

Algoritmo de Viola-Jones

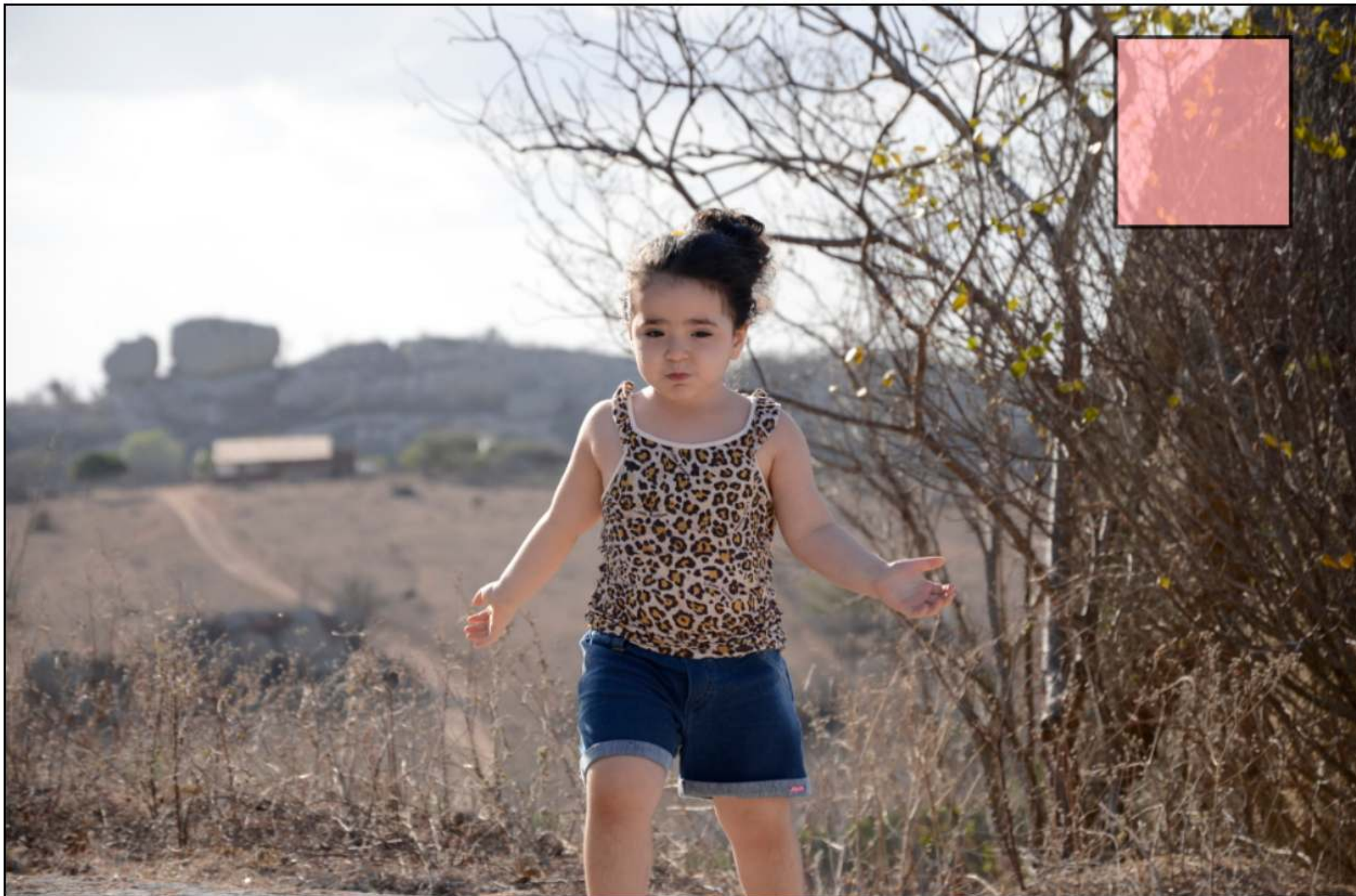
Agostinho Brito

2022

Algoritmo de Viola-Jones

- Proposto em 2001 por Paul Viola e Michael Jones
- Amplamente usado para detecção em tempo real de objetos em imagens cinzentas.
- Treinamento lento, mas de rápida detecção.
- Operando em um **Pentium III de 700MHz**, o algoritmo foi capaz de processar até 15 fps em imagens de 384 por 288 pixels.

