Edge-guided CNN for Denoising Images from Portable Ultrasound Devices

Yingnan Ma, Fei Yang, Anup Basu

Dept. of Computing Science, University of Alberta, Canada
Ultrasound diagnostic technology is popular in the medical field because of its lower cost compared to MRI and CT devices.

Ultrasound is non-invasive and does not subject patients to harmful radiations, like X-rays and CT.

Ultrasound generates images of structures within the human body by using low-power but high-frequency sound waves.

Portable ultrasound devices have been widely utilized to diagnose diseases in the human internal organs, such as the heart and liver.
Limitation:

• The quality of the scanned images is relatively poor compared to standard equipment in hospitals.
• Handheld devices may produce more noise, which can bring difficulties for humans to distinguish the image information for diagnosis.
• There are many sources that can cause noise:
  • Technical limitations of the equipment.
  • Environmental interference.
  • Body factors, such as body fat.
Overview of deep learning methods

• **DnCNN**
  DnCNN is modified from the VGG network with pooling layers removed. It sets the size of the convolutional filters to 3×3 to enlarge the receptive field. It also involves batch normalization to boost the training speed.

• **IRCNN**
  IRCNN includes seven dilated convolutional layers to enlarge the receptive field. The dilation factor is increased from 1 to 4 and decreased back to 1. It also involves Batch normalization to accelerate training.

• **FFDNet**
  FFDNet is a non-blind model. It designs a tunable noise level map to control the trade-off between the denoising performance and the detail preservation. It also involves down-sampling of sub-images to guarantee a good trade-off between inference speed and denoising performance.

• **Noise2Self**
  It can remove blind-level noise without using the noise estimation. It can learn with a single image. Noise2Self utilizes groups of features with conditional independent noise to predict one another. After the learning, a self-supervised loss can be obtained to train a deep convolutional neural network, such as DnCNN.
Proposed Methods

- Edge detection
- Noise addition
- Combination with DnCNN
- Combination with IRCNN
Edge Detection

- Canny edge detector can extract edges based on the difference in intensity. It uses an edge detection operator to obtain the intensity gradients and apply non-maximum suppression to roughly determine edges.
- Finally, double threshold is applied to remove weak potential edges.
Noise Addition

• In the state of the art, Gaussian noise is uniformly added to the original images. However, the random noise added on the edges can mislead a neural network into mapping the edges as residuals. This situation is common when we train the network with high noise levels.

• We use edge information to guide the noise addition. We generate random Gaussian noise and re-position noise that appears on the edges to obtain the edge-guided noise.

• Given the unknown noise level in ultrasound images, we generate noise within a moderate noise range instead of a fixed noise level.
• We take the noise-added patches as the input and use the original patches as the ground truth to provide references.
• To make the DnCNN model work for ultrasound images with a blind noise level, the depth of the neural network is set to 20 so that the size of the receptive field can be increased to 41 x 41.
• Initially, Conv + ReLU is taken as the first layer of the DnCNN model, where 64 feature maps are generated by 64 filters with size 3 x 3 x 1. Following Conv, rectified linear units (ReLU) are involved for non-linearity. From depth 2 to 19, DnCNN takes Conv + BN + ReLU as the layers, which involves 64 filters with size 3 x 3 x 64. Finally, the DnCNN model takes Conv as the last layer, where the output is reconstructed by a 3 x 3 x 64 filter.
Combination with IRCNN

- IRCNN contains seven layers. The first layer is a Dilated Convolution (the dilated convolution is 3×3, and the dilation factor is 1) + ReLU.
- The second to sixth layers are Dilated Convolution (the dilated convolutions are 3×3, and the dilation factors are 2, 3, 4, 3, and 2) + BN + ReLU.
- The last layer is a Dilated Convolution with a dilation factor of 1. The number of feature maps for each middle layer is 64.
Data Arrangement

In our experiments, we utilize our ultrasound dataset to train and test various learning approaches. Data was collected using our Clarius handheld ultrasound device. Our data includes 500 ultrasound images involving 400 ultrasound images of size 180 x 180 for training, 46 ultrasound images of size 321 x 481 for validation and 54 ultrasound images of size 440 x 380 for testing.
Comparison

Denoising results for different methods on sample image: (a) Original image (b) Edge-guided DnCNN (c) DNCNN (d) Edge-guided IRCNN (e) IRCNN (f) Noise2Self (g) FFDNet (h) Anisotropic Diffusion Filter (i) Median Filter (j) Wavelet Filter (k) Non-local Means Filter (l) Bilateral Filter.
Comparison

**PSNR AND SSIM RESULTS OF DIFFERENT METHODS ON TESTING SET.**

<table>
<thead>
<tr>
<th>Methods</th>
<th>E-DnCNN (Blind)</th>
<th>DnCNN (Blind)</th>
<th>E-IRCNN (Blind)</th>
<th>IRCNN (Blind)</th>
<th>Noise2Self (Blind)</th>
<th>FFDNet (Sig15)</th>
<th>Anisotropic Filter</th>
<th>Median Filter</th>
<th>Wavelet Filter</th>
<th>NLM Filter</th>
<th>Bilateral Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>34.9980</td>
<td>33.5115</td>
<td><strong>37.0241</strong></td>
<td>33.3894</td>
<td>29.4837</td>
<td>34.1458</td>
<td>31.3950</td>
<td>32.3649</td>
<td>28.2906</td>
<td>33.8018</td>
<td>34.6056</td>
</tr>
<tr>
<td>SSIM</td>
<td><strong>0.9221</strong></td>
<td>0.9093</td>
<td><strong>0.9433</strong></td>
<td>0.9067</td>
<td>0.9086</td>
<td>0.8434</td>
<td>0.6702</td>
<td>0.6955</td>
<td>0.1973</td>
<td>0.8183</td>
<td>0.8674</td>
</tr>
</tbody>
</table>

**SUBJECTIVE USER EVALUATION RESULTS OF DIFFERENT METHODS ON TESTING SET.**

<table>
<thead>
<tr>
<th>Methods</th>
<th>E-DnCNN (Blind)</th>
<th>DnCNN (Blind)</th>
<th>E-IRCNN (Blind)</th>
<th>IRCNN (Blind)</th>
<th>Noise2Self (Blind)</th>
<th>FFDNet (Sig15)</th>
<th>Anisotropic Filter</th>
<th>Median Filter</th>
<th>Wavelet Filter</th>
<th>NLM Filter</th>
<th>Bilateral Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td><strong>253</strong></td>
<td>206</td>
<td><strong>250</strong></td>
<td>211</td>
<td>192</td>
<td>150</td>
<td>41</td>
<td>39</td>
<td>83</td>
<td>92</td>
<td>133</td>
</tr>
<tr>
<td>Average</td>
<td><strong>10.12</strong></td>
<td><strong>8.24</strong></td>
<td><strong>10.00</strong></td>
<td>8.44</td>
<td>7.68</td>
<td>6.00</td>
<td>1.64</td>
<td>1.56</td>
<td>3.32</td>
<td>3.68</td>
<td>5.32</td>
</tr>
</tbody>
</table>
References:


References:


Thank You For Listening