Few-Shot Learning for Defect Classification

Matteo Minardi (238789)
Computer Vision, University of Trento, Trento, 2023

I. Introduction

This report shows the exploration of the effects of various key parameters on the performance of a few-shot learning model for defect classification. The aim was to unravel the optimal configurations for a range of variables including the patch sizes, number of episodes, query and support sets sizes, and datasets used for training and testing.
A distinguishing aspect of the research is the exclusive reliance on the basic Prototypical Neural Network architecture with only *4* convolutional blocks and on solely *CPU* processing, eschewing GPU acceleration, to demonstrate that exceptional results can still be attained with simpler models and more widely accessible hardware.

II. Results

The tables show the results obtained with the different configurations, take a look at the notebook for a more in-depth analysis. The experiments are divided initially based on the datasets that have been used for training and testing, then for each case they are subdivided by patch size. In order to enhance the efficiency of execution and facilitate a more comprehensive analysis, I opted for the utilization of smaller patch sizes compared to the original implementation, hence why they have been set to *60* and *30*.

In each table it is possible to see each configuration used in each testing case (Cfg) in the form *n_way, n_support, n_query*, the number of episodes (#Eps), the recalls of all the classes of each specific task and the general accuracy score (ACC) for that case. This is very useful because we can exploit this information to see which classes are being detected more easily or with more difficulty, and are influencing the general accuracy score.

The rationale behind the selection of the three distinct configurations derives from a deliberate choice: the initial configuration, *(3,1,1)*, shows exceptional efficiency and results, the second one *(6,1,1)* extends the previous by encompassing all six classes, again with a single support sample per class to accentuate its extremeness, while the final configuration, *(6,4,4)*, comprehends again all six classes but with a richer set of support samples, serving as a valuable point of comparison against the second configuration.

Only the last two approaches have been tried with all three different configurations, because those are the only ones with enough samples to divide for training and testing.

A. Default dataset for both training and testing

1) Fixing patch_size to 60px:

Cfg	#Eps	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	0.82	0.70	0.82	1.0	0.88	0.88	0.83
	100	1.0	0.64	0.72	0.96	0.83	0.78	0.83
	250	1.0	0.73	0.67	1.0	0.81	0.86	0.85
	500	1.0	0.77	0.78	1.0	0.91	0.82	0.87
Cfg	#Eps	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	0.94	0.46	0.1	1.0	0.64	0.42	0.59
	100	1.0	0.50	0.20	1.0	0.60	0.46	0.63
	250	0.94	0.44	0.22	0.99	0.63	0.53	0.63
	500	1.0	0.50	0.30	1.0	0.55	0.44	0.63

2) Fixing patch_size to 30px:

Cfg	#Eps	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	1.0	0.86	0.48	0.83	1.0	0.71	0.81
	100	0.91	0.66	0.31	1.0	0.95	0.90	0.79
	250	0.76	0.66	0.58	0.95	0.89	0.88	0.79
	500	0.97	0.70	0.57	0.92	0.91	0.70	0.80
Cfg	#Eps	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	0.44	0.44	0.22	0.90	0.50	0.48	0.50
	100	0.04	0.45	0.23	1.0	0.54	0.56	0.47
	250	0.65	0.48	0.30	0.95	0.50	0.52	0.57
	500	0.54	0.35	0.36	0.98	0.54	0.51	0.56

B. Adding background for both training and testing

1) Fixing patch_size to 60px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	1.0	0.94	0.65	0.81	1.0	0.85	0.61	0.85
	100	1.0	0.97	0.55	0.66	1.0	0.87	0.54	0.80
	250	1.0	1.0	0.69	0.75	1.0	0.88	0.76	0.87
	500	1.0	0.91	0.74	0.81	1.0	0.90	0.70	0.87
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	1.0	1.0	0.44	0.28	1.0	0.71	0.37	0.69
	100	1.0	0.65	0.31	0.15	1.0	0.79	0.42	0.62
	250	0.99	1.0	0.39	0.20	1.0	0.70	0.35	0.67
	500	1.0	0.98	0.41	0.17	1.0	0.68	0.41	0.66

2) Fixing patch_size to 30px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	1.0	0.95	0.68	0.58	1.0	1.0	0.74	0.84
	100	1.0	0.78	0.73	0.54	1.0	0.91	0.80	0.83
	250	1.0	1.0	0.67	0.60	1.0	0.82	0.79	0.84
	500	1.0	0.85	0.81	0.56	1.0	0.86	0.70	0.83
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	1.0	0.46	0.38	0.30	1.0	0.42	0.34	0.55
	100	1.0	0.96	0.42	0.31	1.0	0.57	0.36	0.67
	250	1.0	0.52	0.60	0.36	1.0	0.57	0.32	0.63
	500	1.0	0.95	0.50	0.35	1.0	0.55	0.42	0.68

C. qPlus standard dataset for both training and testing

This dataset was introduced to try to bring more generalization into the process, including more classes than the original one to see how the results would change when comparing it to a simpler case.

