

**Table 1:** Comparative Performance of Data Filtering Methods on CIFAR10

Gen	Model	Acc	Prec	Rec	F1	Gen	Model	Acc	Prec	Rec	F1
<b>2</b>	Baseline	38.55	38.48	38.55	38.41	<b>12</b>	Baseline	38.18	38.44	38.18	38.02
	Wass	<b>42.34</b>	<b>42.47</b>	<b>42.34</b>	<b>42.07</b>		Wass	<b>43.88</b>	<b>43.67</b>	<b>43.88</b>	<b>43.55</b>
	MMD	42.14	41.71	42.14	41.63		MMD	43.43	42.95	43.43	43.02
	TV	39.93	40.20	39.93	39.85		TV	42.70	42.24	42.70	42.27
<b>4</b>	Baseline	39.21	39.31	39.21	38.95	<b>14</b>	Baseline	37.85	37.72	37.85	37.50
	Wass	42.37	42.44	42.37	<b>42.24</b>		Wass	42.28	42.34	42.28	42.01
	MMD	<b>42.90</b>	<b>42.86</b>	<b>42.90</b>	42.67		MMD	42.74	42.31	42.74	42.33
	TV	42.27	42.26	42.27	42.09		TV	<b>44.10</b>	<b>43.69</b>	<b>44.10</b>	<b>43.76</b>
<b>6</b>	Baseline	38.86	38.96	38.86	38.76	<b>16</b>	Baseline	39.65	40.11	39.65	39.45
	Wass	41.97	42.12	41.97	41.79		Wass	<b>44.51</b>	<b>44.28</b>	<b>44.51</b>	<b>44.27</b>
	MMD	<b>44.65</b>	<b>44.52</b>	<b>44.65</b>	<b>44.32</b>		MMD	43.80	43.27	43.80	43.29
	TV	42.75	43.29	42.75	42.73		TV	43.74	43.16	43.74	43.27
<b>8</b>	Baseline	39.12	38.87	39.12	38.59	<b>18</b>	Baseline	38.85	39.03	38.85	38.85
	Wass	43.82	43.63	43.82	43.42		Wass	43.79	43.53	43.79	<b>43.56</b>
	MMD	<b>44.42</b>	<b>44.24</b>	<b>44.42</b>	<b>44.22</b>		MMD	43.06	42.99	43.06	42.79
	TV	42.84	42.66	42.84	42.35		TV	<b>44.22</b>	<b>43.93</b>	<b>44.22</b>	43.74
<b>10</b>	Baseline	38.14	38.00	38.14	37.90	<b>20</b>	Baseline	38.92	38.93	38.92	38.67
	Wass	43.93	43.79	43.93	43.63		Wass	43.52	43.15	43.52	43.15
	MMD	43.79	<b>44.00</b>	43.79	<b>43.67</b>		MMD	<b>43.97</b>	<b>43.45</b>	<b>43.97</b>	<b>43.27</b>
	TV	42.89	42.53	<b>42.89</b>	42.53		TV	42.59	42.41	42.59	42.30

Performance metrics (%) on CIFAR-10 with 1,000 training samples and varying numbers of generated images (Gen) from SD-XL at **strength=0.3**. Data augmentation uses Wass (Wasserstein), MMD (Maximum Mean Discrepancy), and TV (Total Variation) metrics, retaining the top 60% of images. Baseline: original samples (averaged across Gen); Augmented: unfiltered generated data; Wass, MMD, TV: filtered by respective metrics. Metrics: Acc (Accuracy), Prec (Precision), Rec (Recall), F1 (F1-Score). Augmentation improves over the baseline, with Wass, MMD, and TV showing comparable performance.