



**Middle East Technical University
Informatics Institute**

A DEEP NEURAL NETWORK BASED PRODUCT METADATA VALIDATION APPROACH FOR ONLINE MARKETPLACES

**Advisor Name: Prof. Dr. Altan Koçyiğit
(METU)**

**Student Name: Şükrü Alataş
(IS)**

January 2023

**TECHNICAL REPORT
METU/II-TR-2023-**



**Orta Doğu Teknik Üniversitesi
Enformatik Enstitüsü**

**ÇEVİRİMİÇİ PAZARYERLERİ İÇİN
DERİN SİNİR AĞLARINA DAYALI
ÜRÜN ÜSTBİLGİSİ GEÇERLEME YAKLAŞIMI**

**Danışman Adı: Prof. Dr. Altan Koçyiğit
(ODTÜ)**

**Öğrenci Adı: Şükrü Alataş
(IS)**

Ocak 2023

**TEKNİK RAPOR
ODTÜ/II-TR-2023-**

TABLE OF CONTENTS

1.	INTRODUCTION	5
2.	Automatic Validation Model Training Run	6
2.1.	RGB Color Space	6
2.1.1.	CNN AlexNet.....	6
2.2.	HSL Color Space.....	13
2.2.1.	CNN AlexNet.....	13
3.	Unrefined Dataset Improvement Run	20
3.1.	RGB Color Space	20
3.1.1.	CNN AlexNet with Single Color	20
3.1.2.	Baseline Methods with Single Color.....	30
3.1.3.	CNN AlexNet with Multi-color	49
3.1.4.	Baseline Methods with Multi-color	59
3.2.	HSL Color Space.....	78
3.2.1.	CNN AlexNet with Single Color	78
3.2.2.	Baseline Methods with Single Color.....	88
3.2.3.	CNN AlexNet with Multi-color	101
3.2.4.	Baseline Methods with Multi-color	111
4.	Manually Validated Dataset Improvement Run.....	124
4.1.	RGB Color Space	124
4.1.1.	CNN AlexNet with Single Color	124
4.1.2.	Baseline Methods with Single Color.....	134
4.1.3.	CNN AlexNet with Multi-color	153
4.1.4.	Baseline Methods with Multi-color	163
4.2.	HSL Color Space.....	182
4.2.1.	CNN AlexNet with Single Color	182
4.2.2.	Baseline Methods with Single Color.....	192
4.2.3.	CNN AlexNet with Multi-color	205
4.2.4.	Baseline Methods with Multi-color	215
5.	Refined Dataset Improvement Run.....	228
5.1.	RGB Color Space	228
5.1.1.	CNN AlexNet with Single Color	228

5.1.2.	Baseline Methods with Single Color.....	238
5.1.3.	CNN AlexNet with Multi-color	257
5.1.4.	Baseline Methods with Multi-color	267
5.2.	HSL Color Space.....	286
5.2.1.	CNN AlexNet with Single Color	286
5.2.2.	Baseline Methods with Single Color.....	296
5.2.3.	CNN AlexNet with Multi-color	309
5.2.4.	Baseline Methods with Multi-color	319

INTRODUCTION

This technical report serves as the complimentary appendix for our thesis, which presents a new approach for automated product metadata validation in e-commerce. Our proposal includes an approach using an AI-based automated image validation model for validating product images and an AI-based classification model to validate product metadata in an automated fashion. Our approach offers several advantages over traditional methods, including handling complex and noisy data and adapting to various challenging product categories, such as fashion items. We demonstrate the effectiveness of this approach through comparisons with traditional methods and in different settings, ultimately showing strong support for the use of AI in product metadata validation for online marketplaces.

This technical report provides the outputs of 26 experiments conducted during our research.

AUTOMATIC VALIDATION MODEL TRAINING RUN

2.1. RGB Color Space

2.1.1. CNN AlexNet

Automatic Validation Model Training Run

CNN AlexNet

Total Duration: 372.05 sec.

Learning Rate: 0.000070

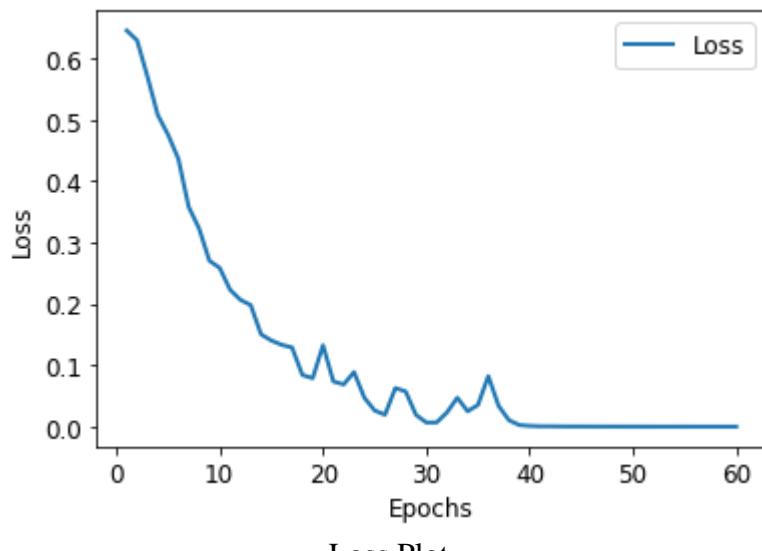
Dropout: 0.10

Batch Size: 32

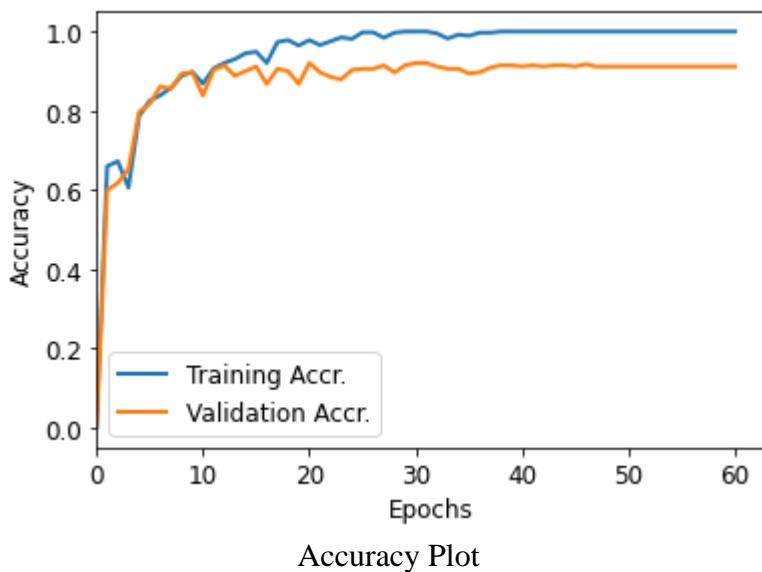
Epoch: 60

Best Performing Epoch: 31

Summary	
Method	Accuracy
Training Dataset	1.0000
Validation Dataset	0.9206
Test Dataset	0.9353



Loss Plot

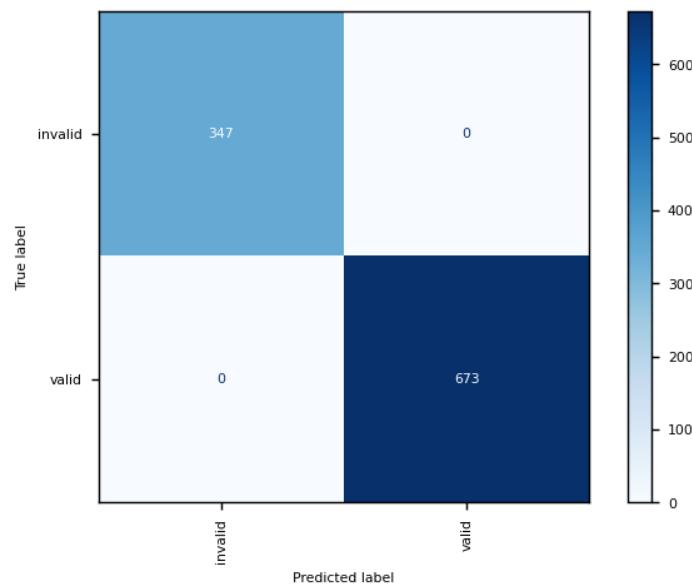


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	1.0000	1.0000	1.0000	347
valid	1.0000	1.0000	1.0000	673
accuracy			1.0000	
macro avg	1.0000	1.0000	1.0000	1020
weighted avg	1.0000	1.0000	1.0000	1020

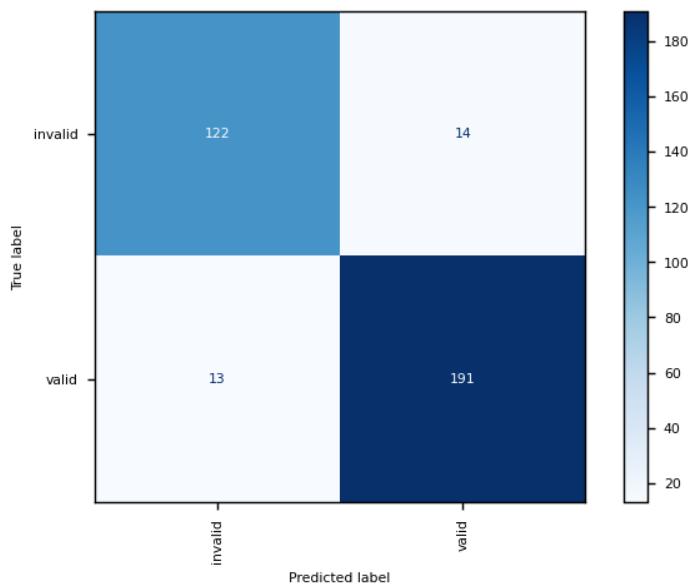


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	0.9037	0.8971	0.9004	136
valid	0.9317	0.9363	0.9340	204
accuracy			0.9206	
macro avg	0.9177	0.9167	0.9172	340
weighted avg	0.9205	0.9206	0.9205	340

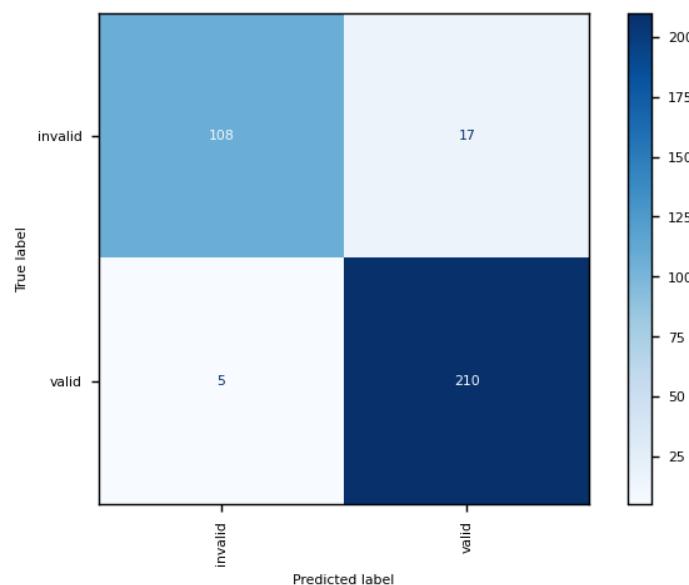


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	0.9558	0.8640	0.9076	125
valid	0.9251	0.9767	0.9502	215
accuracy			0.9353	
macro avg	0.9404	0.9204	0.9289	340
weighted avg	0.9364	0.9353	0.9345	340



Test Dataset Confusion Matrix

valid (100.0%) invalid (100.0%) valid (100.0%) valid (100.0%) valid (100.0%) invalid (100.0%)
>> valid >> invalid >> valid >> valid >> valid >> invalid



valid (100.0%) invalid (100.0%) valid (98.6%) valid (100.0%) valid (96.7%) valid (100.0%)
>> valid >> invalid >> valid >> valid >> valid >> valid



valid (100.0%) invalid (100.0%) valid (100.0%) valid (100.0%) valid (95.8%) valid (100.0%)
>> valid >> invalid >> valid >> valid >> valid >> valid



valid (100.0%) invalid (57.7%) valid (100.0%) valid (100.0%) valid (100.0%) valid (99.5%)
>> valid >> invalid >> valid >> valid >> valid >> valid



valid (99.8%) valid (100.0%) valid (100.0%) invalid (100.0%) valid (100.0%) invalid (100.0%)
>> valid >> valid >> valid >> invalid >> valid >> invalid



valid (100.0%) valid (100.0%) valid (96.6%) valid (100.0%) valid (100.0%) valid (100.0%)
>> valid >> valid >> valid >> valid >> valid >> valid



Sample Images from Test Dataset

invalid (57.7%) valid (99.8%) valid (98.5%) valid (98.4%) valid (85.0%) valid (96.3%)
>> valid >> invalid >> invalid >> invalid >> invalid >> invalid



valid (99.5%) valid (100.0%) valid (98.8%) valid (100.0%) invalid (99.4%) valid (100.0%)
>> invalid >> invalid >> invalid >> invalid >> valid >> invalid



invalid (76.5%) valid (85.0%) valid (100.0%) valid (66.8%) valid (100.0%) valid (99.9%)
>> valid >> invalid >> invalid >> invalid >> invalid >> invalid



invalid (100.0%) valid (99.9%) invalid (71.5%) valid (100.0%)
>> valid >> invalid >> valid >> invalid



Sample Images that cannot be accurately predicted in the Test Dataset

2.2. HSL Color Space

2.2.1. CNN AlexNet

Automatic Validation Model Training Run

CNN AlexNet

Total Duration: 677.43 sec.

Learning Rate: 0.000070

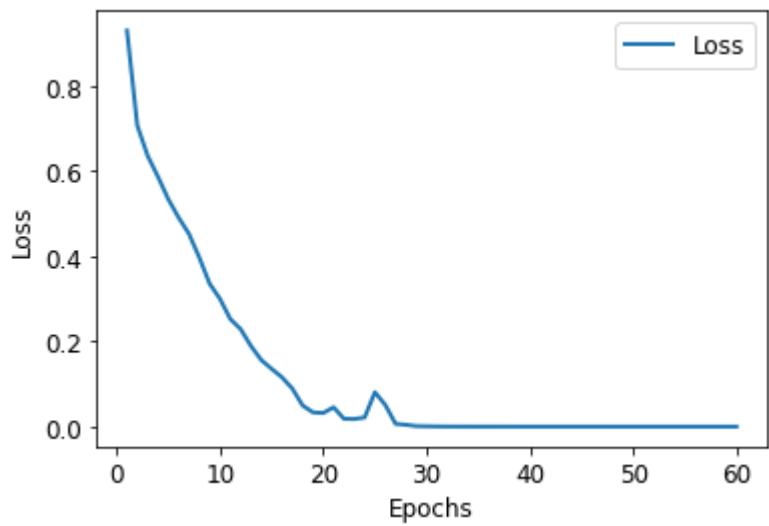
Dropout: 0.10

Batch Size: 32

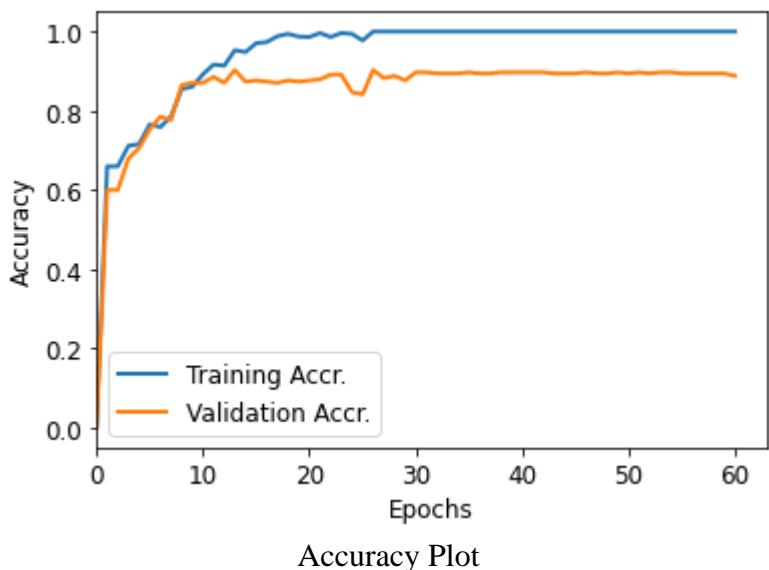
Epoch: 60

Best Performing Epoch: 26

Summary	
Method	Accuracy
Training Dataset	1.0000
Validation Dataset	0.9029
Test Dataset	0.9029



Loss Plot

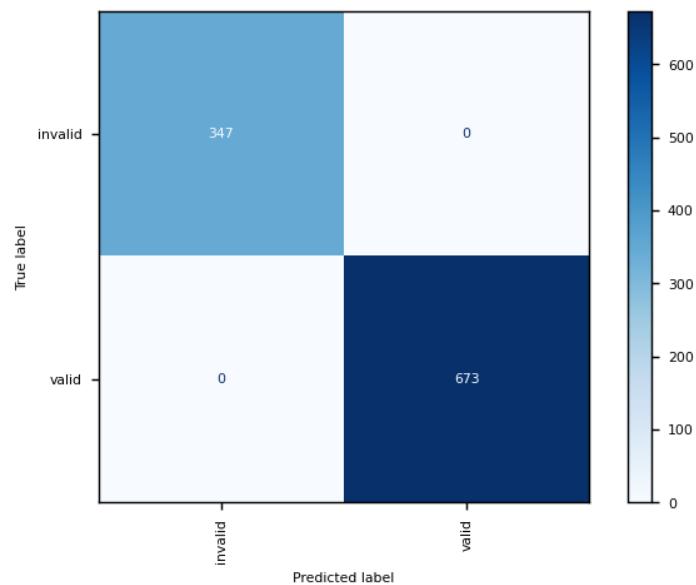


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	1.0000	1.0000	1.0000	347
valid	1.0000	1.0000	1.0000	673
accuracy			1.0000	
macro avg	1.0000	1.0000	1.0000	1020
weighted avg	1.0000	1.0000	1.0000	1020

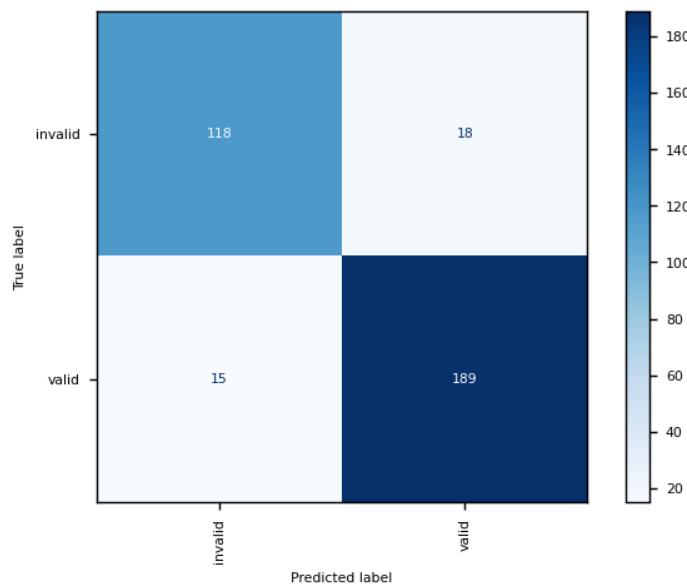


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	0.8872	0.8676	0.8773	136
valid	0.9130	0.9265	0.9197	204
accuracy			0.9029	
macro avg	0.9001	0.8971	0.8985	340
weighted avg	0.9027	0.9029	0.9028	340

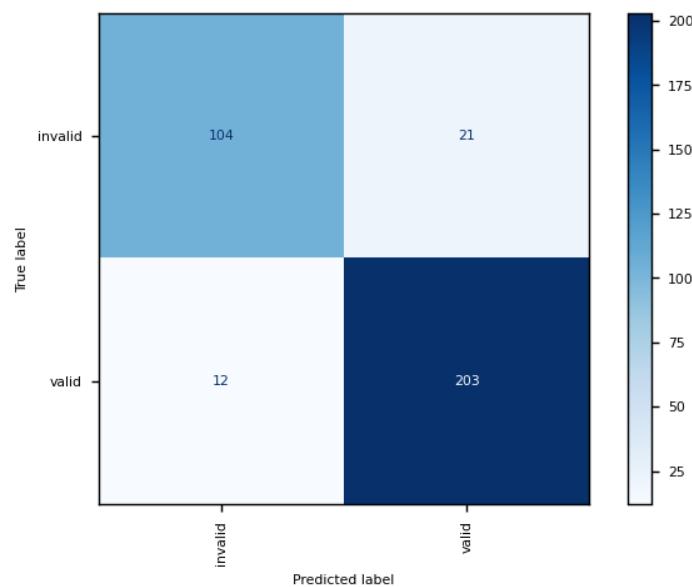


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
invalid	0.8966	0.8320	0.8631	125
valid	0.9062	0.9442	0.9248	215
accuracy			0.9029	
macro avg	0.9014	0.8881	0.8939	340
weighted avg	0.9027	0.9029	0.9021	340



Test Dataset Confusion Matrix

invalid (96.8%) valid (100.0%) invalid (95.4%) invalid (100.0%) valid (100.0%) valid (100.0%)
>> invalid >> valid >> invalid >> invalid >> valid >> valid



valid (100.0%) valid (99.4%) invalid (99.6%) valid (99.9%) valid (100.0%) valid (78.7%)
>> valid >> valid >> invalid >> valid >> valid >> valid



valid (100.0%) invalid (99.9%) valid (100.0%) valid (100.0%) invalid (100.0%) valid (100.0%)
>> valid >> invalid >> valid >> valid >> invalid >> valid



valid (100.0%) invalid (99.8%) valid (99.7%) valid (94.2%) valid (100.0%) valid (100.0%)
>> valid >> invalid >> valid >> valid >> valid >> valid



valid (99.8%) valid (100.0%) valid (100.0%) valid (99.9%) valid (100.0%) invalid (99.0%)
>> valid >> valid >> valid >> valid >> valid >> invalid



valid (99.8%) valid (99.6%) valid (90.8%) valid (91.9%) valid (100.0%) valid (99.1%)
>> valid >> valid >> invalid >> valid >> valid >> valid



Sample Images from Test Dataset

valid (99.5%) invalid (85.0%) valid (99.0%) invalid (99.1%) valid (97.6%) valid (90.8%)
 >> invalid >> valid >> invalid >> valid >> invalid >> invalid



invalid (89.6%) invalid (88.9%) invalid (63.7%) valid (77.0%) valid (64.4%) valid (80.3%)
 >> valid >> valid >> valid >> invalid >> invalid >> invalid



valid (98.8%) invalid (98.6%) valid (99.7%) valid (94.4%) invalid (98.8%) valid (88.5%)
 >> invalid >> valid >> invalid >> invalid >> valid >> invalid



valid (96.6%) invalid (94.0%) invalid (70.3%) valid (82.8%) invalid (90.3%) valid (99.7%)
 >> invalid >> valid >> valid >> invalid >> valid >> invalid



valid (99.6%) valid (56.1%) invalid (51.9%) valid (67.0%) valid (96.7%) invalid (93.4%)
 >> invalid >> invalid >> valid >> invalid >> valid >> valid



valid (97.2%) valid (80.9%) valid (99.8%)
 >> invalid >> invalid >> invalid



Sample Images that cannot be accurately predicted in the Test Dataset

UNREFINED DATASET IMPROVEMENT RUN

3.1. RGB Color Space

3.1.1. CNN AlexNet with Single Color

Unrefined Dataset Improvement Run

CNN AlexNet Single Color with RGB

Total Duration: 2617.87 sec.

Learning Rate: 0.000070

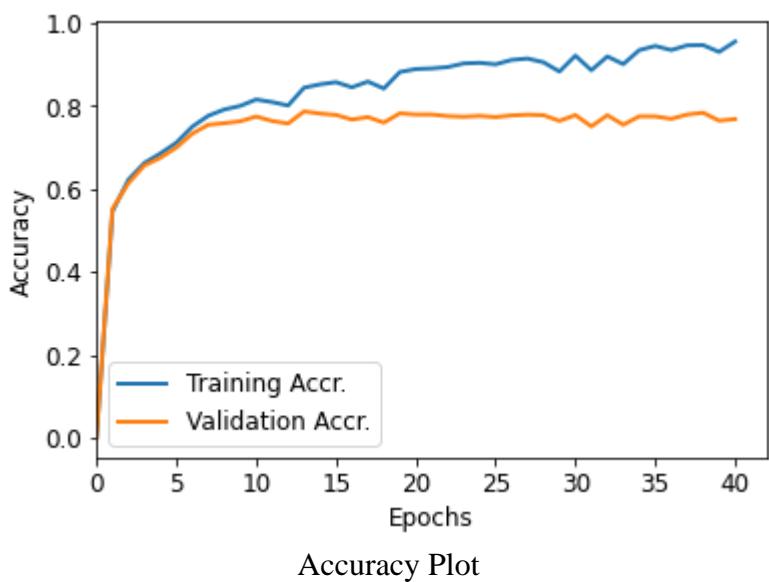
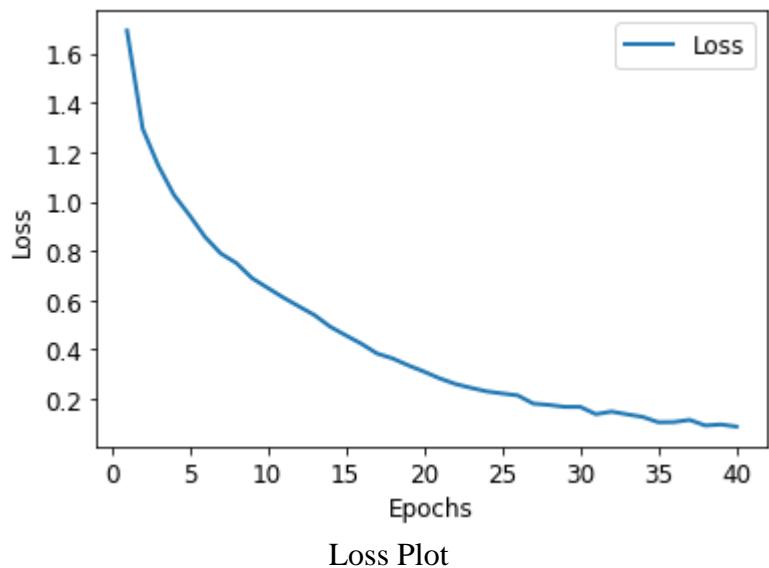
Dropout: 0.10

Batch Size: 32

Epoch: 40

Best Performing Epoch: 13

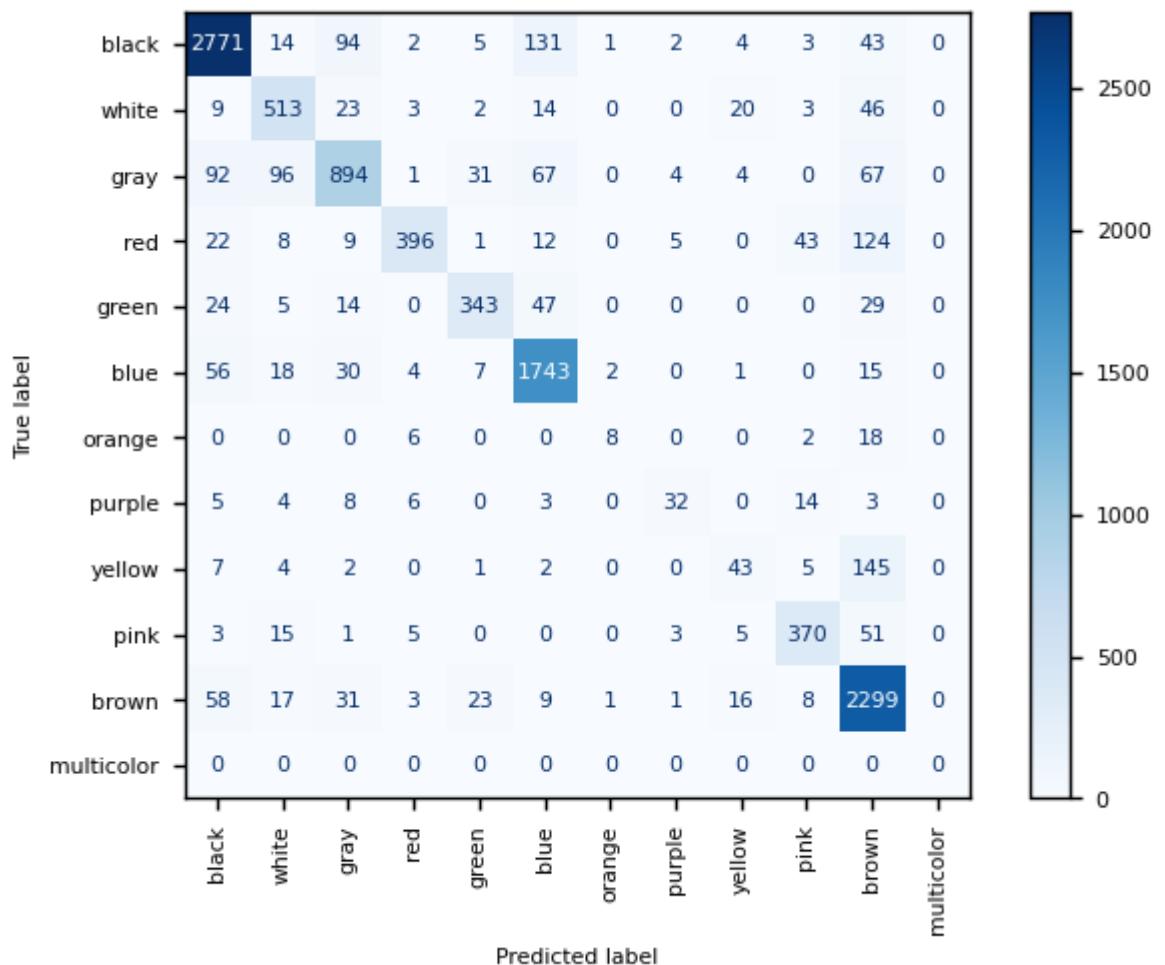
Summary	
Method	Accuracy
Training Dataset	0.8438
Validation Dataset	0.7866
Test Dataset	0.7804



Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9094	0.9026	0.9060	3070
white	0.7392	0.8104	0.7732	633
gray	0.8083	0.7118	0.7570	1256
red	0.9296	0.6387	0.7572	620
green	0.8305	0.7424	0.7840	462
blue	0.8595	0.9291	0.8929	1876
orange	0.6667	0.2353	0.3478	34
purple	0.6809	0.4267	0.5246	75
yellow	0.4624	0.2057	0.2848	209
pink	0.8259	0.8168	0.8213	453
brown	0.8095	0.9323	0.8666	2466
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.8438	0.8438	0.8438	11154
macro avg	0.7101	0.6127	0.6429	11154
weighted avg	0.8417	0.8438	0.8381	11154

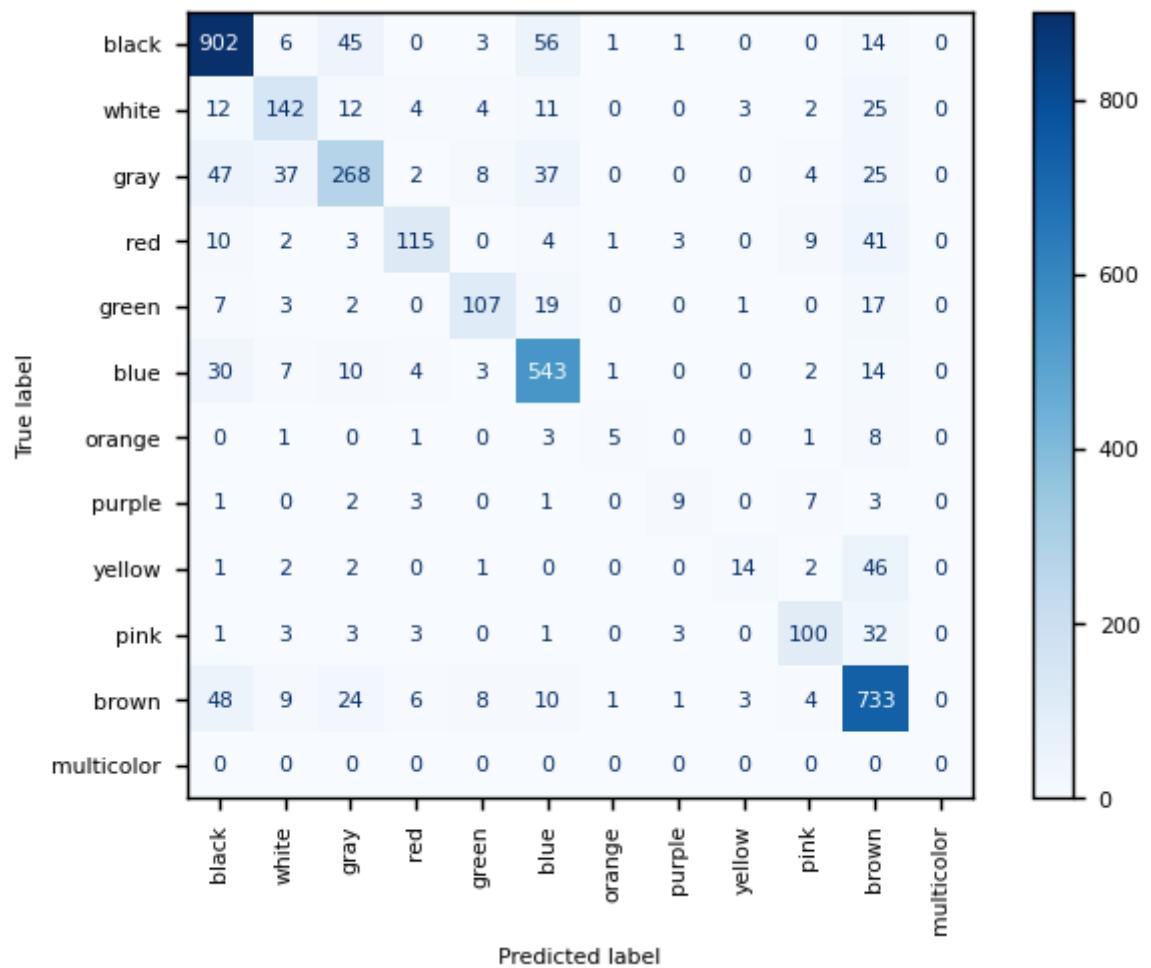


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8517	0.8774	0.8644	1028
white	0.6698	0.6605	0.6651	215
gray	0.7224	0.6262	0.6708	428
red	0.8333	0.6117	0.7055	188
green	0.7985	0.6859	0.7379	156
blue	0.7927	0.8844	0.8360	614
orange	0.5556	0.2632	0.3571	19
purple	0.5294	0.3462	0.4186	26
yellow	0.6667	0.2059	0.3146	68
pink	0.7634	0.6849	0.7220	146
brown	0.7651	0.8654	0.8122	847
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7866	0.7866	0.7866	3735
macro avg	0.6624	0.5593	0.5920	3735
weighted avg	0.7834	0.7866	0.7797	3735

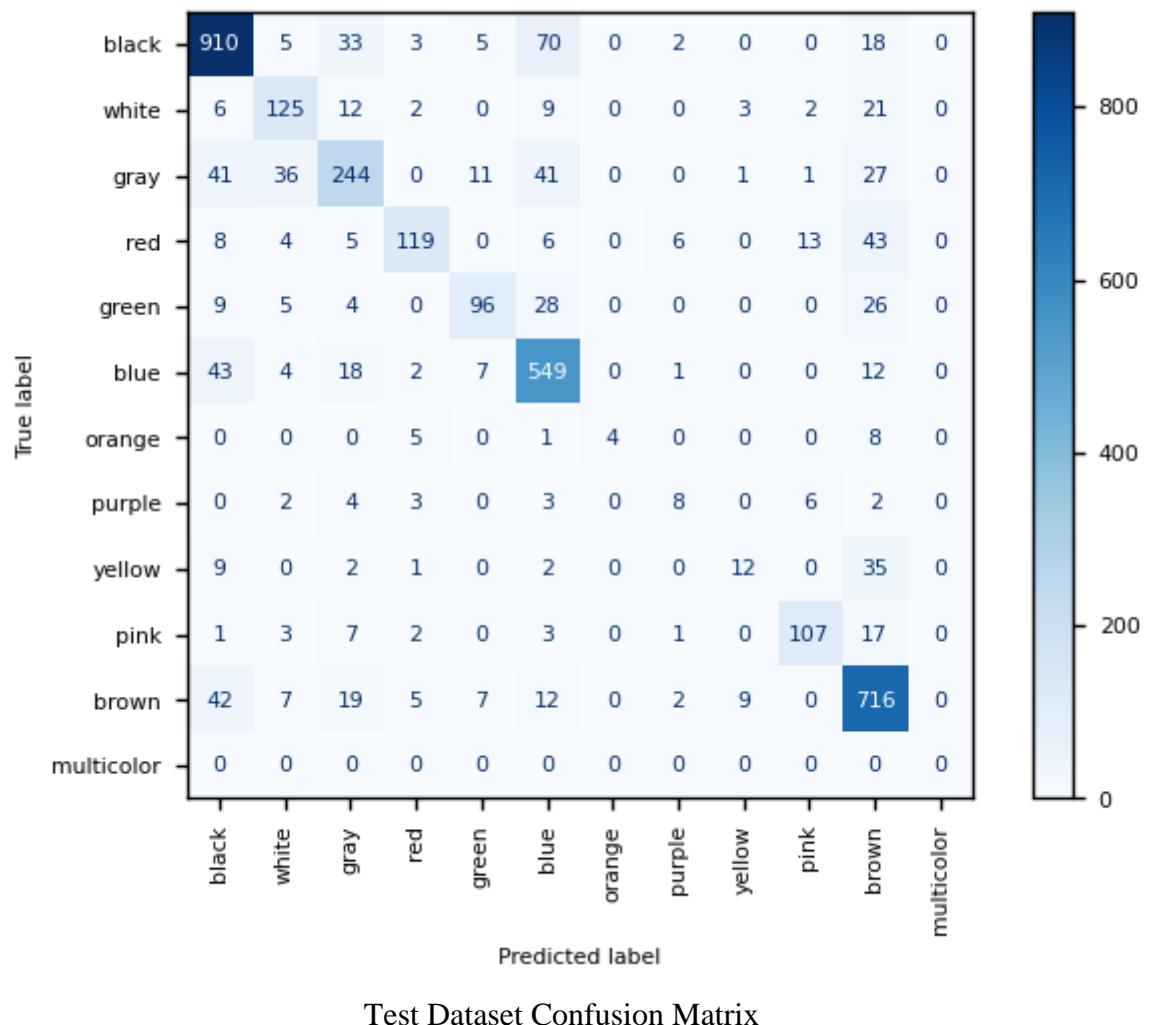


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8513	0.8700	0.8605	1046
white	0.6545	0.6944	0.6739	180
gray	0.7011	0.6070	0.6507	402
red	0.8380	0.5833	0.6879	204
green	0.7619	0.5714	0.6531	168
blue	0.7583	0.8632	0.8074	636
orange	1.0000	0.2222	0.3636	18
purple	0.4000	0.2857	0.3333	28
yellow	0.4800	0.1967	0.2791	61
pink	0.8295	0.7589	0.7926	141
brown	0.7741	0.8742	0.8211	819
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7804	0.7804	0.7804	3703
macro avg	0.6707	0.5439	0.5769	3703
weighted avg	0.7779	0.7804	0.7733	3703



gray (80.7%) blue (100.0%) black (72.1%) black (23.7%) blue (96.7%) green (72.9%)
 >> gray >> blue >> brown >> black >> blue >> blue



black (63.3%) brown (98.1%) black (79.4%) black (98.7%) blue (94.5%) blue (82.5%)
 >> black >> brown >> black >> black >> blue >> blue



black (58.9%) black (99.7%) black (99.9%) black (52.3%) brown (93.9%) brown (31.7%)
 >> black >> black >> black >> black >> brown >> brown



blue (23.8%) black (98.7%) black (96.2%) black (97.4%) brown (92.3%) black (82.4%)
 >> blue >> black >> black >> black >> brown >> black



blue (77.8%) brown (87.2%) brown (97.1%) brown (47.9%) blue (99.4%) gray (72.0%)
 >> brown >> red >> yellow >> brown >> blue >> gray



gray (44.4%) brown (24.6%) black (42.0%) black (43.3%) brown (89.0%) brown (52.4%)
 >> gray >> blue >> brown >> brown >> brown >> blue



Sample Images from Test Dataset

blue (36.4%) white (57.9%) black (42.0%) green (54.0%) brown (97.1%) brown (68.6%)
 >> black >> purple >> brown >> gray >> yellow >> yellow



brown (63.3%) brown (48.7%) brown (52.4%) blue (38.3%) black (34.9%) black (59.5%)
 >> pink >> orange >> blue >> gray >> gray >> blue



black (72.1%) brown (87.2%) gray (95.9%) blue (52.9%) brown (42.7%) black (37.1%)
 >> brown >> red >> black >> black >> red >> brown



brown (24.6%) black (68.4%) green (72.9%) brown (21.4%) yellow (29.1%) gray (60.5%)
 >> blue >> blue >> blue >> gray >> white >> brown



brown (44.2%) blue (77.8%) black (87.5%) blue (42.5%) white (29.1%) brown (36.6%)
 >> red >> brown >> yellow >> red >> brown >> green



purple (51.9%) brown (73.0%) white (75.1%) yellow (39.1%) brown (41.8%) red (58.0%)
 >> red >> yellow >> gray >> brown >> yellow >> brown



Sample Images that cannot be accurately predicted in the Test Dataset

3.1.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Quantization Fastoctree	98.80 sec.	0.4076
Quantization Mmcq	834.05 sec.	0.3918
Quantization Libimagequant	400.85 sec.	0.3451
Pairwise Argmin	18.93 sec.	0.3520
Pairwise Mean	17.21 sec.	0.3255
Array Mean Argmin	11.39 sec.	0.1240
Pairwise First Quantile	52.94 sec.	0.3434
Pairwise Median	53.89 sec.	0.3481
Pairwise Third Quantile	50.11 sec.	0.3061

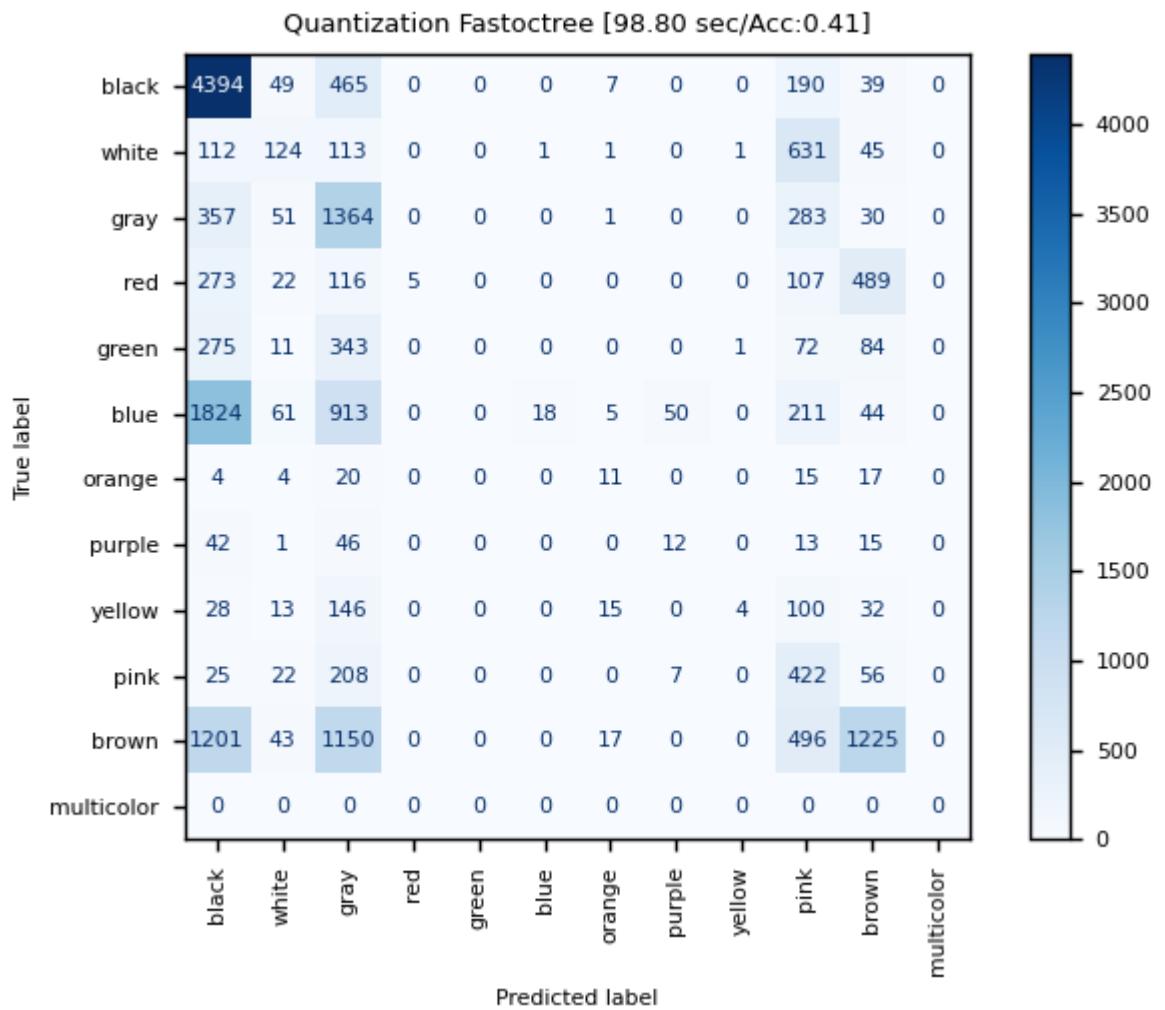
Unrefined Dataset Improvement Run

Single Color with RGB / Quantization Fastoctree Result

Total Duration: 98.80 sec.

