

Estudo sobre um classificador adaptativo para análise de dados de imagens em alta dimensionalidade com um conjunto limitado de amostras de treinamento

Domênica da Silva Sequeira¹, Delnir Monteiro de Lemos^{1*}
*Orientador

¹Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) -
Campus Rio Grande. Rio Grande, RS, Brasil.

Em cenas naturais, ocorrem com certa frequência classes espectralmente muito similares, isto é, os vetores média são muito próximos. Em situações como esta dados de baixa dimensionalidade (*LandSat-TM*, *Spot*) não permitem uma classificação acurada da cena. Por outro lado, sabe-se que dados em alta dimensionalidade tornam possível a separação destas classes, desde que as matrizes covariância sejam suficientemente distintas. Neste caso, o problema de natureza prática que surge é o da estimação dos parâmetros que caracterizam a distribuição de cada classe. Na medida em que a dimensionalidade dos dados cresce, aumenta o número de parâmetros a serem estimados, especialmente na matriz covariância. Contudo, é sabido que, no mundo real, a quantidade de amostras de treinamento disponíveis, é frequentemente muito limitada, ocasionando problemas na estimação dos parâmetros necessários ao classificador, degradando, portanto a acurácia do processo de classificação, na medida em que a dimensionalidade dos dados aumenta. O *Efeito de Hughes*, como é chamado este fenômeno, já é bem conhecido no meio científico, e estudos vêm sendo realizados com o objetivo de mitigar este efeito. Entre as alternativas propostas com a finalidade de mitigar o *Efeito de Hughes*, encontram-se as técnicas que utilizam amostras não rotuladas e amostras semi-rotuladas para reduzir o problema do tamanho reduzido das amostras de treinamento. Neste estudo teve-se por objetivos investigar uma metodologia que permita uma estimação mais confiável dos parâmetros do classificador quando se utiliza dados em alta dimensionalidade e desenvolver uma ferramenta computacional para implementar e testar essa metodologia. Para tanto, foram utilizadas funções de decisão mais ajustadas à realidade da cena analisada. Os resultados mostram que, com um número limitado de amostras de treinamento, as técnicas que utilizam algoritmos adaptativos mostraram-se eficientes em reduzir o *Efeito de Hughes*. Quanto a acurácia, em todos os casos o modelo quadrático mostra-se eficiente através do algoritmo adaptativo, apesar do *Efeito de Hughes*.

Palavras-chave: Alta Dimensionalidade. Classificador Adaptativo. Reconhecimento de Padrões. Sensoriamento Remoto.

Trabalho executado com recursos do Edital PROPPI Nº 014/2015 – Fomento Interno 2016/2017, da Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação.