

# Deep Generative model that uses physical quantities to generate and retrieve solar magnetic active regions

SUBHAMOY CHATTERJEE <sup>1</sup>, ANDRÉS MUÑOZ-JARAMILLO,<sup>1</sup> AND ANNA MALANUSHENKO<sup>2</sup>

<sup>1</sup>*Southwest Research Institute, Boulder, CO 80302, USA*

<sup>2</sup>*High Altitude Observatory, NCAR, Boulder, CO, USA.*

## ABSTRACT

Deep generative models have shown immense potential in generating unseen data that has properties of real data. These models learn complex data-generating distributions starting from a smaller set of latent dimensions. However, generative models have encountered great skepticism in scientific domains due to the disconnection between generative latent vectors and scientifically relevant quantities. In this study, we integrate three types of machine learning models to generate solar magnetic patches in a physically interpretable manner and use those as a query to find matching patches in real observations. We use the magnetic field measurements from Space-weather HMI Active Region Patches (SHARPs) to train a Generative Adversarial Network (GAN). We connect the physical properties of GAN-generated images with their latent vectors to train Support Vector Machines (SVMs) that do mapping between physical and latent spaces. These produce directions in the GAN latent space along which known physical parameters of the SHARPs change. We train a self-supervised learner (SSL) to make queries with generated images and find matches from real data. We find that the GAN-SVM combination enables users to produce high-quality patches that change smoothly only with a prescribed physical quantity, making generative models physically interpretable. We also show that GAN outputs can be used to retrieve real data that shares the same physical properties as the generated query. This elevates Generative Artificial Intelligence (AI) from a means-to-produce artificial data to a novel tool for scientific data interrogation, supporting its applicability beyond the domain of heliophysics.

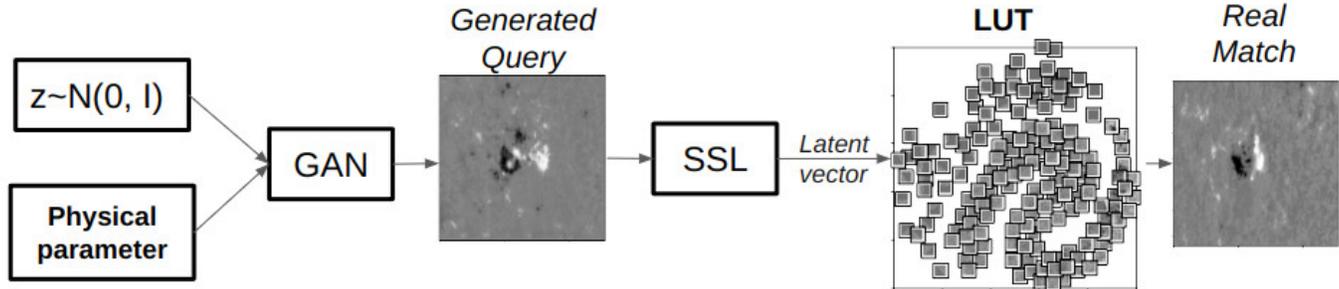
*Keywords:* Solar active regions (1974) — Solar magnetic fields (1503) — Support vector machine (1936)  
— Convolutional neural networks (1938)

## 1. INTRODUCTION

Modern astronomical observatories produce hundreds of petabytes of data during their lifetime. Manually labeling and sifting through such a scale of data is becoming impossible in a human lifetime. A bigger problem is to process these data and retrieve information hidden in those large datasets. Self-Supervised-Learning (SSL) (Chen & He 2020) has emerged as an effective machine learning approach to retrieve information by querying large unlabeled datasets. While it is easy to use an existing image as a query, sometimes the researcher may wish to find data for which there is no readily available physical reference. The ability to generate synthetic data with desired physical characteristics to be used as search query may greatly enhance scientific return from any dataset in question.

Generative Artificial Intelligence (AI) models use complex architectures to learn data-generating distributions and have been highly successful in producing synthetic data with relevant characteristics similar to those in real data. This makes them ideal for empowering users to find structures efficiently without sifting through all data. In this work we show how Generative AI created data can be used as a query for SSLs to retrieve matches from real data. As part of this process, we show how Generative AI has to be modified so that it can create data based on physical quantities.

One of the main challenges of using Generative AI in scientific applications is the obscure nature of hidden data-generating vectors within the model, also known as latent vectors. In Fourier space, for example, an image is uniquely represented by a set of coefficients in the space of the basis functions, which, in turn, correspond to different spatial frequencies, therefore the Fourier dimension space makes physical sense. But in the space that Generative AI operates



**Figure 1. Graphical illustration of the designed and implemented solar image generation and retrieval pipeline.** The process involves a generative model (GAN), supervised learning to utilize physical parameter space, and self-supervised learning (SSL) for reverse real image-search. The Look-Up-Table (LUT) describes the mapping between an image and SSL-derived latent representation.

in, the components of a latent vector that uniquely determines an image are non-interpretable, and also they interact with each other in a non-linear manner during the transformation that creates the image.

Our contribution focuses on addressing the lack of physicality in Generative AI that acts as a barrier to its adoption as a scientific tool. We use supervised learning to map desired physical characteristics related to the obscure latent spaces of generative models, enabling the use of physical quantities to generate or modify physical queries and find real matching observations. This addresses a need that will only become worse as more astronomical data are produced and elevates Generative AI from a means-to-produce artificial data to a novel tool for scientific data interrogation.

