17 February 2022

ECS795P DEEP LEARNING AND COMPUTER VISION, 2022

COURSE WORK 1: IMAGE SUPER-RESOLUTION USING DEEP LEARNING

1. What are the main trends on the topic since the publication of the paper discussed?

In the paper, the CNN and Super resolution are two of the few key concepts that have seen major trends since its publication. The trends for each of them have been discussed below:

- CNN(Convolutional Neural Network):Predicting Road Accident Risk Using Google Maps and a Convolutional Neural Network:- The paper proposes a solution to train a convolutional neural network (CNN) model on past accident data and Google Maps images of accident road segments. The CNN trains by associating accidents with features in a Google Maps image (road characteristics, location features). Finally, the model generates a reliable accident risk score for any Google Maps image.[1]
- Super Resolution: Subspace decomposition-based methods (e.g., MUSIC[1]) and compressed sensing-based algorithms (e.g., SAMV[2]) are used to achieve SR over standard periodogram algorithms in some radar and sonar imaging applications (e.g., magnetic resonance imaging (MRI), high-resolution computed tomography).[13]

2. What are the key ideas of related published works since the original publication?

   Super Resolution: Super-resolution imaging (SR) refers to a group of techniques that improve (increase) an imaging system's resolution. The diffraction limit of systems is transcend in optical SR, whereas the resolution of digital imaging sensors is improved in geometrical SR. Deep convolutional networks are being used to perform super-resolution, according to promising research. It has been demonstrated that it can transform a 20x microscope image of pollen grains into a 1500x scanning electron microscope image. While this technique can boost an image's information content, there is no guarantee that the upscaled features are present in the original image, and deep convolutional upscales should not be used in analytical applications with ambiguous inputs.[2][3]

   Deep learning: Deep-learning architectures such as recurrent neural networks, and convolutional neural networks have been used in medical image analysis producing results that are comparable to, and in some cases superior to, traditional approaches. Image classification was then extended to the more challenging task of generating descriptions (captions) for images, often as a combination of CNNs and LSTMs. In image recognition systems, CNNs are frequently utilised. When used to facial recognition, CNNs resulted in a significant reduction in error rates. Another study found that "5,600 still photos of more than 10 participants" had a recognition rate of 97.6%. After manual training, CNNs were utilised to objectively judge video quality; the resulting system had a very low root mean square error.[4][5].

   Sparse coding: sparse signal recovery theory based sparce coding [11, 12], random forest [10] and artificial neural networks; to analyse statistical relationships between the LR and its corresponding HR image from substantial training examples.

   Mean squared error: Apart from GANs, which includes Adversarial Loss, Mean Square Error is used as a loss function. PSNR is used as an evaluation metric in all of these strategies.

3. What are the main problems solved or improvements over the original work?

   Dong et al. [6] proposed the first CNN-based SISR method, but the strategy required a long training time due to LR image interpolation. FSRCNN [9] is a compact hourglass-shaped CNN structure that uses deconvolution operations to reduce training time and improve performance. In FSRCNN, costly nonlinear mapping is avoided by shrinking and expanding the mapping layer at the beginning and end separately to limit mapping in a lowdimensional feature space [9].

4. What are the remaining problems from the published works so far?
   SRCNN [6] has a problem with increase in the number of layers, but Kim et al.[7]proposed DRCN with up to 16 recursions. Using image level frame of reference when there are more than 16 recursions is still a challenge. It's also worth noting from [6] that while a larger filter size improves performance, deployment speed decreases. The methods described above assume that high-resolution images are down-sampled to produce low-resolution images. Low-resolution images, on the other hand, are degraded more complicatedly in the real world. As a result, when existing methods are used directly to solve real-world low-resolution images in practice, the performance suffers.

5. What is an unsolved problem on the topic most interesting to you to solve and why?
   The encoder-decoder structure with coarse-to-fine methods was introduced by EDRN [8]. Because of the larger receptive field, the encoder-decoder structure can extract features with more context information. The coarse-to-fine structure can gradually restore lost information and mitigate noise effects. To summarise, there are many interesting problems that have been solved in the domain of super-resolution, and one of the interesting use cases that I would like to work with is the issue of working with real-life low images from various sources.

References:

1. *Aarya Agarwal , PREDICTING ROAD ACCIDENT RISK USING GOOGLE MAPS IMAGES AND A CONVOLUTIONAL NEURAL NETWORK ,International Journal of Artificial Intelligence & Applications (IJAIA) Vol.10, No.6, November 2019*
2. *Johnson, Justin; Alahi, Alexandre; Fei-Fei, Li (2016-03-26). "Perceptual Losses for Real-Time Style Transfer and Super-Resolution".*
3. *Grant-Jacob, James A; Mackay, Benita S; Baker, James A G; Xie, Yunhui; Heath, Daniel J; Loxham, Matthew; Eason, Robert W; Mills, Ben (2019-06-18). "A neural lens for super-resolution biological imaging". Journal of Physics Communications.*
4. *Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). "Subject independent facial expression recognition with robust face detection using a convolutional neural network" (PDF). Neural Networks.*
5. *Lawrence, Steve; C. Lee Giles; Ah Chung Tsoi; Andrew D. Back (1997). "Face Recognition: A Convolutional Neural Network Approach". IEEE Transactions on Neural Networks.*
6. *Dong, Chao, et al. "Image super-resolution using deep convolutional networks" IEEE transactions on pattern analysis and machine intelligence 38.2 (2015)*
7. *J. Kim, J. Kwon Lee, and K. Mu Lee, "Deeply-recursive convolutional network for image super-resolution," in CVPR, 2016.*
8. *G. Cheng, A. Matsune, Q. Li, L. Zhu, H. Zang, and S. Zhan. Encoder-decoder residual network for real super resolution. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2019.*
9. *C. Dong, C. C. Loy, and X. Tang, "Accelerating the superresolution convolutional neural network," in ECCV, 2016.*
10. *S. Schulter, C. Leistner, and H. Bischof. Fast and accurate image upscaling with super-resolution forests. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.*
11. *R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in International conference on curves and surfaces, 2010.*
12. *J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Trans. Image Process., vol. 19, no. 11, pp. 2861–2873, Nov. 2010.*
13. https://en.wikipedia.org/wiki/Super-resolution_imaging#Basic_concepts