

HOG descriptors

December 2025

1 Histogram of Oriented Gradients (HOG)

Originally proposed by Dalal and Triggs (2005), Histogram of Oriented Gradients (HOG) descriptors capture the distribution of local gradient orientations rather than raw pixel intensities. This representation is considerably more robust to noise, small spatial deformations, variations in stroke thickness and illumination — all of which are common in handwritten character datasets. Dalal and Triggs demonstrated that HOG descriptors “significantly outperform existing feature sets” for object detection with linear classifiers, which strongly suggests that HOG is also well suited for handwritten character recognition under linear models.

Devido à natureza do problema, vemo-nos obrigados a criar um preditor capaz de distinguir entre 26 letras romanas escritas à mão a partir de um vetor de intensidades correspondentes aos píxeis da imagem. Embora seja possível treinar um modelo logístico diretamente sobre o vetor achatado de $28 \times 28 = 784$ píxeis, esta representação é altamente redundante e extremamente sensível a ruído, variações de iluminação e pequenos deslocamentos da escrita.

Ao utilizar uma técnica de extração de características como o HOG, conseguimos reduzir a dimensionalidade de cada amostra de 784 características para apenas 144 descritores, ao mesmo tempo que fornecemos ao modelo uma representação muito mais informativa. Em vez de depender de intensidades absolutas dos píxeis que são facilmente afetadas por ruído e perturbações locais o HOG capta bordas, orientações e traços da letra, que se mantêm relativamente invariantes perante pequenas transformações geométricas. Tal robustez torna o modelo mais resistente às variações naturais da escrita manual.

Este método revelou-se eficaz: o tempo de treino diminuiu para menos de metade devido à redução da dimensionalidade, e a capacidade de previsão aumentou em cerca de 8% em média quando comparado com o modelo logístico básico baseado apenas em píxeis. Deste modo, o HOG apresenta-se como uma representação de características simultaneamente mais eficiente e mais discriminativa para este tipo de tarefa.