

Relatório das práticas de aprendizado de máquina.

Marcus V. S. Maziero¹, Paulo R. K. Nakaima², Vitor Hugo Borges Basseto³

Abstract—Este relatório apresenta os resultados encontrados em 5 práticas de aprendizado de máquinas para classificação de imagens, mais especificamente, classificação de emoções com base em expressões faciais. A partir da Seção V no final de cada uma das seções encontram-se os links para o código fonte para desenvolvimento e exibição, sendo que o último possibilita ao leitor a execução imediata dos experimentos em qualquer navegador moderno.

Keywords—Inteligência Artificial, Emoções, Imagens, Machine Learning.

I. INTRODUÇÃO

As práticas de aprendizado de máquina - ML(Machine Learning) em inglês, é uma das maneiras de utilizar IA(Inteligência Artificial) para aprimorar o conhecimento de aplicações sobre diversos assuntos, proporcionando um treinamento mais natural para a máquina.

O aprendizado de máquina consiste em promover para a aplicação experiências com diversos cenários para que o mesmo após determinado tempo e quantidade de cenários aprenda e modifique seu comportamento.

Existem diversas empresas que já utilizam da aprendizagem de máquina para alcançar resultados benéficos entre elas: Facebook, Google, Uber e Apple. Em algumas dessas empresas o ML é utilizado para recomendações de produtos e análise de tráfego.

Neste contexto foi realizado esse relatório com a proposta de demonstrar os resultados alcançados por intermédio de práticas propostas sobre o aprendizado de máquina.

¹Marcus V. S. Maziero, ²Paulo R. K. Nakaima e ³Vitor Hugo Borges Basseto estão vinculados à Universidade Tecnológica Federal do Paraná, Av. Alberto Carazzai, 1640, Cornélio Procopio, Brasil. marcus.maziero@outlook.com, nakaima@alunos.utfpr.edu.br, vitorhugobasseto@gmail.com

II. OBJETIVO

III. RESUMO - DEEP LEARNING

IV. PRÁTICA 01

Nesta prática busca-se selecionar as bases de imagens, os descritores para extração de características e os classificadores.

A. Base de dados

Foi selecionada a base de imagens CKPLUS. A qual pode ser encontrada em: <https://www.kaggle.com/shawon10/ckplus>.

Contém 981 imagens de rostos em escalas de cinza com as expressões de raiva, nojo, desprezo, medo, felicidade, tristeza, surpresa. Para este trabalho foram selecionadas apenas as expressões raiva, medo, felicidade, tristeza e surpresa.

A segunda base de imagens selecionada foi a Yale Face Database. A qual pode ser encontrada em: <http://vision.ucsd.edu/content/yale-face-database>. Contém 165 imagens em tons de cinza de rostos de 15 indivíduos fazendo 6 expressões faciais, normal, tristeza, sonolência, surpresa e piscando.

B. Extração de características

Para extração de características foi utilizado os descritores Local Binary Patterns (LBP) e o Gabor. Sendo que o primeiro extrai 256 características de textura é não paramétrico, possui baixo custo de processamento [1]. Seu funcionamento consiste em dividir a imagem em sub-blocos sendo que para cada sub-bloco é calculado o histograma, o vetor de características é formado com a concatenação desses histogramas [1]. O descritor Gabor extrai 60 características de textura como linhas e contornos. Os arquivos finais podem ser encontrados em: <https://drive.google.com/drive/folders/1ZP-CkuoP2mQggsR2sd9QGcbks1hrt68u?usp=sharing>.

C. Seleção dos classificadores

Os classificadores utilizados nas práticas seguintes foram selecionados com base em [1]. Os classificadores são Regressão Logística (Logistic Regression - LR), K vizinhos mais próximos (k-nearest neighbors - KNN), Máquina de Vetores Suporte (Support Vectors Machine - SVM), rede perceptron multicamadas (Multi-layer Perceptron - MLP).

V. PRÁTICA 02

Nesta prática busca-se comparar a acurácia obtida por diferentes classificadores a partir dos arquivos gerados na Seção IV.

A. Descritor LBP

1) *Seleção da técnica de normalização:* A partir dos arquivos gerados com os descritores foram alteradas as técnicas de normalizações para verificar qual delas atinge melhor resultado para acurácia. Para a base de dados CKPLUS com o descritor LBP as diferenças são mínimas como podem ser vistas na Figura 1. Os classificadores utilizados foram: LR, KNN, SVM, MLP.