Single Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.5148	0.8542	0.6424	5144
white	0.3092	0.1206	0.1735	1028
gray	0.2793	0.6539	0.3914	2086
red	1.0000	0.0049	0.0098	1012
green	0.0000	0.0000	0.0000	786
blue	0.9474	0.0058	0.0114	3126
orange	0.1930	0.1549	0.1719	71
purple	0.1739	0.0930	0.1212	129
yellow	0.6667	0.0118	0.0233	338
pink	0.1661	0.5703	0.2573	740
brown	0.5901	0.2965	0.3947	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4076	0.4076	0.4076	18592
macro avg	0.4034	0.2305	0.1831	18592
weighted avg	0.5564	0.4076	0.3336	18592



Single Color with RGB / Quantization Fastoctree Confusion Matrix

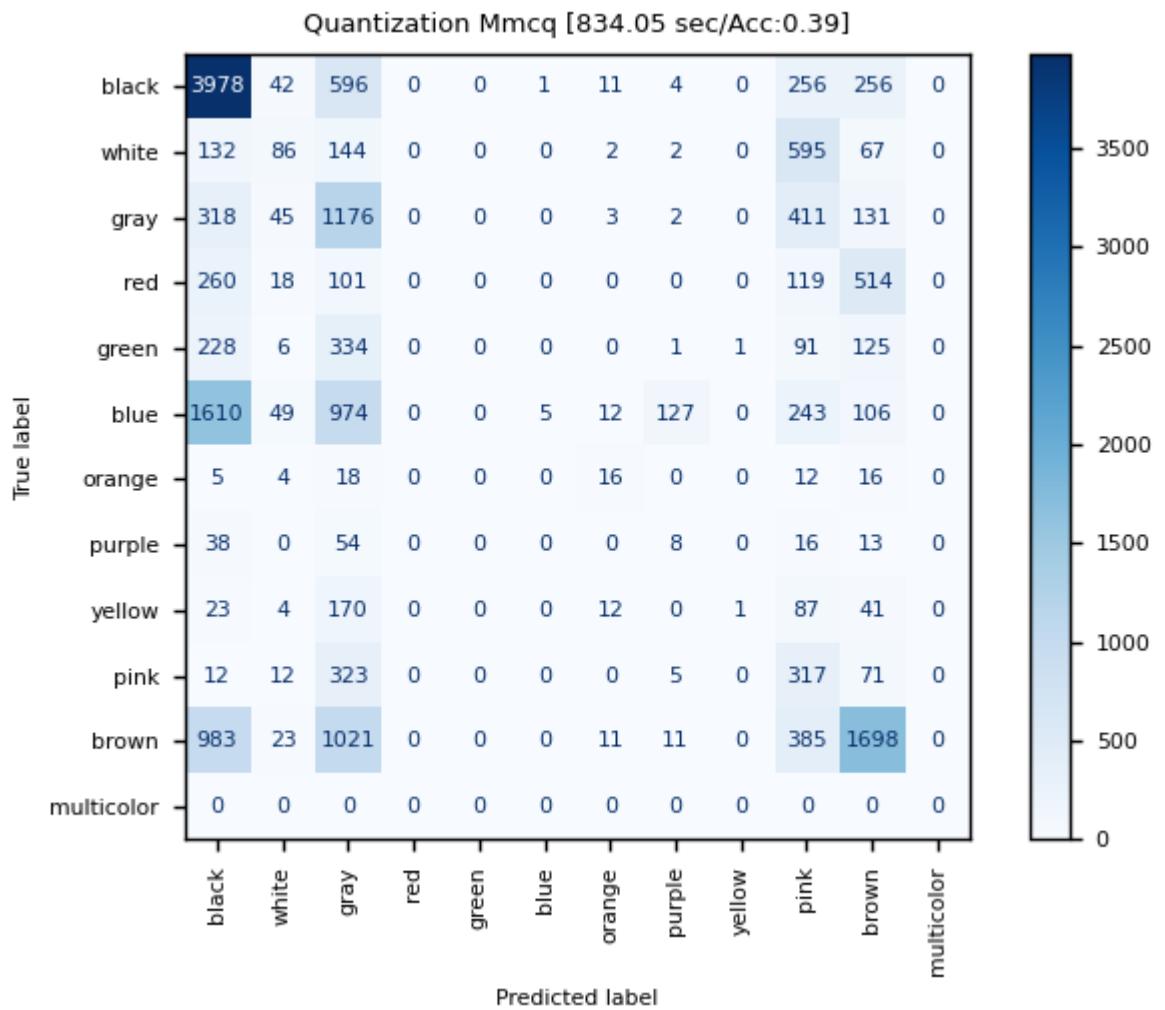
Unrefined Dataset Improvement Run

Single Color with RGB / Quantization Mmcq Result

Total Duration: 834.05 sec.

Single Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.5243	0.7733	0.6249	5144
white	0.2976	0.0837	0.1306	1028
gray	0.2395	0.5638	0.3361	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.8333	0.0016	0.0032	3126
orange	0.2388	0.2254	0.2319	71
purple	0.0500	0.0620	0.0554	129
yellow	0.5000	0.0030	0.0059	338
pink	0.1252	0.4284	0.1938	740
brown	0.5589	0.4109	0.4736	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3918	0.3918	0.3918	18592
macro avg	0.2806	0.2127	0.1713	18592
weighted avg	0.4681	0.3918	0.3327	18592



Single Color with RGB / Quantization Mmcq Confusion Matrix

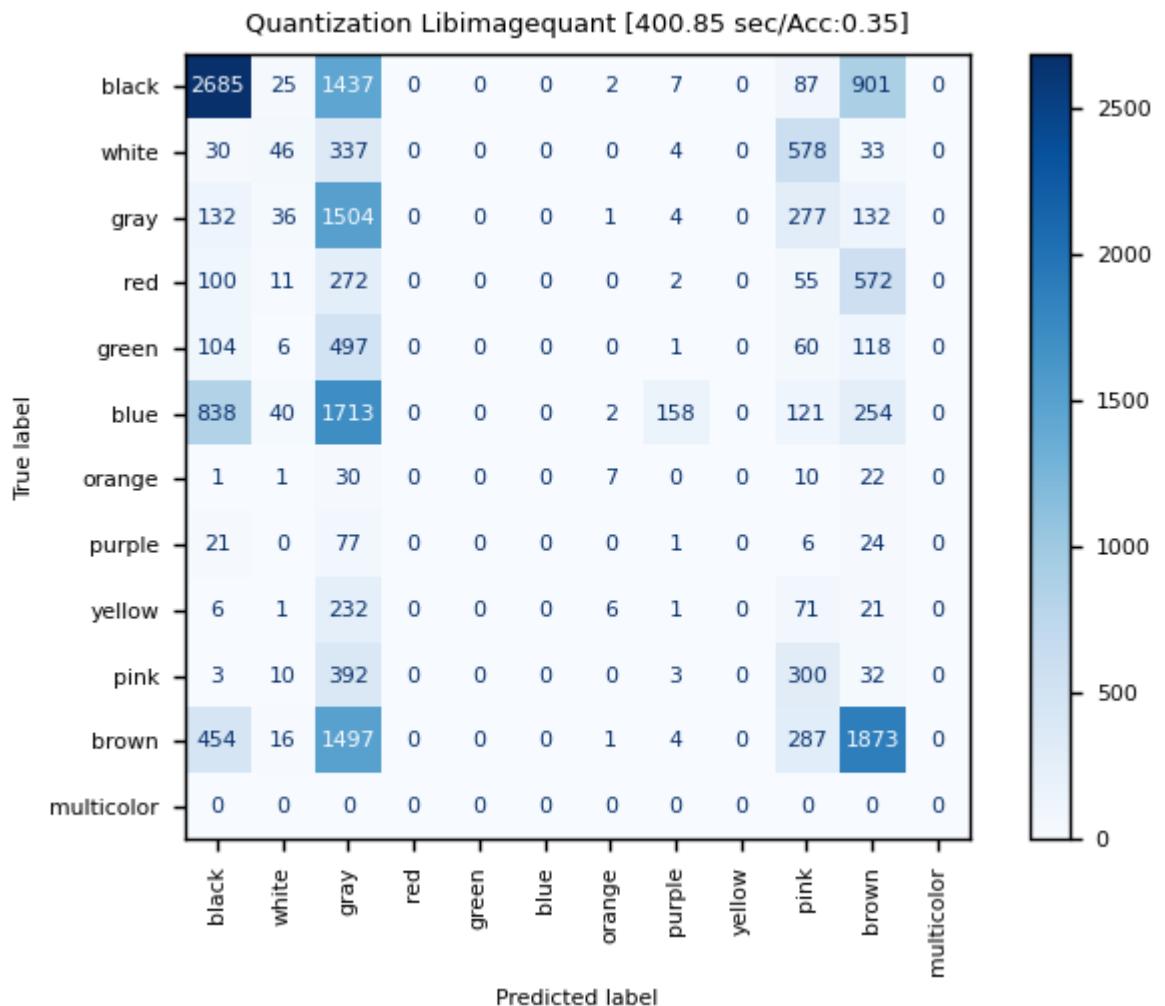
Unrefined Dataset Improvement Run

Single Color with RGB / Quantization Libimagequant Result

Total Duration: 400.85 sec.

Single Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.6139	0.5220	0.5642	5144
white	0.2396	0.0447	0.0754	1028
gray	0.1883	0.7210	0.2986	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.3684	0.0986	0.1556	71
purple	0.0054	0.0078	0.0064	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1620	0.4054	0.2315	740
brown	0.4704	0.4533	0.4617	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3451	0.3451	0.3451	18592
macro avg	0.1707	0.1877	0.1494	18592
weighted avg	0.3166	0.3451	0.3062	18592



Single Color with RGB / Quantization Libimagequant Confusion Matrix

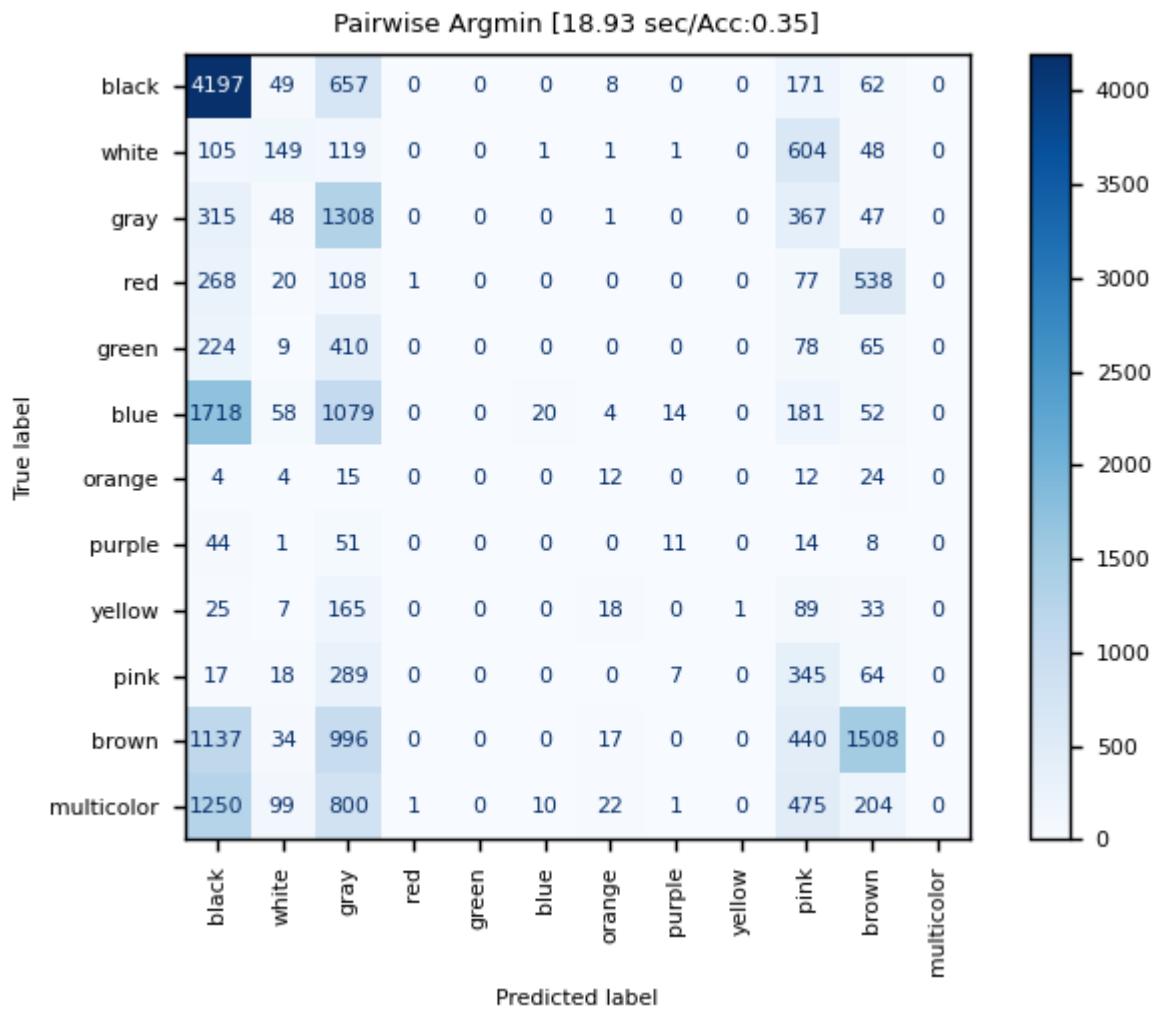
Unrefined Dataset Improvement Run

Single Color with RGB / Pairwise Argmin Result

Total Duration: 18.93 sec.

Single Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.4511	0.8159	0.5810	5144
white	0.3004	0.1449	0.1955	1028
gray	0.2181	0.6270	0.3236	2086
red	0.5000	0.0010	0.0020	1012
green	0.0000	0.0000	0.0000	786
blue	0.6452	0.0064	0.0127	3126
orange	0.1446	0.1690	0.1558	71
purple	0.3235	0.0853	0.1350	129
yellow	1.0000	0.0030	0.0059	338
pink	0.1209	0.4662	0.1920	740
brown	0.5684	0.3650	0.4445	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.3520	
macro avg	0.3560	0.2236	0.1707	21454
weighted avg	0.3932	0.3520	0.2757	21454



Single Color with RGB / Pairwise Argmin Confusion Matrix

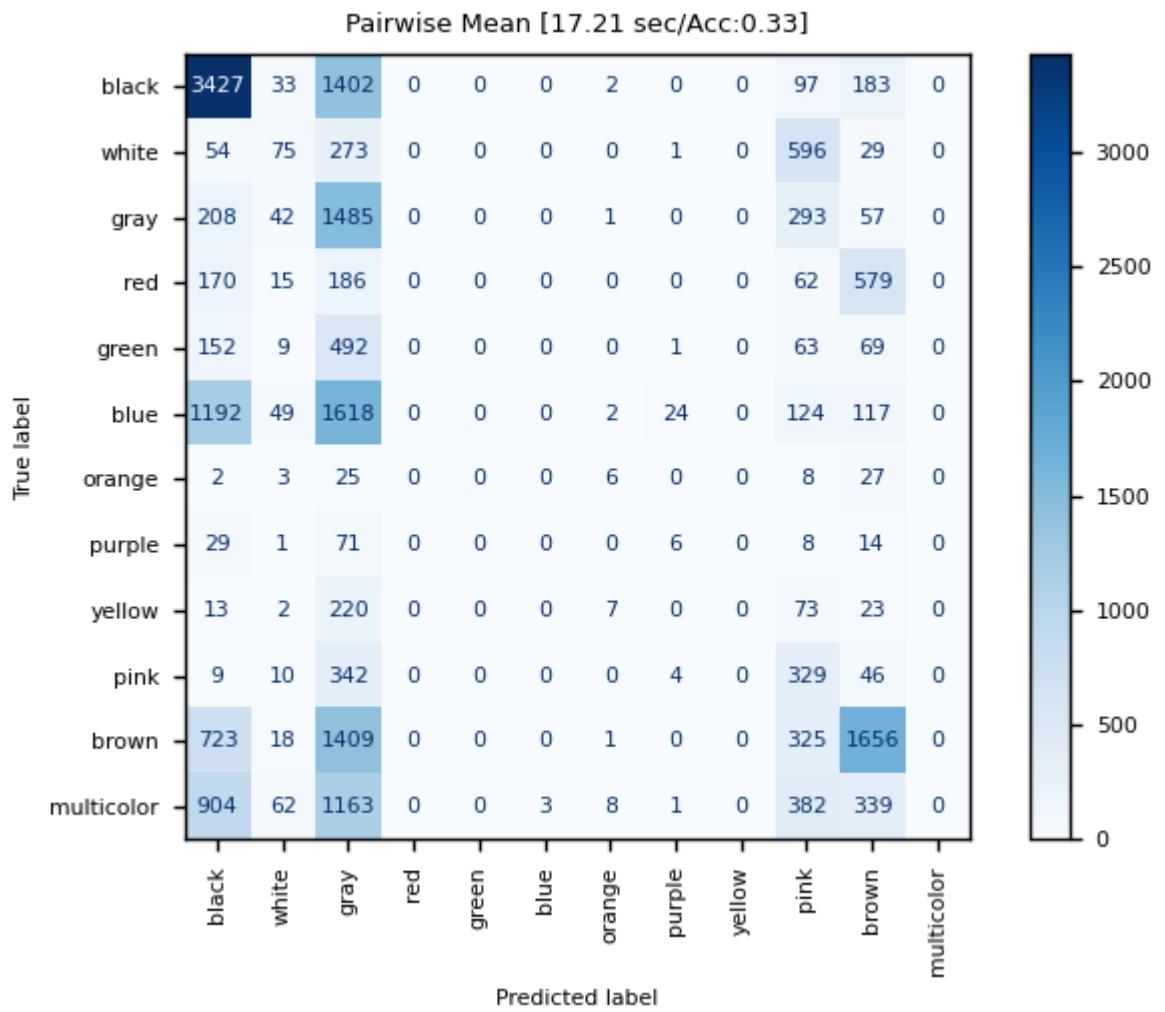
Unrefined Dataset Improvement Run

Single Color with RGB / Pairwise Mean Result

Total Duration: 17.21 sec.

Single Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.4979	0.6662	0.5699	5144
white	0.2351	0.0730	0.1114	1028
gray	0.1710	0.7119	0.2757	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.2222	0.0845	0.1224	71
purple	0.1622	0.0465	0.0723	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1394	0.4446	0.2123	740
brown	0.5276	0.4008	0.4555	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.3255	
macro avg	0.1629	0.2023	0.1516	21454
weighted avg	0.2554	0.3255	0.2647	21454



Single Color with RGB / Pairwise Mean Confusion Matrix

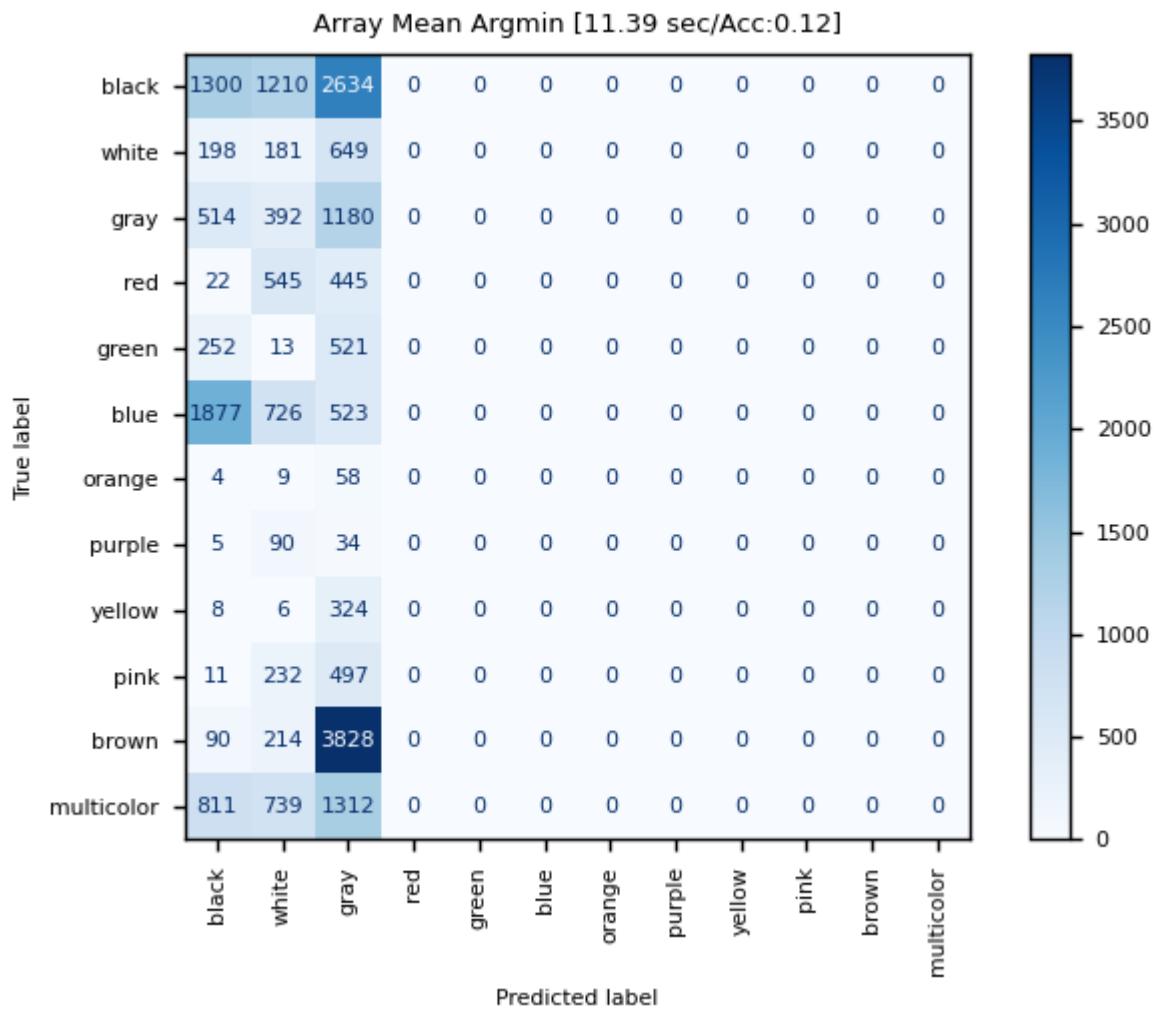
Unrefined Dataset Improvement Run

Single Color with RGB / Array Mean Argmin Result

Total Duration: 11.39 sec.

Single Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2553	0.2527	0.2540	5144
white	0.0415	0.1761	0.0672	1028
gray	0.0983	0.5657	0.1675	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.0000	0.0000	0.0000	740
brown	0.0000	0.0000	0.0000	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.1240	
macro avg	0.0329	0.0829	0.0407	21454
weighted avg	0.0728	0.1240	0.0804	21454



Single Color with RGB / Array Mean Argmin Confusion Matrix

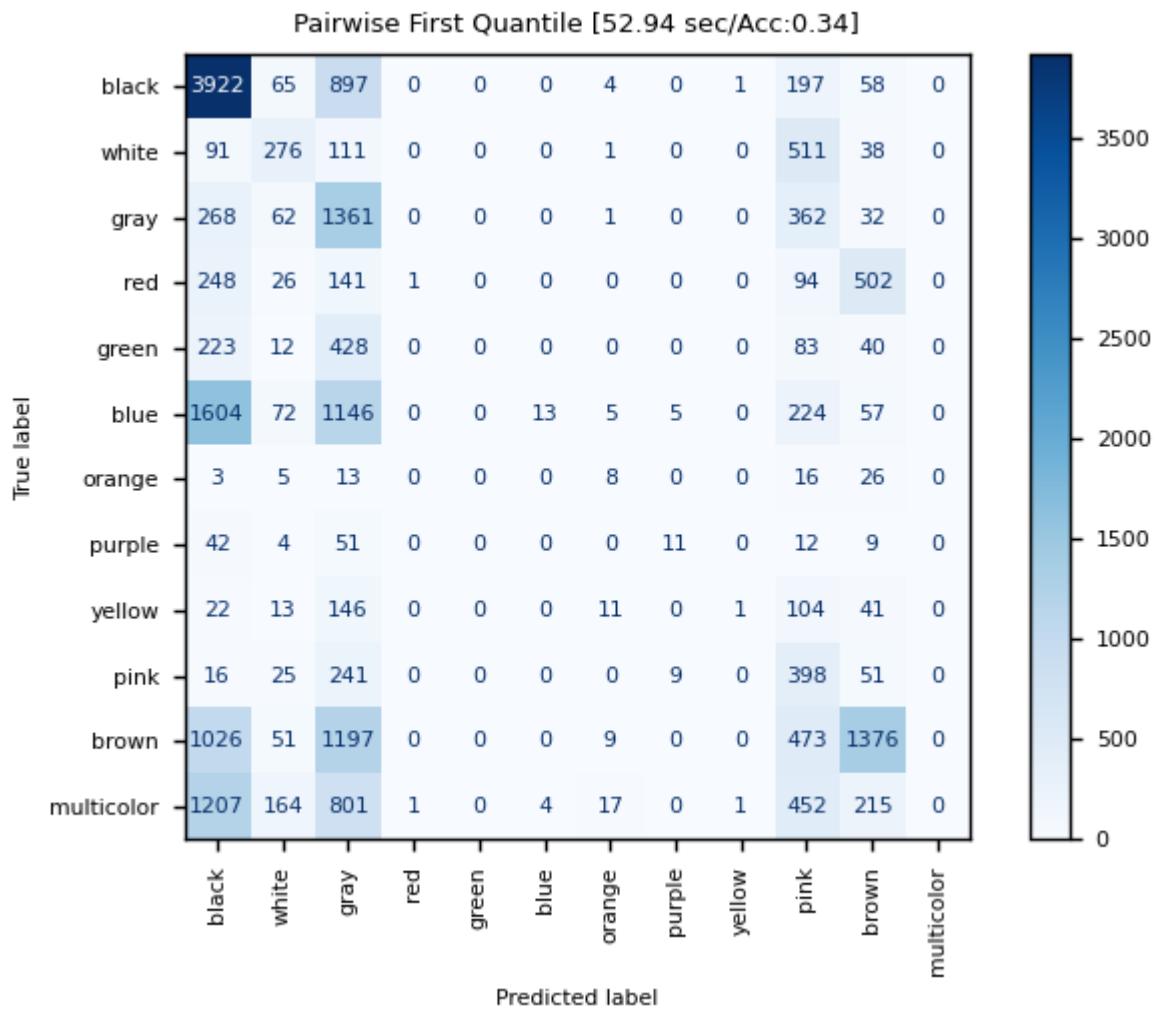
Unrefined Dataset Improvement Run

Single Color with RGB / Pairwise First Quantile Result

Total Duration: 52.94 sec.

Single Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4523	0.7624	0.5677	5144
white	0.3561	0.2685	0.3062	1028
gray	0.2083	0.6524	0.3158	2086
red	0.5000	0.0010	0.0020	1012
green	0.0000	0.0000	0.0000	786
blue	0.7647	0.0042	0.0083	3126
orange	0.1429	0.1127	0.1260	71
purple	0.4400	0.0853	0.1429	129
yellow	0.3333	0.0030	0.0059	338
pink	0.1360	0.5378	0.2171	740
brown	0.5628	0.3330	0.4184	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.3434	
macro avg	0.3247	0.2300	0.1759	21454
weighted avg	0.4022	0.3434	0.2722	21454



Single Color with RGB / Pairwise First Quantile Confusion Matrix

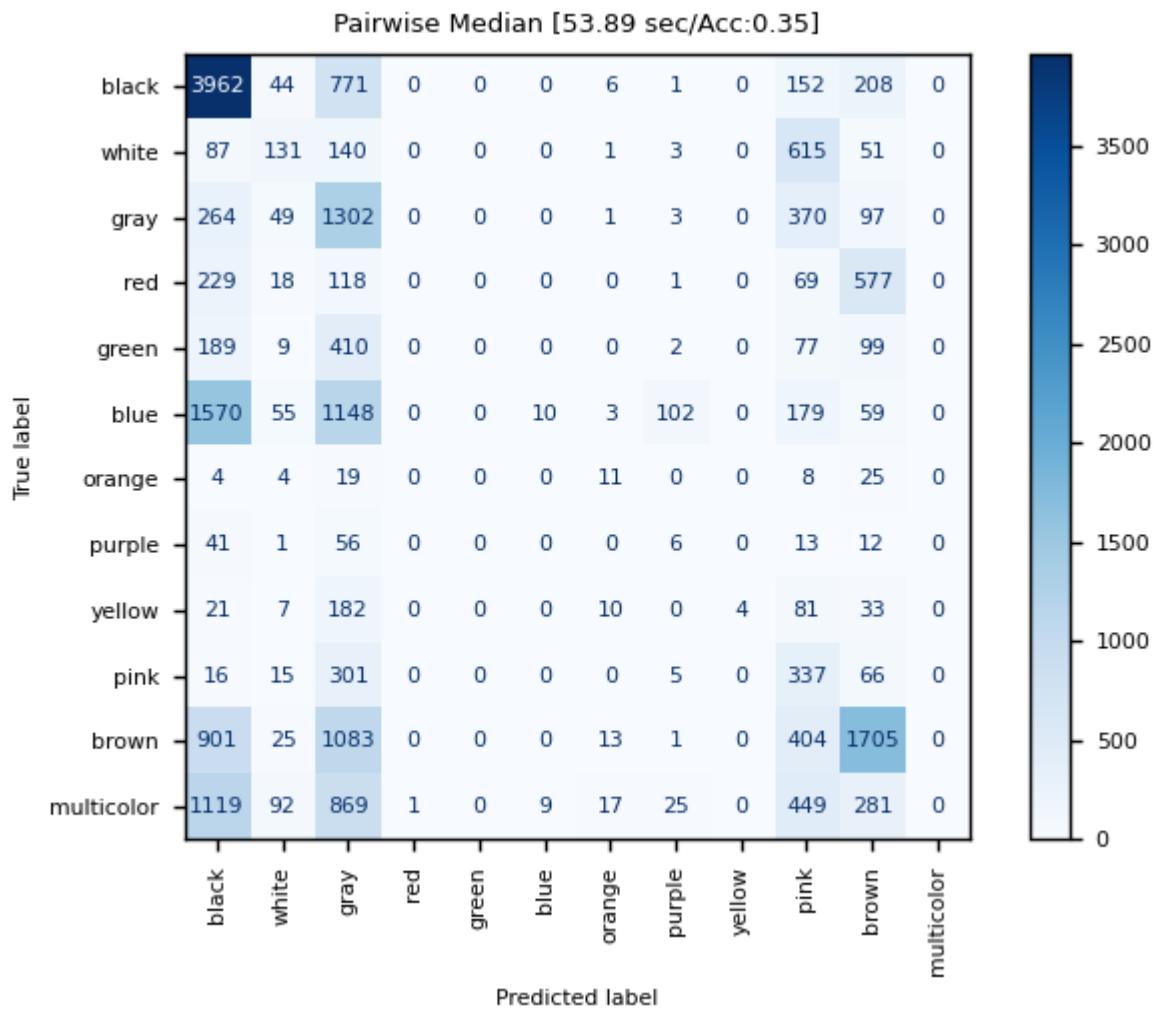
Unrefined Dataset Improvement Run

Single Color with RGB / Pairwise Median Result

Total Duration: 53.89 sec.

Single Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4715	0.7702	0.5849	5144
white	0.2911	0.1274	0.1773	1028
gray	0.2035	0.6242	0.3069	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.5263	0.0032	0.0064	3126
orange	0.1774	0.1549	0.1654	71
purple	0.0403	0.0465	0.0432	129
yellow	1.0000	0.0118	0.0234	338
pink	0.1224	0.4554	0.1929	740
brown	0.5307	0.4126	0.4643	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.3481	
macro avg	0.2803	0.2172	0.1637	21454
weighted avg	0.3465	0.3481	0.2768	21454



Single Color with RGB / Pairwise Median Confusion Matrix

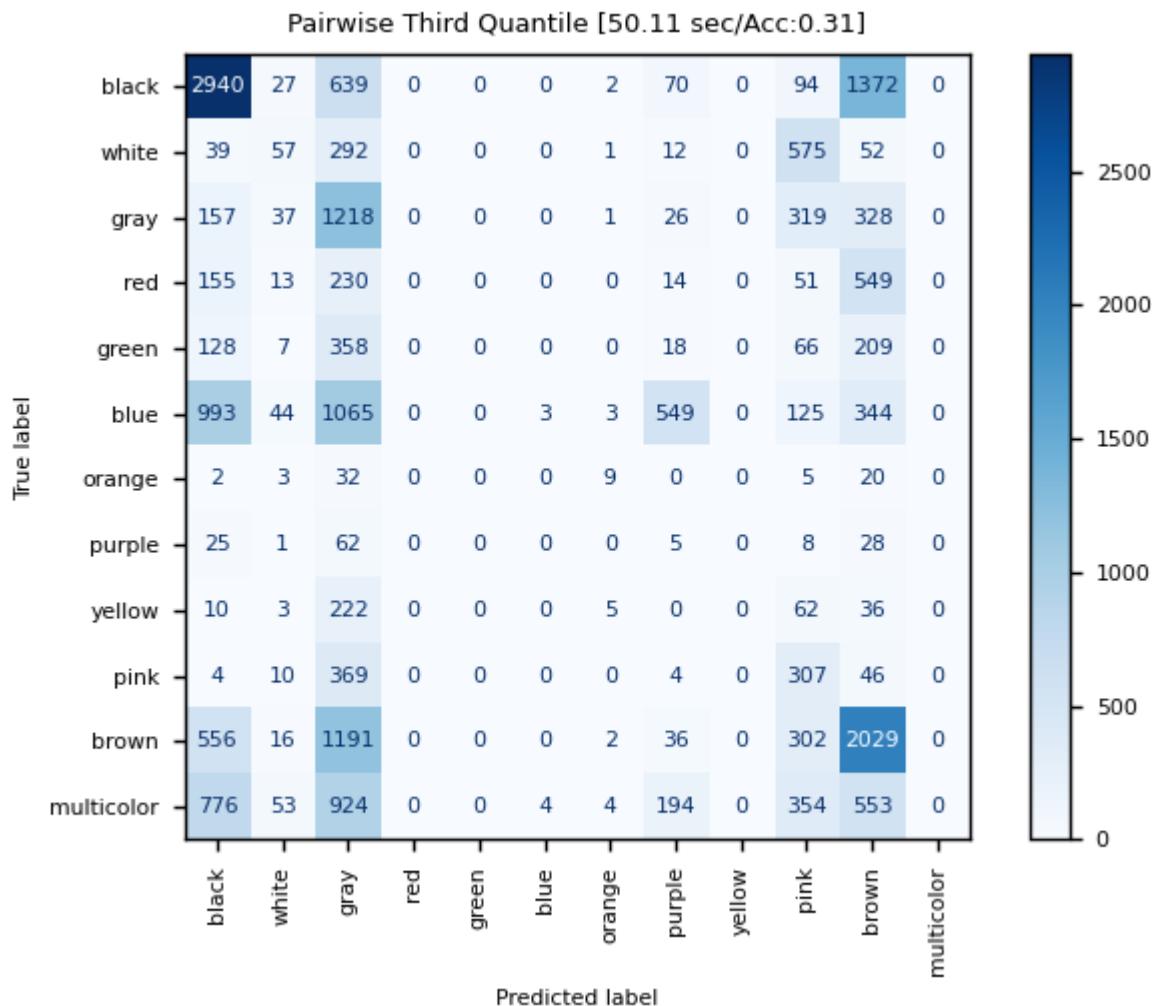
Unrefined Dataset Improvement Run

Single Color with RGB / Pairwise Third Quantile Result

Total Duration: 50.11 sec.

Single Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5082	0.5715	0.5380	5144
white	0.2103	0.0554	0.0878	1028
gray	0.1845	0.5839	0.2804	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.4286	0.0010	0.0019	3126
orange	0.3333	0.1268	0.1837	71
purple	0.0054	0.0388	0.0095	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1354	0.4149	0.2041	740
brown	0.3645	0.4910	0.4184	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.3061	
macro avg	0.1809	0.1903	0.1436	21454
weighted avg	0.2883	0.3061	0.2490	21454



Single Color with RGB / Pairwise Third Quantile Confusion Matrix

3.1.3. CNN AlexNet with Multi-color

Unrefined Dataset Improvement Run

CNN AlexNet Multi-Color with RGB

Total Duration: 2946.28 sec.