Our approach can be easily applied to many astronomical datasets such as emulating exoplanet transients and sifting through a large set of light curves for similar characteristics or generating astronomical objects (such as galaxies and nebulae) and finding similar examples in wide-field surveys.

The focus of our study is solar magnetic active regions (ARs). ARs are the major source of space weather events that impact Earth and drive variability in the solar system. As such, they are of major interest in solar physics. We use a set of physical parameters, used in numerous studies in Heliophysics to describe and study the properties of ARs.

Approaches exist that convert AR parameters to AR images and use those for data-driven simulations. For example, simulations of the solar dynamo use ARs that are modeled as idealistic simple dipoles. However, such dipoles lack the properties of real ARs such as shape, texture, etc. In this study, we make use of a state-of-the-art approach (Shen et al. 2019) utilizing Generative Adversarial Networks (GANs) (Goodfellow et al. 2014) in a supervised manner to modify generated images along directions that correspond to known AR parameters.

We design a pipeline (see Figure 1) that uses GAN-generated data as a query and utilizes a SimSiam (Chen & He 2020) SSL model to retrieve matches from observed data (which in our case are solar magnetic active regions). The data used, different components of the pipeline, and the findings are described in the remainder of this paper.

## DATA

We use Space-weather HMI Active Region Patches (SHARPs)(Bobra et al. 2014) for this work. We acquire about 44,000 patches of radial magnetic field on the solar surface (we use time sequences of **1839** unique patches). The patches are in cylindrical equal area projection. It must be noted that the SHARPs come in various sizes, and to make them ready as input to deep learning models we interpolate the patches to the same image size of  $128 \text{ pix} \times 128 \text{ pix}$ . We also scale the value of magnetic field, dividing by 1000 G, and clipping values to the range  $[-1, 1] \times 1000 \text{ G}$ .

## METHODS

**GAN.** GAN(Goodfellow et al. 2014) works as an adversarial two-player game between two modules, namely generator and discriminator, to learn the data-generating distribution ( $p_x$ ). With training, the generator ( $G$ ) becomes superior in generating unseen data ( $G(z)$ ) from just a noise vector ( $z$ , sampled from a standard normal distribution  $p_z$ ) that has properties of real data ( $x$ ) and the discriminator ( $D$ ) simultaneously improves in distinguishing generated data from real data. Training stops when the Nash equilibrium is reached, making the data-generating distribution indistinguishable from the generative distribution. The shape of  $G(z)$  is the same as the real data  $x$ .  $D$  outputs a probability with  $D(x)$  signifying the probability that  $x$  is a real data point. The optimization of  $G$  and  $D$  are performed as a two-step

process per training iteration by minimizing binary cross-entropy losses  $D_{loss}$  and  $G_{loss}$  respectively as described by the following equations-

$$G_{loss} = -\mathbb{E}_{z \sim p_z} \log D(G(z)) \quad (1)$$

$$D_{loss} = -\frac{1}{2} [\mathbb{E}_{z \sim p_z} \log(1 - D(G(z))) + \mathbb{E}_{x \sim p_x} \log D(x)] \quad (2)$$

We train GAN on Space-weather HMI Active Region Patches (SHARPs)(Bobra et al. 2014) solar magnetic patches. The parameters of  $D$  are not updated when  $G$  is trained. When  $D$  is trained the output of  $G$  is used and its optimization is stopped by disconnecting it from the model graph.

**SVM.** Support Vector Machines (SVMs)(Cortes & Vapnik 1995) fall under a class of supervised Machine Learning models that offer flexibility in dealing with high-dimensional datasets. They are used in both classification and regression. SVMs offer different kernel functions to draw decision boundaries for classification and use a subset of training data (also known as ‘support vectors’) to derive the decision boundary. We train a set of linear Support Vector Machines (SVMs) to learn the decision boundaries on latent space that separate regimes of higher and lower values of physical parameters related to magnetic patches.

**SimSiam.** Simsiam(Chen & He 2020) is a state-of-the-art self-supervised learning model that learns lower-dimensional representations (known as ‘latent vectors’) of images such that the representations are invariant across a specified set of augmentations. It works by augmenting an image twice and passing it through a backbone architecture shared by two branches. The backbone architecture is implemented by a Convolutional Neural Network – ResNet. The outcome of the backbone is then passed through a projection head in both branches and a prediction head in one of the branches. Multilayer perceptrons implement the projection and prediction heads. The outcome of projection ( $\vec{p}$ ) and prediction heads ( $\vec{z}$ ) from two branches are then compared by a negative cosine similarity loss  $L = -\frac{\vec{z} \cdot \vec{p}}{\|\vec{z}\| \|\vec{p}\|}$ . The model is trained by minimizing  $L$ . We train SimSiam on the SHARP data to learn 100-dimensional latent representations are invariant on translation, zoom, flip, and a restricted range of rotation.

