Structured Uncertainty in Generative Imaging Models

Neill Campbell

Joint work with Era Dorta, Margaret Duff, Teo Deveney, Ivor Simpson, Sara Vicente, Lourdes Agapito, Matthias Ehrhardt, Tony Shardlow, and Carl Henrik Ek

Department of Computer Science, University College London and University of Bath

Storytime...(Priors and No Free Lunches...)

“breaking the ubiquitous ML assumption in image and vision computing that errors and uncertainties at neighbouring pixels are independent, despite their demonstrable spatial structure”

Is unsupervised learning a thing?

Overview...

Is unsupervised learning a thing?

Generative models

Structured Uncertainty Prediction Networks (SUPN)

SUPN as a prior for inverse problems

Non-Gaussian likelihoods

Thanks!

Figure 1: Stable Diffusion: *"The manifold of cats."*

Unsupervised learning → generative models

Figure 1: Stable Diffusion: *"The manifold of cats."*

- "Find me some $p(\mathbf{z})$ and $f(\mathbf{z})$ such that $\mathbf{x} \sim f(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.."

Figure 1: Stable Diffusion: *"The manifold of cats."*

- "Find me some $p(\mathbf{z})$ and $f(\mathbf{z})$ such that $\mathbf{x} \sim f(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.."
- This has trivial solutions

Figure 1: Stable Diffusion: *"The manifold of cats."*

- "Find me some $p(\mathbf{z})$ and $f(\mathbf{z})$ such that $\mathbf{x} \sim f(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.."
- This has trivial solutions
  - Need constraints

Unsupervised learning $\rightarrow$ generative models

Figure 1: Stable Diffusion: *"The manifold of cats."*

- "Find me some $p(\mathbf{z})$ and $f(\mathbf{z})$ such that $\mathbf{x} \sim f(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.."
- This has trivial solutions
  - Need constraints
  - Utility $\leftrightarrow$ use-case

Unsupervised learning $\to$ generative models

Figure 1: Stable Diffusion: *"The manifold of cats."*

- "Find me some $p(\mathbf{z})$ and $f(\mathbf{z})$ such that $\mathbf{x} \sim f(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.."
- This has trivial solutions
    - Need constraints
    - Utility $\leftrightarrow$ use-case
- Generative models as priors

Inverse problem setup

- Inverse problem $\mathbf{y} = A\,\mathbf{x} + \varepsilon$ for some forward model $A : \mathcal{X} \to \mathcal{Y}$ and noise ε

Inverse problem setup

- Inverse problem $\mathbf{y} = A\,\mathbf{x} + \varepsilon$ for some forward model $A : \mathcal{X} \to \mathcal{Y}$ and noise ε
- *Variational* regularisation framework (for some similarity $D(\cdot,\cdot)$)

$$\mathbf{x}^* \in \arg\min_{\mathbf{x}\in\mathcal{X}} D(\mathbf{y}, A\,\mathbf{x}) + \lambda\,R(\mathbf{x})$$

Inverse problem setup

- Inverse problem $\mathbf{y} = A\,\mathbf{x} + \varepsilon$ for some forward model $A : \mathcal{X} \to \mathcal{Y}$ and noise ε
- *Variational* regularisation framework (for some similarity $D(\cdot, \cdot)$)

$$\mathbf{x}^* \in \arg\min_{\mathbf{x} \in \mathcal{X}} D(\mathbf{y}, A\,\mathbf{x}) + \lambda\, R(\mathbf{x})$$

- Regulariser from an *explicit* prior distribution, $R(\mathbf{x}) := \log p(\mathbf{x} \,|\, \boldsymbol{\theta})$

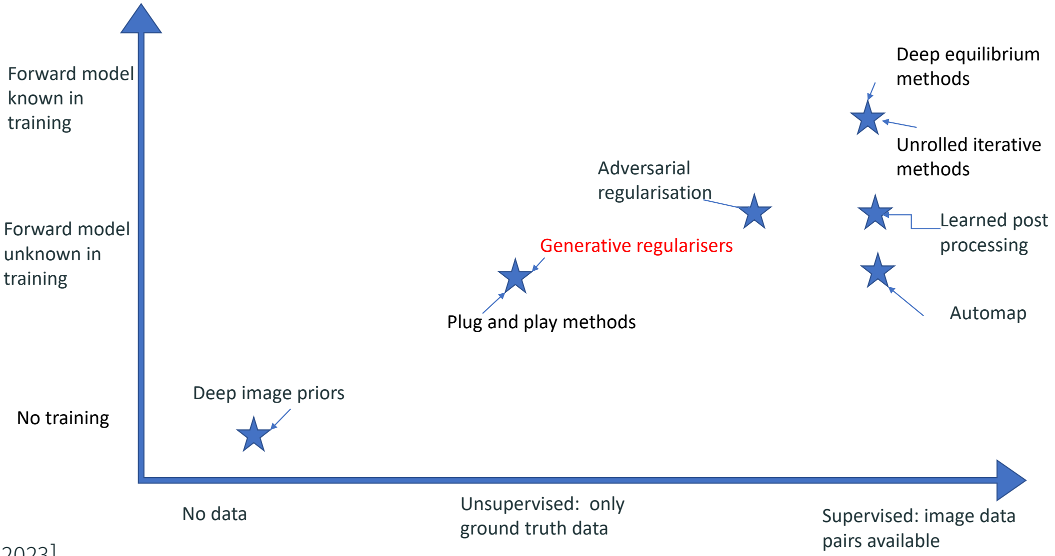
Inverse problem setup

- Inverse problem $\mathbf{y} = A\,\mathbf{x} + \varepsilon$ for some forward model $A : \mathcal{X} \to \mathcal{Y}$ and noise ε
- *Variational* regularisation framework (for some similarity $D(\cdot,\cdot)$)

$$\mathbf{x}^* \in \arg\min_{\mathbf{x}\in\mathcal{X}} D(\mathbf{y}, A\,\mathbf{x}) + \lambda\, R(\mathbf{x})$$

- Regulariser from an *explicit* prior distribution, $R(\mathbf{x}) := \log p(\mathbf{x}\,|\,\boldsymbol{\theta})$
- $\mathbf{x}^*$ considered a MAP estimate if $D(\mathbf{y}, A\,\mathbf{x}) := \log p(\mathbf{y}\,|\,f(A\,\mathbf{x}), \ldots)$

Deep learning approaches for inverse problems



[Duff 2023]

Generative models

Overview...

Is unsupervised learning a thing?

Generative models

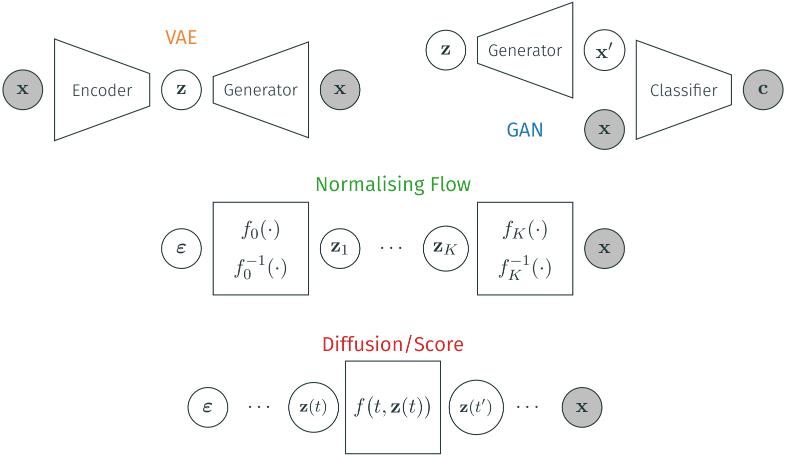
Structured Uncertainty Prediction Networks (SUPN)

SUPN as a prior for inverse problems

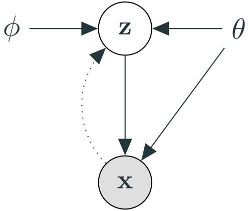
Non-Gaussian likelihoods

Thanks!