Learning Rate: 0.000070

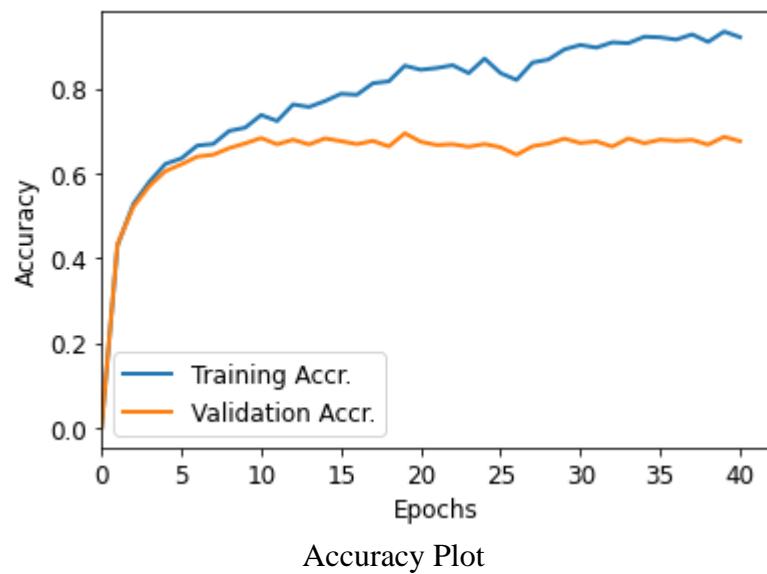
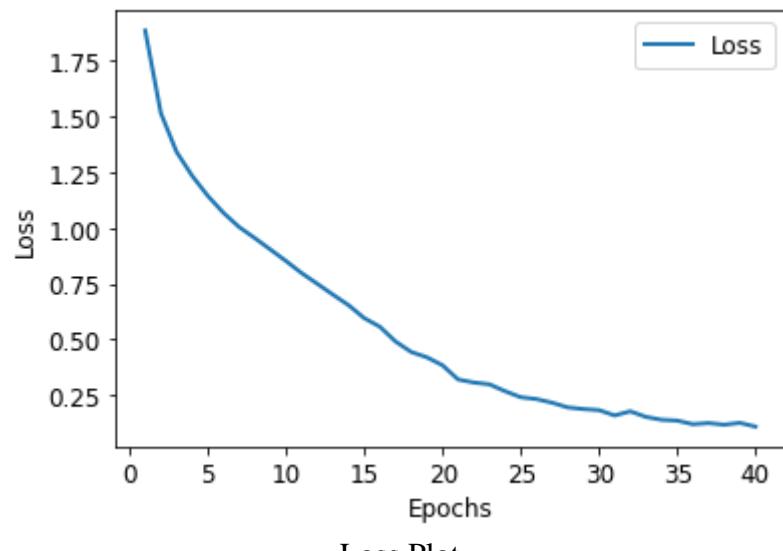
Dropout: 0.10

Batch Size: 32

Epoch: 40

Best Performing Epoch: 19

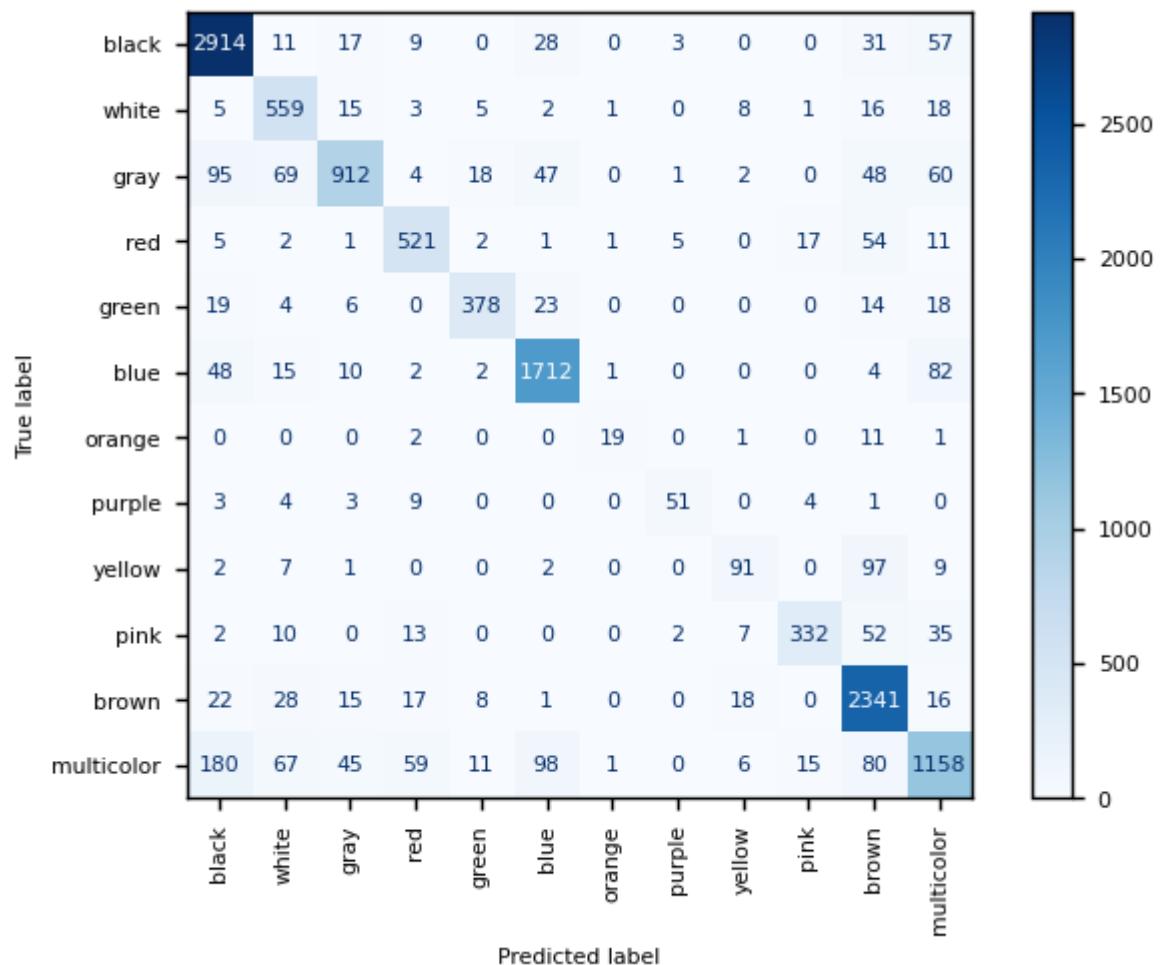
Summary	
Method	Accuracy
Training Dataset	0.8535
Validation Dataset	0.6942
Test Dataset	0.6867



Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8844	0.9492	0.9156	3070
white	0.7204	0.8831	0.7935	633
gray	0.8898	0.7261	0.7996	1256
red	0.8153	0.8403	0.8276	620
green	0.8915	0.8182	0.8533	462
blue	0.8945	0.9126	0.9034	1876
orange	0.8261	0.5588	0.6667	34
purple	0.8226	0.6800	0.7445	75
yellow	0.6842	0.4354	0.5322	209
pink	0.8997	0.7329	0.8078	453
brown	0.8516	0.9493	0.8978	2466
multicolor	0.7904	0.6733	0.7272	1720
accuracy			0.8535	
macro avg	0.8309	0.7633	0.7891	12874
weighted avg	0.8532	0.8535	0.8498	12874

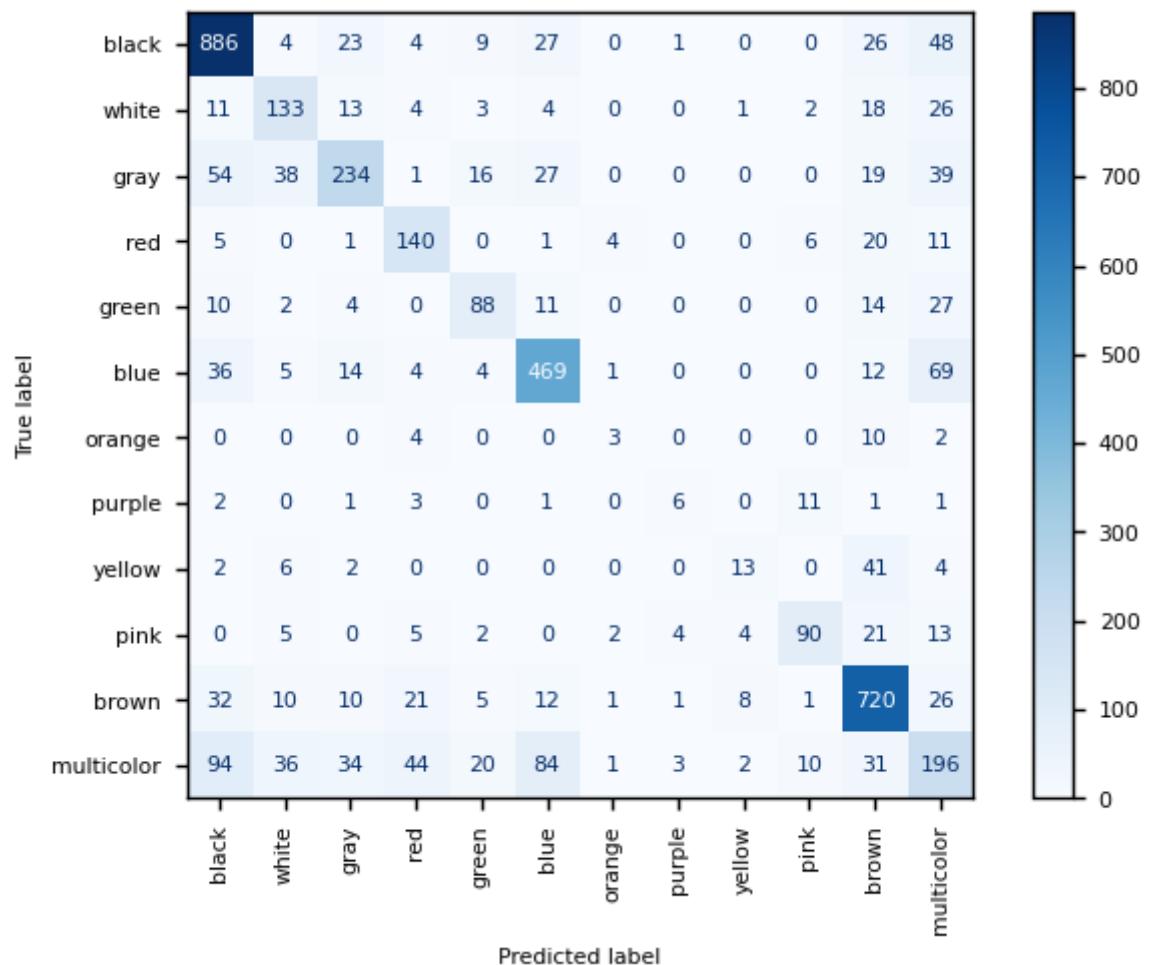


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7827	0.8619	0.8204	1028
white	0.5565	0.6186	0.5859	215
gray	0.6964	0.5467	0.6126	428
red	0.6087	0.7447	0.6699	188
green	0.5986	0.5641	0.5809	156
blue	0.7374	0.7638	0.7504	614
orange	0.2500	0.1579	0.1935	19
purple	0.4000	0.2308	0.2927	26
yellow	0.4643	0.1912	0.2708	68
pink	0.7500	0.6164	0.6767	146
brown	0.7717	0.8501	0.8090	847
multicolor	0.4242	0.3532	0.3854	555
accuracy			0.6942	
macro avg	0.5867	0.5416	0.5540	4290
weighted avg	0.6826	0.6942	0.6845	4290

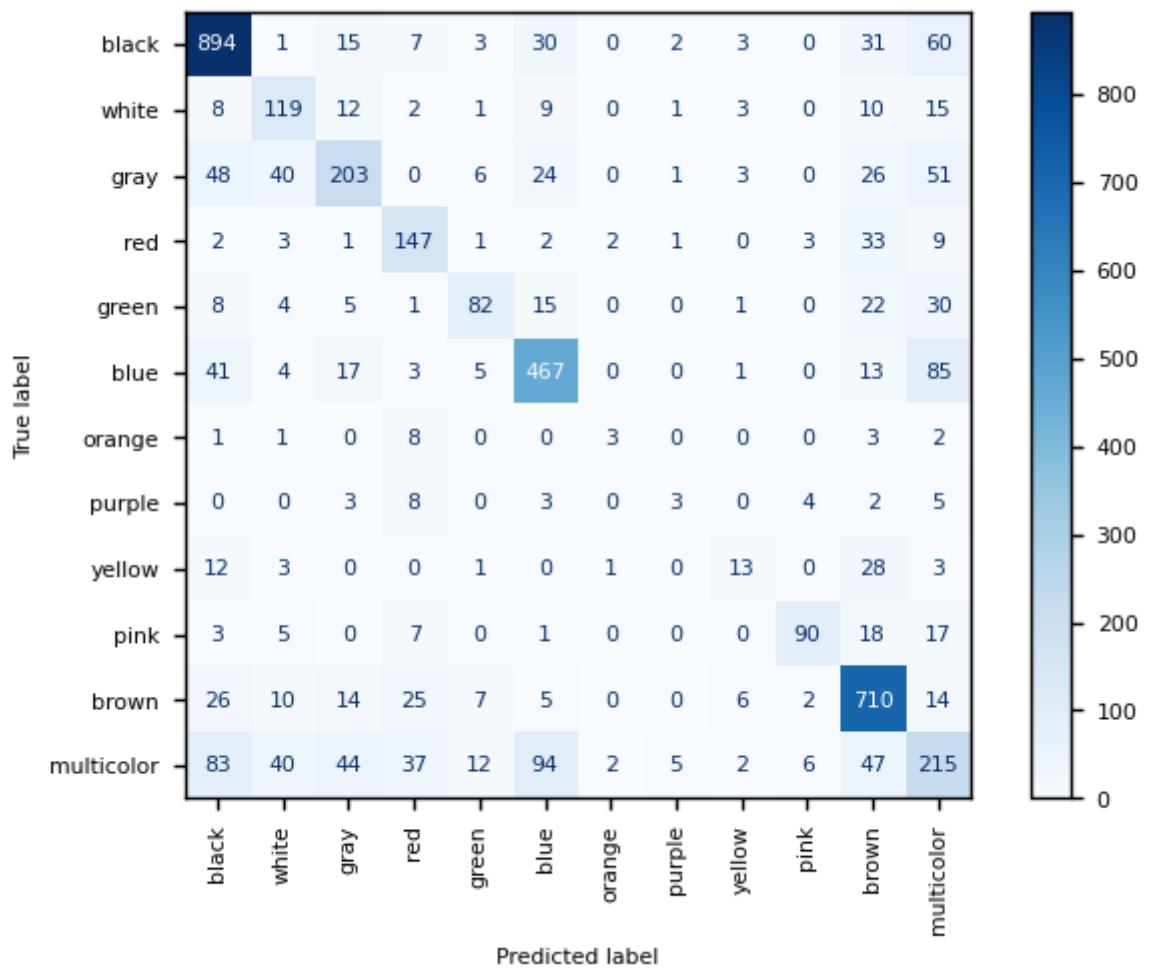


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7940	0.8547	0.8232	1046
white	0.5174	0.6611	0.5805	180
gray	0.6465	0.5050	0.5670	402
red	0.6000	0.7206	0.6548	204
green	0.6949	0.4881	0.5734	168
blue	0.7185	0.7343	0.7263	636
orange	0.3750	0.1667	0.2308	18
purple	0.2308	0.1071	0.1463	28
yellow	0.4062	0.2131	0.2796	61
pink	0.8571	0.6383	0.7317	141
brown	0.7529	0.8669	0.8059	819
multicolor	0.4249	0.3663	0.3934	587
accuracy			0.6867	
macro avg	0.5849	0.5268	0.5427	4290
weighted avg	0.6770	0.6867	0.6771	4290



Test Dataset Confusion Matrix

blue (66.6%) black (96.0%) black (99.7%) black (98.9%) multicolor (96.5%) black (86.9%)
 >> blue >> black >> black >> black >> multicolor >> black



white (94.3%) black (58.7%) black (98.6%) gray (58.6%) white (79.4%) black (97.7%)
 >> white >> multicolor >> black >> gray >> brown >> black



brown (62.0%) black (99.8%) black (87.9%) brown (97.0%) blue (86.3%) white (38.0%)
 >> brown >> black >> black >> brown >> blue >> white



brown (70.9%) multicolor (59.0%) brown (100.0%) brown (99.8%) black (73.5%) blue (97.9%)
 >> yellow >> multicolor >> blue >> brown >> black >> blue



multicolor (40.6%) brown (89.1%) brown (99.7%) red (86.9%) multicolor (99.9%) brown (89.5%)
 >> gray >> brown >> brown >> orange >> multicolor >> brown



brown (63.0%) black (99.0%) brown (86.8%) blue (34.9%) gray (60.6%) black (80.1%)
 >> red >> black >> red >> black >> gray >> black



Sample Images from Test Dataset

purple (71.4%) black (79.2%) white (96.6%) blue (34.9%) brown (100.0%)
 >> red >> gray >> gray >> black >> blue >> multicolor



white (79.4%) brown (99.9%) black (98.3%) blue (17.6%) green (50.0%) multicolor (99.5%)
 >> brown >> yellow >> multicolor >> brown >> brown >> black



black (58.7%) multicolor (59.6%) yellow (86.9%) multicolor (74.3%) yellow (75.0%) gray (44.9%)
 >> multicolor >> blue >> orange >> black >> gray >> blue



black (93.8%) multicolor (42.2%) brown (86.8%) black (24.5%) gray (83.5%) gray (56.2%)
 >> gray >> purple >> yellow >> pink >> multicolor >> multicolor



multicolor (65.2%) multicolor (94.7%) multicolor (41.8%) multicolor (50.6%) black (66.0%) black (70.7%)
 >> blue >> purple >> gray >> black >> brown >> yellow



multicolor (55.6%) black (93.7%) multicolor (81.9%) brown (63.0%) brown (86.8%) blue (45.0%)
 >> white >> gray >> black >> red >> red >> gray



Sample Images that cannot be accurately predicted in the Test Dataset

3.1.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Quantization Fastoctree	53.09 sec.	0.3468
Quantization Mmcq	880.56 sec.	0.3280
Quantization Libimagequant	386.91 sec.	0.2825
Pairwise Argmin	17.85 sec.	0.2734
Pairwise Mean	14.30 sec.	0.3001
Array Mean Argmin	10.34 sec.	0.1240
Pairwise First Quantile	51.70 sec.	0.3218
Pairwise Median	52.39 sec.	0.3315
Pairwise Third Quantile	48.80 sec.	0.3090

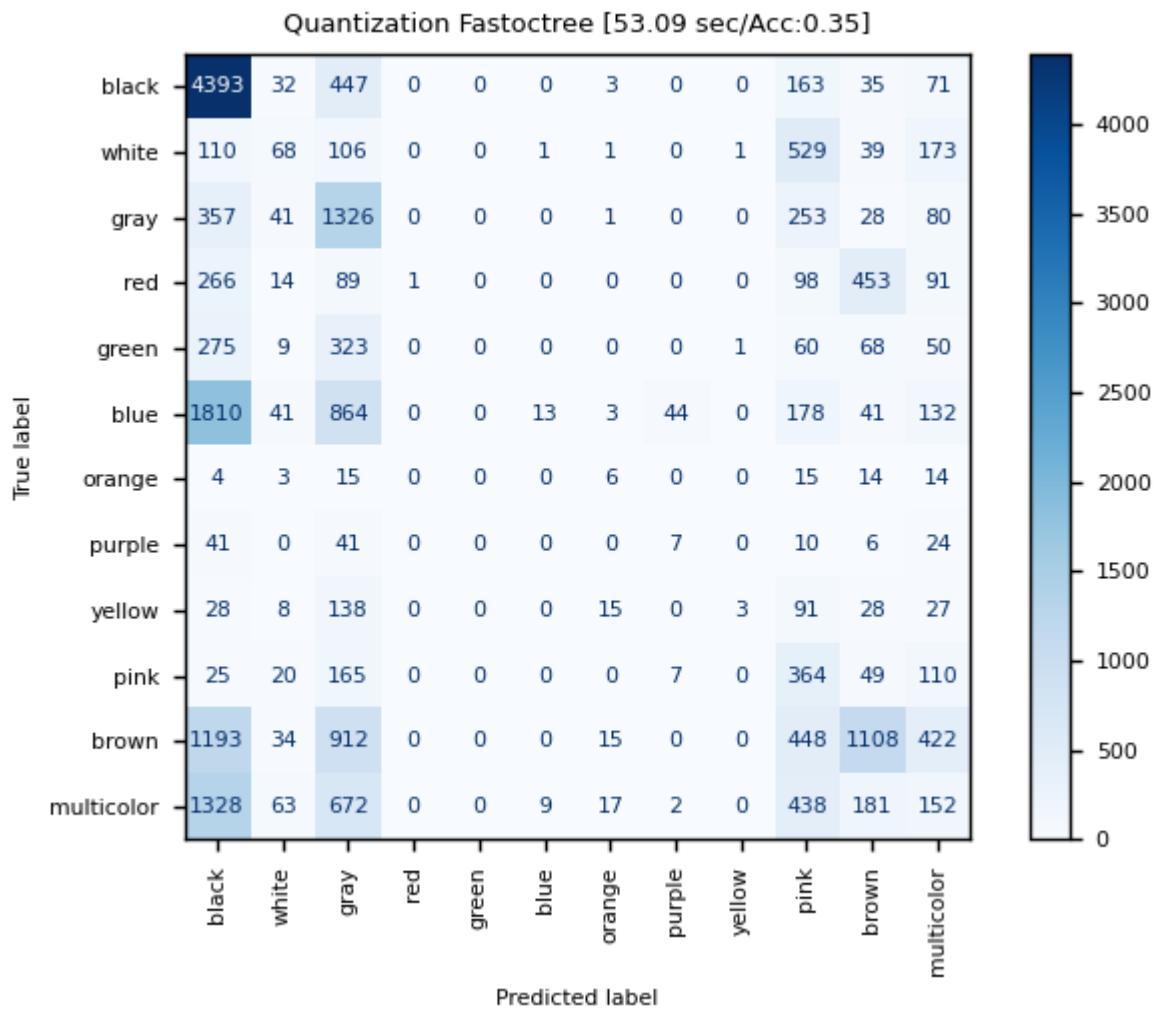
Unrefined Dataset Improvement Run

Multi-Color with RGB / Quantization Fastoctree Result

Total Duration: 53.09 sec.

Multi-Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.4469	0.8540	0.5868	5144
white	0.2042	0.0661	0.0999	1028
gray	0.2601	0.6357	0.3692	2086
red	1.0000	0.0010	0.0020	1012
green	0.0000	0.0000	0.0000	786
blue	0.5652	0.0042	0.0083	3126
orange	0.0984	0.0845	0.0909	71
purple	0.1167	0.0543	0.0741	129
yellow	0.6000	0.0089	0.0175	338
pink	0.1375	0.4919	0.2149	740
brown	0.5405	0.2682	0.3585	4132
multicolor	0.1129	0.0531	0.0722	2862
accuracy			0.3468	
macro avg	0.3402	0.2101	0.1578	21454
weighted avg	0.4061	0.3468	0.2698	21454



Multi-Color with RGB / Quantization Fastoctree Confusion Matrix

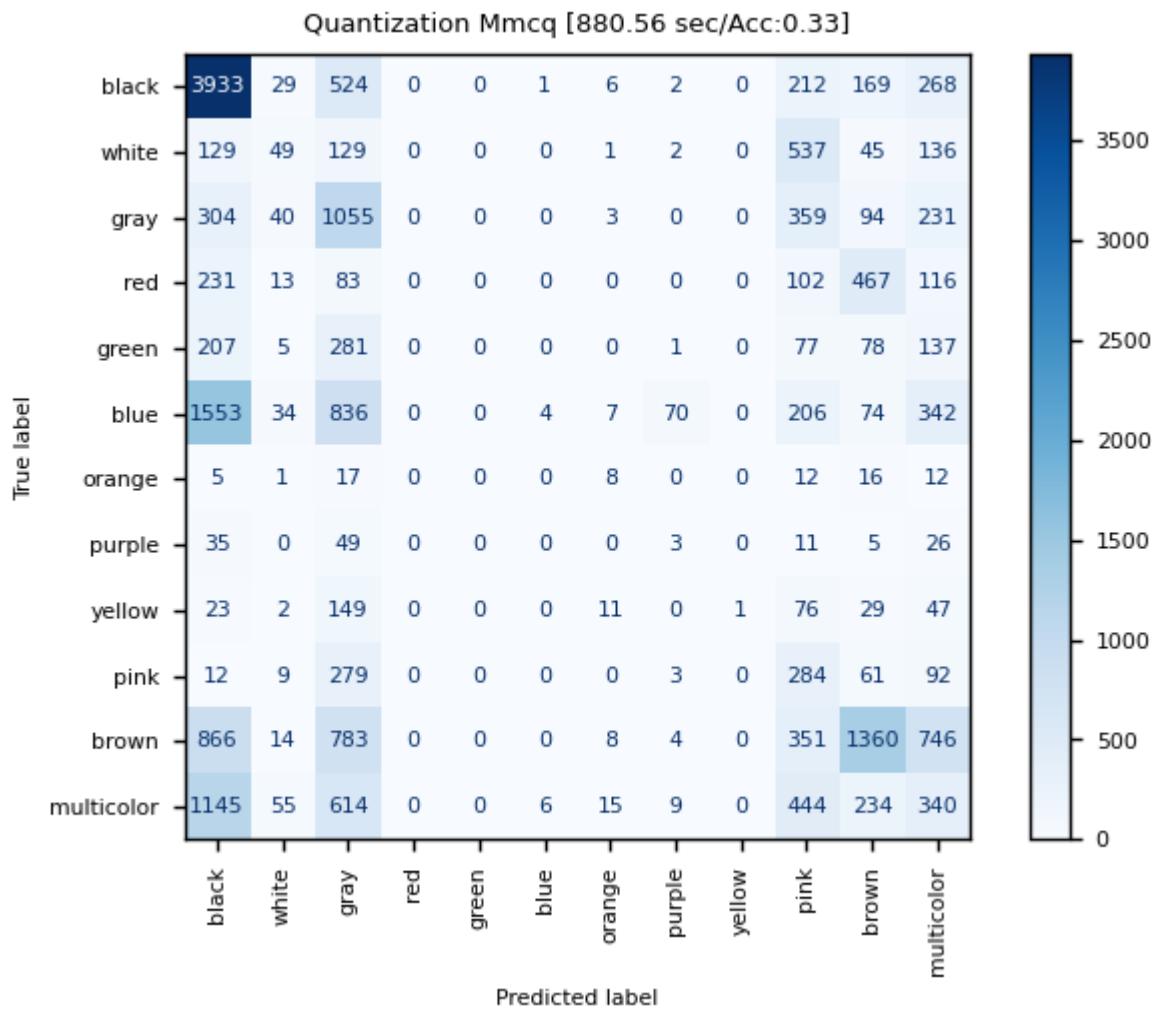
Unrefined Dataset Improvement Run

Multi-Color with RGB / Quantization Mmcq Result

Total Duration: 880.56 sec.

Multi-Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.4658	0.7646	0.5789	5144
white	0.1952	0.0477	0.0766	1028
gray	0.2198	0.5058	0.3065	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.3636	0.0013	0.0026	3126
orange	0.1356	0.1127	0.1231	71
purple	0.0319	0.0233	0.0269	129
yellow	1.0000	0.0030	0.0059	338
pink	0.1063	0.3838	0.1665	740
brown	0.5167	0.3291	0.4021	4132
multicolor	0.1364	0.1188	0.1270	2862
accuracy			0.3280	
macro avg	0.2643	0.1908	0.1513	21454
weighted avg	0.3332	0.3280	0.2734	21454



Multi-Color with RGB / Quantization Mmcq Confusion Matrix

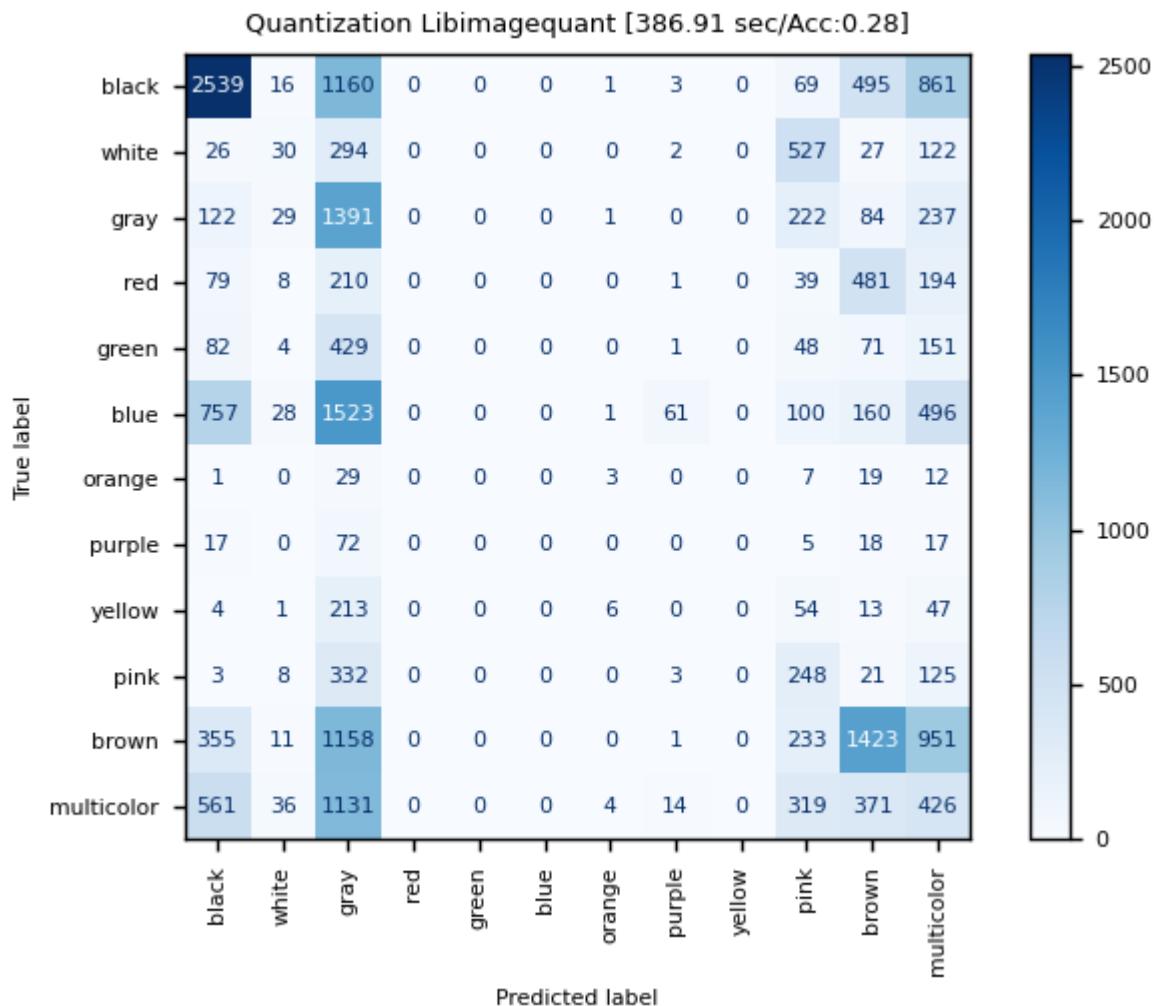
Unrefined Dataset Improvement Run

Multi-Color with RGB / Quantization Libimagequant Result

Total Duration: 386.91 sec.

Multi-Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.5585	0.4936	0.5240	5144
white	0.1754	0.0292	0.0500	1028
gray	0.1751	0.6668	0.2774	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.1875	0.0423	0.0690	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1325	0.3351	0.1900	740
brown	0.4471	0.3444	0.3891	4132
multicolor	0.1171	0.1488	0.1311	2862
accuracy			0.2825	
macro avg	0.1494	0.1717	0.1359	21454
weighted avg	0.2663	0.2825	0.2542	21454



Multi-Color with RGB / Quantization Libimagequant Confusion Matrix

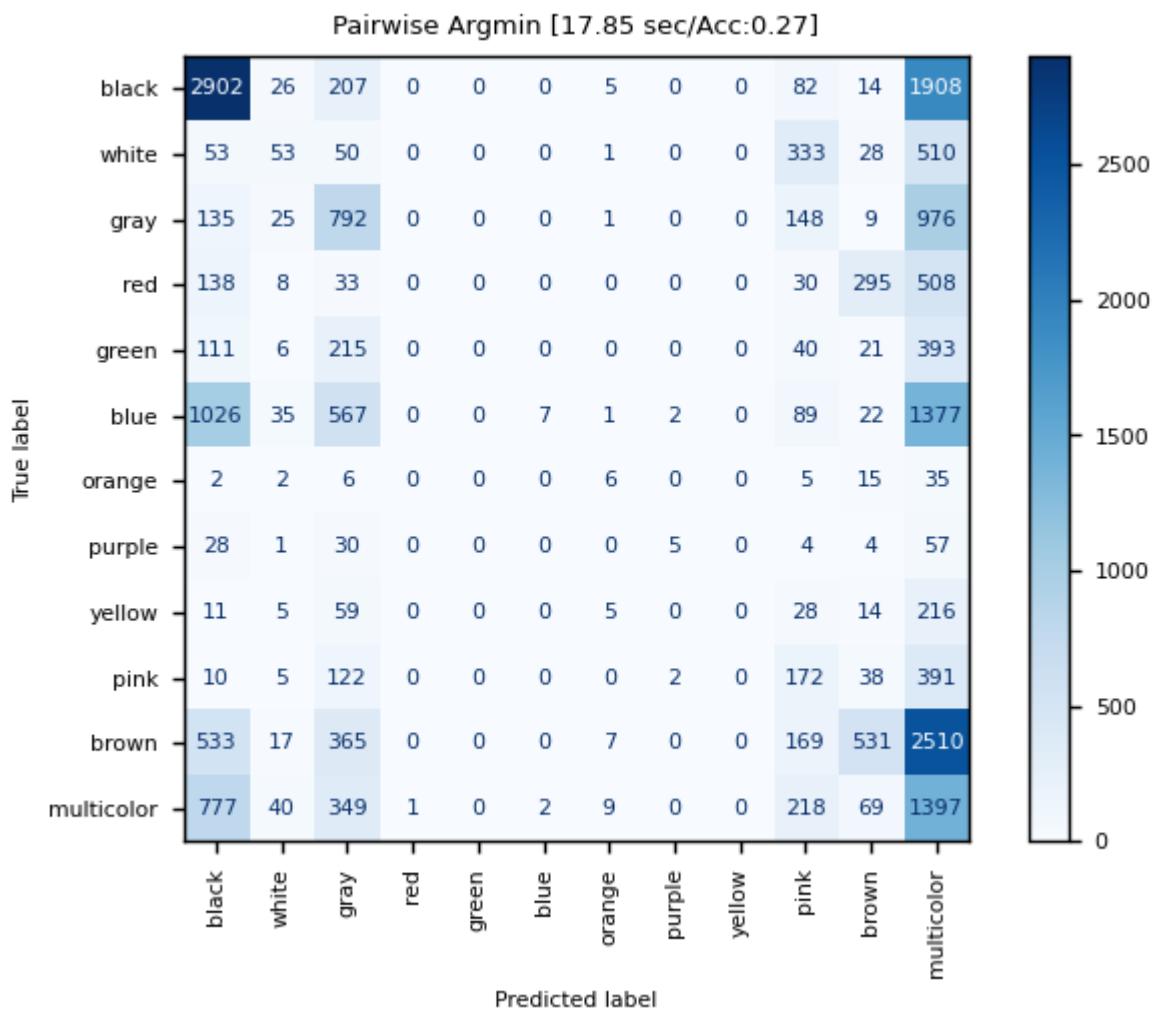
Unrefined Dataset Improvement Run

Multi-Color with RGB / Pairwise Argmin Result

Total Duration: 17.85 sec.

Multi-Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5068	0.5642	0.5339	5144
white	0.2377	0.0516	0.0847	1028
gray	0.2834	0.3797	0.3245	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.7778	0.0022	0.0045	3126
orange	0.1714	0.0845	0.1132	71
purple	0.5556	0.0388	0.0725	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1305	0.2324	0.1672	740
brown	0.5009	0.1285	0.2045	4132
multicolor	0.1359	0.4881	0.2126	2862
accuracy			0.2734	
macro avg	0.2750	0.1642	0.1431	21454
weighted avg	0.3968	0.2734	0.2386	21454



Multi-Color with RGB / Pairwise Argmin Confusion Matrix

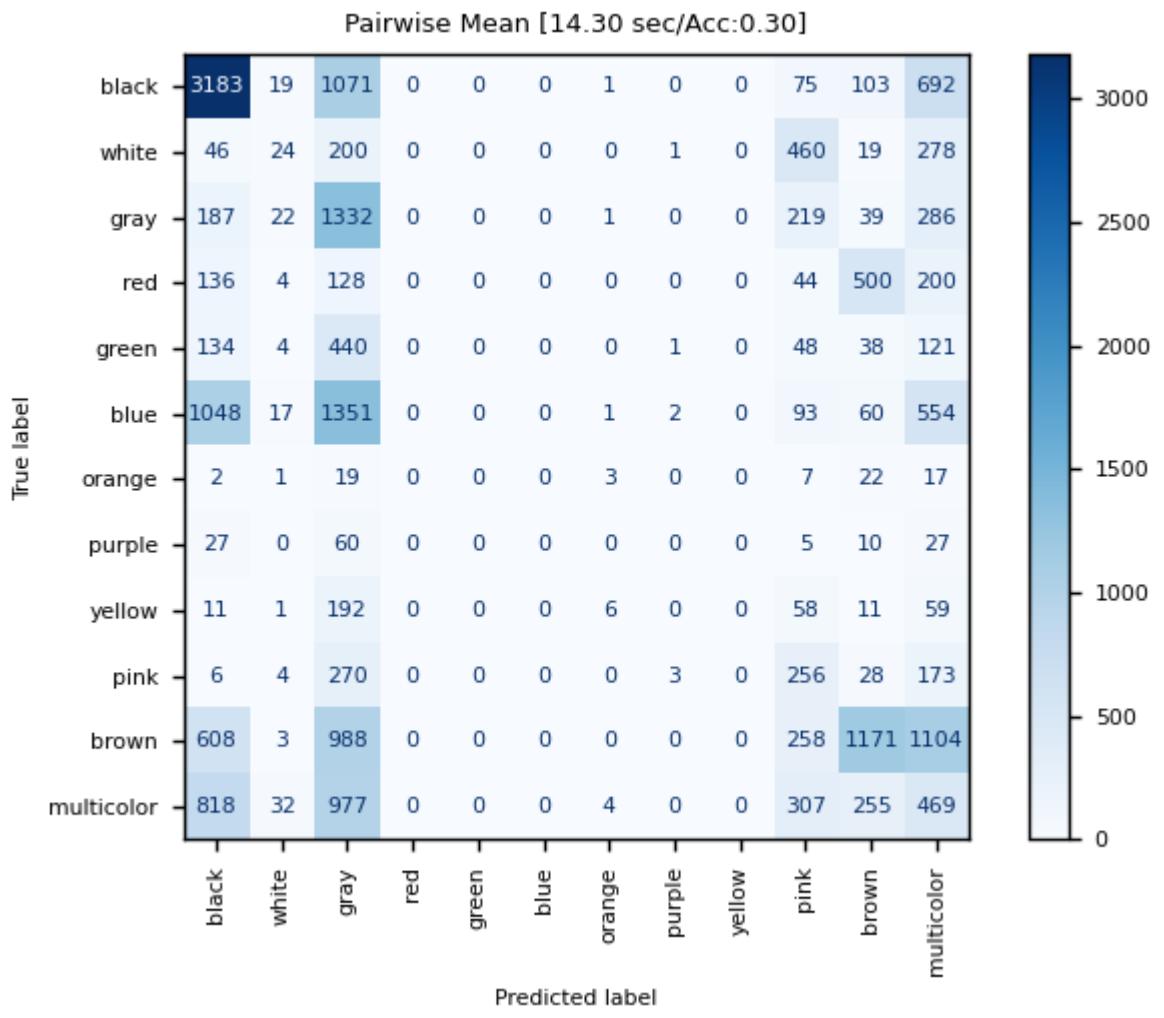
Unrefined Dataset Improvement Run

Multi-Color with RGB / Pairwise Mean Result

Total Duration: 14.30 sec.

Multi-Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.5129	0.6188	0.5609	5144
white	0.1832	0.0233	0.0414	1028
gray	0.1895	0.6385	0.2923	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.1875	0.0423	0.0690	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1399	0.3459	0.1992	740
brown	0.5191	0.2834	0.3666	4132
multicolor	0.1178	0.1639	0.1371	2862
accuracy			0.3001	
macro avg	0.1542	0.1763	0.1389	21454
weighted avg	0.2713	0.3001	0.2609	21454



Multi-Color with RGB / Pairwise Mean Confusion Matrix

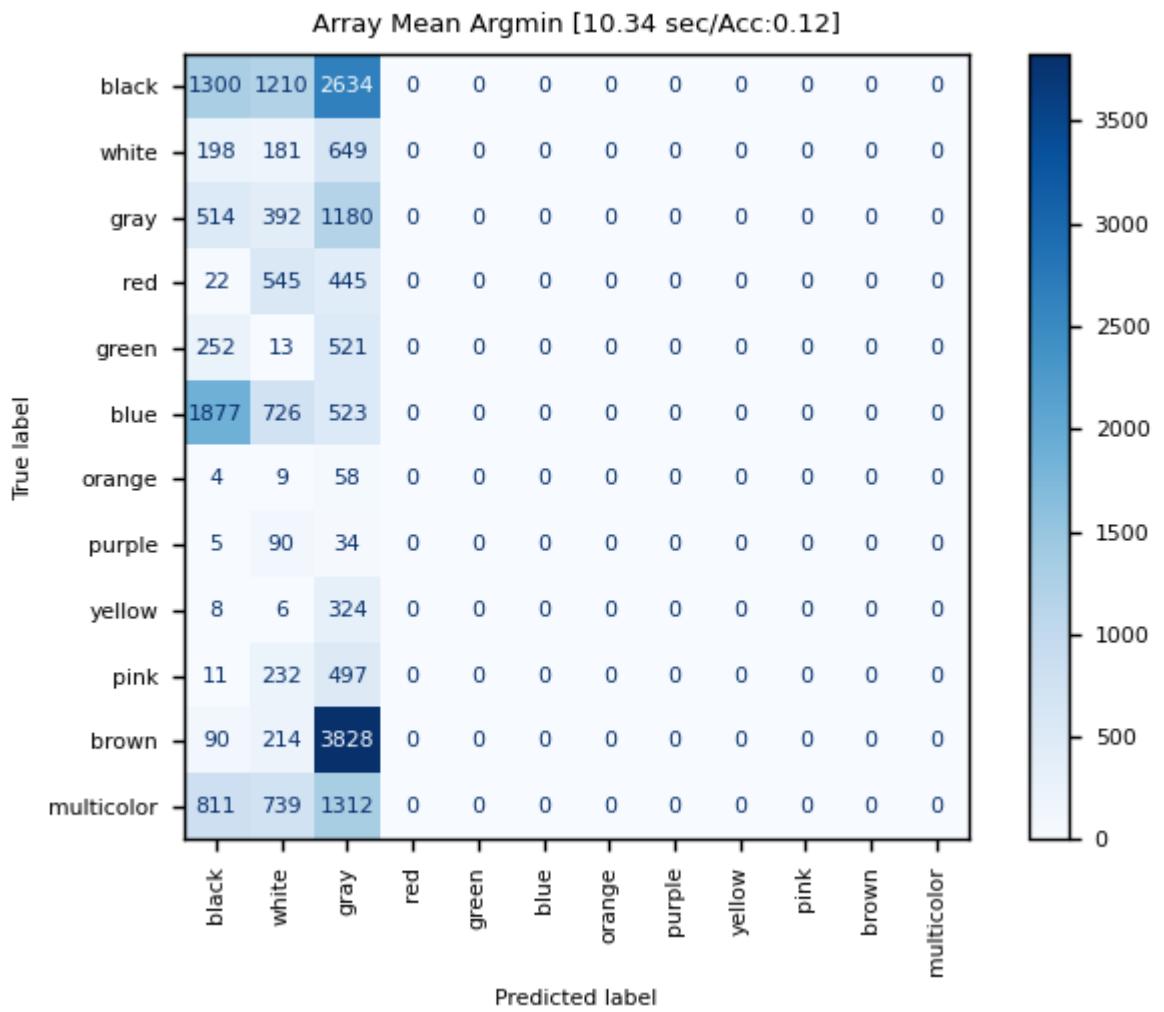
Unrefined Dataset Improvement Run

Multi-Color with RGB / Array Mean Argmin Result

Total Duration: 10.34 sec.

