

Corruption Type	Standard (CNN)	TTA (CNN)	Standard (GMU-CNN)	TTA (GMU-CNN)
Brightness	0.2274	0.9887	0.9913	0.9933
Canny Edges	0.6149	0.9431	0.8998	0.9813
Dotted Line	0.9791	0.9873	0.9896	0.9902
Fog	0.1188	0.6074	0.9234	0.9609
Impulse Noise	0.4529	0.9098	0.9377	0.9683
Motion Blur	0.9675	0.9858	0.9614	0.9666
Shot Noise	0.9226	0.9677	0.9449	0.9757
Spatter	0.9834	0.9857	0.9759	0.9755
Zigzag	0.7826	0.8707	0.9447	0.9608
Average	0.67213333	0.916244	0.952078	0.974733

Table 1: Accuracy comparison of standard (no TTA) and Test-Time Adaptation (TTA) approaches across various corruption types on the MNIST dataset.

Corruption Type	Standard (CNN)	TTA (CNN)	Standard (GMU-CNN)	TTA (GMU-CNN)
Brightness	0.4535	0.8281	0.8250	0.8550
Canny Edges	0.3709	0.5322	0.7058	0.7575
Dotted Line	0.8786	0.8953	0.9118	0.9161
Fog	0.2712	0.6060	0.7442	0.8373
Impulse Noise	0.2056	0.6614	0.6703	0.7837
Motion Blur	0.7188	0.7865	0.6831	0.7830
Shot Noise	0.5959	0.7059	0.4896	0.6870
Spatter	0.8835	0.8902	0.8868	0.8908
Zigzag	0.8131	0.8606	0.8806	0.8941
Average	0.6159	0.7518	0.7558	0.8227

Table 2: Accuracy comparison of standard (no TTA) and Test-Time Adaptation (TTA) approaches across various corruption types on the Fashion-MNIST dataset.

Corruption Type	Standard (VGG)	TTA (VGG)	Standard (GMU(3))	TTA (GMU(3))	Standard (GMU(8))	TTA (GMU(8))
Brightness	0.83642	0.82376	0.8540	0.85266	0.86188	0.86362
Contrast	0.5952	0.74014	0.7489	0.8016	0.75478	0.81086
Defocus Blur	0.68148	0.79334	0.7811	0.82878	0.7977	0.83792
Elastic Transform	0.71436	0.74232	0.7493	0.7808	0.7534	0.78318
Fog	0.74422	0.77798	0.8176	0.8283	0.8311	0.8393
Frost	0.7386	0.7386	0.7410	0.77508	0.7620	0.78332
Gaussian Blur	0.5929	0.77528	0.7410	0.81844	0.7620	0.82662
Impulse Noise	0.59338	0.65226	0.6003	0.75382	0.5608	0.73398
Motion Blur	0.6211	0.74056	0.7142	0.79238	0.7325	0.79896
Pixelate	0.6826	0.7591	0.7192	0.80666	0.7273	0.81328
Saturate	0.8147	0.81246	0.8211	0.8283	0.8298	0.83882
Shot Noise	0.6366	0.72318	0.6372	0.77136	0.6163	0.76632
Spatter	0.7440	0.73732	0.7861	0.79432	0.7874	0.7989
Average	0.691966	0.7551	0.747	0.8025	0.752074	0.807314

Table 3: Accuracy comparison of standard (no TTA) and Test-Time Adaptation (TTA) approaches across various corruption types on the CIFAR-10 dataset.