Como achar a face?



Como achar a face?

- Em cada sub-janela o algoritmo precisa decidir se há uma face lá dentro.
- Redes neurais podem ajudar.... :-/
- Problemas:
 - A solução precisa operar em tempo real.
 - Iluminação.
 - Indivíduos diferentes
 - Posicionamento da face
 - Faces são raras por imagens
- Proposta: usar combinações de classificadores fracos (e simples) para realizar a detecção da face.

- Tipos de características



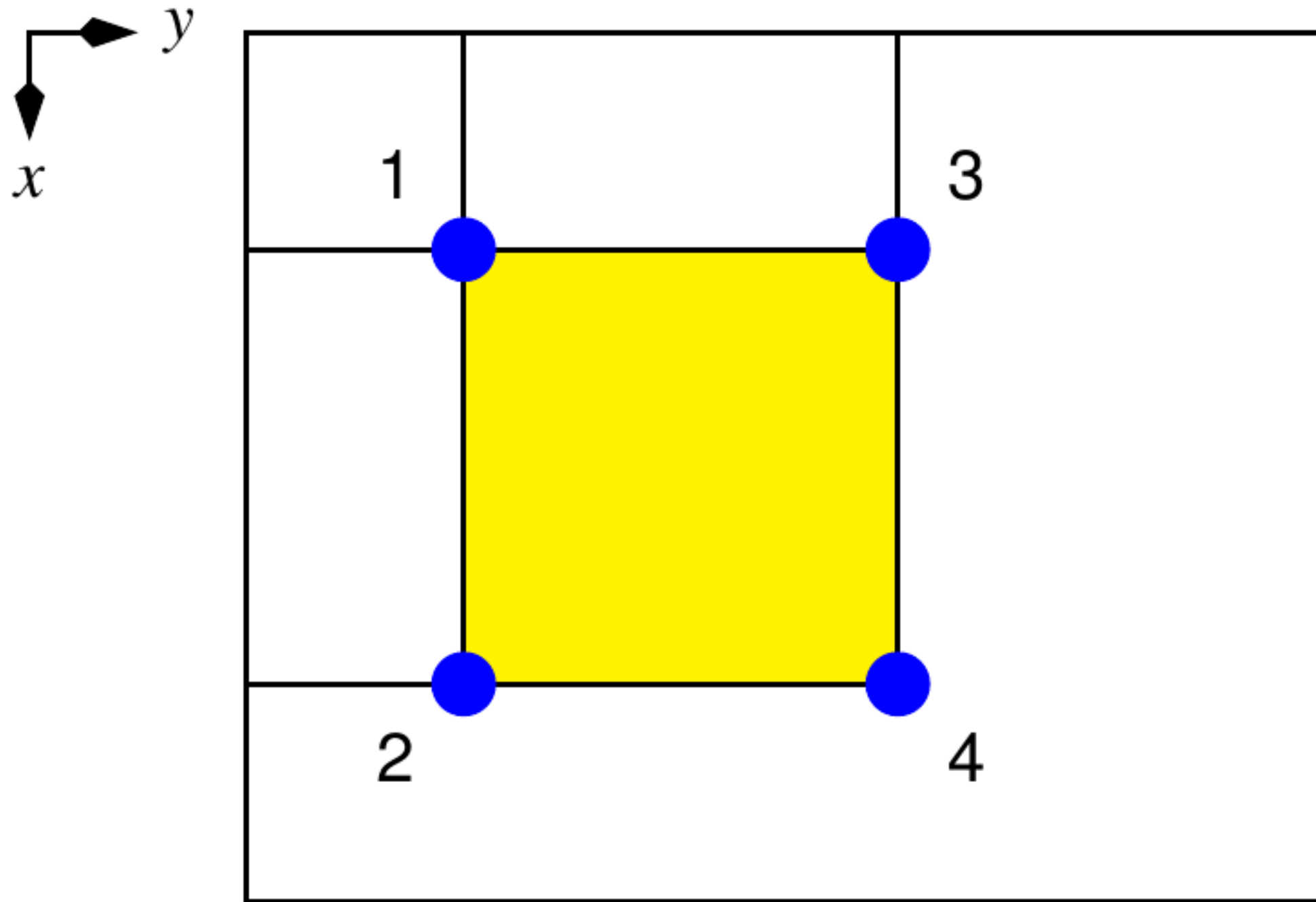
- Medida da característica (*feature*): soma dos pixels nos quadrados pretos menos a soma dos pixels nos quadrados brancos.

$$\text{característica} = \sum \blacksquare - \sum \square$$

- Problema: muitas combinações para um detector com resolução de 24×24 pixels.
- Solução: encontrar as melhores características. Realizar cálculos rápidos. Uso de Adaboost e imagem integral.

