

# Tarefa 01: Detecção de Placa de Veículos

Alexandre Xavier Falcão

Entrega parcial, 02/10/18, e final, 16/10/18

## Resumo

Em sala de aula vimos as operações de normalização divisiva, convolução, ativação, *max-pooling*, e *min-pooling*. Vimos que uma sequência dessas operações usando um filtro (*kernel*) adequado na convolução é capaz de realçar pixels de uma placa de veículo em uma dada imagem. Esta tarefa, portanto, consiste em selecionar um banco de filtros adequados ao problema; projetar um classificador por combinação ponderada de imagens de treinamento filtradas por esses *kernels*, seguida de limiarização e pós-processamento para segmentação de um componente de placa; e depois avaliar a efetividade da detecção de uma região de interesse em torno do componente de placa selecionado em um conjunto de imagens de teste. A efetividade da operação pode ser medida pelo erro médio e o desvio padrão do número de pixels de placa fora da região de interesse somado ao número de pixels de fundo dentro da região de interesse. Menor o erro, maior a efetividade.

## 1 Introdução

Para realizar esta tarefa, um banco com 200 imagens cinza contendo placas de veículos e as respectivas máscaras de cada placa foi selecionado e essas imagens foram divididas em 5 pares de conjuntos de treinamento e teste, cada um contendo 50% das imagens do banco, selecionadas de forma aleatória e com intersecção vazia entre o treinamento e o teste em cada par. Sua avaliação, portanto, deve apresentar a efetividade média e o desvio padrão sobre esses conjuntos de dados (i.e., uma validação cruzada  $5 \times 2$ -fold).

Sejam  $Z_{tr}$  e  $Z_{ts}$  um par, conjunto de treinamento e conjunto de teste, cada um contendo 100 imagens  $\hat{I}_i = (D_I, I_i)$ ,  $i = 1, 2, \dots, 100$ , e as respectivas máscaras  $\hat{M}_i = (D_I, M_i)$  da placa, com  $M_i(p) = 255$  para pixels de placa e  $M_i(p) = 0$  para pixels de fundo. Inicialmente, selecione um certo número  $k$  de *kernels*  $\hat{K}_j$ ,  $j = 1, 2, \dots, k$ , que lhe pareçam adequados ao problema e ajuste os parâmetros da sequência de operações vista em sala para algumas imagens de treinamento. Muito embora esses parâmetros possam ser aprendidos, vamos selecioná-los de forma empírica e mantê-los fixos por simplicidade. Usem primeiro a implementação da sequência de operações, como visto em sala, e depois otimizem o processamento com o uso de multiplicação matricial. Para cada

imagem cinza de treinamento  $\hat{I}_i = (D_I, I_i) \in Z_{tr}$ , você deve aplicar a sequência de operações dada gerando uma imagem  $\hat{J}_{i,j} = (D_I, J_{i,j})$ ,  $j = 1, 2, \dots, k$ , (bandas da imagem  $\hat{J}_i = (D_I, J_i)$ ) resultante da sequência de operações aplicada à imagem  $\hat{I}_i$  usando cada filtro  $\hat{K}_j$ . As bandas  $\hat{J}_{i,j}$ ,  $j = 1, 2, \dots, k$ , de  $\hat{J}_i$  podem ser normalizadas no intervalo  $[0, 255]$ .

O classificador consistirá de uma limiarização  $T \in (0, 255)$  aplicada a uma imagem  $\hat{C}_i = (D_I, C_i)$  definida por

$$\hat{C}_i = \sum_{j=1}^k w_j \hat{J}_{i,j}, \quad (1)$$

onde  $0 \leq w_j \leq 1$  é um peso proporcional a efetividade do filtro  $\hat{K}_j$ . A limiarização gera uma imagem binária  $\hat{B}_i = (D_I, B_i)$ , onde  $\forall p \in D_I$ ,  $B_i(p) = 1$  (classificado como placa) se  $C_i(p) \geq T$ , e  $B_i(p) = 0$ , no caso contrário. O componente correspondente à placa será então selecionado por algum pós-processamento que iremos identificar com os resultados parciais a serem entregues no dia 02/10/18. A região de interesse em torno da placa poderá então ser definida como uma região retangular em torno do centro do componente selecionado.

## 2 Treinamento

O treinamento consiste em encontrar os parâmetros  $w_j$ ,  $j = 1, 2, \dots, k$ , e o limiar  $T$  que minimizam o erro médio de classificação de pixels de placa como fundo e de pixels de fundo como placa. Para cada filtro  $\hat{K}_j$ , você deve encontrar inicialmente o limiar  $T_j$  que minimiza o erro médio de classificação de pixels de placa como fundo e de pixels de fundo como placa quando as imagens  $\hat{J}_{i,j}$ ,  $i = 1, 2, \dots, 100$ , são limiarizadas por  $T_j$ . Isto é, seja  $\hat{B}_{i,j} = (D_I, B_{i,j})$  uma imagem binária definida por  $B_{i,j}(p) = 1$ , se  $J_{i,j}(p) \geq T_j$ , e  $B_{i,j}(p) = 0$ , no caso contrário. O erro  $e_{i,j}(T_j)$  será dado por

