

Iniciado em	sábado, 19 Fev 2022, 10:28
Estado	Finalizada
Concluída em	sábado, 19 Fev 2022, 16:28
Tempo empregado	6 horas
Avaliar	8,60 de um máximo de 10,00(86%)

Questão 1

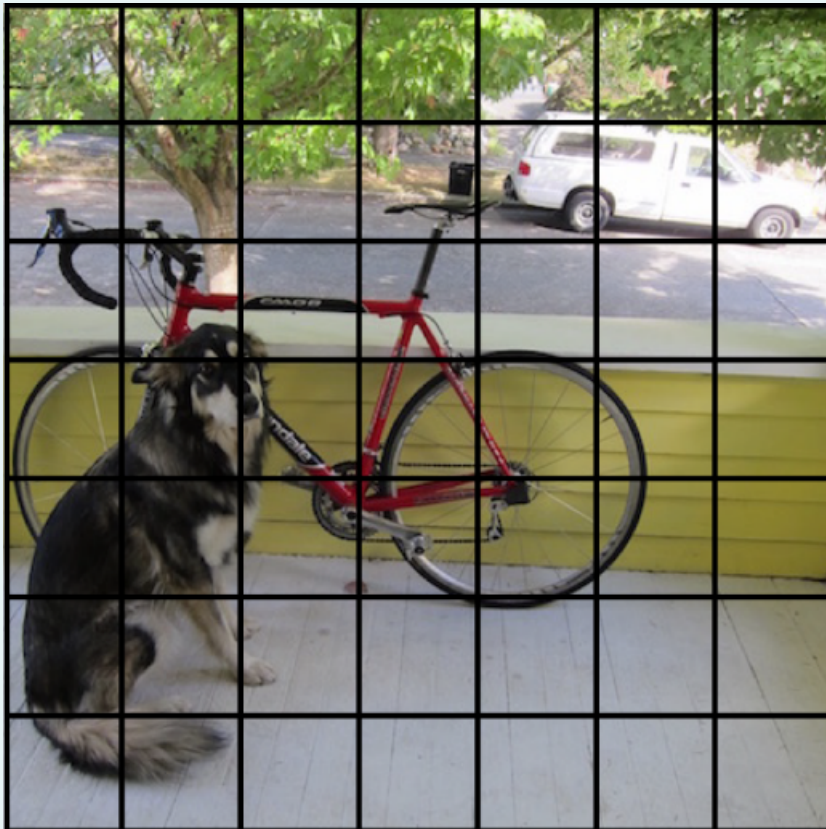
Correto

Atingiu 2,00 de 2,00

Considere o uso do algoritmo Yolo com a seguinte configuração para de detecção de objetos: um gride de 24 x 24, 27 classes e 3 *anchor boxes*.

Qual será a dimensão do volume de saída da rede?

Exemplo de imagem com gride 7x7:



Resposta: ✓

A resposta correta é: 55296.

Questão 2

Parcialmente correto

Atingiu 1,00 de 2,00

Assinale as sentenças verdadeiras. Opções incorretas reduzem a nota final.

- ☐ a. O método GeMOS (Vendramini et al., 2021) é bastante efetivo para segmentação semântica, sendo capaz de explorar qualquer rede convolucional pré-treinada (backbone).
- ☒ b. Apesar de muito simples, o método OpenPixel (Silva et al., 2020) é bastante efetivo para a identificação de pixels de classes desconhecidas e supera a abordagem Closed-Set no geral. ✖
- ☒ c. O método OpenPCS (Oliveira et. al., 2021) realiza segmentação semântica em cenário aberto sendo capaz de explorar qualquer backbone clássico. Para isso, o método extrai características nas camadas do *decoder*, realiza redução de dimensionalidade com o PCA e faz a modelagem gaussiana das amostras das classes conhecidas. Para reduzir o custo computacional da modelagem das distribuições, o método faz uma amostragem dos pixels, reduzindo o número de pontos no espaço de características. ✔
- ☒ d. O método CGDL (Sun et al., 2020) combina baseados em Inferência Variacional e Teoria dos Valores Extremos para para forçar a rede a aprender features que facilitem a modelagem das classes conhecidas. O treinamento da rede envolve uma função de perda composta por três partes: classificação, reconstrução da imagem e similaridade das distribuições (divergência KL). A arquitetura encoder-decoder proposta facilitaria sua adaptação para o problema de segmentação semântica. ✖
- ☐ e. O método C2AE (Oza e Patel, 2019) realiza segmentação semântica por meio de um auto encoder convolucional que reconstrói amostras positivas (*match*) e negativas (*unmatch*) que auxiliam na modelagem das distribuições das classes conhecidas (KKCs).

