Combining Physics-Derived and Machine-Learned Features for Probabilistic Solar Flare Forecasting

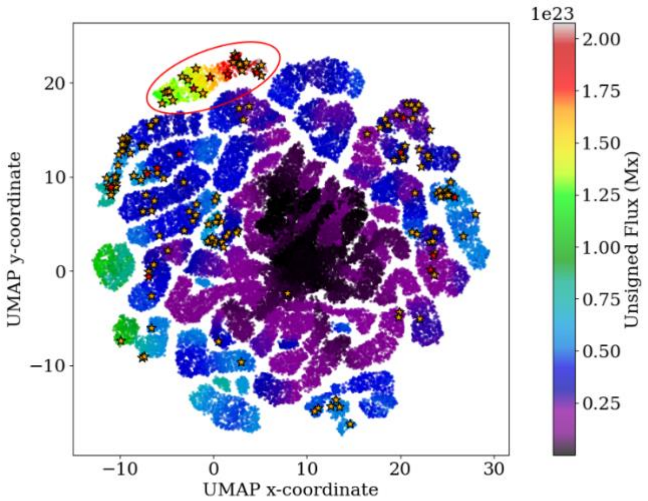

Figure 1. UMAP provides a convenient low-dimensional embedding of the 6D latent vectors extracted by the VAE. Each point corresponds to a SHARP magnetic field map represented in latent space, color-coded by total unsigned flux. Flare events are marked with stars. The embedding shows well-defined clusters for individual regions, with higher-flux regions located toward the periphery of the map.

Solar activity is the primary driver of space weather variability, with solar flares being among the most disruptive phenomena due to their sudden onset and immediate geoeffective impacts. Traditional forecasting methods rely on physics-based parameters derived from solar observations, i.e., Space-weather HMI Active Region Patches (SHARP) dataset. In recent years, machine learning (ML) and deep learning (DL) methods have shown promise for extracting structure from complex, high-dimensional solar datasets. Variational Autoencoders (VAEs), in particular, provide probabilistic and compact representations of images, making them attractive for application in solar activity studies. However, machine-learned features alone often lack interpretability. We propose a hybrid strategy that combines the **predictive power of SHARP parameters** with **disentangled features learned from SDO/HMI magnetograms using β-VAEs**, aiming to leverage the best of both approaches. Our work uses time series of vector magnetic field cutouts from the SHARP data as input to a disentangled β-VAE. The β-VAE produces factorized latent space where each feature corresponds to distinct generative aspects of the active region's magnetic morphology or dynamics. From each magnetogram, only six latent features are extracted, providing a compact but information-rich description per frame. Benchmarking against SHARP parameters such as unsigned flux (see Figure 1) and helicity shows that the learned features evolve in parallel with empirical indicators but also capture additional variability, demonstrating their complementary value. The forecasting pipeline integrates both SHARP and VAE features. A Long Short-Term Memory (LSTM) network models the temporal evolution of the features across different time windows, followed by logistic regression layers to produce probabilistic flare forecasts. Both binary (flare vs. no flare; Alert (M,X class) vs. No Alert (C class)) and multi-class (C, M, X class) probabilistic predictions are generated. Results show that the **combined dataset outperforms either SHARP or VAE features alone**. The inclusion of machine-learned features reduces false alarms and provides better reliability in probabilistic forecasts. Importantly, the disentangled latent space ensures computational efficiency, enabling operational use without prohibitive resource demands. Moreover, the empirical relationships identified between VAE features and physical parameters ensure that a consistently informative input set will be available for predictive models, even in situations where SHARP parameters may be missing, such as in near–real-time data. In addition, latent space parametrization can be extended to provide a unified representation of SDO/AIA EUV observations, which are essential for developing a robust and accurate flare prediction framework. Overall, this study demonstrates that combining physics-derived parameters with machine-learned representations enables a comprehensive characterization of active region magnetic field properties. Integrating empirical SHARP parameters with disentangled VAE features significantly improves probabilistic flare forecasting skill, underscoring the value of bridging physical knowledge and data-drive ML space weather prediction.