1) Fixing patch_size to 60px:

Cfg	#Eps	Altro	Break	Bubble	Coating	Dirt	Dust
3,1,1	50	0.75	0.15	1.0	1.0	0.76	1.0
	100	0.84	0.33	0.92	1.0	0.88	1.0
	250	0.89	0.17	0.88	1.0	0.78	1.0
	500	0.85	0.21	0.97	1.0	0.71	1.0
	#Eps	Glass	Halo	Mark	Point	Scratch	ACC
	50	0.93	0.92	0.33	0.80	0.75	0.76
	100	0.81	0.95	0.44	0.72	0.85	0.79
	250	0.94	0.92	0.44	0.60	0.84	0.76
	500	0.91	0.89	0.55	0.67	0.72	0.77
Cfg	#Eps	Altro	Break	Bubble	Coating	Dirt	Dust
6,1,1	50	0.55	0.0	0.96	0.91	0.64	0.66
	100	0.57	0.0	0.88	0.90	0.35	0.80
	250	0.40	0.0	0.92	0.84	0.5	0.72
	500	0.46	0.01	0.95	0.88	0.55	0.84
	#Eps	Glass	Halo	Mark	Point	Scratch	ACC
	50	0.28	0.33	0.58	0.36	0.45	0.53
	100	0.35	0.24	0.58	0.43	0.46	0.51
	250	0.16	0.33	0.52	0.38	0.48	0.47
	500	0.36	0.37	0.57	0.31	0.50	0.53

2) Fixing patch_size to 30px:

Cfg	#Eps	Altro	Break	Bubble	Coating	Dirt	Dust
3,1,1	50	0.91	0.0	1.0	1.0	0.73	1.0
	100	0.66	0.0	0.93	1.0	0.87	1.0
	250	0.66	0.19	0.97	1.0	0.82	1.0
	500	0.65	0.24	0.99	1.0	0.74	1.0
	#Eps	Glass	Halo	Mark	Point	Scratch	ACC
	50	0.93	0.93	0.85	0.46	0.40	0.76
	100	0.92	1.0	0.71	0.75	0.79	0.79
	250	0.90	1.0	0.64	0.61	0.80	0.79
	500	0.80	0.85	0.58	0.70	0.64	0.74
Cfg	#Eps	Altro	Break	Bubble	Coating	Dirt	Dust
6,1,1	50	0.57	0.0	0.92	0.87	0.28	0.76
	100	0.43	0.53	0.86	0.90	0.41	0.75
	250	0.45	0.06	0.88	0.95	0.32	0.79
	500	0.54	0.44	0.92	0.85	0.24	0.75
	#Eps	Glass	Halo	Mark	Point	Scratch	ACC
	50	0.22	0.44	0.28	0.37	0.39	0.46
	100	0.20	0.65	0.54	0.55	0.74	0.60
	250	0.18	0.47	0.36	0.42	0.60	0.50
	500	0.21	0.45	0.46	0.40	0.63	0.54

D. qPlus detailed dataset for both training and testing

Compared to the previous version, this updated version of the dataset exhibits enriched representations of its existing classes, resulting in the emergence of multiple distinct classes.

1) Fixing patch_size to 60px:

Cfg	#Eps	Altro	Break	Bubble	B_Hole	Coating	Dirt	D_Halo	Dust
3,1,1	50	0.80	0.08	1.0	0.75	1.0	0.78	0.90	1.0
	100	0.79	0.10	1.0	0.77	1.0	0.80	0.90	1.0
	250	0.76	0.10	0.96	0.56	1.0	0.78	0.81	1.0
	500	0.84	0.07	0.96	0.66	1.0	0.79	0.80	0.98
	#Eps	GlassID	Halo	Mark	Point	S_Heavy	S_Light	S_Multi	ACC
	50	0.70	0.81	0.60	0.62	0.76	0.90	0.66	0.75
	100	0.86	1.0	0.50	0.75	0.58	0.72	0.55	0.76
	250	0.85	0.94	0.54	0.61	0.57	0.96	0.60	0.75
	500	0.72	0.92	0.65	0.69	0.66	0.95	0.53	0.75
Cfg	#Eps	Altro	Break	Bubble	B_Hole	Coating	Dirt	D_Halo	Dust
6,1,1	50	0.62	0.0	0.76	0.47	0.95	0.68	0.38	0.53
	100	0.45	0.0	0.87	0.70	0.85	0.50	0.36	0.80
	250	0.51	0.01	0.90	0.70	0.72	0.53	0.54	0.75
	500	0.50	0.0	0.92	0.53	0.80	0.60	0.51	0.82
	#Eps	GlassID	Halo	Mark	Point	S_Heavy	S_Light	S_Multi	ACC
	50	0.14	0.38	0.50	0.38	0.57	0.30	0.48	0.47
	100	0.23	0.25	0.65	0.46	0.53	0.35	0.53	0.51
	250	0.20	0.38	0.49	0.41	0.58	0.46	0.50	0.51
	500	0.18	0.31	0.62	0.37	0.49	0.38	0.55	0.50

2) Fixing patch_size to 30px:

Cfg	#Eps	Altro	Break	Bubble	B_Hole	Coating	Dirt	D_Halo	Dust
3,1,1	50	1.0	0.40	1.0	0.93	1.0	0.58	0.72	0.93
	100	0.63	0.07	1.0	0.56	1.0	1.0	0.89	0.95
	250	0.65	0.36	1.0	0.57	1.0	0.73	0.91	0.96
	500	0.71	0.06	0.96	0.58	1.0	0.78	0.83	0.92
	#Eps	GlassID	Halo	Mark	Point	S_Heavy	S_Light	S_Multi	ACC
	50	0.80	0.88	0.50	0.70	0.57	1.0	0.66	0.77
	100	0.94	0.96	0.59	0.90	0.66	0.92	0.71	0.80
	250	0.77	0.81	0.66	0.52	0.79	0.91	0.66	0.75
	500	0.88	0.93	0.66	0.69	0.70	0.86	0.64	0.75
Cfg	#Eps	Altro	Break	Bubble	B_Hole	Coating	Dirt	D_Halo	Dust
6,1,1	50	0.48	0.0	0.87	0.61	0.80	0.55	0.56	0.78
	100	0.48	0.13	0.90	0.56	0.78	0.39	0.63	0.82
	250	0.46	0.10	0.82	0.52	0.82	0.38	0.57	0.71
	500	0.50	0.0	0.83	0.64	0.82	0.27	0.58	0.67
	#Eps	GlassID	Halo	Mark	Point	S_Heavy	S_Light	S_Multi	ACC
	50	0.04	0.31	0.33	0.26	0.60	0.45	0.55	0.48
	100	0.20	0.52	0.35	0.29	0.54	0.33	0.51	0.50
	250	0.35	0.41	0.52	0.35	0.53	0.41	0.64	0.52
	500	0.07	0.46	0.41	0.48	0.69	0.42	0.57	0.51

E. qPlus standard dataset for training and testing on Default (+ background) dataset

1) Fixing patch_size to 60px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	0.66	0.44	0.71	0.61	0.45	0.72	0.76	0.62
	100	0.65	0.42	0.86	0.71	0.57	0.58	0.78	0.66
	250	0.69	0.39	0.69	0.72	0.63	0.66	0.58	0.62
	500	0.69	0.42	0.65	0.68	0.60	0.68	0.71	0.64
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	0.82	0.25	0.47	0.36	0.57	0.43	0.36	0.47
	100	0.75	0.21	0.35	0.34	0.55	0.50	0.36	0.45
	250	0.80	0.25	0.44	0.39	0.60	0.45	0.42	0.48
	500	0.80	0.25	0.43	0.33	0.56	0.45	0.44	0.47
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,4,4	50	0.55	0.75	0.62	0.49	0.59	0.66	0.75	0.63
	100	0.53	0.79	0.62	0.52	0.64	0.60	0.68	0.62
	250	0.60	0.74	0.61	.49	0.61	0.61	0.72	0.63
	500	0.60	0.75	0.60	0.49	0.64	0.59	0.70	0.63