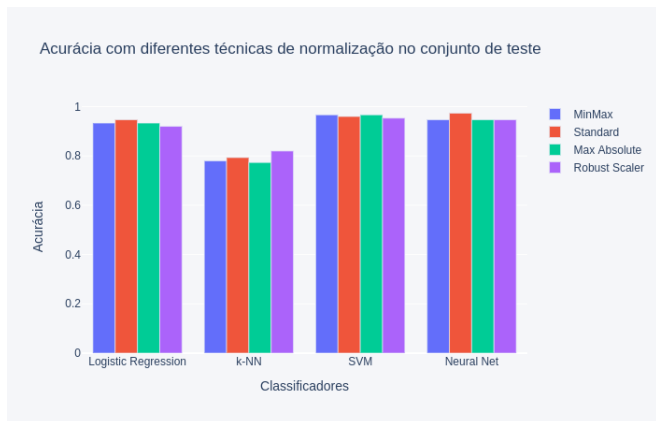


Fig. 1. Comparação de técnicas de normalização com descritor LBP.

O classificador MLP com a normalização Standard foi a que obteve o melhor resultado, 97% de acurácia. O pior resultado foi obtido pelo classificador KNN com a normalização Max Absolute, 77%.

B. Descritor Gabor

Para o descritor Gabor as diferenças entre os resultados obtidos foi alta para o classificador

KNN, contudo nenhum dos classificadores obtiveram 50% de acurácia. A técnica de normalização Min-Max que consiste em transformar cada característica com valor mínimo em 0 e os valores máximos em 1, sendo que o restando é transformado em um valor decima entre 0 e 1, obteve a melhor acurácia com o classificador KNN, 41%. O pior desempenho foi obtido pelo classificador KNN com a normalização Standard, 21%. Os resultados podem ser vistos na Figura 2. Foram mantidos os classificadores utilizados no último experimento.

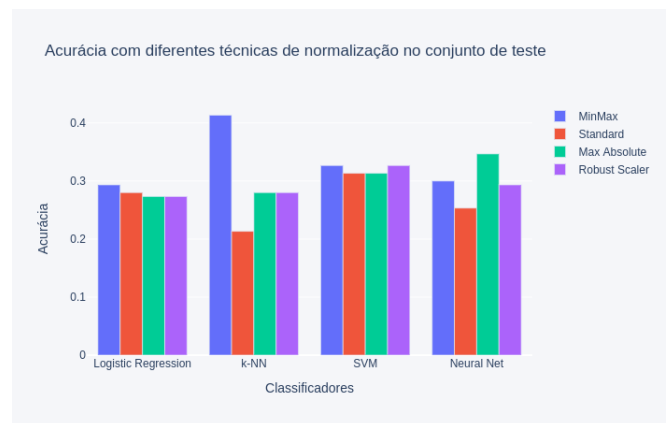


Fig. 2. Comparação de técnicas de normalização com descritor Gabor.

1) *Desenvolvimento:* Prática 02 - ckplus - LBP - all - norm; Prática 02 - ckplus - Gabor - all - norm

2) *Exibição e Execução:* ATUALIZAR

VI. PRÁTICA 03

Neste experimento busca-se utilizar aprendizado não supervisionado com o classificador *k-means* de modo a obter o melhor valor de *k*. Além disto busca-se reduzir a dimensão do conjunto de dados de modo a reter 90% de variância.

A. Descritor LBP

1) *Redução de dimensionalidade:* Buscou-se reduzir a dimensionalidade do conjunto de dados obtidos com o descritor LBP de modo a reter 90% de variância. Primeiramente os dados foram normalizados com a técnica Standard, ou seja, são calculados a média e o desvio padrão da conjunto de amostras, em seguida é subtraída de cada amostra a média, o resultado então é dividido

pelo desvio padrão. Esta técnica foi selecionada pois obteve resultado levemente superior as outras técnicas na Seção V.

Com a técnica PCA para redução de dimensionalidade foi selecionado o menor número de componentes que retinha 90% de variância. Ver Figura 3. Foi possível reduzir de 256 para 133 característica.

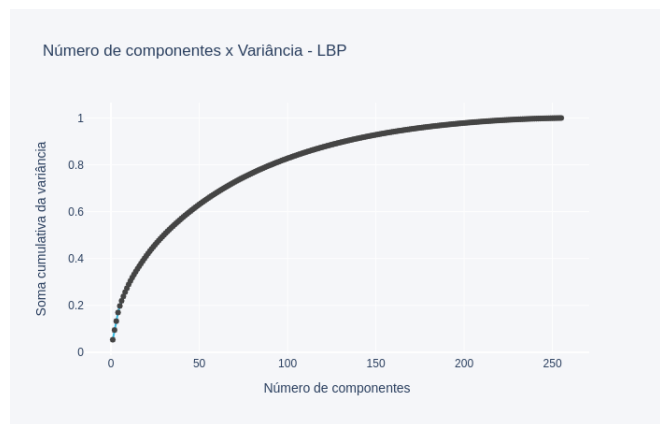


Fig. 3. Selecionar o menor número de componentes retendo 90% de variância.

