CHARLOTTE

How to pick a metric to classify a surface?

Useful resources:

- B. Mullany, E. Savio, H. Haitjema, R. Leach, “The implication and evaluation of geometrical imperfections on manufactured surfaces “, Annals of the CIRP, Vol. 71 (2), pp. 717-739, 2022
- J. Redford, B. Mullany, “Construction of a Multi-Class Discrimination Matrix and Systematic Selection of Areal Texture Parameters for Quantitative Surface and Defect Classification”, Journal of Manufacturing Systems, Volume 71, December, Pages 131-143, 2023.
- J. Redford, B. Mullany, “Classification of Visual Smoothness Standards Using Multi-Scale Areal Texture Parameters and Low-Magnification Coherence Scanning Interferometry”, Materials, 2024, 17, 1653.

Surface Quality and Inspection Descriptors (SQuID)

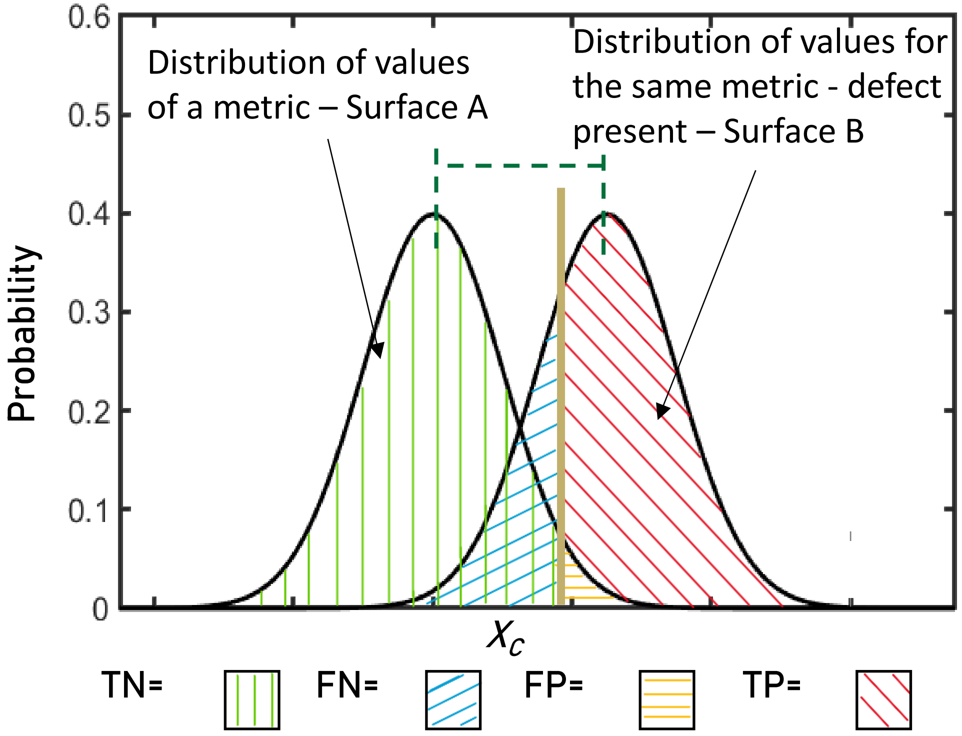
- J. Redford, B. Mullany, "Construction of a Multi-Class Discrimination Matrix and Systematic Selection of Areal Texture Parameters for Quantitative Surface and Defect Classification", Journal of Manufacturing Systems, Volume 71, December, Pages 131-143, 2023.

Why do you need a metric?

- Surface classification?
  - Good/bad
  - Type A or Type B?
- Process control? (is the process drifting)
  - Is the tool wearing?
  - Has the laser power changed?
- Process insights for optimization/understanding?
- For machine learning applications

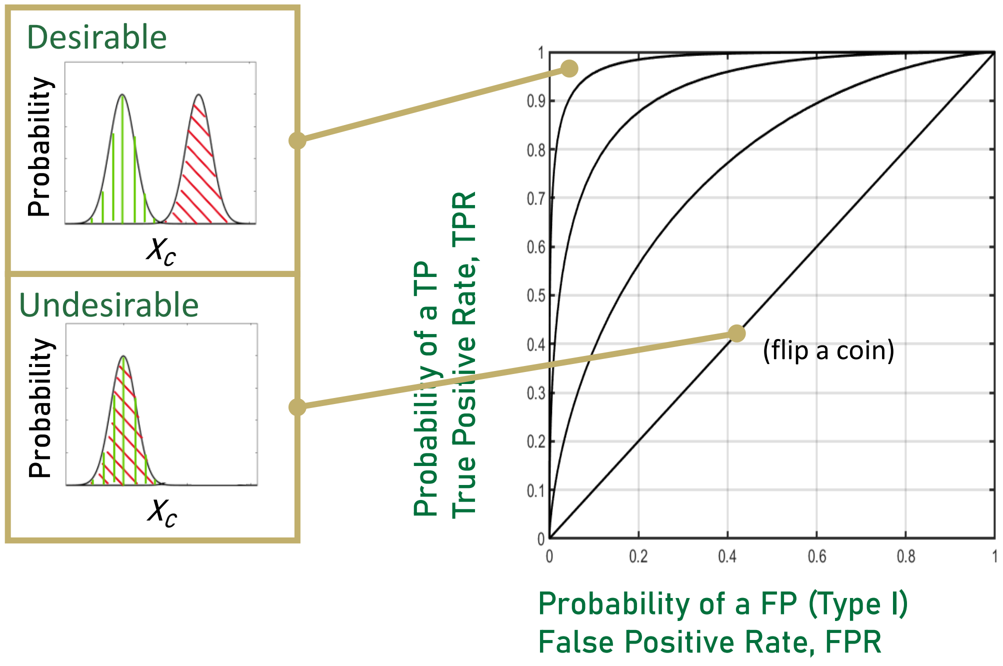
Basic requirements of a 'good' metric

	Defective	Defective-Free
Reject	True Positive (TP)	False Positive (FP) Type I Error
Accept	False Negative (FN) Type II Error	True Negative (TN)



(Green & Swets, 1966)