## RESULTS

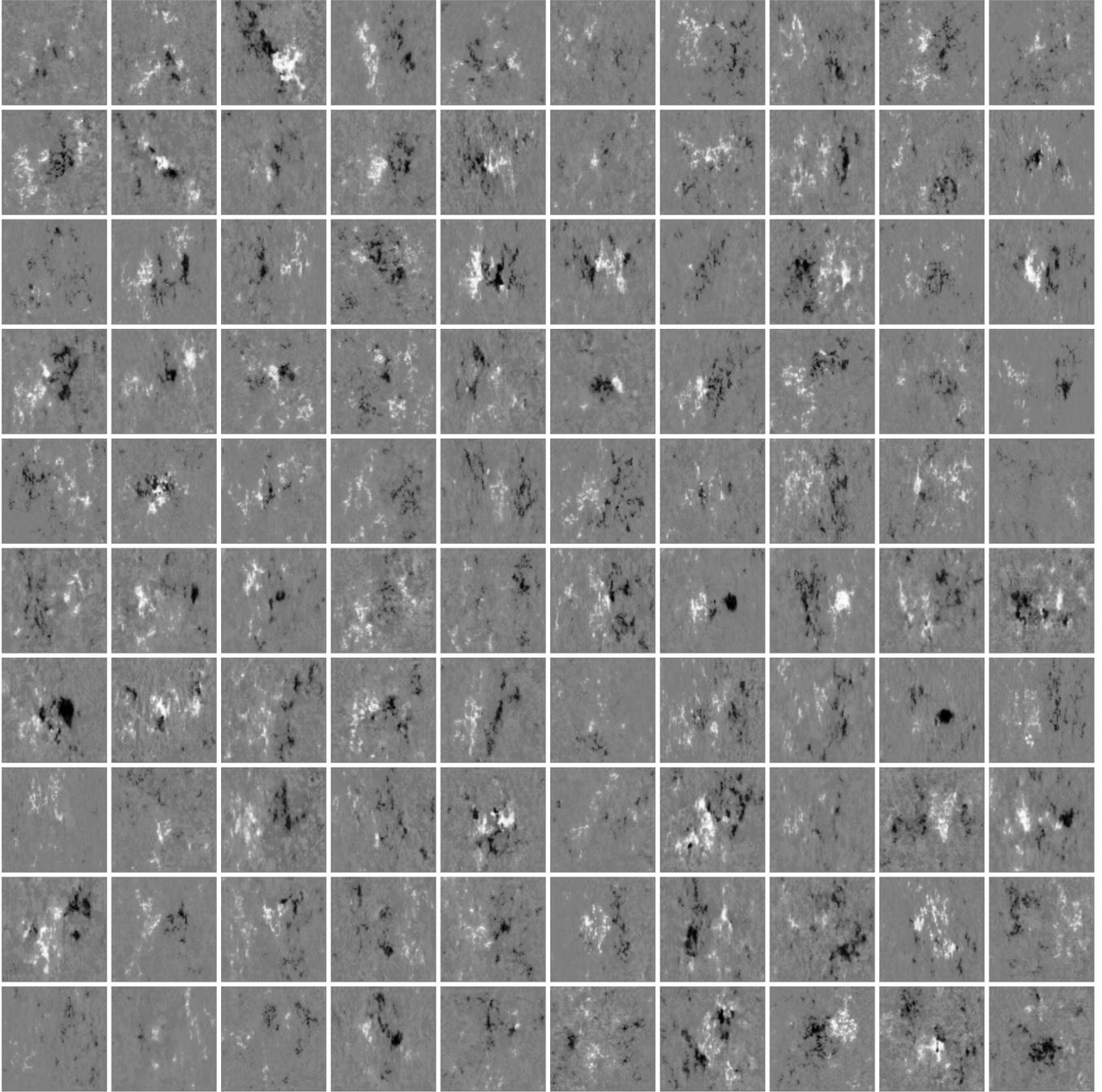
The results from our designed and implemented pipeline (Figure 1) are described sequentially in the following subsections.

**Training Generative Model.** We train a GAN to generate 128 pix  $\times$  128 pix magnetic patches from a 100-dimensional noise vector ( $z$ ; also referred to as ‘latent vector’ throughout the paper) sampled from a standard normal distribution. We implement the generator as a simple Convolutional Neural Network (CNN)-based decoder architecture and use tanh transformation as the last step of the generator architecture, thus making the generated images bounded within  $[-1, 1]$ . We construct a discriminator, based on a CNN-based encoder architecture that we built from scratch, to classify the generated magnetic patches as ‘fake’ and SHARP patches as ‘real’. The last layer in the discriminator is a sigmoid-function transformation.

We train the GAN for 200 epochs with a batch size of 32 and a learning rate of 0.0005. For further inference, we pick the epoch for which the discriminator shows the least accuracy in distinguishing generated data from real. Figure 2 illustrates the results from our GAN as a 10 $\times$ 10 matrix of generated magnetic patches. We note that the flux in generated images ranges from low to high. This is because GAN is trained on time sequences of magnetic patches during their emergence and evolution.

