

Avaliação do desempenho de classificadores em metalografia quantitativa

Classifiers performance evaluation in quantitative metallography

Bernardo Jordão Moreira Sarruf¹, Rafael de Azevedo Cidade²,
Vitor Py Braga³, Gregório Jório Santana³

¹Laboratório de Hidrogênio - Labh2 – PEMM/COPPE/UFRJ CP: 68505 – 21941-972, Rio de Janeiro, RJ
e-mail: sarruf@labh2.coppe.ufrj.br

²Laboratório de Compósitos - LaCOM – PEMM/COPPE/UFRJ CP: 68505 – 21941-972, Rio de Janeiro, RJ
e-mail: rafaelcidade@metalmat.ufrj.br

³SIM Consultoria em Tecnologia Ltda. – Incubadora de Empresas da COPPE/UFRJ – 21941-614, Rio de Janeiro, RJ

RESUMO

As técnicas convencionais para aplicação da metalografia quantitativa têm sido cada vez mais substituídas por processos digitais devido à crescente necessidade de ganho de velocidade nas análises industriais e o processamento de imagens tem então ocupado um lugar importante nesse campo da metalurgia.

Este trabalho teve como objetivo a implementação e avaliação de técnicas de inteligência e visão computacional na resolução de problemas de determinação de área de fases em aços AISI 1020. Para a comparação, três técnicas de decisão automática foram avaliadas: k-vizinhos mais próximos (KNN), que classifica os pontos da imagem através de correlação com a sua vizinhança e duas técnicas de classificação por aprendizado supervisionado: Redes Neurais Artificiais (RNA) e Máquinas de Vetores Suporte (SVM). Índices de indicação de qualidade de classificação também foram avaliados.

Os resultados mostraram superioridade geral na técnica SVM no que tange, principalmente, o tempo de classificação e acurácia relativa, porém em todos os casos houve concordância com os valores de fração de área esperados para esse tipo de material, segundo análises metalúrgicas teóricas.

Palavras-chave: Classificadores de Imagem, Metalografia quantitativa, Processamento de imagens, Caracterização microestrutural.

ABSTRACT

As the need for increasing speed takes place at industrial processes in general, the use of digital techniques such as image processing and automatic classification have been playing an important role at materials characterization and quantitative metallography fields.

The aim of this work was to develop and evaluate computational vision techniques on solving ordinary problems as the area fraction of phases determination of AISI 1020 steels. Three techniques were implemented and evaluated: k-Nearest Neighbors (KNN) which consists in classifying pixels based on their neighborhood information, Artificial Neural Networks and Support Vectors Machine, these last two centered on supervised machine learning processes. Indexes that denote classification quality were then evaluated.

Concerning classification time and relative accuracy, the SVM results have shown superiority. Nevertheless, in all cases, the classification values have agreed with the area fraction values expected for this type of steel based on theoretical metallurgical analysis.

Keywords: Image classifiers; Quantitative metallography, Image analysis, Microstructural characterization.

1. INTRODUÇÃO

O uso de ferramentas de classificação automática tem se tornado importante para metalografia quantitativa [1].

Na indústria, o uso de técnicas que disponibilizem maior celeridade nos processos e reprodutibilidade

em resultados sem introdução de viés por erro de operação tem sido demandado. Este trabalho visa à comparação entre técnicas de classificação de imagens aplicadas à metalografia, em particular, à determinação de frações de área de constituintes.

A quantificação de parâmetros microestruturais, como tamanho de grão e percentual de fases, é de grande importância no estudo dos materiais a fim de correlacioná-los às propriedades dos mesmos. Tradicionalmente a análise é feita através de um limiar simples na escala de cinza. As técnicas a seguir descritas são capazes de operar análises multidimensionais, incorporando ao processo de quantificação elementos, elementos variados como cor, textura e propriedades estatísticas das micrografias.