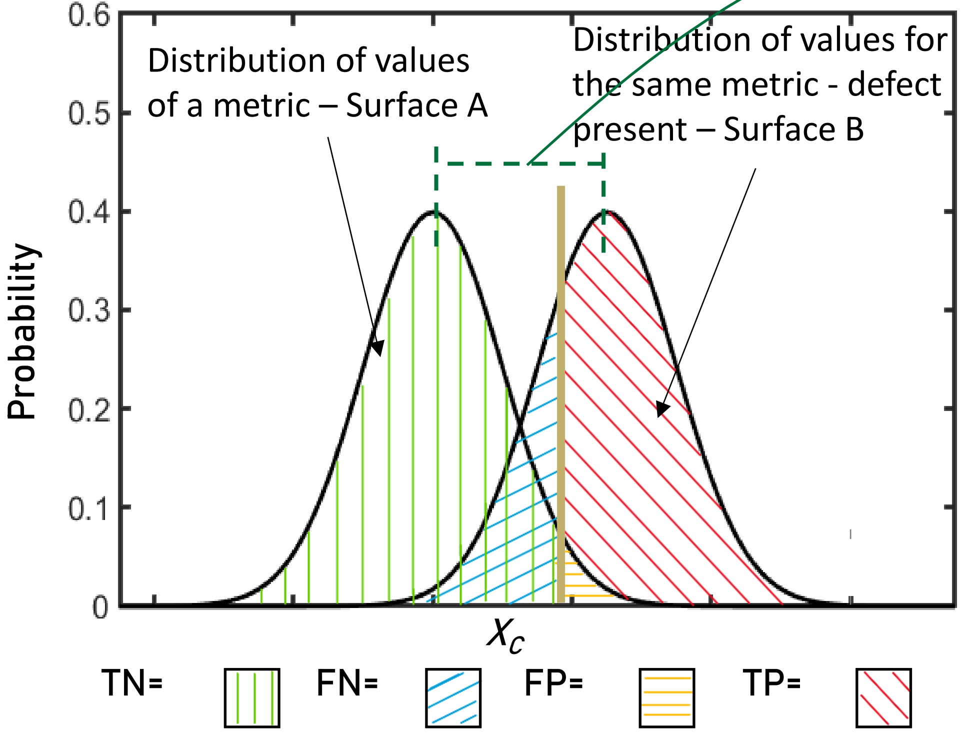
Receiving Operating Curves (ROC)



Quantification of a ‘good’ metric – *d′*

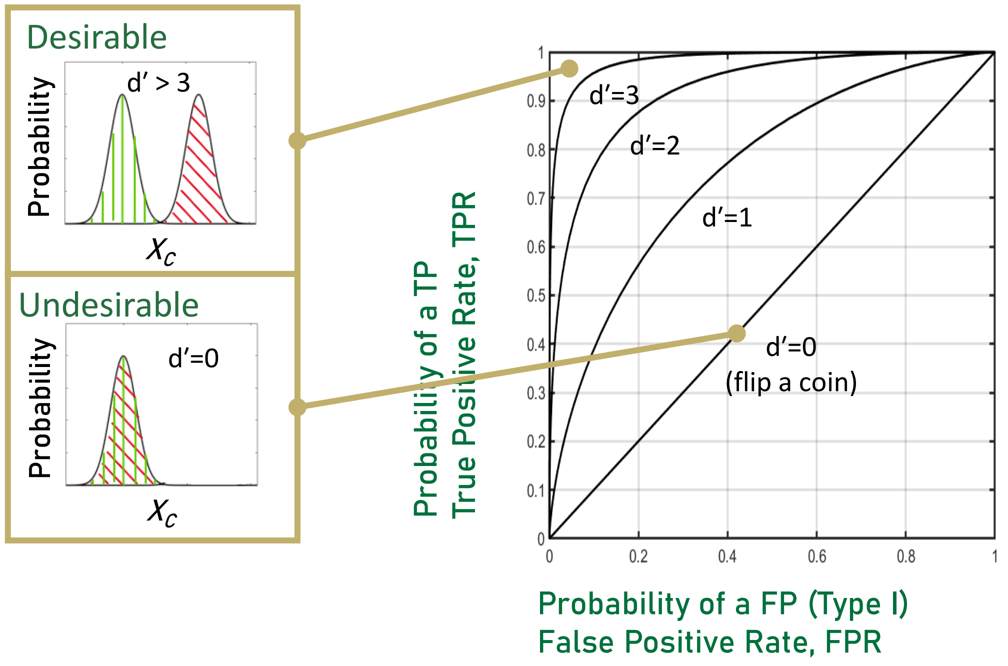
	Defective	Defective-Free
Reject	True Positive (TP)	False Positive (FP) Type I Error
Accept	False Negative (FN) Type II Error	True Negative (TN)

$$d' = \frac{|\mu_A - \mu_B|}{\sqrt{(\sigma_A^2 + \sigma_B^2)/2}}$$

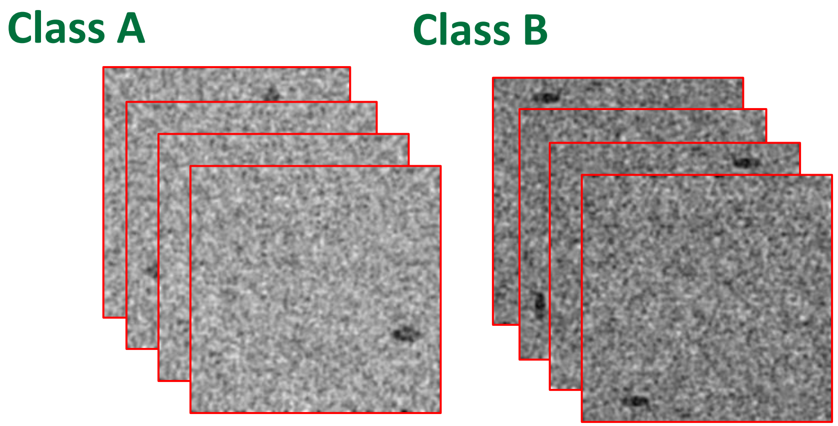


(Green & Swets, 1966)

Receiving Operating Curves (ROC)



