**Table 2:** Evaluation of Wasserstein-Filtered Data Augmentation on ISIC 2018 Dataset. Baseline: original samples (averaged across tasks); Augmented: unfiltered generated data; Wass: Wasserstein-filtered data, retaining the top 60% of images.

Gen	Model	Acc	Prec	Rec	<b>F</b> 1	Gen	Model	Acc	Prec	Rec	F1
3	Augmented Wass			60.00 <b>62.86</b>	57.79 <b>61.07</b>	18	Augmented Wass			48.57 <b>58.57</b>	
6	Augmented Wass			45.71 <b>57.14</b>		21	Augmented Wass			47.14 <b>64.29</b>	
9	Augmented Wass			50.00 <b>55.71</b>		24	Augmented Wass		52.91 <b>65.37</b>	55.71 <b>64.29</b>	51.67 <b>63.91</b>
Bas	seline (Avg.	) 52.32	56.64	52.32	51.88						

Performance metrics (%) on a 7-class skin cancer image dataset (ISIC 2018) [1] with 1,257 original training samples and varying numbers of generated images (Gen) from SD-XL, mixed at strength=0.15 and strength=0.8 (default: 0.75). Higher strength increases diversity but introduces suboptimal samples, requiring Wass filtering. Wass consistently enhances performance over the baseline.

[1] Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review