Bits-Back Efficient Deep Image Compression

Gergely Flamich

June 5, 2019

Proposal

In this project we aim to create a novel state-of-the-art image compression scheme based on Variational Auto-Encoders [5] (VAEs). The images can be mapped to a latent space via the encoder of the VAE, so compression can be achieved by producing an efficient description of these latent representations. From these the decoder should be able to then reconstruct the original images with at most minor distortions.

The project proposes a Minimum Description Length (MDL) based variational coding scheme (based on the "bits-back argument", see [3]) using an approach very similar to β-VAEs [2] to compress images. The starting point for this coding scheme is MIRACLE [1] in which the authors demonstrate both promising theoretical results as well as good empirical evidence of their method for compressing neural networks.

The advantages of this method are that it is

- **principled**, as it is backed up by solid information-theoretical results that give us nice guarantees,
- **efficient** (in the sense of bits-back efficiency), in that it allows us to compress the data close to the theoretical limit,
- **differentiable** (which sets it apart from most contemporary approaches), meaning it can be trained end-to-end.

The following 4 items have been identified as the key challenges of the project[1]

1. Finding the correct benchmarks and metrics to use for evaluating our results.

   After some preliminary research, the dataset for the Challenge in Learned Image Compression (CLIC) 2018 [7] might be a good resource for state-of-the-art benchmarks. It also seems that `bits per pixel (bpp)` for compression efficiency and `Multiscale Structural Similarity (MS-SSIM)` [6] and `Peak Signal-to-Noise Ratio (PSNR)` [4] for image quality are reasonably well-established metrics within the field.

2. Dividing the latent space of the VAE into blocks so that the rejection sampling-based method proposed for MIRACLE becomes computationally feasible and can be adapted for this setting.
3. Finding a suitable encoder/decoder architecture.
4. Finding a suitable training loss for the VAE.

   As a starting point, we will likely start with a simple MSE loss.

The aim of this project is to at least match, but hopefully surpass the state-of-the-art compression techniques on appropriate datasets according to the established metrics.

[1]These have been identified in a meeting with M. Havasi on 25 March 2019.

Workplan

To ensure a smooth and continuous progression, we have opted for weekly half-hour meetings with our first supervisor (M. Havasi) and occasional meetings with our second (J. M. Hernández-Lobato). We are also planning on occasionally consulting F. Huszár as he is an expert in the field and can provide key insights for the project. Our work plan for the project is as follows:

12 April - 6 May: Due to deadlines and exams, this period will be the least productive, our goal during this time is to begin the literature review and look for/ implement some benchmarks.

7 May - 20 May: Finishing literature review, finishing the implementation of benchmarks, implementing first MIRACLE-based model, testing on a tractable dataset (e.g. MNIST). This avoids having to worry about point 2) and to some extent 3) on the list in the above section.

21 May - 3 June: Experimentation with different losses (e.g. moving to VGG loss or some more sophisticated perceptual loss) and hyperparameters. Attempt to upscale architecture for medium to high-resolution images.

4 June - 17 June: Depending on the success of the previous two weeks, carry on experimenting with different architectures/hyperparameters/losses or attempt to address the chunking problem depending on its severity. Preparation of poster and talk for the Industry Day to be held on 17 June.

18 June - 1 July: Depending on success so far, either keep attempting to overcome the difficulties caused by dealing with high-res images or attempt some extensions to the project, e.g. lifting some constraints/ assumptions made at the start, make the code more time/energy efficient.

2 July - 15 July: Start writeup, add literature review, theoretical discussion. Set up and run (or re-run) all required experiments and add results we have so far.

16 July - 30 July: Finish up all remaining experiments, finishing first draft of the dissertation.

31 July - 9 August: Finalising all results, finalising writeup, submitting.

Resource Declaration

As the project's fundamental tool will be a convolutional neural network (CNN), **access to a GPU / GPUs** will be crucial for the success of the project as it will enable quick experimentation.

Currently, it does not seem that the project would involve computation that would heavily benefit from parallelisation on several CPUs hence it is currently unlikely that we would make use of the MLMI grid engine.

No studies involving human participants is planned for this project.

References

[1] Marton Havasi, Robert Peharz, and José Miguel Hernández-Lobato. "Minimal Random Code Learning: Getting Bits Back from Compressed Model Parameters". In: *arXiv preprint arXiv:1810.00440* (2018).

[2] Irina Higgins et al. "beta-vae: Learning basic visual concepts with a constrained variational framework". In: *International Conference on Learning Representations*. 2017.

[3] Geoffrey Hinton and Drew Van Camp. "Keeping neural networks simple by minimizing the description length of the weights". In: *in Proc. of the 6th Ann. ACM Conf. on Computational Learning Theory*. Citeseer. 1993.

[4] Quan Huynh-Thu and Mohammed Ghanbari. "Scope of validity of PSNR in image/video quality assessment". In: *Electronics letters* 44.13 (2008), pp. 800–801.

[5] Diederik P Kingma and Max Welling. “Auto-Encoding Variational Bayes”. In: *arXiv preprint arXiv:1312.6114* (2013).

[6] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. “Multiscale structural similarity for image quality assessment”. In: *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*. Vol. 2. Ieee. 2003, pp. 1398–1402.

[7] *Workshop and Challenge on Learned Image Compression*. `https://www.compression.cc`. Accessed: 2019-03-25.