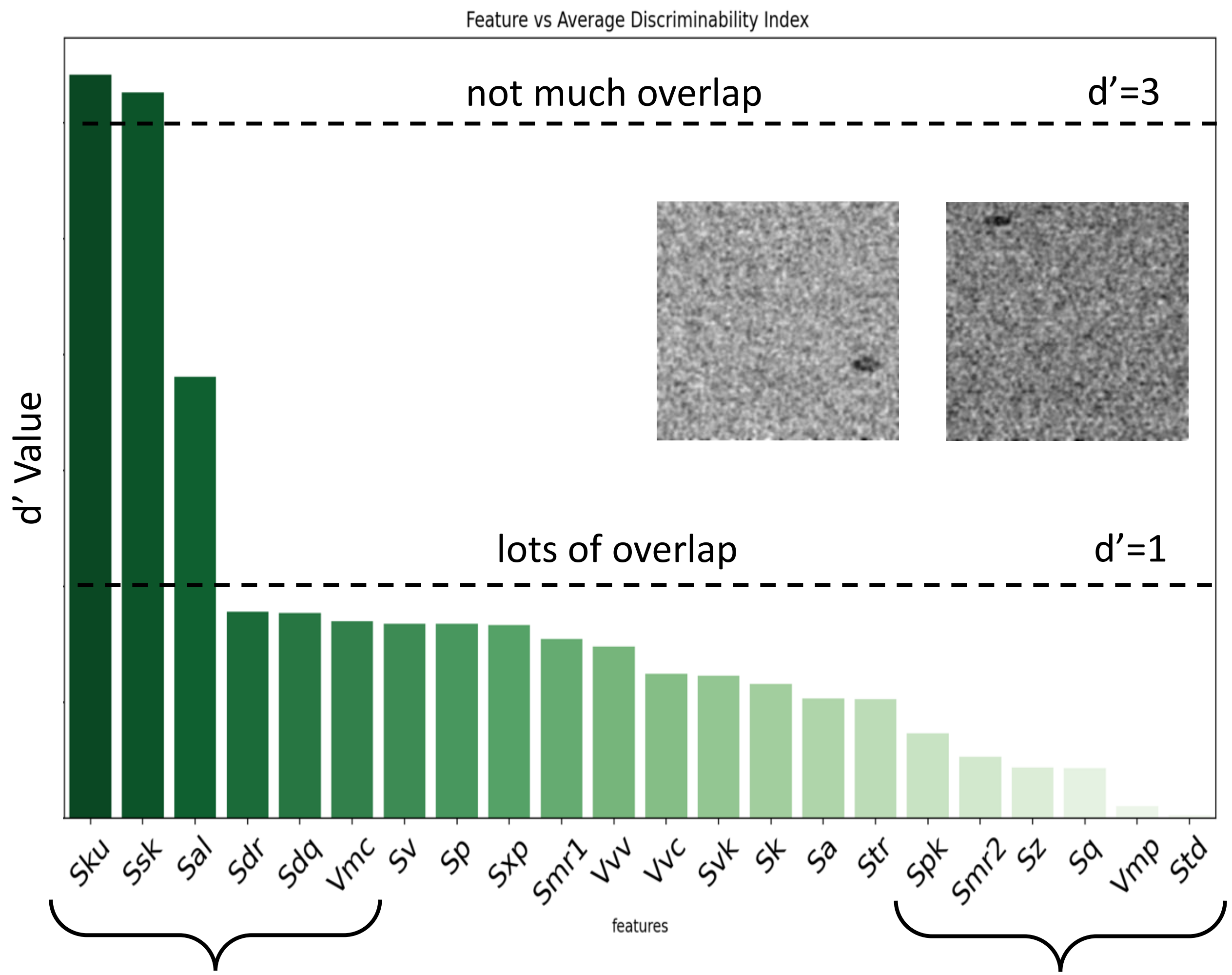
SQuID - Example how it works



- Take multiple measurements on the surface
- Calculate ISO metrics on the areal maps
- Look at the distribution of the metric values for each surface class
- Calculate d' for each metric
- Assess the d' values for each metric - isolate ones that have "good separation" $d' > 4.5$