Generative model zoo



Unreasonable expectations of generative models?



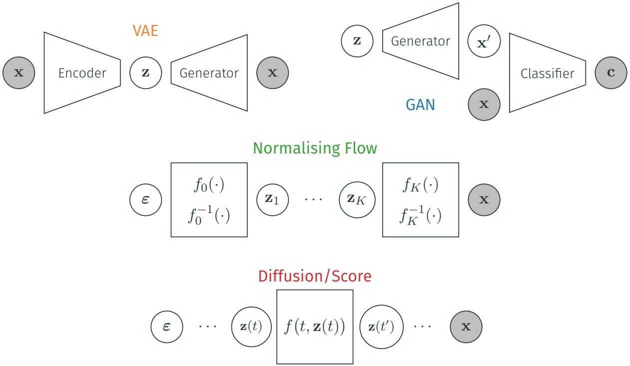
e.g. VAE with:

$$\mathbf{z} \in \mathbb{R}^M,$$
$$\mathbf{x} \in [0, 1]^{3 \times N \times N}$$

Figure 2: How many degrees of freedom are there in the image?

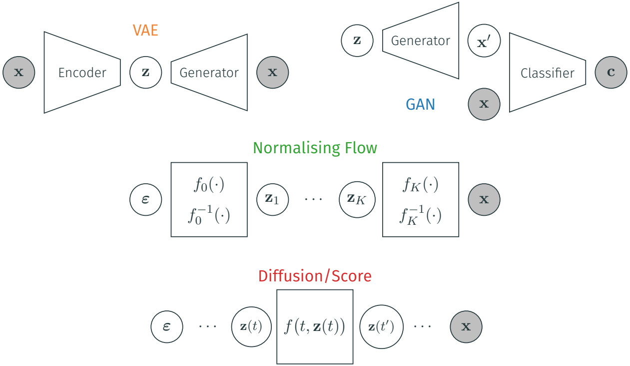
Properties we would like

- Span the data space



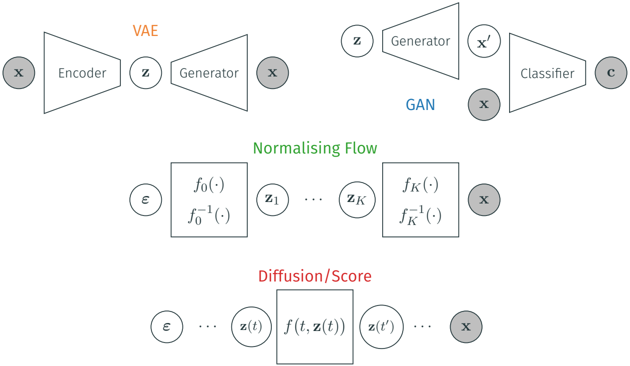
Properties we would like

- Span the data space
- Representative samples



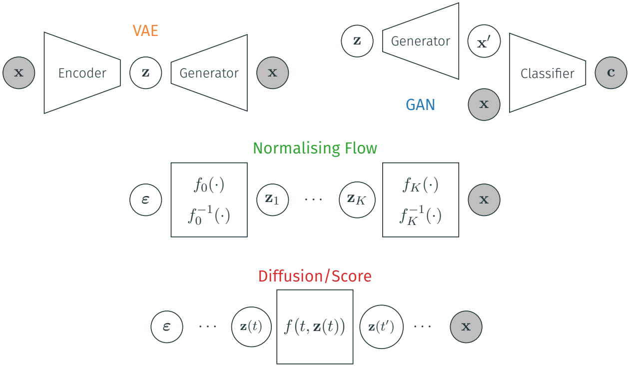
Properties we would like

- Span the data space
- Representative samples
- Conditions on mapping (e.g. "smooth")



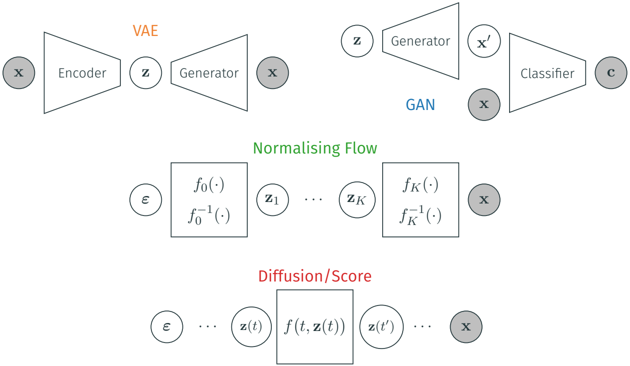
Properties we would like

- Span the data space
- Representative samples
- Conditions on mapping (e.g. "smooth")
- Evaluate densities (e.g. take likelihood)



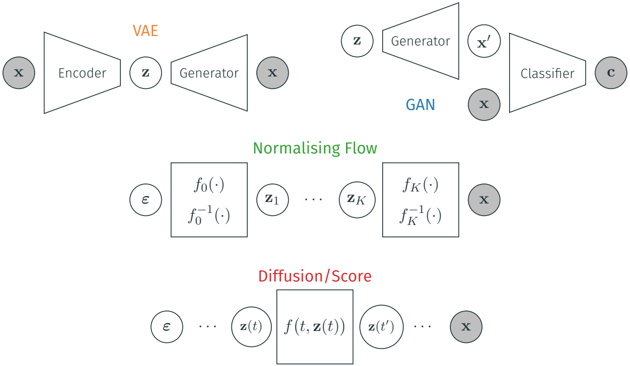
Properties we would like

- Span the data space
- Representative samples
- Conditions on mapping (e.g. "smooth")
- Evaluate densities (e.g. take likelihood)
- Uncertainty (e.g. account for failure to model)



Properties we would like

- Span the data space
- Representative samples
- Conditions on mapping (e.g. “smooth”)
- Evaluate densities (e.g. take likelihood)
- Uncertainty (e.g. account for failure to model)
- Introspection



Structured Uncertainty Prediction Networks (SUPN)

Overview...

Is unsupervised learning a thing?

Generative models

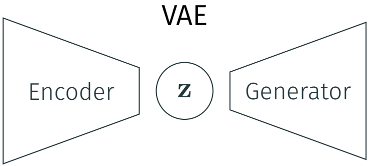
Structured Uncertainty Prediction Networks (SUPN)

SUPN as a prior for inverse problems

Non-Gaussian likelihoods

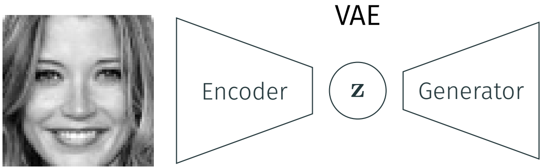
Thanks!

“VAEs produce overly smooth output”



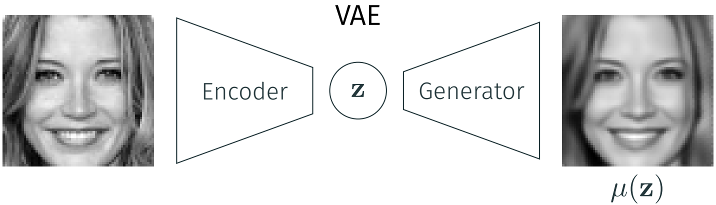
[Dorta et al. 2018]

“VAEs produce overly smooth output”



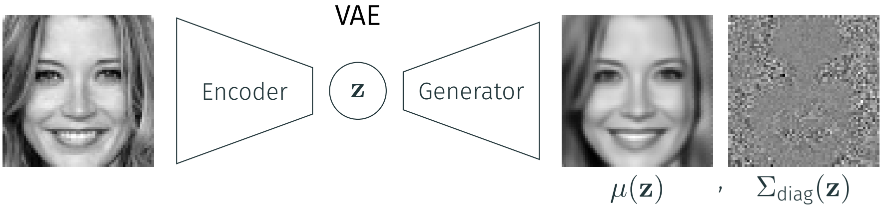
[Dorta et al. 2018]

