

EEE928 - Técnicas Clássicas de Reconhecimento de Padrões

Exercício Template Matching

José Geraldo Fernandes
Escola de Engenharia
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
josegeraldof@ufmg.br

I. ATIVIDADE 1

Implementou-se um método de *template matching* para reconhecimento de padrões em imagens. A entrada alvo é tratada como *template* e faz-se uma varredura na entrada ambiente, avaliando uma métrica de desempenho para cada janela, como na Equação 1, com a função δ sendo uma métrica qualquer.

$$\text{custo} = \delta(\text{janela}, \text{template}) \quad (1)$$

Utilizou-se a Figura 1 como entrada ambiente e as Figuras 2 e 3 como entradas alvo para reconhecimento.



Figura 1. Figura de ambiente para reconhecimento das placas.



Figura 2. Placa de entrada alvo 1.



Figura 3. Placa de entrada alvo 2.

Nessa etapa, utilizou-se como métrica simplesmente as normas L1 e L2 das imagens, *template* e janela, em escala cinza. Também registrou-se o custo pela superfície do ambiente. O resultado é como nas Figuras 4, 5, 6 e 7, a região encontrada é traçada em branco na imagem ambiente.

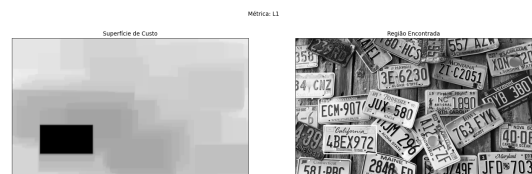


Figura 4. Resultado do processo de *template matching* para o alvo 1 com norma L1.

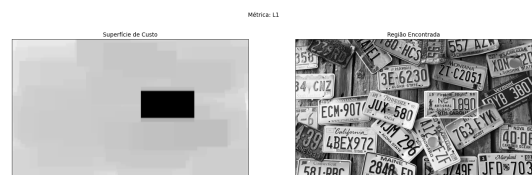


Figura 5. Resultado do processo de *template matching* para o alvo 2 com norma L1.

Note que para todos os casos o resultado é exato, isso é esperado uma vez que é utilizado o próprio recorte do

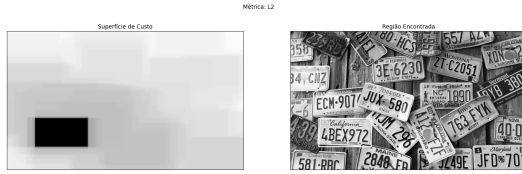


Figura 6. Resultado do processo de *template matching* para o alvo 1 com norma L2.



Figura 7. Resultado do processo de *template matching* para o alvo 2 com norma L2.

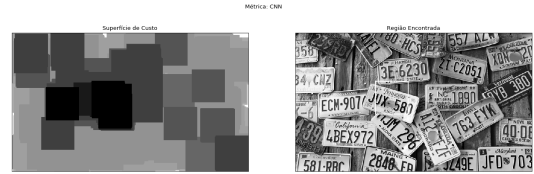


Figura 8. Resultado do processo de *template matching* para a primeira entrada, 'JUX580', com uma CNN.

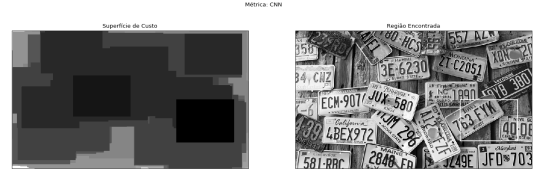


Figura 9. Resultado do processo de *template matching* para a primeira entrada, '763FYK', com uma CNN.

ambiente como *template*. Dessa forma, basta comparar de forma direta os valores dos *pixels* na varredura.

II. ATIVIDADE 2

Em seguida, o problema é alterado de forma significativa. A entrada passa a ser em formato de texto, isso significa que não há um *template* direto para comparação na varredura. Também, e por consequência do anterior, a função de métrica deve considerar uma carga semântica, qual imagem melhor representa o texto, além de resistir a complicações como variação angular e de amplitude.

Esses obstáculos podem ser tratados com a utilização de uma rede neural convolucional (CNN). Assim, aplicou-se a seguinte abordagem: uma CNN busca caracteres na janela de varredura e entrega um texto, reconhecimento ótico de caracteres (OCR); essa saída é comparada com a entrada de texto alvo, a função de métrica. Modifica-se o custo, portanto, como na Equação 2. Note que nessa abordagem as métricas utilizadas anteriormente não são aplicáveis.

$$\text{custo} = \delta(\text{OCR}(\text{janela}), \text{texto}) \quad (2)$$

Utilizou-se uma CNN treinada em reconhecimento de placas de carro. O tamanho da janela de varredura é uma decisão de projeto, testou-se um formato quadrado e retangular, o melhor resultado foi preservado.

As entradas testadas foram: 'JUX580' e '763FYK'. Para comparação dos textos, utilizou-se uma função de similaridade de *strings* nativa de Python, SequenceMatcher do módulo difflib. Essa função foi desenhada para valorizar sequências de caracteres concordantes, dessa forma penalizando erros de leitura dentro de subconjuntos.

O resultado é como nas Figuras 8 e 9, novamente, a região encontrada é traçada em branco na imagem ambiente.

Note que, diferente do resultado da Seção I, a superfície de custo não é lisa e uniforme. Em verdade, longe disso,

em ambos exemplos há regiões concorrentes, de baixo custo, apesar de o modelo acertar placa correta.

Há alguns obstáculos que podem explicar esse resultado. O gargalo maior está na acurácia da extração de caracteres da CNN. Mesmo que caracteres parecidos tenham um custo baixo na rede neural, são penalizados uniformemente pela métrica de comparação. É possível que uma métrica mais inteligente signifique um grande aumento de performance.

Outro ponto importante é o tamanho da janela. Mesmo que as CNNs sejam resistentes a variação de amplitude e rotação, essas distorções implicam um aumento no tamanho da janela de varredura. Esse fenômeno é prejudicial já que a densidade de caracteres na imagem ambiente é muito grande, expondo caracteres invasores, situação que provavelmente não foi vista no treinamento pela sua atipicidade.