2) *Agrupamento:* Com a redução de dimensionalidade obtida foi utilizado o classificador k-means agrupando as amostras de 2 a 30 agrupamentos de forma iterativa. Ao fim, foi compilado e plotado a variância para cada quantidade de agrupamentos para aplicar o método elbow com intuito de definir a melhor quantidade de agrupamentos. O ponto mais distante da reta formada pela ligação do ponto inicial ao final foi 9, o melhor número de agrupamentos. Ver Figura 4. O agrupamento final foi plotado destacando-se os centróides na Figura 5.

B. Descritor Gabor

1) *Redução de dimensionalidade:* Foi aplicada a técnica PCA para as características extraídas com o descritor Gabor. Primeiramente os dados foram normalizados com a técnica MinMax, melhor resultado obtido no experimento anterior. Foi possível reduzir as características de 60 para duas retendo 100% de variância. Ver Figura 6.

2) *Agrupamento:* Na sequência foi utilizado o classificador k-means agrupando as amostras de 2 a 30 agrupamentos de forma iterativa. Ao fim, foi compilado e plotado a variância para cada

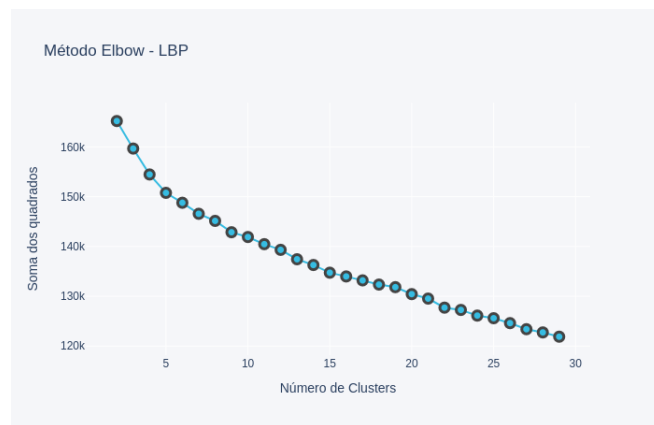


Fig. 4. Selecionar o melhor número de agrupamentos.

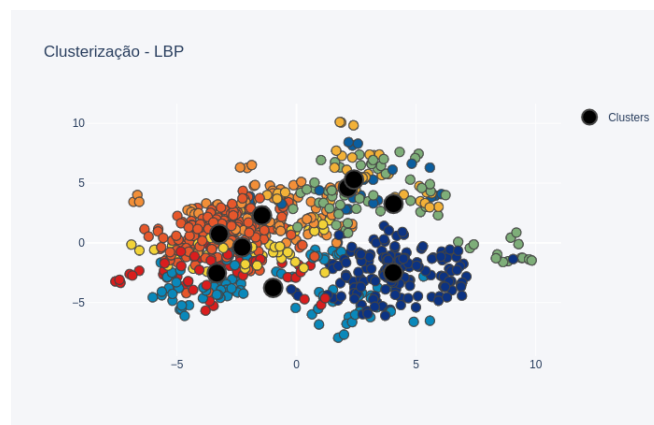


Fig. 5. Conjunto de amostras agrupados em 9.

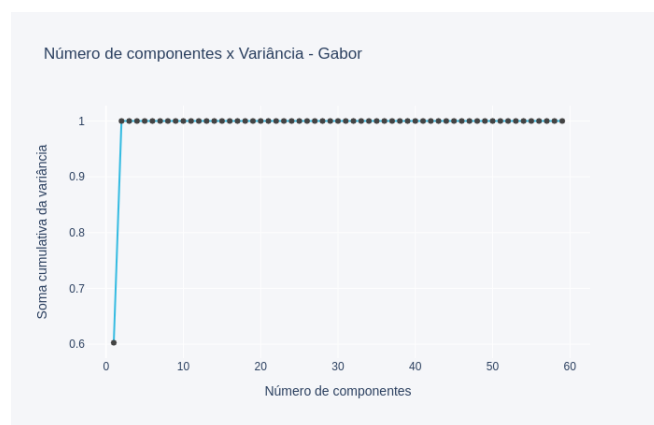


Fig. 6. Selecionar o menor número de componentes retendo 100% de variância.

quantidade de agrupamentos para aplicar o método elbow com intuito de definir a melhor quantidade de agrupamentos. O ponto mais distante da reta formada pela ligação do ponto inicial ao final foi 8, o melhor número de agrupamentos. Ver Figura 7. O agrupamento final foi plotado destacando-se os centróides na Figura 8.