Example how it works

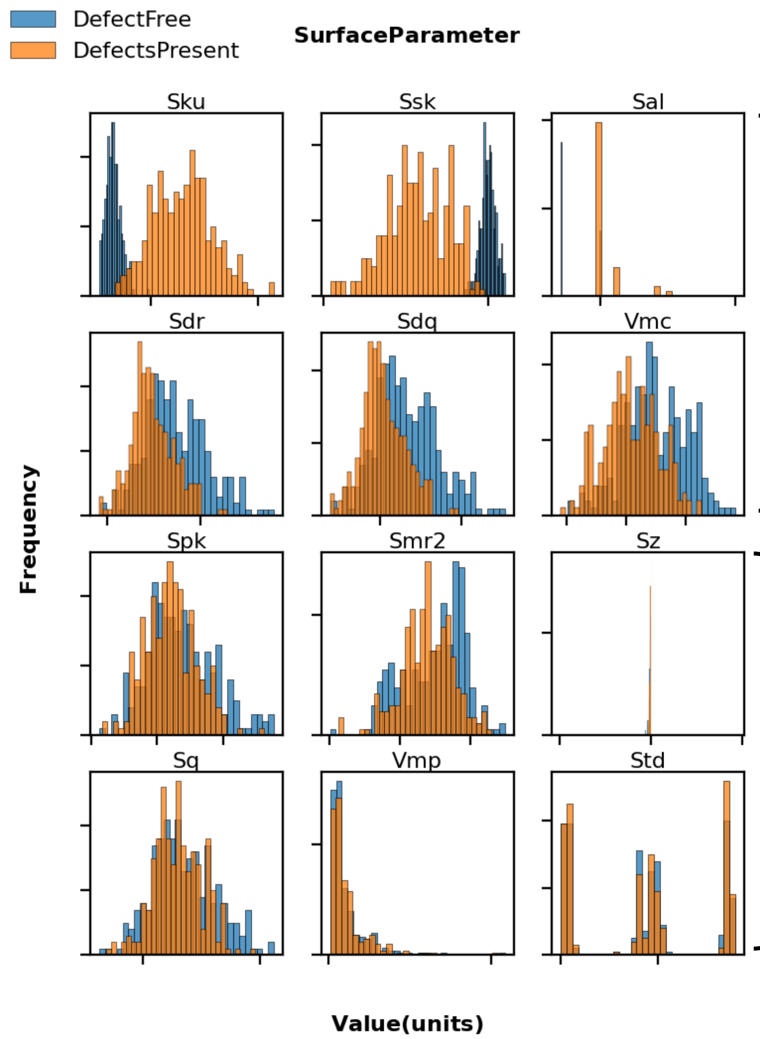
High Contrast Defects



Top 6 Parameters

Bottom 6 Parameters

Feature Histogram



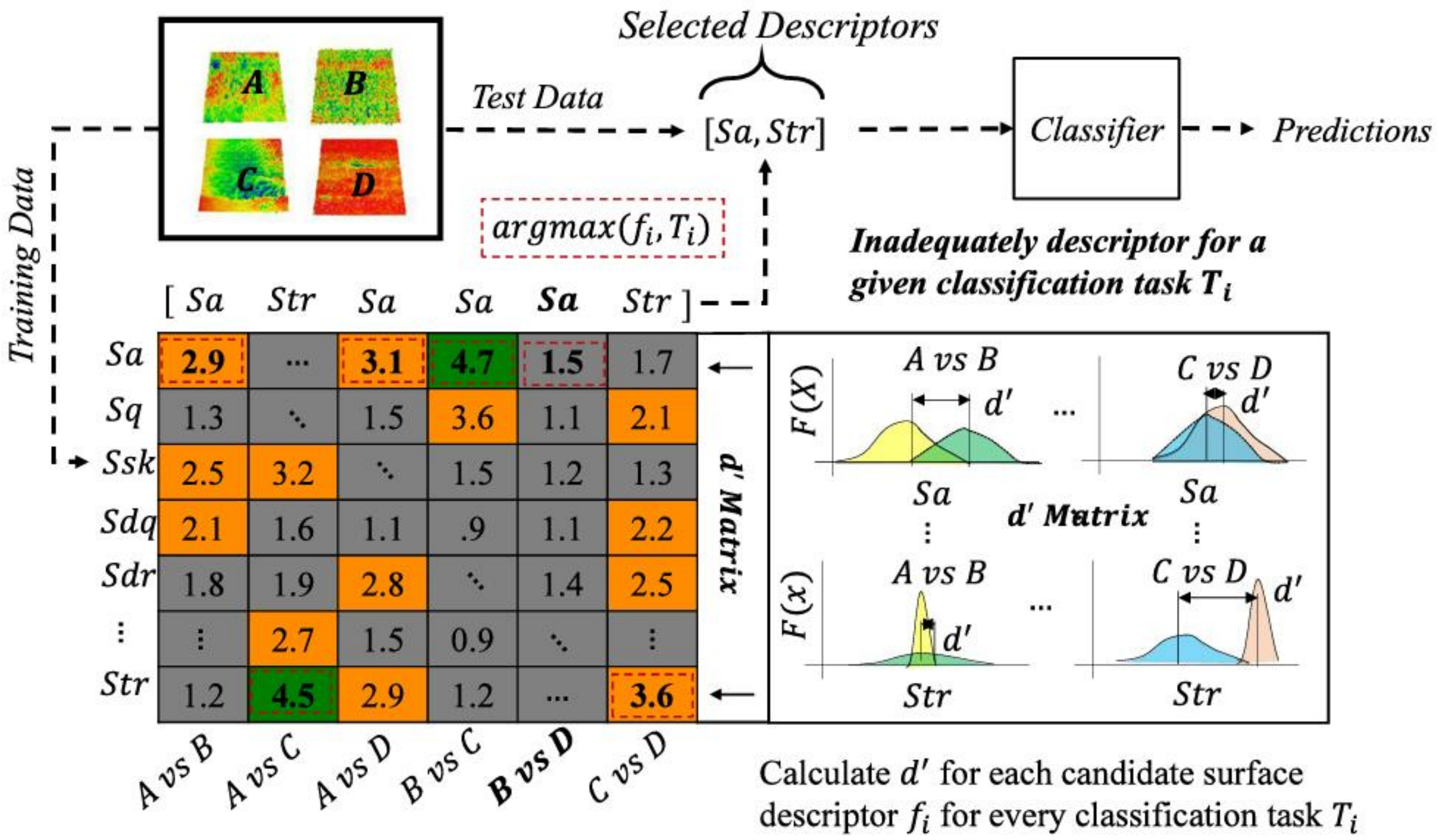
Top 6 Parameters

Bottom 6 Parameters

More than two classes?

More than two classes?

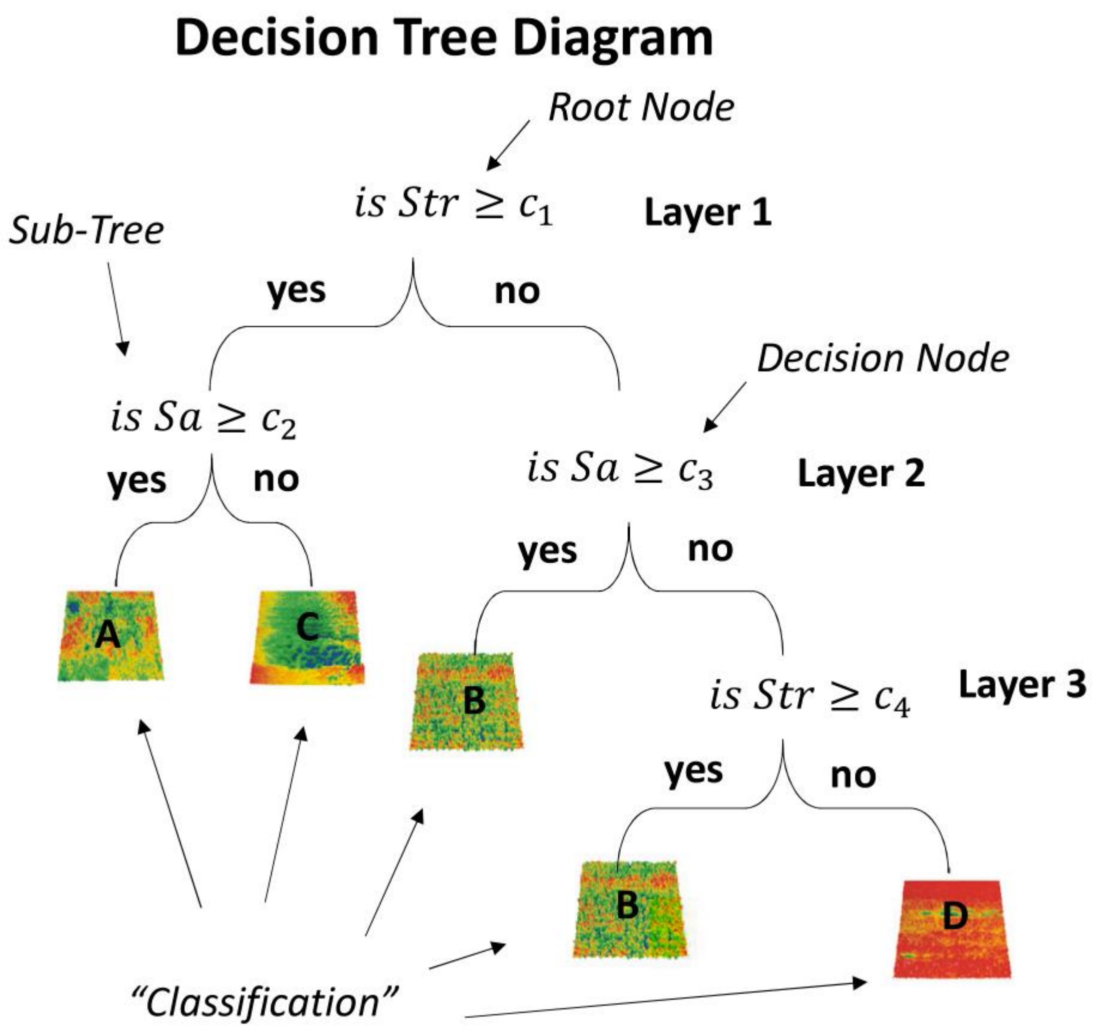
- Compare each class to every other class
- Number of classification challenges = $L(L-1)/2$ (where L = number of classes)