Multi-Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2553	0.2527	0.2540	5144
white	0.0415	0.1761	0.0672	1028
gray	0.0983	0.5657	0.1675	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.0000	0.0000	0.0000	740
brown	0.0000	0.0000	0.0000	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.1240	
macro avg	0.0329	0.0829	0.0407	21454
weighted avg	0.0728	0.1240	0.0804	21454



Multi-Color with RGB / Array Mean Argmin Confusion Matrix

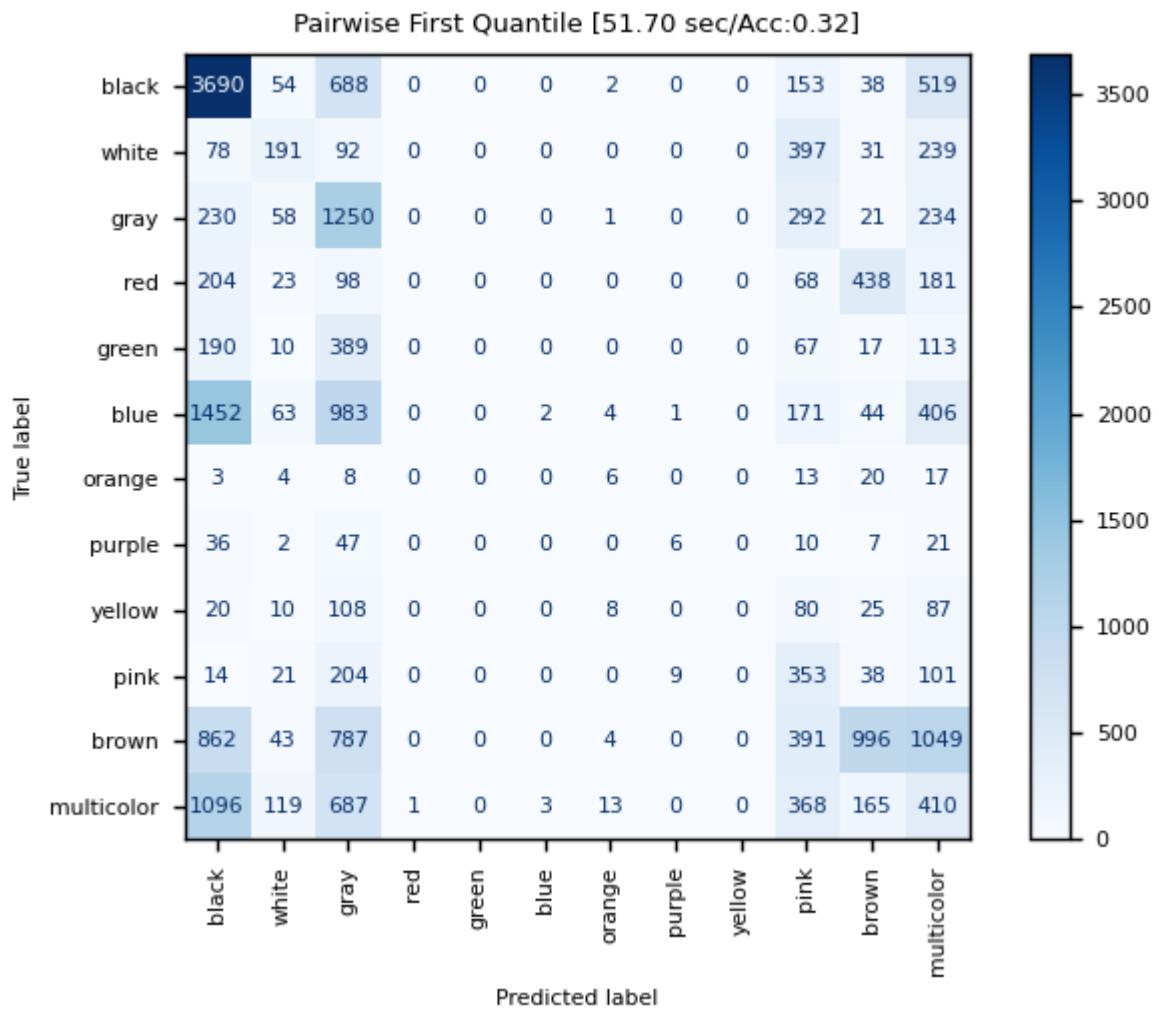
Unrefined Dataset Improvement Run

Multi-Color with RGB / Pairwise First Quantile Result

Total Duration: 51.70 sec.

Multi-Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4686	0.7173	0.5669	5144
white	0.3194	0.1858	0.2349	1028
gray	0.2340	0.5992	0.3366	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.4000	0.0006	0.0013	3126
orange	0.1579	0.0845	0.1101	71
purple	0.3750	0.0465	0.0828	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1494	0.4770	0.2275	740
brown	0.5413	0.2410	0.3336	4132
multicolor	0.1214	0.1433	0.1314	2862
accuracy			0.3218	
macro avg	0.2306	0.2079	0.1688	21454
weighted avg	0.3371	0.3218	0.2706	21454



Multi-Color with RGB / Pairwise First Quantile Confusion Matrix

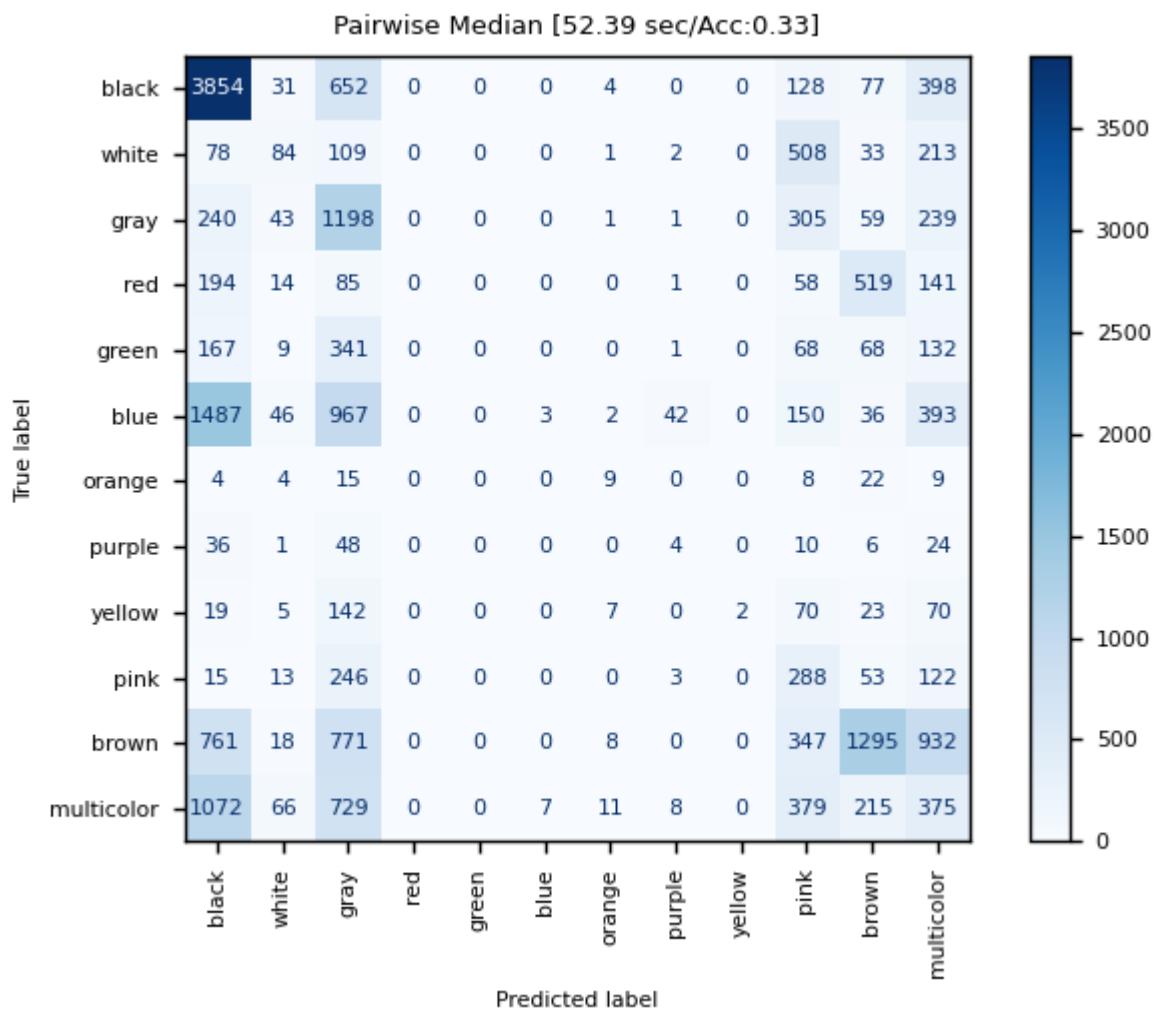
Unrefined Dataset Improvement Run

Multi-Color with RGB / Pairwise Median Result

Total Duration: 52.39 sec.

Multi-Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4862	0.7492	0.5897	5144
white	0.2515	0.0817	0.1233	1028
gray	0.2259	0.5743	0.3243	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.3000	0.0010	0.0019	3126
orange	0.2093	0.1268	0.1579	71
purple	0.0645	0.0310	0.0419	129
yellow	1.0000	0.0059	0.0118	338
pink	0.1242	0.3892	0.1883	740
brown	0.5382	0.3134	0.3961	4132
multicolor	0.1230	0.1310	0.1269	2862
accuracy			0.3315	
macro avg	0.2769	0.2003	0.1635	21454
weighted avg	0.3355	0.3315	0.2798	21454



Multi-Color with RGB / Pairwise Median Confusion Matrix

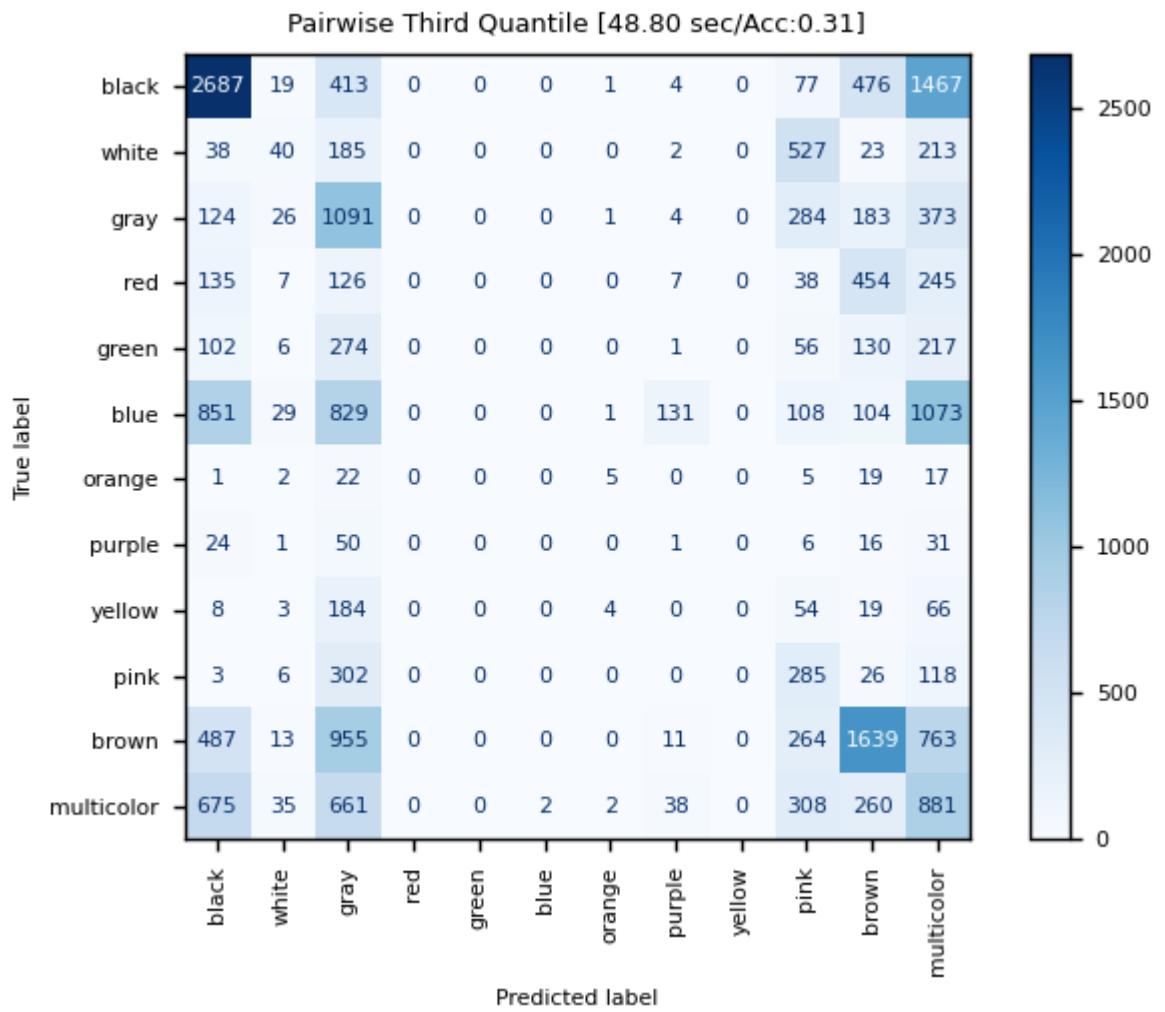
Unrefined Dataset Improvement Run

Multi-Color with RGB / Pairwise Third Quantile Result

Total Duration: 48.80 sec.

Multi-Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5233	0.5224	0.5228	5144
white	0.2139	0.0389	0.0658	1028
gray	0.2143	0.5230	0.3040	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.3571	0.0704	0.1176	71
purple	0.0050	0.0078	0.0061	129
yellow	0.0000	0.0000	0.0000	338
pink	0.1417	0.3851	0.2071	740
brown	0.4894	0.3967	0.4382	4132
multicolor	0.1612	0.3078	0.2116	2862
accuracy			0.3090	
macro avg	0.1755	0.1877	0.1561	21454
weighted avg	0.2784	0.3090	0.2783	21454



Multi-Color with RGB / Pairwise Third Quantile Confusion Matrix

3.2. HSL Color Space

3.2.1. CNN AlexNet with Single Color

Unrefined Dataset Improvement Run

CNN AlexNet Single Color with HSL

Total Duration: 4770.03 sec.

Learning Rate: 0.000070

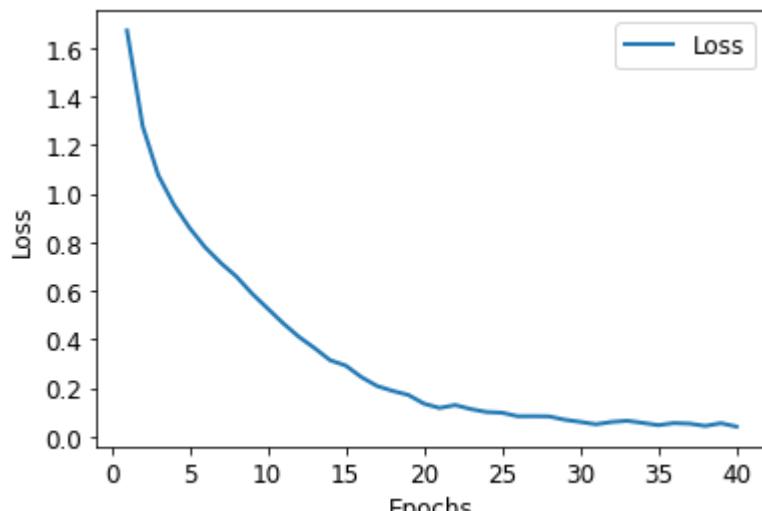
Dropout: 0.10

Batch Size: 32

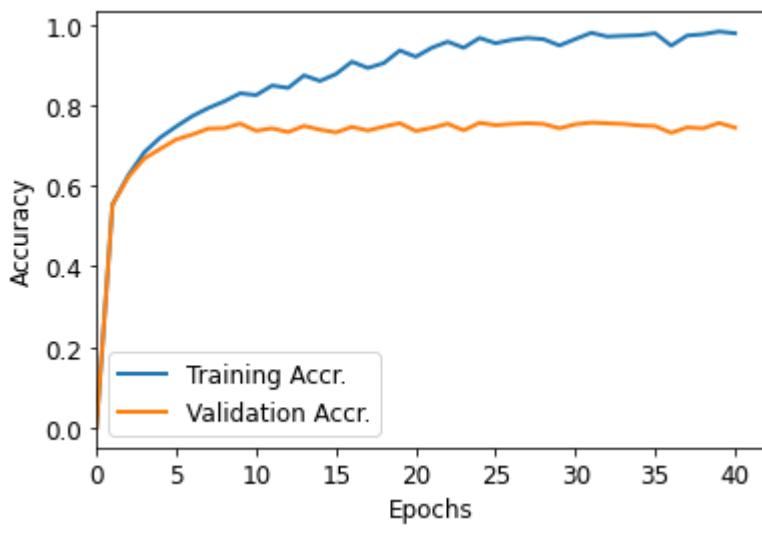
Epoch: 40

Best Performing Epoch: 31

Summary	
Method	Accuracy
Training Dataset	0.9799
Validation Dataset	0.7572
Test Dataset	0.7567



Loss Plot

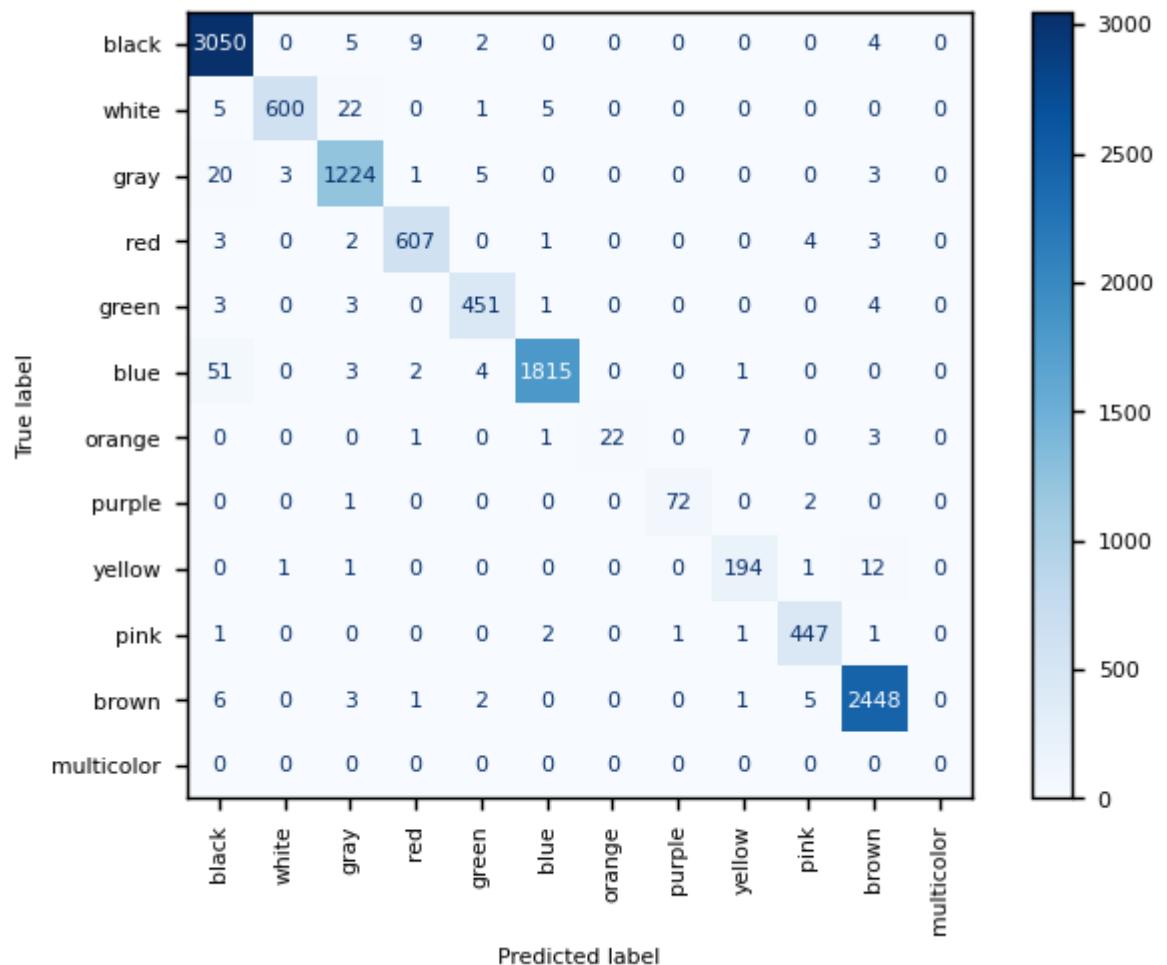


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9716	0.9935	0.9824	3070
white	0.9934	0.9479	0.9701	633
gray	0.9684	0.9745	0.9714	1256
red	0.9775	0.9790	0.9782	620
green	0.9699	0.9762	0.9730	462
blue	0.9945	0.9675	0.9808	1876
orange	1.0000	0.6471	0.7857	34
purple	0.9863	0.9600	0.9730	75
yellow	0.9510	0.9282	0.9395	209
pink	0.9739	0.9868	0.9803	453
brown	0.9879	0.9927	0.9903	2466
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.9799	0.9799	0.9799	11154
macro avg	0.8979	0.8628	0.8771	11154
weighted avg	0.9801	0.9799	0.9798	11154

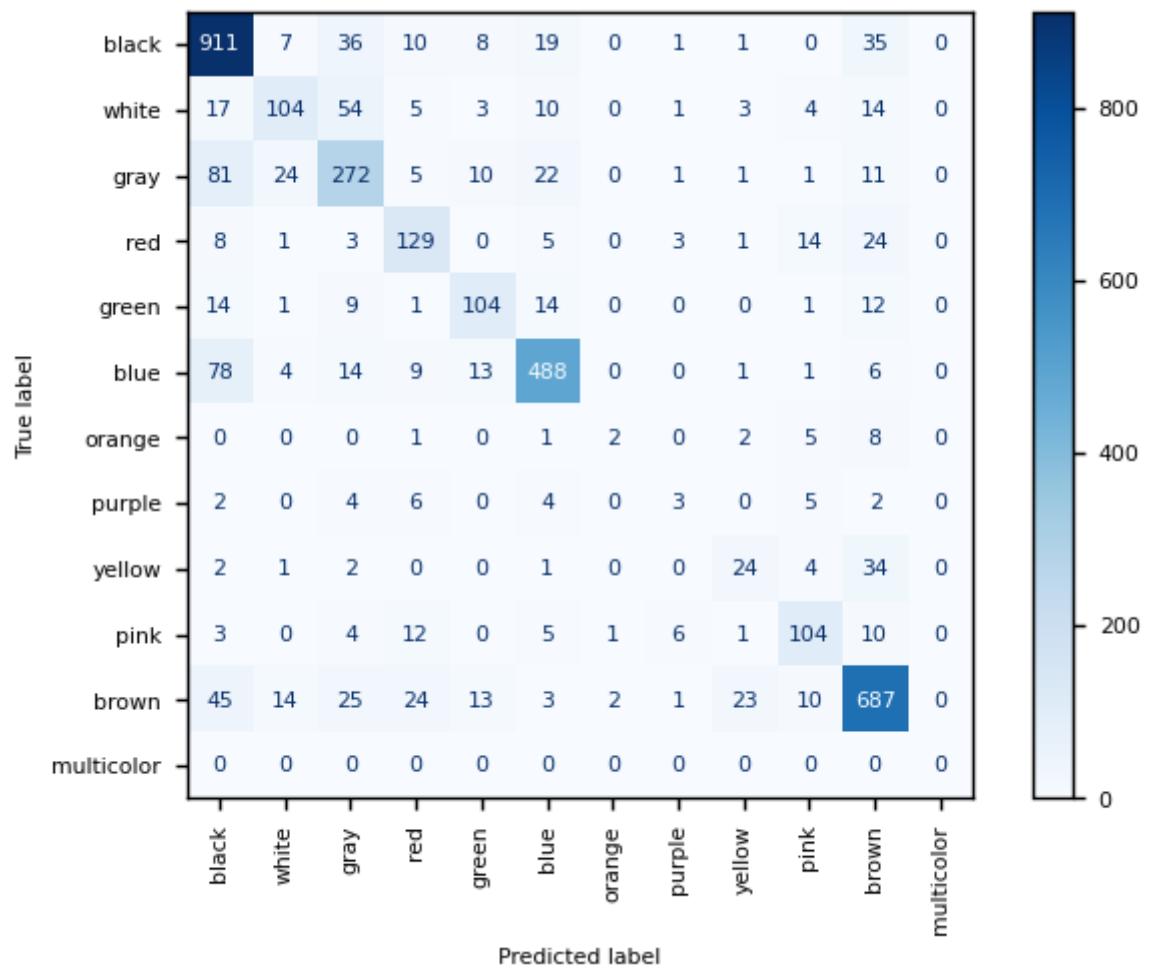


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7847	0.8862	0.8323	1028
white	0.6667	0.4837	0.5606	215
gray	0.6430	0.6355	0.6392	428
red	0.6386	0.6862	0.6615	188
green	0.6887	0.6667	0.6775	156
blue	0.8531	0.7948	0.8229	614
orange	0.4000	0.1053	0.1667	19
purple	0.1875	0.1154	0.1429	26
yellow	0.4211	0.3529	0.3840	68
pink	0.6980	0.7123	0.7051	146
brown	0.8149	0.8111	0.8130	847
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7572	0.7572	0.7572	3735
macro avg	0.5664	0.5208	0.5338	3735
weighted avg	0.7523	0.7572	0.7523	3735

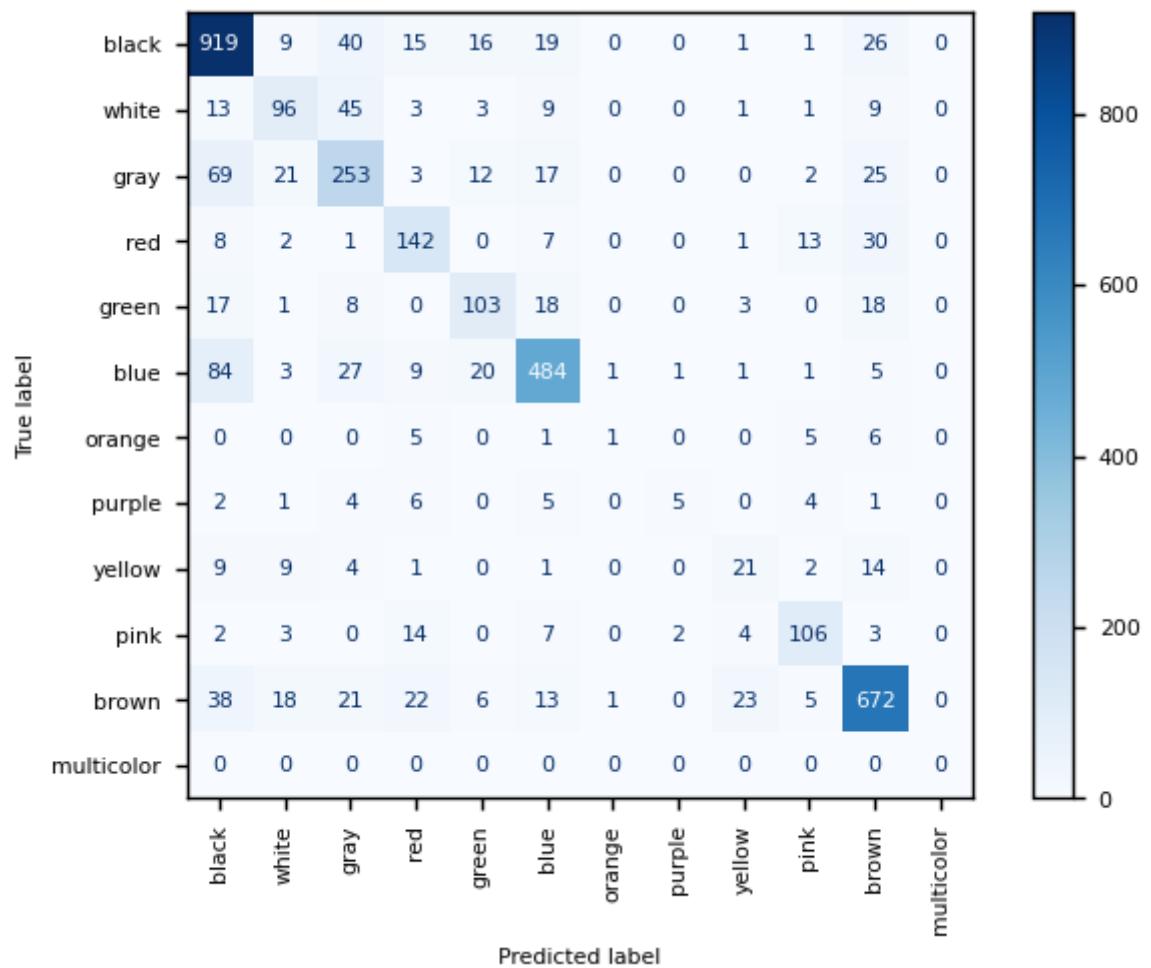


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7916	0.8786	0.8328	1046
white	0.5890	0.5333	0.5598	180
gray	0.6278	0.6294	0.6286	402
red	0.6455	0.6961	0.6698	204
green	0.6438	0.6131	0.6280	168
blue	0.8330	0.7610	0.7954	636
orange	0.3333	0.0556	0.0952	18
purple	0.6250	0.1786	0.2778	28
yellow	0.3818	0.3443	0.3621	61
pink	0.7571	0.7518	0.7544	141
brown	0.8307	0.8205	0.8256	819
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7567	0.7567	0.7567	3703
macro avg	0.5882	0.5218	0.5358	3703
weighted avg	0.7534	0.7567	0.7525	3703



red (99.8%) >> red black (100.0%) >> black brown (100.0%) >> brown blue (81.6%) >> blue gray (100.0%) >> gray green (94.5%) >> blue



black (100.0%) >> black gray (100.0%) >> gray black (100.0%) >> brown black (100.0%) >> black black (57.3%) >> blue black (100.0%) >> black



white (98.5%) >> white gray (59.8%) >> white brown (100.0%) >> brown black (100.0%) >> black white (98.8%) >> black black (99.9%) >> black



brown (100.0%) >> brown black (71.0%) >> blue red (100.0%) >> red black (100.0%) >> black red (99.4%) >> red brown (100.0%) >> yellow



blue (99.9%) >> blue brown (100.0%) >> brown black (99.3%) >> black green (53.1%) >> brown gray (100.0%) >> gray brown (100.0%) >> brown



blue (99.9%) >> gray brown (52.0%) >> black brown (100.0%) >> brown brown (100.0%) >> brown green (69.3%) >> brown white (98.9%) >> white



Sample Images from Test Dataset

white (98.8%) green (94.5%) brown (98.2%) brown (88.0%) black (57.3%) gray (98.2%)
 >> black >> blue >> gray >> blue >> blue >> blue



brown (52.0%) yellow (82.9%) blue (99.9%) red (58.7%) green (71.3%) green (68.5%)
 >> black >> brown >> gray >> orange >> black >> blue



black (71.0%) black (100.0%) black (100.0%) black (89.5%) gray (48.3%) brown (100.0%)
 >> blue >> gray >> brown >> blue >> blue >> yellow



gray (57.3%) black (99.2%) brown (99.9%) green (69.3%) black (100.0%) brown (100.0%)
 >> black >> yellow >> green >> brown >> blue >> yellow



white (93.2%) black (67.4%) black (100.0%) yellow (47.1%) gray (98.6%) white (56.3%)
 >> purple >> gray >> red >> brown >> blue >> yellow



gray (59.8%) black (99.5%) green (81.8%) red (99.6%) gray (91.8%) black (94.3%)
 >> white >> gray >> brown >> brown >> brown >> blue



Sample Images that cannot be accurately predicted in the Test Dataset

3.2.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Pairwise Argmin	44.44 sec.	0.4277
Pairwise Mean	13.03 sec.	0.3766
Array Mean Argmin	63.09 sec.	0.2766
Pairwise First Quantile	102.39 sec.	0.4207
Pairwise Median	62.68 sec.	0.4159
Pairwise Third Quantile	56.50 sec.	0.3564

* Quantization based methods are available only in RGB color space

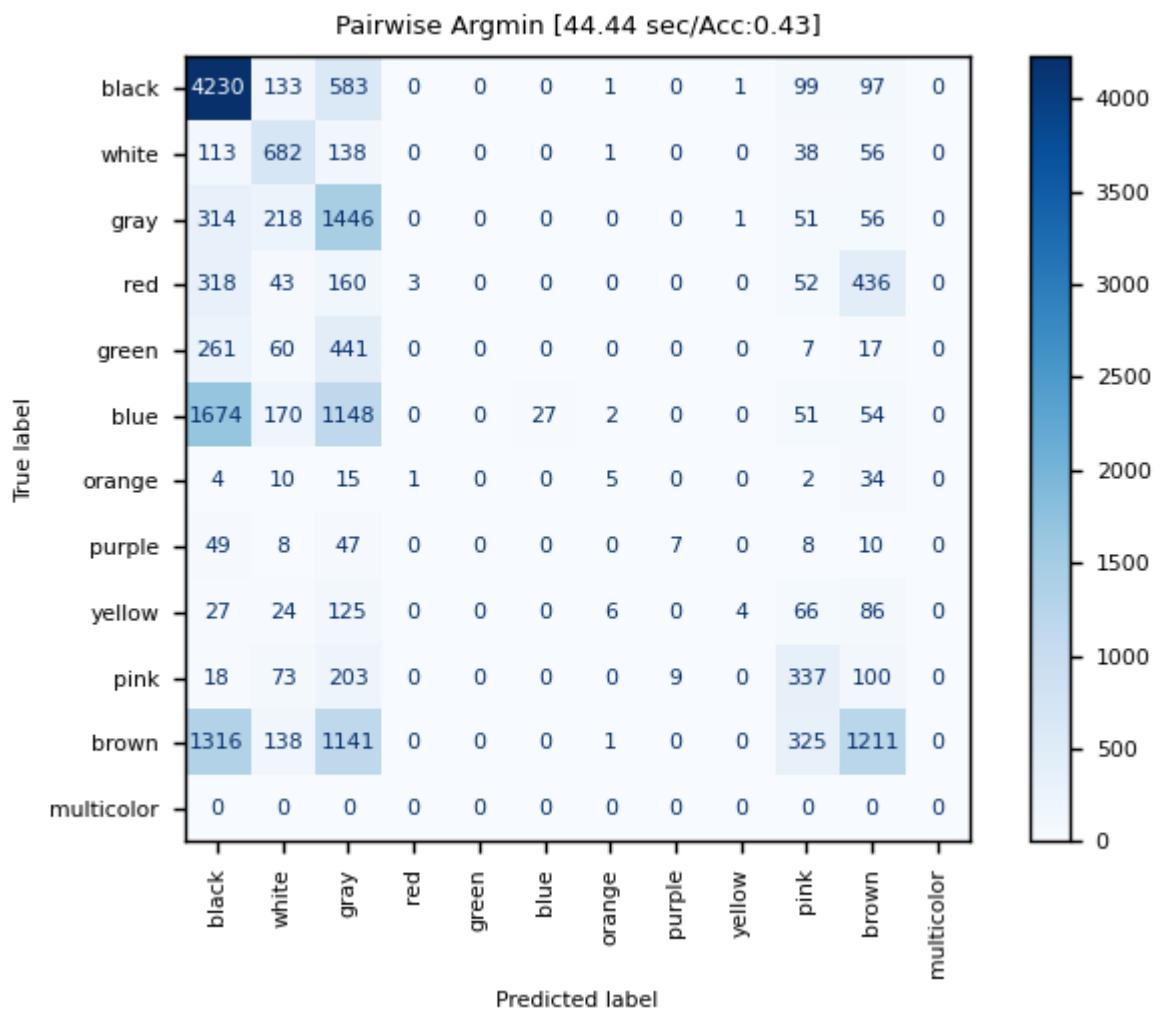
Unrefined Dataset Improvement Run

Single Color with HSL / Pairwise Argmin Result

Total Duration: 44.44 sec.

Single Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5082	0.8223	0.6282	5144
white	0.4375	0.6634	0.5273	1028
gray	0.2655	0.6932	0.3839	2086
red	0.7500	0.0030	0.0059	1012
green	0.0000	0.0000	0.0000	786
blue	1.0000	0.0086	0.0171	3126
orange	0.3125	0.0704	0.1149	71
purple	0.4375	0.0543	0.0966	129
yellow	0.6667	0.0118	0.0233	338
pink	0.3253	0.4554	0.3795	740
brown	0.5614	0.2931	0.3851	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4277	0.4277	0.4277	18592
macro avg	0.4387	0.2563	0.2135	18592
weighted avg	0.5576	0.4277	0.3515	18592



Single Color with HSL / Pairwise Argmin Confusion Matrix

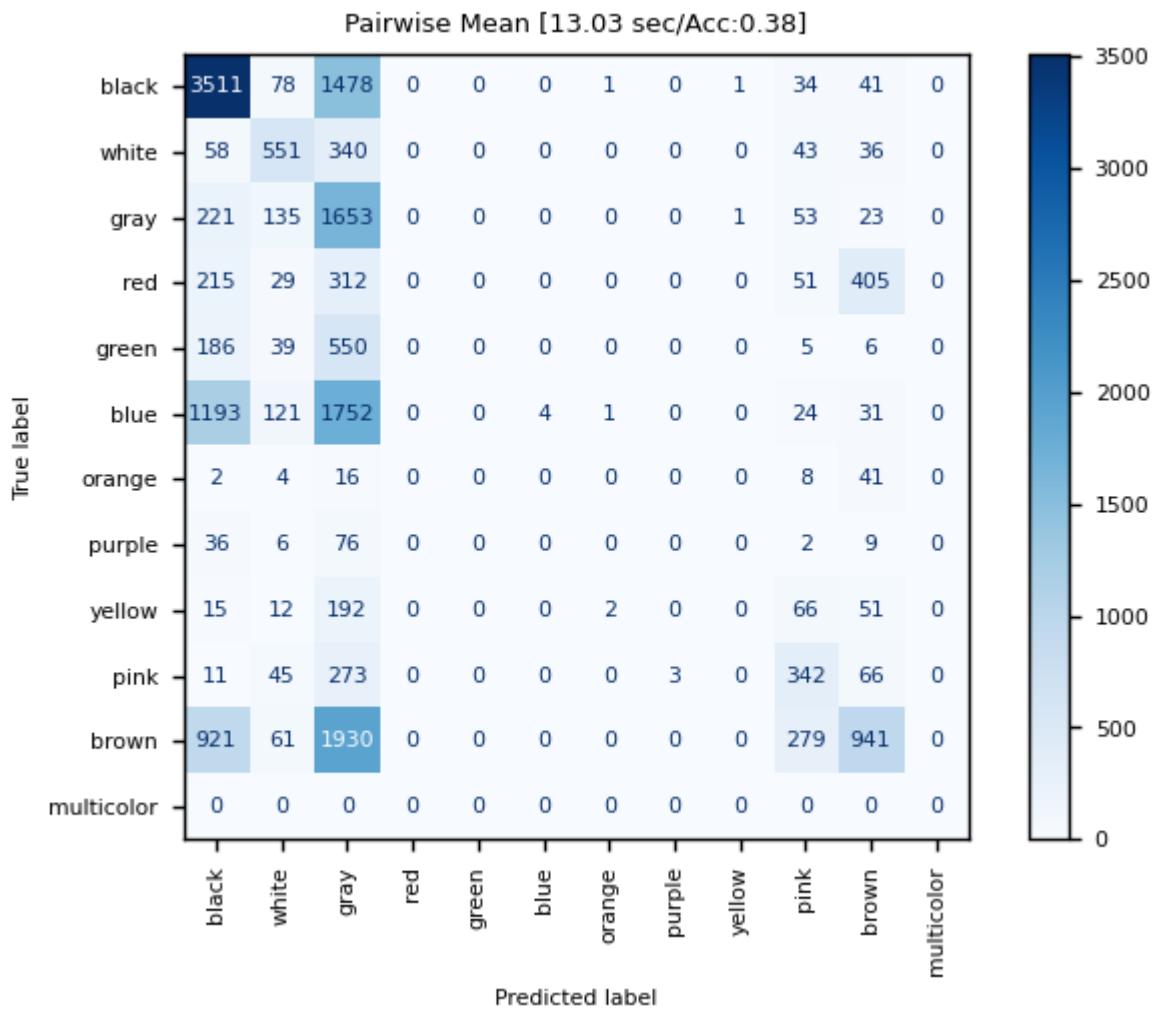
Unrefined Dataset Improvement Run

Single Color with HSL / Pairwise Mean Result

Total Duration: 13.03 sec.