“VAEs produce overly smooth output”



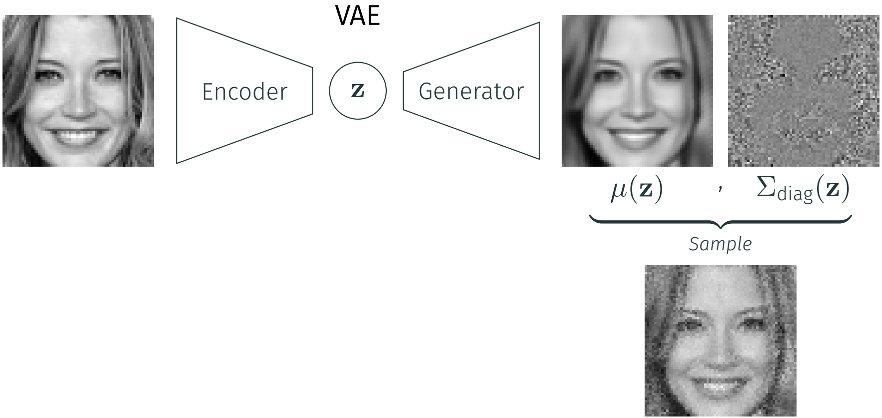
[Dorta et al. 2018]

“VAEs produce overly smooth output”



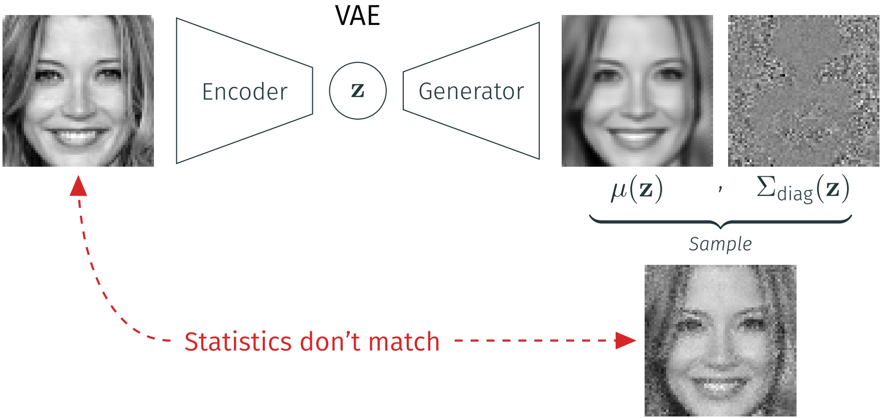
[Dorta et al. 2018]

“VAEs produce overly smooth output”



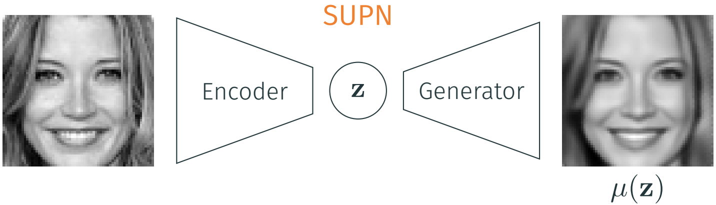
[Dorta et al. 2018]

“VAEs produce overly smooth output”



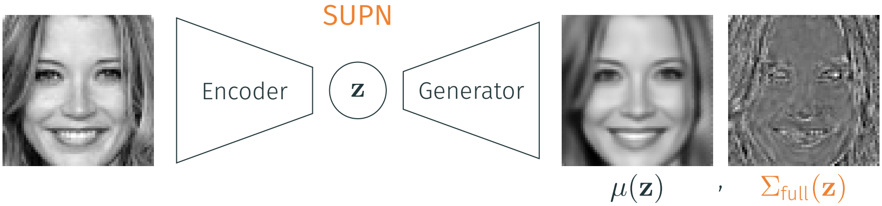
[Dorta et al. 2018]

“VAEs produce overly smooth output”



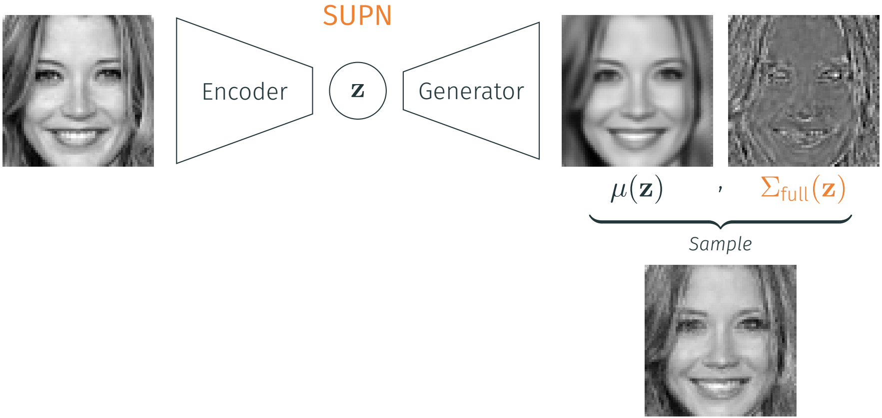
[Dorta et al. 2018]

“VAEs produce overly smooth output”



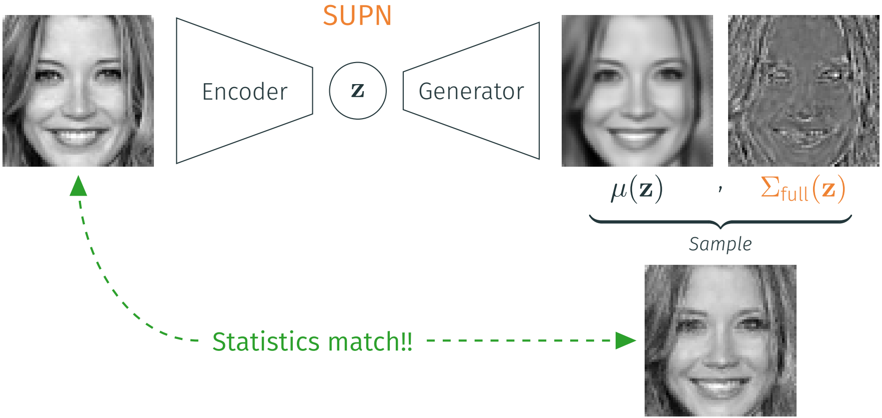
[Dorta et al. 2018]

“VAEs produce overly smooth output”



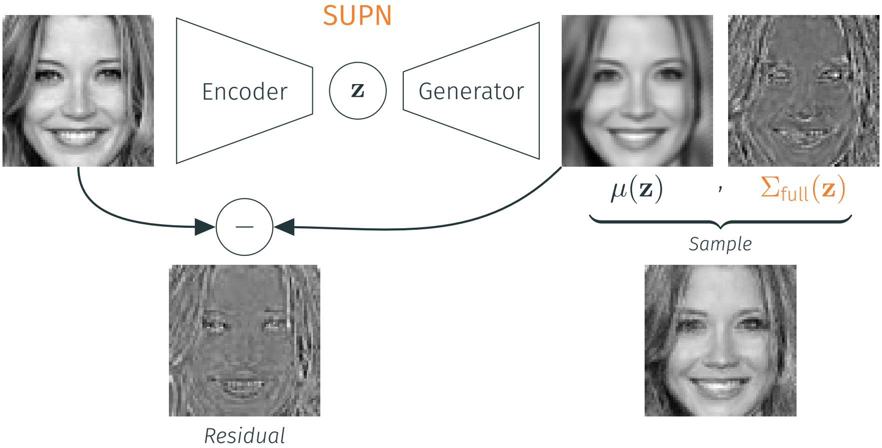
[Dorta et al. 2018]

“VAEs produce overly smooth output”