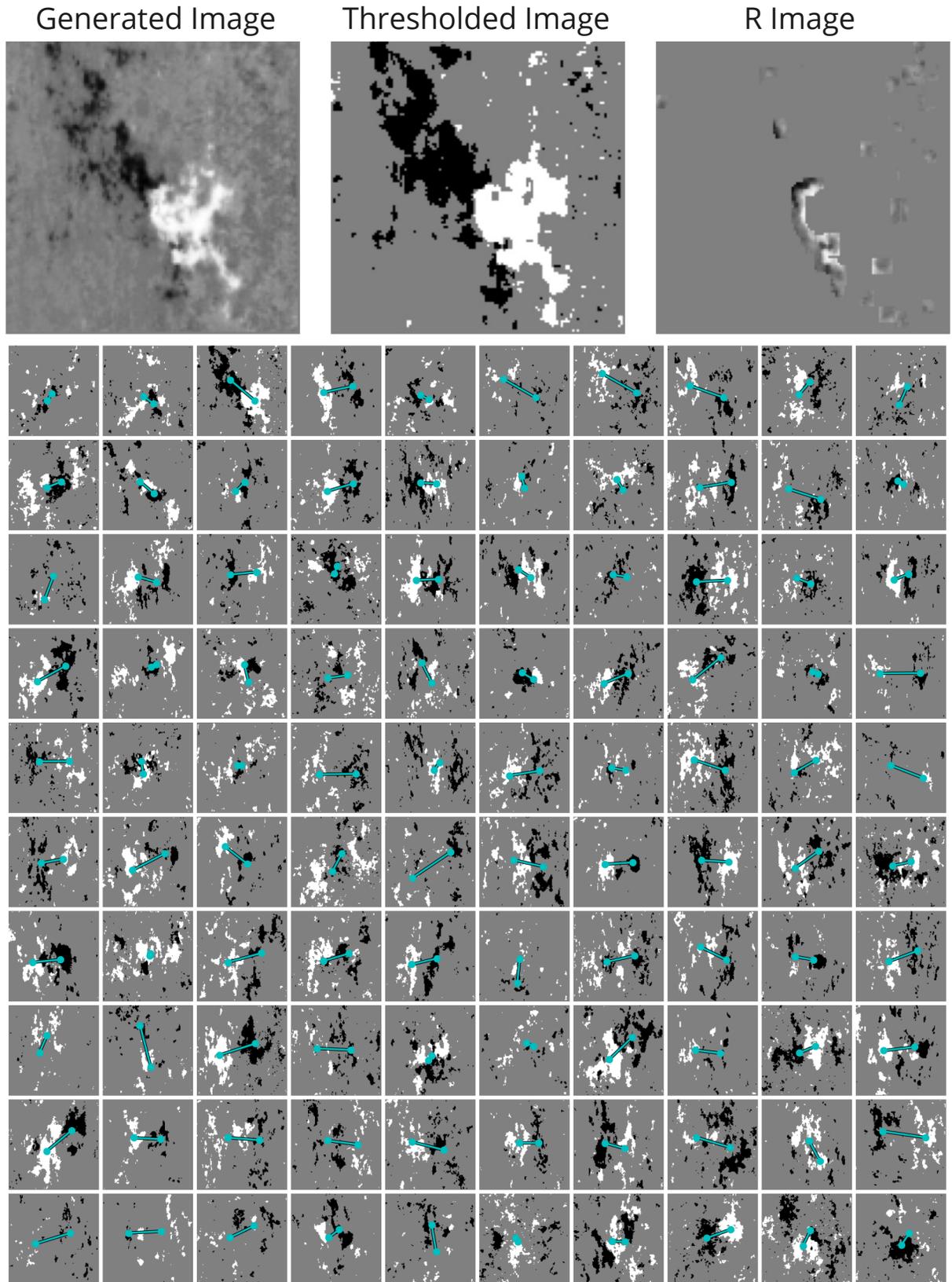
**Interpretable Generation of Magnetic Patches.** Although the GAN can generate realistic solar magnetic patches, interpretable manipulation of the images using the embedding space is non-trivial. Gradually changing each value of each dimension in noise vector  $z$  gives rise to smooth changes in the generated images, but those changes are often not physically interpretable, as multiple features tend to evolve simultaneously. To control the generation in physically meaningful space we found it is most useful to use a supervised approach to learn decision boundaries in the latent space of GAN. The decision boundary, for a given physical parameter, separates low and high values of the parameter calculated from GAN-generated images. This approach is illustrated in the following paragraphs.

We first generate 10,000 magnetic patches using the GAN using randomly sampled noise vector  $z$ . We then calculate physical parameters such as the total unsigned field strength( $TUF$ ), pole-to-pole separation ( $PSEP$ ), and the parameter  $R$ .  $R$  represents a total unsigned field in subregions with a high gradient of the field with polarity transition



**Figure 2. A zoo of GAN-generated magnetic Active Regions that depending on complexity play an important role in driving space weather events.** The shown AR images are generated using 100 latent vectors randomly sampled from a standard normal distribution. The images with diffused flux are seen as the temporal evolution of SHARPs is used during the training of the GAN.

(Figure 3). Following a paper by Schrijver (Schrijver 2007), we calculate  $R$  by dilating the pixels with field strength above 150 G or below -150 G, and use the intersection of those regions as a mask for high gradient regions in the patches. The unsigned field is then integrated within these regions to get the value of  $R$ . Using the same threshold of  $\pm 150$  G we also calculate  $PSEP$  as the distance between field strength-weighted centroids of pixels within positive and negative polarity regions (shown as cyan lines in Figure 3). To calculate  $TUF$  we sum unsigned field strength over all the image pixels.



