Colon Polyps Detection based on Deep Learning

陈承勃

School of Mathematics
Sun Yat-sen University

May 12, 2018

Outline

1. Task Description
2. Model Selection
   - Existing State-of-the-art Detection Systems
   - Advantages of YOLO architecture
3. Colon Polyps Detection System
4. Experiments
   - Training Settings
   - Testing
   - Evaluation of Classification
5. Conclusions

Real-time Colon Polyps Detection Task

- Accurate classification
- Precise localization
- Real-time performance

Dataset

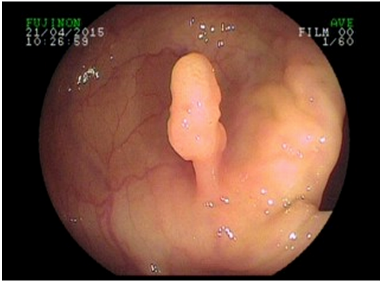


Figure 1: An example of colon endoscopy(结肠内镜）images. The whole dataset consists of 1,164 images of various sizes from different patients. 324 among them are discriminated to have at least one colon polyps(结肠息肉). The dataset comes from Doctor Zhong Dong.


```
- <annotation>
    <folder>FFOutput</folder>
    <filename>100001~19.png</filename>
    <path>E:\FFOutput\100001~19.png</path>
  - <source>
      <database>Unknown</database>
    </source>
  - <size>
      <width>568</width>
      <height>484</height>
      <depth>3</depth>
    </size>
    <segmented>0</segmented>
  - <object>
      <name>Colon polyps</name>
      <pose>Unspecified</pose>
      <truncated>0</truncated>
      <difficult>0</difficult>
    - <bndbox>
        <xmin>291</xmin>
        <ymin>9</ymin>
        <xmax>432</xmax>
        <ymax>203</ymax>
      </bndbox>
    </object>
  </annotation>
```


Figure 2: An example of labels.

Outline

1. Task Description
2. Model Selection
   - Existing State-of-the-art Detection Systems
   - Advantages of YOLO architecture
3. Colon Polyps Detection System
4. Experiments
   - Training Settings
   - Testing
   - Evaluation of Classification
5. Conclusions

- DPM (Deformable Parts Model)
  - Sliding window approach
  - Disjoint detection pipeline

- DPM (Deformable Parts Model)
  - Sliding window approach
  - Disjoint detection pipeline
- R-CNN series (R-CNN, Fast R-CNN, Faster R-CNN)
  - Region proposals approach
  - Disjoint detection pipeline

- DPM (Deformable Parts Model)
  - Sliding window approach
  - Disjoint detection pipeline
- R-CNN series (R-CNN, Fast R-CNN, Faster R-CNN)
  - Region proposals approach
  - Disjoint detection pipeline
- YOLO series (YOLOv1, YOLOv2, YOLOv3)
  - Region proposals approach
  - A single fully convolutional network

Outline

Advantages of YOLO architecture

- End-to-end Architecture

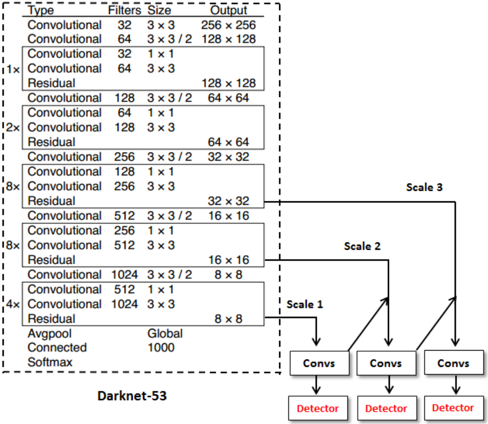


Figure 3: YOLOv3 architecture

Advantages of YOLO architecture

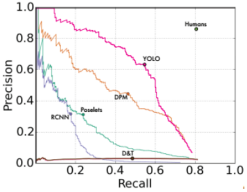
- End-to-end Architecture
- Real-time Performance

Real-Time Detectors	Train	mAP	FPS
100Hz DPM [31]	2007	16.0	100
30Hz DPM [31]	2007	26.1	30
Fast YOLO	2007+2012	52.7	**155**
YOLO	2007+2012	**63.4**	45
Less Than Real-Time			
Fastest DPM [38]	2007	30.4	15
R-CNN Minus R [20]	2007	53.5	6
Fast R-CNN [14]	2007+2012	70.0	0.5
Faster R-CNN VGG-16[28]	2007+2012	73.2	7
Faster R-CNN ZF [28]	2007+2012	62.1	18
YOLO VGG-16	2007+2012	66.4	21

Figure 3: Real-time detection systems testing on PASCAL VOC 2007. This table is from [6].

