RG and Sampling
Executive Summary

Iman Faisal (igf23)
Word Count: 675

Department of Physics, Cambridge

This project aimed to reproduce and critically analyse the RG-Flow model, a machine learning architecture that integrates concepts from theoretical physics, particularly the renormalisation group (RG) and tensor network methods. The central objective was to rigorously investigate and validate the findings presented by Hu et al. (2022), who introduced RG-Flow as a hierarchical and interpretable generative modelling approach based on normalising flows [1].

The renormalisation group is a theoretical tool developed to study complex physical systems exhibiting multiple interacting scales. It simplifies intricate systems by applying a systematic process known as "coarse-graining," wherein detailed microscopic features are progressively grouped into broader-scale characteristics [2]. This process effectively highlights essential underlying phenomena by discarding irrelevant microscopic details. RG has demonstrated remarkable success in explaining phase transitions and critical phenomena across various areas in physics, forming the theoretical backbone of many contemporary physical models.

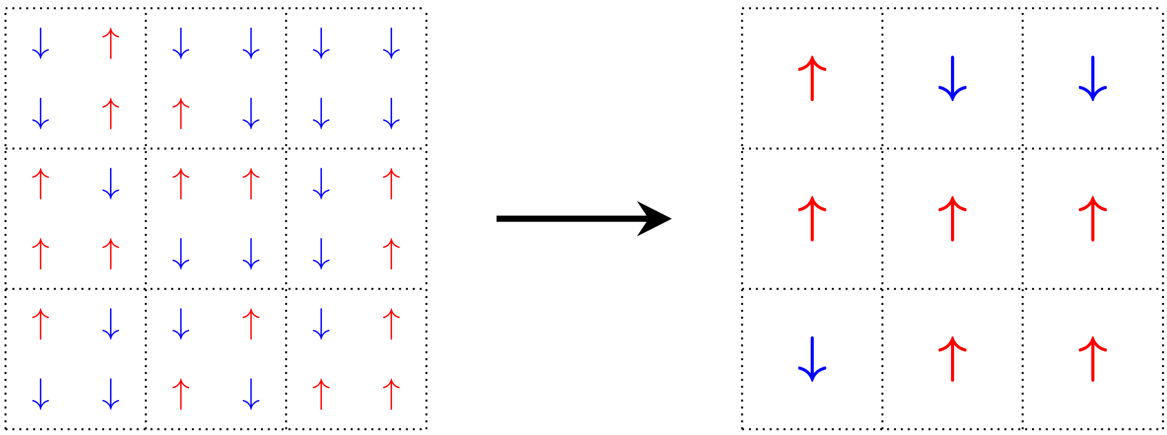

Figure 1: Illustration of Kadanoff block-spin coarse-graining transformation. The original lattice is divided into blocks and the majority vote defines the coarse-grained block-spin. In the case of a tie, ↑ wins.

The RG-Flow model draws extensively from the Multiscale Entanglement Renormalisation Ansatz (MERA), a tensor network approach designed to manage the challenges posed by long-range entanglement in quantum many-body systems [3]. MERA operates through two core procedures: disentanglers, which systematically simplify local interactions, and decimators, which merge multiple variables into more compact and manageable representations. This structure enables efficient coarse-graining, significantly simplifying the complexity of quantum states.

Within RG-Flow, these quantum-inspired processes are translated into classical transformations tailored explicitly for machine learning purposes. These transformations, termed coupling layers, operate in two complementary steps analogous to MERA's disentangler and decimator phases. Each RG-Flow layer progressively distils detailed image data into a hierarchy of increasingly abstract and compact representations. This hierarchical decomposition simplifies data complexity while maintaining interpretability, making it especially suitable for detailed feature rich, high-dimensional datasets.

Two distinct prior distributions, Gaussian and Laplacian, were evaluated to encode the data into latent variables effectively. The Gaussian prior, characterised by its smooth, bell-shaped distribution, facilitates dense, entangled latent representations. Conversely, the Laplacian prior, exhibits a sharp peak and heavier tails. It is thought to promote sparsity in latent representations, enhancing their interpretability. A temperature parameter was introduced to modulate the regularisation strength.

The RG-Flow model underwent rigorous testing on three distinct datasets: CelebA, a widely utilised dataset of facial images, and two synthetic datasets, MSDS1 and MSDS2, explicitly designed to evaluate the model's hierarchical disentanglement capabilities [4]. The CelebA images underwent meticulous preprocessing, including cropping, resizing, and data augmentation, ensuring uniformity and quality. The synthetic datasets provided controlled environments for precise assessments of RG-Flow's capacity to separate and manage hierarchical features.

Evaluation involved two primary metrics: Bits-per-Dimension (BPD) and the Fréchet Inception Distance (FID). BPD measures the efficiency with which the model encodes the data by quantifying the negative log-likelihood per data dimension. FID assesses the realism and perceptual quality of generated images relative to real images, utilising feature representations extracted from a pre-trained neural network. The results from these metrics indicated that RG-Flow effectively captured complex data distributions and produced realistic synthetic images.

Dataset	**BPD (Laplacian)**	**BPD (Gaussian)**	**FID (Laplacian)**	**FID (Gaussian)**
CelebA	3.7238	3.7432	83.815	80.418
MSDS1	1.8562	1.8505	103.164	103.637
MSDS2	1.9375	1.9457	105.798	106.914

Table 1: Comparison of bits per dimension (BPD) and Fréchet Inception Distance (FID) across different datasets using Laplacian and Gaussian priors.