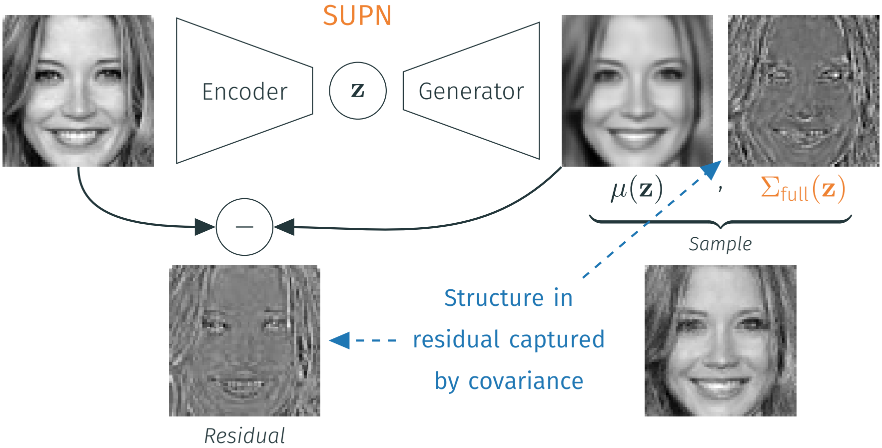
[Dorta et al. 2018]

“VAEs produce overly smooth output”



[Dorta et al. 2018]

"VAEs produce overly smooth output"



[Dorta et al. 2018]

Problem! Dense covariance $\mathcal{O}(N^2)$...

- **Problem**: $\Sigma_{\text{full}}(\mathbf{z})$ is quadratic in the number of pixels

Problem! Dense covariance $\mathcal{O}(N^2)$...

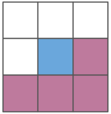
- **Problem**: $\Sigma_{\text{full}}(\mathbf{z})$ is quadratic in the number of pixels
- **Solution**: Sparse parameterisation of the Cholesky factor of the precision

$$\Sigma(\mathbf{z}) := \left[\Lambda(\mathbf{z})\right]^{-1} := \left[L_\Lambda(\mathbf{z})\, L_\Lambda^\top(\mathbf{z})\right]^{-1}$$

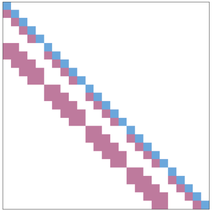
Problem! Dense covariance $\mathcal{O}(N^2)$...

- **Problem**: $\Sigma_{\text{full}}(\mathbf{z})$ is quadratic in the number of pixels
- **Solution**: Sparse parameterisation of the Cholesky factor of the precision

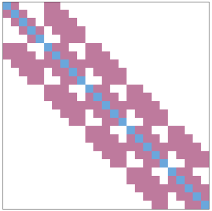
$$\Sigma(\mathbf{z}) := [\Lambda(\mathbf{z})]^{-1} := [L_\Lambda(\mathbf{z})\, L_\Lambda^\top(\mathbf{z})]^{-1}$$



Neighbourhood in image domain



Sparsity in the precision Cholesky matrix L_Λ



Sparsity in the precision matrix $\Lambda(\mathbf{z}) := \Sigma^{-1}(\mathbf{z})$

- Sparse parameterisation of the Cholesky factor of the precision

$$\Sigma(\mathbf{z}) := [\Lambda(\mathbf{z})]^{-1} := [L_\Lambda(\mathbf{z})\, L_\Lambda^\top(\mathbf{z})]^{-1}$$

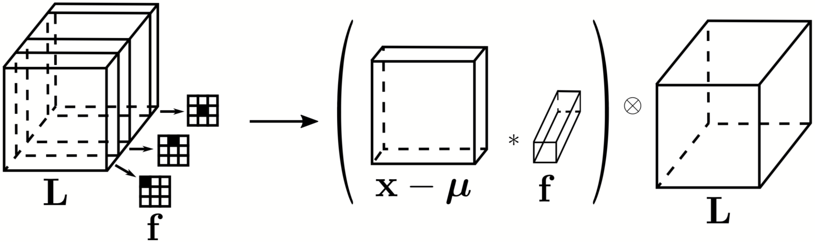


Figure 3: Implementation through convolutional structure: matrix-vector product in $\mathcal{O}(N)$

Examples of samples

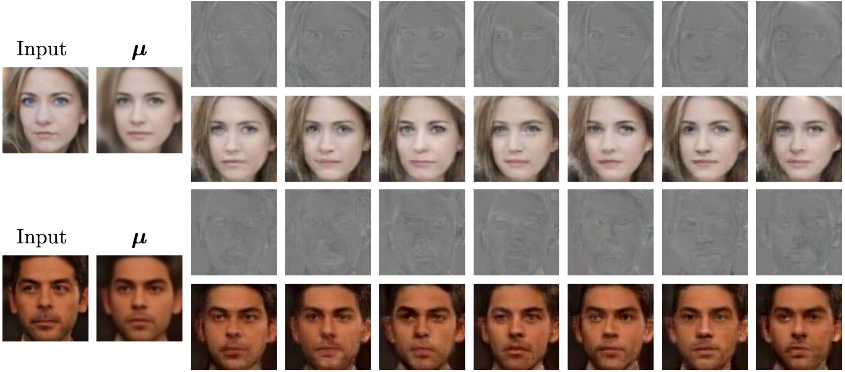


Figure 4: Variation in samples from the model on test data

Introspection of the captured covariance structure

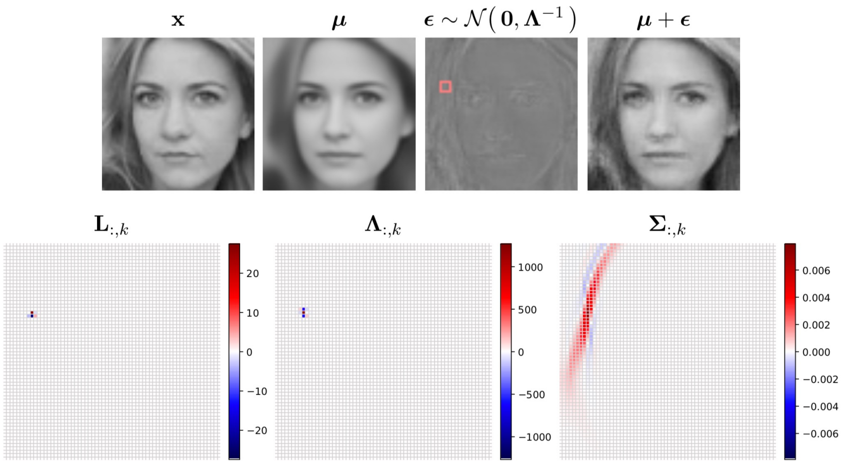


Figure 5: Visualisation of the learned correlations

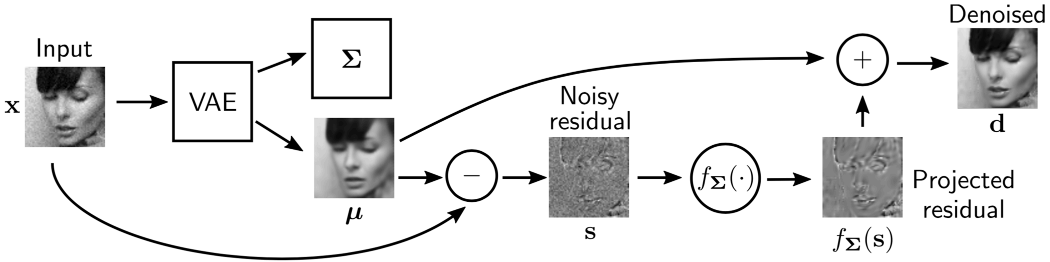
Links to established concepts...

