**Cross-channel Communication Networks**

**Jianwei Yang**, Zhile Ren, Chuang Gan, Hongyuan Zhu, Devi Parikh

1. Georgia Tech; 2. Facebook AI Research (FAIR); 3. MIT-IBM Watson AI Lab; 4. Institute for Infocomm Research, A*Star

**Motivation and Contributions**
1. The neurons at each layer typically respond to input data independently and do not share any connections.
2. Redundant information could be accumulated among neurons.
3. Overlook the communication among neurons within the same level.

**Cross-channel Communication**

**Feature Encoder**
- Step 1: Flatten each channel response
- Step 2: Pass flattened feature into bottleneck layer

**Message Passing**
- Step 1: Averaging the whole feature map across channels
- Step 2: Computing affinities across channels
- Step 3: Weighted sum messages for other channels

**Feature Decoder**
- Step 1: Pass feature to bottleneck layer
- Step 2: Reshape output feature to channel shape
- Step 3: Add the reshaped feature to the original response map

**Measure the Neurons**
- Correlations between channel responses:
  - \( c_{ij} = \frac{1}{n} \sum_{m=1}^{n} (y_{i}[m] - \bar{y}_i)(y_{j}[m] - \bar{y}_j) \)
  - The lower correlations, the more diverse neurons become.
  - Use this to measure how neurons learn during the training with C3 block.

**Ablations Studies**
- C3 block helps to reduce the depth of neural networks while retain the performance.

**Experiments**

**Performance on vision tasks**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.28</td>
<td>41.7</td>
<td>67.73</td>
</tr>
<tr>
<td>Baseline + SE</td>
<td>0.28</td>
<td>41.8</td>
<td>66.57</td>
</tr>
<tr>
<td>Baseline + C3</td>
<td>0.35</td>
<td>46.00</td>
<td>69.34</td>
</tr>
</tbody>
</table>

**Table 1:** Classification accuracies (%) on CIFAR-100 [17] with different models.

**Visualizations**
- Neurons excites on the object regions in the image, either when there are a single or multiple objects. Also, different neurons have different response regions.

**Contributions**
1. The neurons at each layer typically respond to input data independently and do not share any connections.
2. Redundant information could be accumulated among neurons.
3. Achieve better performance across different vision tasks, such as image classification, object detection, and segmentation.