TOWARD LEARNING LATENT-VARIABLE REPRESENTATIONS OF MICROSTRUCTURES BY OPTIMIZING IN SPATIAL STATISTICS SPACE

Anonymous authors
Paper under double-blind review

ABSTRACT

In Materials Science, material development involves evaluating and optimizing the internal structures of the material, generically referred to as microstructures. Microstructures structure is stochastic, analogously to image textures. A particular microstructure can be well characterized by its *spatial statistics* Paulson et al. (2017), analogously to image texture being characterized by the response to a Fourier-like filter bank Varma & Zisserman (2002).

Material design would benefit from low-dimensional representation of microstructure Paulson et al. (2017).

In this work, we train Variational Autoencoders (VAE) to produce reconstructions of textures that preserve the spatial statistics of the original texture, while not necessarily reconstructing the same image in data space. We accomplish this by adding a differentiable term to the cost function in order to minimize the distance between the original and the reconstruction in spatial statistics space.

Our experiments indicate that it is possible to train a VAE that minimizes the distance in spatial statistics space between the original and the reconstruction of synthetic images. In future work, we will apply the same techniques to microstructures, with the goal of obtaining low-dimensional representations of material microstructures.

1 INTRODUCTION

Computer-aided materials design holds the promise of designing materials with favorable properties by enabling researchers to computationally optimize potential materials.

Material microstructure is often represented as a large 3D image. 2D images are often also used. In research, both real images (e.g., X-ray) and simulated microstructures are used.

In "data space," the dimensionality of the microstructure representation is the number of pixels/voxels, which can be quite large. This project is motivated by the need for working in low-dimensional latent space when optimizing material properties. Previous work includes representing the microstructure using the first principal components of the spatial statistics representation Paulson et al. (2017). However, such a representation is not reversible. A reversible low-dimensional representation allows one to optimize the material in low-dimensional space, and to then reconstruct the material in data space.

In this paper, we modify the Variational Autoencoder (VAE) Kingma & Welling (2014) to learn a low-dimensional representation that preserves the statistical properties of the material by making the original and the reconstruction be similar in spatial statistics space. We work with 2D images.

2 BACKGROUND

Generation of images with a specific texture that is specified by the response to a Fourier-like filter bank (or indeed the Fourier transform) goes back decades Matsuyama et al. (1983). Recently, the literature on neural style transfer Gatys et al. (2016) constrained the image generator with two cost functions: one that controlled the content and one that controlled the style. The style cost function constrained the response to a ConvNet in the lower layers, some of which are known to be Fourier-response-like Zeiler & Fergus (2014).

In Materials Science, the microstructure m is often characterized using spatial statistics Cecen et al. (2016),

$$f\left(h, h' \mid \boldsymbol{r}\right)=\frac{1}{\operatorname{Vol}\left(\Omega_{\boldsymbol{r}}\right)} \int_{\Omega_{\boldsymbol{r}}} m(h, \boldsymbol{x}) m\left(h', \boldsymbol{x}+\boldsymbol{r}\right) d \boldsymbol{x}.$$

Here, $m(h, \boldsymbol{x})$ is the local state h at location $\boldsymbol{x}$ in the microstructure. The spatial statistics (or correlations) can be interpreted as the set of correlations between all locations at distance r, in state h. In practice the spatial statistics can be computed as the inverse Fourier transform of the magnitudes of the Fourier transform of the microstructure Paulson et al. (2017).

3 METHODS

We train a VAE to reconstruct images of texture such that, in spatial statistics space, the original and the reconstruction are close. That is, the VAE loss we use is

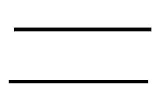
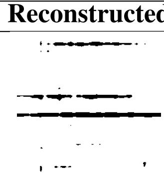
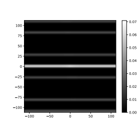
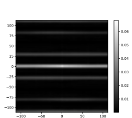
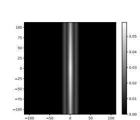
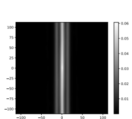
$$\mathcal{L}=\alpha \cdot ||f(\mathbf{x})-f(\mathbf{x_{recon}})||_2+\beta \cdot \text{Loss}_{KL}. \tag{1}$$