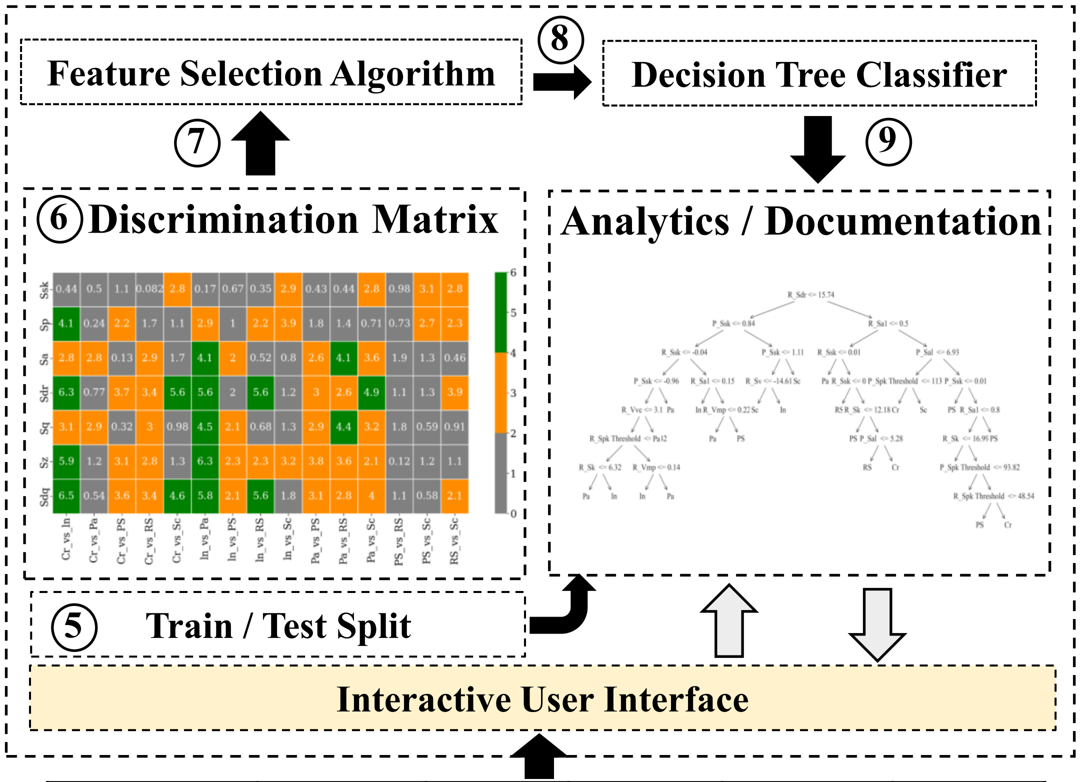
What to do with the best metrics?

Best metrics …

- Metrics -> Features for a Decision Tree Algorithm



SQuID™- Surface Quality and Inspection Descriptors



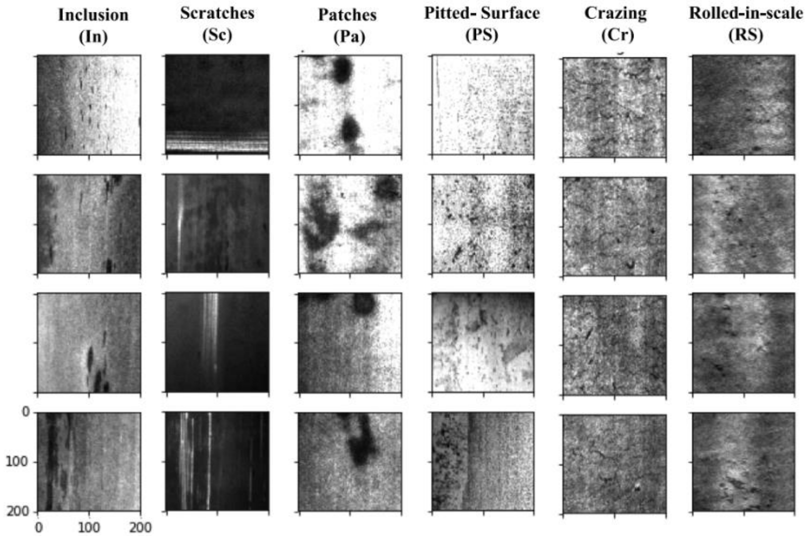
File	Label	*Sa*	*Sq*	*Ssk*	...
Cr_1.xyz	Crazing	24.5			
Cr_2.xyz	Crazing	24.2			
…	….	25.2			
Pitted_1.xyz	Pitted	28.6			
…					

Standardized Dataset Format (.csv)

B. Mullany

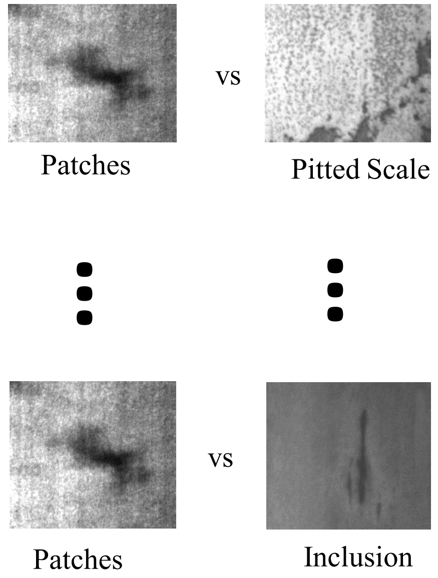
An example using the NEU

NEU-CLS Dataset

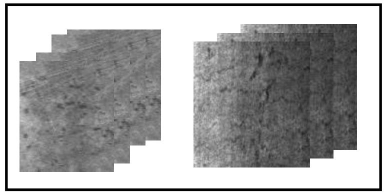


- Collection of 1800 Images
- 6 Defect Types (300 images/type)
- Low resolution 200 x 200 pixel bitmap images

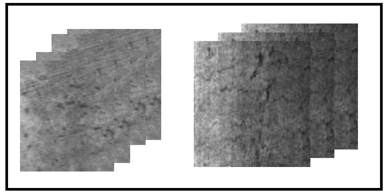
15 Binary Combination Tasks



Training Images (50%)



Test Images (50%)



F-Operator

Computed on Training Set — ***Decision Threshold*** — ***Computed on Testing Set***

