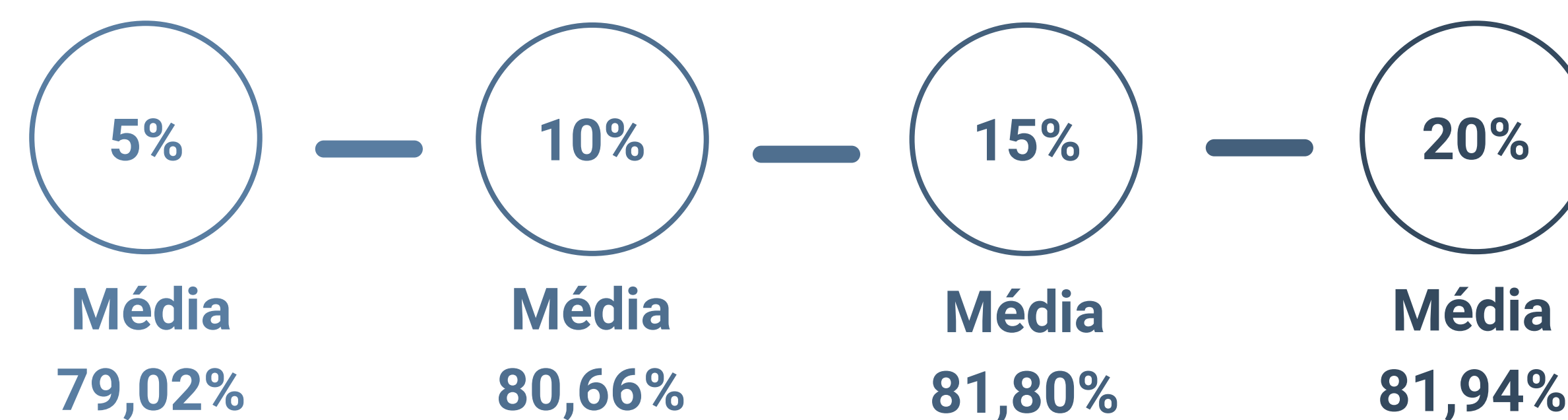


Fig. 7: Acurácia média de acordo com a % inicial de exemplos rotulados.

Comparação dos classificadores utilizados no WEKA:

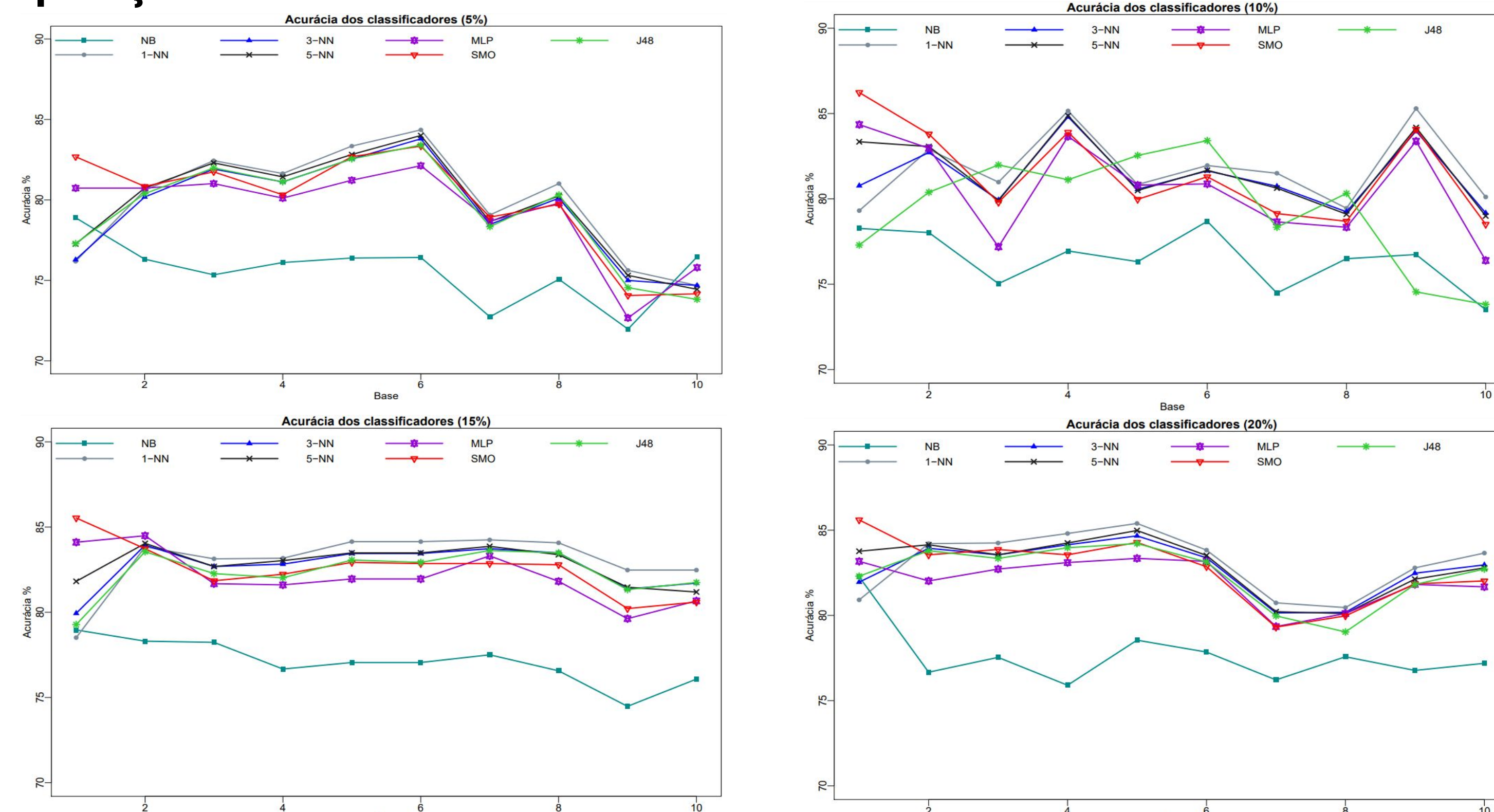


Fig. 8: Acurácia dos classificadores a partir da % inicial de exemplos rotulados.

Deep Learning e variação de parâmetros das Redes Neurais:

Dada as bases de DA com maior resultado variou-se alguns parâmetros da Rede Neural usada no WEKA descritos na Figura 9. O melhor resultado foi:

DA Imagens obteve **84,46%** no WEKA e **84,28%** no *TensorFlow*.

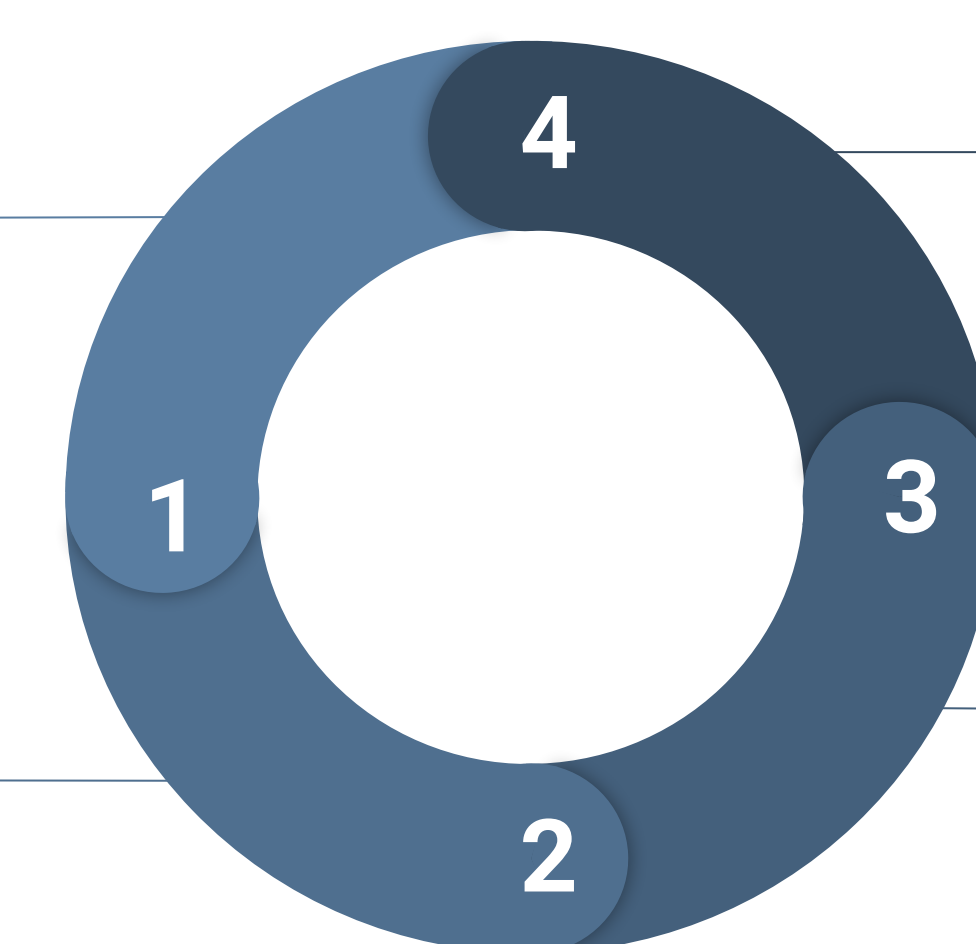
DA Grafos-SSL obteve **84,11%** no WEKA e **83,92%** no *TensorFlow*.

Número de camadas ocultas

Ambos com 3 camadas ocultas.

Learning Rate (LR)

Ambos com LR = 0.3.



Número de Neurônios

Ambos com 64 neurônios em cada camada oculta.

Tempo de Treinamento (Épocas)

Imagens 40 épocas, para Grafos-SSL 600 épocas.

Fig. 9: Diagrama dos parâmetros dos melhores resultados encontrados no WEKA.

Conclusão

Conclui-se que ambas as técnicas são eficientes para a aplicação escolhida e que técnicas de DA podem ser aplicadas em casos com poucos exemplos rotulados para melhorar o modelo gerado pelos algoritmos supervisionados.

Referências

- [1] Penatti, Otávio AB, Keiller Nogueira, and Jefersson A. dos Santos. "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?." *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*. IEEE, 2015.
- [2] Zhou, Denny, et al. "Learning with local and global consistency." *Advances in neural information processing systems*. 2004.

Apoio e Agradecimentos

Esse projeto foi apoiado e financiado pela Universidade Federal de São Paulo.