That is, instead of minimizing the distance between the input x and the reconstruction x_{recon}, we minimize the distance between the spatial statistics f of x and x_{recon}. The spatial statistics function can be efficiently computed using FFT and is differentiable[1]

4 RESULTS

On a dataset 100,000 randomly-placed vertical and horizontal lines (between 1 and 4) of various lengths on a 224x224 images, we demonstrate that our network reconstructs images with similar spatial statistics, but which are quite different in terms of L2 in image space. That is, we are able to obtain a low-dimensional representation of the image that contains information about the spatial statistics. Our next step is to apply the method to simulated and real microstructures. We train a VAE with a pretrained ResNet-152-based encoder with a bottleneck of size 9, and untrained decoder. On a manually-labelled test set of 100 images, the reconstructions have the correct orientation in 98% of cases. On average, line volume differed by 17% between the original and the reconstruction.

Table 1: Original, reconstructed, and spatial statistics difference

Data		**Spatial statistics**	
Original	**Reconstructed**	**Original**	**Reconstructed**
			
			

URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of ICLR 2024 Tiny Papers Track.

[1] https://pytorch.org/docs/stable/fft.html

REFERENCES

Ahmet Cecen, Tony Fast, and Surya R Kalidindi. Versatile algorithms for the computation of 2-point spatial correlations in quantifying material structure. *Integrating Materials and Manufacturing Innovation*, 5:1–15, 2016.

Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423, 2016.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *Proc. ICLR*, 2014.

Takashi Matsuyama, Shu-Ichi Miura, and Makoto Nagao. Structural analysis of natural textures by fourier transformation. *Computer vision, graphics, and image processing*, 24(3):347–362, 1983.

Noah H Paulson, Matthew W Priddy, David L McDowell, and Surya R Kalidindi. Reduced-order structure-property linkages for polycrystalline microstructures based on 2-point statistics. *Acta Materialia*, 129:428–438, 2017.

Manik Varma and Andrew Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part III 7*, pp. 255–271. Springer, 2002.

Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*, pp. 818–833. Springer, 2014.

A EVALUATION

The performance of the VAE was assessed with 100 image pairs from the test set. From the 100 reconstructions, 98 had the correct line orientation, and 63 had the correct number of lines as the original image

The reconstructions are characterized by randomly arranged lines, except one line remains constant across all images, varying only in size and orientation in response to the source image. The black pixels in the reconstructions, which represent lines, exhibited an average difference in volume fraction of 17%.

A.1 CALCULATION OF VOLUME FRACTION DIFFERENCE

The volume fraction difference percentage is a metric in evaluating the accuracy of the reconstructed images in terms of the black pixel content. This metric quantifies the difference in the number of black pixels (representing lines in images) between the original and the thresholded (or reconstructed) images. The calculation of this percentage is vital for understanding the degree to which the reconstructed images deviate from the original in terms of pixel composition.

The formula for calculating the volume fraction difference percentage is as follows:

$$\left| \frac{\text{original\_black\_pixel\_count} - \text{thresholded\_reconstruction\_black\_pixel\_count}}{\text{original\_black\_pixel\_count}} \right| \times 100 \quad (2)$$

After calculating the percentage difference, the value is rounded to four decimal places for precision and ease of interpretation. This rounding process helps in standardizing the results for comparison and further analysis.

Note: The calculated volume fraction percentage is rounded to four decimal places.

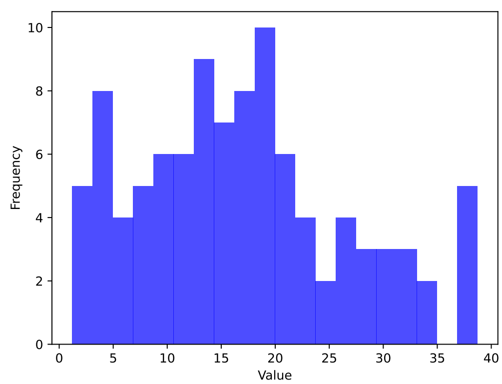


Figure 1: Volume fraction difference of black pixels in original-reconstruction pairs in percentage.