Sua resposta está parcialmente correta.

Você selecionou muitas opções.

A resposta correta é:

O método OpenPCS (Oliveira et. al., 2021) realiza segmentação semântica em cenário aberto sendo capaz de explorar qualquer backbone clássico. Para isso, o método extrai características nas camadas do *decoder*, realiza redução de dimensionalidade com o PCA e faz a modelagem gaussiana das amostras das classes conhecidas. Para reduzir o custo computacional da modelagem das distribuições, o método faz uma amostragem dos pixels, reduzindo o número de pontos no espaço de características..

Indique as sentenças verdadeiras abaixo. Respostas incorretas podem reduzir sua nota final.

- ☒ a. Um sistema de reconhecimento 'open world' precisa ser atualizado continuamente (aprendizado incremental), ser robusto a categorias não vistas inicialmente (*open set*) e ter um tempo de inatividade mínimo (escalabilidade). ✓
- ☒ b. O Open-Set Recognition (OSR) descreve um cenário em que amostras de novas classes, não vistas no treinamento (UUCs), aparecem durante o teste. Os modelos OSR devem classificar corretamente as amostras de classes conhecidas e identificar as entradas UUC. ✓
- ☐ c. Um dos principais desafios na realização do reconhecimento openset é que a falta de informações sobre amostras de novas classes cria a necessidade de superajustar os modelos à distribuição de classes conhecidas e isso reduz a generalização. O método proposto por Sun et al (2020) chamado *Conditional Gaussian Distribution Learning* (CGDL) reduz esse efeito forçando o aprendizado de distribuições Weibull em diferentes níveis de representação, gerando um espaço de características mais fácil de ser modelado.
- ☒ d. O problema de segmentação semântica de imagens de sensoriamento remoto é essencialmente aberto, pois a presença de novas classes é inevitável na maioria das aplicações. ✓
- ☐ e. Embora o problema de Open Set Recognition (OSR) seja semelhante a muitas outras tarefas de aprendizado de máquina, não há uma relação clara com os métodos propostos para tarefas de zero-shot learning.

Sua resposta está correta.

As respostas corretas são:

O Open-Set Recognition (OSR) descreve um cenário em que amostras de novas classes, não vistas no treinamento (UUCs), aparecem durante o teste. Os modelos OSR devem classificar corretamente as amostras de classes conhecidas e identificar as entradas UUC.,

Um sistema de reconhecimento 'open world' precisa ser atualizado continuamente (aprendizado incremental), ser robusto a categorias não vistas inicialmente (*open set*) e ter um tempo de inatividade mínimo (escalabilidade).,

O problema de segmentação semântica de imagens de sensoriamento remoto é essencialmente aberto, pois a presença de novas classes é inevitável na maioria das aplicações.

Questão 4

Parcialmente correto

Atingiu 1,60 de 2,00

Indique as sentenças verdadeiras abaixo. Respostas incorretas reduzem a nota final.