2. MATERIAIS E MÉTODOS

Foram selecionadas para teste dos classificadores duas imagens de um aço AISI 1020 que se apresenta bifásico (ferrita + perlita). Na Figura 1.a e 1.b, as imagens apresentam aumentos de 100 e 200 vezes, respectivamente.

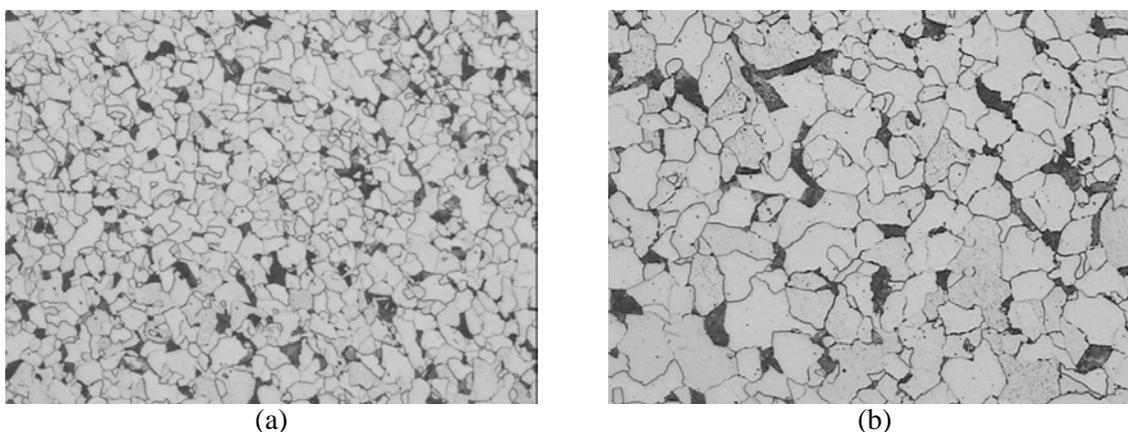


Figura 1: Amostras metalográficas de aço AISI 1020 com aumento de 100 e 200 vezes, respectivamente.

2.1 Preparação de Amostras

A preparação destas amostras contou, basicamente, com três etapas: lixamento, polimento e ataque químico. As amostras, previamente embutidas em baquelite, foram lixadas em lixas d'água de granulometrias 220, 320, 400 e 600, sendo devidamente rotacionadas de 90° entre cada lixa para a completa remoção dos riscos da lixa anterior.

Posteriormente, as amostras foram polidas com alumina, utilizando um disco rotatório coberto com feltro. O ataque químico consistiu de um banho de 20 segundos em NITAL 2% (solução de álcool e ácido nítrico com 2% em peso) a fim de revelar a microestrutura.

A aquisição das imagens foi realizada através de uma câmera fotográfica, acoplada a um microscópio ótico, conectada a um computador por uma placa de captura. Desta maneira, as imagens capturadas, com iluminação em campo claro pelo microscópio, puderam ser exportadas em formato digital.

2.2 Classificadores

O problema da classificação das imagens obtidas através da microscopia foi abordado do ponto de vista da classificação estatística. Algumas técnicas foram escolhidas, com base na popularidade de sua aplicação, e implementadas pela equipe. O mesmo problema foi proposto a cada uma das técnicas de aprendizado de máquina.

Um software comparativo foi desenvolvido pelos autores em linguagem C# e imagens de treino em RGB com matiz zero e baixos níveis de saturação, foram apresentadas aos diferentes classificadores. Portanto os classificadores operaram em um subconjunto do espaço HSB.

2.2.1 K-Nearest Neighborhood (K-NN)

A técnica de k-vizinhos mais próximos classifica objetos de acordo com sua proximidade a dados de treino

em um determinado espaço de atributos (feature space). Neste trabalho, a referida técnica foi implementada tomando um espaço bidimensional composto pelas componentes de saturação e brilho de um espaço de cores HSB.