A.2 Image thresholding function

The threshold function is implemented to modify the image based on a predefined threshold value, setting pixel values below the threshold to zero and those equal to or above the threshold to one. This process is crucial for enhancing image clarity in the VAE's reconstructed outputs.

The mathematical representation of the threshold function is as follows:

$$\text{threshold\_pixel}(pixel, threshold) = \begin{cases} 0 & \text{if } pixel < threshold, \\ 1 & \text{otherwise.} \end{cases} \tag{3}$$

A.3 Plots of number of lines in the original images vs. the number of lines in reconstructions

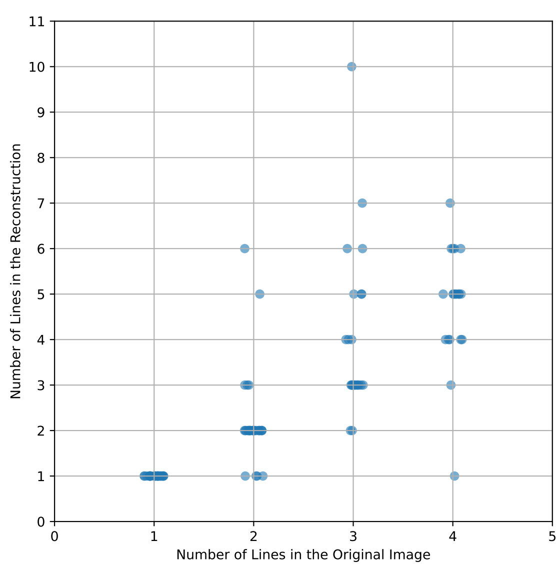


Figure 2: Number of lines in the original images vs. the number of lines in reconstructions.

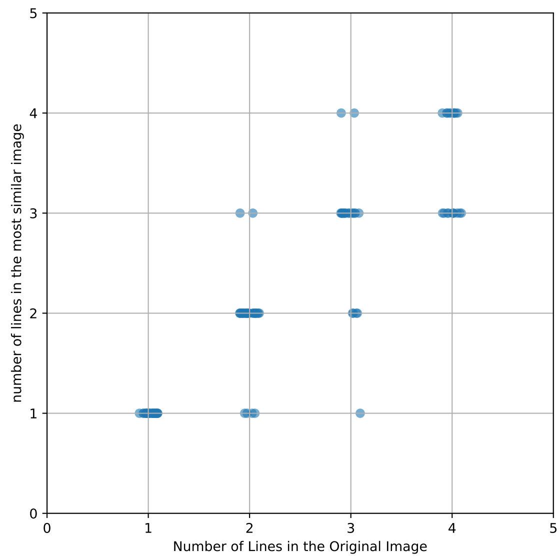


Figure 3: Number of lines in the original images vs. the number of lines in the most similar images, based on spatial statistics, to the reconstructions.

B More illustrative examples from the training set and their spatial statistics

Table 2: Original, reconstructed, and spatial statistics difference

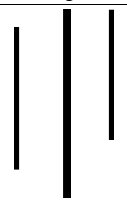
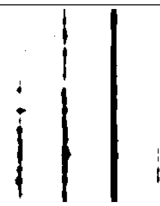
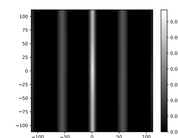
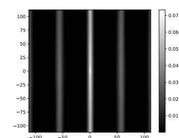
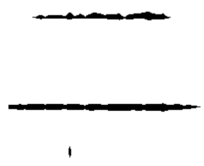
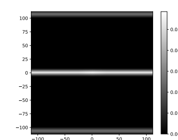
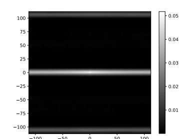
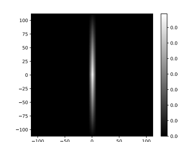
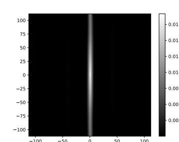
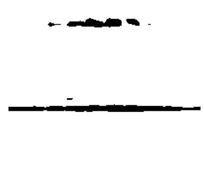
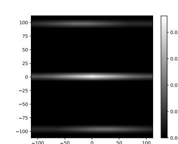
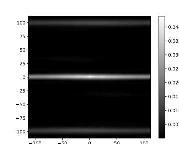
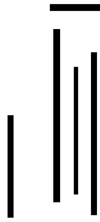
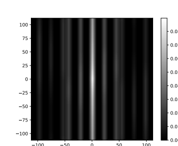
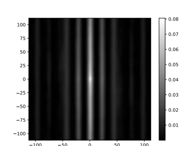
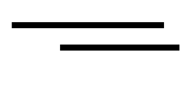
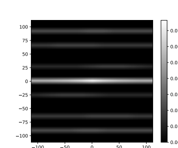
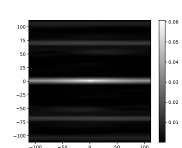
Data		Auto-correlations	
Original	**Reconstructed**	**Original**	**Reconstructed**
			