ISO 25178-2 Parameters

$$d' = \frac{|\mu_A - \mu_B|}{\sqrt{(\sigma_A^2 + \sigma_B^2)/2}}$$

$$c = \frac{\mu_A - \mu_B}{2}$$

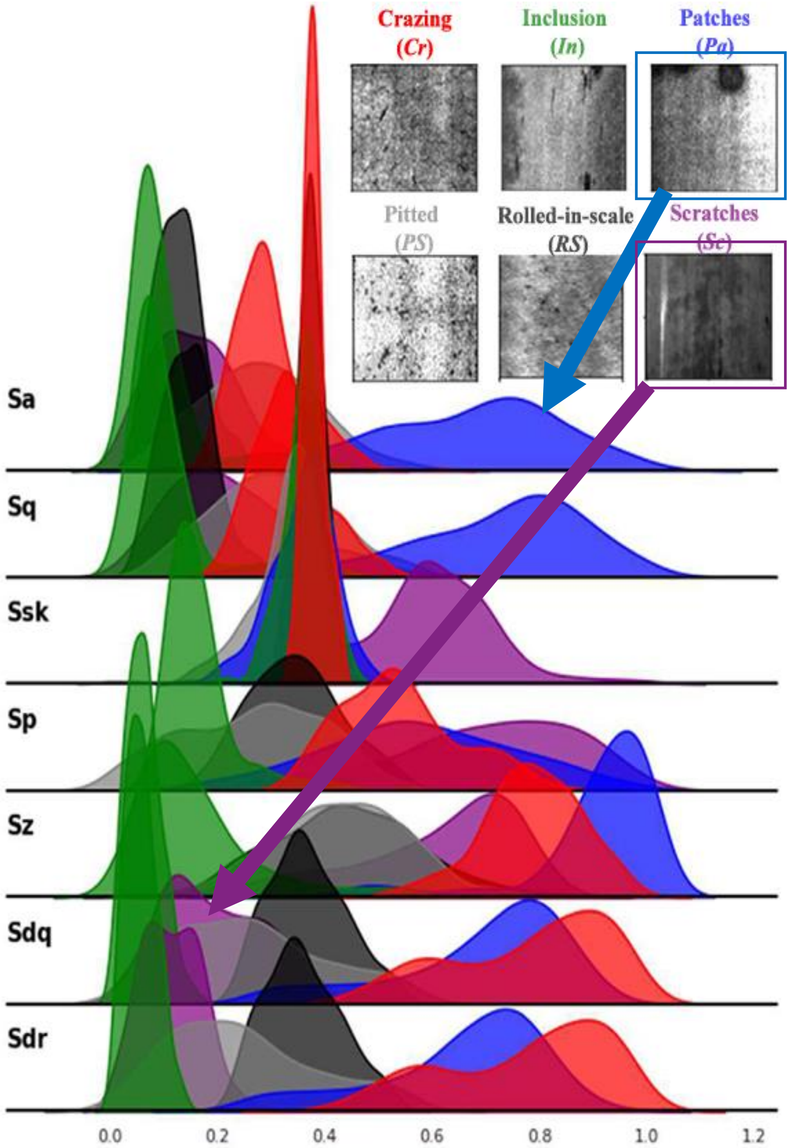
$$Acc = \frac{\#\ Correct\ Predictions}{\#\ Total\ Predictions}$$

NEU example cont.

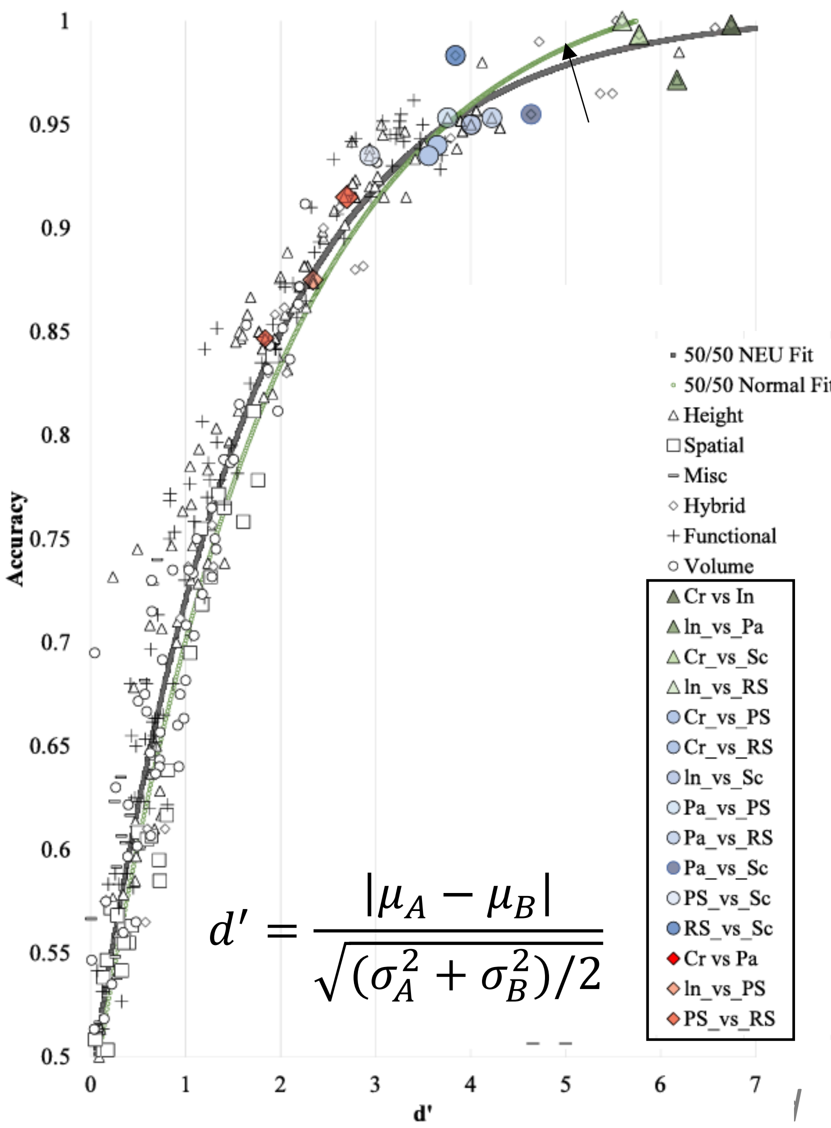
d' Matrix

	Cr_vs_In	Cr_vs_Pa	Cr_vs_PS	Cr_vs_RS	Cr_vs_Sc	In_vs_Pa	In_vs_PS	In_vs_RS	In_vs_Sc	Pa_vs_PS	Pa_vs_RS	Pa_vs_Sc	PS_vs_RS	PS_vs_Sc	RS_vs_Sc
Sa	2.8	2.8	0.13	2.9	1.7	4.1	2	0.52	0.8	2.6	4.1	3.6	1.9	1.3	0.46
Sq	3.1	2.9	0.32	3	0.98	4.5	2.1	0.68	1.3	2.9	4.4	3.2	1.8	0.59	0.91
Ssk	0.44	0.5	1.1	0.082	2.8	0.17	0.67	0.35	2.9	0.43	0.44	2.8	0.98	3.1	2.8
Sku	0.81	1.1	0.3	0.73	1.7	0.36	0.083	1.3	1.8	0.29	1.5	1.8	0.55	1.7	1.7
Sp	4.1	0.24	2.2	1.7	1.1	2.9	1	2.2	3.9	1.8	1.4	0.71	0.73	2.7	2.3
Sv	3.8	0.69	1.3	2.3	3.9	3	2.2	1.9	0.029	1.6	2.1	3	0.68	2.2	2
Sz	5.9	1.2	3.1	2.8	1.3	6.3	2.3	2.3	3.2	3.8	3.6	2.1	0.12	1.2	1.1
Sal	1.8	0.84	1.7	0.27	0.4	1.4	0.13	1.9	1.3	1.3	0.73	0.3	1.7	1.2	0.21
Str	0.62	0.19	0.69	0.16	0.52	0.36	0.016	0.44	1	0.4	0.046	0.62	0.49	1.1	0.63
Std	0.46	0.15	0.25	0.69	0.019	0.54	0.17	0.18	0.34	0.34	0.76	0.12	0.36	0.18	0.52
Sdq	6.5	0.54	3.6	3.4	4.6	5.8	2.1	5.6	1.8	3.1	2.8	4	1.1	0.58	2.1
Sdr	6.3	0.77	3.7	3.4	5.6	5.6	2	5.6	1.2	3	2.6	4.9	1.1	1.3	3.9
Smr1	0.19	0.41	0.45	0.57	0.081	0.26	0.5	0.49	0.21	0.61	0.59	0.41	0.21	0.33	0.24
Smr2	0.32	0.55	0.8	0.59	2.6	0.35	0.51	0.63	1.8	0.052	0.71	1.4	1	1.9	2.2
Sk	3.2	2.2	0.63	3.1	2.8	3.9	1.9	0.88	1.2	2.4	3.6	3.5	1.4	1.2	0.44
Spk	3.2	2.1	0.58	3.2	2.5	3.7	1.7	0.87	1.3	2.2	3.5	3.2	1.3	0.93	0.64
Svk	3.3	2.3	0.65	3	2.3	3.9	2	0.89	1.2	2.5	3.7	3.4	1.5	1.1	0.48
Sxp	3.4	1.3	1.2	2.7	0.24	2.9	1.3	1.1	1.6	1.9	2.5	0.72	0.67	0.82	1.2
Vvv	0.01	1.9	0.8	0.66	2.2	1.9	0.84	0.85	2.4	1.3	2.1	0.76	1.3	0.97	3.1
Vvc	0.19	0.048	0.0019	0.18	1.5	0.093	0.1	0.61	1.6	0.044	0.012	0.51	0.08	0.95	2
Vmp	0.62	1.5	0.83	1	0.24	1.5	1	0.46	0.52	1.1	1.6	1.4	1.1	0.56	0.64
Vmc	0.58	0.34	0.86	0.11	0.7	0.076	0.56	1.2	2	0.41	0.52	1	1.3	1.9	0.88