Advantages of YOLO architecture

- End-to-end Architecture
- Real-time Performance
- Better Generalizability



	VOC 2007	Picasso		People-Art
	AP	AP	Best F_1	AP
YOLO	**59.2**	**53.3**	**0.590**	**45**
R-CNN	54.2	10.4	0.226	26
DPM	43.2	37.8	0.458	32
Poselets [2]	36.5	17.8	0.271	
D&T [4]	-	1.9	0.051	

Figure 3: Generalization results on Picasso and People-Art work. The figure comes from [6].

Advantages of YOLO architecture

- End-to-end Architecture
- Real-time Performance
- Better Generalizability
- Less False Positive Errors

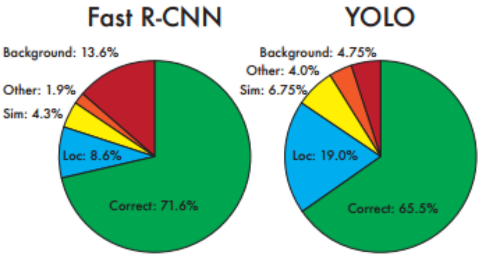


Figure 3: Error analysis: Fast R-CNN vs. YOLO. The figure from [6].

Advantages of YOLO architecture

- End-to-end Architecture
- Real-time Performance
- Better Generalizability
- Less False Positive Errors
- Comparable .5AP accuracy

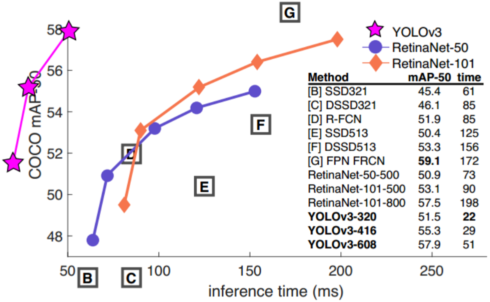


Figure 3: Speed and mAP-50 accuracy comparisons. This figure is a screenshot from [8]

Colon Polyps Detection System

- YOLOv3 Architecture

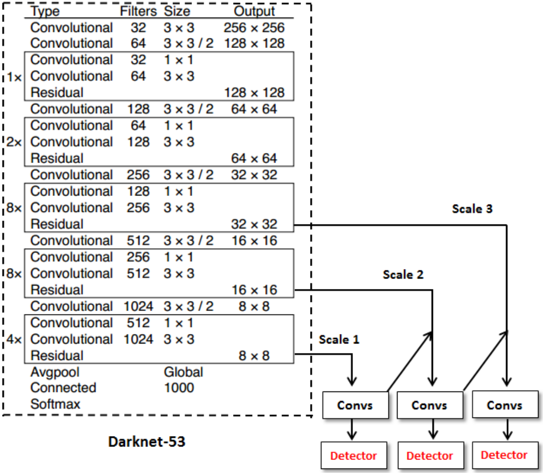
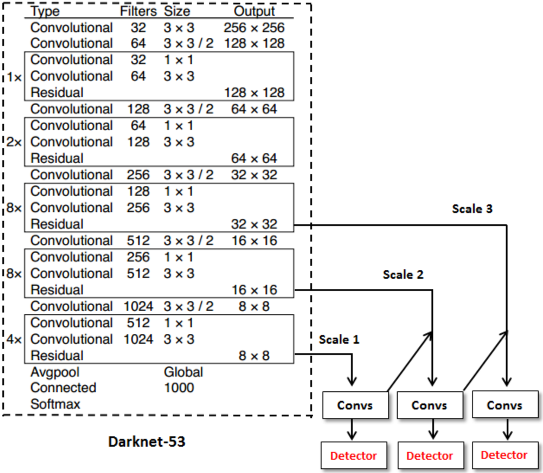