			
			
			
			
			

Table 3: Original, reconstructed, and spatial statistics difference

Data		**Auto-correlations**	
Original	**Reconstructed**	**Original**	**Reconstructed**

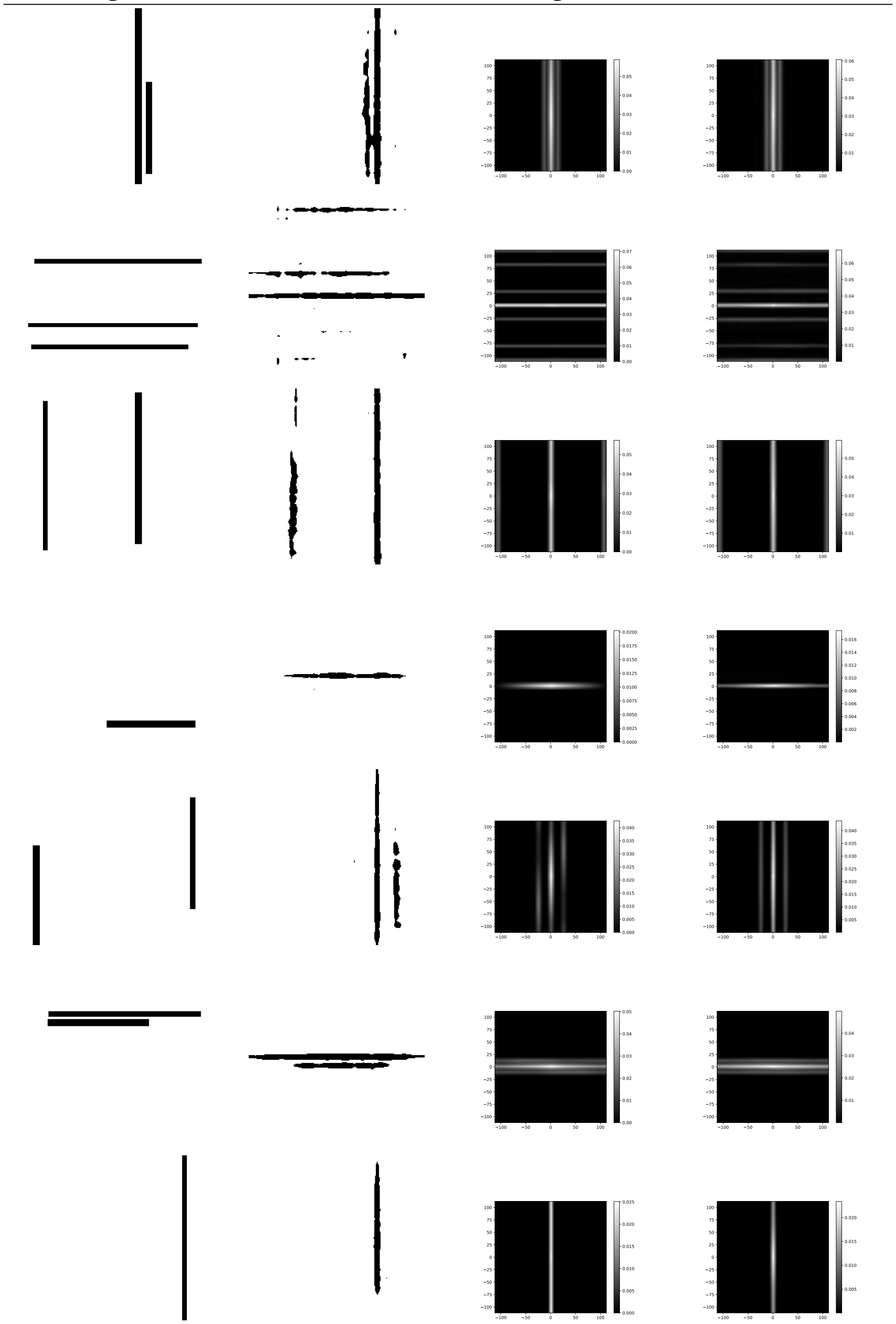
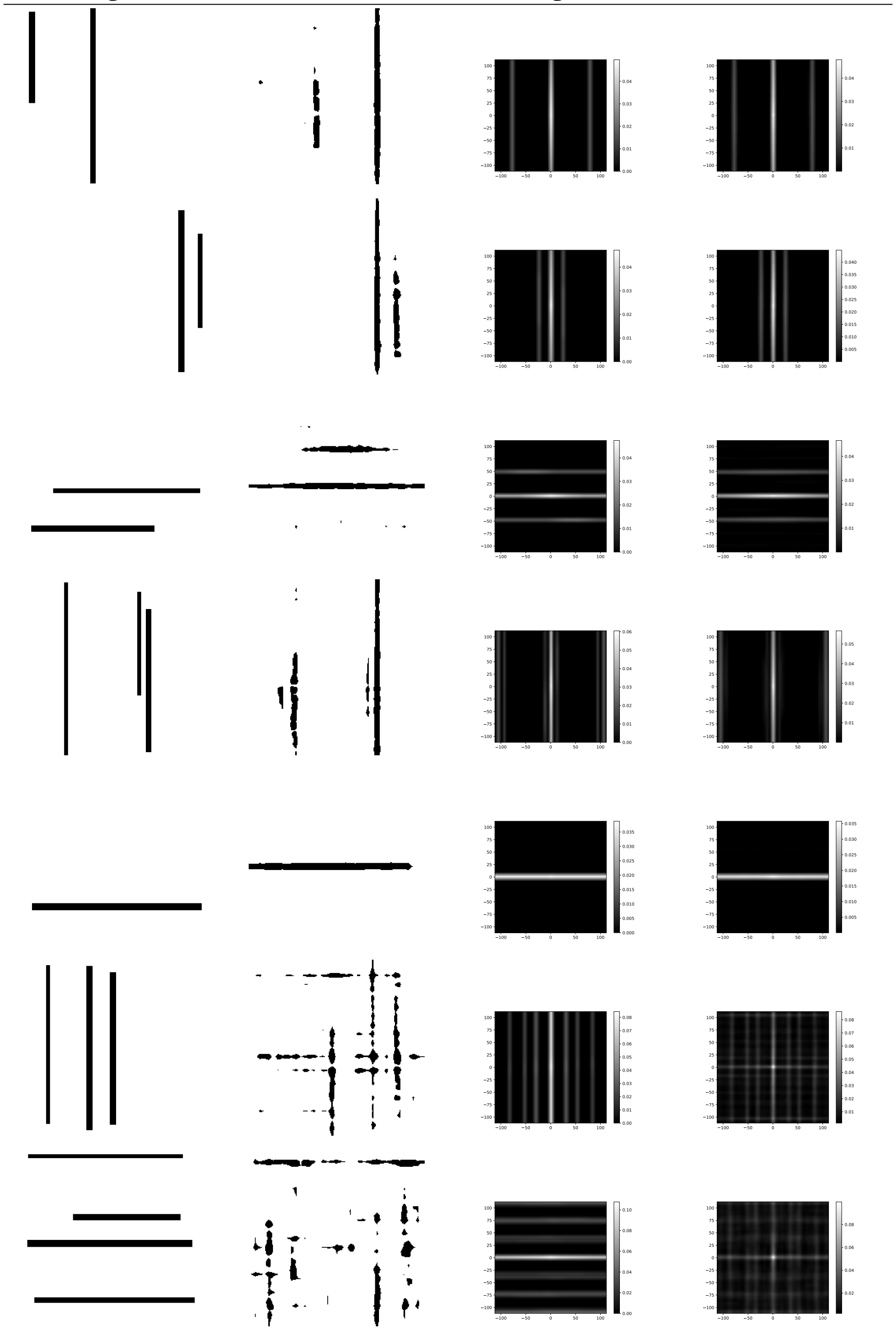


