

Comparação de Classificadores

Igor Wandermurem Dummer¹

Vitória, ES

UFES[†]

Abstract

O trabalho consiste na análise e construção de classificadores utilizados na detecção de classes de uma base de dados de lâmpadas. Na proposta do trabalho, foram utilizados 5 classificadores: Dummy Zero R (ZR), Bagging (BA), AdaBoost (AB), RandomForest (RF) e Heterogeneous Pooling (HP), este último sendo necessária sua construção. Para questão de avaliação, suas acurácias são comparadas e são feitas observações gerais de como os classificadores se performaram, empregando matrizes de confusão e gráficos como *boxplots*.

Keywords: classificadores, HP, análise de acurácia

1. Introdução

Este trabalho consiste em realizar uma comparação experimental entre um conjunto pré-definido de técnicas de aprendizado e classificação automática aplicadas a um problema de classificação. As técnicas escolhidas são: ZeroR (ZR), Bagging (BA), AdaBoost (AB), RandomForest (RF) e Heterogeneous Pooling (HP).

A base de dados é descrita na Seção 2. Os resultados são apresentados na Seção 4 e analisados na Subseção 5.1.

2. Base de dados

A base de dados usada no trabalho foi obtida de um projeto de pesquisa que visa utilizar
10 informações de imagens de lâmpadas e luminárias de iluminação pública para atualização de cadastro das concessionárias de energia. A base completa contém 297 exemplos de imagens de lâmpadas e luminárias.

[†]Matrícula: 2019109389

2.1. Descrição do domínio

A partir da imagem colorida do dispositivo e suas representações, é realizado o processamento das imagens e são extraídas características para compor a base de dados. Três tipos de técnicas foram aplicadas nessa extração. Descritores de Fourier utilizam a imagem binária e são usados para descrever o contorno do objeto. Descritores de Hu usam a imagem em escala de cinza e definem um conjunto de momentos invariantes para descrever a imagem. Os descritores de Haralick usam uma abordagem estatística para descrever as texturas da
20 imagem com base na distribuição e relacionamento da escala de cinza da imagem. A base completa obtida contém 10 descritores de Fourier, 7 descritores de Hu e 6 descritores de Haralick, totalizando 23 características.

O domínio foi decidido a partir do último dígito da matrícula (2019109389). Foi decidido que, para o dígito 9, seria considerado os 7 descritores de Hu e 6 descritores de Haralick.

2.2. Definição das classes e das características

As classes são caracterizadas pelo tipo: Lâmpada de vapor de mercúrio (Mercury Vapor - MV), lâmpada de vapor de sódio de alta pressão (High Pressure Sodium - HPS) e lâmpada de vapor metálico (Metal Halide – MH) e pela potência: 70, 100, 125, 150, 250 ou 400W da lâmpada. Nota-se que a base é razoavelmente balanceada.

30 2.3. Número de instâncias

A base de dados contém um total de 297 exemplos de imagens de lâmpadas e luminárias de iluminação pública. Esses exemplos são usados para treinar e testar os classificadores nas diferentes etapas do experimento.

3. O método Heterogeneous Pooling

O Heterogeneous Pooling (HP) é um método de classificação que combina diferentes classificadores para realizar a tarefa de classificação de instâncias. Os classificadores utilizados como base são a Árvore de Decisão (DT), o Naive Bayes Gaussiano (NB) e o K Vizinhos Mais Próximos (KNN). O HP adota uma estratégia de votação majoritária para determinar a classe final de uma instância. Cada classificador base emite uma predição para a instância
40 e a classe mais votada entre os classificadores é selecionada como a classe final. Em situações

em que ocorre um empate na votação, a classe mais frequente na base de treinamento original, entre as classes empatadas, é escolhida. A quantidade de classificadores base utilizados no HP é definida pelo parâmetro *n_samples*. Esse parâmetro indica o número de vezes que os classificadores base serão usados para compor o conjunto combinado.

4. Descrição dos experimentos realizados e seus resultados

O procedimento experimental será dividido em duas etapas. A primeira etapa consiste no treino e teste com 3 rodadas de validação cruzada estratificada de 10 folds do classificador que não possui hiperparâmetros, isto é, o classificador ZR. A segunda etapa consiste no treino, validação e teste dos classificadores que precisam de ajuste de hiperparâmetros, isto é, os classificadores BA, AB, RF e HP. Neste caso o procedimento de treinamento, validação e teste será realizado através de 3 rodadas de ciclos aninhados de validação e teste, com o ciclo interno de validação contendo 4 folds e o externo de teste com 10 folds. A busca em grade (grid search) do ciclo interno deve considerar os seguintes valores de hiperparâmetros de cada técnica de aprendizado: A busca na grade do loop interno considerou os seguintes valores de hiperparâmetros de cada técnica de aprendizado:

Bagging	AdaBoost	RandomForest	HeterogeneousPooling
3	3	3	1
9	9	9	3
15	15	15	5
21	21	21	7

Table 1: Valores de hiperparâmetros utilizados em *grid search*.

