Session:

GO BEYOND



# Structural Material Property Tailoring of Dual Phase Titanium Alloy Microstructures Using Deep Neural Networks

**Ryan Noraas, Nagendra Somanath**
***Pratt and Whitney***
***E.Hartford CT-06108.***

**Michael Giering, Olusegun Oshin**
***United Technologies Research Center***
***E.Hartford CT-06108.***


JANUARY 10 2019

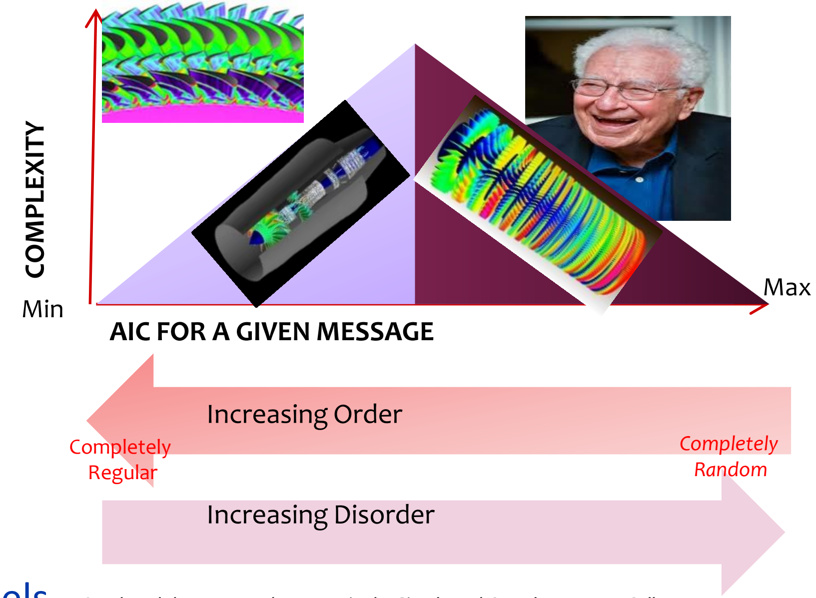
A UNITED TECHNOLOGIES COMPANY

# OVERVIEW

Understand, Train and inculcate a deep learning neural network to predict high cycle fatigue curves from m-structure images of structural alloy materials.

- Introduction
- Alloys, Manufacturing & Turbomachinery
- Mathematics of Deep Machine Learning
- Algorithms in Deep Learning
- Cutting Edge of Machine Learning
- Structural Material Tailoring
- Exploring Material Design Space using Generative Models
- Summary

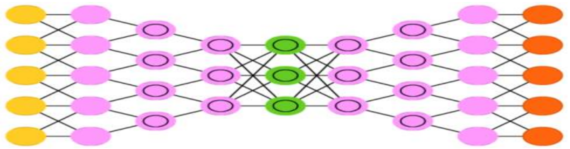
AIC – Algorithmic Information Content




Quark and the Jaguar-Adventures in the Simple and Complex, Murray Gell-Mann, 1995.

*"I believe that, one day, information will come to be viewed as being as fundamental as energy and matter."*
*-Demis Hassabis (DeepMind)*

Objective: Understand, Train and inculcate a deep learning neural network to predict high cycle fatigue curves from m-structure images of structural alloy materials

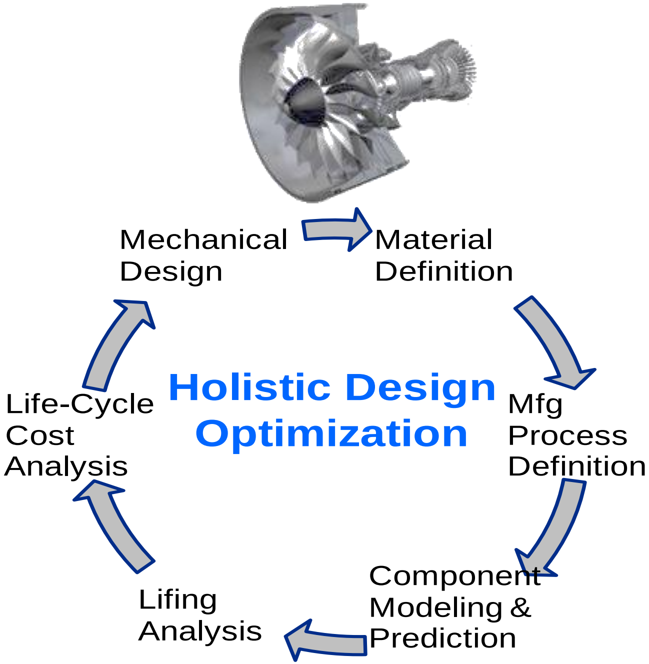


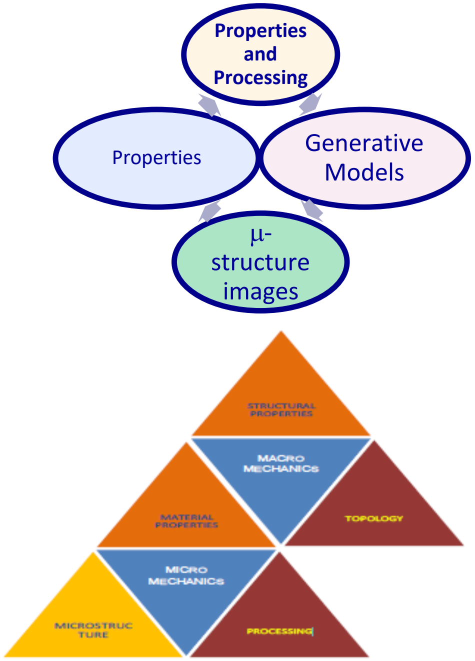
- HCF is one of the primary causes of failure in high speed turbomachinery applications.
- Fatigue (HCF/LCF) and Creep issues continue to lead to unpredictable failure in turbomachinery components due to crack propagation under complex multidisciplinary loading conditions.

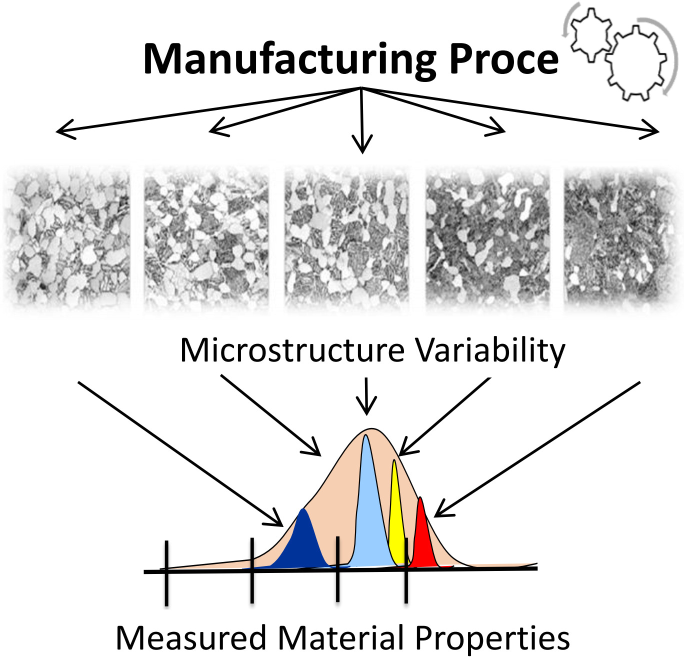

- There has been significant industrial effort in developing μ–structure-sensitive models for fatigue prediction.
- Alloy microstructure can vary significantly during thermal and mechanical processing –driving fatigue properties.
- High cycle fatigue of Ti-6Al-4V (various pedigrees)

# Structural Material Property Tailoring Using Deep Neural Networks

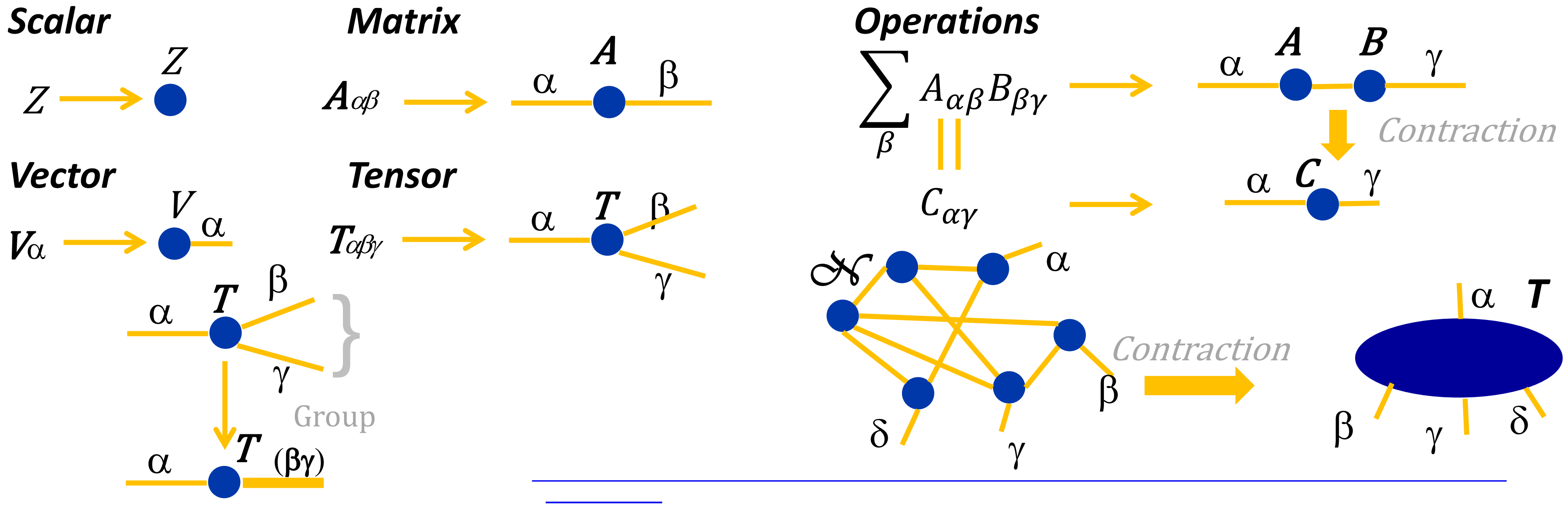
Modeling and optimization of material *μ–structure*, mechanical properties and manufacturing processes to enable *validated analysis* of alloy component behavior for ***design***, ***development***, ***production*** and ***use.***



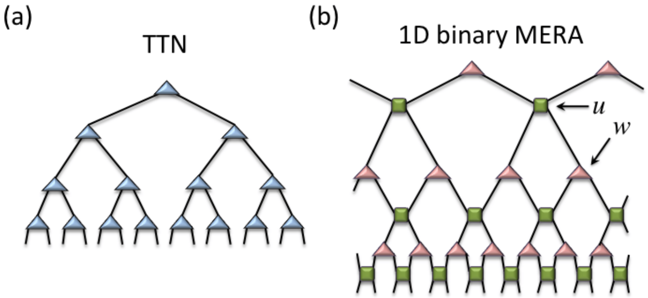



*The exceptional simplicity of physics-based functions hinges on properties such as symmetry, locality, compositionality and polynomial log-probability :- These properties translate to exceptionally simple neural networks approximating both natural phenomena such as images and abstract representations thereof!*



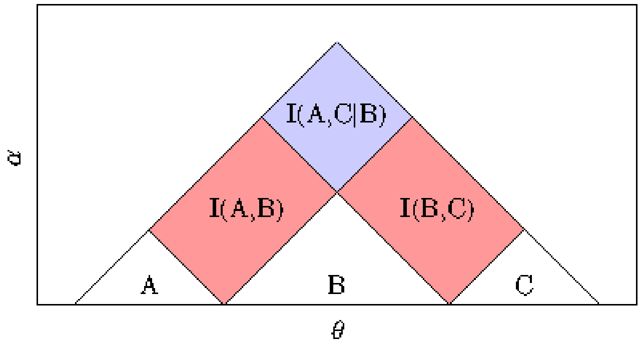

Neural network Learning proceeds through the *"Emergence of a 'thin manifold' of optimal Information"* in high-dimensional space that represents the natural phenomena that it is trained on. Mathematically there exists an *optimization mechanism* that reduces the *"degrees of freedom (or problem dimension)"* so that computation is feasible. *"Learning compression mechanism"* is found in the architecture of the DL network, (e.g. MERA tensor network).

