# Image Reconstruction by VAE

Xu Yingche*[1], Deng Xinjian[2], Wu Yakun[3], Zhai Penglin[4], and Xu Xinpeng[5]

[1]College of Electronic Science and Engineering, Jilin University, Changchun 130015, China
[2]School Of Machinery And Vehicles, Beijing Institute Of Technology, Beijing 100081, China
[3]School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan,430070, China
[4]School of Electrical Engineering, Chongqing University, Chongqing 400030, China
[5]School of underwater acoustic engineering, Harbin Engineering University, Harbin 150000, China
*Correspondence: Yingchexu@outlook.com; Tel: +86 13614267948

December 20, 2023

**Abstract**

Variational Autoencoders (VAEs) have become a recognized tool for learning data distributions in various applications [2]. Building on previous research, this paper investigates image reconstruction by incorporating Planar Flows into VAEs. As a form of Normalizing Flows, Planar Flows potentially aid VAEs in modeling more intricate distributions, which may enhance the quality of reconstructed images. Our study employs a well-established architecture that integrates Planar Flows with VAEs. Experiments on MNIST are provided to compare our replication efforts with traditional VAEs. The results suggest that utilizing Planar Flows can offer certain improvements in image reconstruction. This work further solidifies the understanding of the combined use of VAEs and Planar Flows in image reconstruction tasks.
**Keywords: Machine learning, VAEs, Normalizing Flows, Image Reconstruction.**

## 1 Introduction

The Inference methods of a probabilistic model are divided into two categories, exact inference and approximate inference. Normally we are willing to obtain the true posterior, yet with the expansion of dimensions, the calculation complexity has an exponential growth [5].

That's why we come up with approximate inference where we are calculating the approximate solutions of the problem on a smaller complexity. Two major methods are Markov Sampling and Variational Inference, which we discuss in this article.

The choice of approximate posterior distribution is one of the core problems in Variational Inference [4]. Traditional ways of making this approximation focus on simple distributions (Mostly Gaussian). Yet when it comes to complex distributions, sometimes bimodal distribution, The Gaussian family is unable to make accurate approximations.

Normalizing Flow is among the various methods available to address this problem. It provides a general mechanism for defining expressive probability distributions, generated from simple distributions [3]. By combining the Normalizing Flow with a Variational Auto Encoding Network, we expect to have a better performance in the task of image generation.

# 2 Variational Auto Encoding

At its core, a Variational Autoencoder (VAE) is a generative model that aims to learn a compressed and continuous representation of input data, often referred to as a "latent space", while also enabling the generation of new data samples that resemble the original data distribution.

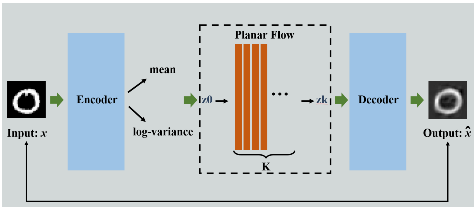


Figure 1: Variational Auto Encoder Method

VAEs combine elements from probabilistic modeling, neural networks, and variational inference. Here's a simplified description of the fundamental theory behind VAEs.

## 2.1 Space Representation

VAEs focus on learning a lower-dimensional representation of input data. This representation is known as the "latent space" or "latent variables".Each point in this space represents a different configuration of the data. The idea is to capture the essential features or patterns of the data in this compressed space.

## 2.2 Encoder and Decoder

VAEs consist of two main components: an encoder and a decoder.

**Encoder** The encoder takes input data and maps it to a distribution in the latent space. This distribution is typically a Gaussian distribution defined by its mean and standard deviation. The encoder's role is to learn how to compress the input data into the latent space.

**Decoder** The decoder takes a point from the latent space and maps it back to the data space to reconstruct the original input. The decoder learns how to generate data samples from the latent space.

### 2.3 Variational Inference

VAEs use variational inference to train the model and learn the parameters of the encoder and decoder. Variational inference involves approximating a complex distribution (the true posterior) with a simpler distribution (the variational distribution) that is easier to work with. In the case of VAEs, the variational distribution is the distribution produced by the encoder.