Podemos vislumbrar os resultados na Tabela 2. O classificador ZeroR tem a pior performance em decorrência da base ser rasoavelmente balanceada, visto que este sempre classifica com a classe predominante. Em seguida, temos uma performance também insatisfatória do classificador AdaBoost, que é superado pelo Bagging, RandomForest e HeterogeneousPooling. Entre esses três, HeterogeneousPooling, que foi implementado, mostrou ter pior performance em relação aos outros dois, que possuem valor de média próximos, mas destacando o RandomForest que possui menor desvio padrão. Portanto, os melhores classificadores são Bagging e RandomForest (para o problema proposto).

Método	Média	Desvio Padrão	Limite Inferior	Limite Superior
ZR	0.165	0.011	0.161	0.169
BA	0.649	0.090	0.617	0.681
AB	0.338	0.031	0.327	0.349
RF	0.649	0.066	0.625	0.673
HP	0.580	0.095	0.546	0.613

Table 2: Medidas estatísticas dos resultados dos métodos.

Na tabela 3 podemos visualizar os p-values obtidos através das métricas *wilcoxon* (matriz triangular inferior) e *t pareado* (matriz triangular superior).

Os p-values que são menores que o nível de significância (0.05) são marcados em negrito, mostrando que existe uma diferença significativa estatisticamente entre os classificadores, em termos de desempenho. Como dito anteriormente, os classificadores que tiveram melhor performance para o problema proposto foi o Bagging e o RandomForest, e isso fica visível na tabela.

ZeroR	0.0	0.0	0.0	0.0
0.0	BG	0.0	0.9834	0.000576
0.0	0.0	AB	0.0	0.0
0.0	0.9635	0.0	RF	0.000129
0.0	0.001650	0.0	0.000474	HP

Table 3: Testes estatísticos entre cada par de modelo.

É possível observar as diferenças e semelhanças entre os classificadores por meio de um gráfico boxplot. Essa representação permite destacar os detalhes mencionados anteriormente, como a baixa performance do ZeroR e melhor performance entre Bagging e RandomForest.

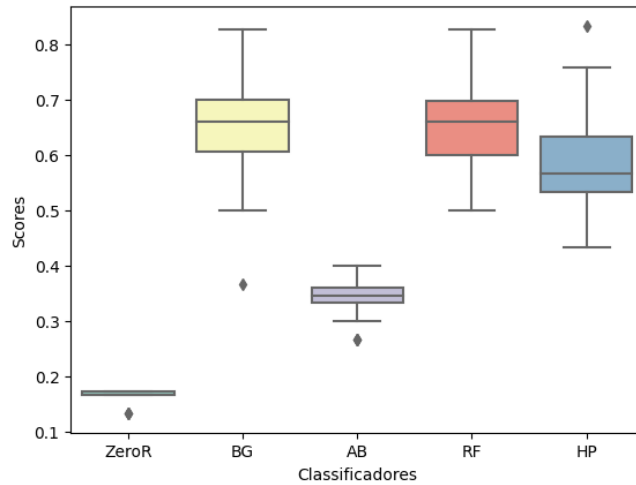


Figure 1: Boxplot dos resultados estatísticos dos 5 classificadores.

5. Conclusões

5.1. Análise geral dos resultados

De forma geral, concluímos que o classificador construído obteve resultados consideráveis, tendo melhor performance que o ZeroR e AdaBoost, mas performance pior que Random-Forest e Bagging, estes que obtiveram resultados muito similares.

80 5.2. Contribuições do trabalho

O trabalho abordou a comparação entre diferentes classificadores, mostrando na prática a diferença de comportamento entre eles. Temos, principalmente, o foco no Heterogeneous Pooling, que é um classificador menos frequentemente utilizado em relação aos demais.

5.3. Melhorias e trabalhos futuros

Para trabalhos futuros podemos utilizar de outras bases de dados e diferentes classificadores, também implementados.

References

Todo o referencial teórico foi retirado do material disponibilizado pelo professor Flávio Varejão na plataforma Google Classroom. Em especial, a própria especificação do trabalho.