Semi-Supervised Learning with Deep Generative Models

Ondřej Jakš

2025

Outline

1. Introduction
2. Background
3. Proposed Models
4. Experiments and Results
5. Strengths and Limitations
6. Relation to Own Work
7. Conclusion
8. References

Motivation

- Labeled data is often scarce and expensive to obtain.
- Unlabeled data is abundant.
- Goal: leverage generative models for semi-supervised learning.
- This paper proposes methods combining deep generative models (VAE) with classifiers.

Paper Overview

- Authors: Kingma, Rezende, Mohamed, Welling (NeurIPS 2014)
- Title: Semi-Supervised Learning with Deep Generative Models
- Main contributions:
  - Introduces semi-supervised learning using VAEs.
  - Proposes two models: M1 and M2.
  - Combines generative and discriminative objectives.

Generative Models

- Learn a probabilistic model $p(x)$ of the data.
- Variational Autoencoders (VAE) approximate $p(x)$ using latent variables z.
- Variational inference: approximate posterior $q(z|x)$.

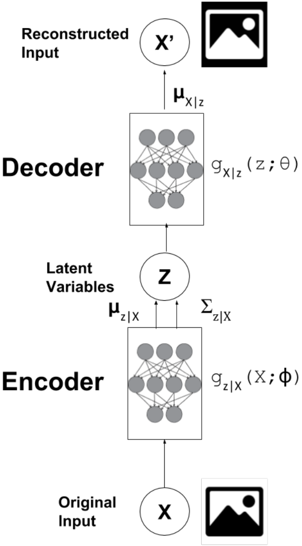
Semi-Supervised Learning

- Use a small labeled dataset and a large unlabeled dataset.
- Traditional methods: require heavy assumptions or are hard to scale.
- Proposed approach: leverage deep generative models to learn a latent space for classification.

Model M1: Latent Feature Model

- Simple VAE trained unsupervised on all data.
- Obtain latent representation z for each data point.
- Train a separate classifier on labeled z's.
- Advantages: simple, modular.

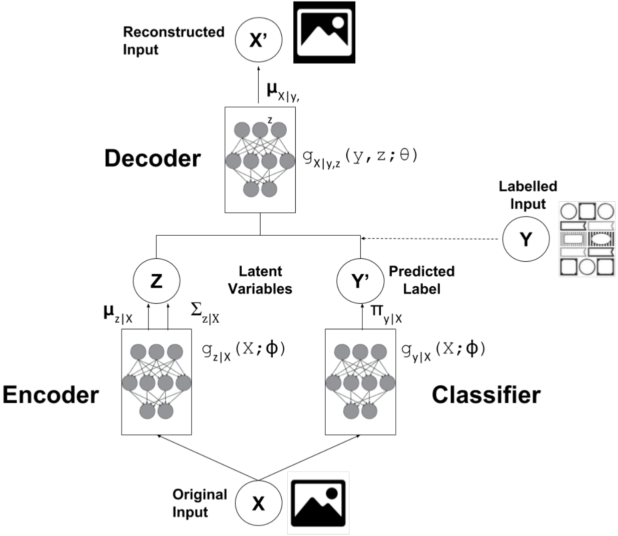
Model M1 Diagram



Model M2: Generative Semi-Supervised Model

- Explicitly models class label y jointly with latent variable z: $p(x, y, z) = p(y)p(z)p_\theta(x|y, z)$
- Inference model $q_\phi(z|x, y)$ approximates posterior.
- Enables semi-supervised learning directly in generative model.
- Combines generative loss with classification loss.

Model M2 Diagram



- Diagram shows flow: $x \rightarrow z \rightarrow \hat{x}$ and y inference

Model Training

Objective: minimization of a unified loss combining labeled and unlabeled data.

- For **unlabeled** samples, the model sums over all possible labels:

$$\mathcal{J}_{\text{unlabeled}} = \sum_{x \in D_{\text{unlabeled}}} \left[\sum_{y} q_\phi(y|x)\, \mathcal{L}(x, y) - \mathcal{H}(q_\phi(y|x)) \right]$$

- For **labeled** samples:

$$\mathcal{J}_{\text{labeled}} = \sum_{(x,y) \in D_{\text{labeled}}} [\mathcal{L}(x, y) - \alpha \log q_\phi(y|x)]$$

- The total training objective:

$$\mathcal{J} = \mathcal{J}_{\text{unlabeled}} + \mathcal{J}_{\text{labeled}}$$

- Optimized using stochastic gradient descent and the reparameterization trick.

Datasets

- MNIST, SVHN, NORB
- Varying numbers of labeled data
- Evaluate classification accuracy and generative quality

Classification Results

- M2 outperforms M1 and other baselines with few labels.
- Example: MNIST with 1000 labels:
  - M2: 97.5% accuracy
  - M1 + classifier: 90.5%
- Demonstrates effectiveness of joint generative-discriminative training.

Generative Quality

- Model can generate realistic samples.
- Can perform analogical reasoning: change label y while keeping z fixed.
- Useful for data augmentation and understanding latent space.

Strengths

- Innovative combination of generative models and classification.
- Scalable and flexible framework.
- Strong performance with very few labeled examples.

Limitations

- Simpler models may not capture all data structure.
- Some comparisons with other methods are missing.
- Optimization can be challenging for larger datasets.

Relevance to My Project

- My project adopts the same principle as the M1/M2 models: combining a VAE with a classifier to improve classification.
- The latent representation z learned by the VAE serves as the input to the classifier, similar to the feature-learning motivation in the paper.

Relevance to My Project

- My model employs a joint loss function that integrates reconstruction error, KL divergence, and a classification objective.
- I incorporate a self-organizing principle to structure the latent space.
- I hypothesize that these mechanisms will lead to a more coherent geometric structure of the latent space, potentially improving classification performance.

Conclusion

- Semi-supervised learning can be effectively performed using deep generative models.
- M1 and M2 offer complementary approaches.
- Generative modeling improves classification, especially with few labels.
- Provides a framework for further exploration of latent representations.

References

- Kingma, D.P., Rezende, D.J., Mohamed, S., Welling, M. (2014). *Semi-Supervised Learning with Deep Generative Models.* NeurIPS 2014. https://proceedings.neurips.cc/paper/2014/hash/6d42b1217a6996997ead5a8398c1f944-Abstract.html
- M1 diagram: Adapted from Bounded Rationality Blog, 2018. https://bjlkeng.io/posts/semi-supervised-learning-with-variational-autoencoders/
- M2 diagram: Adapted from Bounded Rationality Blog, 2018. https://bjlkeng.io/posts/semi-supervised-learning-with-variational-autoencoders/

Questions?

Thank you for your attention!
Any questions?