O eixo de matiz foi desconsiderado nesse caso para tornar a análise bidimensional. A métrica euclidiana é utilizada, através da implementação do kd-tree [2], na tomada de decisão.

2.2.2 Rede Neural Artificial (RNA)

Redes neurais artificiais são um conjunto de modelos matemáticos que funcionam de maneira análoga às redes neurais biológicas. O modelo escolhido para este estudo foi o modelo do perceptron multicamadas retropropagado (multilayer perceptron) [3].

2.2.3 Support Vector Machine (SVM)

O método de máquinas de vetores de suporte realiza a classificação através da construção de um hiperplano n-dimensional que separa de forma ótima dados entre duas categorias [4]. A técnica tem grande similaridade às redes neurais artificiais. Um modelo SVM utilizando uma função sigmóide como kernel, equivale, formalmente, a uma rede neural perceptron de duas camadas [5].

Para a SVM, o princípio da minimização do risco estrutural é aplicado minimizando um limitante superior do risco esperado. Por outro lado, no MLP (multilayer perceptron) a minimização do erro nos dados de treino é feita da forma empírica tradicional. A principal diferença na minimização do risco é melhorar o desempenho da generalização do SVM em comparação a um MLP [6].

A técnica de treinamento em ambas as metodologias foi supervisionada pelo usuário, utilizando-se escolha de pontos aleatórios na imagem e atribuindo ao constituinte metalúrgico correspondente.

2.3 Análise Estatística

Com objetivo de comparar o desempenho dos classificadores adotados, foi criado um roteiro para realização de uma avaliação estatística dos resultados. Esta estatística permitiu a elucidação dos resultados obtidos por cada classificador bem como a eficiência, confiabilidade e qualidade de classificação.

As imagens metalográficas apresentavam duas formações, identificadas como ferrita e perlita, em relação ao nível de cinza. Os testes foram divididos em quatro etapas para cada uma das duas imagens testadas.

Na primeira etapa foram marcados, deliberadamente e em consistência com a observação metalográfica, três pontos da imagem de cada formação, visando o treino da máquina, e realizados dez execuções de classificação. Foram tomadas as quantidades fornecidas como resultado por cada classificador (em %), os índices relacionados à qualidade de classificação, assim como a acurácia média fornecida pela validação cruzada. Posteriormente foram calculadas as médias e desvios padrão dos resultados desta primeira etapa.

O procedimento descrito foi repetido utilizando 6, 12 e 18 pontos de treino para cada imagem (segunda, terceira e quarta etapas, respectivamente). Uma vez executadas as 4 etapas, foram calculadas médias e desvios padrão dos resultados das médias das 4 etapas.