### 2.4 Sampling and Reparameterization

To train VAEs, a key technique is the reparameterization trick. Instead of directly sampling from the distribution produced by the encoder, which makes backpropagation difficult, the trick involves introducing a random noise term that is multiplied by the standard Gaussian.

## 3 Normalizing Flows

Normalizing Flows provide a mechanism to transform a simple distribution (like a Gaussian) into a more complex one, while still allowing for tractable density estimation and sample generation. This is done through a sequence of invertible transformations. The power of Normalizing Flows lies in their ability to model intricate and multi-modal distributions. By stacking multiple transformations, we can iteratively refine the distribution and better capture the underlying data distribution.

The math behind Normalizing Flows involves the change of variables formula, which allows the density of the transformed variable to be computed in terms of the density of the original variable. This keeps the density estimation tractable even after many transformations.

### 3.1 Planar Flows

Planar Flows, as a simple member of the family of Normalizing Flows, play a pivotal role in transforming basic probability distributions into more complex, nuanced distributions. Their primary objective is to enable sophisticated distribution shaping without compromising tractability.

The transformation mechanism of Planar Flows can be articulated by equation (1).

$$f(z) = z + uh(\omega^\top z + b) \tag{1}$$

Here:
$z$ represents the original latent variable.
$u$ and $\omega$ are weight vectors, with $b$ acting as the bias term.
$h$ is an element-wise non-linearity, typically embodied by the hyperbolic tangent function.

### 3.2 Jacobian Determinant

To fathom the density transformation in Planar Flows, understanding the Jacobian determinant is indispensable. For Planar Flows, the determinant of the Jacobian of the transformation is given by equations (2) and (3):

$$\psi(\mathbf{z}) = h^{'}(\omega^{\top}\mathbf{z} + b)\omega. \tag{2}$$

$$\left| det\frac{\partial f}{\partial \mathbf{z}} \right| = \left| det(\mathbf{I} + \mathbf{u}\psi(\mathbf{z})^{\top}) \right| = \left| 1 + \mathbf{u}^{\top}\psi(\mathbf{z}) \right|. \tag{3}$$

### 3.3 Log Probability Calculation via Jacobian

The transformed log probability density is not only shaped by the initial log density but is also influenced by the log determinant of the Jacobian. Mathematically, the transformed distribution and log probability after the transformation are expressed as equations (4) and (5):

$$\mathbf{z}_K = f_K \circ f_{K-1} \circ \cdots \circ f_1(\mathbf{z}). \tag{4}$$

$$\ln q_K(\mathbf{z}_K) = \ln q_0(\mathbf{z}) - \sum_{k=1}^{K} \ln \left| 1 + \mathbf{u}_k^{\top}\psi_k(\mathbf{z}_{k-1}) \right|. \tag{5}$$

In conclusion, Planar Flows emerge as a powerful tool in the domain of deep generative modeling. By facilitating the shaping of intricate probability distributions from basic foundational distributions, they bridge the divide between simplicity and complexity, ensuring that we have the flexibility to model real-world data distributions while retaining computational feasibility.

## 4 Experiments and Results

### 4.1 Distribution Approximation by Planar Flows

The objective of this section of the experiment is to practically test and verify the capability of Planar Flows in transforming distributions.

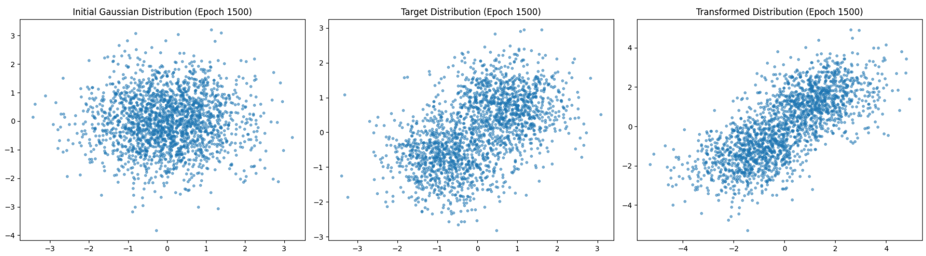


Figure 2: Transformation Results Using Planar Flows

We adopted the standard Gaussian distribution as our initial distribution, and a superimposed normal distribution with two distinct means (located at (-1.5, -1.5) and (1.5, 1.5) respectively) as our target distribution. We trained Planar Flows for the distribution approximation. Using K=15 layers of Planar Flows, and undergoing 1500 Epochs, the following figure presents the results of this experiment with a Loss value of 0.153.