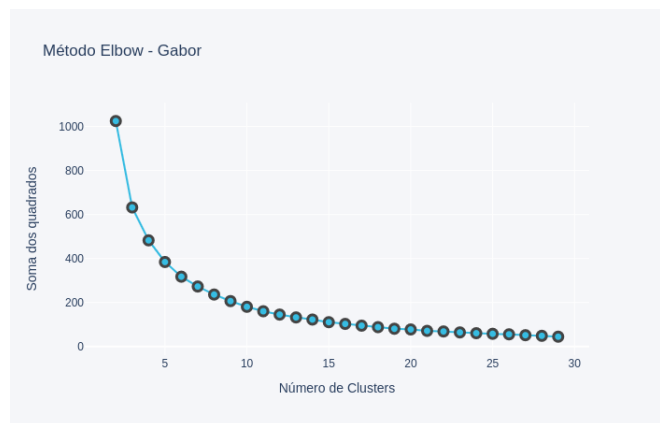


Fig. 7. Selecionar o melhor número de agrupamentos.

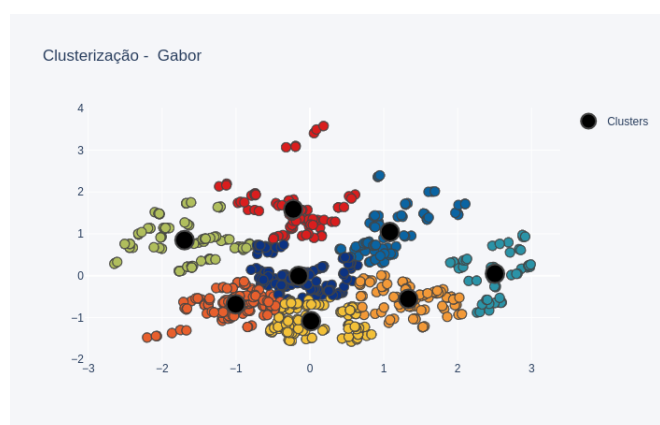


Fig. 8. Conjunto de amostras agrupados em 8.

Nota-se que o agrupamento plotado na Figura 8 possui fronteiras mais definidas do que na Figura 5.

3) *Desenvolvimento*: Prática 03 - ckplus - LBP - pca - kmeans; Prática 03 - ckplus - Gabor - pca - kmeans

4) *Exibição e Execução*: ATUALIZAR

VII. PRÁTICA 04

Nesta prática busca-se comparar os resultados de aprendizado supervisionado realizado nos

experimentos anteriores com o método Label Propagation.

A. *Descritor LBP*

1) *Seleção dos parâmetros para o método Label Propagation*: O método Label Propagation pode ser utilizado com o kernel de k vizinhos mais próximos (K-Nearest Neighbors - KNN) e função de base radial (Radial Basis Function - RBF). Para o primeiro caso deve-se definir o número máximo de iterações o número de vizinhos e outros. Para o RBF deve-se definir o valor gamma, a quantidade máxima de iterações e outros.

Para este experimento variou-se o kernel entre KNN e RBF, o valor gama em 20, 25, 30, a quantidade máxima de iterações em 300, 500, 1000 e o número de vizinhos em 7, 14, 21 e 28. Para cada combinação foi feita a validação cruzada em 5 partições. O critério para definição dos melhores parâmetros foi a média da acurácia. Os melhores parâmetros para o kernel KNN e RBF estão descritos na Tabela I. Para todos os classificadores o desempenho da acurácia apresentou queda quando comparado as técnicas supervisionadas. Contudo existe a redução de 600 imagens rotuladas para 360. A comparação com os resultados das práticas supervisionadas podem ser vistas na Figura 9.

TABLE I
MÉDIA E DESVIO PADRÃO PARA ACURÁCIA

Kernel	Iteração	Vizinhos	Gamma	Média	Desv. Pad.	Pos.
KNN	300	7	n/a	0.39	0.021	1
RBF	300	n/a	20	0.10	0.003	13

2) *Desenvolvimento*: Prática 02 - ckplus - LBP - standard; Prática 04 - ckpuls - LBP - gridsearch - both; Prática 04 - ckpuls - LBP - supervisionado vs semi-supervisionado

3) *Exibição e Execução*: ATUALIZAR

VIII. PRÁTICA 05

Nesta prática busca-se

IX. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi realizado o reconhecimento de expressões faciais com aprendizado de



Fig. 9. Supervisionado vs Semi-supervisionado.

máquina. Foram comparadas com base na acurácia o melhor desempenho para os descritores LBP e Gabor. O primeiro foi que obteve melhor resultado. Foram comparadas diferentes técnicas de normalização sendo que os testes realizados com o descritor Gabor foi o que apresentou maior sensibilidade para este tipo de alteração. Além disto foi aplicada a técnica de redução de dimensionalidade com o PCA, novamente os resultados baseados no descritor Gabor foi o mais afetado.

REFERENCES

- [1] S. Rajan, P. Chenniappan, S. Devaraj, and N. Madian. Facial expression recognition techniques: a comprehensive survey. *IET Image Processing*, 13(7):1031–1040, 2019.