## Mutual Information Chain Rule (MICR):




(a) Tree Tensor Network (TTN). (b) 1d1d binary MERA, Multi Scale Entanglement Renormalization Ansatz (MERA), with unitaries u and isommetries w.

In MERA -**circles** depict ***"disentanglers"***, -**triangles** ***"isometries".*** Mathematically the nodes of a mapping (i.e., *circles*) map ***matrices*** to other ***matrices***. *Triangles* take a ***matrix*** and map it to a ***vector***. The key though here is to realize that the 'compression' capability arises from the hierarchy and the entanglement.




- **MICR** : As *Information / Data* moves from the bottom to the top of the network, the information entanglement increases.
- Deep Learning (DL) networks are *tensor Information / Data* networks.
- As *Information / Data* flows from input to output in either a fully connected network or a convolution network, *Information / Data* are ***similarly*** entangled.
- "Holographic Principle" driven by quantum computation, reveals the capability of representing high dimensional problems using a relatively low number of model parameters.

*The key is dimensionality. Problems which are difficult to solve in low dimensional spaces become easier when "lifted" into a higher dimensional space. Think how much easier your day would be if you could move freely in the extra dimension we call time. Data points intertwined in their native, low-dimensional state can become linearly separable when given the extra breathing room of more dimensions.*

# Chaldni Sand Patterns - Modal Shapes Reflect System Free Energy Optima

On the Figures Obtained by Strewing Sand on Vibrating Surfaces, Commonly Called Acoustic Figures Charles Wheatstone , Philosophical Transactions of the Royal Society of London, Vol. 123 (1833), pp. 593-633



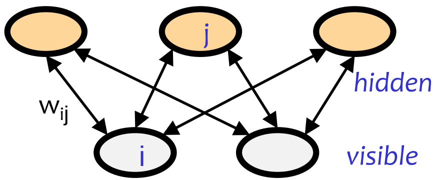
*The maximum likelihood principle can equivalently be viewed as selecting parameters to minimize the Kullback-Leibler divergence.*

$$KL(P_e||P_\theta) = \sum_{x \in X} P_e(x)\left(\log P_e(x) - \log P_\theta(x)\right)$$

## Restricted Boltzmann Machines

$$E_\theta = -(x^T W h + x^T b + h^T c)$$

$$P_\theta(x,h) = {}^{1}/_{Z}\, e^{-(E_\theta(x,h))}$$




Inductive Principles for Restricted Boltzmann Machine Learning', Benjamin M. Marlin, Kevin Swersky, Bo Chen and Nando de Freitas, Journal of Machine Learning Research – Proceedings Track for Artificial Intelligence and Statistics (AISTATS), pp. 509–516, Vol: 9, 2010.

## Contrastive Divergence

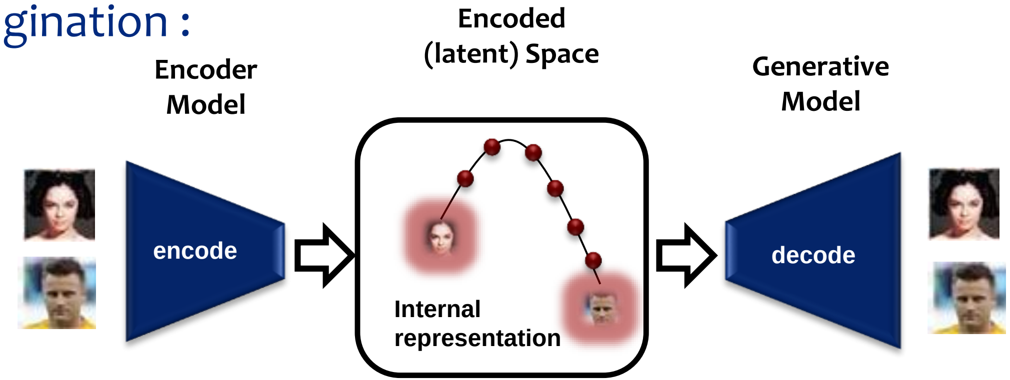
$$f^{CD}(\theta) = \sum_{x \in X} P_e(x) \log\left(\frac{P_e(x)}{P_\theta(x)}\right) - Q_\theta(x) \log\left(\frac{Q_\theta(x)}{P_\theta(x)}\right)$$

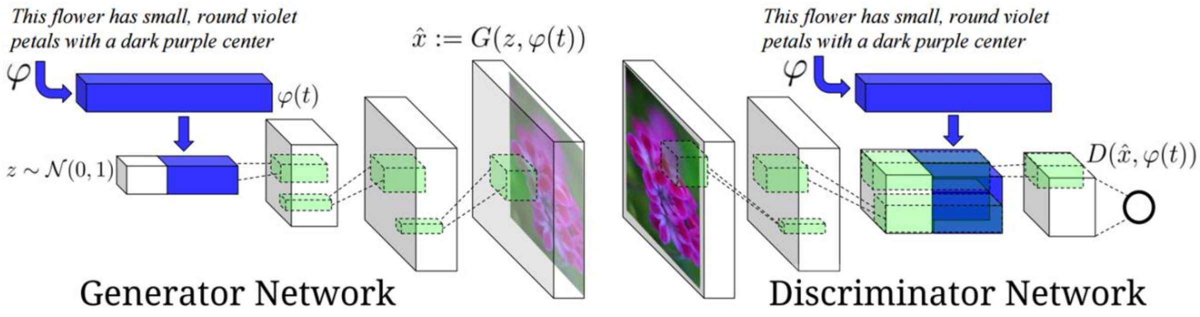
$$\nabla f^{CD} \approx \frac{-1}{N} \sum_{n=1}^{N} \left[\nabla F_\theta(x_n) - \sum_{x \epsilon X} Q_\theta(x) \nabla F_\theta(x)\right]$$

*'Training products of experts by minimizing contrastive divergence.', Hinton, G., Neural Computation, pp. 1771-1800, Vol. 14(8), 2002.*

# Machines with The Gift of Imagination :

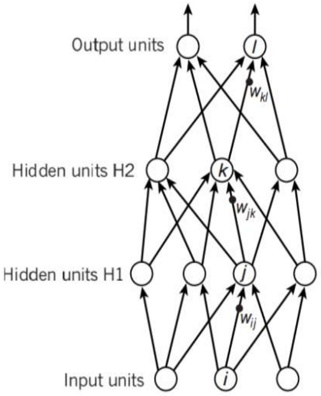
**The internal representation of the models are unexpectedly rich.**



## DERIVATIVE FLOW IN THE NET

**FORWARD PASS**




$y_l = f(z_l)$

$z_l = \sum_{k \varepsilon H2} w_{kl}\, y_k$

$y_k = f(z_k)$

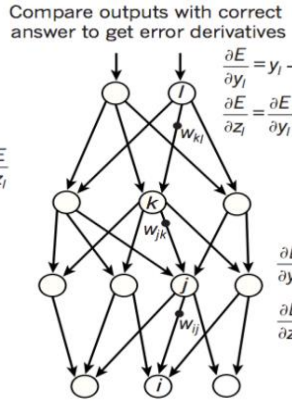
$z_k = \sum_{j \varepsilon H1} w_{jk}\, y_j$

$y_j = f(z_j)$

$z_j = \sum_{i \varepsilon \text{Input}} w_{ij}\, x_i$

$\frac{\partial E}{\partial y_k} = \sum_{l \varepsilon \text{out}} w_{kl} \frac{\partial E}{\partial z_l}$

$\frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k}$



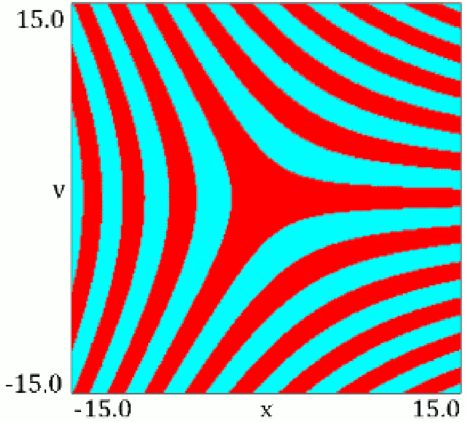

$\frac{\partial E}{\partial y_l} = y_l - t_l$

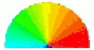
$\frac{\partial E}{\partial z_l} = \frac{\partial E}{\partial y_l} \frac{\partial y_l}{\partial z_l}$

$\frac{\partial E}{\partial y_j} = \sum_{k \varepsilon H2} w_{jk} \frac{\partial E}{\partial z_k}$

$\frac{\partial E}{\partial z_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial z_j}$

**REVERSE PASS**



### GIBBS SAMPLING

Sample $p(z_1, z_2, \ldots, z_M)$

1. Initialize $\{z_i : i = 1, \ldots, M\}$
2. For $\tau = 1, \ldots, T$:
   - Sample $z_1^{(\tau+1)} \sim p(z_1 | z_2^{(\tau)}, z_3^{(\tau)}, \ldots, z_M^{(\tau)})$.
   - Sample $z_2^{(\tau+1)} \sim p(z_2 | z_1^{(\tau+1)}, z_3^{(\tau)}, \ldots, z_M^{(\tau)})$.

     ⋮
   - Sample $z_j^{(\tau+1)} \sim p(z_j | z_1^{(\tau+1)}, \ldots, z_{j-1}^{(\tau+1)}, z_{j+1}^{(\tau)}, \ldots, z_M^{(\tau)})$.

     ⋮
   - Sample $z_M^{(\tau+1)} \sim p(z_M | z_1^{(\tau+1)}, z_2^{(\tau+1)}, \ldots, z_{M-1}^{(\tau+1)})$.