2) Fixing patch_size to 30px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	0.72	0.26	0.75	0.79	0.53	0.54	0.75	0.62
	100	0.80	0.36	0.66	0.74	0.42	0.74	0.85	0.65
	250	0.72	0.43	0.74	0.71	0.45	0.84	0.69	0.66
	500	0.80	0.44	0.74	0.76	0.45	0.70	0.72	0.66
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	0.84	0.12	0.47	0.36	0.46	0.56	0.40	0.47
	100	0.82	0.09	0.44	0.41	0.56	0.49	0.45	0.47
	250	0.85	0.16	0.45	0.32	0.51	0.51	0.49	0.47
	500	0.87	0.12	0.46	0.30	0.47	0.53	0.40	0.45
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,4,4	50	0.82	0.31	0.52	0.52	0.65	0.59	0.60	0.57
	100	0.81	0.29	0.51	0.51	0.59	0.60	0.65	0.57
	250	0.77	0.35	0.50	0.50	0.56	0.55	0.62	0.55
	500	0.82	0.32	0.52	0.49	0.56	0.57	0.62	0.55

F. qPlus detailed dataset for training and testing on Default (+ background) dataset

1) Fixing patch_size to 60px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	0.63	0.34	0.62	0.80	0.40	0.75	0.72	0.61
	100	0.70	0.31	0.54	0.78	0.63	0.72	0.57	0.60
	250	0.77	0.30	0.58	0.74	0.50	0.72	0.65	0.62
	500	0.77	0.22	0.59	0.74	0.50	0.74	0.61	0.59
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	74	0.35	0.52	0.31	0.65	0.50	0.48	0.51
	100	0.83	0.40	0.40	0.29	0.56	0.41	0.38	0.47
	250	0.80	0.40	0.35	0.29	0.52	0.43	0.41	0.46
	500	0.79	0.39	0.42	0.33	0.55	0.42	0.40	0.47
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,4,4	50	0.55	0.73	0.50	0.41	0.63	0.50	0.61	0.57
	100	0.57	0.81	0.56	0.39	0.70	0.52	0.59	0.60
	250	0.63	0.81	0.51	0.42	0.62	0.55	0.58	0.59
	500	0.62	0.81	0.49	0.45	0.65	0.52	0.58	0.59

2) Fixing patch_size to 30px:

Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
3,1,1	50	0.84	0.46	0.79	0.72	0.56	0.52	0.53	0.66
	100	0.79	0.33	0.63	0.61	0.62	0.61	0.65	0.61
	250	0.84	0.34	0.64	0.72	0.70	0.57	0.54	0.62
	500	0.76	0.37	0.61	0.71	0.65	0.58	0.64	0.62
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,1,1	50	0.81	0.17	0.53	0.30	0.34	0.45	0.60	0.45
	100	0.74	0.15	0.55	0.26	0.48	0.48	0.53	0.46
	250	0.77	0.20	0.45	0.37	0.53	0.56	0.50	0.49
	500	0.86	0.22	0.47	0.31	0.51	0.51	0.45	0.48
Cfg	#Eps	BG	Break	Bubble	Dirt	Mark	Point	Scratch	ACC
6,4,4	50	0.81	0.44	0.61	0.35	0.55	0.56	0.70	0.57
	100	0.80	0.49	0.49	0.35	0.59	0.57	0.56	0.61
	250	0.76	0.47	0.55	0.37	0.57	0.58	0.70	0.57
	500	0.73	0.47	0.53	0.33	0.59	0.57	0.69	0.56

III. CONCLUSION

Notably, the accuracy using the smaller patch size was sometimes even able to surpass that of the larger counterpart, presenting a possible approach for scenarios where computational efficiency is more desired. Remarkably, even with the executions conducted solely on a CPU, no substantial obstacles were encountered, especially when exploiting the smaller patch size.

Overall, the first configuration prominently outperformed the others. This outcome can be attributed to its efficiency, likely by focusing on a subset of classes rather than attempting to deal with the entire spectrum concurrently, as observed in the subsequent configurations. The drop in performance of the latter configurations can be attributed to the higher model complexity resulting from the increased abundance of classes. The only big issue that appeared is the inability for the model trained on the *qPlus* dataset to consistently detect the *"Break"* class, often resulting in almost zero recall, probably because it lacks enough evidence.

Experiments with less classes at a time seem more accurate, but they have less chance to make mistakes and are not actually that much better. In fact, when comparing those results with the ones with higher number of *n_way*, we can see that the difference is not always too substantial.

As expected, increasing the number of support samples shows a great improvement in these Meta-Learning tasks.