FPD-M-net: Fingerprint Image Denoising Using M-Net Based Convolutional Neural Networks

July 20, 2018

1 Team details

- Team name:
  MIPture
- Team leader name:
  Sukesh Adiga V
- Team leader address, phone number and email:
  CVIT,
  IIIT-Hyderabad,
  India
  +91 9743493614
  sukesh.adigav@research.iiit.ac.in
- Rest of the team members:
  Jayanthi Sivaswamy
- Team website URL (if any)
- Affiliation:

2 Contribution details

- Title of the contribution:
  FPD-M-net: Fingerprint Image Denoising Using M-Net Based Convolutional Neural Networks
- Final score:
  The final score for validation and test set are given in below table. Validation results are update with latest results.

set	MSE	PSNR	SSIM
validation	0.0270	16.5149	0.8255
test	0.0268	16.5534	0.8261

- General method description:
  We pose the fingerprint denoising as a segmentation (foreground) task. We use a deep learning based architecture. Our architecture is based on the M-net which was proposed for brain segmentation. We modify this architecture and call the resulting architecture as $FPD\text{-}M\text{-}net$. It consists of Convolutional Layer(CONV), Maxpooling layer, Dropout layer, Batch Normalization layer(BN) and ReLU Activation Functions with Encoder and Decoder style of architecture which are very similar to M-net. The schematic representation of the architecture is shown in Figure 1. Dropout layer is used between two CONV-BN-ReLU block. A combination of a per-pixel loss and structure similarity loss are used for the loss function.

- References:

  - Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
  - Mehta, Raghav, and Jayanthi Sivaswamy. "M-net: A Convolutional Neural Network for deep brain structure segmentation." Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on. IEEE, 2017.

- Representative image / diagram of the method:

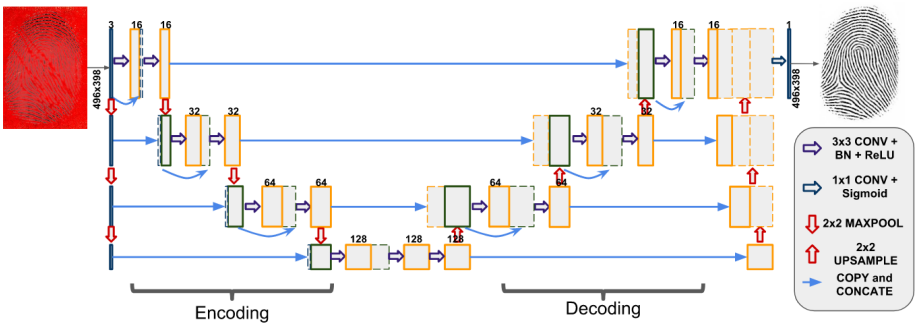

Figure 1: The schematic representation of the FPD-M-net architecture. Solid yellow boxes represent output of CONV-BN-ReLU block. Dashed boxes represent copied feature maps. Number of feature maps is denoted on the top of the box.

- Describe data preprocessing techniques applied (if any):
  Input and groundtruth images are padded with the edge values and images are normalized to $[0, 1]$ range.

3 Method description

3.1 Features / Data representation

Describe features used or data representation model (if any): N/A

3.1.1 Dimensionality reduction

Dimensionality reduction technique applied (if any): N/A

3.1.2 Compositional model

Compositional model used, i.e. pictorial structure (if any): N/A

3.1.3 Learning strategy

Learning strategy applied (if any):
We use a deep learning based architecture, which learn the model weights using conventional backpropagation mechanism.

3.1.4 Other techniques

Other technique/strategy used not included in previous items (if any): N/A

3.1.5 Method complexity

Estimated method complexity:
The training, validation and testing of our method is as simple as any other deep learning reconstruction/enhancing method.

3.2 Data Fusion Strategies

List data fusion strategies (how different feature descriptions are combined) for learning the model / network: Single frame, early, slow, late. (if any): N/A

3.3 Global Method Description

- Which pre-trained or external methods have been used (for any stage, if any): N/A
- Which additional data has been used in addition to the provided ChaLearn training and validation data (at any stage, if any): N/A

- Qualitative advantages of the proposed solution:
  The mean squared error(MSE) is arguably the dominant error measure across many diverse fields. In deep learning also, it is widely used as loss function for many application. But MSE and consequently the Peak Signal-to-Noise Ratio(PSNR), do not correlate well with human's perception of image quality. Structure similarity index(SSIM) is the other reference-based metric which is designed to improve on traditional methods such as MSE/PSNR and it correlate better with human's perception. So we have used a combination per-pixel loss and structure similarity loss. Though our method doesn't beat the others methods in challenge, we believe that our results could be perceptual better than other method.

- Results of the comparison to other approaches (if any): N/A

- Novelty degree of the solution and if is has been previously published:
  We pose the fingerprint denoising as a segmentation (foreground) task. We used the used M-net architecture which was proposed for brain segmentation. The skip connections introduced between convolution filters and deep super-vision functionality in our network which allows it to learn better features. We modify this architecture to suit for fingerprint denoising task. A combination of a per-pixel loss and structure similarity loss are used for the loss function.

4 Other details

- Language and implementation details (including platform, memory, parallelization requirements):
  The FPD-M-Net was implemented on a NVIDIA GTX 1080 GPU, with 12GB of GPU RAM on a core i7 processor. The entire architecture was implemented in Keras library using Theano backend. The model is trained for 75 epochs. A stochastic gradient descent(SGD) optimizer as used to minimize the per-pixel loss and structure similarity loss. The training parameters were: learning rate of 0.1; Nesterov momentum was set to 0.75; decay rate was set 1e-5. After 50 epochs learning rate is reduced to 0.01; Nesterov momentum was set to 0.95. The batch size was chosen as 8.

- Human effort required for implementation, training and validation?
  Our architecture is very easy to implement and it is end-to-end. The model doesn't use any pre-processing and post-processing stage, So training and validation doesn't require much human effort.

- Training/testing expended time?
  Training for 75 epochs took about a week. The testing time for each images is on an average of 0.72seconds on GPU machine.

- General comments and impressions of the challenge? what do you expect from a new challenge in face and looking at people analysis?

Overall the challenge was very interesting. Some of the information which we sent through e-mail could have been uploaded directly in competition site itself.