### CONTRASTIVE DIVERGENCE

$$f^{CD}(\theta) = \sum_{x \in X} P_e(x) \log\left(\frac{P_e(x)}{P_\theta(x)}\right) - Q_\theta(x) \log\left(\frac{Q_\theta(x)}{P_\theta(x)}\right)$$

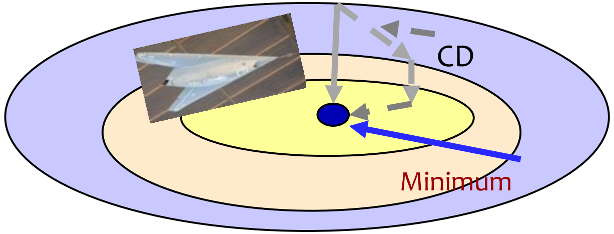
$$\nabla f^{CD} \approx \frac{-1}{N} \sum_{n=1}^{N} \left[\nabla F_\theta(x_n) - \sum_{x \epsilon X} Q_\theta(x) \nabla F_\theta(x)\right]$$

On contrastive divergence learning., Carreira-Perpignan, M. A. and Hinton, G. E. In Artificial Intelligence and Statistics, 2005

### FAST CONTRASTIVE DIVERGENCE (CD)

$$\theta_{t+1} = \theta_t + \lambda \frac{\partial L(X; \theta)}{\partial \theta}$$

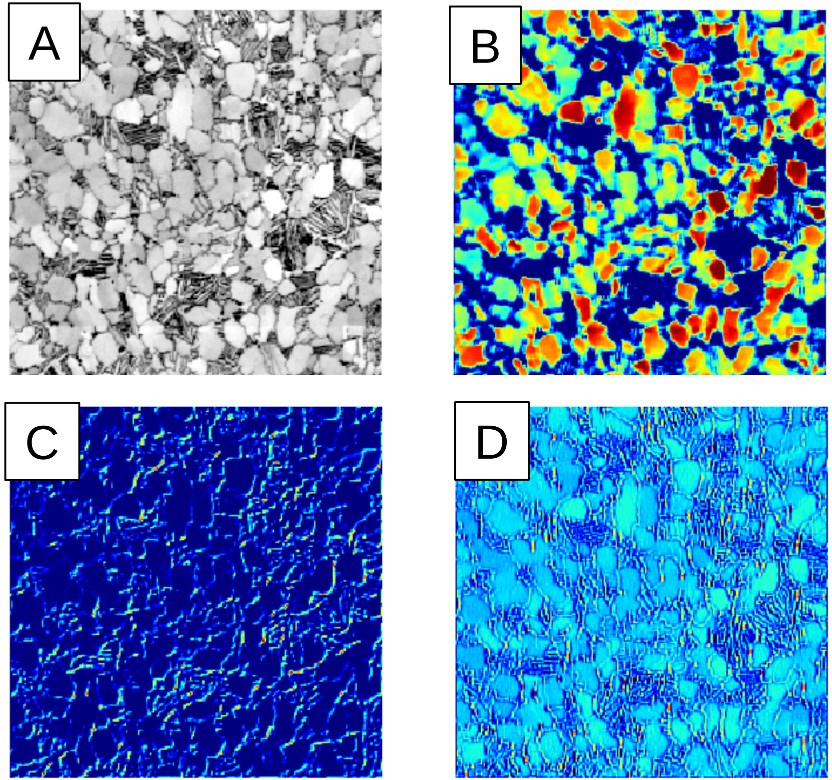
Approximate but fast gradient



Alloy System Ti-6Al-4V

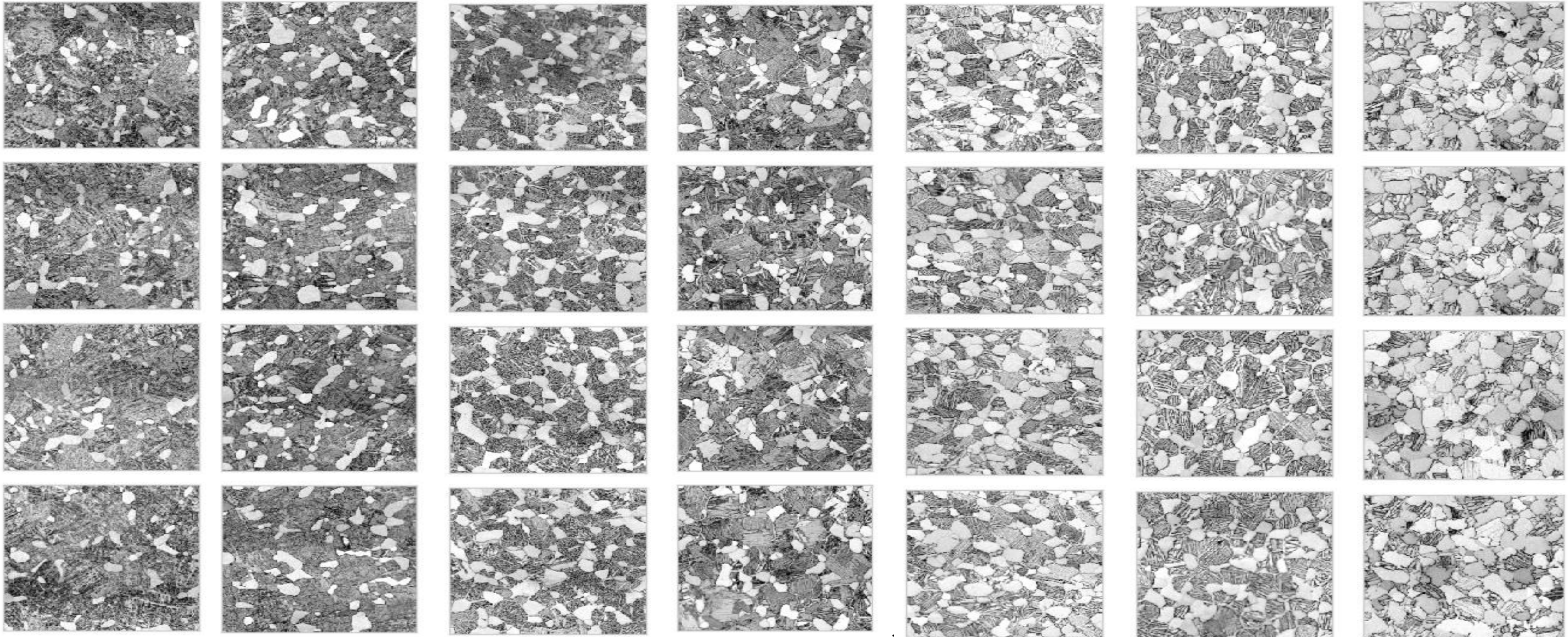
CNN Image Representations and Material Physics

- Material properties are driven by the constitutive components of the dual phase titanium μ–structure (α and β phase) (e.g. image A).
- CNNs automatically learn and capture physical contributions from individual phases (B, C) as well as interaction effects (D).
- Higher fidelity predictions come at the expense of reduced interpretability .

