

1 The Surface Defect Inspection Dataset

The Surface Defect Inspection (SDI) dataset consists of 20,414 images at various resolutions. It contains three kinds of products—product A, product B and product C. Each can be further classified into three foreground defect domains—Normal, Scratches and Spots. Fig. 6 shows example images of the SDI dataset. It is worth mentioning that the dataset is highly imbalanced not only between normal and defective samples but also between different products as shown in Table 6. This sets a more challenging task when training deep neural networks like GANs and downstream classifiers.

For each foreground defect and background product, we randomly select 50 images, 30% of which are used for validation and 70% as the test set. All remaining images are used as training sets for GAN and classifier training. We present the distribution of the training set when training DT-GAN in Table 7. Note that the normal samples used in GAN training are only a subset of all available samples in Normal and we keep the rest of them for generating defective samples at test time. For classifier training, we show the statistics in Table 8, where the number of normal samples involved in classifier training increases incrementally. The validation set is used to select the best model during classifier training while the test set is left untouched until the final evaluation. Both the validation and test set are inaccessible by DT-GAN.

Table 6: Distribution of the full SDI dataset.

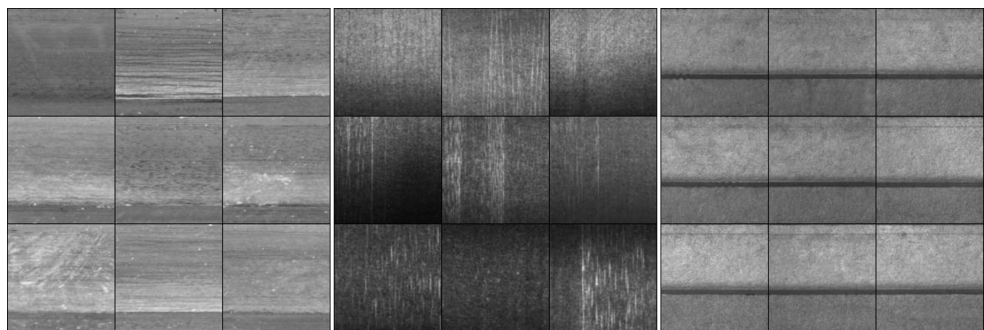
	Overview		
	A	B	C
Normal	6,250	6,250	6,250
Scratches	340	167	121
Spots	108	670	258

Table 7: The training set for DT-GAN and the baseline image synthesis methods.

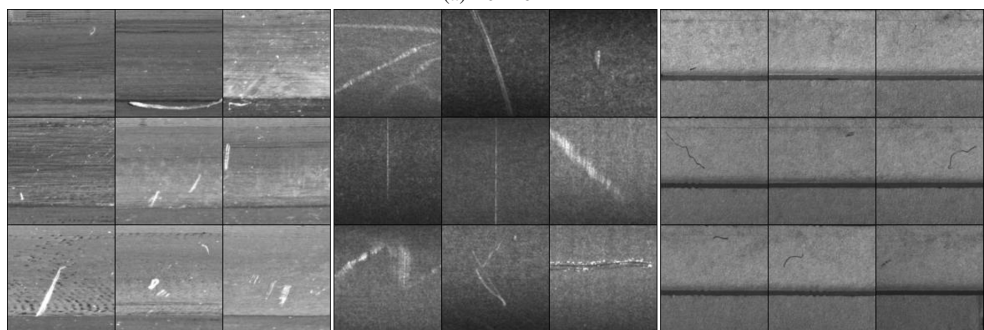
	Overview		
	A	B	C
Normal	700	700	700
Scratches	290	117	71
Spots	58	620	208

Table 8: The training, validation, and test set for classifier training, where N increases incrementally—1,500, 2,200, 4,000, and 6,200.

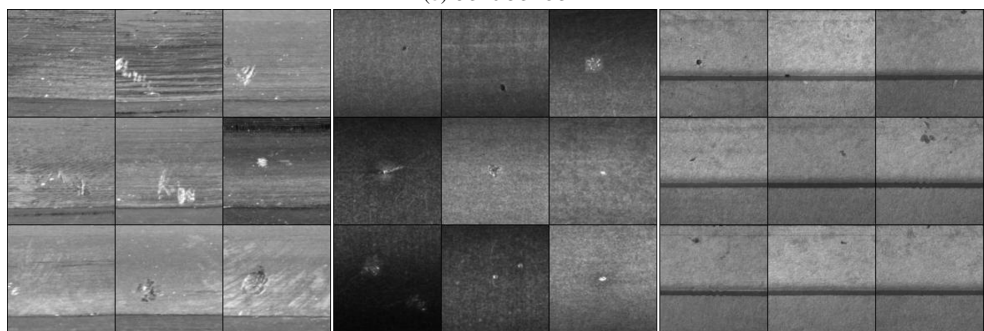
	Train			Validation			Test		
	A	B	C	A	B	C	A	B	C
Normal	N	N	N	12	18	15	38	32	35
Scratches	290	117	71	14	16	15	36	34	35
Spots	58	620	208	14	16	15	36	34	35



(a) Normal



(b) Scratches



(c) Spots

Figure 6: Overview of the SDI dataset. (Best viewed in color.)