Imagem integral

- A imagem integral de uma imagem é construída com soma de todos os pixels acima e à esquerda de um dado ponto.

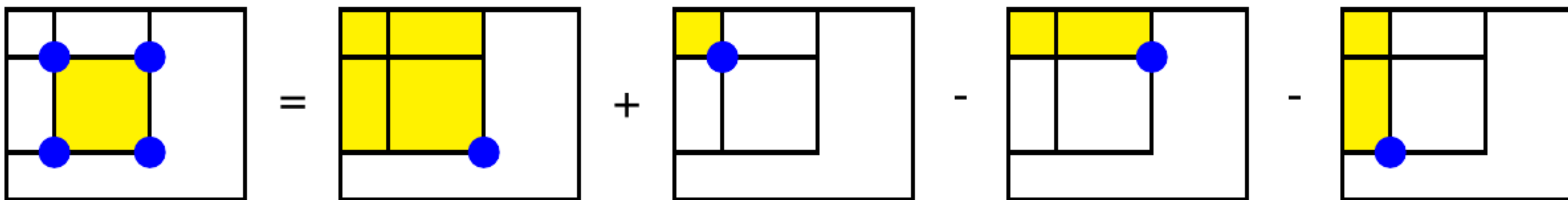


$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

$$s(x, y) = s(x, y - 1) + i(x, y)$$
$$ii(x, y) = ii(x - 1, y) + s(x, y)$$

Pode ser calculada com uma única passada na imagem.

$$soma(\blacksquare) = \sum (ii_4 + ii_1 - ii_2 - ii_3)$$

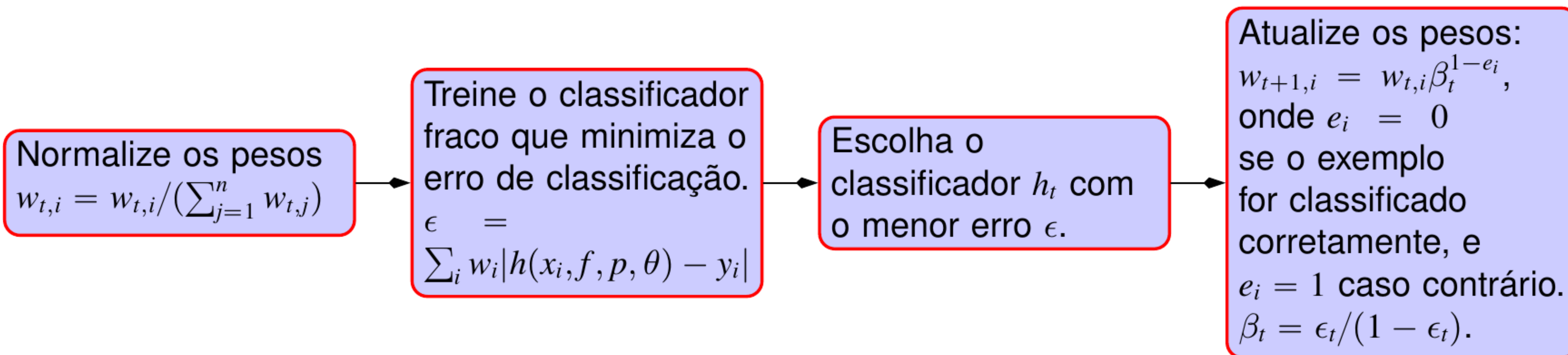


- Um classificador fraco é dado por

$$h_j(x) = \begin{cases} 1 & \text{se } p_j f_j(x) < p_j \theta_j \\ 0 & \text{caso contrário} \end{cases}$$

- Na equação, x é uma janela de 24×24 pixels.
- $f_j(x)$ é a característica (feature) calculada para o retângulo.
- θ é um threshold. p_j é uma polarização (+1 ou -1).
- Na escolha dos classificadores, a característica mais apropriada é selecionada.
- No treinamento, todos os filtros possíveis são testados em múltiplas rodadas.
- Para uma janela de 24×24 pixels, há um total de 162.336 características.