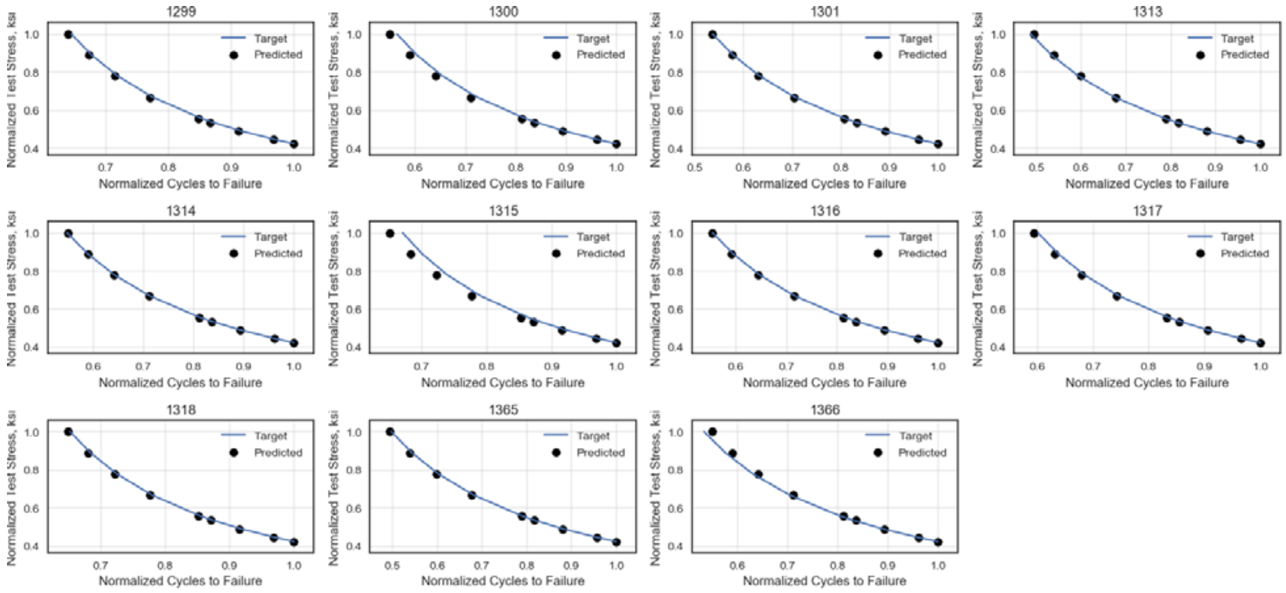


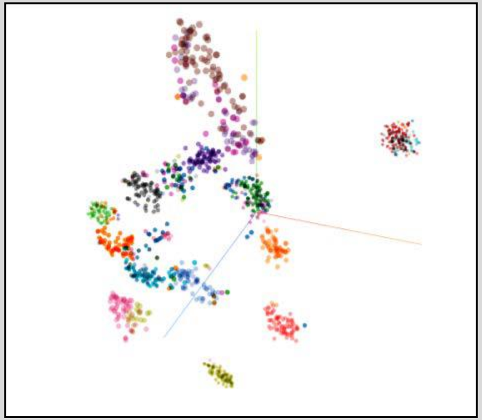
## Ti-6Al-4V GLOW Generated Pedigrees -

14 6 18 16 17 12 2



## HCF model predictions for 11 pedigrees of Ti-6Al-4V.

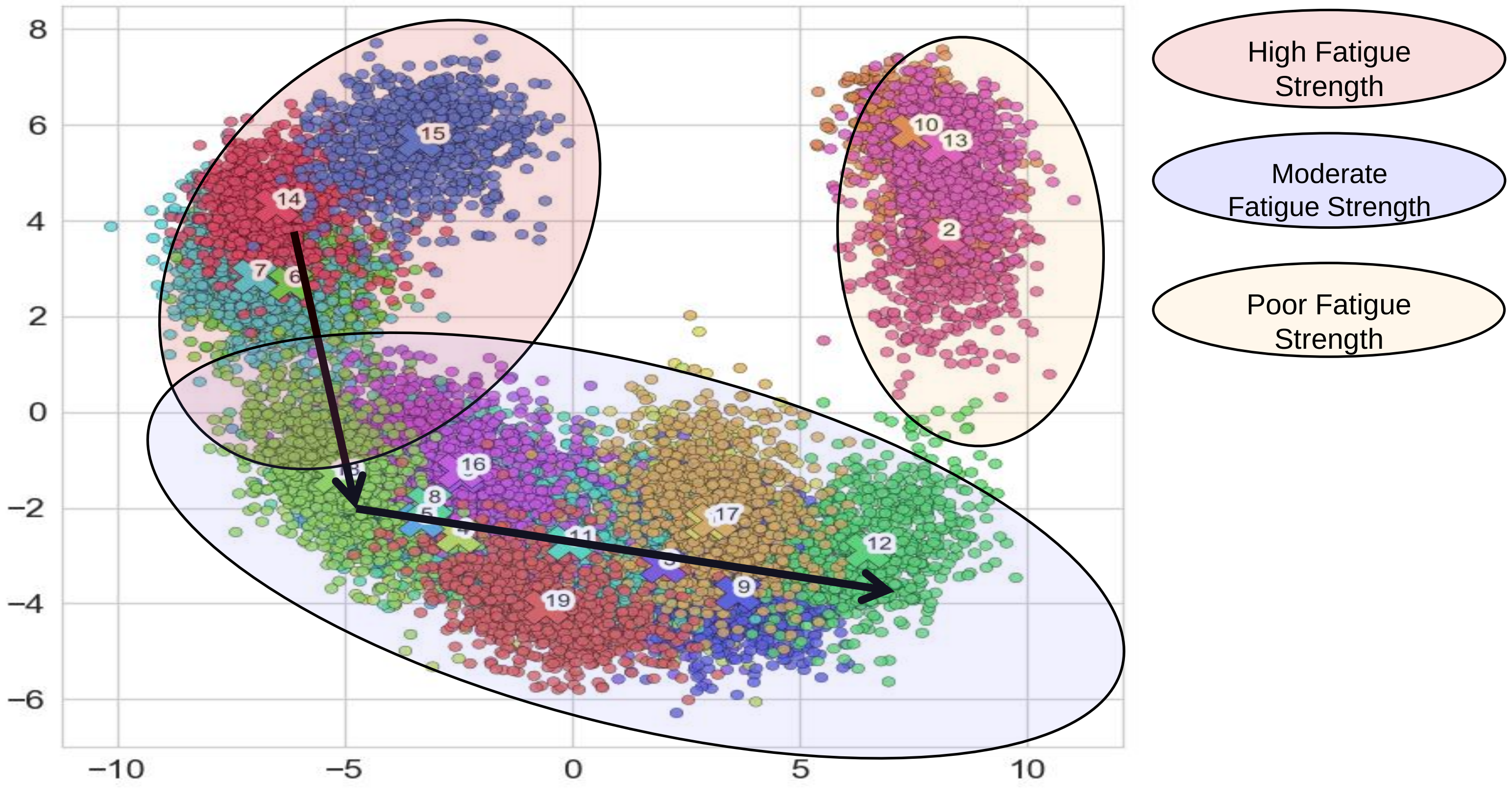


3D visualization of the 150,000-Dim μ-structure image space.

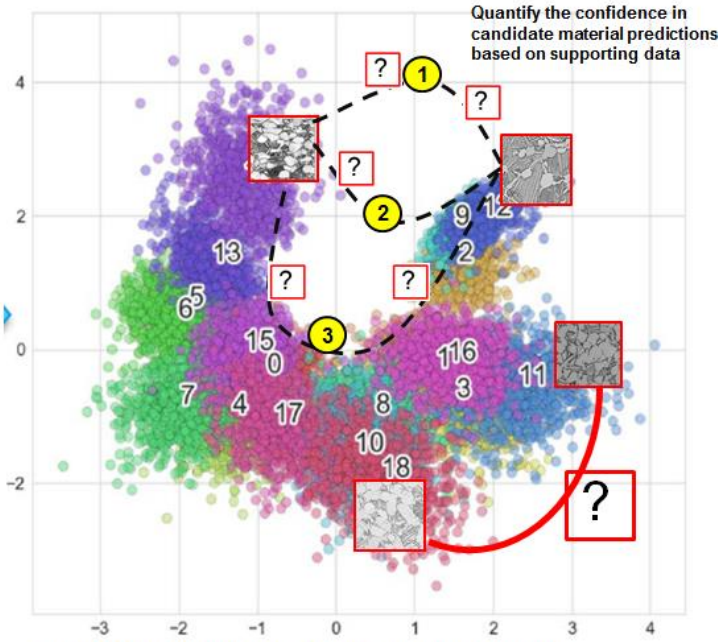
A visualization of the image embedding space (right) highlights the presence of clusters of similar-looking material

# Principal Component Analyses – Latent Space of Ti-6Al-4V pedigrees with HCF correlation

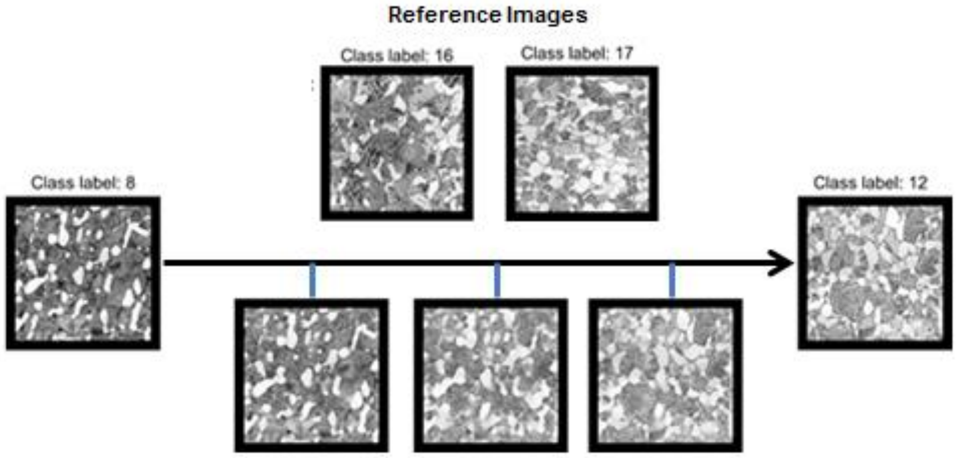


# Structural Material Property Tailoring Using Deep Neural Networks

*GLOW model embeds $\mu$–structure images in a latent space in which new materials can be discovered and optimized!*




Scatter plot of material samples clustering by image, pedigree and processing information




GLOW generated images

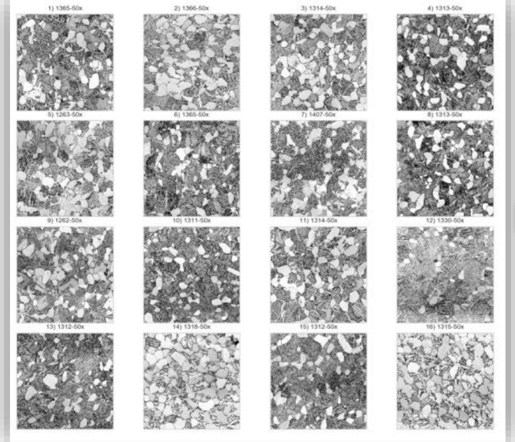
GLOW can accurately interpolate between distinct material pedigrees (visible clusters) and generate photo-realistic images along this path. Generated images can be used for material prediction and optimzation.

## Exploring Material Design Space using Generative Models

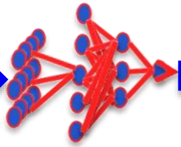
## Structural Material Property Tailoring Using Deep Neural Networks

*Deep learning presents an approach to scalable analytics for materials engineering, driving down cost and increasing quality.*

**Images**



Neural Network



**Latent Space Representation**

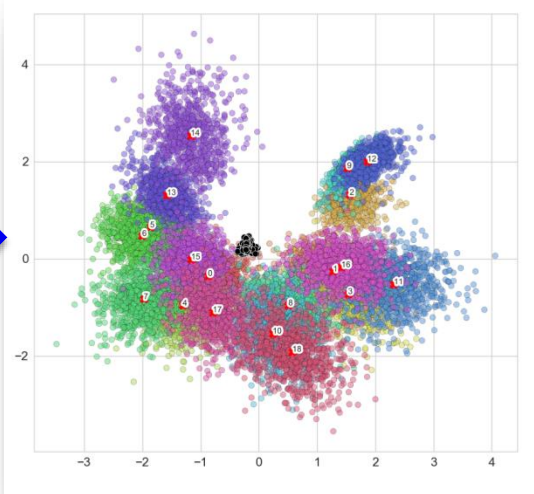
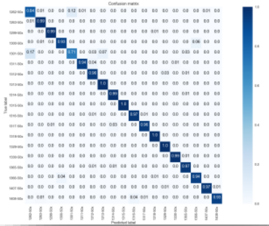
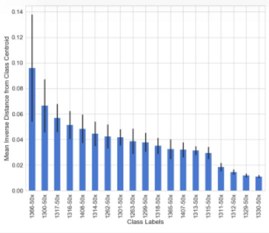


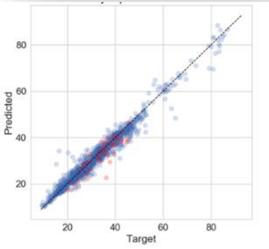

***Image Classification***



***Visual Similarity Ranking***

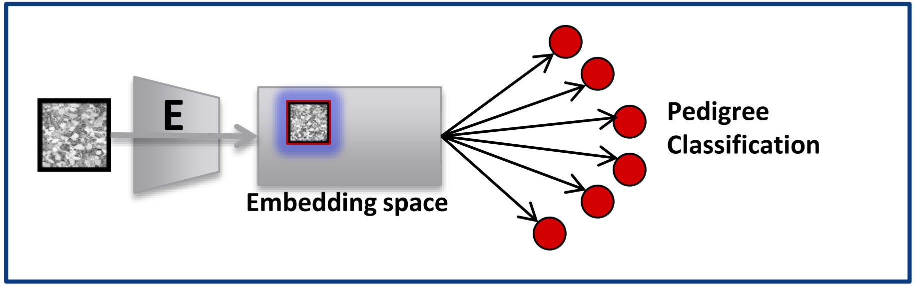


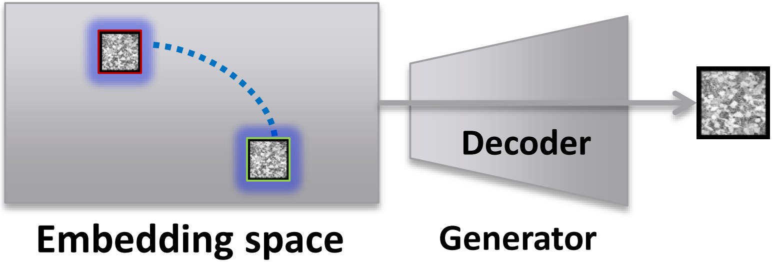
***Property Prediction***

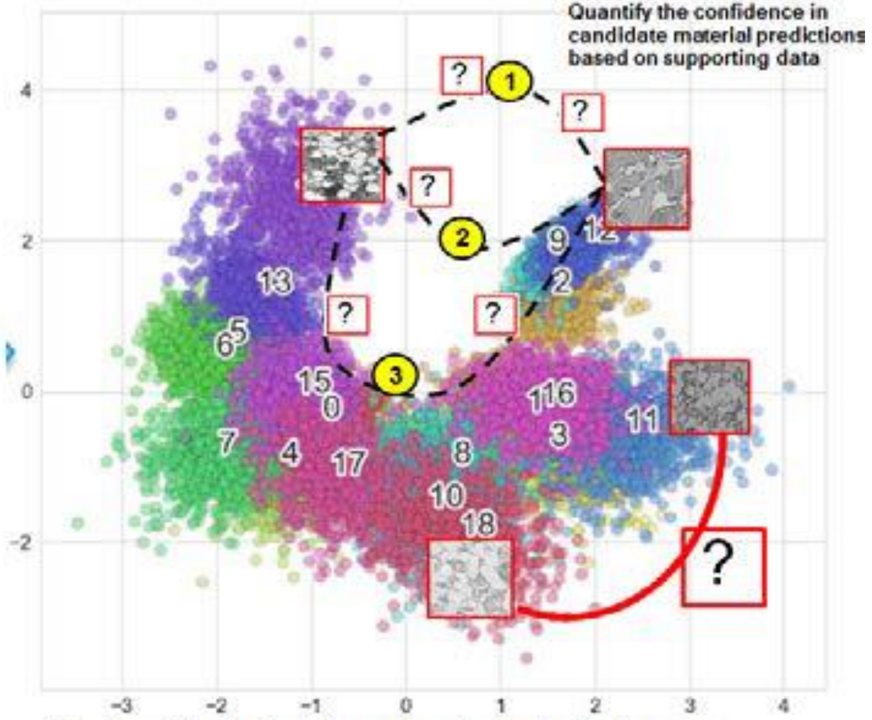



- *Material Definition.*
- *Process Quality Control*
- *Outlier detection*
- *Determining statistical representations of microstructures*

