

 E-mail de  
Universidade de São  
Paulo

Pedro Henrique Barbosa de Almeida <pedro.hba@usp.br>

## [IC] Patch extraction e classificadores

Nina S. T. Hirata <nina@ime.usp.br>

21 de maio de 2019 22:28

Responder a: nina@ime.usp.br

Para: Pedro Henrique Barbosa de Almeida <pedro.hba@usp.br>

Oi Pedro,

### 1.

Bem, basicamente, binarizo a imagem e passo os componente conexo cada um por vez. O objetivo é pegar uma vizinhança tal que, quando redimensionada para 41x41, o "retângulo" que contém o componente conexo na imagem de escala normal fique com dimensão 9x9.

Acho que está quase correto. A escala deve ser definida em função daquele que é maior entre altura e largura. Ou seja, em vez de considerar o retângulo que contém o componente conexo, a gente considera o menor quadrado que contém ele (por exemplo, se for um componente com  $h > w$ , significa que vai ficar uma margem na esquerda e na direita, enquanto que acima e abaixo o box fica justo).

Se considerarmos o retângulo e depois normalizar para 9x9, o conteúdo vai ficar deformado.

Vou apagar o skip :-), e ir para a próxima pergunta.

### 2.

Bem, com os *patches* em mãos, agora devo me voltar ao treinamento dos classificadores. Andei vendo o que que a Ana me mandou, mas algumas coisas estão um pouco confusas. A começar pelo mais básico: o que são, na verdade, aqueles .xml? São imagens? Rótulos? Além disso, o que ela chamou de imagens rotuladas são, na verdade, imagens escuras com alguns elementos um pouquinho mais claros. A intensidade desses pixels é que são os rótulos (0: 'Other', 1: 'Text', 2: 'Graphic', 3: 'Image', 4: 'Separator', 5: 'Table')?

É o seguinte: o ground truth original é dado na forma de polígonos envoltórios da região. Por exemplo, cada bloco de texto, cada imagem, cada tabela, na imagem é marcada por meio de polígonos e cada um dos polígonos tem uma classe (y) associada. O que a Ana faz é verificar a qual polígono um componente conexo está contido para definir a classe do componente onexo. O xml define esses polígonos e a classe deles.

Sim, imagens rotuladas são isso mesmo que você falou. Só não sei se as imagens que ela passou contém polígonos rotulados ou componentes conexos rotulados. Eu esperaria que fossem polígonos.

### Um ponto que eu não entendi:

Posso ter entendido errado o que você escreveu, mas entendi que você está fazendo o recorte ao redor dos componentes conexos?

O que eu tenho em mente é o seguinte:

- 1) aplicar aquela ideia de binarizar a imagem, calcular os componentes conexos, e aplicar o algoritmo de superpixel apenas nos componentes que são maiores que um dado tamanho T (por exemplo)
- 2) Assim, suas regiões serão os componentes conexos não maiores que T, mais os superpixels
- 3) os recortes deverão ser feitos ao redor dessas regiões
- 4) a classe correspondente a cada recorte pode ser definida como sendo o do polígono que cobre a maioria dos pontos da região (acho que pode acontecer de uma região ser coberta por mais de um polígono).
- 5) classificar uma dessas regiões corresponde a pintar todos os pixels da região com o rótulo devolvido pelo classificador

Sobre o modelo, para escolher o modelo seria necessário fazer um grid-search ou algo parecido variando vários dos hiperparâmetros ...  
Eu sugiro, por enquanto, testar algum modelo simples, do tipo

30/05/2019

E-mail de Universidade de São Paulo - [IC] Patch extraction e classificadores

Google LeNet-5 (não sei se o do cachorro e gato era um desses).

Se algum ponto não ficou claro, avise-me.

[],

Nina