Super-Resolution of Low-Quality Images for Realtime Pothole Detection

CS230 - Autumn 2020

Moo Jin Kim, Sharan Ramjee

Stanford University

Introduction

- Publicly available pothole detection datasets largely consist of high-resolution dashcam footage.
- Given the expensive nature of high-resolution dashcams, drivers instead opt for cheaper, low-resolution alternatives.
- As such, object detection models do not perform well for use with a majority of dashcams due to the domain mismatch between the high-resolution training data and the low-resolution test data.
- We present a novel approach to address this issue for real-time automated pothole detection through the use of Super-Resolution Generative Adversarial Networks (SRGANs).

Related Works

- Past approaches for automated pothole detection involved accelerometers, gyroscopes, wireless IoT sensors, thermal imaging cameras, etc.
  - High cost / complexity for setting up equipment
- Modern approaches leverage deep learning + dashcam footage
  - Various YOLOv3 architectures achieved the quickest and most reliable pothole detection
- However, no past work evaluated the effects of super-resolution on dashcam images
  - Kim et al. used SRGAN on images from CCTV cameras before feeding into CNN-based vehicle model classifier; upscaling from 224x224px to 896x896px led to significant increases in classification accuracy

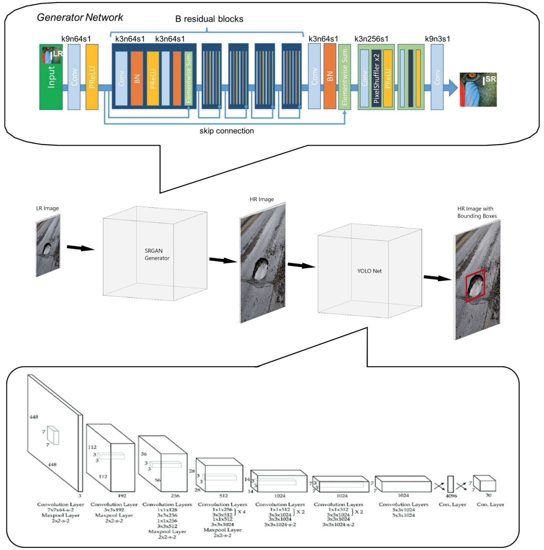
Dataset

3,888 RGB images (from Stellenbosch University)

- 2,634 images in train set (68%); 627 in val set (16%); 627 in test set (16%)
- Originally 3680x2760 pixels
- 3 versions of dataset:
    - Rescaled to 4K (3840x2160px)
    - Downscaled to 720p (1280x720px)
    - Downscaled to 360p (640x360px)

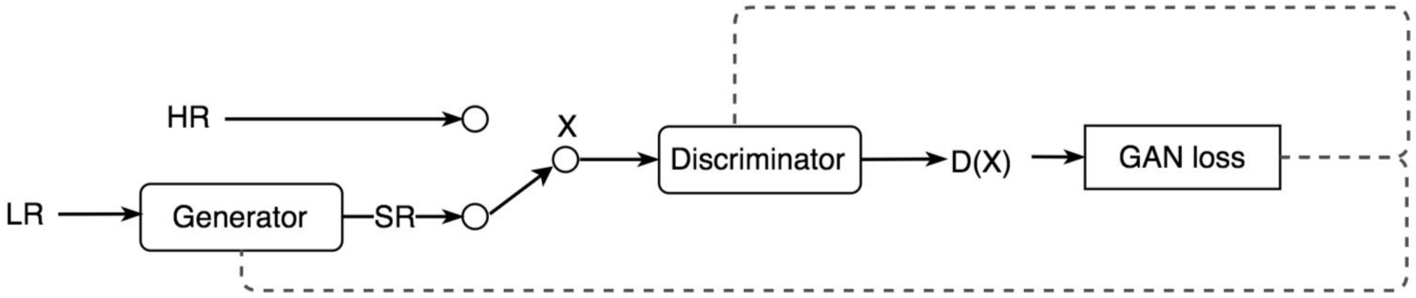
Methods

- The pothole detection pipeline consists of an SRGAN followed by a YOLOv4 model where the SRGAN performs super-resolution on the incoming stream of low-resolution dashcam frames and the YOLO model performs object detection.



SRGAN

- The SRGAN consists of a generator, which upscales the low-resolution (LR) images to super-resolution (SR) images, and a discriminator, which distinguishes between the HR and SR images and backpropagates the GAN loss to train the discriminator and generator.