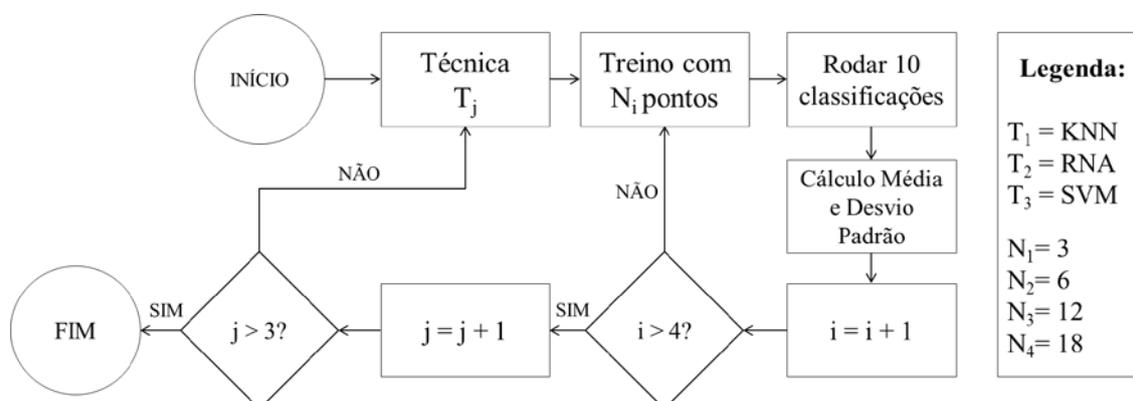


Figura 2: Esquema explicativo para metodologia utilizada nos testes

3. RESULTADOS

As Figuras 3 a 6 mostram os resultados de classificação utilizando as três técnicas e números diferentes de pontos (3 e 18), segundo o roteiro descrito na seção anterior. Observa-se que a técnica K-NN, principalmente, sofre alta influência na quantidade de pontos utilizados na etapa de treinamento da máquina e que o SVM demonstra diferenças quase que imperceptíveis.

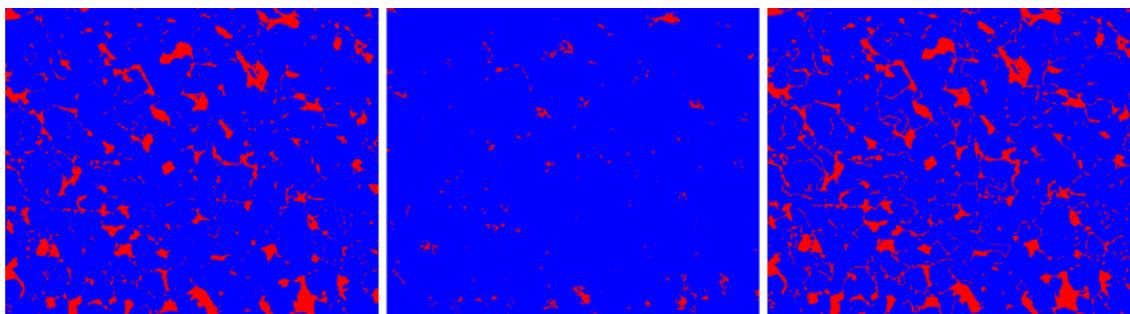


Figura 3: Imagens classificadas da amostra com aumento de 100x com 3 pontos na fase de treinamento para SVM, K-NN e RNA respectivamente.

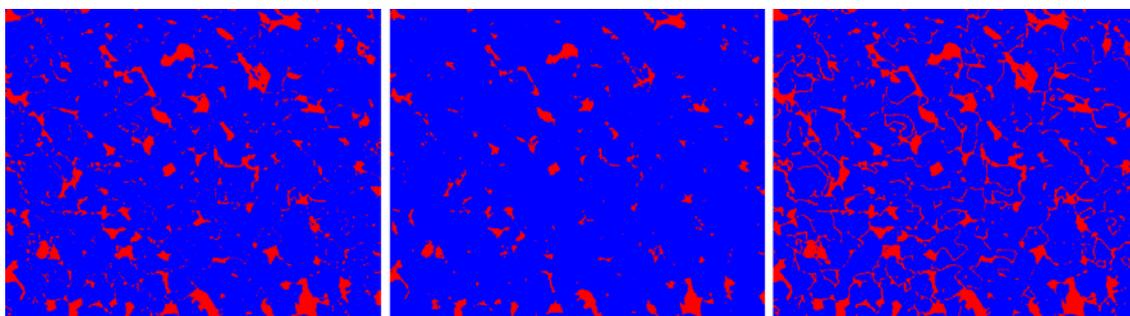


Figura 4: Imagens classificadas da amostra com aumento de 100x com 18 pontos na fase de treinamento para SVM, K-NN e RNA respectivamente.