**Figure 3.** Calculation of physical parameters from thresholded GAN generated magnetic patches. The top row shows a GAN-generated active region, its thresholded version, and the derived regions of polarity inversion with a high field gradient from left to right. The collage starting from the second row shows the thresholded version of all GAN-generated images in Figure 2 with field-strength-weighted centroids of positive and negative polarity connected by a cyan line.

We make a lookup table of the 100-dimensional noise vector  $z$  vs. the calculated parameters for all 10,000 generated images and use the median of each parameter as a threshold to assign a binary label. Using these binary labels we train a Support Vector Machine (SVM) to learn the decision boundary on the 100-dimensional latent space that separates regimes of higher and lower values of the selected parameters. We use 7,000 data points for training and 3,000 for validation. We acquire the normal vectors ( $\hat{\mathbf{n}}^p$ ) for the decision boundaries of three trained SVMs ( $p \in \{\text{TUF}, \text{PSEP}, \text{R}\}$ ).

To visualize the latent space and the SVM-derived decision boundary in 2-dimensions described by  $(z'_1, z'_2)$  we project the 100-dimensional space  $(z_1, z_2, \dots, z_{100})$  using a simple transformation:  $z'_1 = z_1$  and  $z'_2 = \frac{1}{n_2} \sum_{i=2}^{100} n_i z_i$ , where  $(n_1, n_2, \dots, n_{100})$  defines the normal ( $\hat{\mathbf{n}}^p$ ) to the hyperplane defining the decision boundary and  $n_2 \neq 0$ . This transformed the hyperplane to a line in the visualized 2-dimensions. In Figure 4, one can observe a clear demarcation between physical parameters represented by images above and below the decision line.

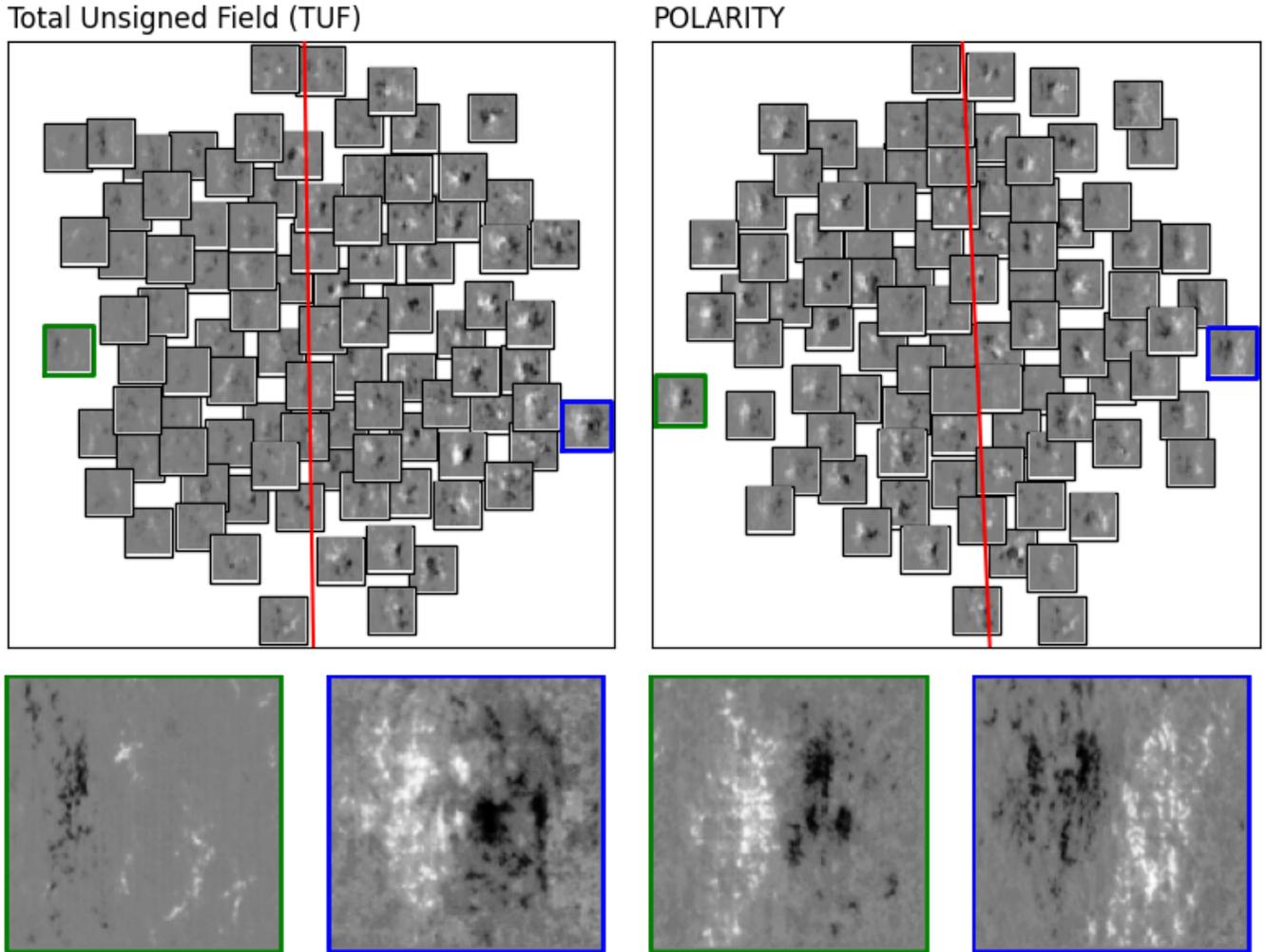
Given a latent vector ( $\mathbf{z}$ ) we shift by a quantity  $\epsilon$  using  $\mathbf{z}_{shift} = \mathbf{z} + \epsilon \hat{\mathbf{n}}$  along the directions estimated by training SVMs. We use  $\mathbf{z}_{shift}$  to generate images for different values of  $\epsilon$  and find the generated images to reflect the change in the physical parameter encoded by  $\hat{\mathbf{n}}$  (see Figure 5). However, physical properties are often entangled, and change in one can cause change in another. As can be observed from Figure 5, positive shifts along  $\hat{\mathbf{n}}_U$  causes an increase in  $R$  (new high gradient polarity inversion region) in generated images. This effect can be partially mitigated through conditional manipulation (Shen et al. 2019). We use a new direction subtracting the projection along  $\hat{\mathbf{n}}^{TUF}$  given by:  $\hat{\mathbf{n}}^{new} = \frac{\hat{\mathbf{n}}^{TUF} - (\hat{\mathbf{n}}^{TUF} \cdot \hat{\mathbf{n}}^R) \hat{\mathbf{n}}^R}{|\hat{\mathbf{n}}^{TUF} - (\hat{\mathbf{n}}^{TUF} \cdot \hat{\mathbf{n}}^R) \hat{\mathbf{n}}^R|}$ . The second row of Figure 5 depicts the change in TUF minimizing the change in  $R$ .