Single Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.5513	0.6825	0.6099	5144
white	0.5097	0.5360	0.5225	1028
gray	0.1928	0.7924	0.3102	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	1.0000	0.0013	0.0026	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3771	0.4622	0.4153	740
brown	0.5703	0.2277	0.3255	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3766	0.3766	0.3766	18592
macro avg	0.2668	0.2252	0.1822	18592
weighted avg	0.5122	0.3766	0.3217	18592



Single Color with HSL / Pairwise Mean Confusion Matrix

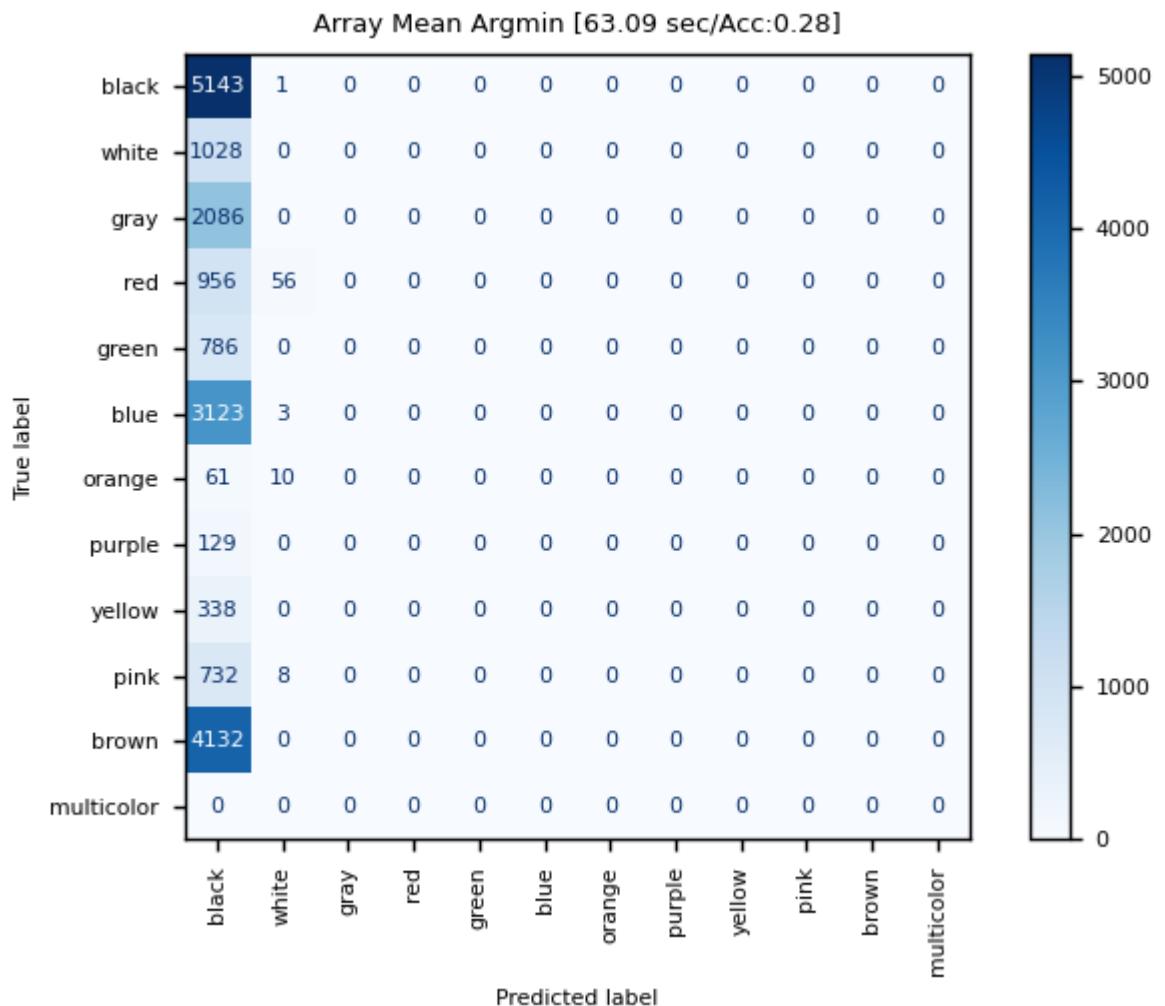
Unrefined Dataset Improvement Run

Single Color with HSL / Array Mean Argmin Result

Total Duration: 63.09 sec.

Single Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2778	0.9998	0.4348	5144
white	0.0000	0.0000	0.0000	1028
gray	0.0000	0.0000	0.0000	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.0000	0.0000	0.0000	740
brown	0.0000	0.0000	0.0000	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2766	0.2766	0.2766	18592
macro avg	0.0231	0.0833	0.0362	18592
weighted avg	0.0769	0.2766	0.1203	18592



Single Color with HSL / Array Mean Argmin Confusion Matrix

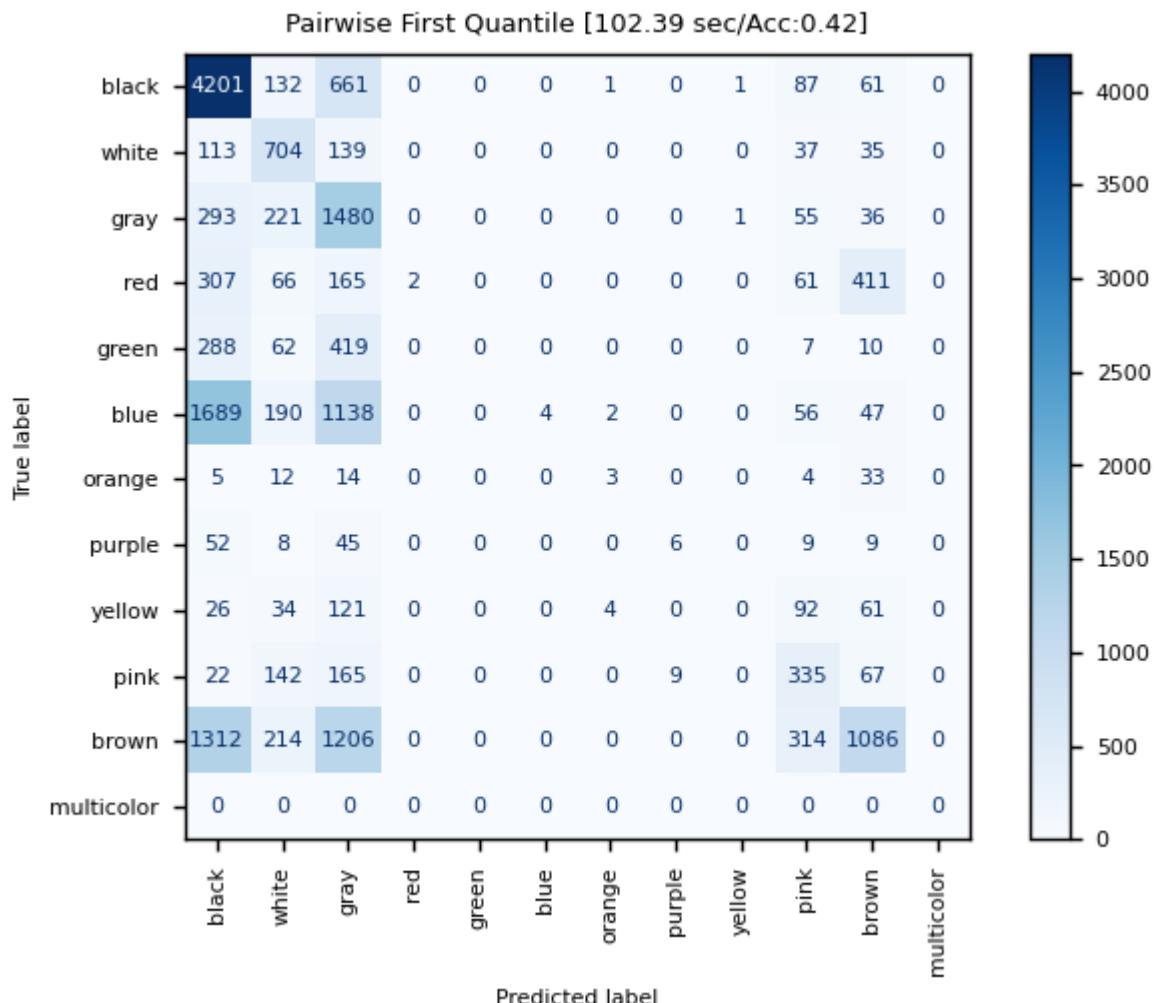
Unrefined Dataset Improvement Run

Single Color with HSL / Pairwise First Quantile Result

Total Duration: 102.39 sec.

Single Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5057	0.8167	0.6246	5144
white	0.3944	0.6848	0.5005	1028
gray	0.2665	0.7095	0.3875	2086
red	1.0000	0.0020	0.0039	1012
green	0.0000	0.0000	0.0000	786
blue	1.0000	0.0013	0.0026	3126
orange	0.3000	0.0423	0.0741	71
purple	0.4000	0.0465	0.0833	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3169	0.4527	0.3728	740
brown	0.5851	0.2628	0.3627	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4207	0.4207	0.4207	18592
macro avg	0.3974	0.2515	0.2010	18592
weighted avg	0.5608	0.4207	0.3409	18592



Single Color with HSL / Pairwise First Quantile Confusion Matrix

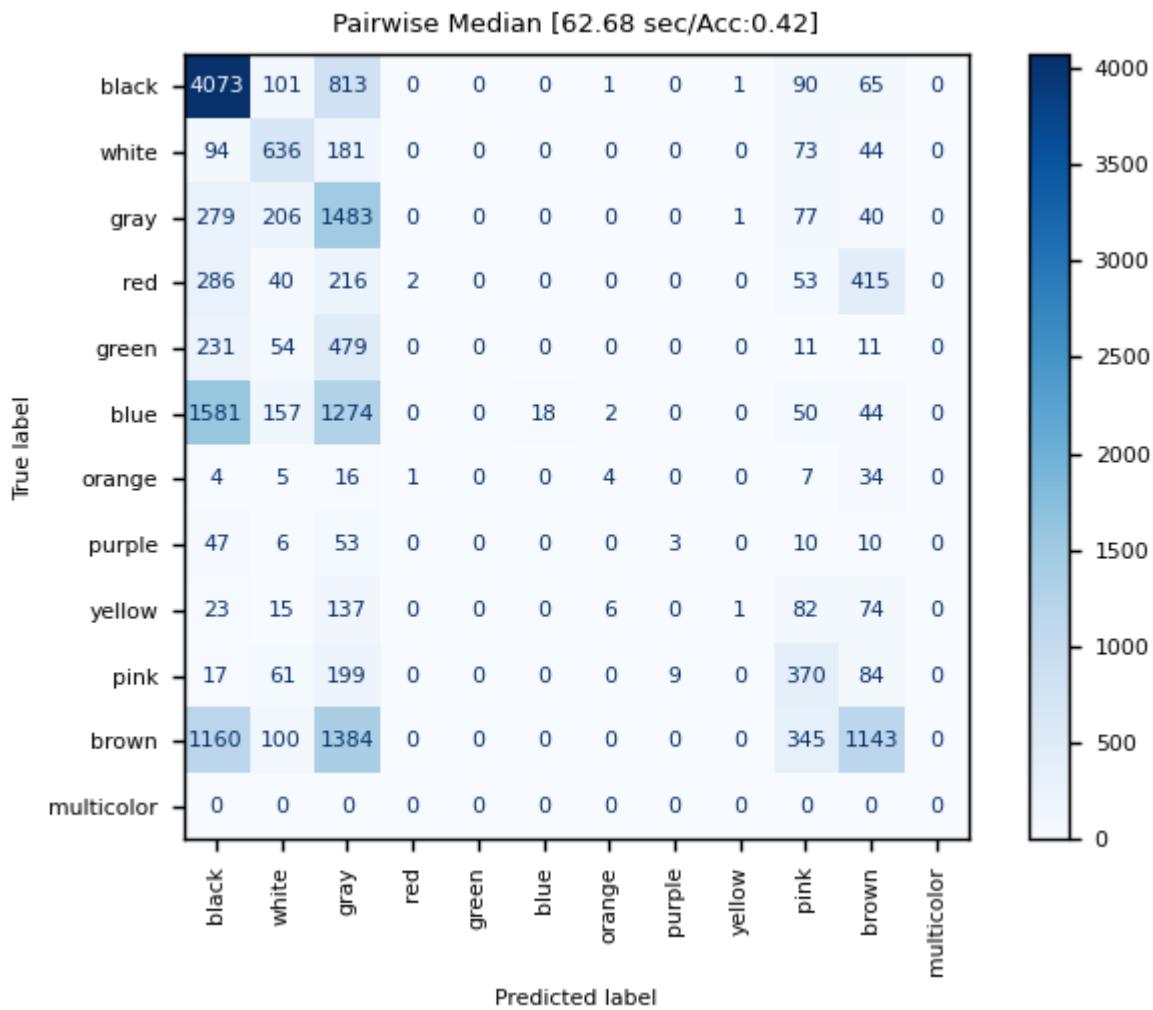
Unrefined Dataset Improvement Run

Single Color with HSL / Pairwise Median Result

Total Duration: 62.68 sec.

Single Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.5225	0.7918	0.6296	5144
white	0.4605	0.6187	0.5280	1028
gray	0.2379	0.7109	0.3564	2086
red	0.6667	0.0020	0.0039	1012
green	0.0000	0.0000	0.0000	786
blue	1.0000	0.0058	0.0115	3126
orange	0.3077	0.0563	0.0952	71
purple	0.2500	0.0233	0.0426	129
yellow	0.3333	0.0030	0.0059	338
pink	0.3168	0.5000	0.3878	740
brown	0.5820	0.2766	0.3750	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4159	0.4159	0.4159	18592
macro avg	0.3898	0.2490	0.2030	18592
weighted avg	0.5521	0.4159	0.3451	18592



Single Color with HSL / Pairwise Median Confusion Matrix

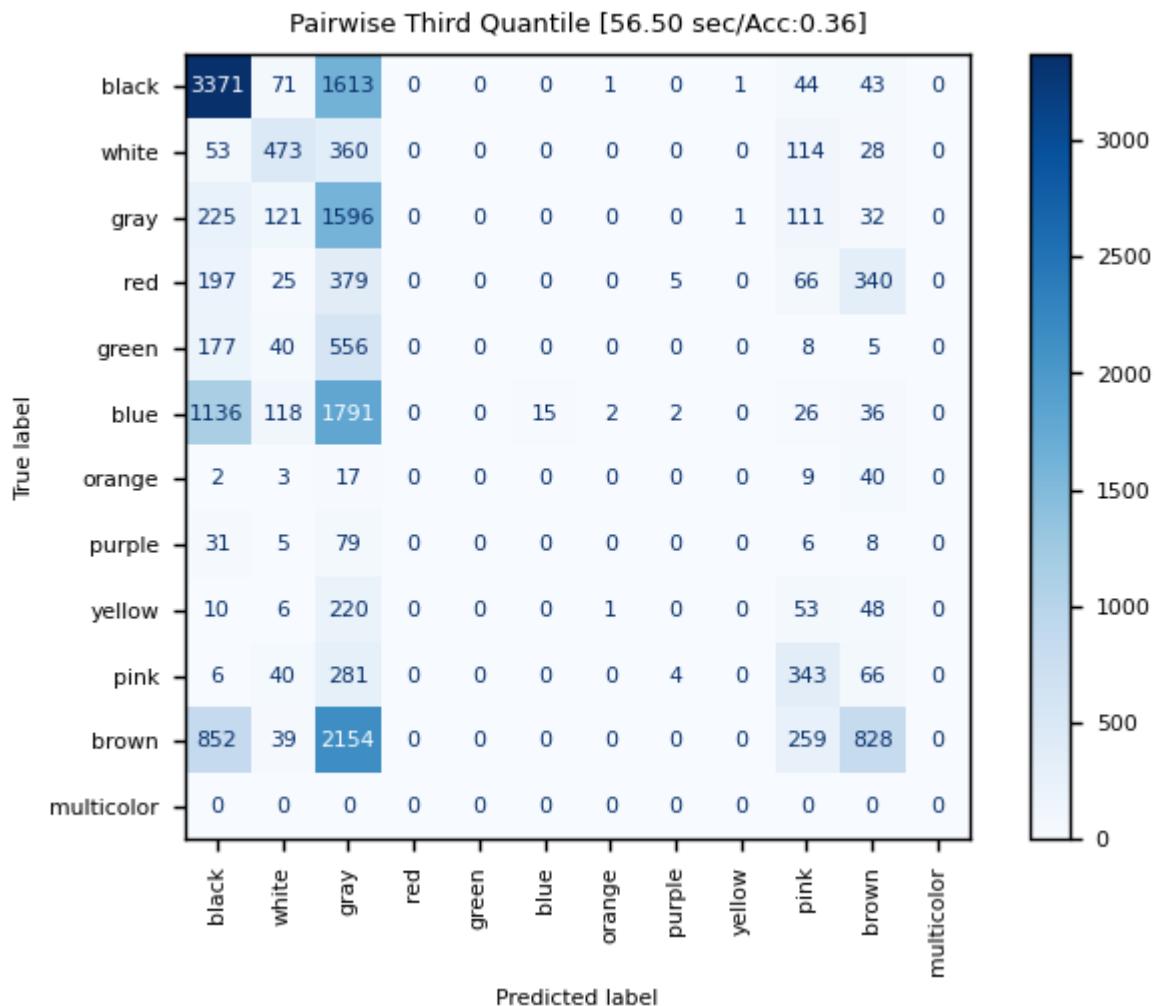
Unrefined Dataset Improvement Run

Single Color with HSL / Pairwise Third Quantile Result

Total Duration: 56.50 sec.

Single Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5563	0.6553	0.6017	5144
white	0.5027	0.4601	0.4804	1028
gray	0.1764	0.7651	0.2867	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	1.0000	0.0048	0.0096	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3301	0.4635	0.3856	740
brown	0.5617	0.2004	0.2954	4132
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3564	0.3564	0.3564	18592
macro avg	0.2606	0.2124	0.1716	18592
weighted avg	0.5076	0.3564	0.3078	18592



Single Color with HSL / Pairwise Third Quantile Confusion Matrix

3.2.3. CNN AlexNet with Multi-color

Unrefined Dataset Improvement Run

CNN AlexNet Multi-Color with HSL

Total Duration: 5415.81 sec.

Learning Rate: 0.000070

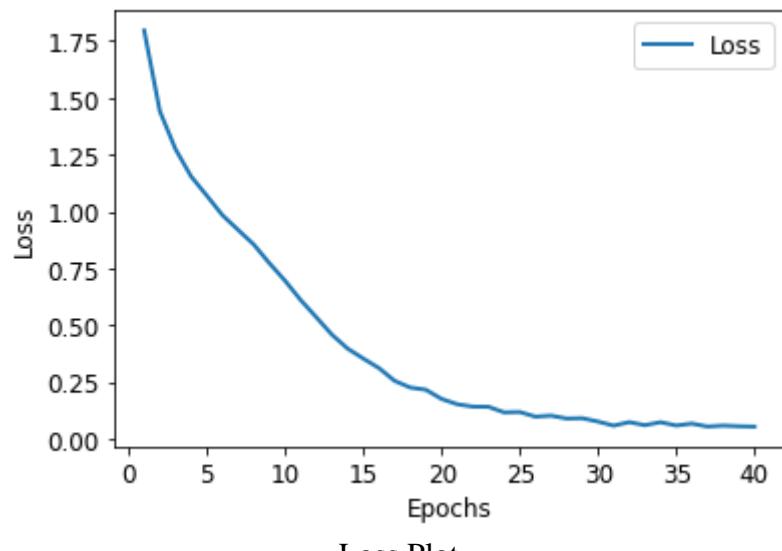
Dropout: 0.10

Batch Size: 32

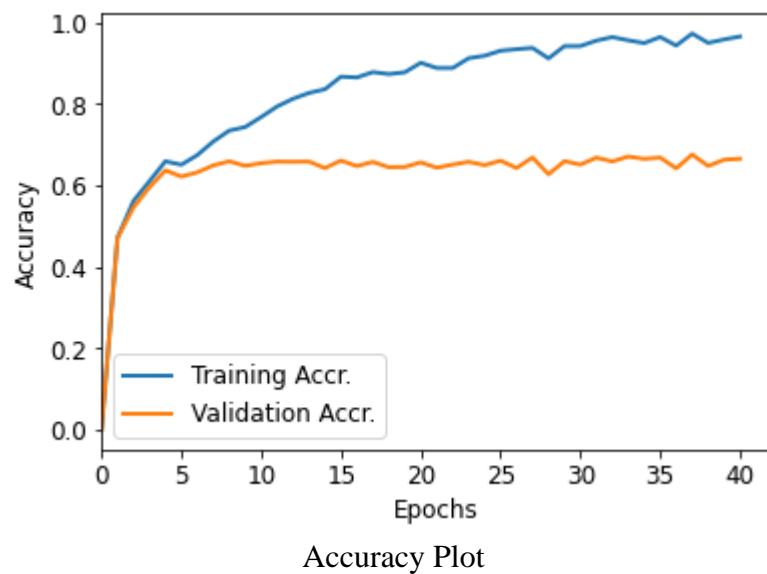
Epoch: 40

Best Performing Epoch: 37

Summary	
Method	Accuracy
Training Dataset	0.9732
Validation Dataset	0.6760
Test Dataset	0.6606



Loss Plot

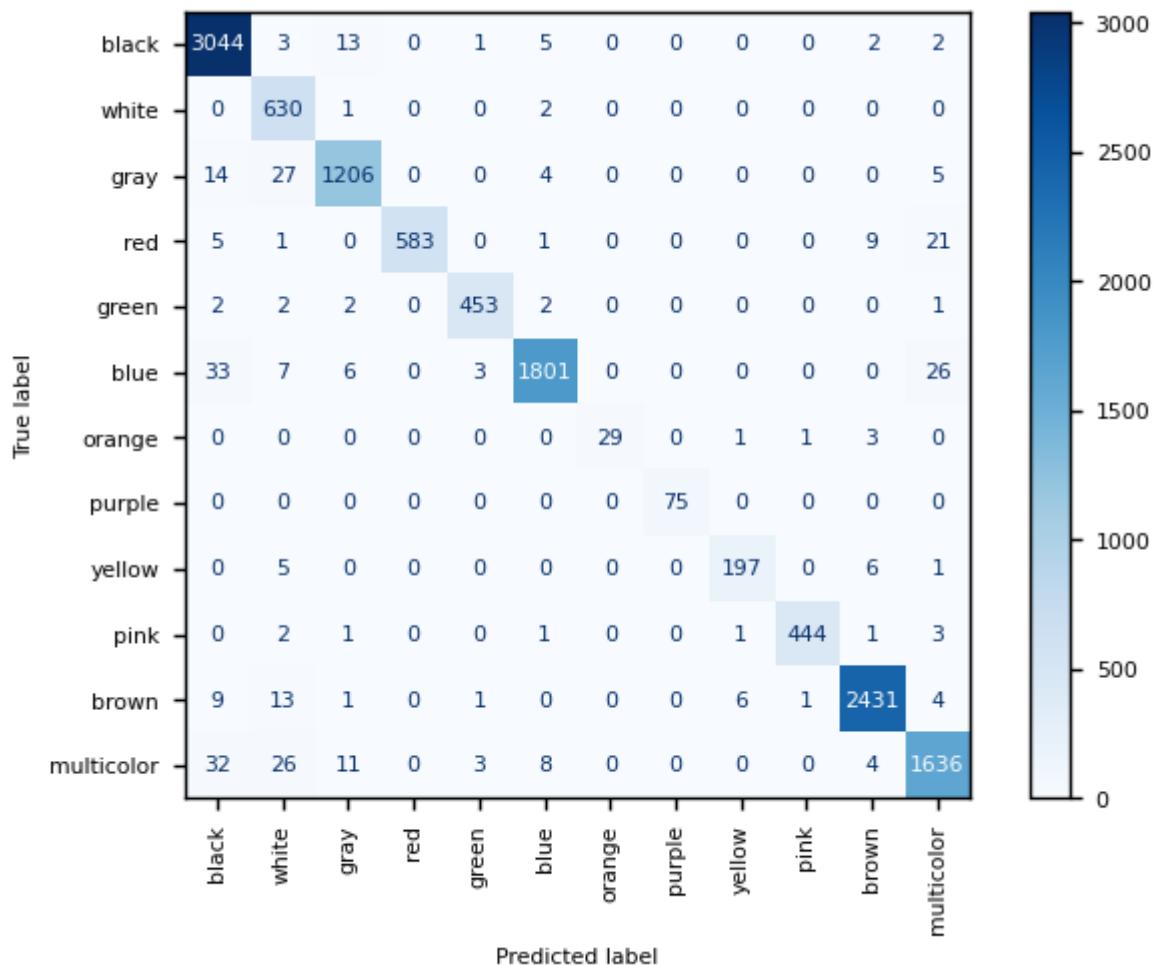


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9697	0.9915	0.9805	3070
white	0.8799	0.9953	0.9340	633
gray	0.9718	0.9602	0.9660	1256
red	1.0000	0.9403	0.9692	620
green	0.9826	0.9805	0.9816	462
blue	0.9874	0.9600	0.9735	1876
orange	1.0000	0.8529	0.9206	34
purple	1.0000	1.0000	1.0000	75
yellow	0.9610	0.9426	0.9517	209
pink	0.9955	0.9801	0.9878	453
brown	0.9898	0.9858	0.9878	2466
multicolor	0.9629	0.9512	0.9570	1720
accuracy			0.9732	
macro avg	0.9751	0.9617	0.9675	12874
weighted avg	0.9740	0.9732	0.9733	12874

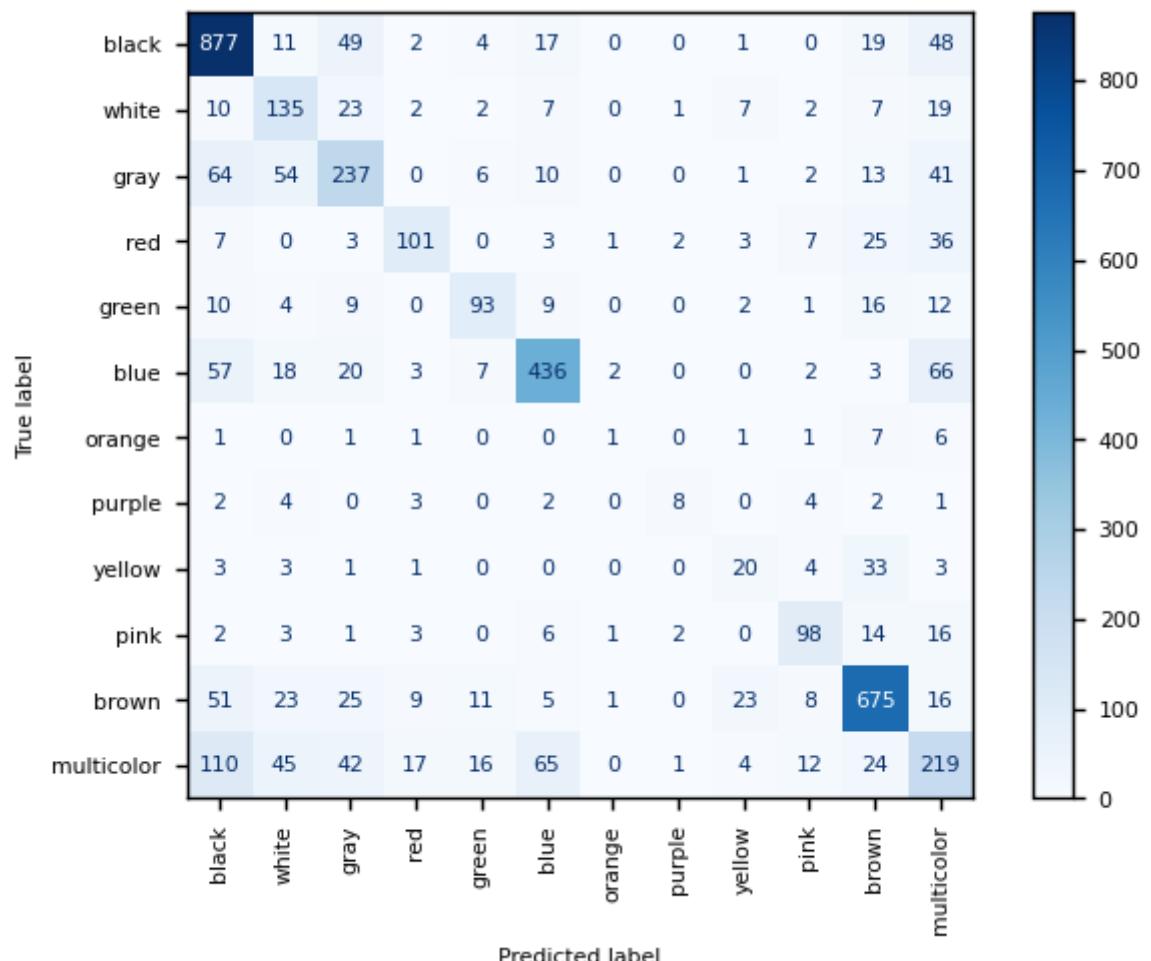


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7345	0.8531	0.7894	1028
white	0.4500	0.6279	0.5243	215
gray	0.5766	0.5537	0.5650	428
red	0.7113	0.5372	0.6121	188
green	0.6691	0.5962	0.6305	156
blue	0.7786	0.7101	0.7428	614
orange	0.1667	0.0526	0.0800	19
purple	0.5714	0.3077	0.4000	26
yellow	0.3226	0.2941	0.3077	68
pink	0.6950	0.6712	0.6829	146
brown	0.8055	0.7969	0.8012	847
multicolor	0.4534	0.3946	0.4220	555
accuracy			0.6760	
macro avg	0.5779	0.5330	0.5465	4290
weighted avg	0.6737	0.6760	0.6715	4290

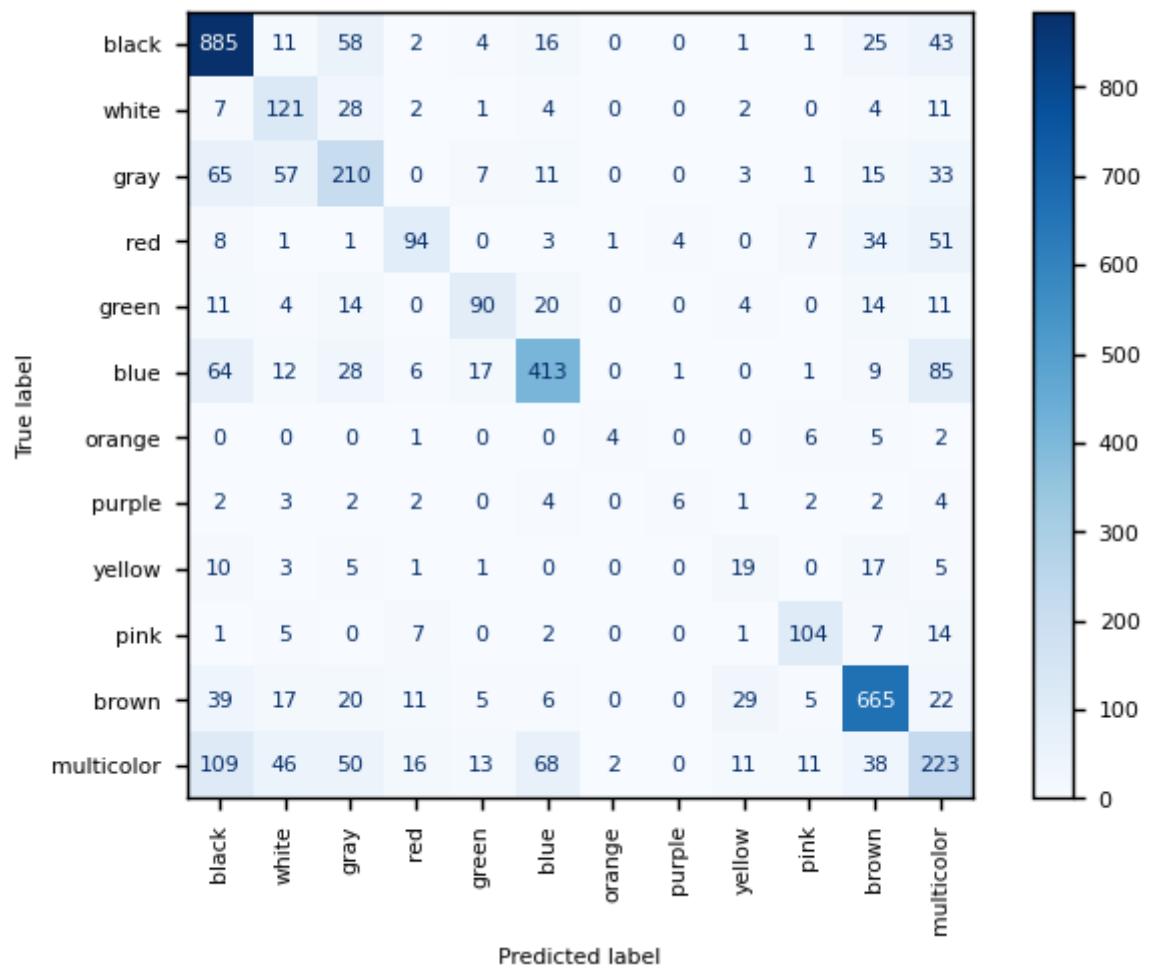


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7369	0.8461	0.7877	1046
white	0.4321	0.6722	0.5261	180
gray	0.5048	0.5224	0.5134	402
red	0.6620	0.4608	0.5434	204
green	0.6522	0.5357	0.5882	168
blue	0.7550	0.6494	0.6982	636
orange	0.5714	0.2222	0.3200	18
purple	0.5455	0.2143	0.3077	28
yellow	0.2676	0.3115	0.2879	61
pink	0.7536	0.7376	0.7455	141
brown	0.7964	0.8120	0.8041	819
multicolor	0.4425	0.3799	0.4088	587
accuracy			0.6606	
macro avg	0.5933	0.5303	0.5443	4290
weighted avg	0.6612	0.6606	0.6560	4290



Test Dataset Confusion Matrix

gray (97.1%)
 >> brown gray (99.2%)
 >> gray multicolor (72.4%)
 >> multicolor black (69.3%)
 >> brown white (100.0%)
 >> brown blue (51.9%)
 >> blue



brown (100.0%)
 >> brown blue (58.8%)
 >> blue black (75.2%)
 >> black black (98.9%)
 >> black black (100.0%)
 >> black brown (67.0%)
 >> blue



gray (97.8%)
 >> gray black (100.0%)
 >> black black (57.1%)
 >> black brown (100.0%)
 >> brown blue (99.8%)
 >> blue brown (100.0%)
 >> brown



pink (84.0%)
 >> pink black (100.0%)
 >> black multicolor (100.0%)
 >> red red (62.6%)
 >> blue blue (100.0%)
 >> blue black (100.0%)
 >> black



multicolor (99.9%)
 >> blue brown (100.0%)
 >> brown multicolor (100.0%)
 >> multicolor white (99.8%)
 >> black black (64.4%)
 >> multicolor brown (100.0%)
 >> brown



green (99.8%)
 >> multicolor black (100.0%)
 >> black blue (86.0%)
 >> gray brown (76.8%)
 >> gray black (99.9%)
 >> black brown (100.0%)
 >> brown



Sample Images from Test Dataset

gray (53.0%) blue (63.6%) brown (66.3%) white (100.0%) brown (76.8%) black (69.3%)
 >> black >> brown >> yellow >> brown >> gray >> brown



white (99.8%) multicolor (81.1%) blue (99.4%) multicolor (100.0%) multicolor (78.9%) blue (94.5%)
 >> black >> blue >> multicolor >> red >> white >> blue



black (84.5%) brown (100.0%) black (99.8%) multicolor (100.0%) white (100.0%) gray (97.1%)
 >> gray >> yellow >> multicolor >> blue >> purple >> brown



black (98.5%) yellow (36.3%) black (100.0%) green (89.0%) gray (100.0%) blue (86.0%)
 >> brown >> gray >> multicolor >> multicolor >> blue >> gray



multicolor (100.0%) multicolor (94.8%) brown (100.0%) brown (57.4%) black (89.5%) gray (72.6%)
 >> blue >> gray >> yellow >> blue >> red >> blue



multicolor (100.0%) brown (98.4%) blue (83.4%) yellow (74.5%) gray (98.8%) black (64.4%)
 >> red >> brown >> multicolor >> brown >> green >> multicolor



Sample Images that cannot be accurately predicted in the Test Dataset

3.2.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Pairwise Argmin	49.51 sec.	0.2954
Pairwise Mean	13.09 sec.	0.2266
Array Mean Argmin	60.24 sec.	0.2397
Pairwise First Quantile	97.38 sec.	0.3535
Pairwise Median	60.55 sec.	0.3454
Pairwise Third Quantile	52.52 sec.	0.2895

* Quantization based methods are available only in RGB color space

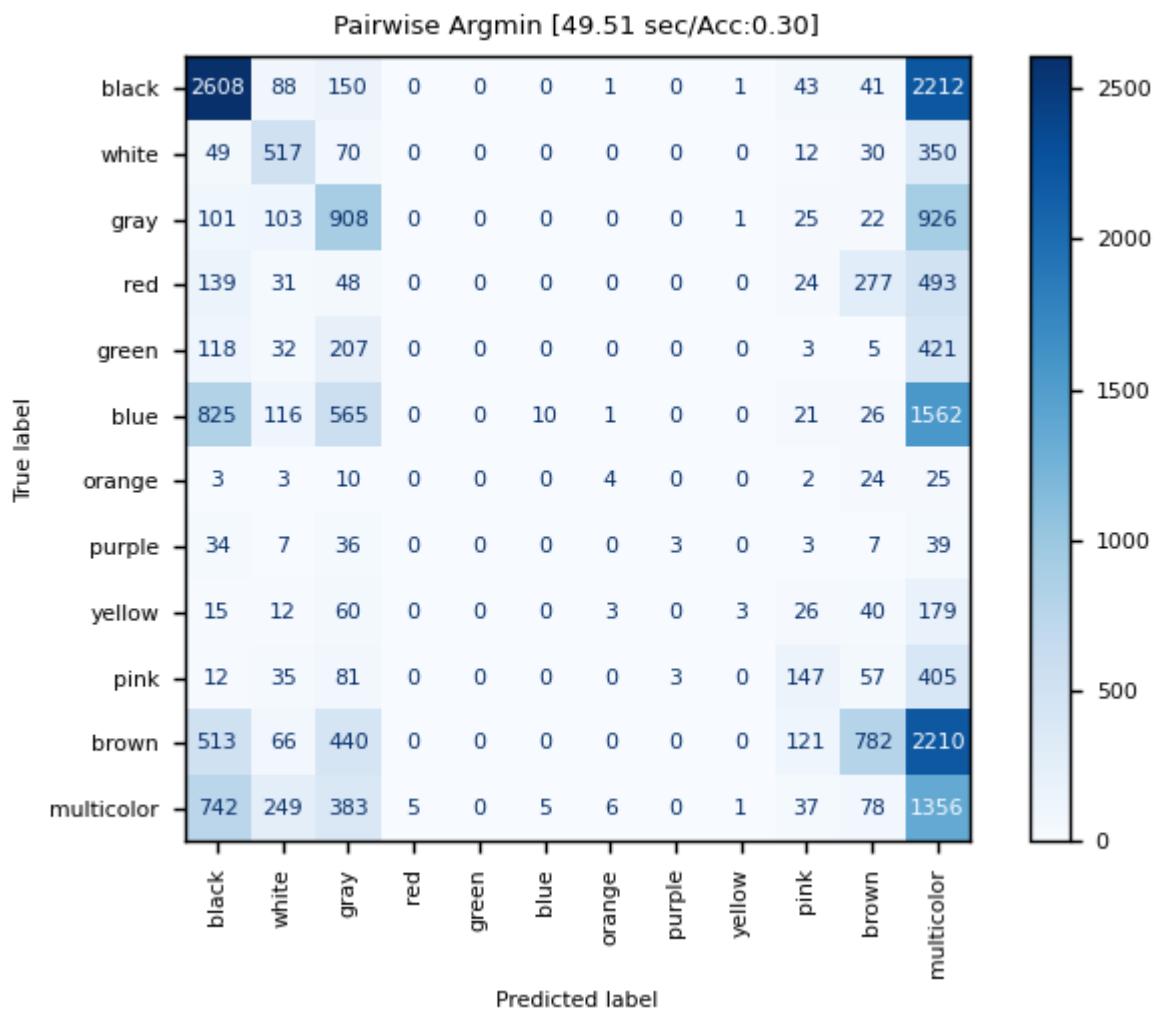
Unrefined Dataset Improvement Run

Multi-Color with HSL / Pairwise Argmin Result

Total Duration: 49.51 sec.