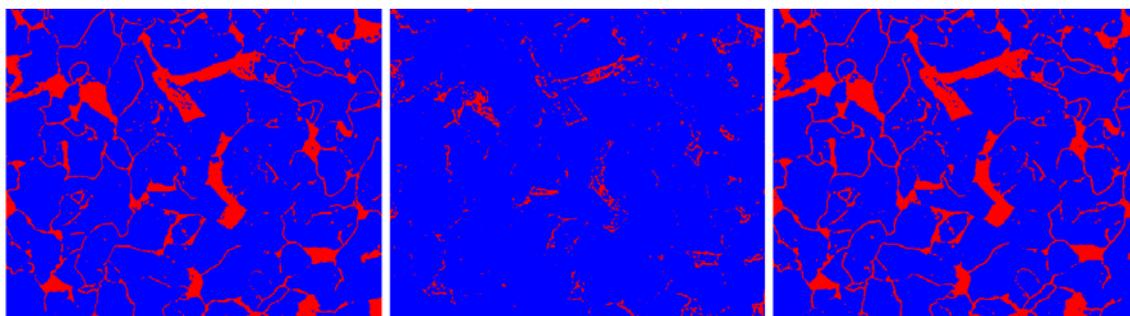


Figura 5: Imagens classificadas da amostra com aumento de 200x com 3 pontos na fase de treinamento para SVM, K-NN e RNA respectivamente.

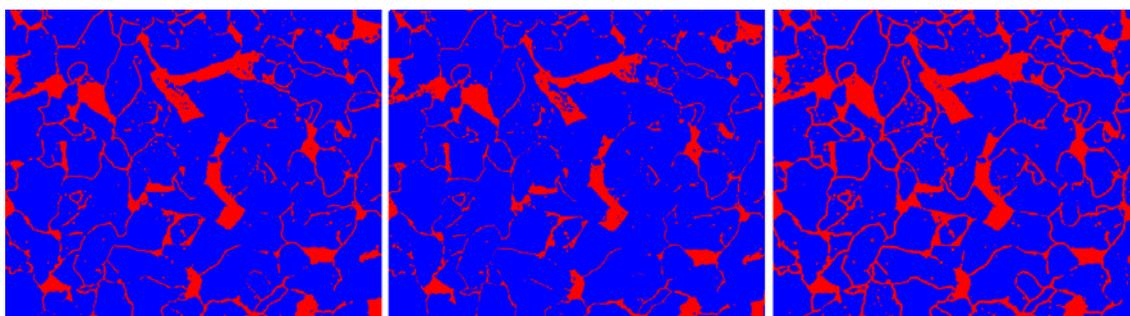


Figura 6: Imagens classificadas da amostra com aumento de 200x com 18 pontos na fase de treinamento para SVM, K-NN e RNA respectivamente.

Os testes foram tratados para cada imagem (Aço AISI 1020 com aumento 100x e aumento de 200x) e para cada técnica de classificação estudada.

As Tabelas 1 a 6 representam a última etapa da metodologia apresentada na Figura 2 e mostram as médias dos resultados de quantidade de fase ferrita (%), quantidade de perlita (%), tempo de execução da classificação, os índices Kappa, Pabak e a acurácia média obtida pela validação cruzada para os vários tipos de testes com os diversos números de pontos de treino utilizados [7].

Tabela 1: Dados experimentais de teste para a imagem com aumento de 100x e utilizando a técnica SVM.

SVM	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	88,11	88,12	88,12	88,12	88,12	0,01
Perlita	11,89	11,88	11,88	11,88	11,88	0,01
Tempo (s)	3,09	3,26	3,50	3,63	3,37	0,24
Kappa	1,00	1,00	1,00	1,00	1,00	0,00
Pabak	1,00	1,00	1,00	1,00	1,00	0,00
Valid. Cruzada	1,00	1,00	1,00	1,00	1,00	0,00

Tabela 2: Dados experimentais de teste para a imagem com aumento de 100x e utilizando a técnica K-NN.