Table 4: Original, reconstructed, and spatial statistics difference

Data		Auto-correlations	
Original	**Reconstructed**	**Original**	**Reconstructed**



B.1 More illustrative examples of our results along with most similar images from the training set

Table 5: Original, reconstructed, and spatial statistics difference

Data		Most similar images	
Original	Reconstructed	Based on image MSE	Based on auto-corr. MSE

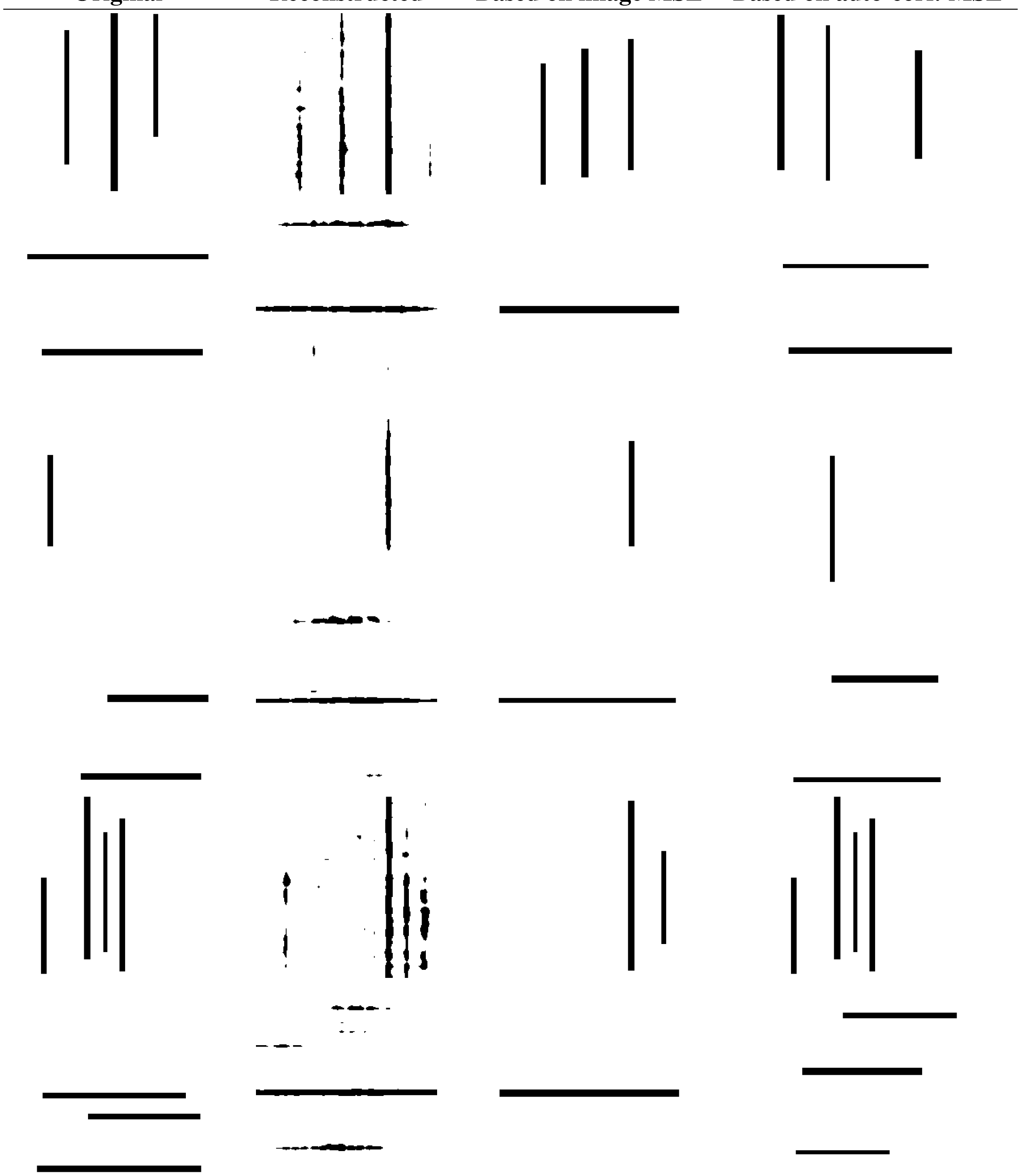


Table 6: Original, reconstructed, and spatial statistics difference

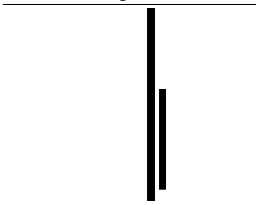
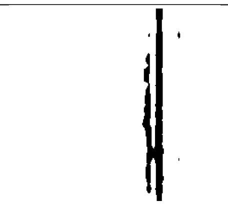
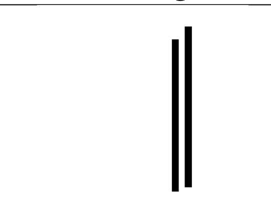
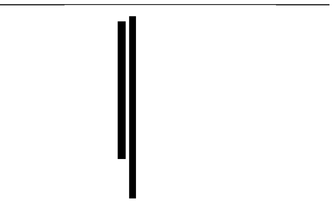