Multi-Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5055	0.5070	0.5063	5144
white	0.4106	0.5029	0.4521	1028
gray	0.3070	0.4353	0.3600	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.6667	0.0032	0.0064	3126
orange	0.2667	0.0563	0.0930	71
purple	0.5000	0.0233	0.0444	129
yellow	0.5000	0.0089	0.0174	338
pink	0.3168	0.1986	0.2442	740
brown	0.5630	0.1893	0.2833	4132
multicolor	0.1332	0.4738	0.2080	2862
accuracy			0.2954	
macro avg	0.3475	0.1999	0.1846	21454
weighted avg	0.4168	0.2954	0.2706	21454



Multi-Color with HSL / Pairwise Argmin Confusion Matrix

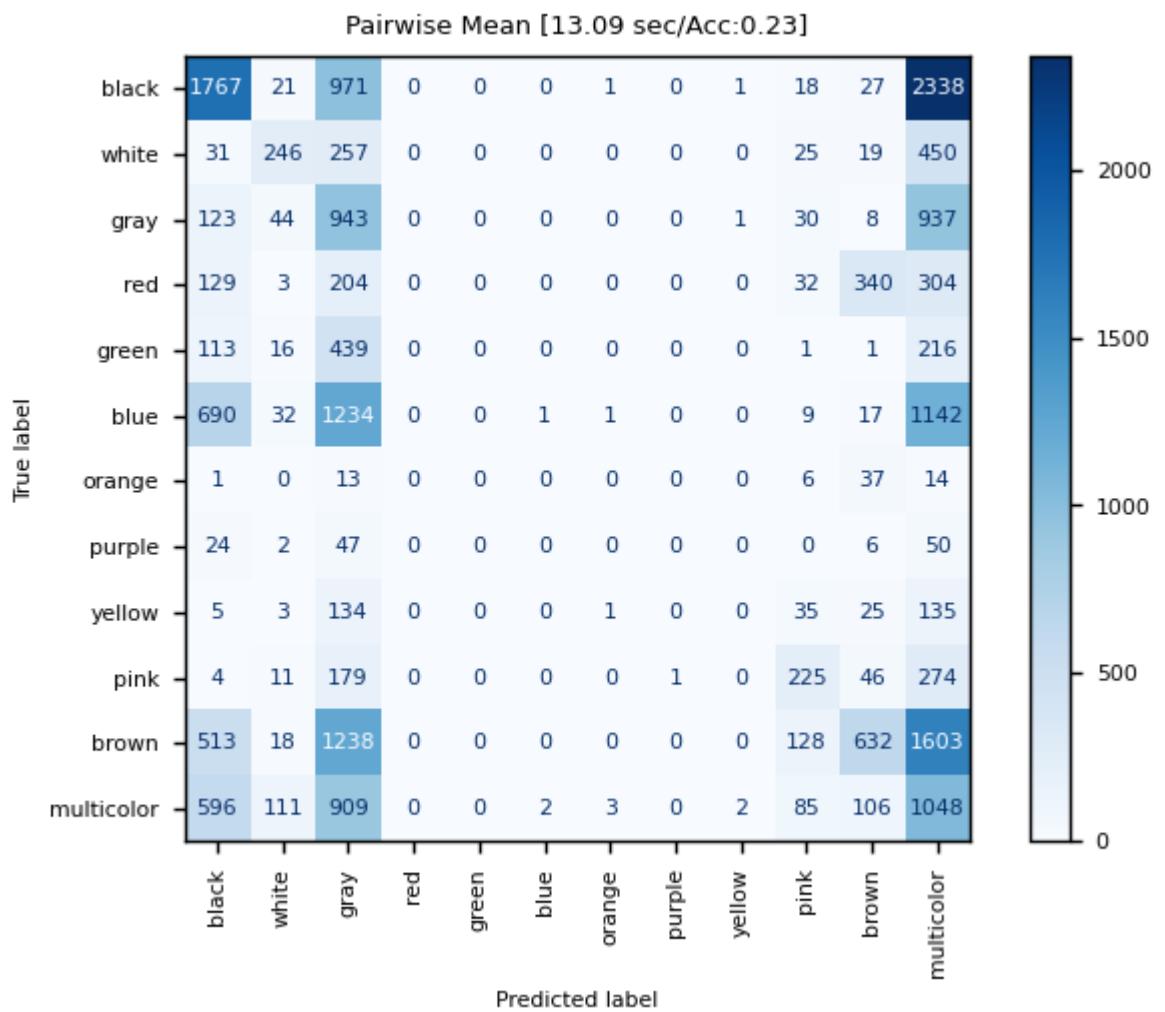
Unrefined Dataset Improvement Run

Multi-Color with HSL / Pairwise Mean Result

Total Duration: 13.09 sec.

Multi-Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.4422	0.3435	0.3867	5144
white	0.4852	0.2393	0.3205	1028
gray	0.1436	0.4521	0.2179	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.3333	0.0003	0.0006	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3788	0.3041	0.3373	740
brown	0.5000	0.1530	0.2342	4132
multicolor	0.1231	0.3662	0.1843	2862
accuracy			0.2266	
macro avg	0.2005	0.1549	0.1401	21454
weighted avg	0.3176	0.2266	0.2107	21454



Multi-Color with HSL / Pairwise Mean Confusion Matrix

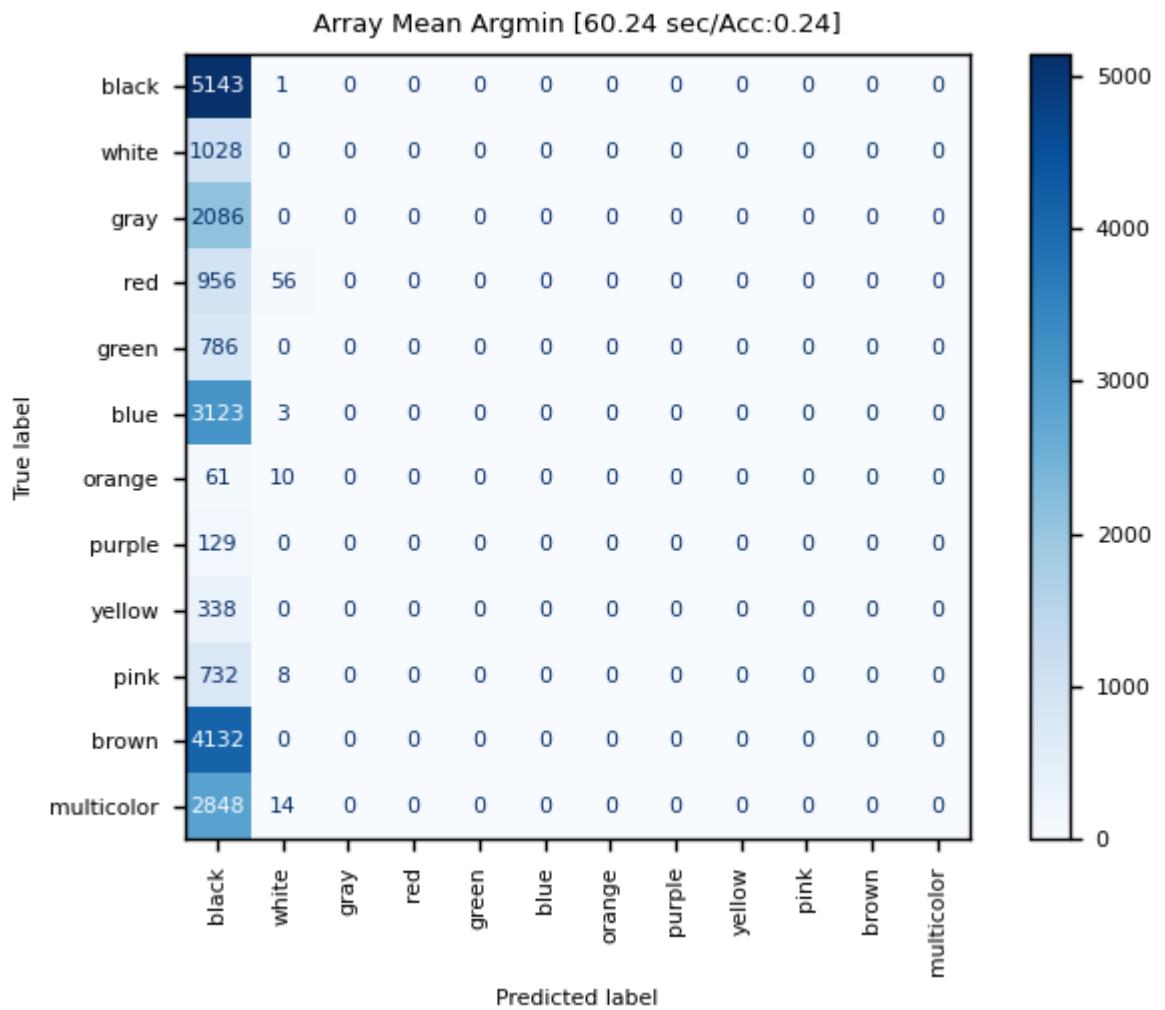
Unrefined Dataset Improvement Run

Multi-Color with HSL / Array Mean Argmin Result

Total Duration: 60.24 sec.

Multi-Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2408	0.9998	0.3881	5144
white	0.0000	0.0000	0.0000	1028
gray	0.0000	0.0000	0.0000	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.0000	0.0000	0.0000	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.0000	0.0000	0.0000	740
brown	0.0000	0.0000	0.0000	4132
multicolor	0.0000	0.0000	0.0000	2862
accuracy			0.2397	
macro avg	0.0201	0.0833	0.0323	21454
weighted avg	0.0577	0.2397	0.0930	21454



Multi-Color with HSL / Array Mean Argmin Confusion Matrix

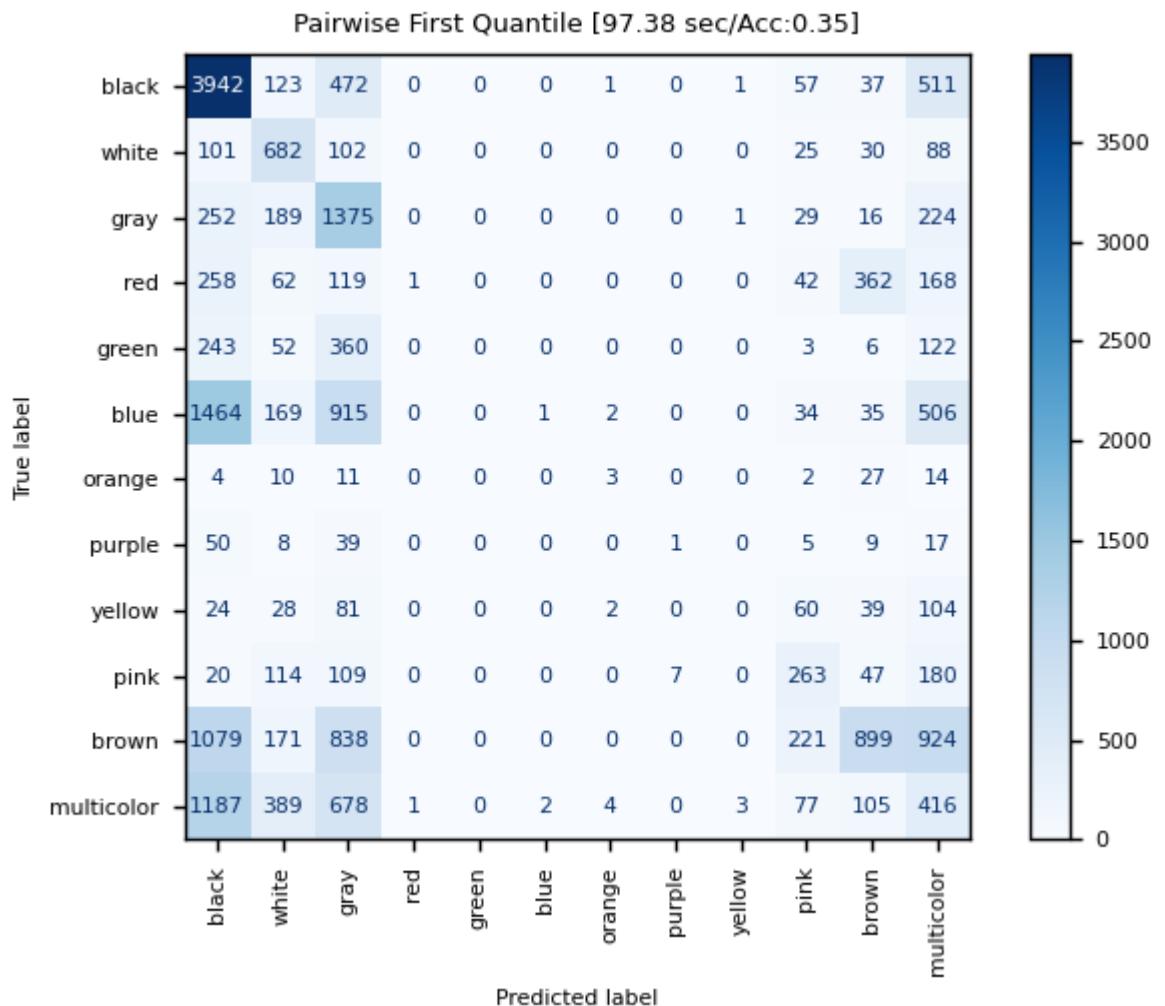
Unrefined Dataset Improvement Run

Multi-Color with HSL / Pairwise First Quantile Result

Total Duration: 97.38 sec.

Multi-Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4571	0.7663	0.5726	5144
white	0.3415	0.6634	0.4509	1028
gray	0.2697	0.6592	0.3827	2086
red	0.5000	0.0010	0.0020	1012
green	0.0000	0.0000	0.0000	786
blue	0.3333	0.0003	0.0006	3126
orange	0.2500	0.0423	0.0723	71
purple	0.1250	0.0078	0.0146	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3215	0.3554	0.3376	740
brown	0.5577	0.2176	0.3130	4132
multicolor	0.1271	0.1454	0.1356	2862
accuracy			0.3535	
macro avg	0.2736	0.2382	0.1902	21454
weighted avg	0.3614	0.3535	0.2867	21454



Multi-Color with HSL / Pairwise First Quantile Confusion Matrix

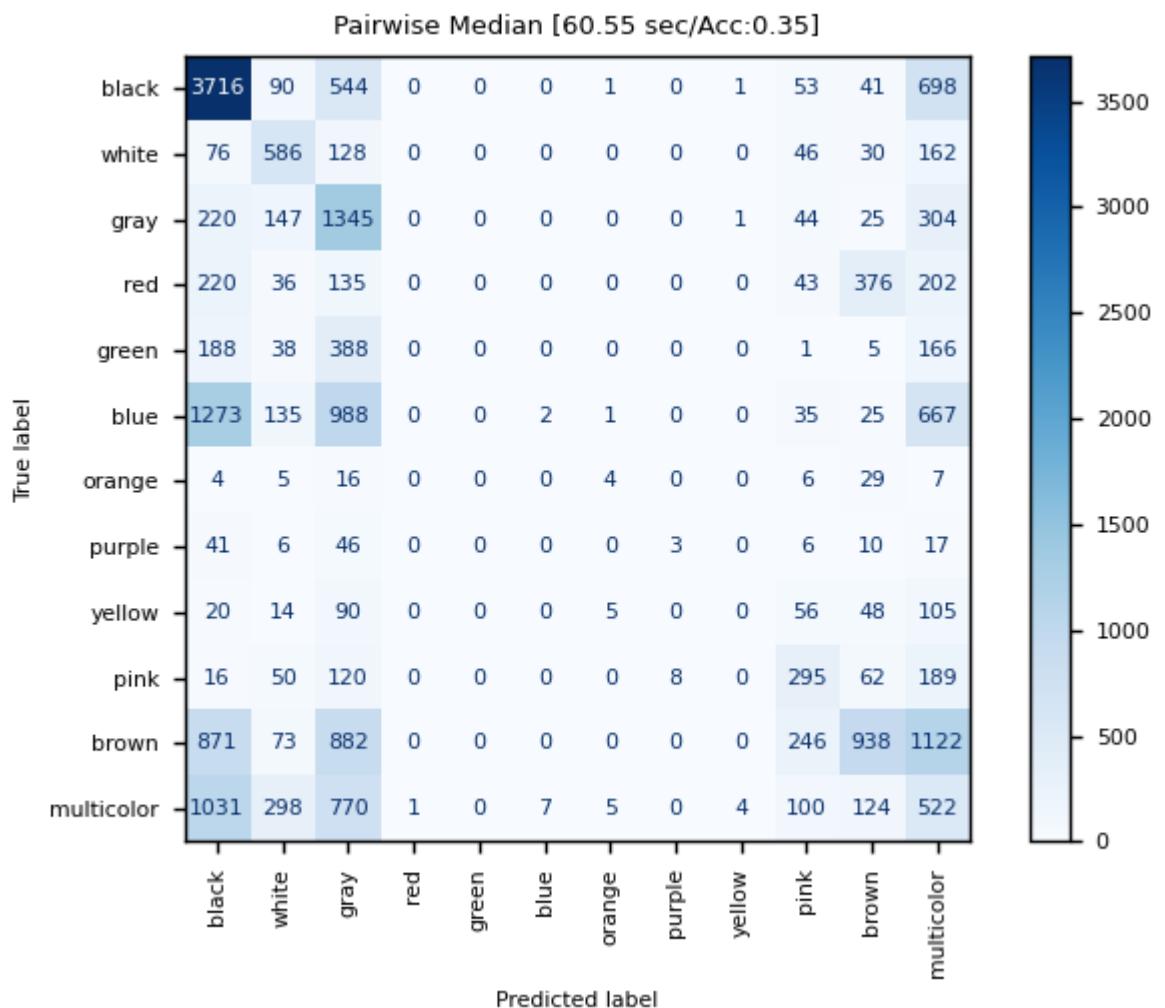
Unrefined Dataset Improvement Run

Multi-Color with HSL / Pairwise Median Result

Total Duration: 60.55 sec.

Multi-Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4841	0.7224	0.5797	5144
white	0.3965	0.5700	0.4677	1028
gray	0.2467	0.6448	0.3569	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.2222	0.0006	0.0013	3126
orange	0.2500	0.0563	0.0920	71
purple	0.2727	0.0233	0.0429	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3169	0.3986	0.3531	740
brown	0.5476	0.2270	0.3210	4132
multicolor	0.1255	0.1824	0.1487	2862
accuracy			0.3454	
macro avg	0.2385	0.2355	0.1969	21454
weighted avg	0.3270	0.3454	0.2907	21454



Multi-Color with HSL / Pairwise Median Confusion Matrix

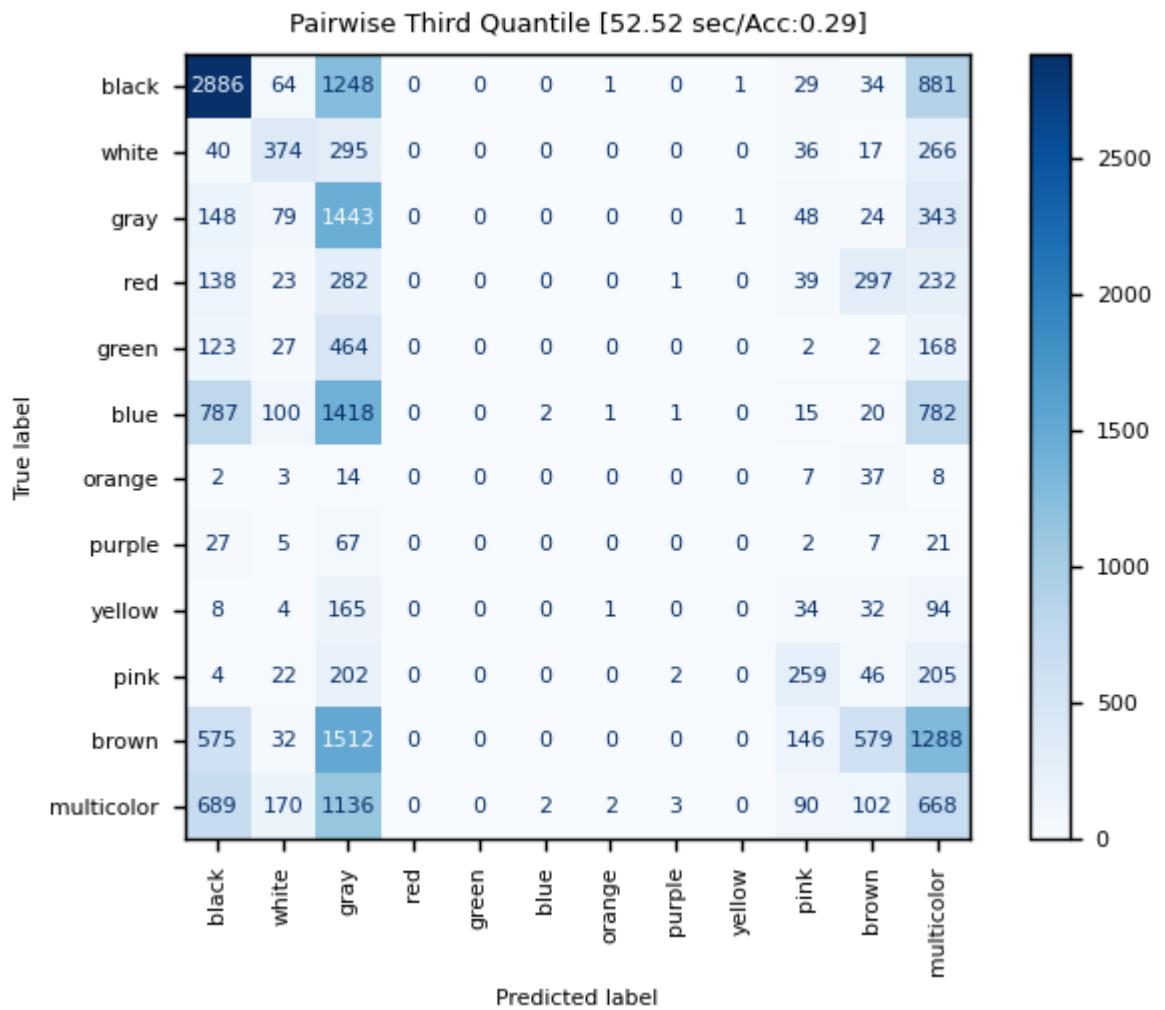
Unrefined Dataset Improvement Run

Multi-Color with HSL / Pairwise Third Quantile Result

Total Duration: 52.52 sec.

Multi-Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5318	0.5610	0.5460	5144
white	0.4142	0.3638	0.3874	1028
gray	0.1750	0.6918	0.2793	2086
red	0.0000	0.0000	0.0000	1012
green	0.0000	0.0000	0.0000	786
blue	0.5000	0.0006	0.0013	3126
orange	0.0000	0.0000	0.0000	71
purple	0.0000	0.0000	0.0000	129
yellow	0.0000	0.0000	0.0000	338
pink	0.3663	0.3500	0.3580	740
brown	0.4837	0.1401	0.2173	4132
multicolor	0.1348	0.2334	0.1709	2862
accuracy			0.2895	
macro avg	0.2171	0.1951	0.1633	21454
weighted avg	0.3610	0.2895	0.2538	21454



Multi-Color with HSL / Pairwise Third Quantile Confusion Matrix

MANUALLY VALIDATED DATASET IMPROVEMENT RUN

4.1. RGB Color Space

4.1.1. CNN AlexNet with Single Color

Manually Validated Dataset Improvement Run

CNN AlexNet Single Color with RGB

Total Duration: 498.66 sec.

Learning Rate: 0.000050

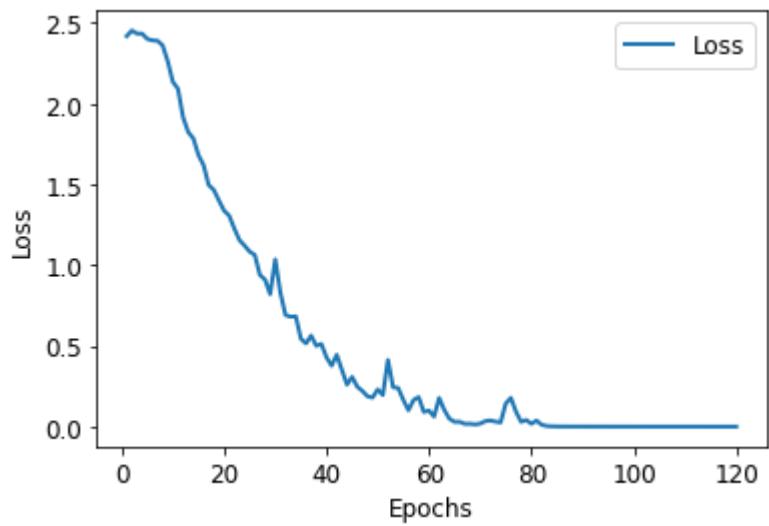
Dropout: 0.05

Batch Size: 16

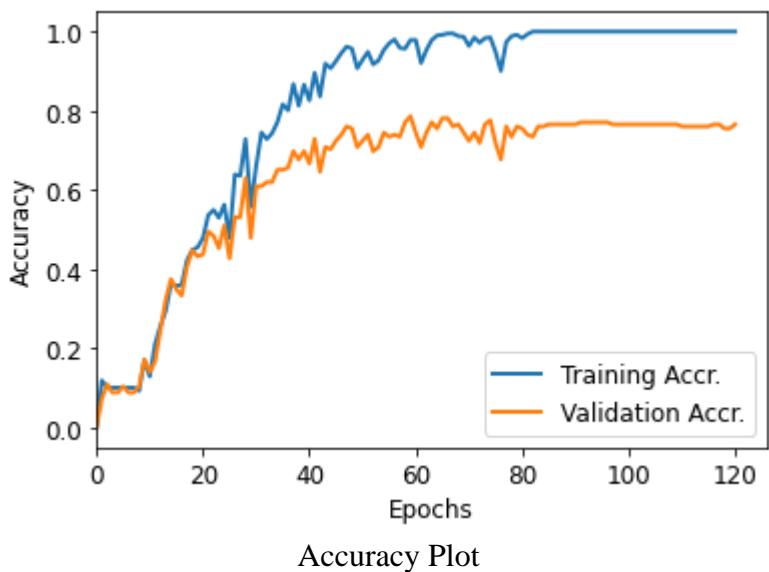
Epoch: 120

Best Performing Epoch: 59

Summary	
Method	Accuracy
Training Dataset	0.9785
Validation Dataset	0.7865
Test Dataset	0.7704



Loss Plot

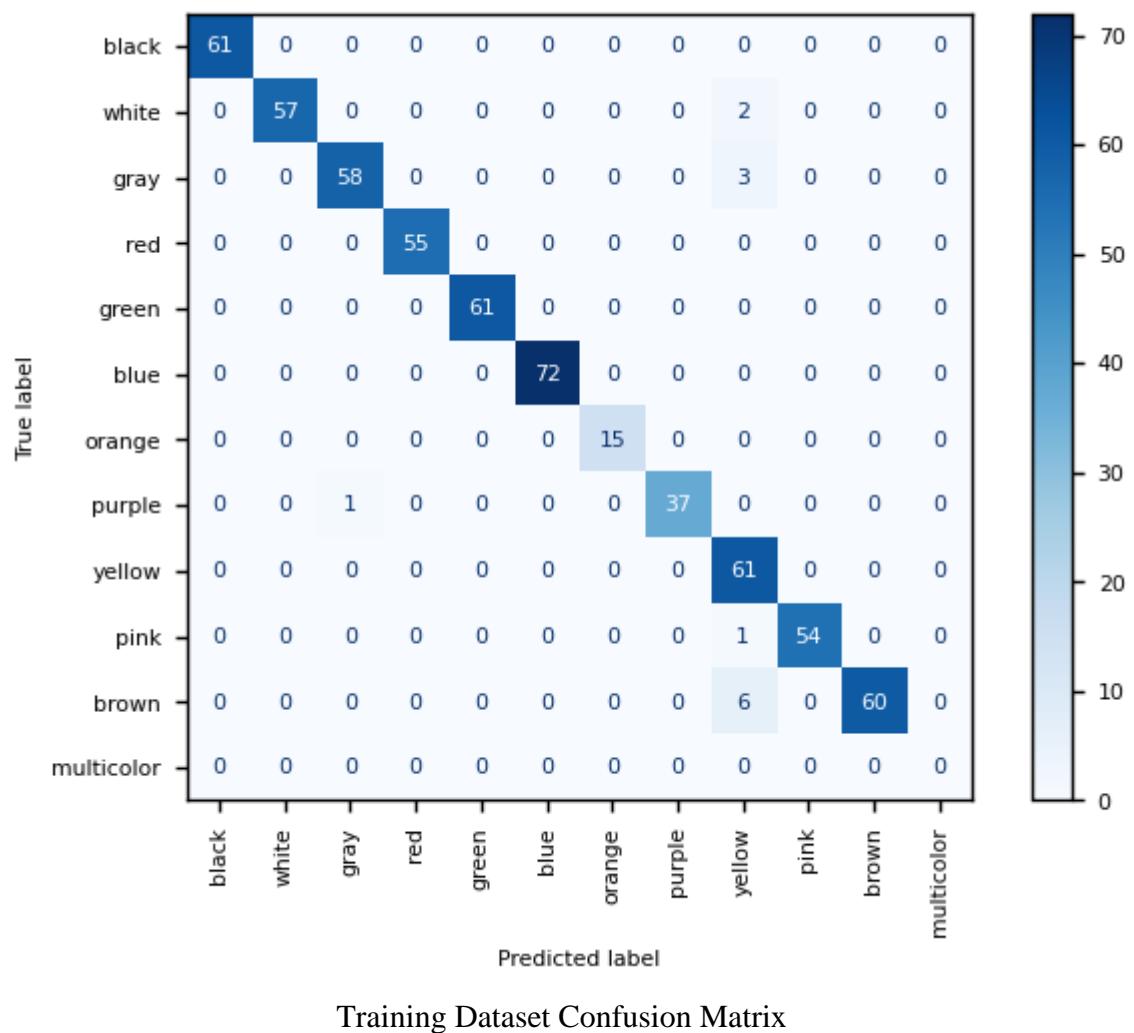


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

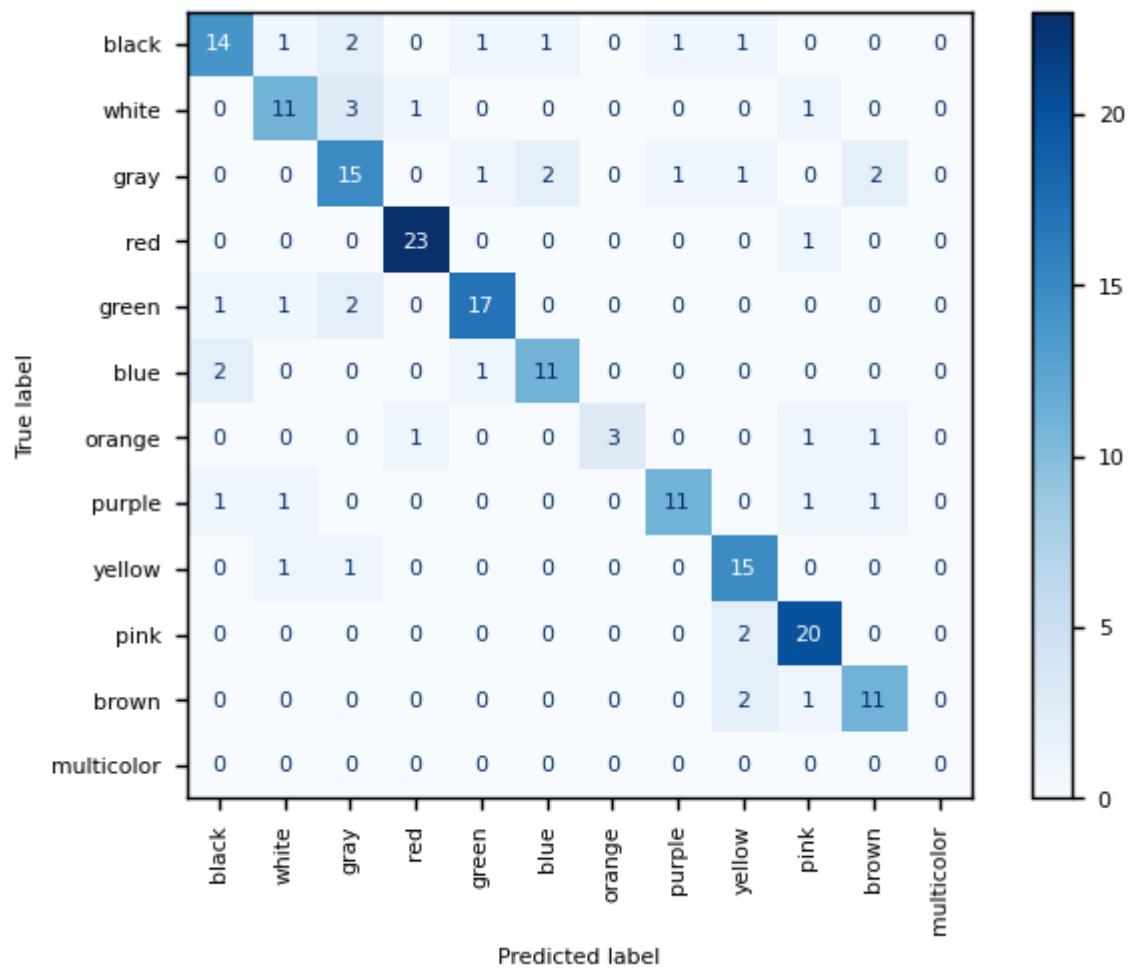
	Precision	Recall	F1 Score	Support
black	1.0000	1.0000	1.0000	61
white	1.0000	0.9661	0.9828	59
gray	0.9831	0.9508	0.9667	61
red	1.0000	1.0000	1.0000	55
green	1.0000	1.0000	1.0000	61
blue	1.0000	1.0000	1.0000	72
orange	1.0000	1.0000	1.0000	15
purple	1.0000	0.9737	0.9867	38
yellow	0.8356	1.0000	0.9104	61
pink	1.0000	0.9818	0.9908	55
brown	1.0000	0.9091	0.9524	66
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.9785	0.9785	0.9785	604
macro avg	0.9016	0.8985	0.8991	604
weighted avg	0.9817	0.9785	0.9790	604



Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7778	0.6667	0.7179	21
white	0.7333	0.6875	0.7097	16
gray	0.6522	0.6818	0.6667	22
red	0.9200	0.9583	0.9388	24
green	0.8500	0.8095	0.8293	21
blue	0.7857	0.7857	0.7857	14
orange	1.0000	0.5000	0.6667	6
purple	0.8462	0.7333	0.7857	15
yellow	0.7143	0.8824	0.7895	17
pink	0.8000	0.9091	0.8511	22
brown	0.7333	0.7857	0.7586	14
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7865	0.7865	0.7865	192
macro avg	0.7344	0.7000	0.7083	192
weighted avg	0.7919	0.7865	0.7843	192

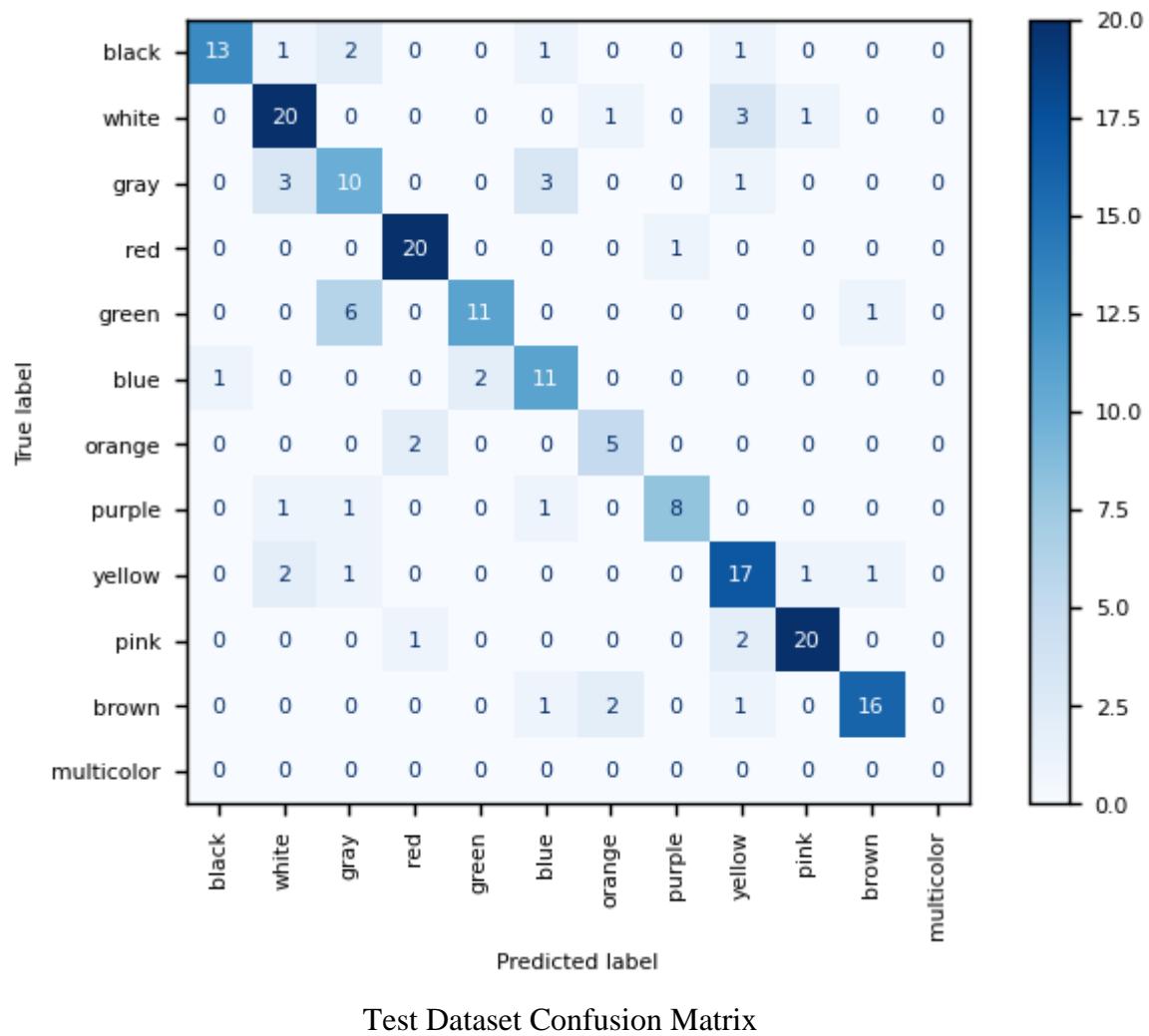


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9286	0.7222	0.8125	18
white	0.7407	0.8000	0.7692	25
gray	0.5000	0.5882	0.5405	17
red	0.8696	0.9524	0.9091	21
green	0.8462	0.6111	0.7097	18
blue	0.6471	0.7857	0.7097	14
orange	0.6250	0.7143	0.6667	7
purple	0.8889	0.7273	0.8000	11
yellow	0.6800	0.7727	0.7234	22
pink	0.9091	0.8696	0.8889	23
brown	0.8889	0.8000	0.8421	20
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.7704	0.7704	0.7704	196
macro avg	0.7103	0.6953	0.6976	196
weighted avg	0.7861	0.7704	0.7730	196



Test Dataset Confusion Matrix

pink (99.7%) gray (98.6%) white (99.8%) orange (69.6%) black (100.0%) black (99.9%)
 >> pink >> gray >> white >> white >> black >> black



white (64.1%) pink (98.0%) black (100.0%) gray (98.6%) blue (56.6%) red (99.9%)
 >> white >> pink >> black >> gray >> brown >> red



white (82.3%) gray (99.8%) pink (89.4%) brown (99.9%) red (100.0%) yellow (94.4%)
 >> gray >> gray >> white >> yellow >> red >> white



blue (100.0%) green (92.3%) blue (100.0%) blue (100.0%) black (100.0%) black (85.3%)
 >> blue >> green >> blue >> blue >> black >> black



orange (72.8%) gray (42.0%) brown (67.0%) brown (100.0%) blue (100.0%) red (100.0%)
 >> brown >> black >> brown >> brown >> blue >> red



blue (100.0%) green (86.1%) black (99.7%) brown (99.9%) gray (50.2%) white (91.4%)
 >> blue >> green >> black >> brown >> green >> white



Sample Images from Test Dataset

green (89.4%) blue (37.0%) gray (66.1%) gray (100.0%) green (83.1%) white (82.3%)
 >> blue >> gray >> green >> black >> blue >> gray



white (98.1%) gray (82.3%) blue (99.3%) yellow (57.0%) brown (78.5%) gray (98.9%)
 >> gray >> green >> black >> black >> green >> green



orange (78.6%) yellow (94.1%) brown (99.9%) blue (77.7%) gray (97.6%) white (80.8%)
 >> brown >> white >> yellow >> gray >> green >> purple



blue (80.7%) gray (50.2%) blue (56.6%) yellow (91.9%) red (80.7%) yellow (93.3%)
 >> gray >> green >> brown >> white >> pink >> pink



gray (74.3%) yellow (94.4%) black (98.6%) red (87.7%) white (47.7%) orange (69.6%)
 >> green >> pink >> blue >> orange >> gray >> white



yellow (56.3%) pink (68.1%) white (56.3%) gray (42.0%) purple (77.2%) pink (89.4%)
 >> gray >> yellow >> yellow >> black >> red >> white



Sample Images that cannot be accurately predicted in the Test Dataset

4.1.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Quantization Fastoctree	2.99 sec.	0.3296
Quantization Mmcq	44.22 sec.	0.2933
Quantization Libimagequant	17.99 sec.	0.2702
Pairwise Argmin	0.47 sec.	0.3196
Pairwise Mean	0.34 sec.	0.2944
Array Mean Argmin	0.34 sec.	0.0988
Pairwise First Quantile	1.91 sec.	0.3347
Pairwise Median	1.91 sec.	0.3145
Pairwise Third Quantile	2.00 sec.	0.2772

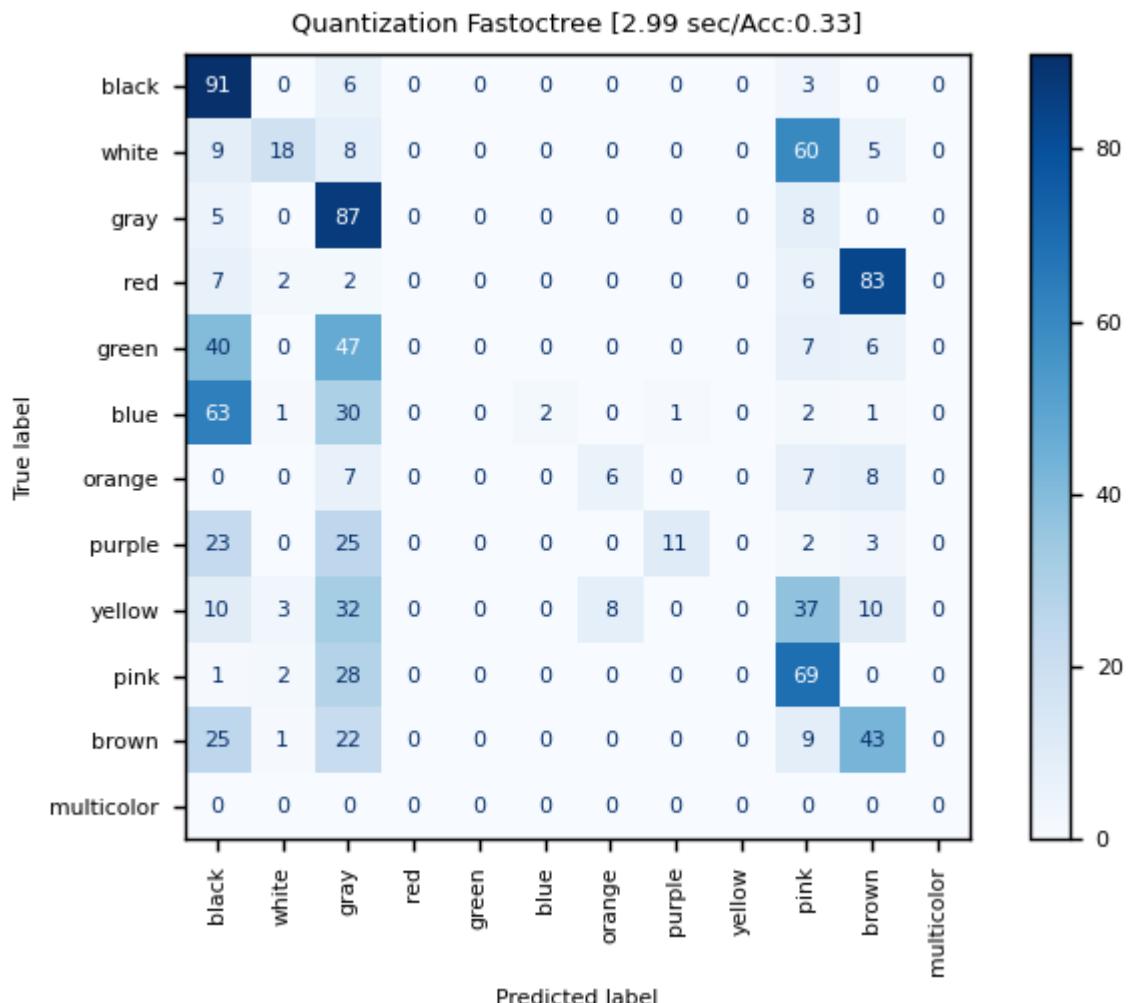
Manually Validated Dataset Improvement Run

Single Color with RGB / Quantization Fastoctree Result

Total Duration: 2.99 sec.