- Links to Conditional Random Field (CRF) models
    - a Gaussian CRF - e.g. “Regression Tree Fields” [Jancsary et al. 2012]
- Links to adaptive local regularisation models
    - e.g. locally adaptive TV or Laplacian based methods
- Links to Wavelet approaches
    - considering hierarchical extensions or combining fixed basis functions

Links to established concepts…

- Links to Conditional Random Field (CRF) models
  - a Gaussian CRF - e.g. "Regression Tree Fields" [Jancsary et al. 2012]
- Links to adaptive local regularisation models
  - e.g. locally adaptive TV or Laplacian based methods
- Links to Wavelet approaches
  - considering hierarchical extensions or combining fixed basis functions
- Things to be careful about
  - priors on sparse precision (consider Cholesky structure)
  - need to bound terms
  - *lots to say about these things…*

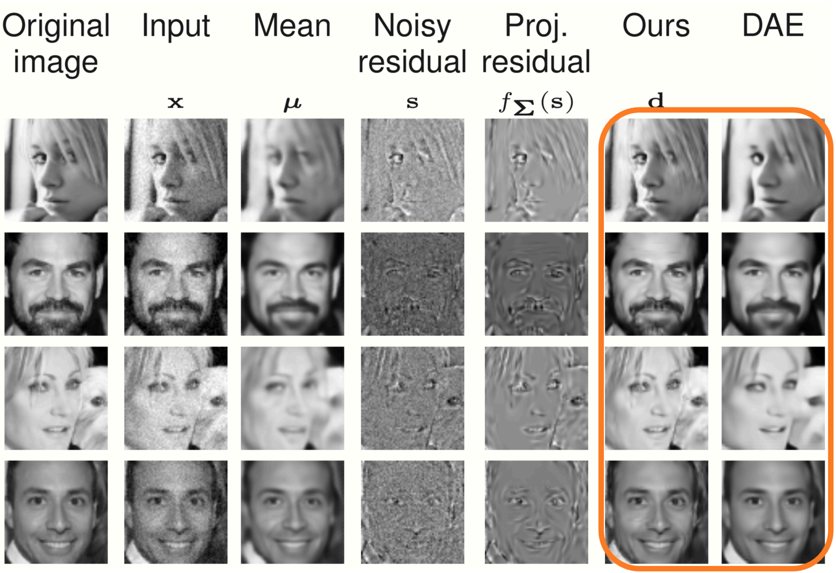
Testing with denoising...



Model	MSE	PSNR	SSIM
DAE	0.005 ± 0.003	28.89 ± 1.69	0.90 ± 0.03
SUPN	$\mathbf{0.003 \pm 0.001}$	$\mathbf{31.38 \pm 0.92}$	$\mathbf{0.92 \pm 0.02}$

Figure 6: Denoising example using SUPN (vs a denoising autoencoder). The SUPN model has only been trained as in a generative manner (i.e. as a prior).

Testing with denoising...



SUPN as a prior for inverse problems

Overview...

Is unsupervised learning a thing?

Generative models

Structured Uncertainty Prediction Networks (SUPN)

SUPN as a prior for inverse problems

Non-Gaussian likelihoods

Thanks!

SUPN as a prior for inverse problems

- Consider a hierarchical model for the inverse problem

$$p(\mathbf{x}, \mathbf{z} \,|\, \mathbf{y}) \propto p(\mathbf{y} \,|\, \mathbf{x})\, p_{\mathcal{G}}(\mathbf{x} \,|\, \mathbf{z})\, p_{\mathcal{Z}}(\mathbf{z})$$

- We will take a MAP estimate for $\mathbf{z}$ rather than marginalising `:-(`

SUPN as a prior for inverse problems

- Consider a hierarchical model for the inverse problem

$$p(\mathbf{x}, \mathbf{z} \,|\, \mathbf{y}) \propto p(\mathbf{y} \,|\, \mathbf{x})\, p_{\mathcal{G}}(\mathbf{x} \,|\, \mathbf{z})\, p_{\mathcal{Z}}(\mathbf{z})$$

- We will take a MAP estimate for $\mathbf{z}$ rather than marginalising `:-(`
- From before (with a Gaussian observation likelihood) and $p_{\mathcal{Z}}(\mathbf{z}) \sim \mathcal{N}(\mathbf{0}, I)$

$$D(\mathbf{y}, A\,\mathbf{x}) := \frac{1}{2\sigma^2}\|A\,\mathbf{x} - \mathbf{y}\|_2^2$$

$$R(\mathbf{x}) := \min_{\mathbf{z}\in\mathcal{Z}} \; \log|\Sigma_\theta(\mathbf{z})| + \frac{1}{2}\|\mathbf{x} - \mu_\theta(\mathbf{z})\|^2_{\Sigma_\theta(\mathbf{z})} + \frac{1}{2}\|\mathbf{z}\|_2^2$$

- Where the *Generator* provides $\mathcal{N}(\mathbf{x} \,|\, \mu_\theta(\mathbf{z}), \Sigma_\theta(\mathbf{z}))$ via a network $[\mu, L_\Lambda] = f(\mathbf{z}; \theta)$ and $\|\mathbf{a}\|^2_\Sigma := \mathbf{a}^\top\,\Sigma^{-1}\,\mathbf{a}$ denotes a Gaussian weighted norm

SUPN as a prior for inverse problems

- Consider a hierarchical model for the inverse problem

$$p(\mathbf{x}, \mathbf{z} \,|\, \mathbf{y}) \propto p(\mathbf{y} \,|\, \mathbf{x}) \, p_{\mathcal{G}}(\mathbf{x} \,|\, \mathbf{z}) \, p_{\mathcal{Z}}(\mathbf{z})$$

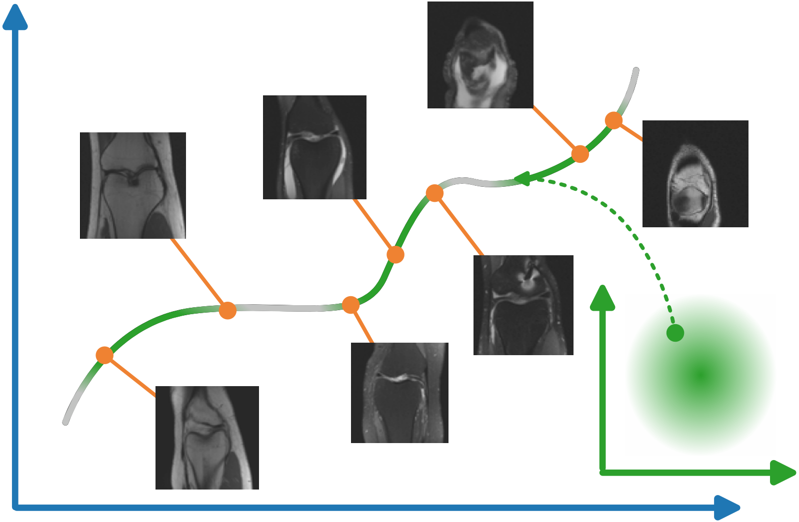
- We will take a MAP estimate for $\mathbf{z}$ rather than marginalising `:-(`
- From before (with a Gaussian observation likelihood) and $p_{\mathcal{Z}}(\mathbf{z}) \sim \mathcal{N}(\mathbf{0}, I)$

$$D(\mathbf{y}, A\,\mathbf{x}) := \frac{1}{2\sigma^2} \|A\,\mathbf{x} - \mathbf{y}\|_2^2$$