## 4.2 Generative Tasks on MNIST

We primarily focused on the reconstruction training of each category of data from the MNIST dataset. We utilized a VAE constituted of an encoder and a decoder, and right after the encoder we employed a Planar Flow. The main purpose of the Planar Flow is to refine the posterior distribution of the variational autoencoder.

The process begins with data from MNIST serving as the input x. This input is fed into the encoder, which consists of two linear layers followed by a ReLU layer. Upon outputting the mean vector and log-variance, an initial $z_0$ is obtained. After passing through k layers of the Planar Flow, we derive a new $z_k$. Finally, this is fed into the decoder, which, like the encoder, is constructed with two linear layers and a ReLU layer, leading to the output of the reconstructed image. We trained by comparing the reconstructed image with the original one.

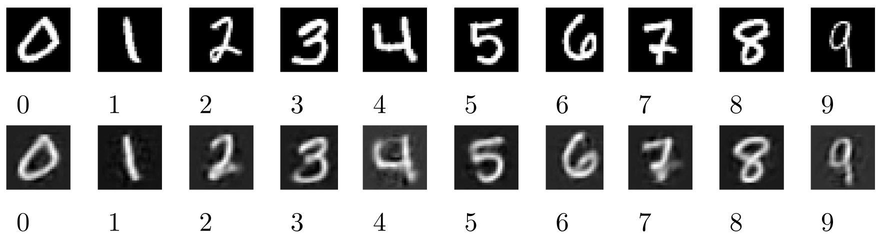


Figure 3: Original Images(above) and Reconstructed Images(below) without Normalizing Flow

We conducted training for 30 epochs and visualized the reconstructed images, shown in Figures 3 and 4. The first two results represent K=0 (Figure 3), meaning no Planar Flow was used to refine the posterior distribution.

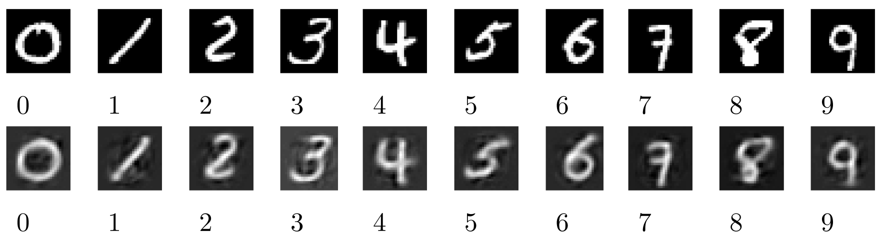


Figure 4: Original Images(above) and Reconstructed Images (below)in MNIST with Normalizing Flow when $k = 5$

In contrast, the second illustrates the outcome for K=5. Remarkably, both methods performed quite well in the task of reconstruction. (Figure 4)

We visualized the changes in our training loss. Due to the exaggerated loss from the initial iteration, the subsequent visual details aren't very distinct. Hence, we separately plotted a graph excluding the loss from that first iteration. It is evident that our loss value started over 200 and consistently decreased, eventually plateauing around 6. (Figure 5,6.)