*Each point in the embedding space corresponds to a synthesized $\mu$–structure image.*



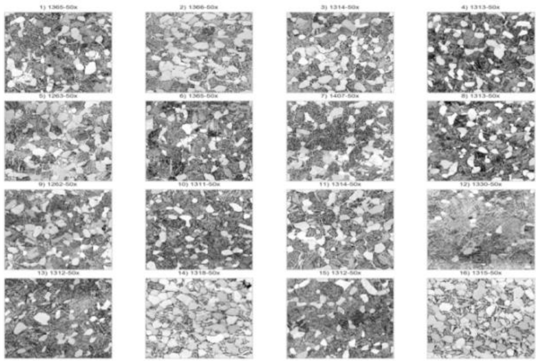




Scatter plot of material samples clustering by image, pedigree and processing information

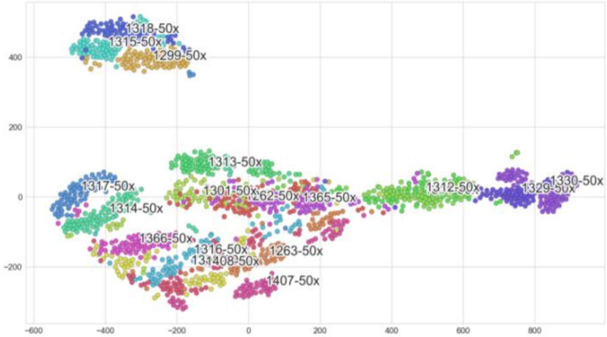
# Structural Material Property Tailoring Using Deep Neural Networks

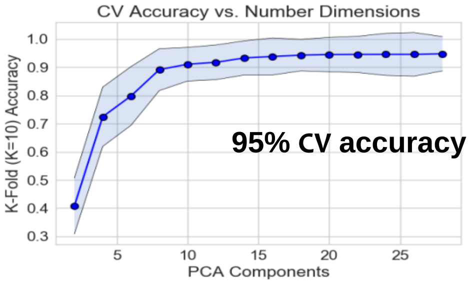
## Convolutional nets demonstrated to accurately classify subtle differences in microstructure

19 Different Heat Treatments

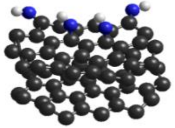


Use Dimensionality Reduction to View in 2D

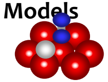


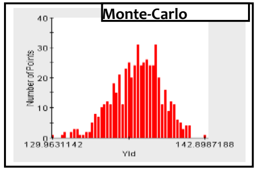



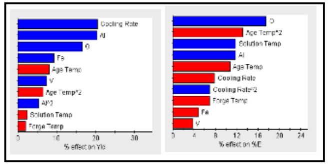

Input Images → Extract Statistics Using CNN → Dimensionality Reduction → Prediction

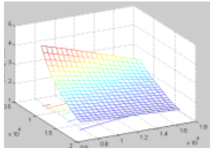


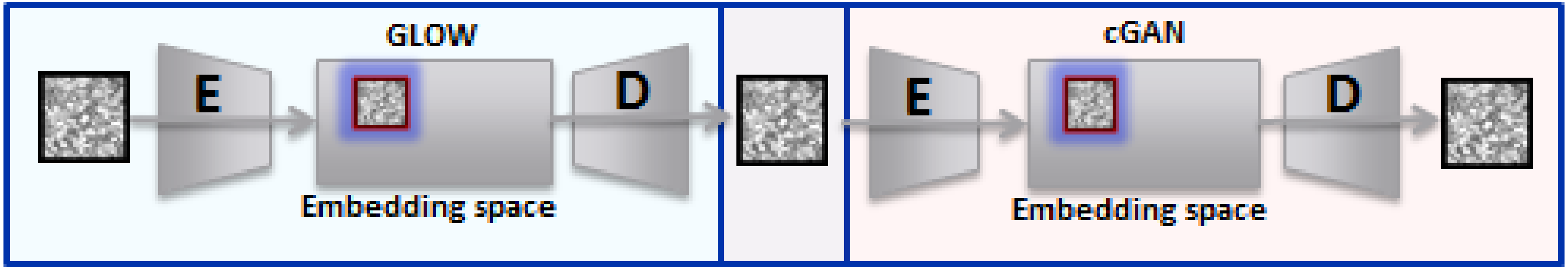
Physics Models Of Atomic Structures





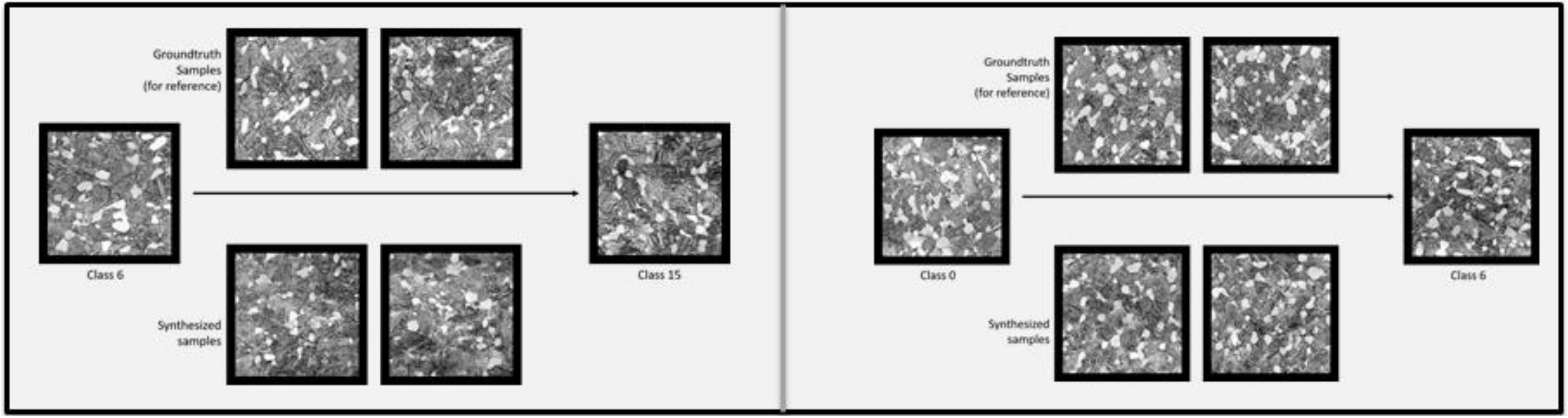





GLOW model:
- trained on raw input micrographs.
- Learns an embedding space.
- Outputs high quality images.

cGAN model:
- trained on GLOW outputs *AND* raw input micrographs.
- Learns an embedding space.
- Improves image grain textures.

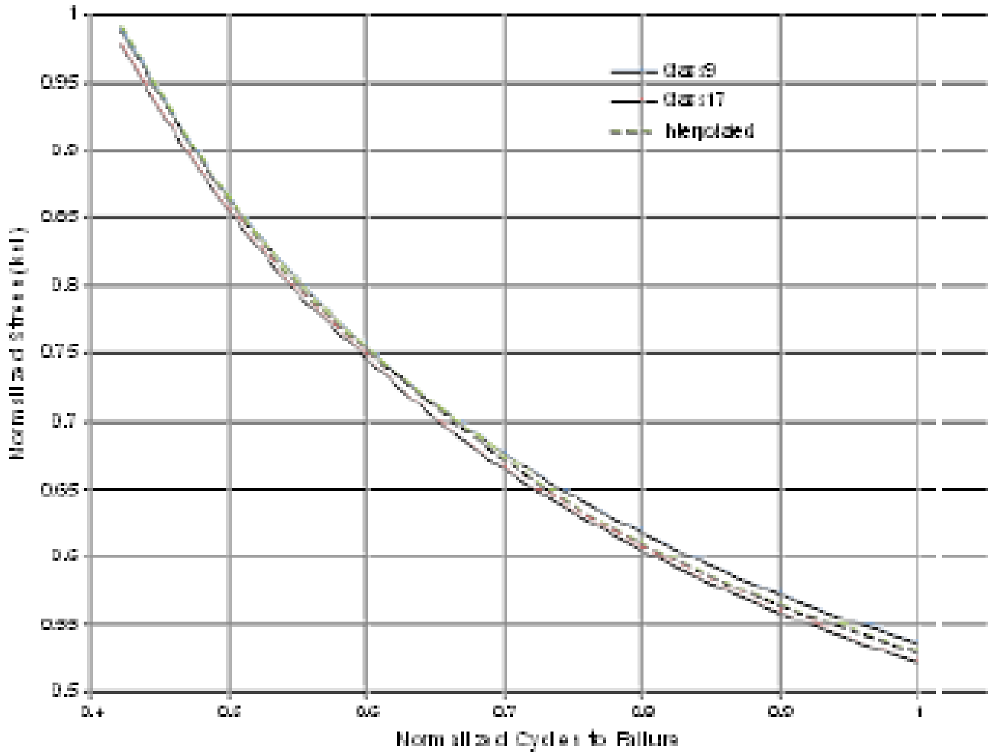
This toolchain requires 1-1 mappings at each step. This drives our choice of models to use.




- Examples of GLOW-cGAN interpolation between two known material classes.
- Center top: Ground-truth images of intermediate material.
- Center bottom: Synthesized images of intermediate material.

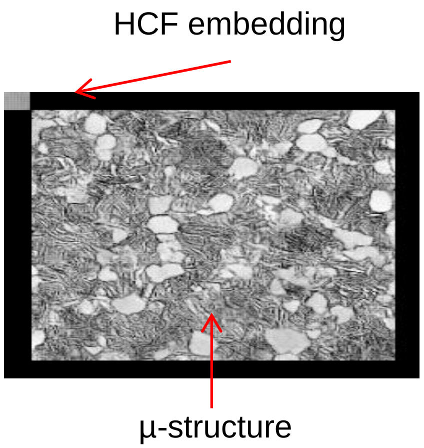
# Structural Material Property Tailoring Using Deep Neural Networks

Predicted fatigue curve for the generated image set closely aligns with the known ground-truth for that alloy.




- Predicted fatigue curve for the generated image set closely aligns with the known ground-truth for Ti-6Al-4V alloy.
- Rigorous, additional validation required, but the results offer credence that the models can be useful for deeplearning framework based interpolation in regions of design interest.

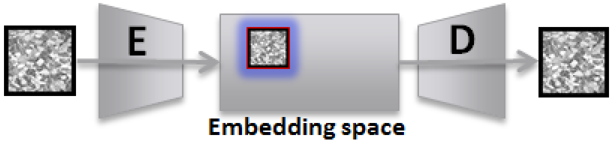
- Matrix augmentation describes the process of generating a constrained labeled matrix.
- Training data by transformation by including the basic pixel image RGB channel data with the labeled data of measured attributes as an additional channel.
- Matrix data augmentation for deep learning is done to form a physics consistent formulation for optimization.
- Matrix data augmentation can be numerically beneficial as an additional regularizer for very deep architectures.
- Alternate way to incorporate known prior knowledge about image attributes, physics and processing constraints.

