SCUBA: Deep Lossless Compression

Samuel Tenka, Rajith Weerasinghe, Aaron Zhao

MSAIL

To Compress, Predict!

We begin with the result of Shannon that the smallest expected compressed size for a sequence of potentially dependent variables is the sum of their conditional entropies [4]. Thus, good compression reduces to prediction: if one can predict the right-hand side of a selfie given the left-hand side, then one can compress that selfie by 2x. Classical compression algorithms such as PNG rely on hand-crafted image priors for prediction. Our work demonstrates that **learned distributions can match or improve lossless image compression ratios**.

Inferred Factorization

We propose a new probabilistic model for predicting pixels. The core idea is to learn to infer an approximate dependency graph between image components.

If $x=(x_t:0\leq t<n)$'s components are ordered, we may view x as sampled in sequence:

  for $0 \leq t < n$:
    sample $x_t \mid (x_s:s<t)$

Hence comes the standard factorization

  $p(x)=p(x_0|x_1...x_{n-1})...p(x_{n-2}|x_{n-1})p(x_{n-1})$

To model x's density, we model each factor above. The natural order often correponds to an approximate Markov property that aids in modeling each factor.

For x not naturally ordered, we propose to permute the components to advantage. The permutation i will be learned, data-dependent, and inductively deducible:

  for $0 \leq t < n$:
    sample $i(t) \mid ((i(s), x_{i(s)}):s<t)$
    sample $x_{i(t)} \mid (x_{i(s)}:s<t)$

When used generatively, this technique of **Inferred Factorization** is akin to DRAW of [2].

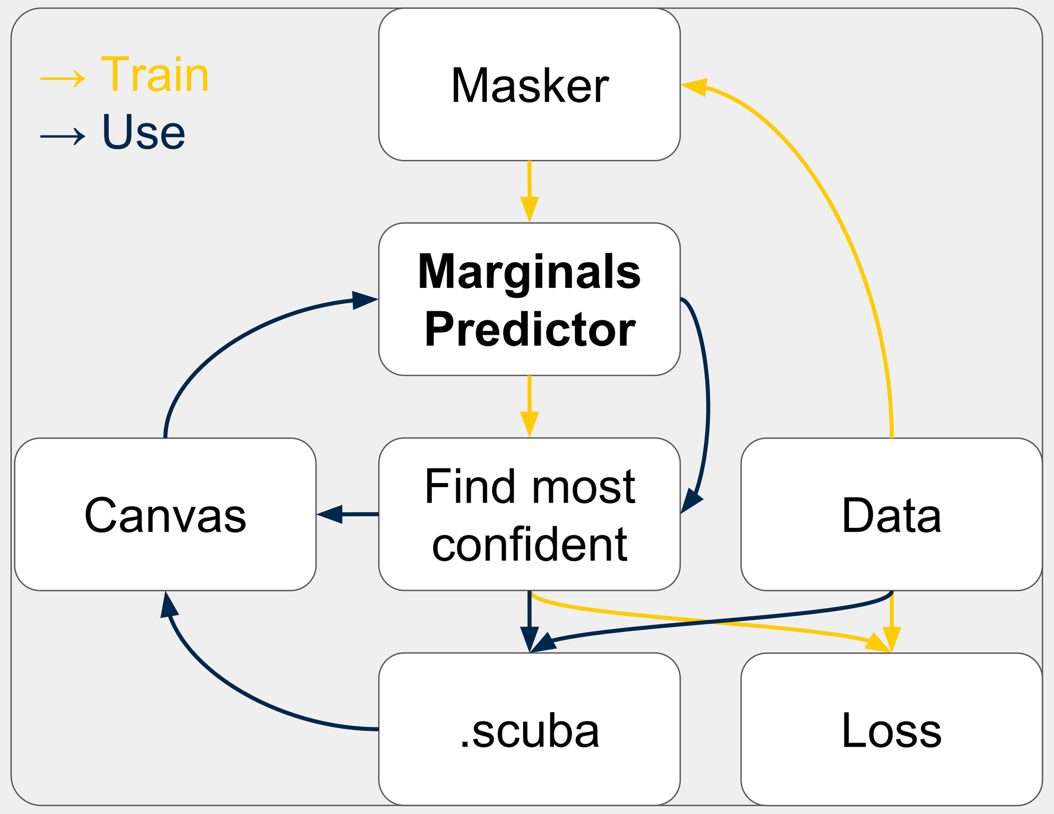
The SCUBA Framework

We propose a novel framework for lossless compression:
***S**hannon-**C**oding via **U**ncertainty-**B**ased **A**daptation*

$$i^\star = \underset{i:u_i=1}{\text{argmax}} \int p_i(\chi)\log(p_i(\chi))d\chi$$

$$\mathcal{L} = -\log(p_{i^\star}(x_{i^\star})) - \gamma \underset{i:u_i=1}{\text{average}} \log(p_i(x_i))$$