K-NN	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	98,52	97,45	96,02	94,36	96,59	1,80
Perlita	1,48	2,55	3,98	5,64	3,41	1,80
Tempo (s)	4,56	7,78	13,04	18,55	10,98	6,14
Kappa	0,36	0,70	0,83	0,90	0,70	0,24
Pabak	0,36	0,70	0,83	0,90	0,70	0,24
Valid. Cruzada	0,50	0,72	0,86	0,88	0,74	0,17

Tabela 3: Dados experimentais de teste para a imagem com aumento de 100x e utilizando a técnica RNA.

RNA	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	82,95	82,94	82,66	82,39	82,73	0,26
Perlita	17,05	17,06	17,34	17,61	17,27	0,26
Tempo (s)	11,07	10,76	11,14	10,90	10,97	0,17
Kappa	1,00	1,00	1,00	1,00	1,00	0,00
Pabak	1,00	1,00	1,00	1,00	1,00	0,00
Valid. Cruzada	0,97	1,00	1,00	1,00	0,99	0,02

Tabela 4: Dados experimentais de teste para a imagem com aumento de 200x e utilizando a técnica SVM.

SVM	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	85,76	85,80	85,85	85,88	85,82	0,05
Perlita	14,24	14,20	14,15	14,12	14,18	0,05
Tempo (s)	3,32	3,52	3,73	3,57	3,53	0,17
Kappa	0,93	0,87	0,96	0,97	0,93	0,05
Pabak	0,93	0,87	0,96	0,97	0,93	0,05
Valid. Cruzada	0,95	0,93	0,98	0,99	0,96	0,02

Tabela 5: Dados experimentais de teste para a imagem com aumento de 200x e utilizando a técnica K-NN.

K-NN	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	96,66	92,58	91,37	90,09	92,67	2,84
Perlita	3,34	7,42	8,63	9,91	7,33	2,84
Tempo (s)	4,13	8,85	14,99	19,37	11,84	6,71
Kappa	0,33	0,67	0,85	0,90	0,69	0,26
Pabak	0,33	0,67	0,85	0,90	0,69	0,26
Valid. Cruzada	0,50	0,69	0,86	0,87	0,73	0,17

Tabela 6: Dados experimentais de teste para a imagem com aumento de 200x e utilizando a técnica RNA.

RNA	3 Pontos	6 Pontos	12 Pontos	18 Pontos	Média	Desvio Padrão
Ferrita	83,14	80,90	82,40	81,91	82,09	0,94
Perlita	16,86	19,10	17,60	18,10	17,91	0,94
Tempo (s)	10,97	10,93	10,94	11,28	11,03	0,17
Kappa	1,00	1,00	0,99	1,00	1,00	0,00
Pabak	1,00	1,00	0,99	1,00	1,00	0,00
Valid. Cruzada	0,97	0,96	0,98	1,00	0,98	0,02

As Tabelas 7 e 9 apresentam os valores médios, obtidos por cada método, dos constituintes das imagens com aumentos de 100x e 200x respectivamente, o valor teórico obtido através do diagrama de equilíbrio Fe-Fe₃C, utilizando a regra da alavanca, para a composição do aço utilizado (AISI 1020) e o valor obtido através de binarização da imagem através da adoção de um valor limiar na escala de cinza. As Tabelas 8 e 10 mostram a discrepância entre os resultados de cada método e aqueles esperados segundo a regra da alavanca.

Tabela 7: Percentual médio dos microconstituintes para cada método (aumento 100x)

	SVM	KNN	RNA	Teórico	Limiar
Ferrita	88,12	96,59	82,73	75,88	73,18
Perlita	11,88	3,41	17,27	24,12	26,82

Tabela 8: Discrepância relativa (%) entre os valores obtidos e os valores teóricos do percentual dos constituintes (aumento 100x)

	SVM	KNN	RNA
Ferrita	14,26	21,44	8,27
Perlita	50,74	85,86	28,39

Tabela 9: Percentual médio dos microconstituintes para cada método (aumento 200x)