For more than one conditional boundary we use orthogonal decomposition of primal direction  $x$  onto subspace  $W$  constructed by the column space of the matrix  $A$ , which consists of conditional boundary vectors as columns. The vector perpendicular to the subspace  $W$  takes the generic form:  $x - A(A^T A)^{-1} A^T x$ . The third and fourth rows of Figure 5 depict the polarity flip and the change in  $R$  with the change in PSEP. The sixth row of Figure 5 shows the effect of conditionally manipulating the change in  $PSEP$  over polarity flip and  $R$ . It shows no polarity flip as  $PSEP$  increases and a minimized change in  $R$ .

**Querying Generated Image to Find Matches in Real Data.** Although GAN-generated images look realistic, they should possess physical properties comparable with real data. We train a Self-Supervised Learning (SSL) model called SimSiam on the resized SHARPs that learns a 100-dimensional latent space that is invariant across augmentations namely shift (within 10 pixels), zoom (within a factor of 0.8 to 1.2), rotation (within  $-20^\circ$  to  $20^\circ$ ), vertical, and horizontal flip. The images and their augmented version are converted to augmentation-invariant latent vectors through two identical CNNs. Once trained, we find the latent representation of the query image and use the negative cosine distance as a metric to find the nearest neighbor in the latent space. The image corresponding to the nearest neighbor in latent space is retrieved as the best match from the set of real images for a query image generated by GAN. We randomly generate 1000 images and retrieve the best-matching real images using the SSL. We compare the distributions of different physical parameters of generated images and real matches. Figure 6 illustrates the retrieval results for a generated query of a bipolar magnetic patch both in its original form and polarity flipped version. The retrieved images seem to conserve the overall polarity arrangement, as well as finer-scale structures. The 2-dimensional histogram of TUF shows a one-to-one correspondence (Pearson correlation  $\sim 0.78$ , Spearman rank correlation  $\sim 0.79$ ) between the unsigned field calculated from 1000 generated queries, and corresponding real matches (Figure 6). The one-dimensional histograms of TUF in Figure 6 depict the match between the generated images, retrieved SHARPs, and the entire space of SHARPs. This validates that our GAN can capture the entire TUF regime of real magnetic patches through its generation and retrieval. To characterize the correspondence further we calculate the integrated field strength within positive (TPF) and negative (TNF) polarity regions. We observe statistically significant positive Pearson (Spearman) correlations of 0.73 (0.76), 0.28 (0.31), 0.69 (0.7), and 0.72 (0.72) between generated and closest real images for the parameters  $R$ ,  $PSEP$ ,  $TPF$ , and  $TNF$  respectively. The relatively poor correlation in  $PSEP$  comes from the occurrence of the small-scale features and complex multipolar regions that drive the calculation of field-strength-weighted centroids for each polarity. This does not indicate an intrinsic morphological mismatch between the general query and the real match. A better polarity separation measure should cause an improvement in the  $PSEP$  correlation.