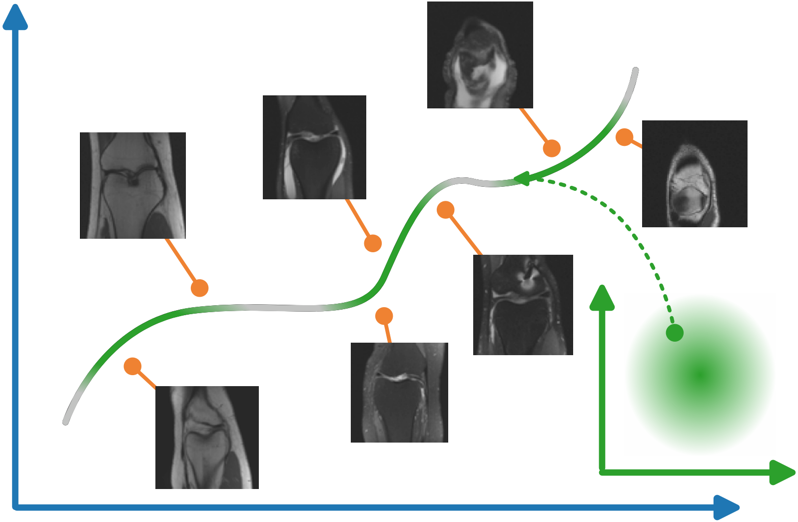
$$R(\mathbf{x}) := \min_{\mathbf{z} \in \mathcal{Z}} \; \log |\Sigma_\theta(\mathbf{z})| + \frac{1}{2} \|\mathbf{x} - \mu_\theta(\mathbf{z})\|_{\Sigma_\theta(\mathbf{z})}^2 + \frac{1}{2} \|\mathbf{z}\|_2^2$$

- Where the *Generator* provides $\mathcal{N}(\mathbf{x} \,|\, \mu_\theta(\mathbf{z}), \Sigma_\theta(\mathbf{z}))$ via a network $[\mu, L_\Lambda] = f(\mathbf{z}; \theta)$ and $\|\mathbf{a}\|_\Sigma^2 := \mathbf{a}^\top \, \Sigma^{-1} \, \mathbf{a}$ denotes a Gaussian weighted norm
- *Note: the network still outputs $\mathcal{O}(N)$ values and evaluation of $R(\mathbf{x})$ can be performed in $\mathcal{O}(N)$ time using L_Λ for the first two terms*

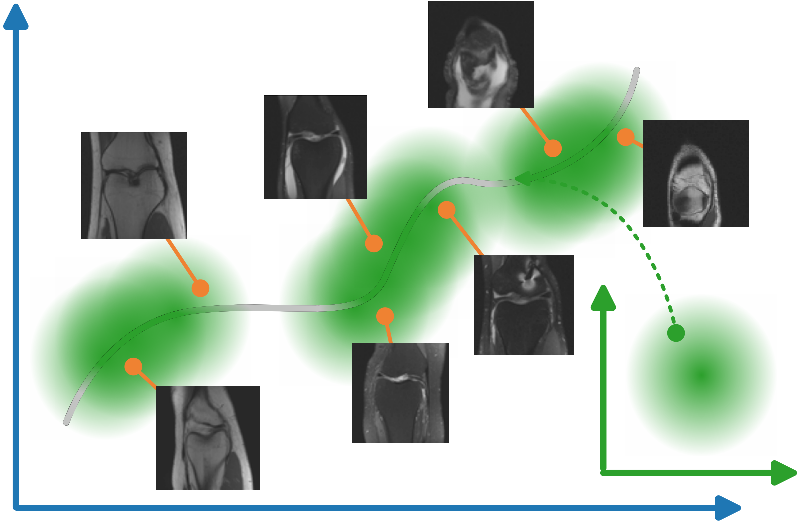
Aside: Images and manifolds



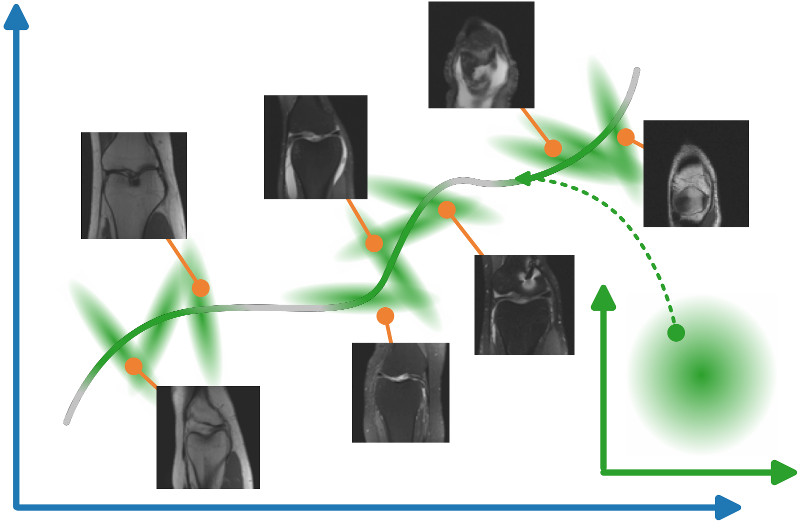
Aside: Images and manifolds



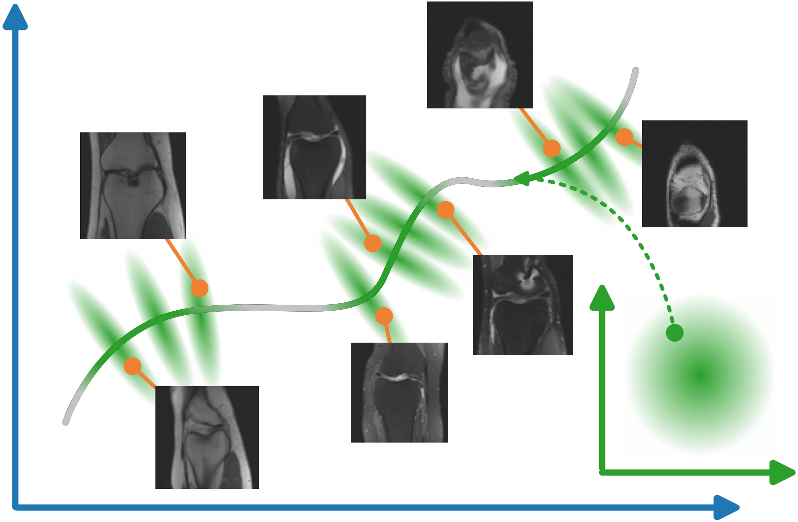
Aside: Images and manifolds



Aside: Images and manifolds



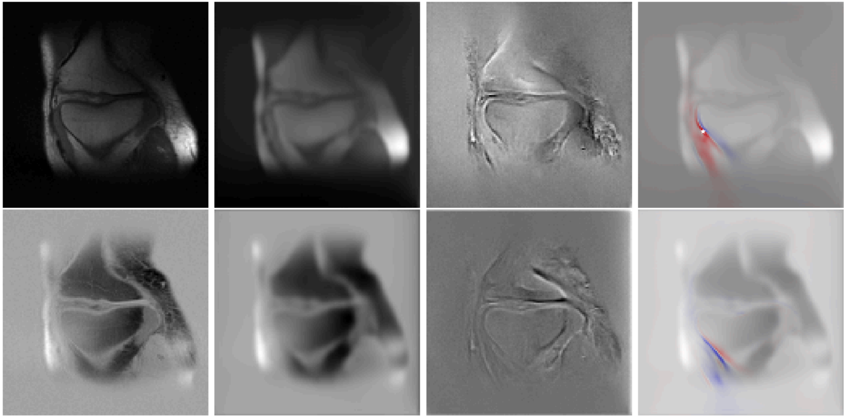
Aside: Images and manifolds



Proof of concept example: NYU fastMRI knee dataset

- Task inspired by the single-coil reconstruction
- Compressed sensing task (progressively reducing the amount of data measured)
- Initialise with $\mathbf{z}^{(0)}$ using the encoding of a rough reconstruction, given by the adjoint of the forward operator, and the corresponding mean output for $\mathbf{x}^{(0)}$
- Can also estimate the uncertainty (e.g. by sampling through)