$$e_{i,j}(T_j) = \alpha N_{i,j}^{(0)} + \beta N_{i,j}^{(1)}, \quad (2)$$

onde  $N_{i,j}^{(0)}$  é o número de pixels de fundo classificados como placa (i.e.,  $B_{i,j}(p) = 1$  quando  $p$  tem valor  $M_i(p) = 0$  em  $\hat{M}_i$ ),  $N_{i,j}^{(1)}$  é o número de pixels de placa classificados como fundo (i.e.,  $B_{i,j}(p) = 0$  quando  $p$  tem valor  $M_i(p) = 255$  em  $\hat{M}_i$ ), e  $\alpha \ll \beta$  são constantes que compensam o fato da placa ser bem menor que o fundo. Isto é, o erro  $N_{i,j}^{(1)}$  precisa ter peso bem maior do que o erro  $N_{i,j}^{(0)}$  (e.g.,  $\beta = 10\alpha$ ). O valor ótimo de  $T_j \in (0, 255)$  deve ser aquele que minimiza o erro médio  $e_j$ ,

$$e_j = \frac{1}{100} \sum_{i=1}^{100} e_{i,j}(T_j). \quad (3)$$

Após encontrar  $T_j$  ótimo para cada filtro  $\hat{K}_j$ ,  $j = 1, 2, \dots, k$ , definimos os pesos  $w_j$  da seguinte forma.

$$w_j = 1 - \frac{e_j}{\sum_{l=1}^k e_l} \quad (4)$$

Após encontrar os valores ótimos de  $w_j$ ,  $j = 1, 2, \dots, k$ , o valor ótimo do limiar  $T$  poderá ser encontrado da seguinte forma. Calcula-se  $\hat{C}_i$  usando a Equação 1 e os pesos  $w_j$  encontrados. Para cada valor de  $T \in (0, 255)$ , você deve calcular o erro  $e_i(T)$  da limiarização de  $\hat{C}_i$  por  $T$ . Isto é,

$$e_i(T) = \alpha N_i^{(0)} + \beta N_i^{(1)}, \quad (5)$$

onde  $N_i^{(0)}$  é o número de pixels de fundo classificados como placa (i.e.,  $B_i(p) = 1$  quando  $p$  tem valor  $M_i(p) = 0$  em  $\hat{M}_i$ ),  $N_i^{(1)}$  é o número de pixels de placa classificados como fundo (i.e.,  $B_i(p) = 0$  quando  $p$  tem valor  $M_i(p) = 255$  em  $\hat{M}_i$ ), e  $\alpha \ll \beta$  são as mesmas constantes acima. Depois, o valor ótimo de  $T$  é aquele que minimiza o erro médio  $e$ ,

$$e = \frac{1}{100} \sum_{i=1}^{100} e_i(T). \quad (6)$$

Os parâmetros aprendidos,  $w_j$ ,  $j = 1, 2, \dots, k$ , e  $T$  devem então ser gravados em um arquivo texto para serem usados pelo teste.

### 3 Teste

Para cada imagem  $\hat{I}_i \in Z_{ts}$  de teste,  $i = 1, 2, \dots, 100$ , você deve executar a sequência de operações acima, seguida da combinação ponderada das imagens das bandas (Equação 1), limiarização  $T$ , e pós-processamento (a ser definido) sobre a imagem resultante  $\hat{B}_i = (D_I, B_i)$ . O erro médio no conjunto de teste deve ser então calculado pela Equação 6 e o mesmo pode ser feito trocando-se  $Z_{tr}$  por  $Z_{ts}$ , e repetindo o procedimento para cada um dos 5 pares de conjuntos  $Z_{tr}$  e  $Z_{ts}$ . Ao final, calcula-se o erro médio final dessas 10 execuções do procedimento e o desvio padrão desta medida.

### 4 Apresentação dos resultados

Os resultados devem ser apresentados em um relatório escrito conforme o seguinte padrão. O relatório pode ter até 20 páginas de texto 11pt, incluindo figuras, tabelas, gráficos, e referências bibliográficas se necessário. Ele deve conter:

- Título da tarefa, código da disciplina, nome do aluno, RA, e data de entrega do relatório.

- Introdução com pesquisa bibliográfica sobre as técnicas estudadas para implementação da tarefa, se for o caso, e resumo do que foi implementado e dos resultados obtidos.
- Descrição teórica das técnicas, discussão sobre as dificuldades práticas encontradas, e explicação sobre como elas foram resolvidas na tarefa.
- Explicação dos experimentos e a apresentação dos resultados.
- Conclusão com uma discussão dos resultados principais.

Este relatório é devido para 16/10/18, mas uma versão parcial deverá ser apresentada para a aula do dia 02/10/18. Nesta aula iremos discutir o andamento da tarefa e decidir o pós-processamento que será adotado para a conclusão da tarefa e do relatório.