HyperGAN:
Generating Diverse, Performant Neural Networks

Neale Ratzlaff, Fuxin Li
Oregon State University
36th ICML 2019
Uncertainty

- High predictive accuracy is not sufficient for many tasks
- We want to know when our models are uncertain about the data
Fixing Overconfidence

- Given many models, each model behaves differently on outlier data
- By averaging their predictions, we can detect anomalies
Fixing Overconfidence

◆ Given many models, each model behaves differently on outlier data

◆ By averaging their predictions, we can detect anomalies
Fixing Overconfidence

- Variational inference gives a model posterior where we can sample many models.
- Ensembles of models from random starts may also detect outliers.
Regularization is too Restrictive

- Learning with VI is restrictive, it cannot model the complex model posterior
- Without regularization, our outputs mode collapse, losing diversity
Implicit Model Distribution

- We learn an implicit distribution over network parameters with a GAN
- We can instantly generate any number of diverse, fully trained networks
Implicit Model Distribution

- With a GAN, we can sample many networks instantly
- However, with just a Gaussian input, the generated networks tend to be similar
Mixer Network for Diverse Ensembles

- Want to generate diverse ensembles, without repeatedly training models.
- Our novel Mixer, transforms the input noise to learn complex structure.
- Mixer outputs are used to generate diverse layer parameters.
Generating Diverse Neural Networks

- Every training step we sample a new batch of networks
- The diversity given by the mixer lets us find many different models which solve the target task
HyperGAN Training: Full Architecture

- Prevent mode collapse by regularizing the Mixer with a Discriminator
- We use the target loss to train HyperGAN
Weight Diversity

HyperGAN learns diverse weight posteriors beyond simple Gaussians imposed by variational inference
Results - Classification

- MNIST 5000: train on 5k example subset.
- CIFAR-5: Restricted subset of CIFAR-10

<table>
<thead>
<tr>
<th>Method</th>
<th>MNIST</th>
<th>MNIST 5000</th>
<th>CIFAR-5</th>
<th>CIFAR-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 network</td>
<td>98.64 ± 0.3</td>
<td>96.69 ± 0.3</td>
<td>84.50 ± 0.6</td>
<td>76.32 ± 0.3</td>
</tr>
<tr>
<td>5 networks</td>
<td>98.75 ± 0.3</td>
<td>97.24 ± 0.14</td>
<td>85.51 ± 0.2</td>
<td>76.84 ± 1.0</td>
</tr>
<tr>
<td>10 networks</td>
<td>99.22 ± 0.09</td>
<td>97.33 ± 0.1</td>
<td>85.54 ± 0.2</td>
<td>77.52 ± 0.09</td>
</tr>
<tr>
<td>100 networks</td>
<td><strong>99.31 ± 0.02</strong></td>
<td><strong>97.71 ± 0.05</strong></td>
<td><strong>85.81 ± 0.02</strong></td>
<td><strong>77.71 ± 0.03</strong></td>
</tr>
<tr>
<td>APD</td>
<td>98.61</td>
<td>96.35</td>
<td>83.21</td>
<td>75.62</td>
</tr>
<tr>
<td>MNF</td>
<td>99.30</td>
<td>97.52</td>
<td>84.00</td>
<td>76.71</td>
</tr>
<tr>
<td>MC Dropout</td>
<td>98.73</td>
<td>95.58</td>
<td>84.00</td>
<td>72.75</td>
</tr>
<tr>
<td>Random Start</td>
<td>99.14</td>
<td>97.09</td>
<td>83.84</td>
<td>74.79</td>
</tr>
</tbody>
</table>
Out of Distribution Experiments

- Outlier detection on CIFAR-10 and MNIST datasets
  - MNIST → notMNIST
  - CIFAR (0-4) → CIFAR (5-9)
- Adversarial Examples: FGSM and PGD

Our increased diversity allows us to outperform other methods
Conclusion

- HyperGAN generates diverse models
- Makes few assumptions about output weight distribution
- Method is straightforward and extensible

Come to our poster for more details!