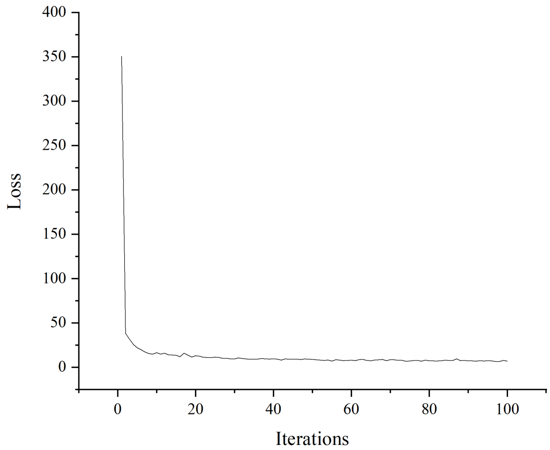


Figure 5: Loss1

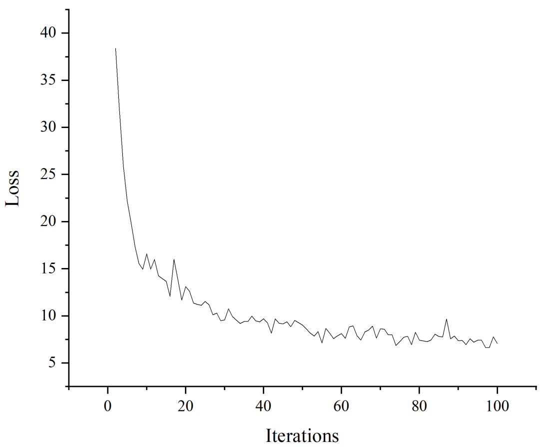


Figure 6: Loss2

# 5 Conclusions and Future Directions

## 5.1 Conclusions

This study investigated the potential of integrating Planar Flows, a subset of Normalizing Flows, into Variational Autoencoders (VAEs) for the purpose of enhancing image reconstruction. Our research built upon the foundational work of prior studies and went into the intricacies of probabilistic modeling, drawing connections between variational inference and deep generative modeling techniques. Experimental findings on the MNIST dataset showed that the amalgamation of VAEs with Planar Flows indeed fosters improvements in image reconstruction, particularly when compared with traditional VAEs. Specifically, the introduction of Planar Flows to the VAE architecture led to more nuanced posterior distributions, further enhancing the reconstruction fidelity.

However, as with many pioneering explorations, our study had its limitations. The stability and consistency of our results could benefit from further refinement. Despite these challenges, our findings hold promise for the future. The incorporation of Normalizing Flows, especially Planar Flows, into VAEs opens up new avenues for modeling complex data distributions, ultimately refining the granularity and quality of generative models.

### 5.2 Future Directions

As the field of machine learning continues to evolve and mature, we foresee a future where more sophisticated versions of these models can be developed. This would not only bolster image reconstruction capabilities but also revolutionize how we approach generative modeling tasks in various domains. As we move forward, it will be of paramount importance to continue bridging the gaps between theoretical constructs and practical implementations to harness the full potential of these models.

## 6 Acknowledgment

We would like to express our gratitude to Professor Jose Hernandez-Lobato and his team for their invaluable guidance and support throughout this project. They provided us with clear directions and constructive feedback on the experimental design and methodology, as well as insightful suggestions on the analysis and interpretation of the results. We learned a lot from their expertise and experience in the field of Machine learning.

We would also like to thank all the scholars whose works we have cited in the references section for their contributions to the existing knowledge base of Variational Inference and Normalizing Flows. Their research has laid a solid foundation for our study and inspired us to explore new perspectives and possibilities.

We are grateful to the University of Cambridge for providing us with the opportunity and resources to conduct this program. We appreciate the assistance and cooperation of the staff and students at Magdalene [3] [1].

## 7 Author Contributions

Conceptualization: X.Y, D.X, W.Y; methodology, validation, formal analysis: X.Y and D.X; resources, data curation: Z.P, D.X and D.X.; writing-original draft preparation, visualization: X.X and W.Y; writing-reviewing and editing: X.Y and W.Y; project administration: X.Y. All authors have read and agreed to the published version of the manuscript.

## 8 Funding

This research received no external funding.

## 9 Research Guidelines

This study followed the research guidelines of the Machine Learning and Machine Intellegence Cambridge Academic Program 2023.

## 10 Informed Consent Statement

Not Applicable.

## 11 Data Availability

Please contact the corresponding author(s) for all reasonable requests for access to the data.

## 12 Conflicts of Interest

The authors declare no conflict of interest.

## 13 Intellectual Property

The authors attest that copyright belongs to them, the article has not been published elsewhere, and there is no infringement of any intellectual property rights as far as they are aware.

## 14 References

[1] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *ArXiv*, abs/1601.00670, 2016.

[2] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. *CoRR*, abs/1312.6114, 2013.

[3] George Papamakarios, Eric T. Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.*, 22:57:1–57:64, 2019.

[4] Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International Conference on Machine Learning*, 2015.

[5] Zhi-Hua Zhou. *Machine learning*. Springer Nature, 2021.