Figure 4: YOLOv3 architecture

Colon Polyps Detection System

- YOLOv3 Architecture
  - Fully convolutional network

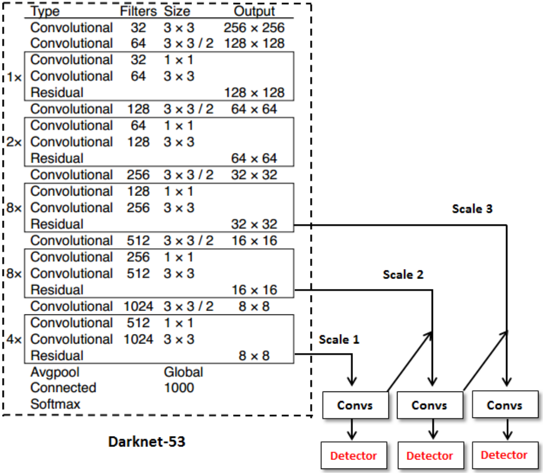


	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 4: YOLOv3 architecture

Colon Polyps Detection System

- YOLOv3 Architecture
  - Fully convolutional network
  - Residual blocks

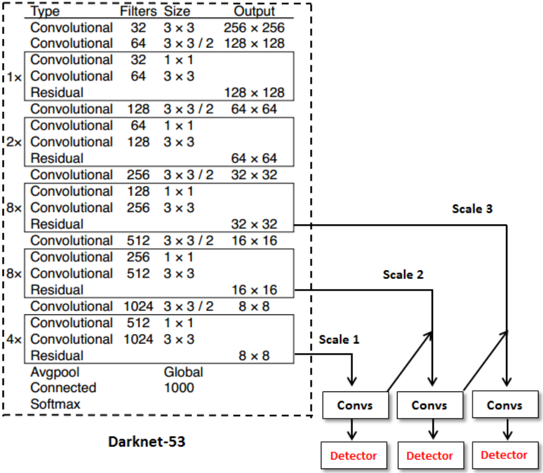


	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 4: YOLOv3 architecture

Colon Polyps Detection System

- YOLOv3 Architecture
  - Fully convolutional network
  - Residual blocks
  - Multi-scale prediction

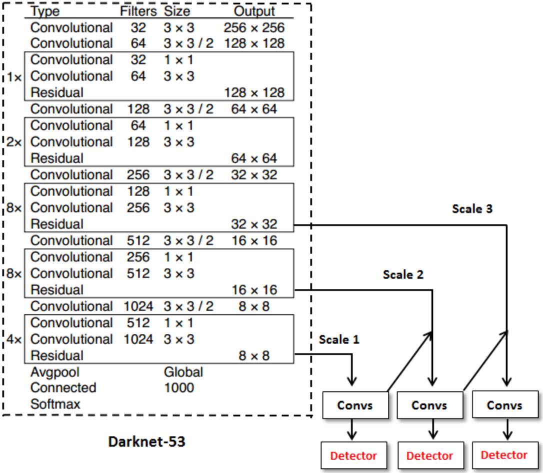


	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 4: YOLOv3 architecture

Colon Polyps Detection System

- YOLOv3 Architecture
  - Fully convolutional network
  - Residual blocks
  - Multi-scale prediction
- Transfer Learning
  numfiters=18



	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 4: YOLOv3 architecture

Outline

Training Settings

- Dataset partitioning
  - Training set: 224 of the 324 labeled images.
  - Test set: the rest of the labeled images.
- Multi-scale training approach
- Data augmentation: rotation, saturation, exposure, etc.
- Learning rate: 0.1, momentum: 0.9, weight decay: 0.0005
- Batch size: 64, Max batches: 15000.
- 15.5 hours for training.
  - A Titan X GPU for the first 2500 batches.
  - Two Titan X GPUs for subsequent iterations.

Average loss curves

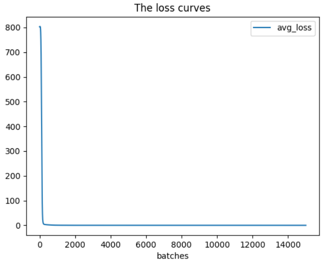


Figure 5: Curve of exponentially weighted average loss with coefficient being 0.9

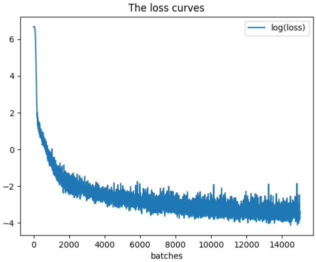


Figure 6: Curve of logarithm of average loss

Outline

Testing

- Testing on a Titan X GPU.
- 75.26% recall and 70.87% precision
- 0.0306 seconds per image (32.6 fps). Real-time performance.