- ☒ a. Na detecção de objetos com janelas deslizantes: aumentar o *stride* melhora a eficiência computacional, mas diminui a precisão. Duas alternativas possíveis são: implementação convolucional de janela deslizante (estratégia da YOLO) ou não classificar todas as janelas (propostas de regiões). ✓
- ☐ b. O R-CNN original usa o algoritmo de segmentação SLIC para propor regiões de interesse, ou seja, regiões que possivelmente contêm um objeto. Cada uma dessas regiões é achatada (*warped*) e passada por uma rede convolucional para obter um mapa de características correspondente (um processo muito caro).
- ☒ c. O Fast R-CNN passa a imagem apenas uma vez pela ConvNet, obtendo um mapa de características. As regiões propostas por meio de segmentação são então "cortadas" no mapa de características e, em seguida, reduzidas em tamanho através da operação RoI Pooling. Passar a imagem uma vez pela ConvNet reduz substancialmente o tempo de detecção de objetos. ✓
- ☒ d. O RoI Pooling transforma caixas delimitadoras propostas (regiões) de tamanhos diferentes em uma lista de matrizes do mesmo tamanho. Isso possibilita o uso de redes convolucionais para classificar regiões. ✓
- ☐ e. No ROI Pooling, os limites de RoI são quantizados (arredondados) para corresponder ao mapa de recursos (snap to grid). Isso pode levar a problemas de alinhamento que causam erros na classificação por pixel. De qualquer forma, os erros de alinhamento não são um problema na detecção de objetos, pois basta fornecer uma caixa delimitadora correta.
- ☒ f. O gargalo do Fast R-CNN é a proposta de regiões por meio de busca seletiva. O Faster R-CNN utiliza uma rede convolucional para propor regiões (RPN), que pode ser vista como uma rede cujo papel é determinar se existe um objeto centrado em um determinado pixel. Isso torna o Faster R-CNN muito mais eficiente. ✓

Sua resposta está parcialmente correta.

Você selecionou corretamente 4.

As respostas corretas são:

Na detecção de objetos com janelas deslizantes: aumentar o *stride* melhora a eficiência computacional, mas diminui a precisão. Duas alternativas possíveis são: implementação convolucional de janela deslizante (estratégia da YOLO) ou não classificar todas as janelas (propostas de regiões).,

O gargalo do Fast R-CNN é a proposta de regiões por meio de busca seletiva. O Faster R-CNN utiliza uma rede convolucional para propor regiões (RPN), que pode ser vista como uma rede cujo papel é determinar se existe um objeto centrado em um determinado pixel. Isso torna o Faster R-CNN muito mais eficiente.,

O RoI Pooling transforma caixas delimitadoras propostas (regiões) de tamanhos diferentes em uma lista de matrizes do mesmo tamanho. Isso possibilita o uso de redes convolucionais para classificar regiões.,

No ROI Pooling, os limites de RoI são quantizados (arredondados) para corresponder ao mapa de recursos (snap to grid). Isso pode levar a problemas de alinhamento que causam erros na classificação por pixel. De qualquer forma, os erros de alinhamento não são um problema na detecção de objetos, pois basta fornecer uma caixa delimitadora correta.,

O Fast R-CNN passa a imagem apenas uma vez pela ConvNet, obtendo um mapa de características. As regiões propostas por meio de segmentação são então "cortadas" no mapa de características e, em seguida, reduzidas em tamanho através da operação RoI Pooling. Passar a imagem uma vez pela ConvNet reduz substancialmente o tempo de detecção de objetos.

Descreva de forma objetiva: quais são as principais limitações e contribuições das Fully Convolutional Networks (CVPR, 2015)?

Limitações:

- Um problema desta implementação é que, ao propagar através de várias camadas convolucionais e de agrupamento alternadas, a resolução dos feature maps de saída é reduzida. Portanto, as previsões diretas de FCNs são tipicamente em baixa resolução, resultando em limites de objetos relativamente difusos;
- Precisamos aprender muitos parâmetros;
- Rodar essas redes em inputs de larga escala (como imagens de alta resolução ou vídeos) pode ser exaustivo em termos de consumo de memória da GPU;
- A classificação de cada pixel não considera totalmente a relação entre os pixels. Ignora a etapa de regularização espacial usada nos métodos usuais de segmentação baseada em classificação de pixels e não tem consistência espacial;

Contribuições:

- São aplicáveis a inputs de qualquer tamanho;
- Popularizou o uso de redes convolucionais end-to-end para segmentação semântica;
- Adaptou redes pré-treinadas com o ImageNet para segmentação semântica;
- Utilização de Up-sampling usando interpolação;
- Introduziu “skip connections” para aumentar a granularidade do up-sampling;

Comentário:

◀ [OpenSetSemanticSegmentation.ipynb](#)

Seguir para...