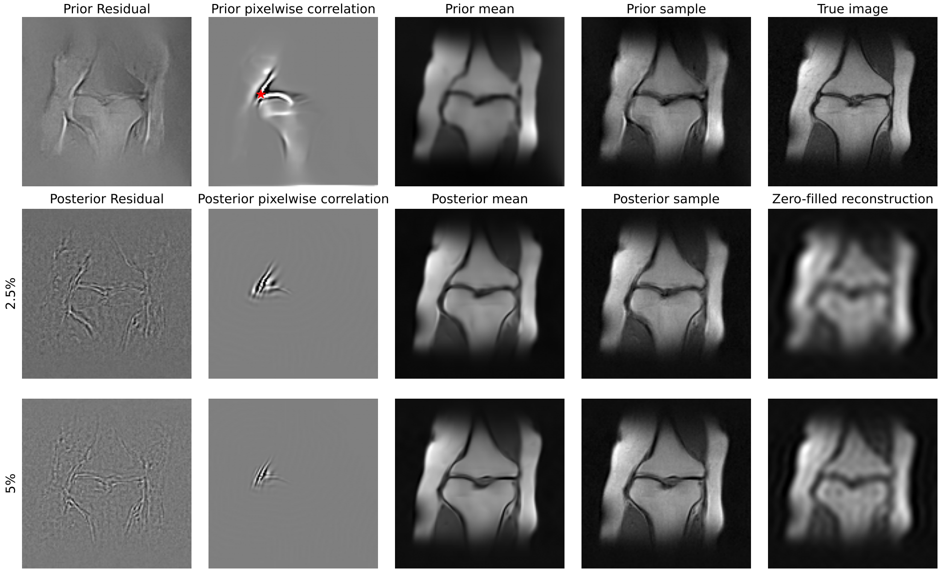
FastMRI knee learned prior covariance...



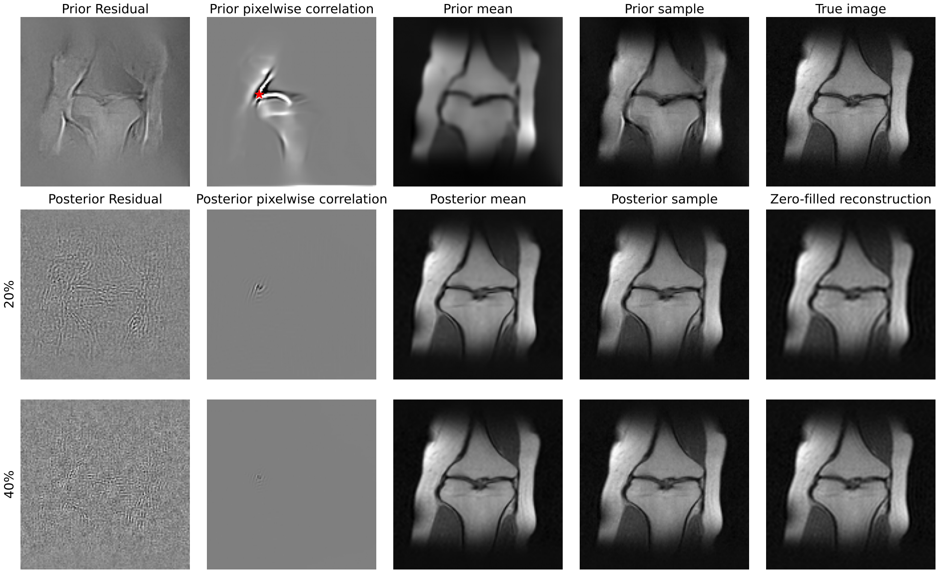
Real (top) and complex (bottom) channels of the learned prior.

Left to right: True Image, Mean, Prior Residual Sample, Pixelwise Correlations

Compressed sensing reconstruction results



Compressed sensing reconstruction results



Comparison vs supervised reconstruction method

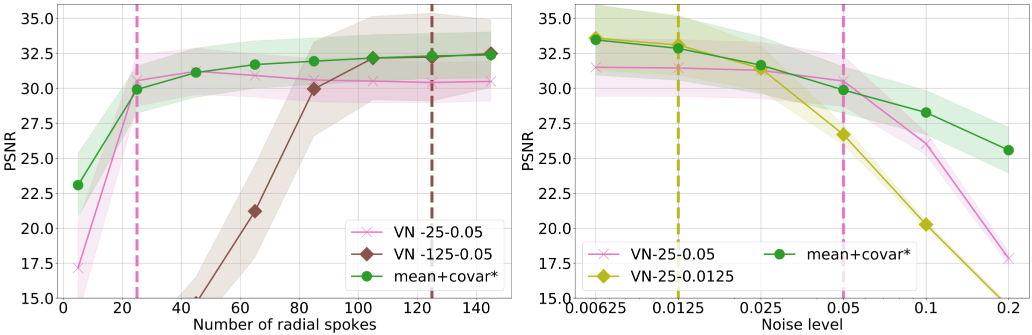
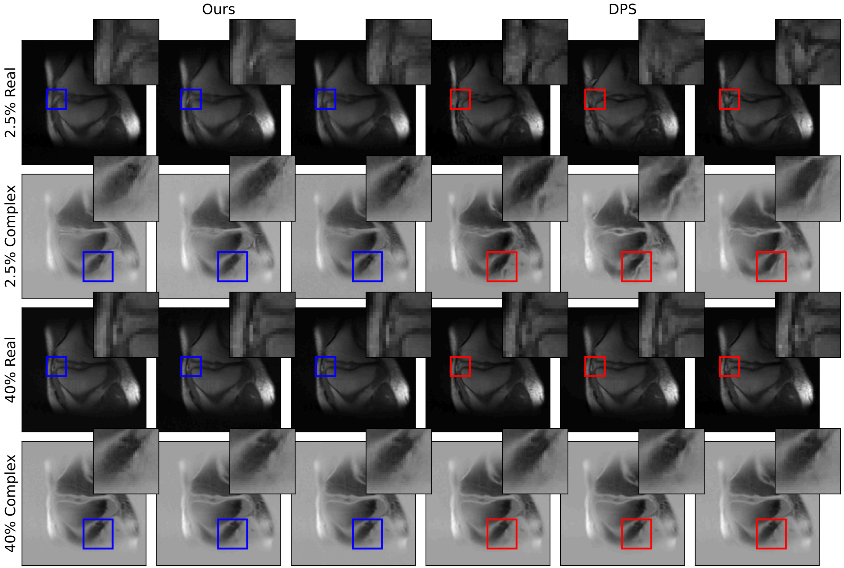


Figure 13: Comparison with the supervised variational networks [Hammernik et al. 2018]. The vertical lines depict the experimental settings the variational networks were trained on.

Reconstruction uncertainty: samples



Open challenges

- Nice introspection but what about dataset bias?
- Convergence rates (e.g. looking at natural gradients)
- Convexity/uniqueness
- Assumption that “ground truth” data available

3rd Workshop on Uncertainty Quantification for Computer Vision

ECCV 2024 Workshop

In the last decade, substantial progress has been made w.r.t. the performance of computer vision systems, a significant part of it thanks to deep learning. These advancements prompted sharp community growth and a rise in industrial investment. However, most current models lack the ability to reason about the confidence of their predictions; integrating uncertainty quantification into vision systems will help recognize failure scenarios and enable robust applications.

In addition to advances in Bayesian deep learning, providing practical approaches for vision problems, the workshop will provide a forum for discussing promising research directions, which have received less attention, as well as advancing current practices to drive future research. Examples include: the development of new metrics that reflect the real-world need for uncertainty when using vision systems with down-stream tasks; and moving beyond point-estimates to address the multi-modal ambiguities inherent in many vision tasks.

This years UNcertainty quantification for Computer Vision (UNCV) Workshop aims to raise awareness and generate discussion regarding how predictive uncertainty can, and should, be effectively incorporated into models within the vision community. The workshop will bring together experts from machine learning and computer vision to create a new generation of well-calibrated and effective methods that *know when they do not know*.

Non-Gaussian likelihoods

Overview...

- Is unsupervised learning a thing?
- Generative models
- Structured Uncertainty Prediction Networks (SUPN)
- SUPN as a prior for inverse problems
- Non-Gaussian likelihoods
- Thanks!