	SVM	KNN	RNA	Teórico	Limiar
Ferrita	85,82	92,67	82,09	75,88	83,46
Perlita	14,18	7,33	17,91	24,12	16,54

Tabela 10: Discrepância relativa (%) entre os valores obtidos e os valores teóricos do percentual dos constituintes (aumento 200x)

	SVM	KNN	RNA
Ferrita	11,58	18,12	7,56
Perlita	41,21	69,61	25,75

4. DISCUSSÃO

Observa-se, nas Figuras 3, 4, 5 e 6, que o classificador define como perlita (nesse caso a formação escura) os contornos de grão e inclusões da micrografia. Esta representa uma tendência errônea das técnicas e metodologias de classificação automática utilizadas que pode ser sanada como a utilização de alguns filtros já aplicados e bem consolidados em processamento digital de imagens. Da mesma forma, é possível que em maiores aumentos, ie. 500x, as lamelas de ferrita eutetóide presentes na perlita sejam classificadas como ferrita proeutetóide, superestimando seu percentual. A solução para a classificação das lamelas de perlita pode partir da inclusão de áreas e formas topológicas como dimensões a serem analisadas pelos classificadores, além do espaço de cores. Porém o desenvolvimento destes algoritmos foge ao escopo deste trabalho.

Em relação às inclusões, estas podem ser classificadas através de pós-processamento, por filtragem de pequenas áreas de maneira a propiciar sua posterior mensuração.

Nos gráficos da Figura 7 é mostrado o módulo da diferença de classificação entre cada uma das técnicas. Conforme o esperado, o SVM e a RNA apresentam resultados semelhantes devido à proximidade de sua formulação matemática.

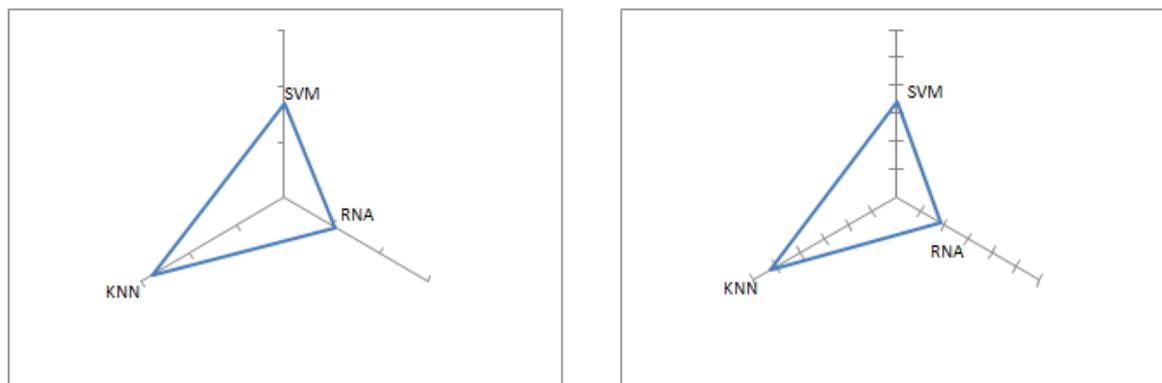


Figura 7: Módulo da diferença de resultado de classificação entre as técnicas para as imagens com aumento de 100x e 200x respectivamente

É perceptível, pelas Tabelas 1 a 6, que o classificador SVM apresenta os menores desvios padrão para os dois tipos de imagem, seguido do classificador que utiliza a Rede Neural Artificial. O classificador que usa a técnica de K vizinhos mais próximos (K-NN) é o que demonstra maior discrepância nos seus resultados que é inversamente proporcional à quantidade de pontos de treino utilizados. As Tabelas 7 e 9 mostram a mais alta dispersão entre aumentos para a técnica de limiarização assistida, exatamente pelo fato de ser a mais subjetiva. Nota-se através das Tabelas 8 e 10 que, em contrapartida ao desvio padrão, o classificador RNA apresenta os menores valores de discrepância relativa em relação aos valores teóricos do percentual dos constituintes para ambas os aumentos (100x e 200x). Observa-se que o desvio padrão para a técnica SVM converge suavemente com aumento do número de pontos de teste e para as outras técnicas parece experimentar um máximo com posterior convergência. Para que essa conclusão seja definitiva, seria necessária uma análise levando-se em consideração outras classificações incluindo cada vez mais pontos.