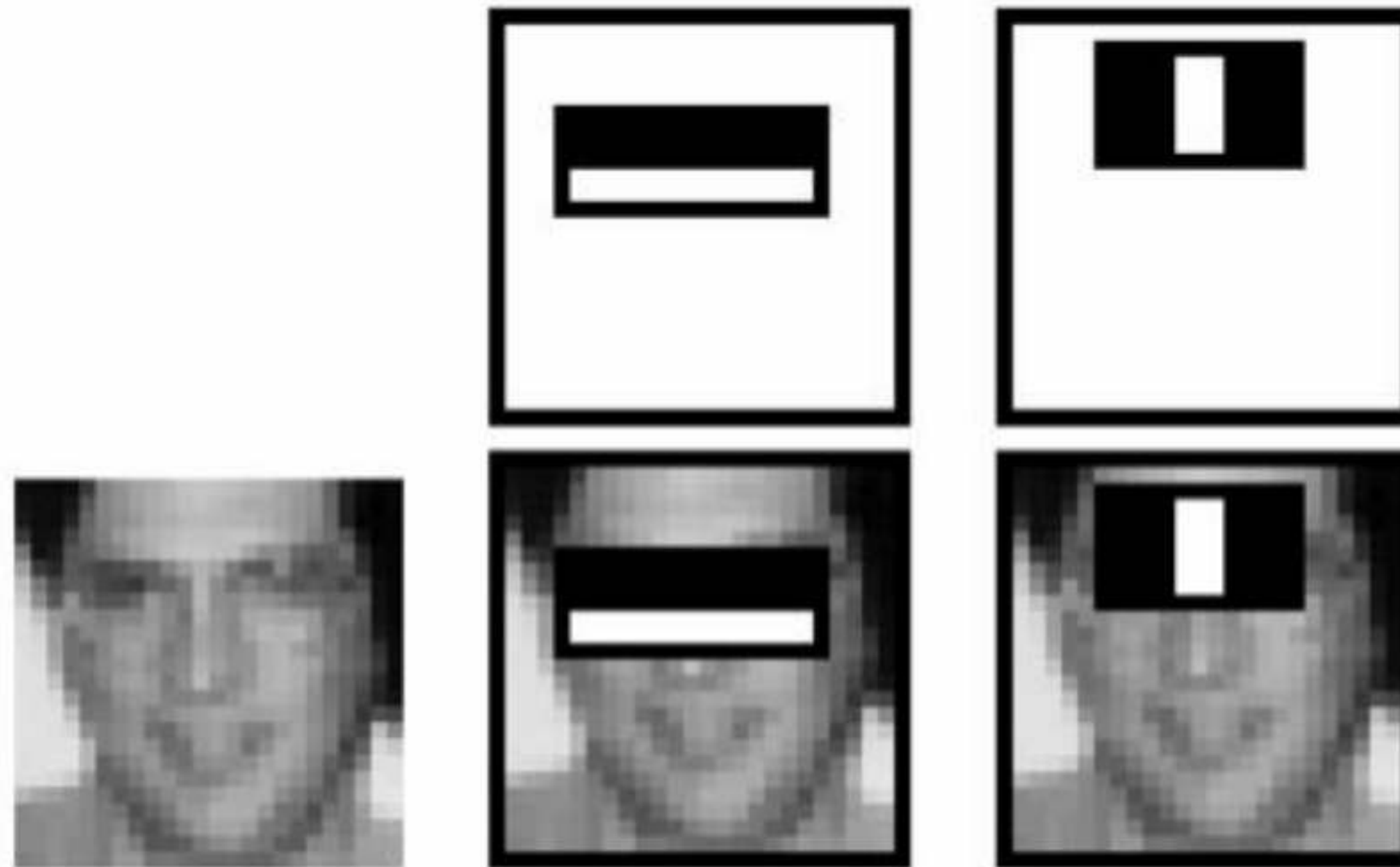
- As N imagens de treinamento são rotuladas $(x_1, y_1), \dots, (x_n, y_n)$ onde $y_i = 0, 1$ para exemplos negativos ou positivos.
- Inicialize os pesos $w_{1,i} = 1/(2m), 1/(2l)$, para $y_i = 0, 1$, onde m e l são os números de exemplos negativos e positivos.
- T é a quantidade de passos de ajustes.
- Para $t = 1, \dots, T$:



O classificador forte final é dado por

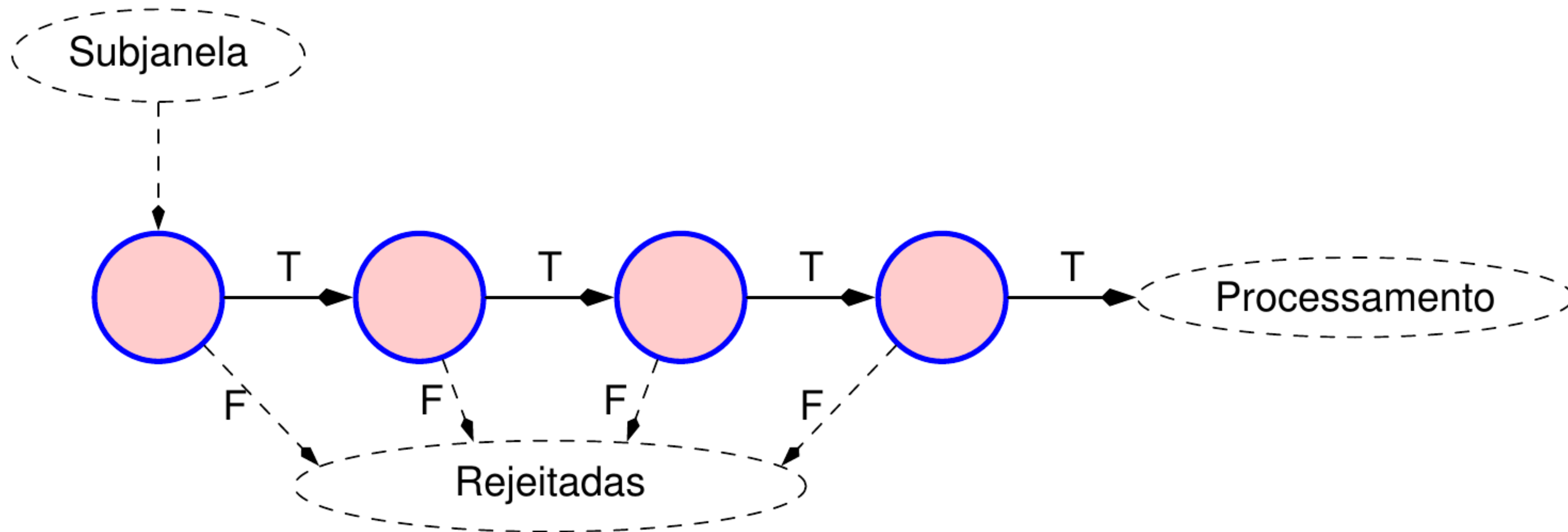
$$h(x) = \begin{cases} 1 & \text{se } \sum_{t=1}^T \alpha_t h_t(x) \geq 1/2 \sum_{t=1}^T \alpha_t, \text{ onde } \alpha_t = \log(1/\beta_t) \\ 0 & \text{caso contrário} \end{cases}$$

Exemplo de características selecionadas



Melhorando o algoritmo

- Eliminar a maior quantidade de características incapazes de realizar classificações corretas.
- Organizar os classificadores em cascata.
- Cada estágio remove todas as classificações incorretas, repassando positivos e falsos-positivos para o estágio seguinte da cascata.
- O detector é treinado com N níveis de cascata.
- Cada nível reduz o número de falsos-positivos e reduz a taxa de detecção, inserindo novas características para serem testadas.



- Número de estágios
- Características
- Thresholds

- Todas as imagens são normalizadas para evitar problemas de iluminação.
- O detector desenvolvido por Viola e Jones tinha 38 níveis de cascata e 6000 características.
- Em um conjunto de dados de 507 faces, e 75 milhões de subjanelas, uma média de 10 avaliações de características é feita para cada subjanela.
- A redução da escala do detector é feita redimensionando o detector e não a própria imagem. Assim, as características podem ser calculadas em qualquer escala com o mesmo custo computacional.
- As cinco primeiras camadas possuíam 1, 10, 25, 25 e 50 características, respectivamente.
- Na biblioteca OpenCV, o classificador `haarcascade_frontalface_default.xml` possui **25** estágios e **2913** características ao todo. Os cinco primeiro estágios avaliam 9, 16, 27, 32 e 52 características, respectivamente.



Obrigado