Single Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.3321	0.9100	0.4866	100
white	0.6667	0.1800	0.2835	100
gray	0.2959	0.8700	0.4416	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0200	0.0392	100
orange	0.4286	0.2143	0.2857	28
purple	0.9167	0.1719	0.2895	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3286	0.6900	0.4452	100
brown	0.2704	0.4300	0.3320	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3296	0.3296	0.3296	992
macro avg	0.3532	0.2905	0.2169	992
weighted avg	0.3629	0.3296	0.2312	992



Single Color with RGB / Quantization Fastoctree Confusion Matrix

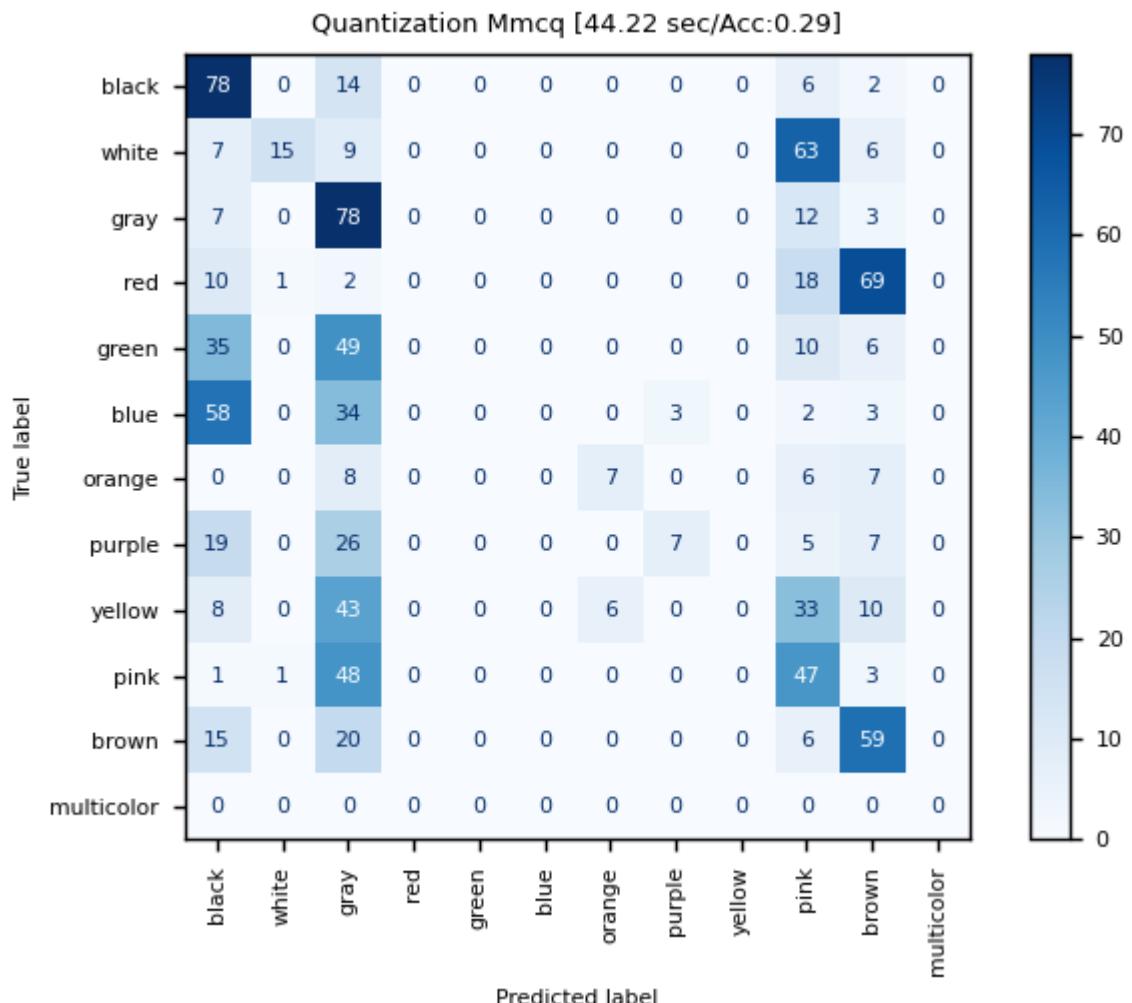
Manually Validated Dataset Improvement Run

Single Color with RGB / Quantization Mmcq Result

Total Duration: 44.22 sec.

Single Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.3277	0.7800	0.4615	100
white	0.8824	0.1500	0.2564	100
gray	0.2356	0.7800	0.3619	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.5385	0.2500	0.3415	28
purple	0.7000	0.1094	0.1892	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2260	0.4700	0.3052	100
brown	0.3371	0.5900	0.4291	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2933	0.2933	0.2933	992
macro avg	0.2706	0.2608	0.1954	992
weighted avg	0.2629	0.2933	0.2047	992



Single Color with RGB / Quantization Mmcq Confusion Matrix

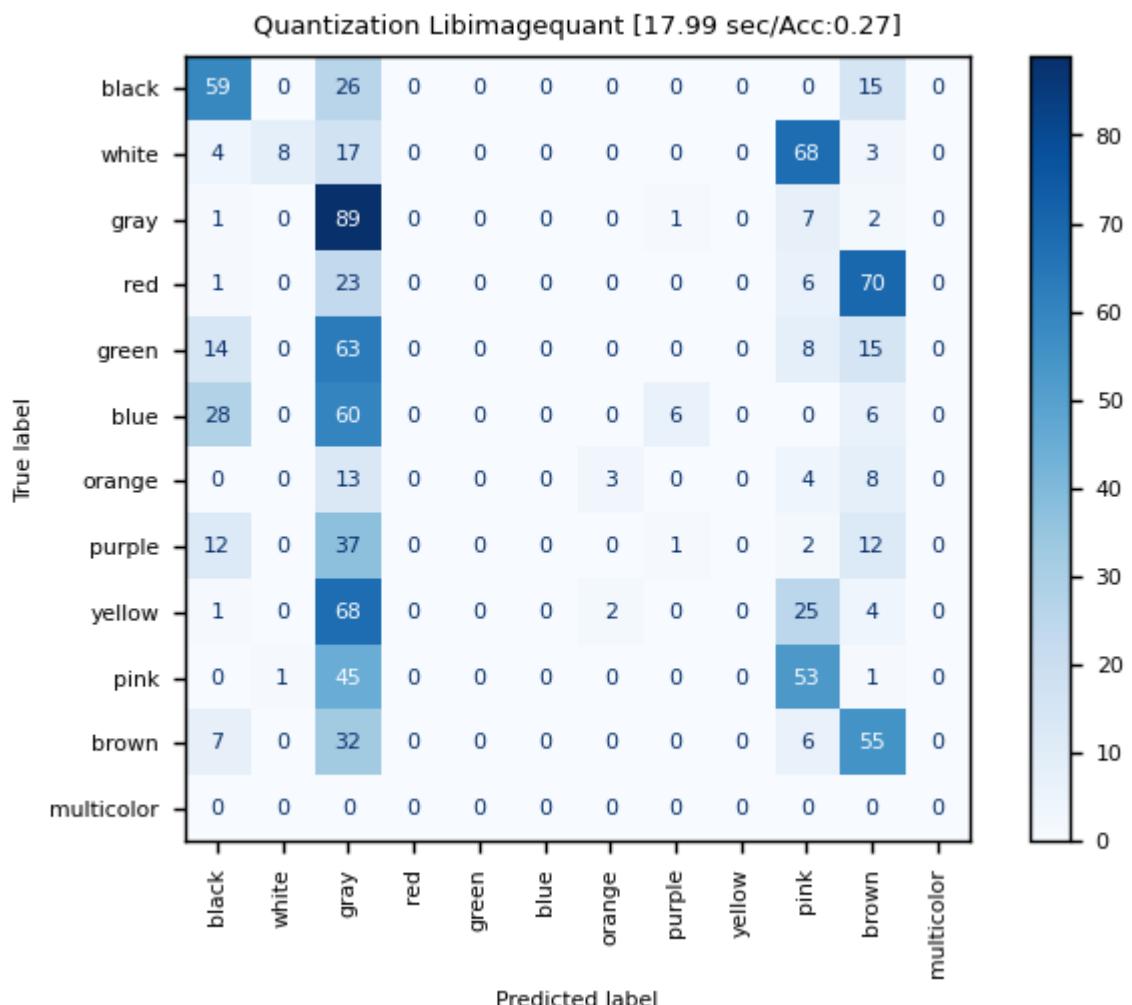
Manually Validated Dataset Improvement Run

Single Color with RGB / Quantization Libimagequant Result

Total Duration: 17.99 sec.

Single Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.4646	0.5900	0.5198	100
white	0.8889	0.0800	0.1468	100
gray	0.1882	0.8900	0.3106	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.6000	0.1071	0.1818	28
purple	0.1250	0.0156	0.0278	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2961	0.5300	0.3799	100
brown	0.2880	0.5500	0.3780	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2702	0.2702	0.2702	992
macro avg	0.2376	0.2302	0.1621	992
weighted avg	0.2393	0.2702	0.1818	992



Single Color with RGB / Quantization Libimagequant Confusion Matrix

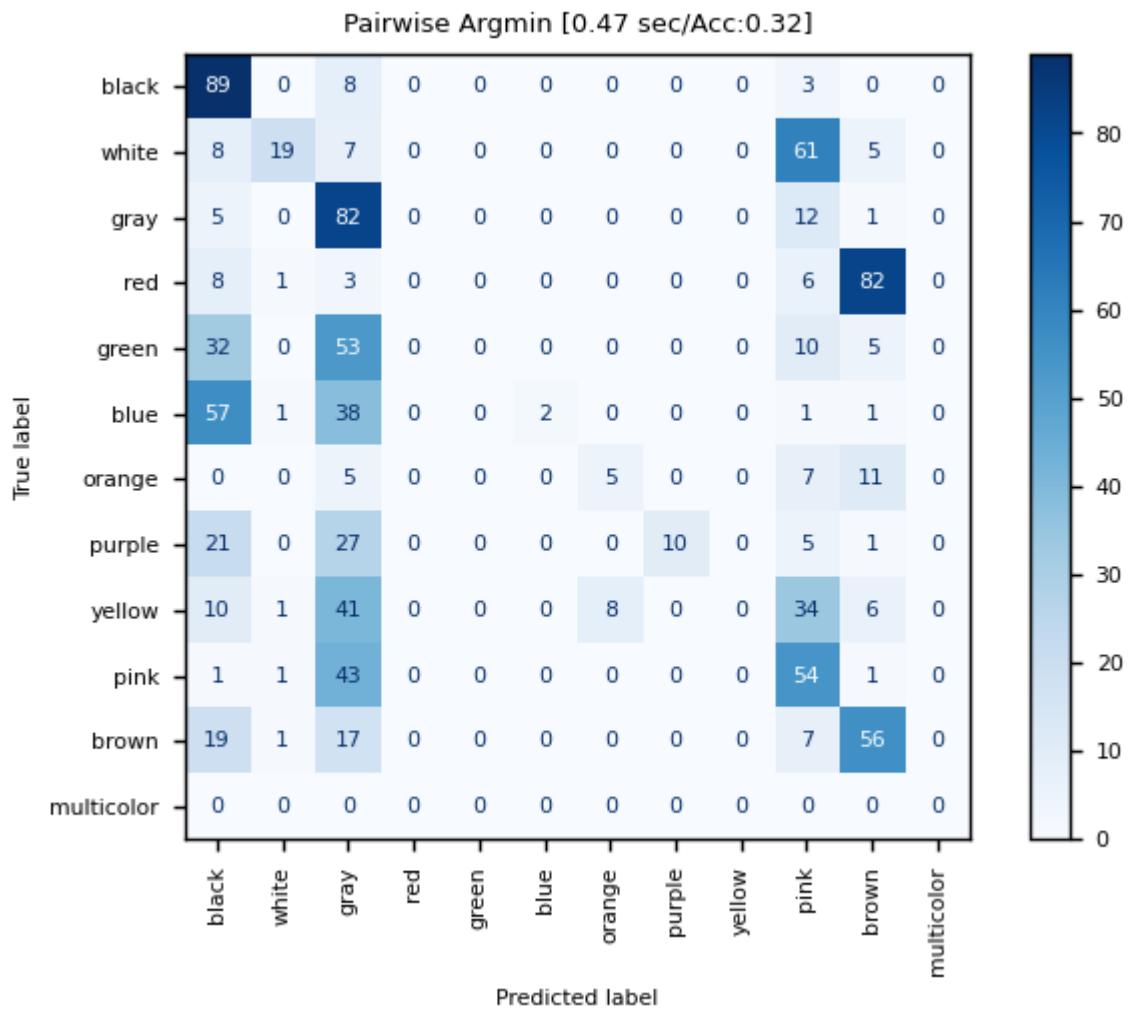
Manually Validated Dataset Improvement Run

Single Color with RGB / Pairwise Argmin Result

Total Duration: 0.47 sec.

Single Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.3560	0.8900	0.5086	100
white	0.7917	0.1900	0.3065	100
gray	0.2531	0.8200	0.3868	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0200	0.0392	100
orange	0.3846	0.1786	0.2439	28
purple	1.0000	0.1562	0.2703	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2700	0.5400	0.3600	100
brown	0.3314	0.5600	0.4164	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3196	0.3196	0.3196	992
macro avg	0.3656	0.2796	0.2110	992
weighted avg	0.3780	0.3196	0.2277	992



Single Color with RGB / Pairwise Argmin Confusion Matrix

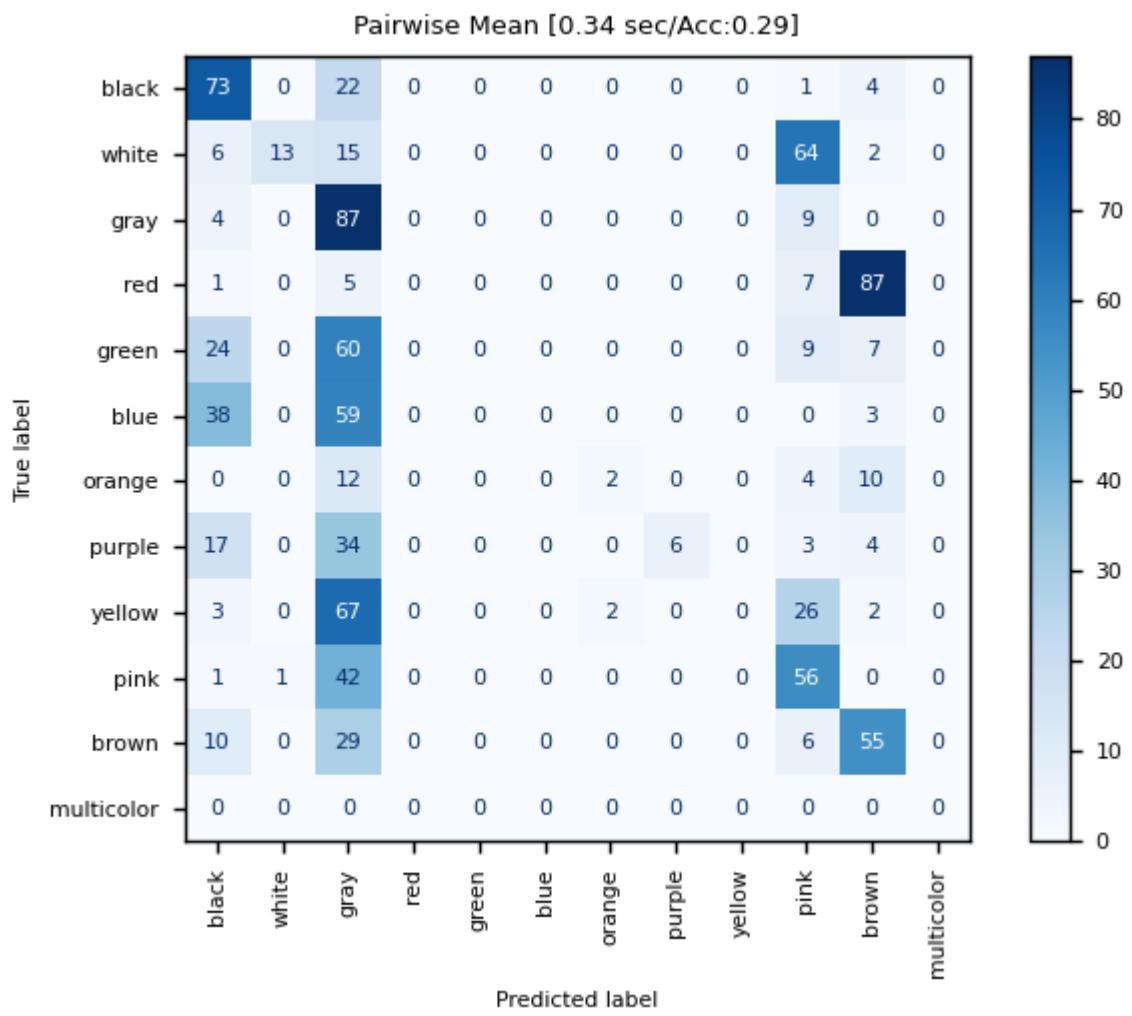
Manually Validated Dataset Improvement Run

Single Color with RGB / Pairwise Mean Result

Total Duration: 0.34 sec.

Single Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.4124	0.7300	0.5271	100
white	0.9286	0.1300	0.2281	100
gray	0.2014	0.8700	0.3271	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.5000	0.0714	0.1250	28
purple	1.0000	0.0938	0.1714	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3027	0.5600	0.3930	100
brown	0.3161	0.5500	0.4015	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2944	0.2944	0.2944	992
macro avg	0.3051	0.2504	0.1811	992
weighted avg	0.2965	0.2944	0.2038	992



Single Color with RGB / Pairwise Mean Confusion Matrix

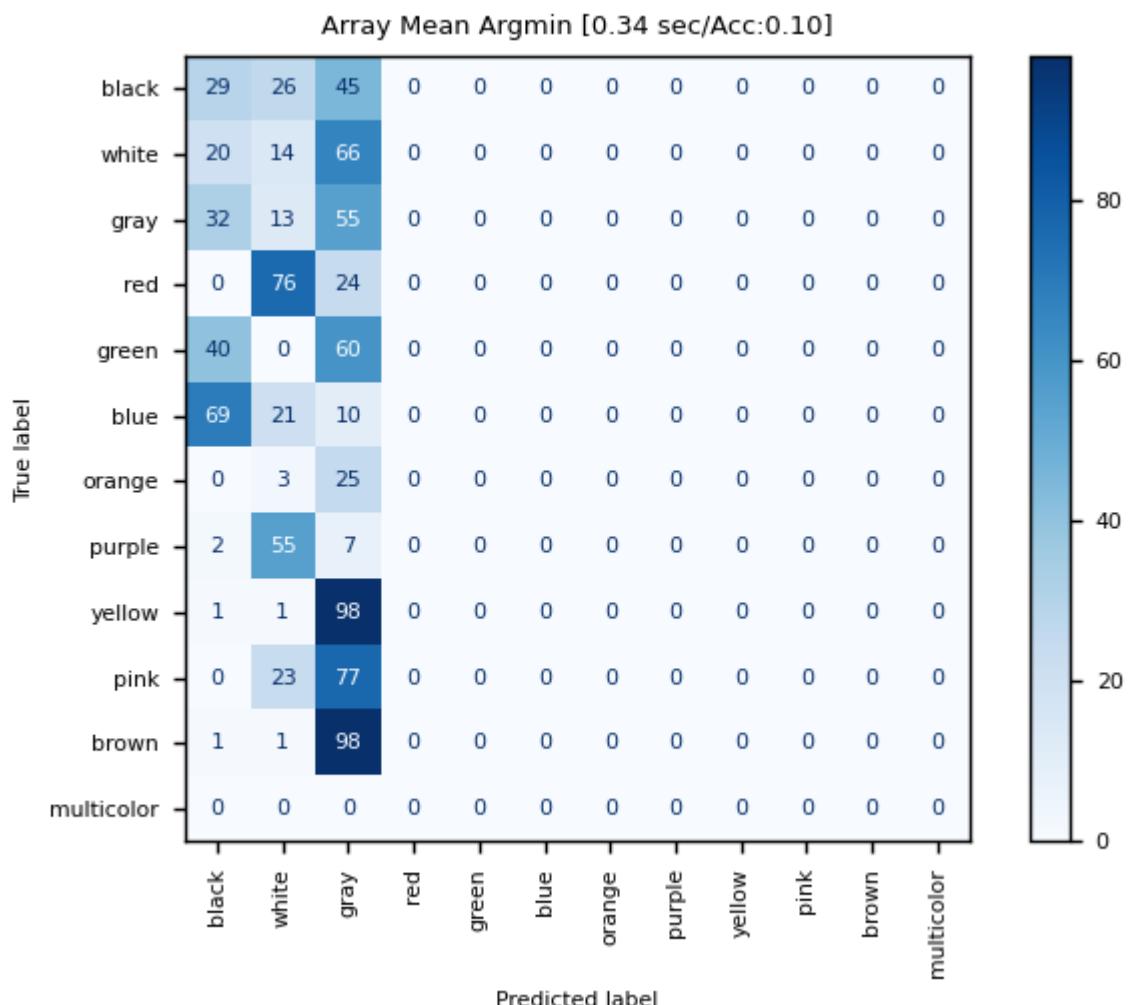
Manually Validated Dataset Improvement Run

Single Color with RGB / Array Mean Argmin Result

Total Duration: 0.34 sec.

Single Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.1495	0.2900	0.1973	100
white	0.0601	0.1400	0.0841	100
gray	0.0973	0.5500	0.1654	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.0000	0.0000	0.0000	100
brown	0.0000	0.0000	0.0000	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.0988	0.0988	0.0988	992
macro avg	0.0256	0.0817	0.0372	992
weighted avg	0.0309	0.0988	0.0450	992



Single Color with RGB / Array Mean Argmin Confusion Matrix

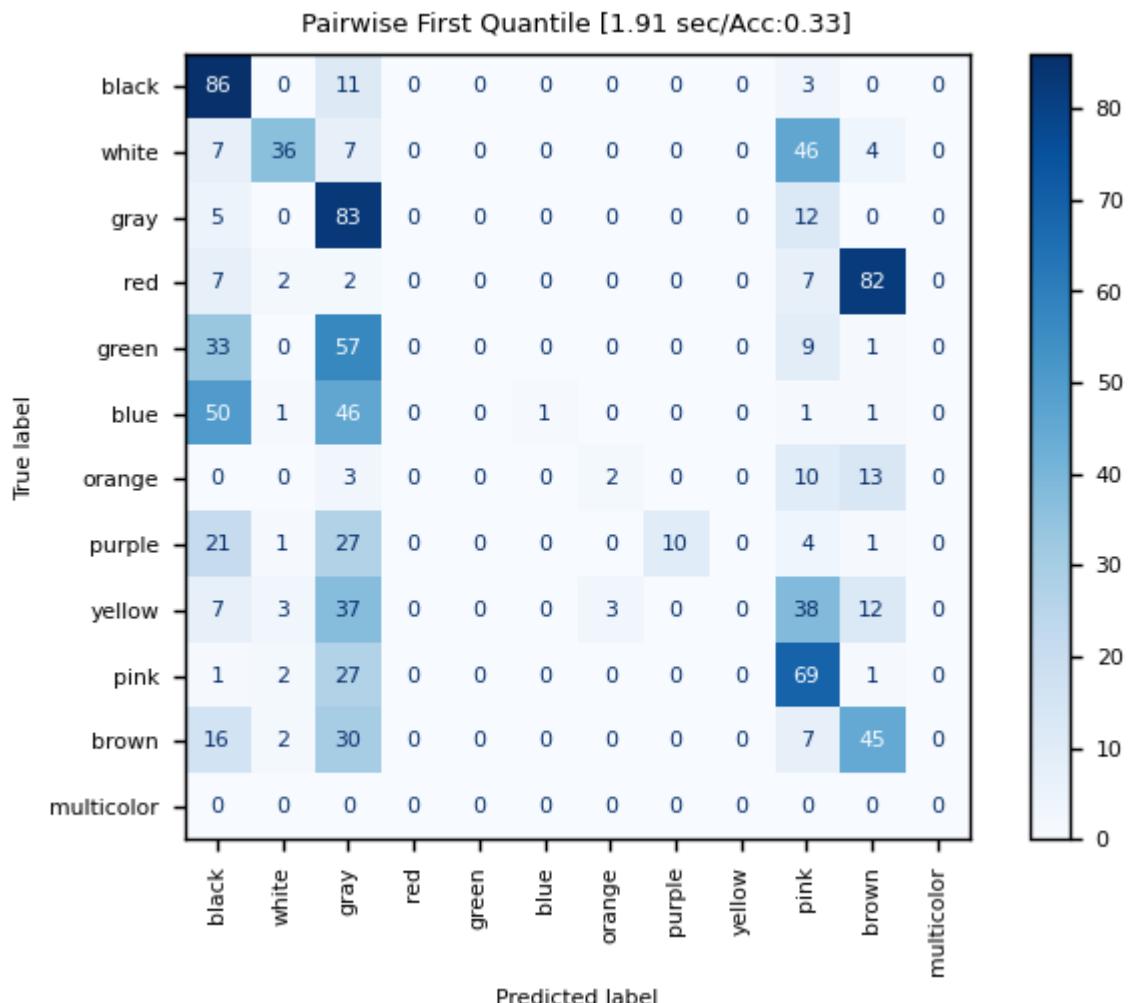
Manually Validated Dataset Improvement Run

Single Color with RGB / Pairwise First Quantile Result

Total Duration: 1.91 sec.

Single Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.3691	0.8600	0.5165	100
white	0.7660	0.3600	0.4898	100
gray	0.2515	0.8300	0.3860	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0100	0.0198	100
orange	0.4000	0.0714	0.1212	28
purple	1.0000	0.1562	0.2703	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3350	0.6900	0.4510	100
brown	0.2812	0.4500	0.3462	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3347	0.3347	0.3347	992
macro avg	0.3669	0.2856	0.2167	992
weighted avg	0.3785	0.3347	0.2436	992



Single Color with RGB / Pairwise First Quantile Confusion Matrix

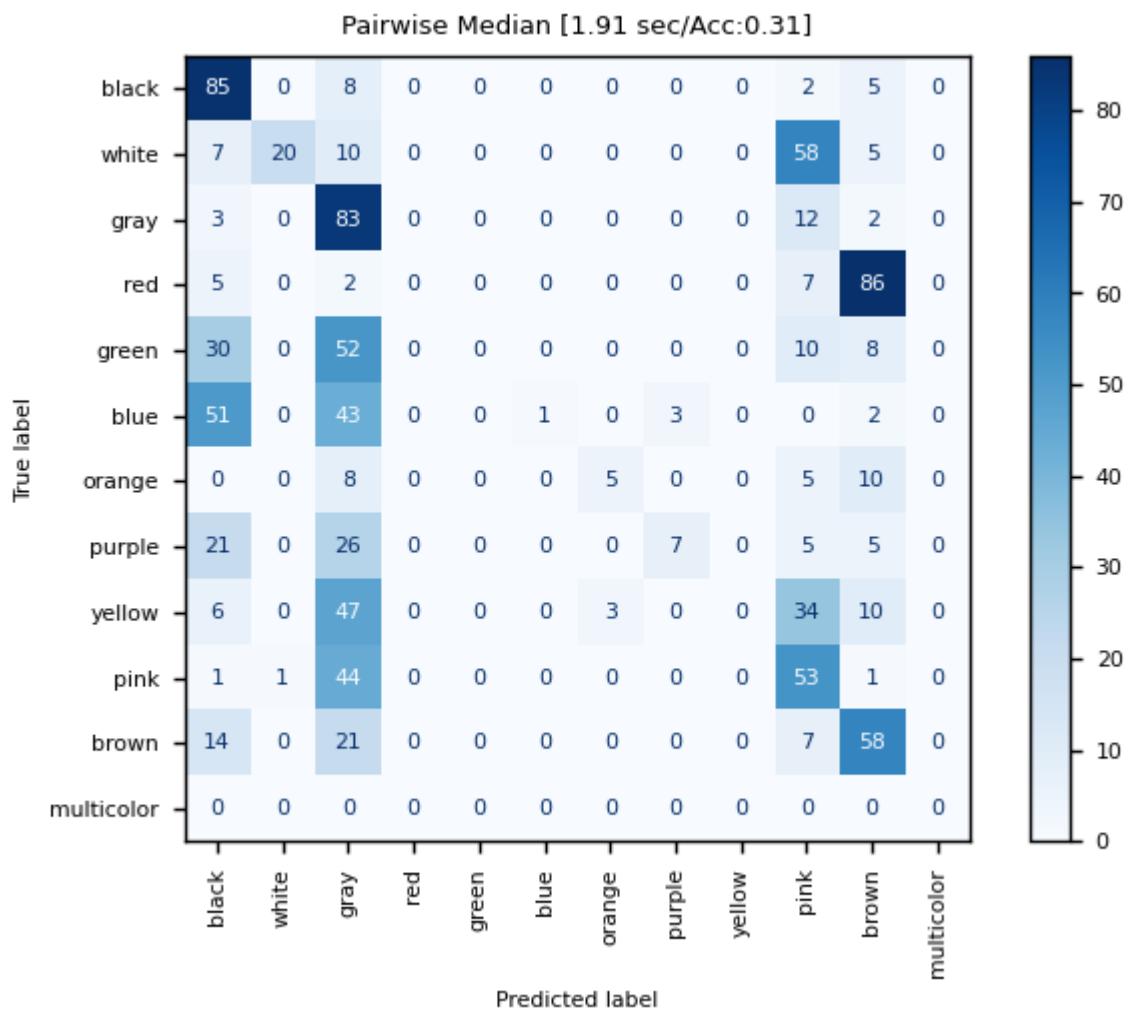
Manually Validated Dataset Improvement Run

Single Color with RGB / Pairwise Median Result

Total Duration: 1.91 sec.

Single Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.3812	0.8500	0.5263	100
white	0.9524	0.2000	0.3306	100
gray	0.2413	0.8300	0.3739	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0100	0.0198	100
orange	0.6250	0.1786	0.2778	28
purple	0.7000	0.1094	0.1892	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2746	0.5300	0.3618	100
brown	0.3021	0.5800	0.3973	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3145	0.3145	0.3145	992
macro avg	0.3730	0.2740	0.2064	992
weighted avg	0.3805	0.3145	0.2226	992



Single Color with RGB / Pairwise Median Confusion Matrix

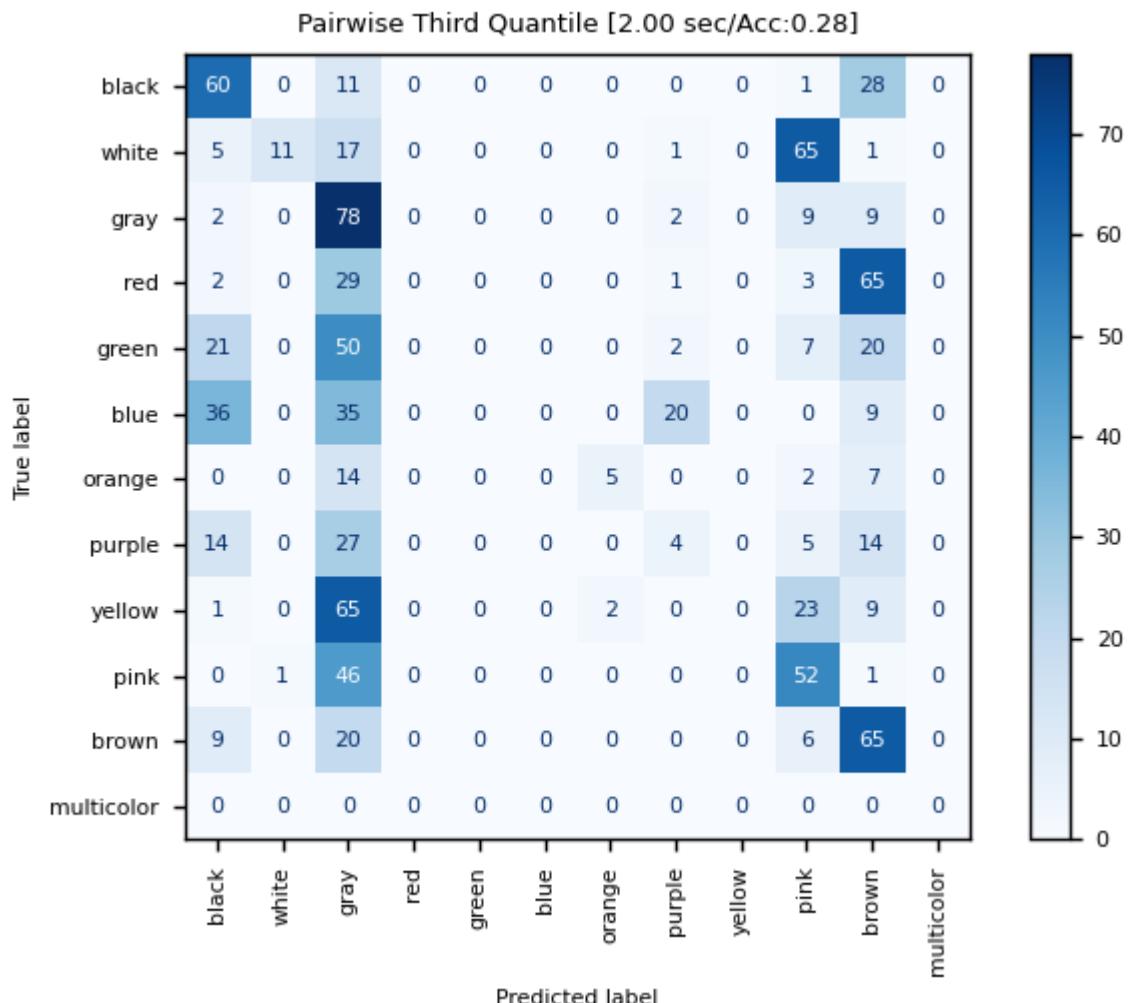
Manually Validated Dataset Improvement Run

Single Color with RGB / Pairwise Third Quantile Result

Total Duration: 2.00 sec.

Single Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4000	0.6000	0.4800	100
white	0.9167	0.1100	0.1964	100
gray	0.1990	0.7800	0.3171	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.7143	0.1786	0.2857	28
purple	0.1333	0.0625	0.0851	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3006	0.5200	0.3810	100
brown	0.2851	0.6500	0.3963	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2772	0.2772	0.2772	992
macro avg	0.2457	0.2418	0.1785	992
weighted avg	0.2406	0.2772	0.1921	992



Single Color with RGB / Pairwise Third Quantile Confusion Matrix

4.1.3. CNN AlexNet with Multi-color

Manually Validated Dataset Improvement Run

CNN AlexNet Multi-Color with RGB

Total Duration: 524.06 sec.

Learning Rate: 0.000050

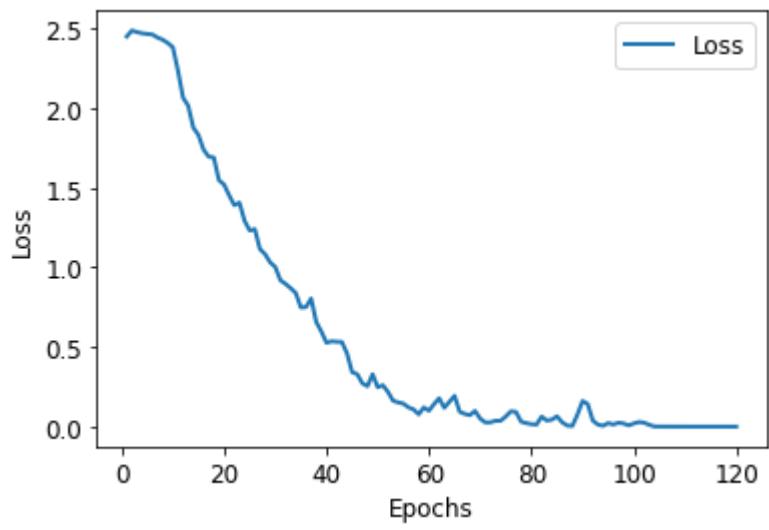
Dropout: 0.05

Batch Size: 16

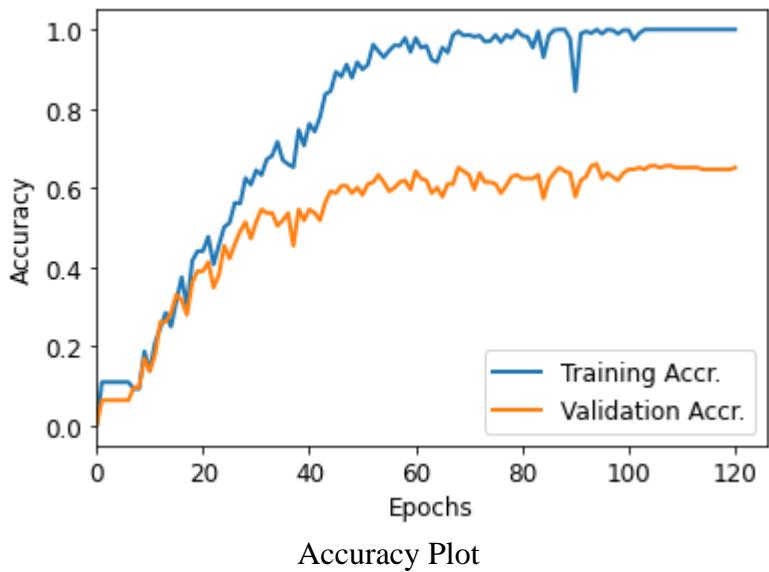
Epoch: 120

Best Performing Epoch: 94

Summary	
Method	Accuracy
Training Dataset	1.0000
Validation Dataset	0.6606
Test Dataset	0.6743