YOLO

- YOLO network splits input image into grid of cells
  - each cell predicts whether center of a pothole is inside of it
  - outputs vector w/ confidence scores of all cells + bounding boxes
- Using YOLOv4 (April 2020) for this project
  - Trained on 3 versions of the training set
    - 360p
    - 720p
    - 4K

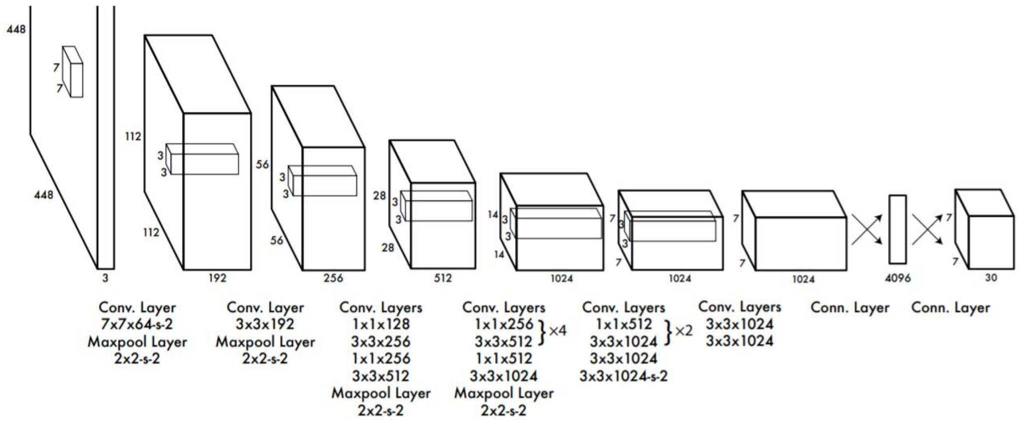


Image source: https://arxiv.org/pdf/1506.02640.pdf

Experiments and Results (Quantitative)

Model	Training set	Test set	Evaluation Metrics	
			mAP (%)	F1-score
YOLOv4 (baseline)	720p	720p	64.01	0.67
YOLOv4	4K	720p	60.67	0.64
YOLOv4	4K	**4K (SRGAN)**	**65.84**	**0.68**
YOLOv4 (upper benchmark)	4K	4K (original)	68.79	0.70
YOLOv4 (baseline)	360p	360p	28.49	0.39
YOLOv4	720p	360p	31.09	0.42
YOLOv4	720p	**720p (SRGAN)**	**60.06**	**0.65**
YOLOv4 (upper benchmark)	720p	720p (original)	64.01	0.67

Experiments and Results (Qualitative): SRGAN

Experiments and Results (Qualitative): YOLO

Demo: SRGAN-to-YOLO Output

Conclusion

Low-resolution dashcams can be used for more accurate real-time pothole detection

Dashcams can record at low resolutions to save memory

Can apply SRGAN for tougher object detection problems

Image sources: Bottom left: https://en.wikipedia.org/wiki/Hard_disk_drive Bottom right: https://www.mentalfloss.com/article/51292/night-sky-6-hours-stars

Future Work

- Assess how pothole detection speed and performance changes with different / newer models
  - SRResNet
  - FaSTGAN
  - YOLOv5
  - PP-YOLO

References

[1] S. Nienaber, M. J. Booysen, and R. S. Kroon, “Detecting potholes using simple image processing techniques and real-world footage,” SATC, July 2015. Pretoria, South Africa.

[2] S. Nienaber, R. S. Kroon, and M. J. Booysen, “A comparison of low-cost monocular vision techniques for pothole distance estimation,” in *IEEE CIVTS*, December 2015. Cape Town, South Africa.

[3] E. N. Ukhwah, E. M. Yuniarno, and Y. K. Suprapto, “Asphalt pavement pothole detection using deep learning method based on yolo neural network,” in *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pp. 35–40, IEEE, 2019.

[4] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681–4690, 2017.

[5] P. Mohan, V. N. Padmanabhan, and R. Ramjee, “Nericell: Rich monitoring of road and traffic conditions using mobile smartphones,” in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, SenSys ’08, (New York, NY, USA), p. 323–336, Association for Computing Machinery, 2008.

[6] G. D. D. Silva, R. S. Perera, C. Keppitiyagama, K. Zoysa, N. M. Laxaman, and K. Thilakarathna, “Automated pothole detection using wireless sensor motes,” 2008.

[7] K. Bansal, K. Mittal, G. Ahuja, A. Singh, and S. S. Gill, “Deepbus: Machine learning based real time pothole detection system for smart transportation using iot,” *Internet Technology Letters*, vol. 3, 03 2020.

[8] Aparna, Y. Bhatia, R. Rai, V. Gupta, N. Aggarwal, and A. Akula, “Convolutional neural networks based potholes detection using thermal imaging,” *Journal of King Saud University - Computer and Information Sciences*, 2019.

[9] P. Ping, X. Yang, and Z. Gao, “A deep learning approach for street pothole detection,” in *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*, (Los Alamitos, CA, USA), pp. 198–204, IEEE Computer Society, aug 2020.

[10] J. Kim, J. Lee, K. Song, and Y.-S. Kim, “Vehicle model recognition using srgan for low-resolution vehicle images,” in *Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition*, AIPR ’19, (New York, NY, USA), p. 42–45, Association for Computing Machinery, 2019.

[11] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” 2020.

[12] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[13] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in neural information processing systems*, pp. 2672–2680, 2014.

[15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[16] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft coco: Common objects in context,” in *European conference on computer vision*, pp. 740–755, Springer, 2014.

[17] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual dense network for image super-resolution,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2472–2481, 2018.