“Learning Structured Gaussians to Approximate Deep Ensembles”

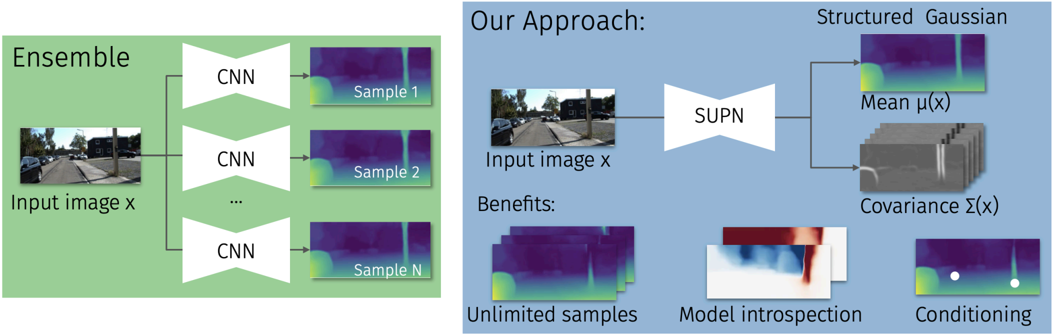


Figure 14: Use the structured Gaussian approach for “ensemble distillation”; approximate the output from a deep ensemble [Poggi et al. 2020, Lakshminarayanan et al. 2017]

[Simpson et al. 2022]

Non-Gaussian likelihood

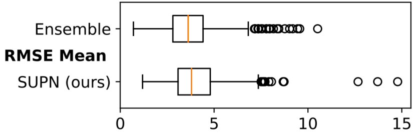
- Use a link function to change to different likelihood (e.g. a depth range through logits)
- Training from ensemble outputs using log-likelihood
- Output distribution captures epistemic and aleatoric uncertainty (via the ensemble)

Advantages

- Efficiency improvement
- Ability to draw unlimited samples
- Introspection
- Conditional sampling

Accuracy and uncertainty results

Accuracy Comparison:
The approximation captures the original ensemble well



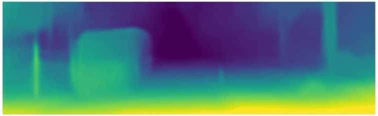
Uncertainty Metrics:
Pixelwise Area Under the Sparsification Error, Area Under the Random Gain and the Log-Likelihood

Model name	RMSE AUSE ↓	RMSE AURG ↑	LL $\times 10^5$ ↑
Ensemble [6]	2.927 (1.327)	0.324 (1.019)	
Diagonal	5.075 (1.924)	-1.697 (0.799)	1.77 (11.48)
SUPN	1.555 (1.307)	1.856 (1.355)	40.60 (1.35)

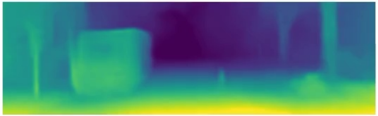
Figure 15: Monocular depth estimation results vs the original ensemble

Samples

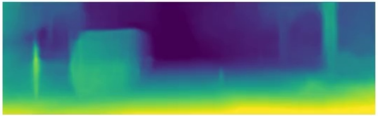
SUPN mean



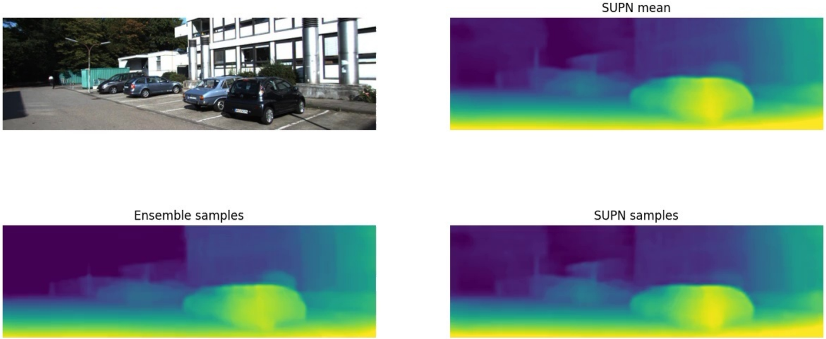
Ensemble samples



SUPN samples

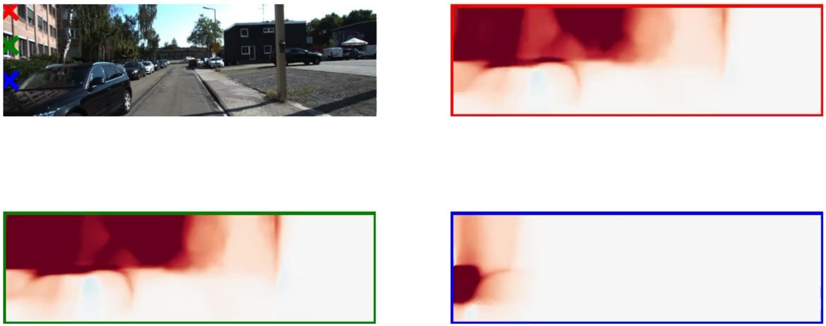


Samples



Introspection

Introspection



Conditional sampling

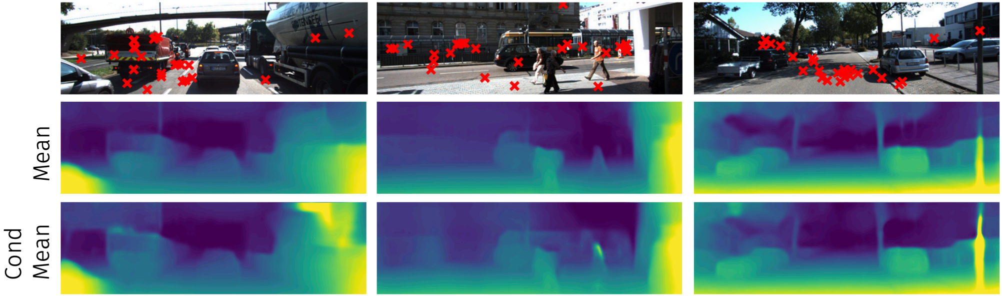


Figure 16: We can also perform *conditional* sampling using efficient sparse precision operations

$$p(\mathbf{d}_{\mathsf{U}} \,|\, \mathbf{d}_{\mathsf{K}} = \boldsymbol{\alpha}) \sim \mathcal{N}(\mathbf{b}, B)$$

$$\mathbf{b} := \boldsymbol{\mu}_{\mathsf{U}} - \Lambda_{\mathsf{UU}}^{-1}\,\Lambda_{\mathsf{UK}}\,(\boldsymbol{\alpha} - \boldsymbol{\mu}_{\mathsf{K}}), B := \Lambda_{\mathsf{UU}}^{-1}$$

Thanks!

Thanks! https://www.ndfcampbell.org neill.campbell@ucl.ac.uk

Joint work with Era Dorta, Margaret Duff, Teo Deveney, Ivor Simpson, Sara Vicente, Lourdes Agapito, Matthias Ehrhardt, Tony Shardlow, and Carl Henrik Ek. Thanks to the EPSRC CAMERA Research Centre, the CDE and SAMBa CDTs, and the Royal Society.

- **VAEs with Structured Image Covariance Applied to Compressed Sensing MRI,** *Margaret Duff, Ivor Simpson, Matthias J. Ehrhardt and Neill D. F. Campbell*, IoP Physics in Medicine and Biology, 2023
- **Regularising Inverse Problems with Generative Machine Learning Models,** *Margaret Duff, Neill D. F. Campbell and Matthias J. Ehrhardt*, Journal of Mathematical Imaging and Vision, 2023
- **Learning Structured Gaussians to Approximate Deep Ensembles,** *Ivor Simpson, Sara Vicente and Neill D. F. Campbell*, CVPR, 2022
- **Structured Uncertainty Prediction Networks,** *Era Dorta, Sara Vicente, Lourdes Agapito, Neill D. F. Campbell and Ivor Simpson*, CVPR, 2018