Selected Features

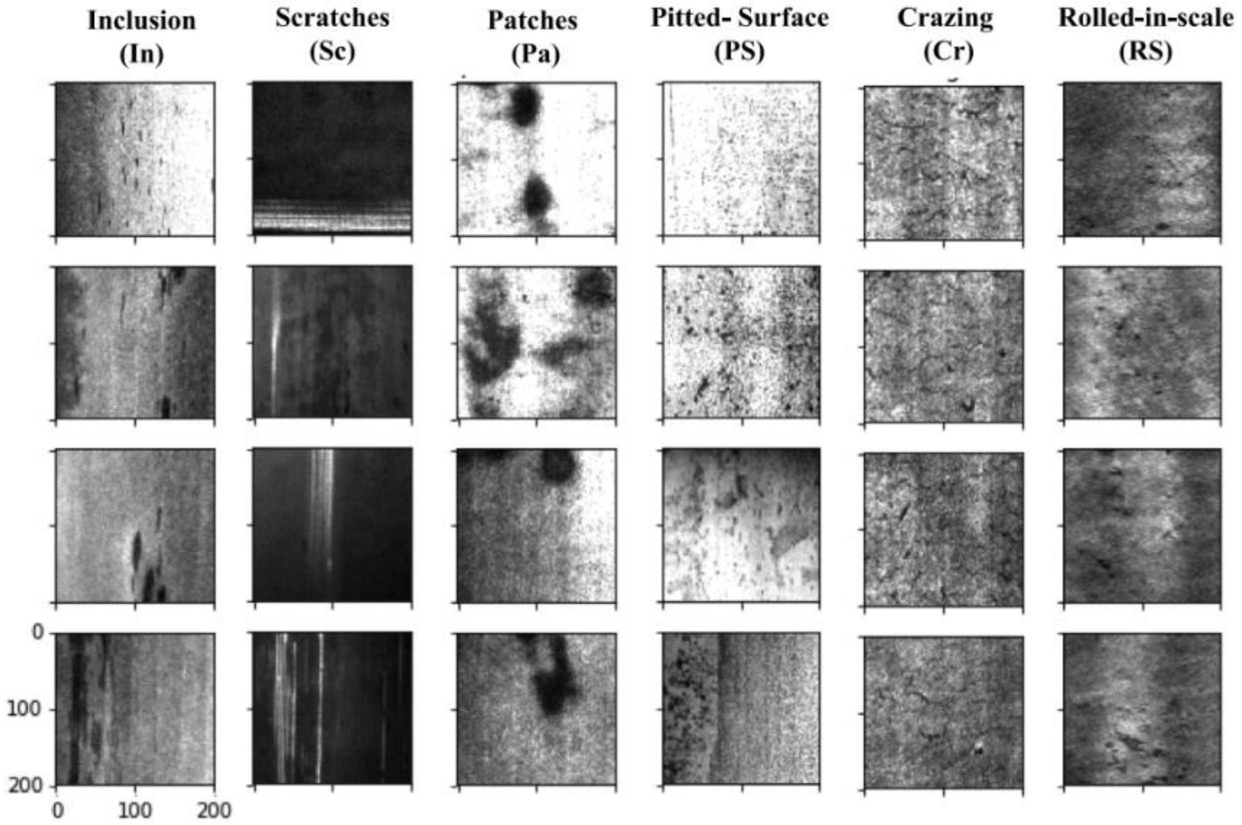


d' vs Accuracy

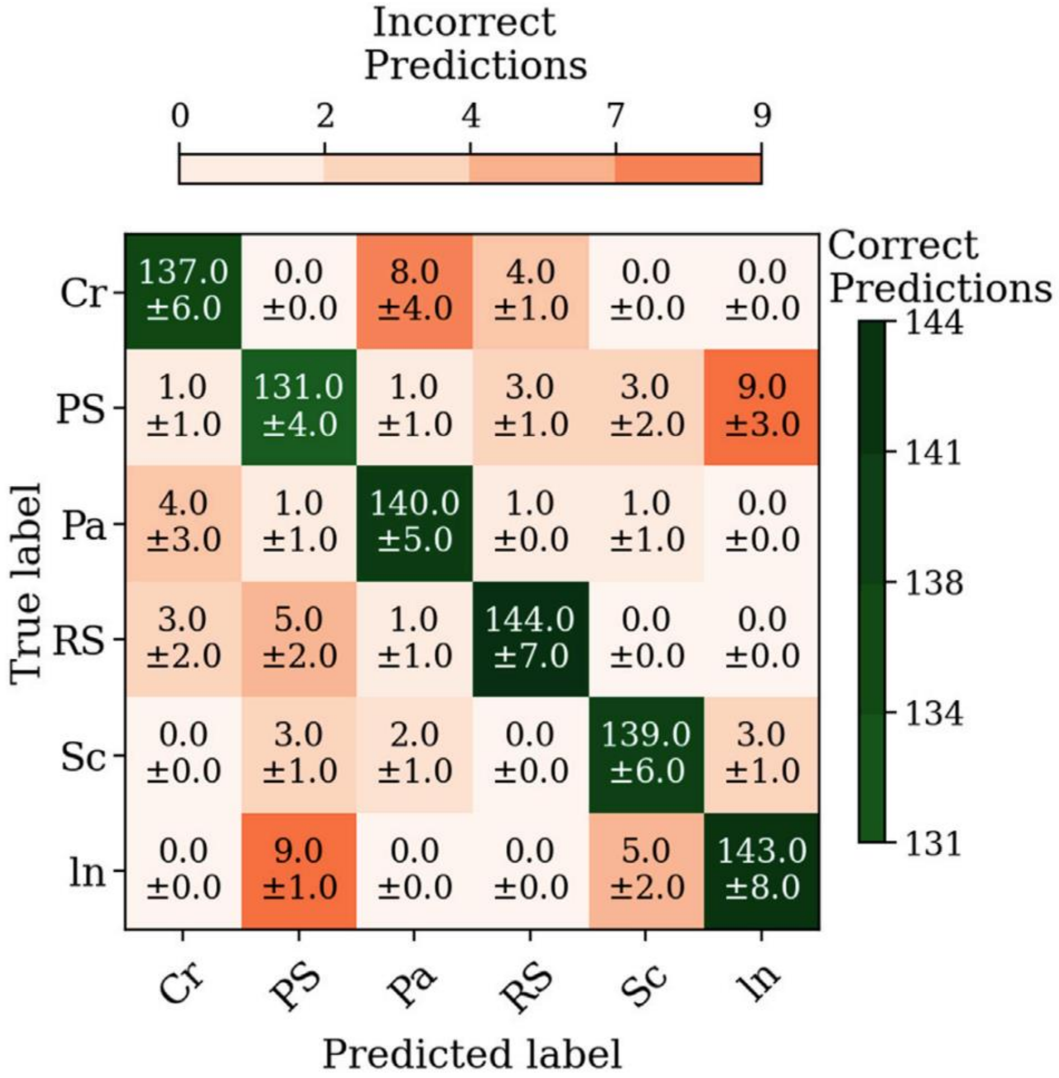


$$d' = \frac{|\mu_A - \mu_B|}{\sqrt{(\sigma_A^2 + \sigma_B^2)/2}}$$

NEU example cont.



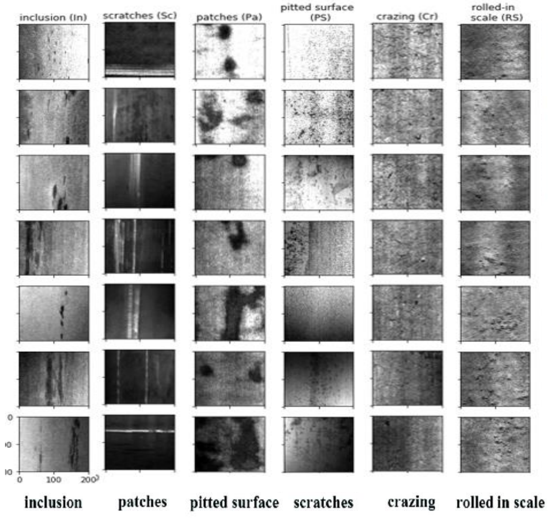
Performance of the Decision Tree – confusion matrix



NEU example cont.

Difficult to link image processing features back to the surface

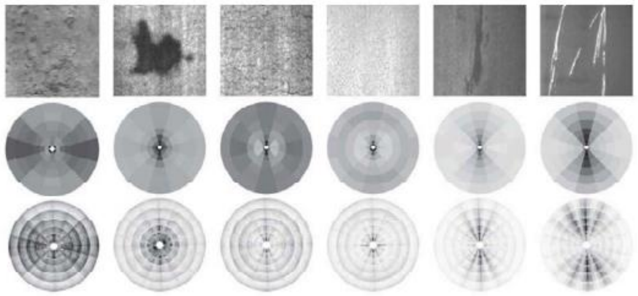
NEU-CLS Dataset



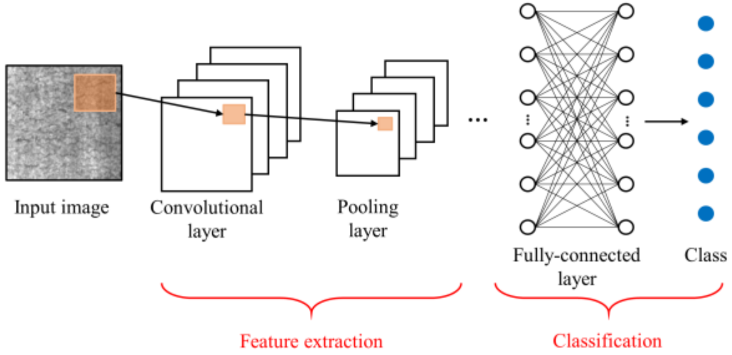
2013

Method	Classifier	Accuracy
LBP	NNC	95.07±0.71
	SVM	97.93±0.66
LTP	NNC	95.93±0.39
	SVM	98.22±0.52
CLBP	NNC	96.91±0.24
	SVM	98.18±0.51
SCN	NNC	97.24±0.27
	SVM	98.60±0.59

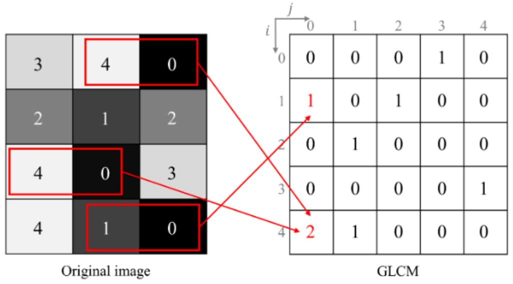
K. Song et al. 2014



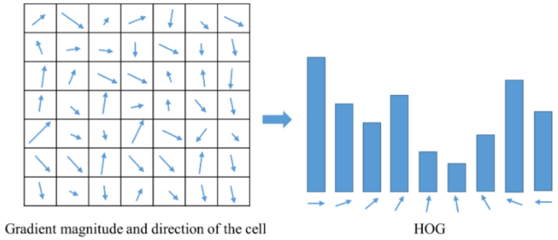
scattering convolution network



2022



(a)



Lee et al 2019

Table 4. A comparative analysis with several existing methods.

Classifier	Algorithm	Feature Extraction Method	Accuracy (%)
SVM [8]	ML	AECLBP	98.93
CNN [12]	DL	Feature learning	99.05
Ensemble of SVMs [16]	ML	ULBP+GLOM+HOG+Gabor filter+Gray level histogram	96.39
PLCNN [17]	DL	Feature learning	90.7
Proposed (CNN)	**DL**	**Feature learning**	**99.44**

Lee et al 2019

Redford & Mullany 2023

DT + ISO 25178-2 > 95%

References

https://ieeexplore.ieee.org/abstract/document/8710501

https://www.mdpi.com/2076-3417/9/24/5449

https://www.mdpi.com/2076-3417/9/24/5449

We know how decisions are made and we know what the parameters mean

Be Careful

Correlation ≠ causation