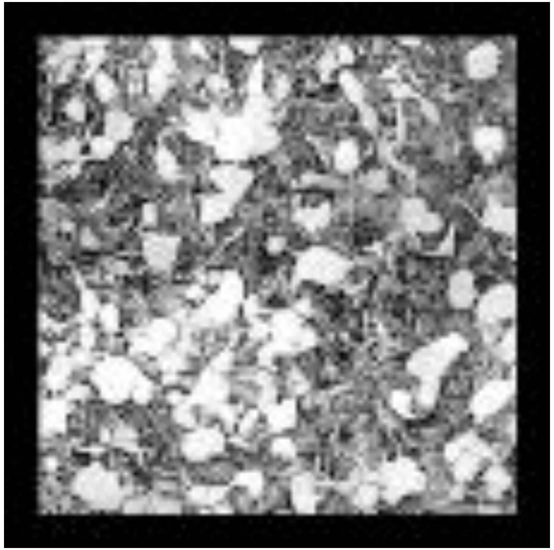



AUGMENTED MATRIX FORMULATION

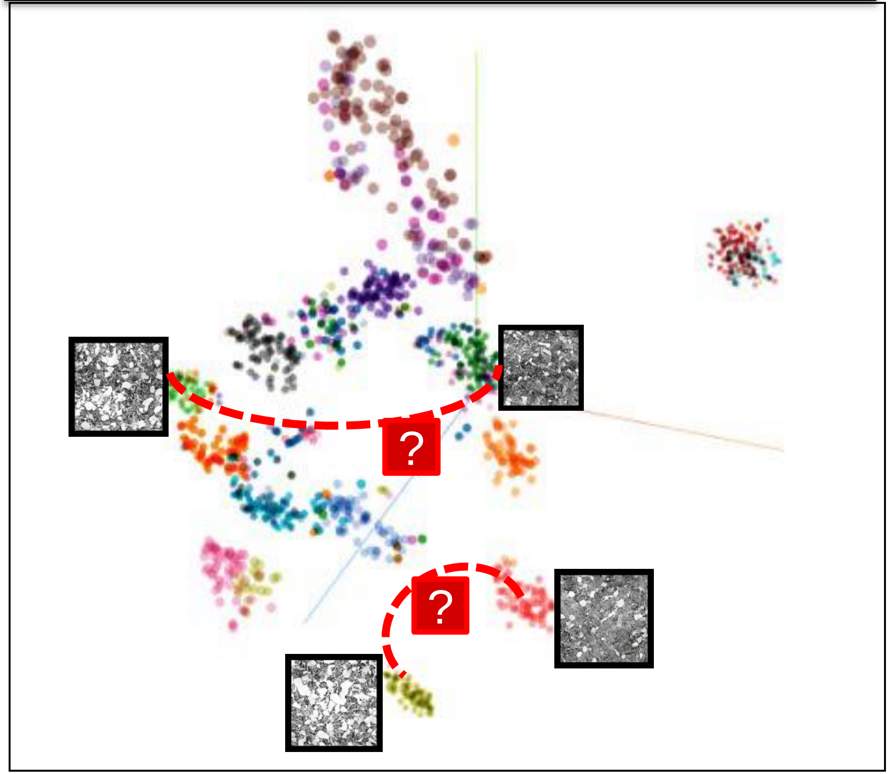
Test images augmented with associated prior information of normalized fatigue strength.

# Generative Images Using GLOW

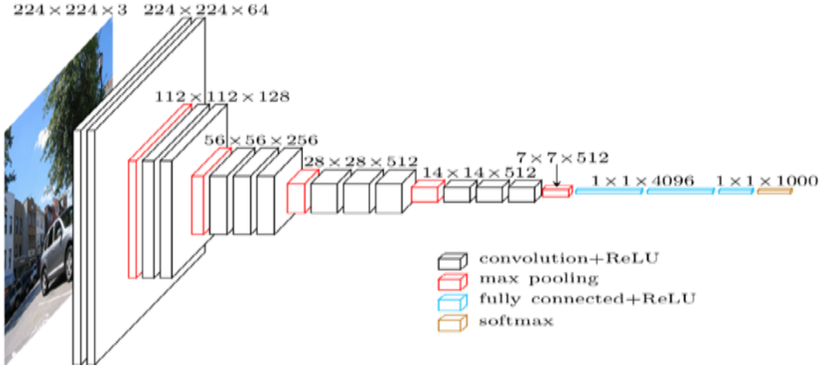


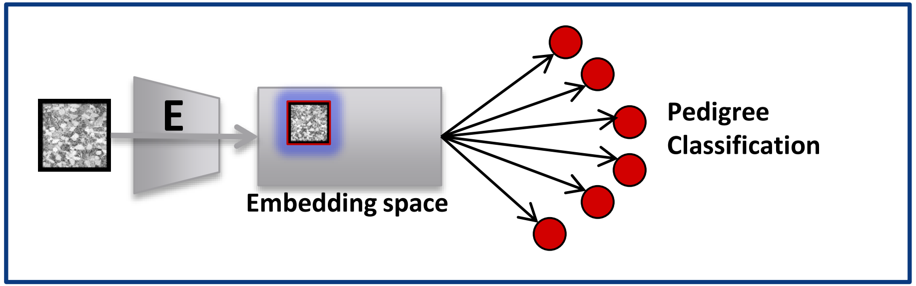



3D visualization of the 150,000-Dim μ–structure design space.

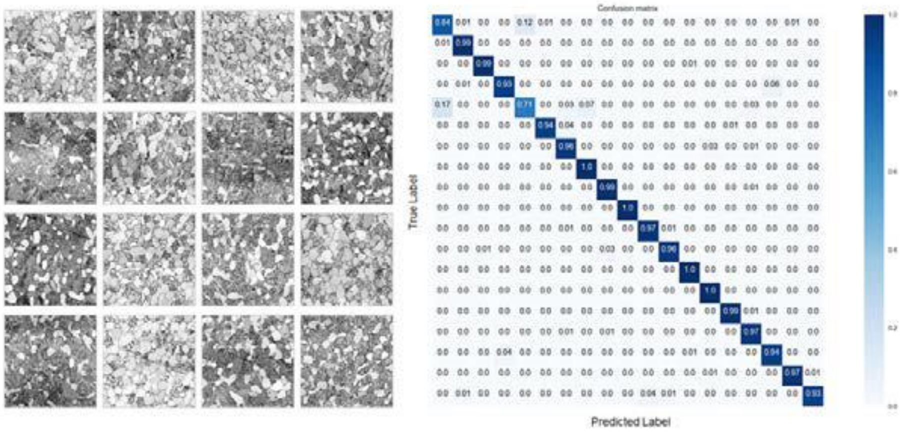


## Convolutional networks for automated material quality and conformance



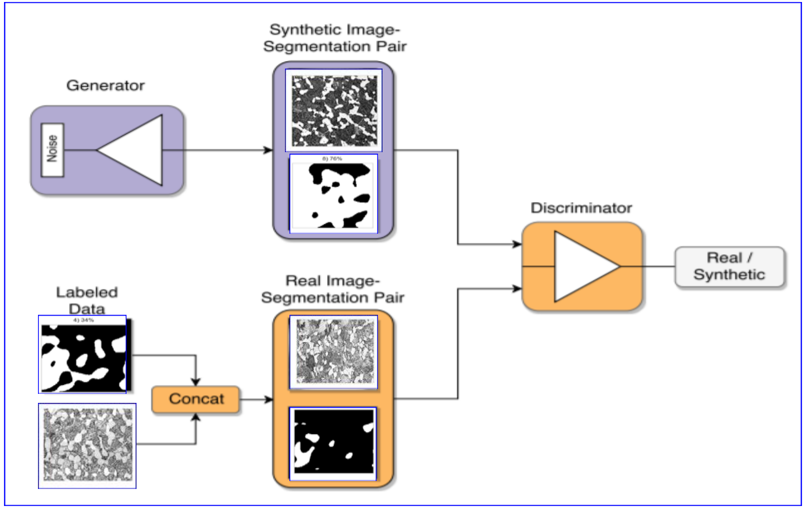

VGG16 model embeddings are supplemented with a classification layer.




Confusion matrix showing classification results from a tuned VGG16 convolution neural network (CNN).

A generative model is tuned to generate additional, synthetic data resembling the real data.

- 2D models render realistic images to increase amount of training data / improve performance and predictability.
- GANs, refine purely rendered data with information from real data.
- Standard GAN definition only allows for generation of images, without respective labels.
- Generated data used for data augmentation, conventional GAN formulation needs to be modified.
- Fuse image and segmentation mask to create an image-segmentation pair.




SYNTHETIC DATA AUGMENTATION

RGB image [W × H × 3] + segmentation mask [W × H × 1] = image segmentation pair [W ×H ×4].

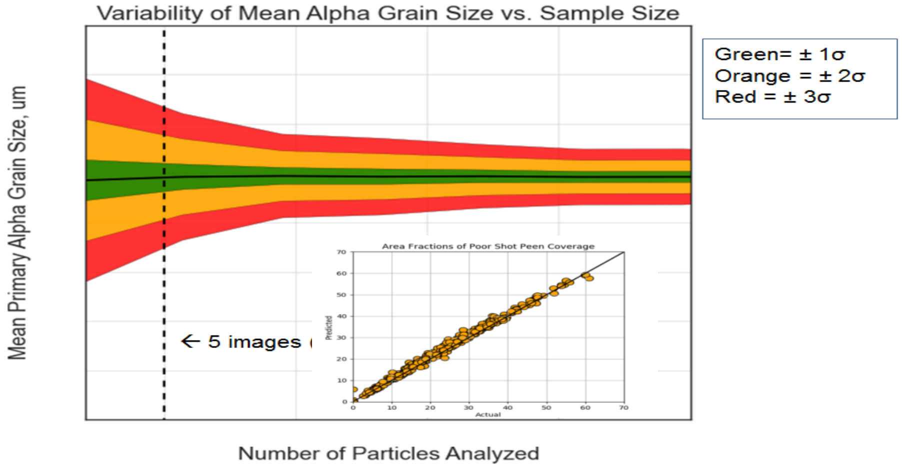
# Structural Material Property Tailoring Using Deep Neural Networks

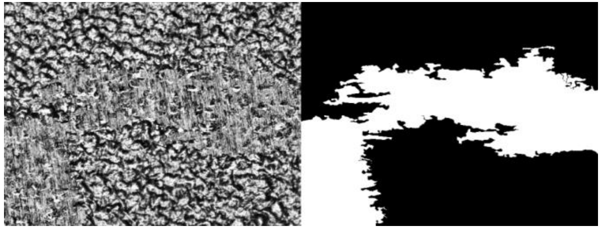
Deep CGAN Neural Networks:
- State-of-the-art in object recognition and classification
- Models able to learn distinguishing features in image during supervised training

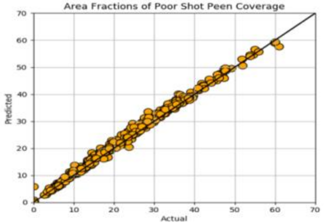
Need:
- Image analysis can take 1-2 hrs/image
- Reduce uncertainty w/ more data

Applications:
- Pixel-wise classification
- Anomaly detection (defects)
- Statistical representations of microstructure