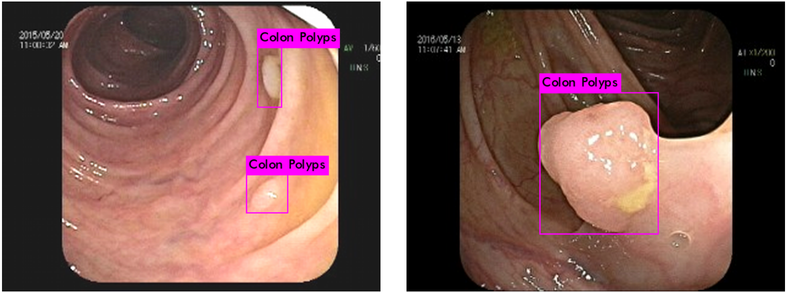


Figure 7: Examples of correctly detected colon polyps.

Outline

- Regarded as a classifier
- Dataset: test set + 100 normal images
- Comments on the classifier
  - Capable of classifying colon polyps images from normal images.
  - Struggling to precisely localize the colon polyps.

	Ground Truth		Total
Prediction	Normal	Colon Polyp(s)	
Normal	93	11	104
Colon Polyp(s)	13	83	96
Total	106	94	200

Table 1: Results of classification evaluation. recall=83/94=88.30%, precision=83/96=86.46%

Analysis of the false positive errors

- Mistakes other polyps as colon polyps.
- Need further improvement to learn more detailed features of colon polyps.

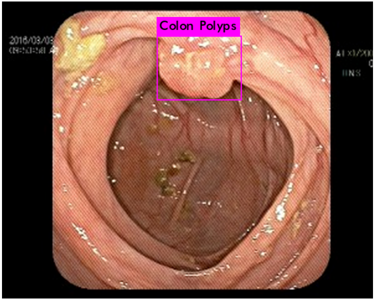


Figure 8: An example of true positive predictions

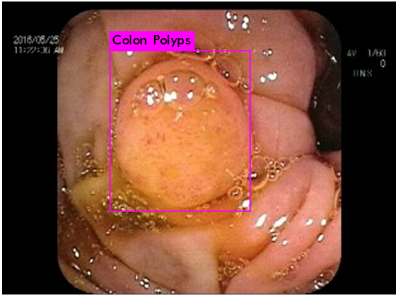


Figure 9: An example of false positive predictions

Conclusions

- Our colon polyps detection system:
  - achieves real-time performance
  - performs quite well in distinguishing images of colon polyps from normal images
  - struggles to figure out the exact lesion(病灶) locations.
  - tend to mistake other polyps as colon polyps

Conclusions

- Our colon polyps detection system:
  - achieves real-time performance
  - performs quite well in distinguishing images of colon polyps from normal images
  - struggles to figure out the exact lesion(病灶) locations.
  - tend to mistake other polyps as colon polyps
- Future work:
  - Collect More training data
  - Try other network architecture or transfer learning architecture

Bibliography I

Shie C K, Chuang C H, Chou C N, et al.
Transfer representation learning for medical image analysis
International Conference of the IEEE Engineering in Medicine & Biology Society. , Conf Proc IEEE Eng Med Biol Soc, 2015:711.

P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan.
Object detection with discriminatively trained part based models.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627 – 1645, 2010. 1, 4

Bibliography II

R. Girshick, J. Donahue, T. Darrell, and J. Malik.
Rich feature hierarchies for accurate object detection and semantic segmentation.
In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 580 – 587. IEEE, 2014. 1, 4, 7

Girshick, Ross.
Fast r-cnn.
Proceedings of the IEEE International Conference on Computer Vision. 2015.

Bibliography III

Ren S, He K, Girshick R, et al.
Faster r-cnn. Towards real-time object detection with region proposal networks
[C]. NIPS,2015.

Redmon, Joseph; Divvala, Santosh; Girshick, Ross; Farhadi, Ali
You Only Look Once: Unified, Real-Time Object Detection.
arXiv:1506.02640. 06/2015.

Redmon, Joseph and Farhadi, Ali.
YOLO9000: Better, Faster, Stronger
arXiv:1612.08242. 2016.

Bibliography IV

Redmon, Joseph and Farhadi, Ali.
YOLOv3: An Incremental Improvement.
arXiv, 2018.

Kermany et al., 2018, Cell 172, 1122 – 1131 February 22, 2018
2018 Elsevier Inc. https://doi.org/10.1016/j.cell.2018.02.010