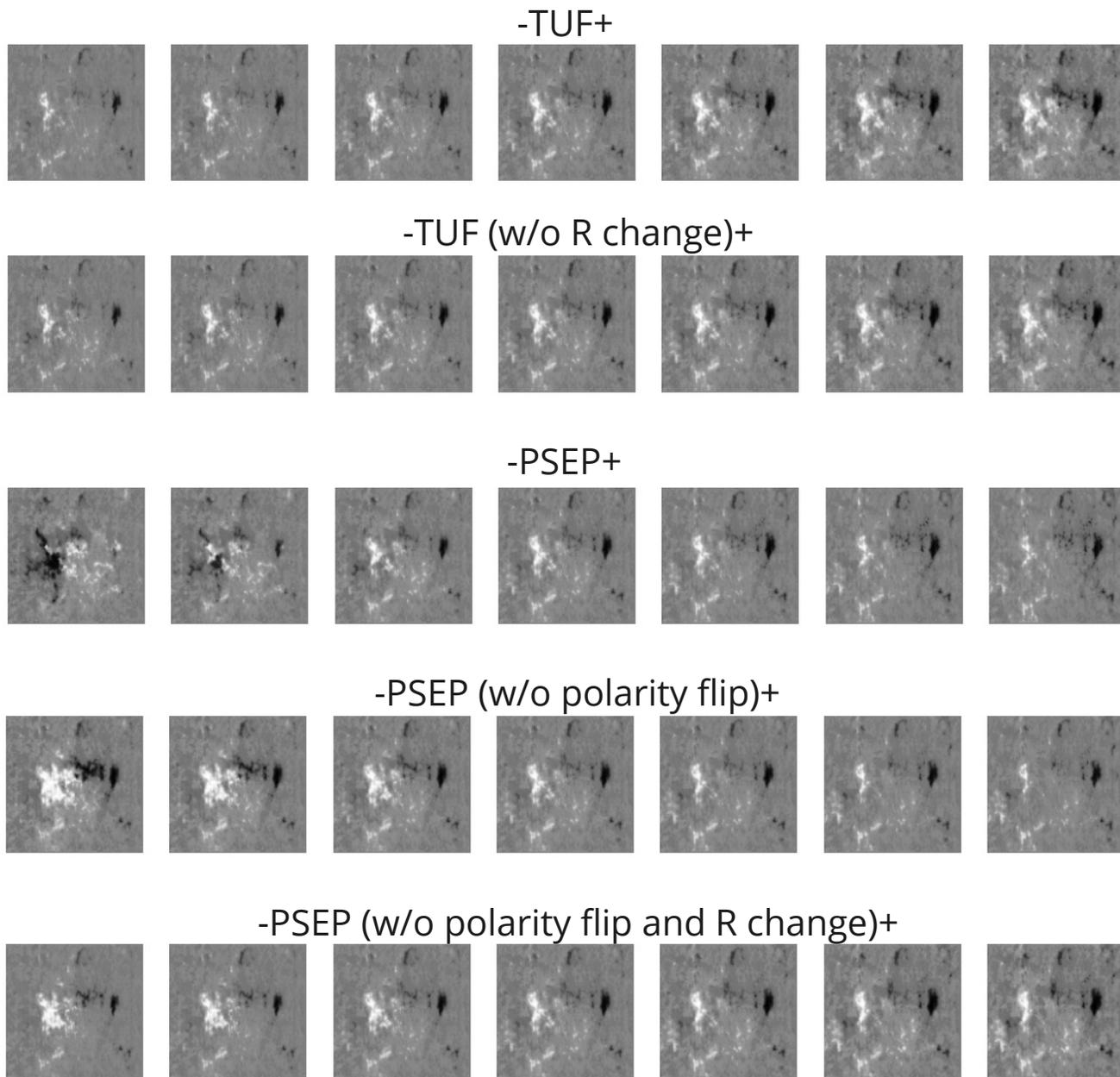
## DISCUSSION AND CONCLUSION

We demonstrated that a Deep Generative model can be harnessed to generate scientifically meaningful queries to find matches in scientific datasets. We achieved this by applying three types of machine learning models — Generative



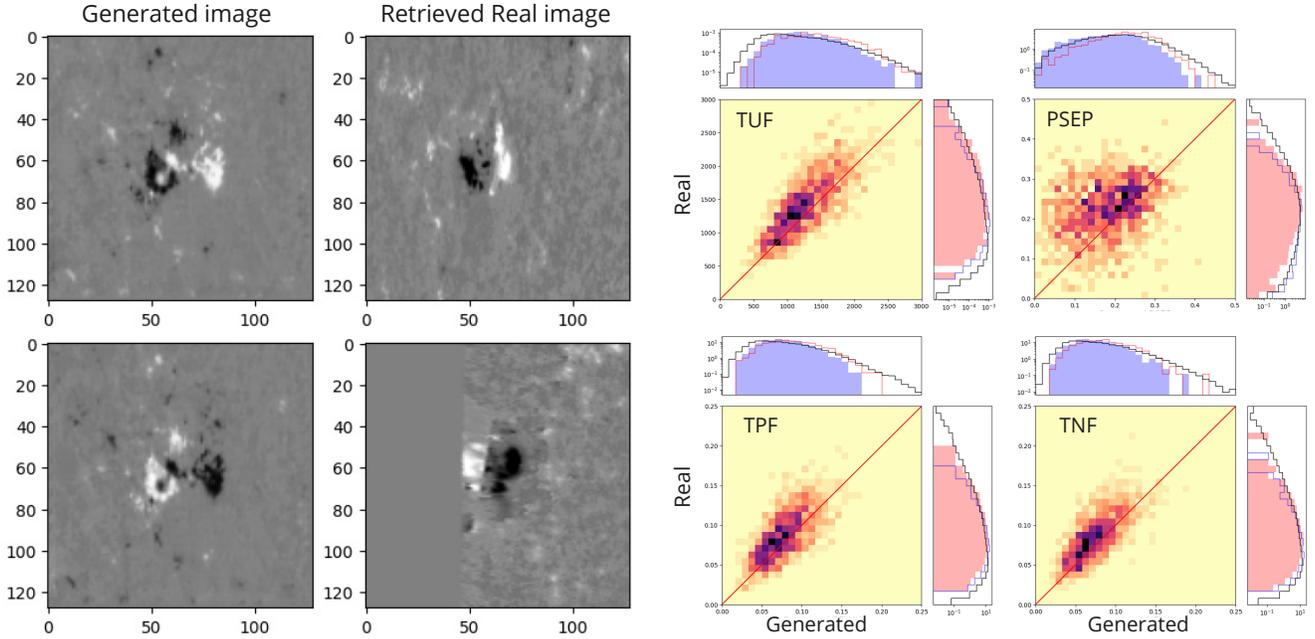
**Figure 4. Generated solar images overlaid on the 2-dimensional projection of latent space.** The projection was performed such that the decision boundary hyperplane becomes a line (marked in red) in the depicted 2d latent space. The decision boundary marks a clear separation between the physical properties (total unsigned field (TUF), and polarity) represented by the generated images. The bottom rows show zoomed-in views (outlined with green and blue rectangles) of the generated images from both extremes across the decision boundaries.