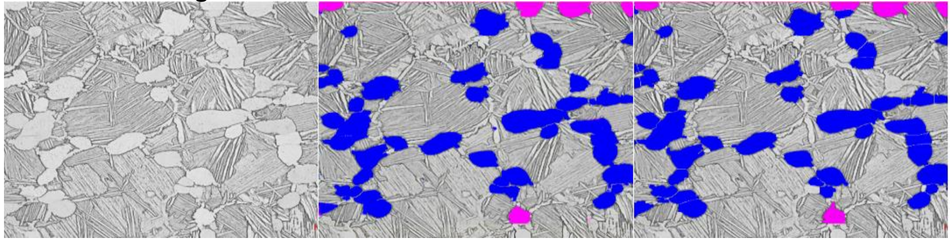
Progress
- Pix2pix models developed capable of performing near-human segmentation of microstructures



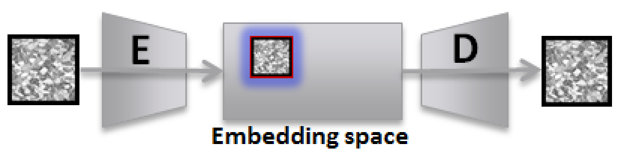


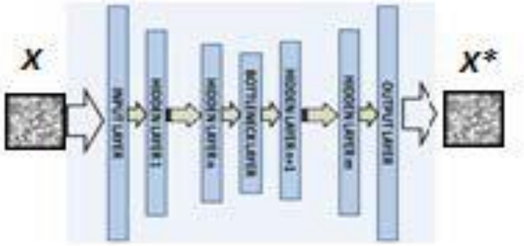

Raw Image Model Technician



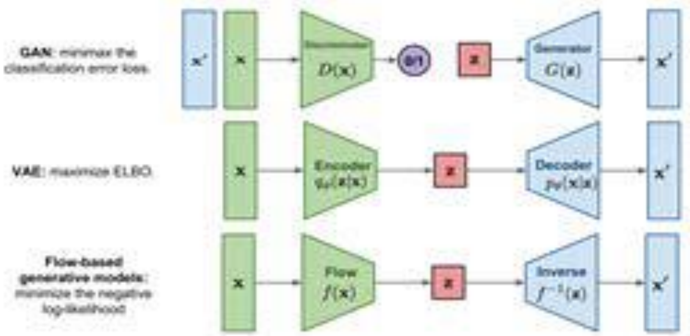
Conditional generative adversarial network (CGAN) (above)

*The generative models used are all encoder-decoder networks.*

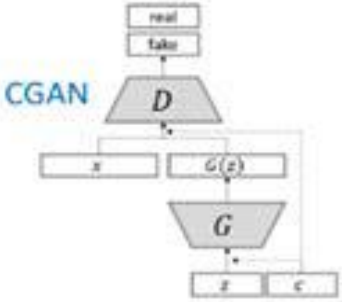



CNN




GAN – VAE – FLOW



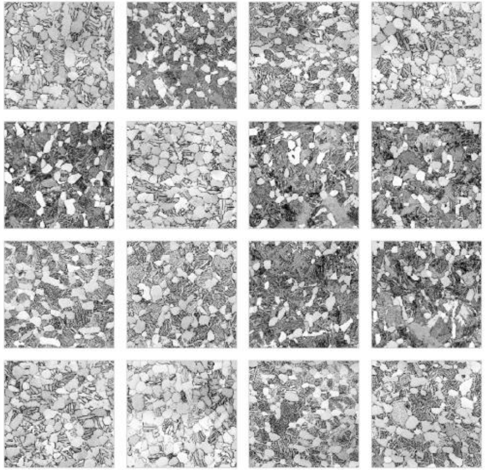

cGAN

The learning process uses all standard techniques, like backpropagation, but train two models: a Generative Model (G) and a Discriminative Model (D). G captures data distribution and generates samples and the G icalculates the probability that a presented sample came from the training data and not from the generative model.

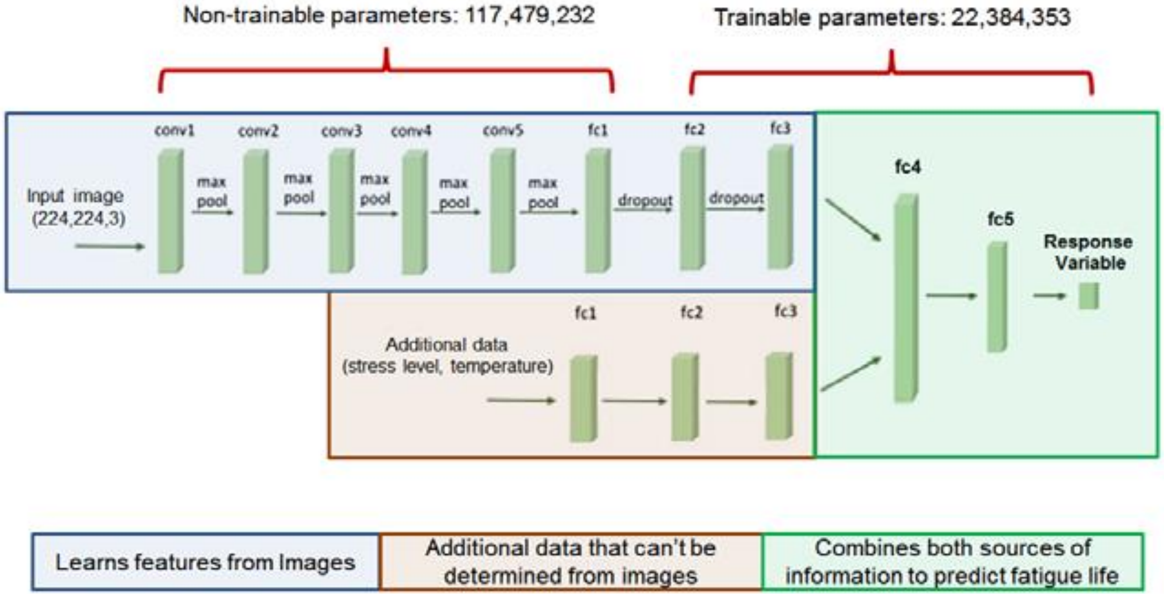
G AND D are competing against each other. The Generative Model (G) will try to generate samples for which Discriminative Model will conclude that they are coming from the training dataset. Essentially, the Generative Model will try to maximize the probability of the Discriminative Model creating a mistake, similar to minmax two-player game.

## Using CNNs to predict material properties from images

Training Input



Modified VGG16 Model Architecture




Modified VGG16 architecture, accepting two inputs (image and vector)

# Exploring Material Design Space using Generative Models

Input: Microstructure and HCF label

GLOW

GLOW Latent Space

GLOW Image Gen

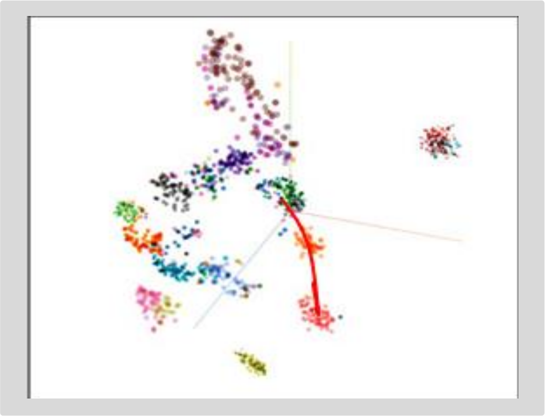
VGG (untuned)

GLOW Compressed Latent Space

UMAP Visualization

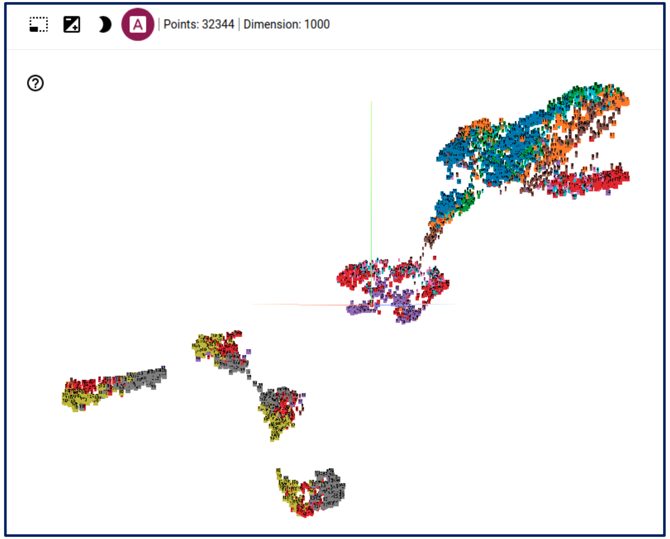
Choose paths (Points P)

Choose non-training data points along P and orbits



# Latent Space Exploration

- The Latent Space dimension of the GLOW model is very high dimensional.
- Gaining intuition of positions of the relative locations of pedigree clusters in the latent space requires dim-reduction and visualization.
- The UMAP visualization method was chosen for:
  - Repeatability of the visualization.
  - Ability to save and embed new data after training.
- We confirmed that the addition of the HCF information patch on the image did not noticeably impact the latent space.
- The latent space of the padded images is less separable, though relative cluster positions remain stable.

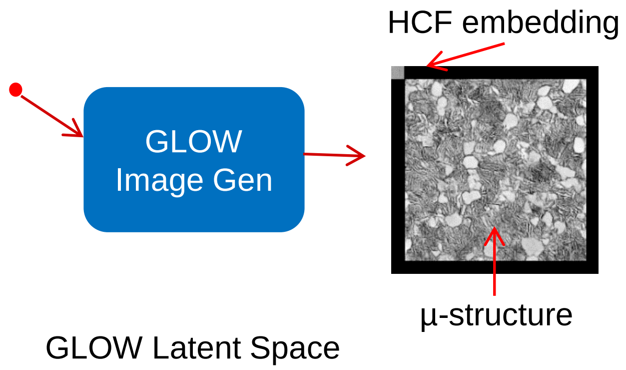


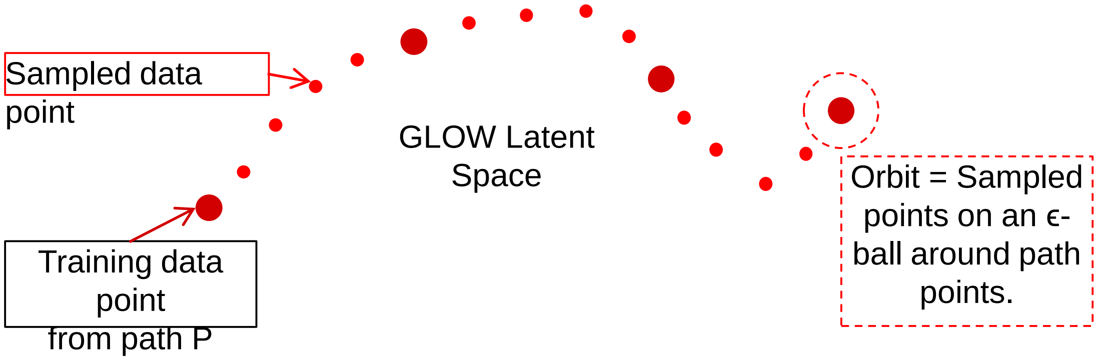
# Sample / Generate / Validate

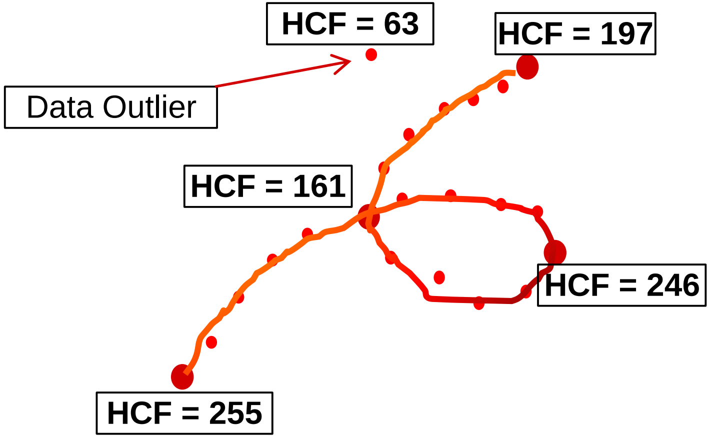
Once a set of points P for a path in the latent space are chosen, we anchor our latent space sampling to it.