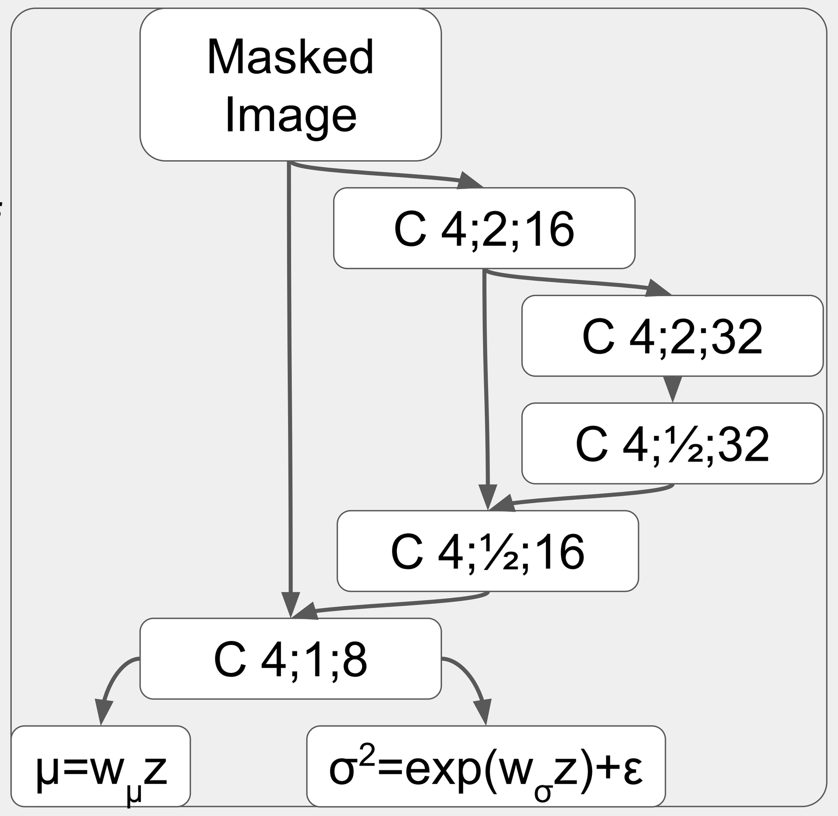
The corresponding generative algorithm iteratively samples from the unknown cell of greatest negentropy. The parameter γ is a unitless regularizer. We set γ=1 throughout.



SCUBA on Images

We use a Convolutional U-Net for predicting Gaussian Marginals of unknown pixel-channel values given a partially known image. Such networks achieve accurate image-to-image translation at local and global scales [1]. We set ε=1 to avoid singularities. We also found that averaging the predicted Gaussians with a weight-(1/16) uniform distribution leads to greater robustness to unfamiliar test distributions.

Here, "C k;s;c" indicates a convolutional layer with kernel kxk, stride s, and c output channels, composed with a Leaky ReLU activation, leak=0.1.