O tempo de execução da classificação é influenciado por fatores de treino e pela técnica escolhida. A complexidade algorítmica da busca em uma kd-tree implica em um tempo de execução cada vez maior em função da quantidade de dados de treino na técnica k-vizinhos mais próximos. Para as técnicas SVM e RNA, onde a construção do hiperplano de separação tem seu custo apenas na inicialização [6], identifica-se que possuem um tempo de busca praticamente constante independente da quantidade de pontos. Nos resultados foi apresentada a possibilidade de inclusão de novos grupos de pontos na estatística de teste, essa ação poderia levar a uma conclusão mais definitiva, mostrando perfeitamente o momento da experimentação de um máximo no comportamento entre desvio padrão/número de pontos de teste (Tabelas 1 a 6) e ainda expor que a técnica SVM continuará com o seu comportamento estável.

5. CONCLUSÕES

Dentre as metodologias utilizadas, o SVM possui maior robustez em relação à quantidade de pontos de teste considerados. No entanto a qualificação do classificador como mais eficaz, em relação aos valores teóricos, demanda um estudo de representatividade do número de campos observados. Contudo, a proximidade dos resultados obtidos pelo SVM e pela RNA leva à avaliação do tempo de execução e a sensibilidade de dados

de treino como critérios de escolha. O que aponta o SVM como o candidato mais indicado para este tipo de análise, inclusive em problemas com maiores dimensões e fatores de análise.

Problemas relacionados aos contornos de grão e inclusões abordados anteriormente, apresentam relativa facilidade de solução, fazendo com que a escolha e a devida aplicação da técnica computacional sejam realmente o ponto de partida para implementação de classificadores mais eficazes.

6. AGRADECIMENTOS

Os autores gostariam de prestar os agradecimentos aos colegas, professores e pesquisadores do Programa de Engenharia Metalúrgica e de Materiais da COPPE/UF RJ.

7. BIBLIOGRAFIA

- [1] GOMES, O. F. M., PACIORNIK, S., “Automatic Classification of Graphite in Cast Iron.” *Microscopy and Microanalysis*, 11, 363–371, 2005
- [2] NIGSCH, F., BENDER, A., VAN BUUREN, B., “Melting Point Prediction Employing k-Nearest Neighbor Algorithms and Genetic Parameter Optimization”, *Journal of Chemical Information and Modeling*, 2006 46 (6), 2412-2422
- [3] GALLANT, S. I., “Perceptron-Based Learning Algorithms.”, *IEEE TRANSACTIONS ON NEURAL NETWORKS*, v. 1, n. 2, Jun. 1990
- [4] CHANG, C., LIN, C., “LIBSVM: a library for support vector machines”, 2001. Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [5] COLLOBERT, R., BENGIO, S.,- ACM International Conference Proceeding Series, v. 69, p. 23, 2004
- [6] FRIAZ-MARTINEZ, E., SANCHEZ, A., VELEZ, J., “Support vector machines versus multi-layer perceptrons for efficient off-line signature recognition.”, *Engineering Applications of Artificial Intelligence*, v. 19, n. 6, pp. 693-704 , 2006.
- [7] LOBÃO, J., FRANÇA-ROCHA, W., SILVA, A., “Aplicação dos Índices KAPPA & PABAK na validação da classificação automática de imagem de satélite em Feira de Santana-BA”, In: Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril 2005, INPE, p. 1207-1214