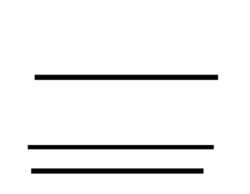
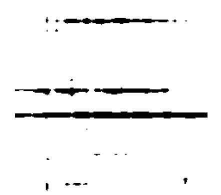
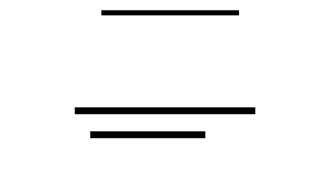
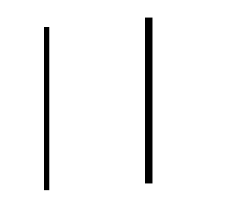
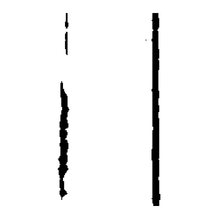
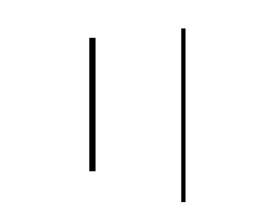
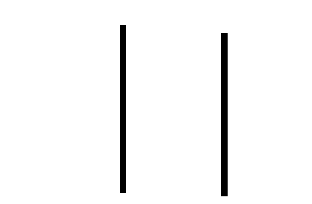
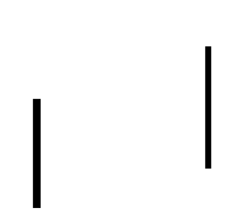
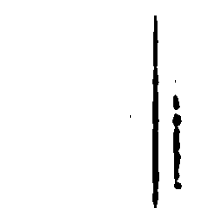
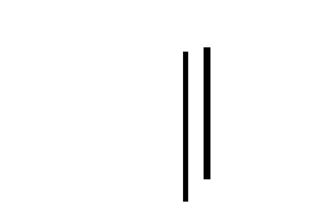
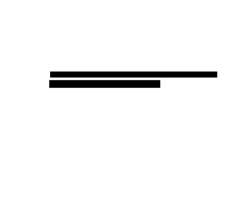
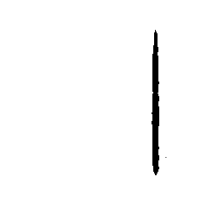
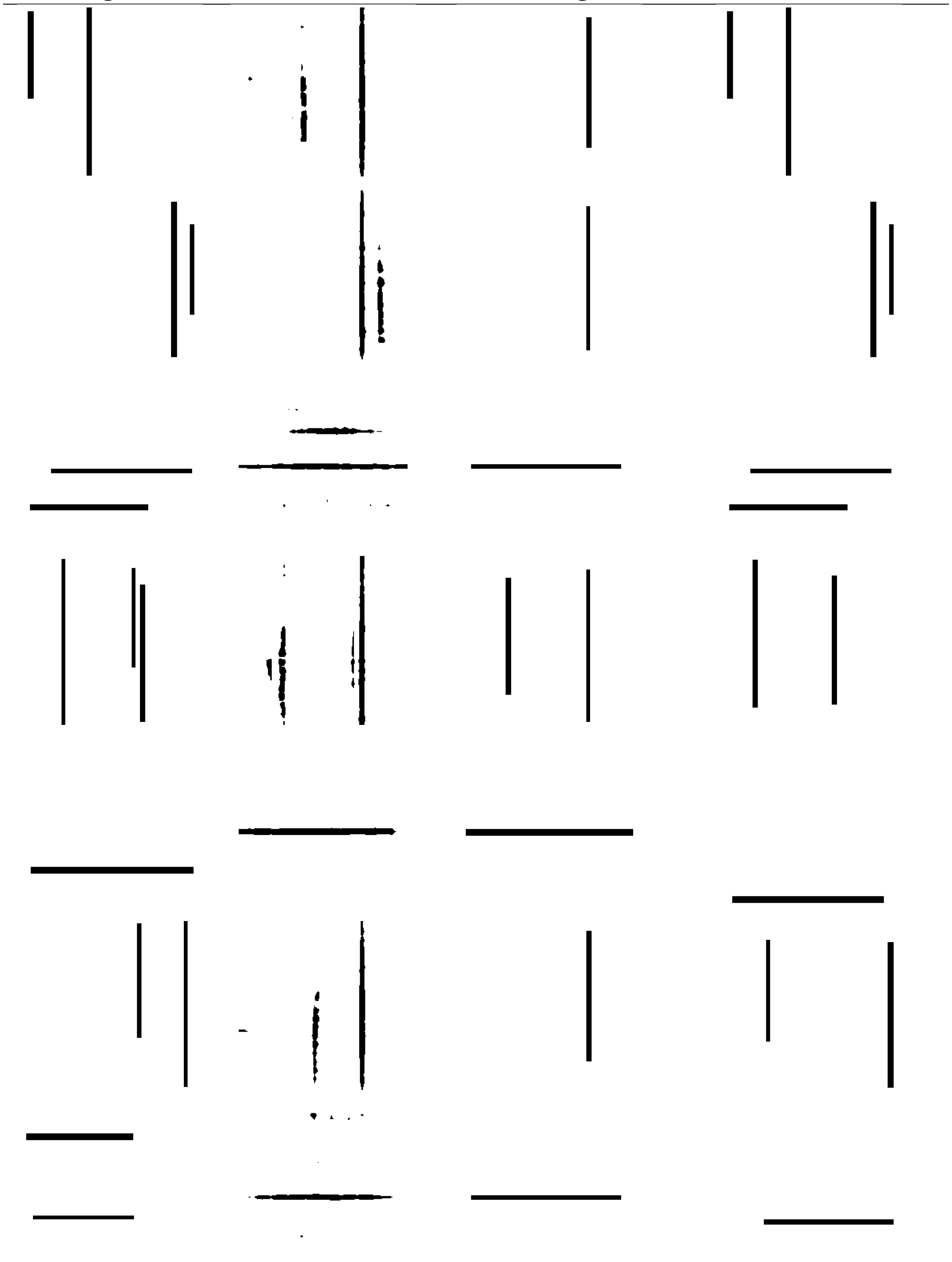
Data		**Most similar images**	
Original	**Reconstructed**	**Based on image MSE**	**Based on auto-corr. MSE**
			
			
			
			
			
			
			

Table 7: Original, reconstructed, and spatial statistics difference

Data		**Most similar images**	
Original	**Reconstructed**	**Based on image MSE**	**Based on auto-corr. MSE**



In the post-processing stage of the VAE reconstructions, a threshold value of 0.05 was utilized on the reconstructed tensors. This thresholding transforms grayscale images into binary ones, a step taken to improve the clarity of the output images. Crucially, this thresholding process preserves the integrity of the reconstructions, ensuring that the resulting binary images are appropriate for further analysis and comparison. For additional details, refer to the Appendix A2.