Adversarial Networks (GANs), Support Vector Machines (SVMs), and a SSL model SimSiam, on the solar magnetic patch dataset, SHARP. Whereas GANs enabled a generation of magnetic active region images from 100-dimensional latent vectors, SVMs helped connect physical parameters (such as total unsigned magnetic field, polarity separation, and total field in the vicinity of polarity inversion lines with high gradients) to the generative latent vectors of GAN. We found that this image generation can be tweaked along directions that reflect changes in the mentioned physical quantities. We showed that the generated images can also be used as a query to the SSL-derived latent space to retrieve matches from real magnetic regions. These retrieved regions matched the query both visually and quantitatively. This approach thus elevates generative models from a means-to-generate-synthetic data to a novel tool for the efficient mining of real scientific data. Even though we demonstrated this approach in a specific domain of solar astronomy, it can be easily adapted to any other field of astronomy dealing with big datasets of any modality and complexity. In the future, we would like to improve our software to use numerical values of physical parameters (e.g., calibrate



**Figure 5. Manipulation of GAN generated image along two different physically interpretable directions.** Two directions are shown – total unsigned field (TUF; top two rows) and polarity separation (PSEP; bottom three rows). The 2nd row from the top shows the result of decoupling TUF from the direction along which the R changes. The 4th row shows the result of decoupling PSEP from the direction along which polarity is flipped and the bottom-most row shows the result of decoupling PSEP from both polarity flip and change in R.

the change in magnetic-flux-influencing parameter to the value in actual physical units in Mx). This will enable more precise control of the generation of queries.



**Figure 6. Retrieval and comparison of physical parameters with real nearest neighbors.** The left-most column shows two generated bipolar active regions with opposite polarity. The 2nd column from the left shows the nearest matching real images to the respective generated queries. The 3rd and 4th columns from the left show the 2D histogram density of the 4 physical parameters (TUF, PSEP, TPF, and TNF) calculated from 1000 randomly generated images and the nearest real neighbors. The marginal histograms of generated and matching real image parameters are plotted in light blue and light pink respectively. The black curves overlaid on the marginals represent the histogram of respective physical parameters evaluated across all real images in the dataset. The marginals for generated images are plotted over real and vice-versa for comparison in respective colors.

## 2. DATA AND CODE AVAILABILITY

The SHARPs, used to train the generative model for this study, are available from the Joint Science Operations Center (<http://jsoc.stanford.edu>).

Our developed codebase is publicly available via a GitHub repository: [https://github.com/subhamoysgit/Generative\\_Interrogator](https://github.com/subhamoysgit/Generative_Interrogator).

## ACKNOWLEDGMENTS

This research was funded by NASA HGIO grant 80NSSC23K0416.

## REFERENCES

- Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, *Solar Physics*, 289, 3549, doi: [10.1007/s11207-014-0529-3](https://doi.org/10.1007/s11207-014-0529-3)
- Chen, X., & He, K. 2020, arXiv e-prints, arXiv:2011.10566, doi: [10.48550/arXiv.2011.10566](https://doi.org/10.48550/arXiv.2011.10566)
- Cortes, C., & Vapnik, V. 1995, *Machine learning*, 20, 273, doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018)
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv e-prints, arXiv:1406.2661, doi: [10.48550/arXiv.1406.2661](https://doi.org/10.48550/arXiv.1406.2661)
- Schrijver, C. J. 2007, *The Astrophysical Journal*, 655, L117, doi: [10.1086/511857](https://doi.org/10.1086/511857)
- Shen, Y., Gu, J., Tang, X., & Zhou, B. 2019, arXiv e-prints, arXiv:1907.10786, doi: [10.48550/arXiv.1907.10786](https://doi.org/10.48550/arXiv.1907.10786)