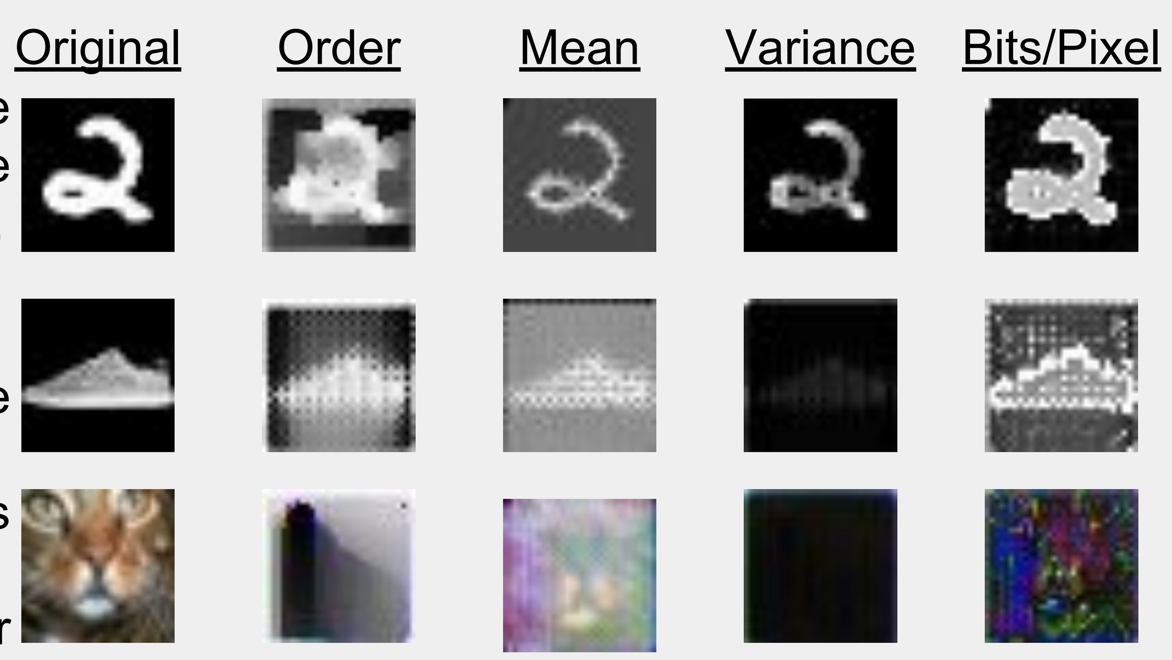


SCUBA Visualized

To the left are SCUBA outputs on three testing images. The bitrate (that is, bits per pixel) peaks at edges, supporting the intuition that SCUBA is able compress constant regions well.

SCUBA's inferred factorizations are diverse:

- for the digit '2', it sweeps through the background before traversing the foreground
- for the shoe photograph, it fills in left and right borders before adopting an interlacing scheme
- for the cat image, it sweeps through the lower triangle for reasons we do not yet understand



Results

SCUBA performs comparably to PNG, the de facto standard in lossless image compression, on both content-specialized datasets such as MNIST and diverse datasets such as CIFAR-10:

Bits/Pixel	bitmap (ideal)	PNG w/o header	SCUBA
MNIST	8.00	3.03	**2.10**
FASHION [5]	8.00	**5.91**	5.98
CIFAR-10 [3]	24.00	**21.39**	22.47

Conclusion and Future Steps

This preliminary work demonstrates the potential for lossless compression through sophisticated prediction techniques. Though our current models perform at the level of current methods, the former represent just the beginning of a promising family of compression algorithms. For instance, we have so far tuned no architecture hyperparameters, and we expect significant gains from tuning network depth, width, and regularizers.

The generality of Inferred Factorization lends itself to other domains such as audio, video, natural language, and structured texts such as HTML; the latter seems especially liable to an inferred factorization approach due to the presence of long-range grammatical dependencies. Thus, we plan to adapt SCUBA to these other domains.

More broadly, we ask:
Alternate Representations: can index permutation generalize to other invertible maps?
Lossy compression: can learned image priors advance the accuracy-size trade-off in *lossy* image compression? We propose to use WGANs [0] to learn human-correlated image metrics.

References

[0] Arjovsky et al. *Wasserstein GAN* (2017)
[1] Isola et al. *Image-to-Image Translation with Conditional Adversarial Networks* (2017)
[2] Gregor et al. *DRAW: A Recurrent Neural Network For Image Generation* (2015)
[3] Krizhevsky. *Learning Multiple Layers of Features from Tiny Images* (2009)
[4] Shannon. *Prediction and Entropy of Printed English* (1951)
[5] Xiao et al. *Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms* (2017)

Sponsors

The Michigan Student Artificial Intelligence Lab (**MSAIL**.github.io) is a reading group on Machine Learning. MSAIL is sponsored by the Michigan Institute for Data Science (**MIDAS**.umich.edu) and mentored by Professor **Laura Balzano**, Professor **Danai Koutra**, and Professor **Jenna Wiens**. The authors thank MIDAS and Professors Balzano, Koutra, and Wiens for providing advice and computational tools for this work.