# Classificação de dígitos manuscritos utilizando HOG e KNN

#### Gabriel de Oliveira Pontarolo

# 1 Introdução

O relatório exibe os resultados do uso de um extrator de características HOG (histogram of oriented gradients) em um classificador KNN para dígitos escritos a mão. Foram feitas comparações variando a divisão de base, tamanho de K e utilizando diferentes métricas.

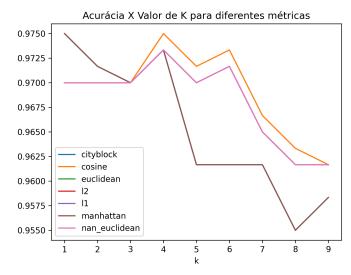
## 2 Extrator de características

Para a extração de características, foi utilizado o descritor HOG, uma técnica faz o uso da orientação dos gradientes em regiões específicas da imagem, focando na estrutura ou forma do objeto [1].

Foi utilizada a implementação do algoritmo disponibilizada dentro da biblioteca  $OpenCV^{[2]}$ , com os valores de parâmetros configurados inicialmente para os mesmos utilizados no *paper* original, por Dalal et al.<sup>[3]</sup>, e ajustados por meio de testes para os valores de tamanho de janela e imagem de 24x24, tamanho de célula de 6x6, tamanho do bloco para 12x12 (2x tamanho da célula) e número de *bins* para 9.

## 3 Testes

Os testes foram feitos comparando a acurácia do KNN variando os valores de K entre 1 e 9, e entre as métricas de distância *cityblock*, similaridade de cossenos, distância euclidiana, norma 12, norma 11, *manhattan* e distância não-euclidiana. A divisão de base nesse passo foi fixada em 70% para treinamento e 30% para teste. Foi obtido o seguinte grafico para o valor da acurácia:



É importante ressaltar que algumas curvas são extremamente similares a outras, como 12 e euclidiana, e por isso não ficaram visíveis no gráfico. Assim, a maior acurácia obtida foi de **0.975** para **K** = **4** e **similaridade de cossenos** com a seguinte matriz de confusão:

54	0	0	0	0	1	0	0	0	0
0	50	0	0	0	0	0	0	0	0
0	0	64	0	0	0	0	1	1	0
0	0	0	66	0	0	0	0	2	0
0	0	0	0	58	0	0	0	0	2
0	0	0	2	0	54	0	0	0	0
0	0	0	0	0	0	65	0	1	0
0	0	0	0	1	0	0	55	0	0
1	0	0	0	1	0	0	0	50	0
0	0	0	0	0	0	0	2	0	69

#### E *f-scores* por classe de:

0	0.98
1	1.00
2	0.98
3	0.97
4	0.97
5	0.97
6	0.99
7	0.96
8	0.94
9	0.97

Por fim, a partir dos parâmetros para o valor de K (4) e métrica (cossenos) encontrados, foram feitas cinco execuções variando a divisão da base. A tabela abaixo apresenta os *f-scores* e a acurácia de cada execução, com a divisão no formato teste:treinamento, e a média para cada classe:

	1:9	2:8	3:7	4:6	5:5	Avg
0	0.98	0.99	0.98	0.99	0.98	0.984
1	0.97	0.98	1.00	0.99	0.98	0.984
2	1.00	0.99	0.98	0.98	0.98	0.986
3	0.96	0.99	0.97	0.96	0.96	0.968
4	0.95	0.97	0.97	0.95	0.96	0.96
5	0.96	0.98	0.97	0.97	0.98	0.972
6	1.00	1.00	0.99	0.99	0.99	0.994
7	0.97	0.95	0.96	0.97	0.96	0.962
8	0.97	0.97	0.94	0.95	0.94	0.954
9	0.98	0.97	0.97	0.97	0.97	0.972
Acc	0.975	0.98	0.975	0.973	0.97	0.974

Onde a maior acurácia de **0.98** foi observada com uma divisão de 20% para teste e 80% para treinamento, e a média de **0.974**.

# 5 Referências

[1] Histogram of Oriented Gradients. Wikipedia. https://en.wikipedia.org/wiki/Histogram\_of\_orient ed gradients

[2] OpenCV: cv::HOGDescriptor Struct Reference. OpenCV. https://docs.opencv.org/4.x/d5/d33/structcv\_1\_1H OGDescriptor.html

[3] Navneet Dalal, Bill Triggs. Histograms of Oriented Gradients for Human Detection. International Conference on Computer Vision & Pattern Recognition (CVPR '05), Jun 2005, San Diego, United States. pp.886–893, 10.1109/CVPR.2005.177. inria-00548512