We found that for relatively large batch size of 512, both priors preform approximately the same. We hypothesis this is due to noise dominating the loss landscape and overcoming any regularizing effects.

A detailed latent space analysis, employing receptive field studies, demonstrated RG-Flow's successful hierarchical learning. This measures the spatial extent to which changes in latent variables affect the output images. At the highest hierarchical levels, latent variables significantly influenced broad, structural aspects of images, such as facial outlines and overall lighting. Conversely, lower-level latent variables governed more specific, detailed features, including facial expressions, eye shapes, and subtle shading nuances. This can be seen in Figure. 2. This hierarchical separation of scales aligns closely with expectations of the renormalisation group. This shows that the model architecture serves as a good inductive bias for problems with scale dependent features.

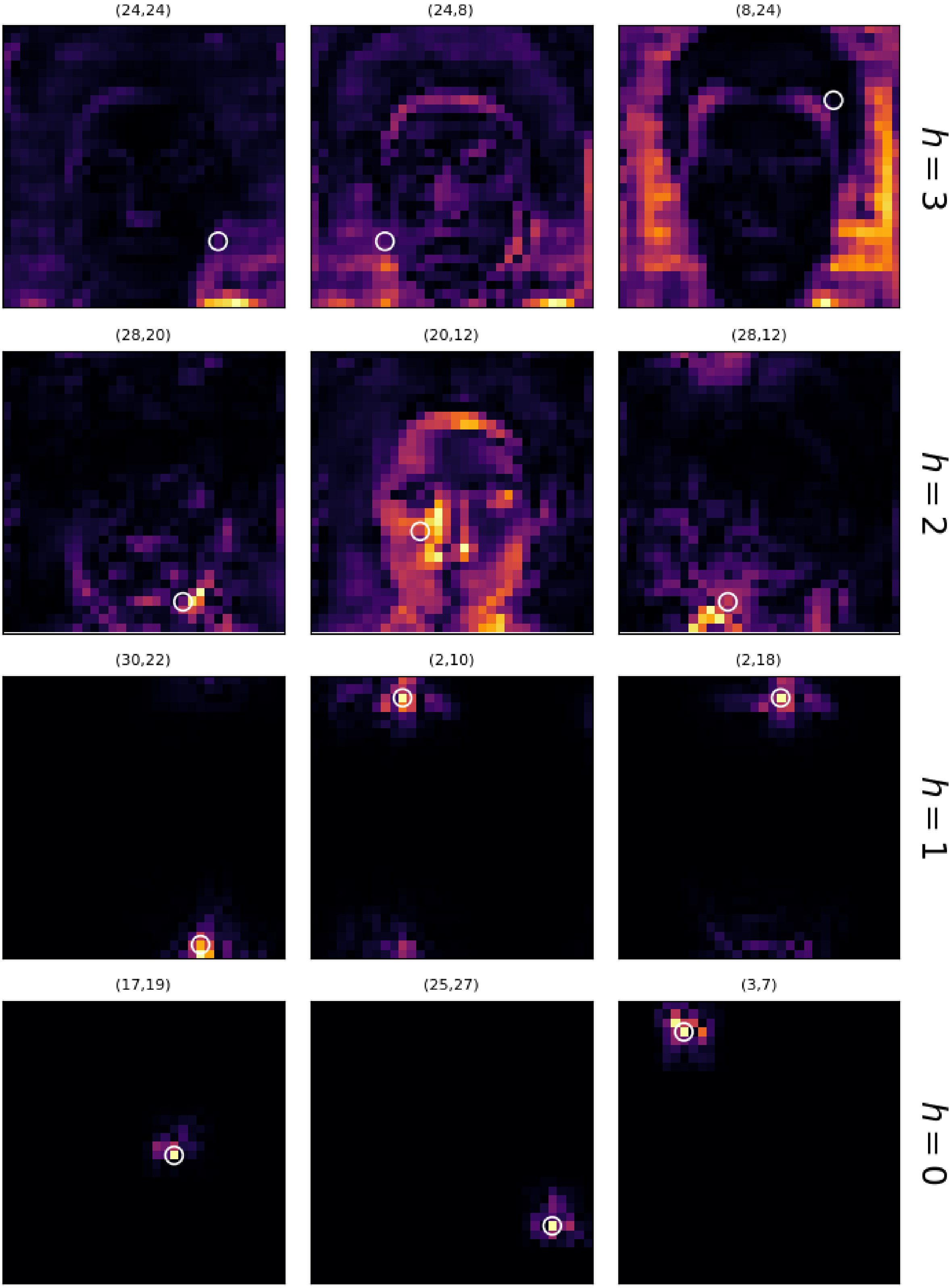

Figure 2: Receptive Fields for Gaussian RG-Flow trained on CelebA.

References

[1] Hu HY, Wu D, You YZ, Olshausen B, Chen Y. RG-Flow: a hierarchical and explainable flow model based on renormalization group and sparse prior. Machine Learning: Science and Technology. 2022 aug;3(3):035009. Available from: https://dx.doi.org/10.1088/2632-2153/ac8393.

[2] Kadanoff LP. Scaling laws for ising models near T_c. Physics Physique Fizika. 1966 Jun;2:263-72. Available from: https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.2.263.

[3] Vidal G. Class of Quantum Many-Body States That Can Be Efficiently Simulated. Physical Review Letters. 2008 Sep;101(11). Available from: http://dx.doi.org/10.1103/PhysRevLett.101.110501.

[4] Liu Z, Luo P, Wang X, Tang X. Deep Learning Face Attributes in the Wild. In: Proceedings of International Conference on Computer Vision (ICCV); 2015. .