Reconstruct sampled points to evaluate the microstructure image and predicted HCF value.

Orbits are a generally accepted method of sampling from generative image models.



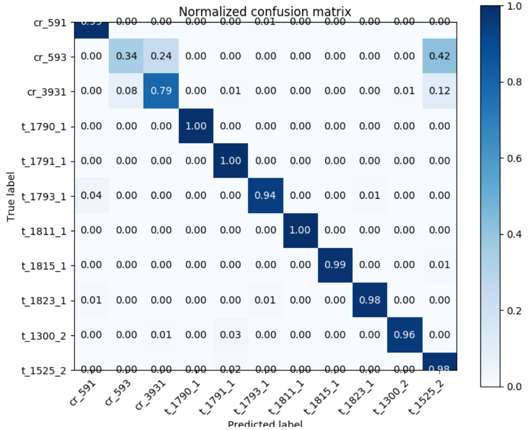



## Structural Material Property Tailoring Using Deep Neural Networks

A VGG16 network tuned to predict the pedigree class of our data and applied to GLOW generated image to see if it would be recognized as valid and of correct class.

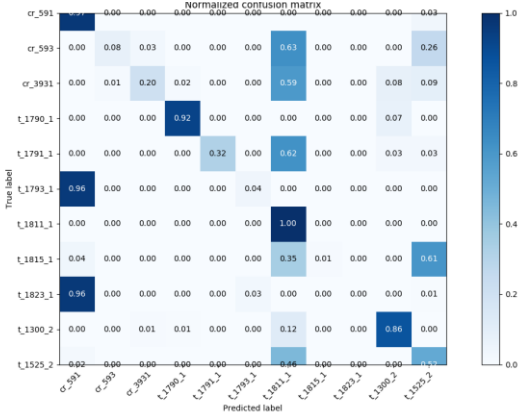
VGG classifier confusion matrix on real data performs perfectly aside from class 2, 3 which are difficult for even experts to discriminate.

Precision=93%, Recall =93%.




Confusion matrix on *translated* images. The VGG's ability to correctly classify the GLOW generated images decreases significantly.

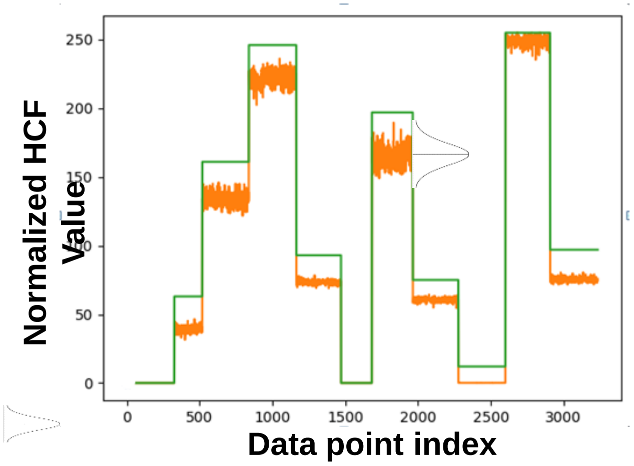
Precision 63%, Recall 47%

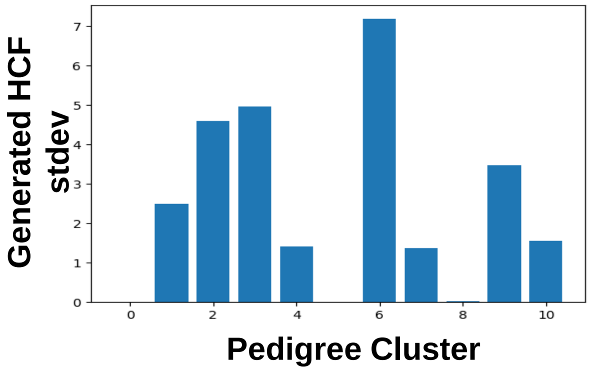



Unexpected given much better results in previous work. Clearly training for the HCF value and microstructure simultaneously has degraded the microstructure.

## Intra-Cluster Paths

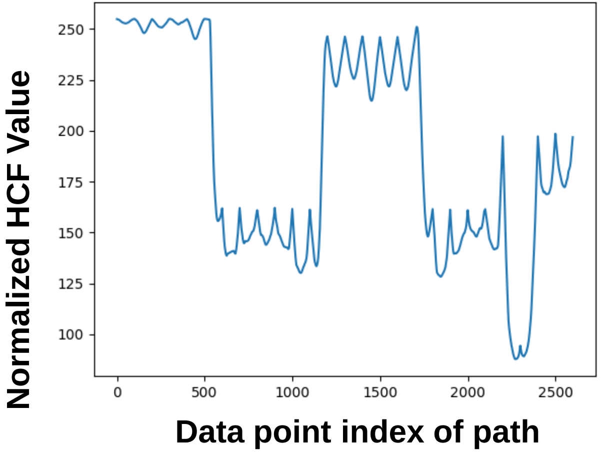
- Green -> normalized HCF values across all 11 classes training data.
- Orange -> mean predicted HCF predicted after data translated in latent space.
- Predicted HCF's -> 99.4% correlated with the raw HCF's.
- Intra-class variance is small compared to differences in cluster HCF values.



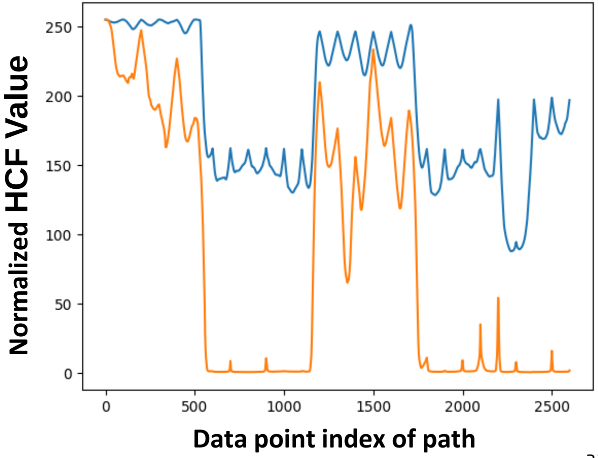
## Predicted HCF Values

Tracing a path through multiple pedigree clusters we see predicted HCF values clustered around actual.




Garland-like artifact is a reflection of the offset bias we saw earlier and will be greatly mitigated. Model used was trained on all 11 pedigrees.

Repeat the previous path experiment, but with **pedigree 8 not present** in the data to see how well the micrograph are predicted for a material absent in training.

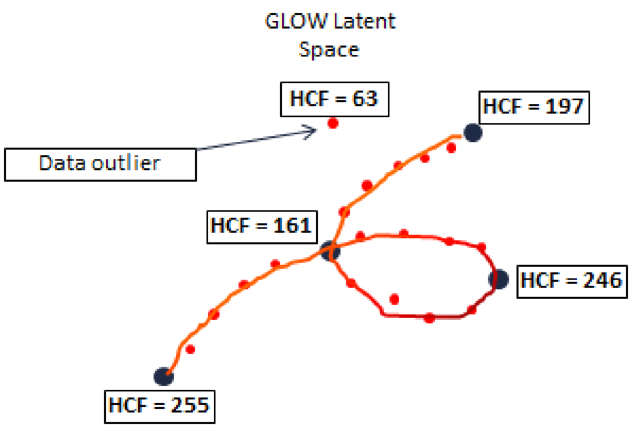


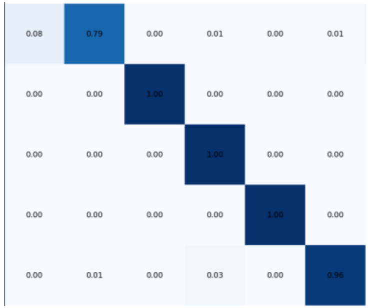

Removing pedigree 8 disabled it's ability to predict HCF in that region and degraded performance in areas that overlapped with it around index 300-500.

This is evidence that the HCF predictions of this method aren't robust far outside of their training exemplars.

# Classification of Generated Micrographs

A VGG16 network was tuned to predict the pedigree class of our data and then applied to the GLOW generated image to see if it would be recognized as valid and of the correct class.

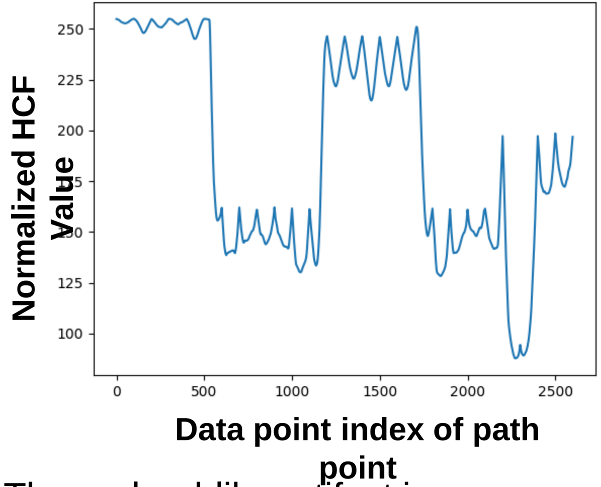



GLOW generated micrographs were classified correctly 99% of the time along the path.

Precision=99%, Recall = 95%

Within an area of strong support, the GLOW model generates excellent microstructures.




The garland-like artifact is a reflection of the offset bias we saw earlier and will be greatly mitigated. The model used was trained on all 11 pedigrees.

# Summary

- Convolutional neural networks (CNNs) used for challenging problems in materials engineering such as image classification, quantitative property / image attribute prediction, and synthetic microstructure image generation.
- Generative and predictive capability of convolutional neural networks were demonstrated on a Ti-6Al-4V high cycle fatigue dataset.
- Model was validated via the generation of new microstructures images by interpolating between known pedigrees and predicting fatigue response.
- CNNs also demonstrated good performance in predicting quantitative attributes from images (phase fraction) – bypassing the need for image segmentation and analysis.
- Synthetic microstructures were generated using the GLOW model and were able to accurately classified by a CNN trained on real images. The GLOW model was then trained on a subset of the data (leave one out validation) in order to assess ability to interpolate between known pedigrees of material.
- The model had limited ability to predict HCF properties in the unknown region. Future work will include collection of additional microstructure images to further populate the sparse latent space and improve ability to interpolate between known pedigrees.

# Something In this World She knows….

"You can't connect the dots looking forward; you can only connect them looking backwards.
So you have to trust that the dots will somehow connect in your future."

**Steve Jobs – From The Tao of Steve**


PRATT & WHITNEY
DEPENDABLE ENGINES

GO BEYOND