Loss Plot

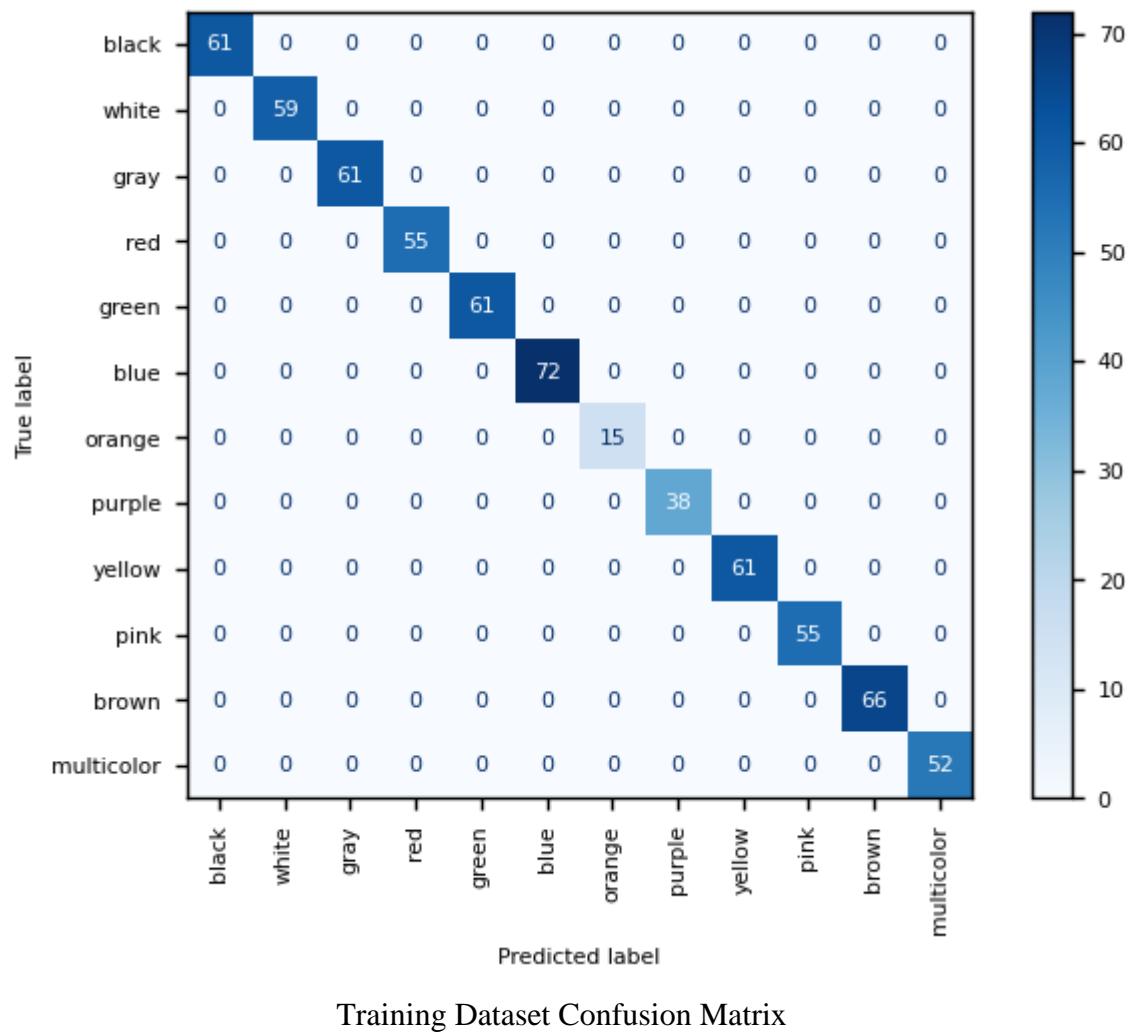


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

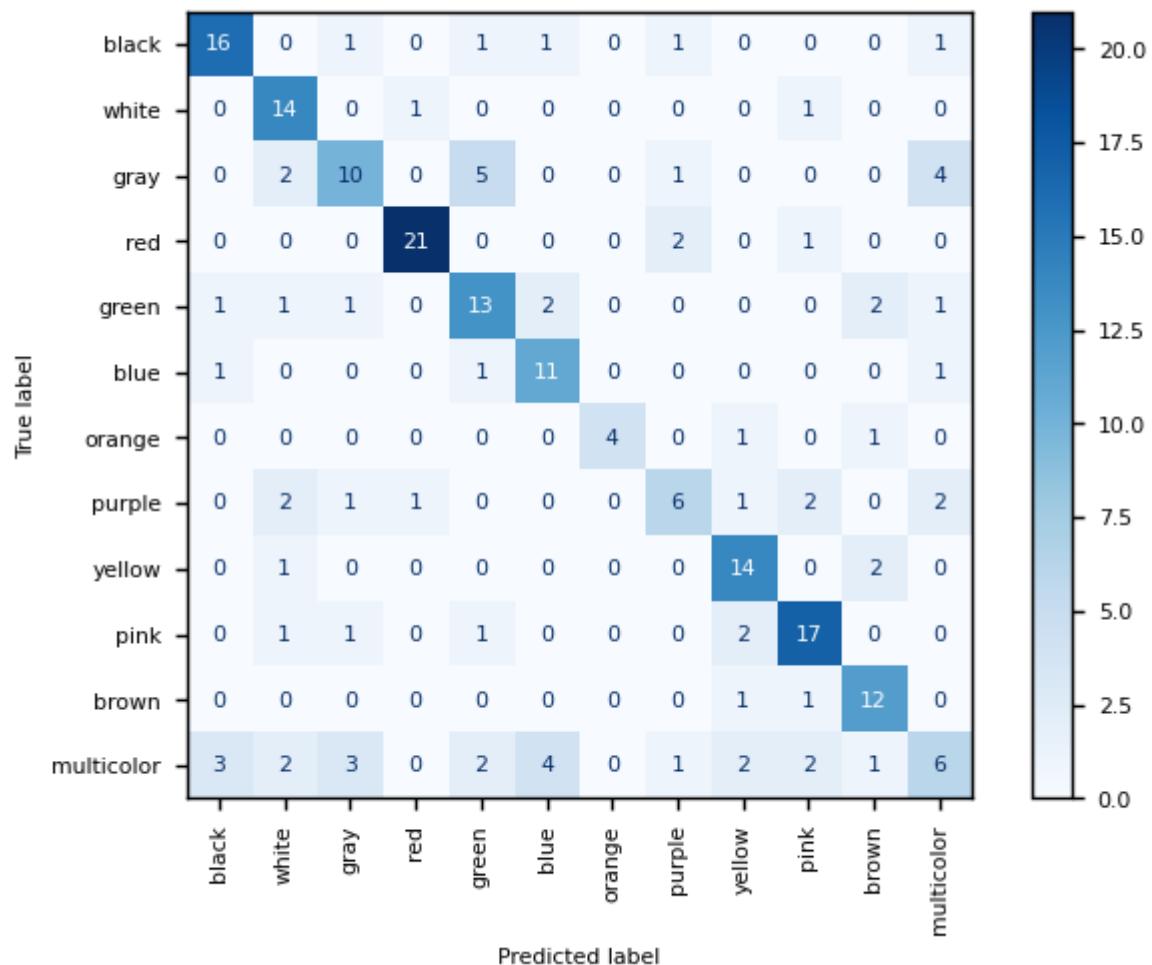
	Precision	Recall	F1 Score	Support
black	1.0000	1.0000	1.0000	61
white	1.0000	1.0000	1.0000	59
gray	1.0000	1.0000	1.0000	61
red	1.0000	1.0000	1.0000	55
green	1.0000	1.0000	1.0000	61
blue	1.0000	1.0000	1.0000	72
orange	1.0000	1.0000	1.0000	15
purple	1.0000	1.0000	1.0000	38
yellow	1.0000	1.0000	1.0000	61
pink	1.0000	1.0000	1.0000	55
brown	1.0000	1.0000	1.0000	66
multicolor	1.0000	1.0000	1.0000	52
accuracy			1.0000	
macro avg	1.0000	1.0000	1.0000	656
weighted avg	1.0000	1.0000	1.0000	656



Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7619	0.7619	0.7619	21
white	0.6087	0.8750	0.7179	16
gray	0.5882	0.4545	0.5128	22
red	0.9130	0.8750	0.8936	24
green	0.5652	0.6190	0.5909	21
blue	0.6111	0.7857	0.6875	14
orange	1.0000	0.6667	0.8000	6
purple	0.5455	0.4000	0.4615	15
yellow	0.6667	0.8235	0.7368	17
pink	0.7083	0.7727	0.7391	22
brown	0.6667	0.8571	0.7500	14
multicolor	0.4000	0.2308	0.2927	26
accuracy			0.6606	
macro avg	0.6696	0.6768	0.6621	218
weighted avg	0.6507	0.6606	0.6462	218

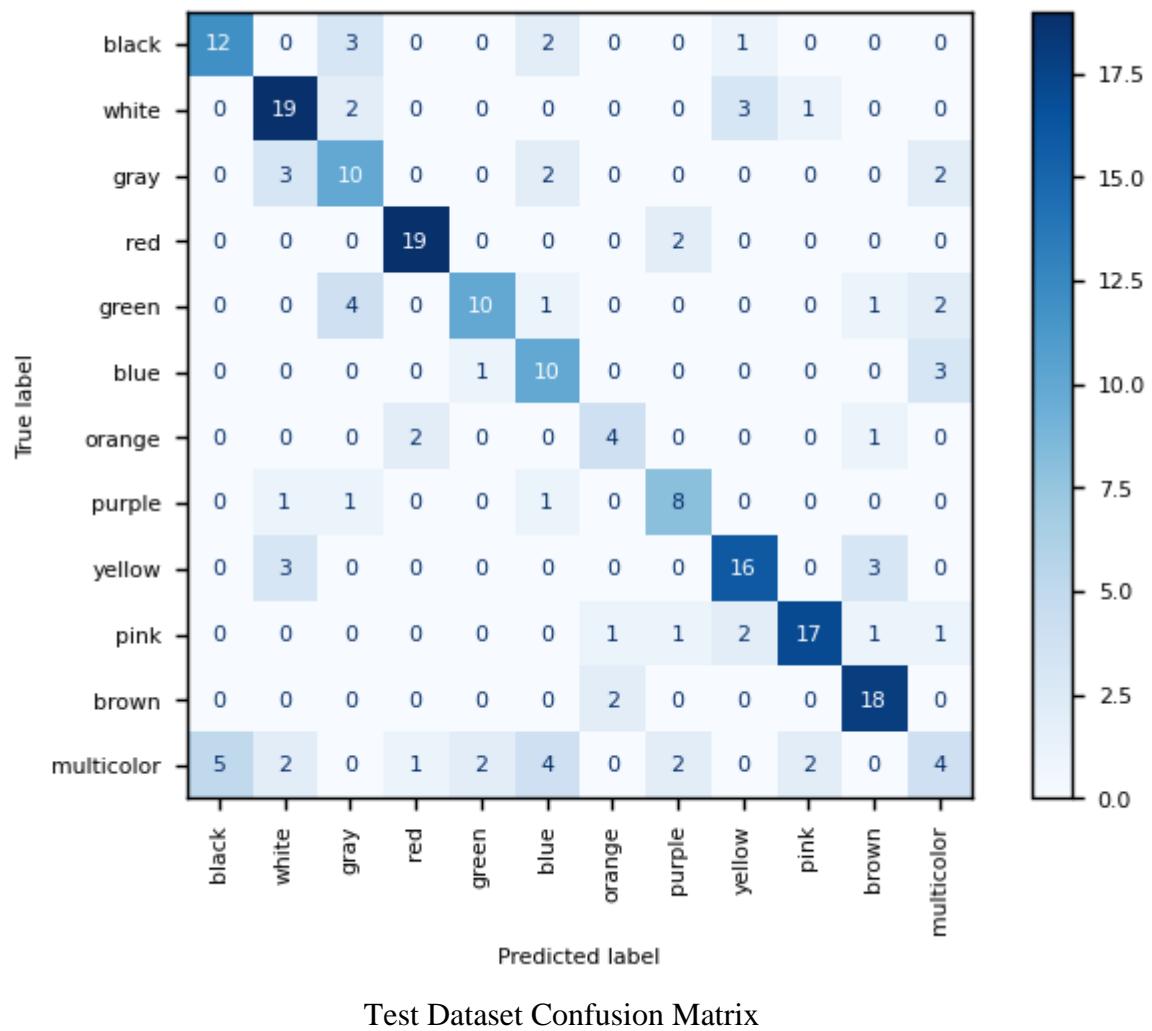


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7059	0.6667	0.6857	18
white	0.6786	0.7600	0.7170	25
gray	0.5000	0.5882	0.5405	17
red	0.8636	0.9048	0.8837	21
green	0.7692	0.5556	0.6452	18
blue	0.5000	0.7143	0.5882	14
orange	0.5714	0.5714	0.5714	7
purple	0.6154	0.7273	0.6667	11
yellow	0.7273	0.7273	0.7273	22
pink	0.8500	0.7391	0.7907	23
brown	0.7500	0.9000	0.8182	20
multicolor	0.3333	0.1818	0.2353	22
accuracy			0.6743	
macro avg	0.6554	0.6697	0.6558	218
weighted avg	0.6688	0.6743	0.6648	218



black (100.0%) black (100.0%) pink (100.0%) brown (100.0%) orange (86.5%) multicolor (100.0%)
 >> black >> black >> pink >> brown >> brown >> green



black (100.0%) white (99.6%) gray (100.0%) brown (100.0%) white (100.0%) red (100.0%)
 >> black >> white >> gray >> brown >> white >> red >> multicolor



white (99.6%) green (71.1%) white (99.3%) purple (87.5%) blue (80.0%) white (90.1%)
 >> white >> green >> white >> purple >> multicolor >> blue >> white



pink (98.3%) black (77.2%) gray (46.2%) brown (100.0%) white (100.0%) green (86.9%)
 >> pink >> multicolor >> gray >> brown >> white >> green >> multicolor



pink (97.4%) purple (100.0%) red (99.9%) white (99.9%) pink (100.0%) pink (100.0%)
 >> multicolor >> red >> red >> white >> pink >> pink



multicolor (96.5%) black (99.4%) orange (99.8%) blue (100.0%) white (100.0%) blue (97.1%)
 >> multicolor >> black >> orange >> pink >> blue >> white >> blue



Sample Images from Test Dataset

pink (100.0%) green (86.9%) pink (88.7%)
 >> multicolor >> multicolor >> white
 orange (67.4%) multicolor (98.5%) blue (80.0%)
 >> brown >> blue >> gray



blue (83.4%) red (100.0%) black (94.1%) multicolor (98.7%) gray (100.0%) pink (97.4%)
 >> multicolor >> multicolor >> multicolor >> gray >> black >> multicolor



black (72.2%) white (82.9%) gray (96.6%) purple (93.2%) multicolor (74.2%) white (59.1%)
 >> multicolor >> gray >> black >> pink >> pink >> multicolor



multicolor (100.0%) beige (99.8%) multicolor (61.8%) grey (99.8%) blue (96.7%) green (66.5%)
 >> green >> pink >> gray >> green >> multicolor >> blue



black (100.0%) white (100.0%) black (77.2%) multicolor (83.5%) orange (86.5%) multicolor (98.9%)
 >> multicolor >> gray >> multicolor >> green >> brown >> blue



white (92.5%) yellow (57.6%) gray (99.9%) yellow (99.2%) blue (93.6%) brown (54.5%)
 >> multicolor >> pink >> white >> black >> gray >> green



Sample Images that cannot be accurately predicted in the Test Dataset

4.1.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Quantization Fastoctree	3.06 sec.	0.2683
Quantization Mmcq	47.41 sec.	0.2445
Quantization Libimagequant	19.40 sec.	0.2280
Pairwise Argmin	0.55 sec.	0.1885
Pairwise Mean	0.35 sec.	0.2409
Array Mean Argmin	0.39 sec.	0.0988
Pairwise First Quantile	2.12 sec.	0.2873
Pairwise Median	2.06 sec.	0.2792
Pairwise Third Quantile	1.95 sec.	0.2550

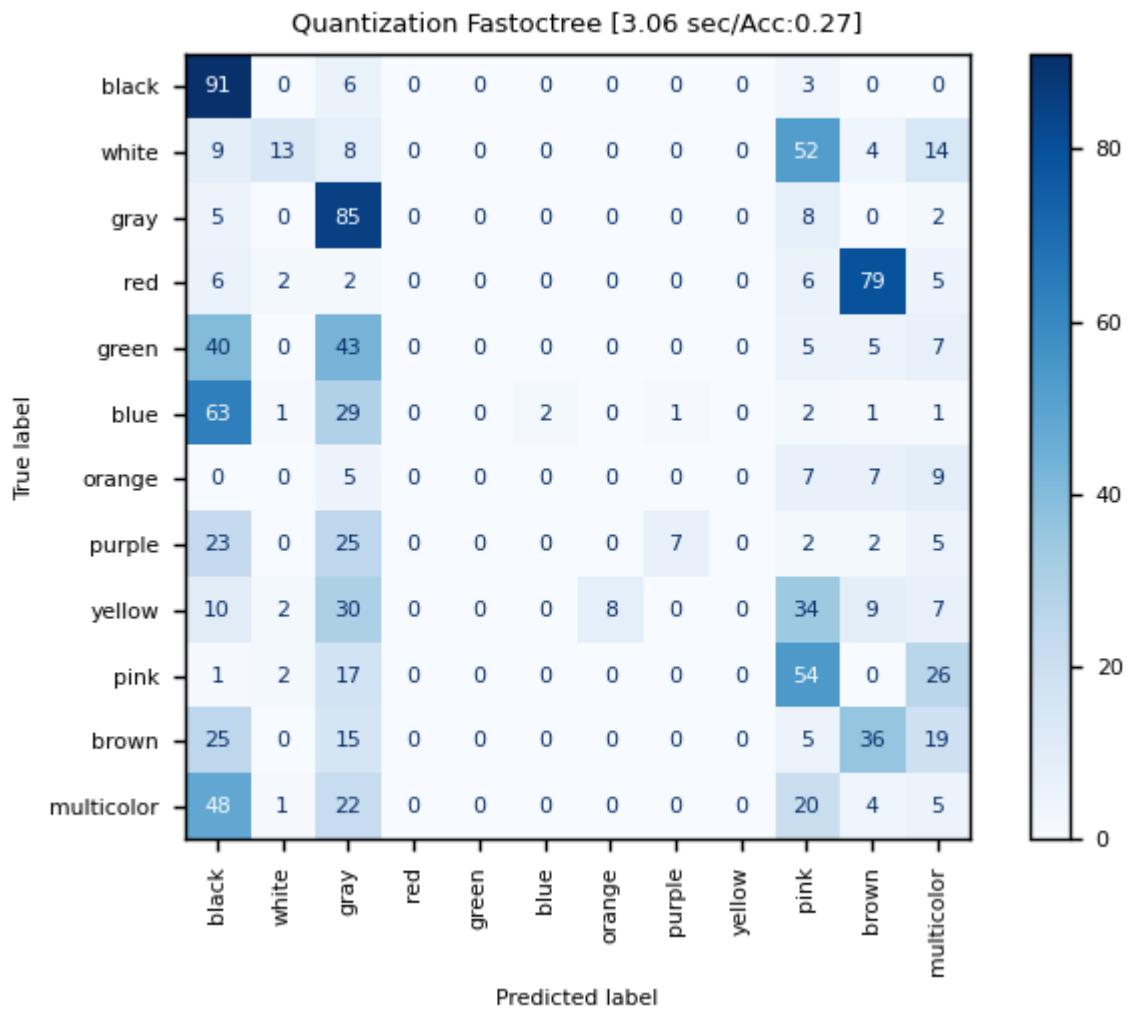
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Quantization Fastoctree Result

Total Duration: 3.06 sec.

Multi-Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.2835	0.9100	0.4323	100
white	0.6190	0.1300	0.2149	100
gray	0.2962	0.8500	0.4393	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0200	0.0392	100
orange	0.0000	0.0000	0.0000	28
purple	0.8750	0.1094	0.1944	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2727	0.5400	0.3624	100
brown	0.2449	0.3600	0.2915	100
multicolor	0.0500	0.0500	0.0500	100
accuracy			0.2683	
macro avg	0.3034	0.2474	0.1687	1092
weighted avg	0.3046	0.2683	0.1789	1092



Multi-Color with RGB / Quantization Fastoctree Confusion Matrix

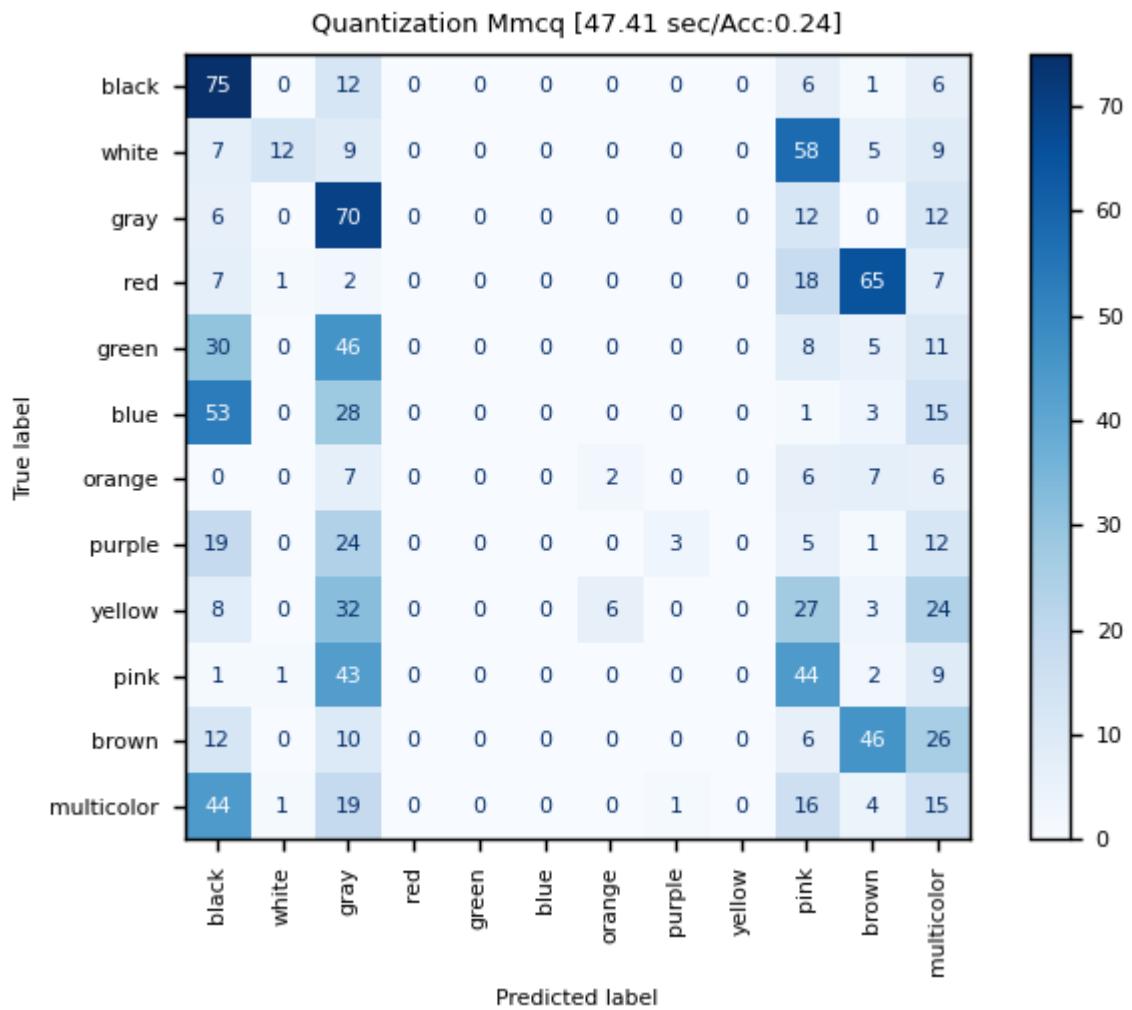
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Quantization Mmcq Result

Total Duration: 47.41 sec.

Multi-Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.2863	0.7500	0.4144	100
white	0.8000	0.1200	0.2087	100
gray	0.2318	0.7000	0.3483	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.2500	0.0714	0.1111	28
purple	0.7500	0.0469	0.0882	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2126	0.4400	0.2866	100
brown	0.3239	0.4600	0.3802	100
multicolor	0.0987	0.1500	0.1190	100
accuracy			0.2445	
macro avg	0.2461	0.2282	0.1630	1092
weighted avg	0.2292	0.2445	0.1689	1092



Multi-Color with RGB / Quantization Mmcq Confusion Matrix

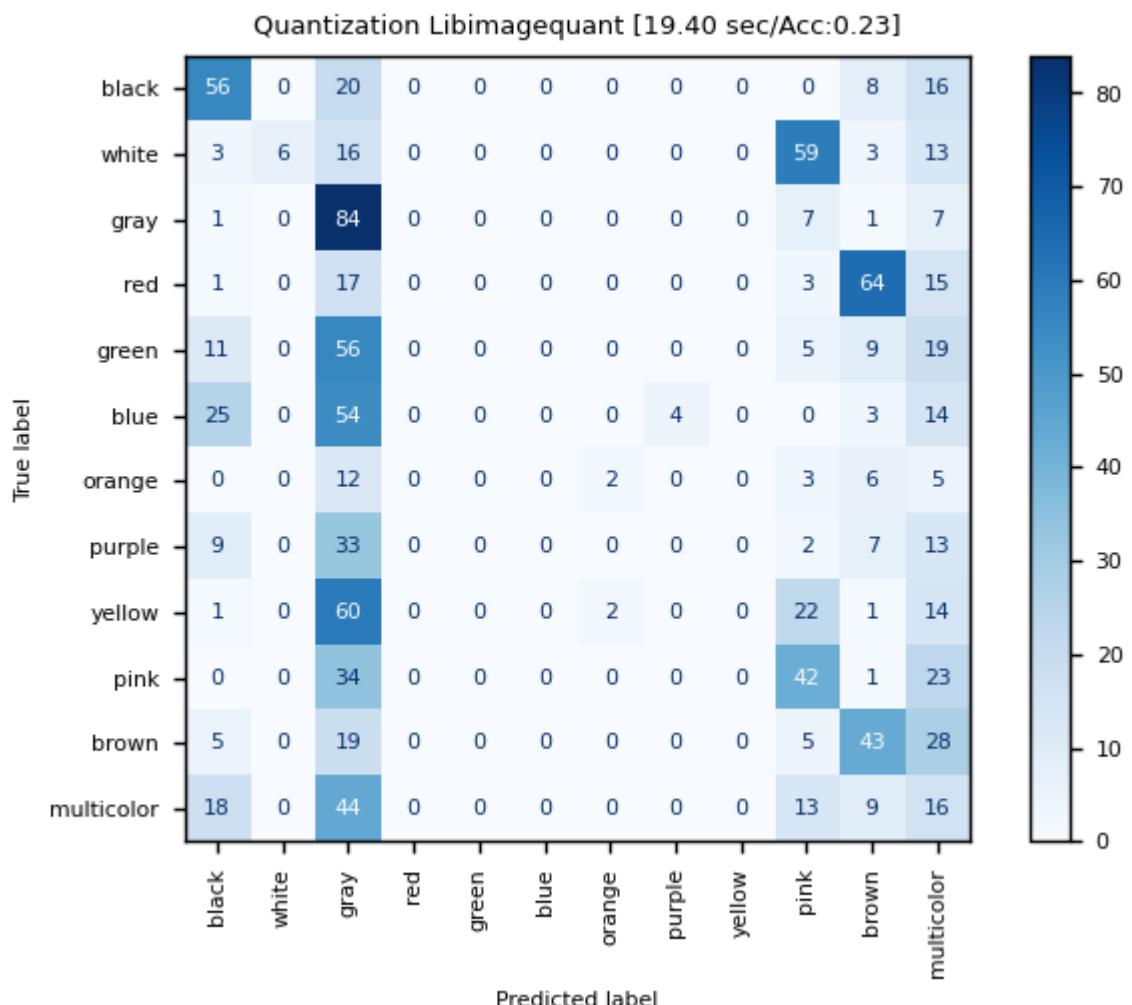
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Quantization Libimagequant Result

Total Duration: 19.40 sec.

Multi-Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.4308	0.5600	0.4870	100
white	1.0000	0.0600	0.1132	100
gray	0.1871	0.8400	0.3060	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.5000	0.0714	0.1250	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2609	0.4200	0.3218	100
brown	0.2774	0.4300	0.3373	100
multicolor	0.0874	0.1600	0.1131	100
accuracy			0.2280	
macro avg	0.2286	0.2118	0.1503	1092
weighted avg	0.2183	0.2280	0.1569	1092



Multi-Color with RGB / Quantization Libimagequant Confusion Matrix

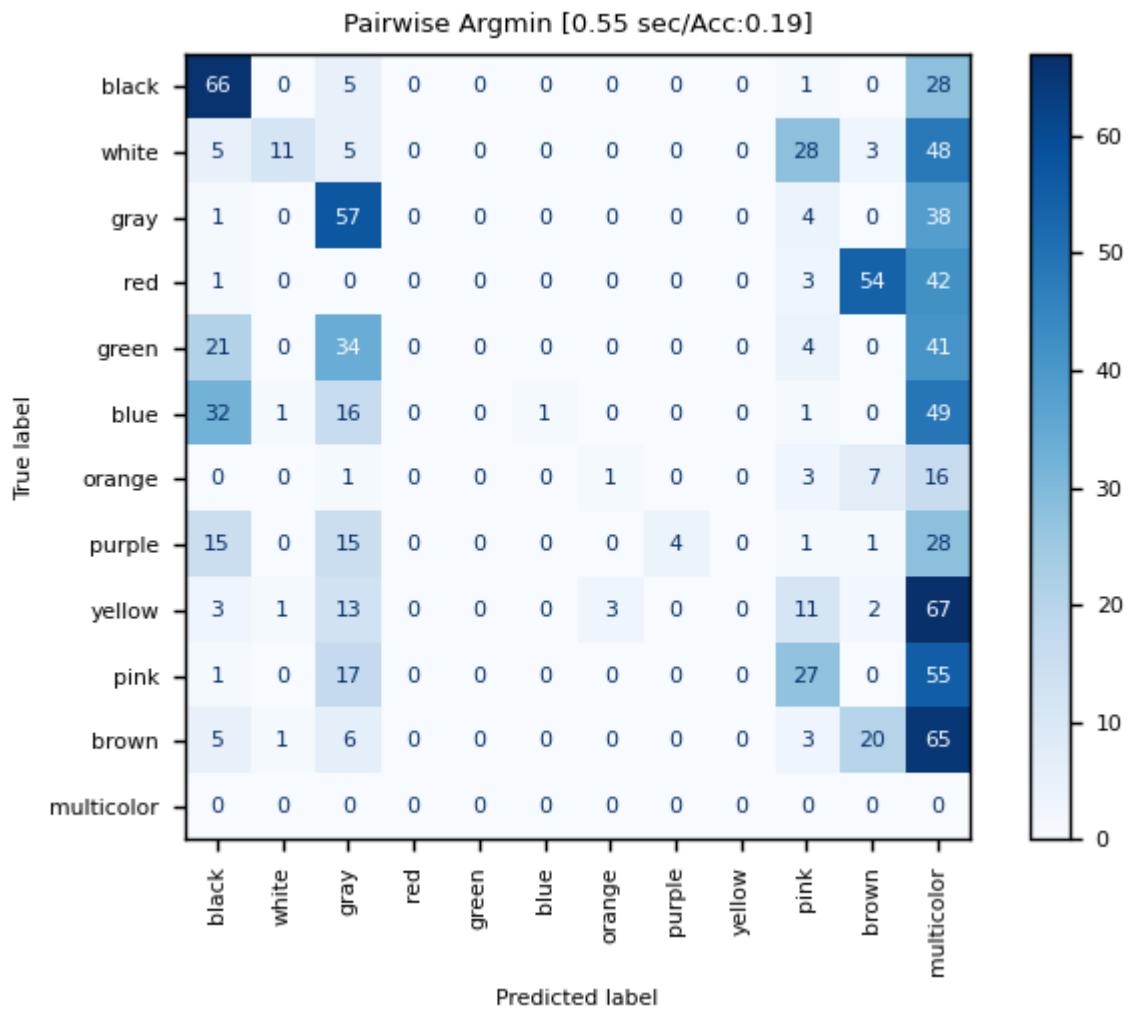
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Pairwise Argmin Result

Total Duration: 0.55 sec.

Multi-Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.4400	0.6600	0.5280	100
white	0.7857	0.1100	0.1930	100
gray	0.3373	0.5700	0.4238	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	1.0000	0.0100	0.0198	100
orange	0.2500	0.0357	0.0625	28
purple	1.0000	0.0625	0.1176	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3140	0.2700	0.2903	100
brown	0.2299	0.2000	0.2139	100
multicolor	0.0000	0.0000	0.0000	0
accuracy			0.1885	
macro avg	0.3631	0.1599	0.1541	992
weighted avg	0.3848	0.1885	0.1776	992



Multi-Color with RGB / Pairwise Argmin Confusion Matrix

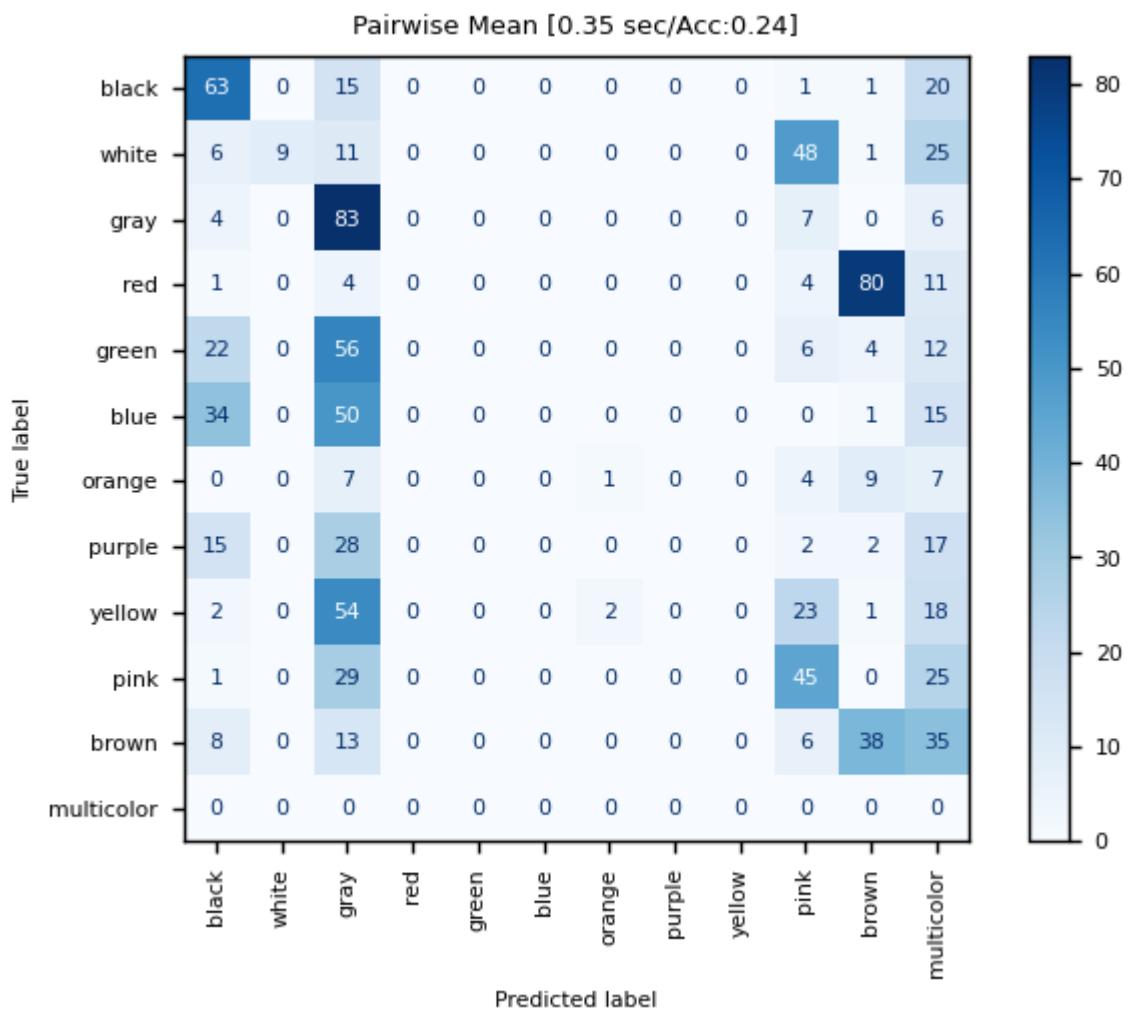
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Pairwise Mean Result

Total Duration: 0.35 sec.

Multi-Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.4038	0.6300	0.4922	100
white	1.0000	0.0900	0.1651	100
gray	0.2371	0.8300	0.3689	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.3333	0.0357	0.0645	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3082	0.4500	0.3659	100
brown	0.2774	0.3800	0.3207	100
multicolor	0.0000	0.0000	0.0000	0
accuracy			0.2409	
macro avg	0.2133	0.2013	0.1481	992
weighted avg	0.2339	0.2409	0.1745	992



Multi-Color with RGB / Pairwise Mean Confusion Matrix

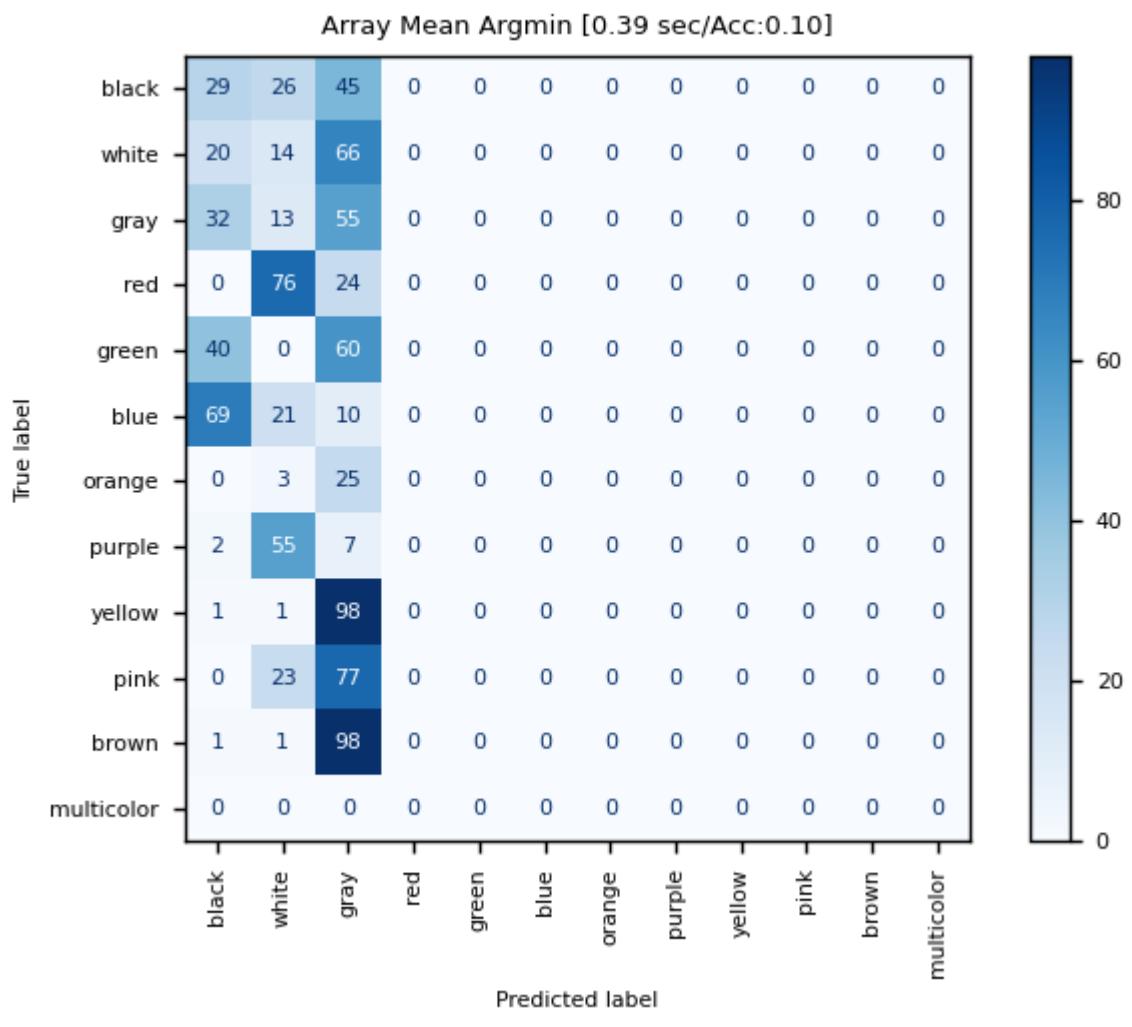
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Array Mean Argmin Result

Total Duration: 0.39 sec.

Multi-Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.1495	0.2900	0.1973	100
white	0.0601	0.1400	0.0841	100
gray	0.0973	0.5500	0.1654	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.0000	0.0000	0.0000	100
brown	0.0000	0.0000	0.0000	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.0988	0.0988	0.0988	992
macro avg	0.0256	0.0817	0.0372	992
weighted avg	0.0309	0.0988	0.0450	992



Multi-Color with RGB / Array Mean Argmin Confusion Matrix

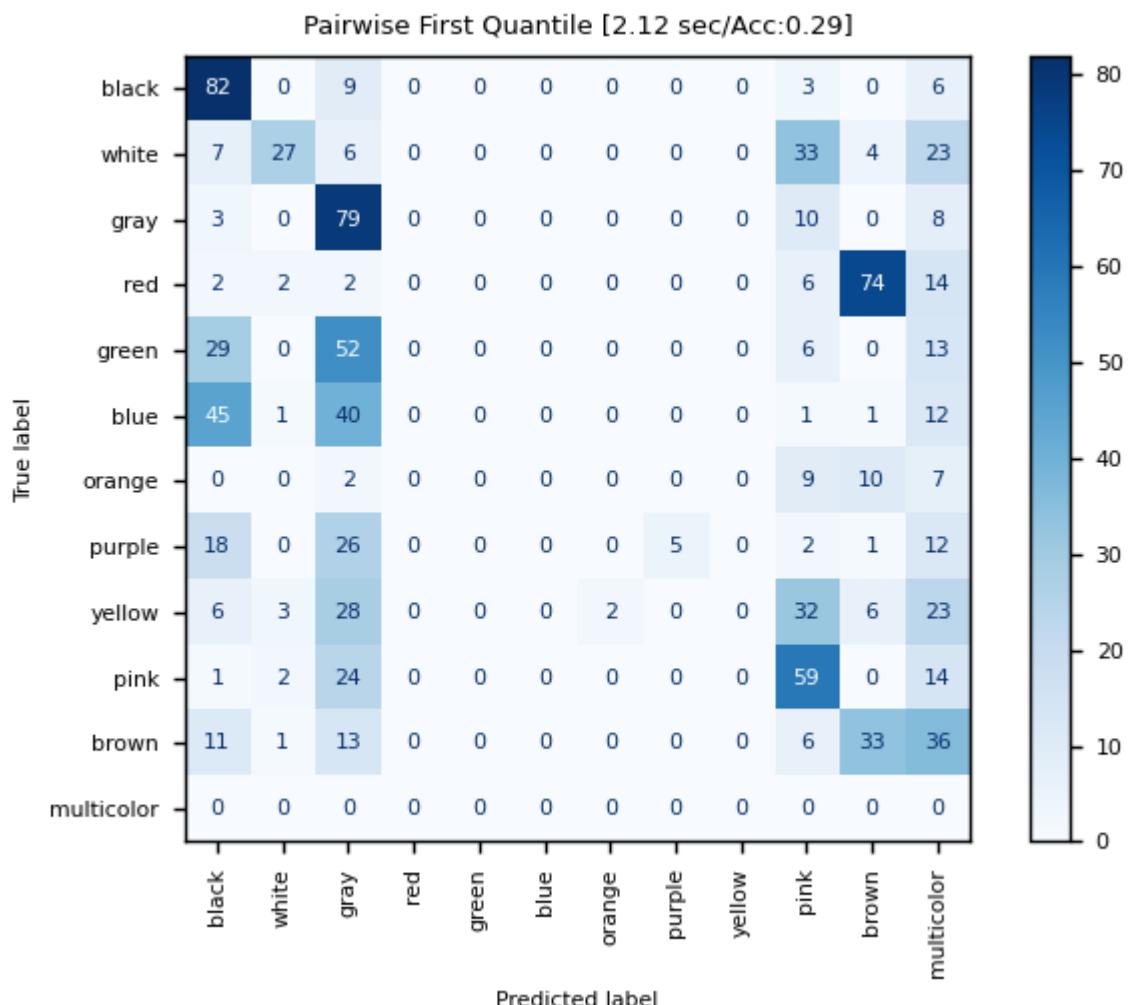
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Pairwise First Quantile Result

Total Duration: 2.12 sec.

Multi-Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4020	0.8200	0.5395	100
white	0.7500	0.2700	0.3971	100
gray	0.2811	0.7900	0.4147	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0781	0.1449	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3533	0.5900	0.4419	100
brown	0.2558	0.3300	0.2882	100
multicolor	0.0000	0.0000	0.0000	0
accuracy			0.2873	
macro avg	0.2535	0.2398	0.1855	992
weighted avg	0.2704	0.2873	0.2192	992



Multi-Color with RGB / Pairwise First Quantile Confusion Matrix

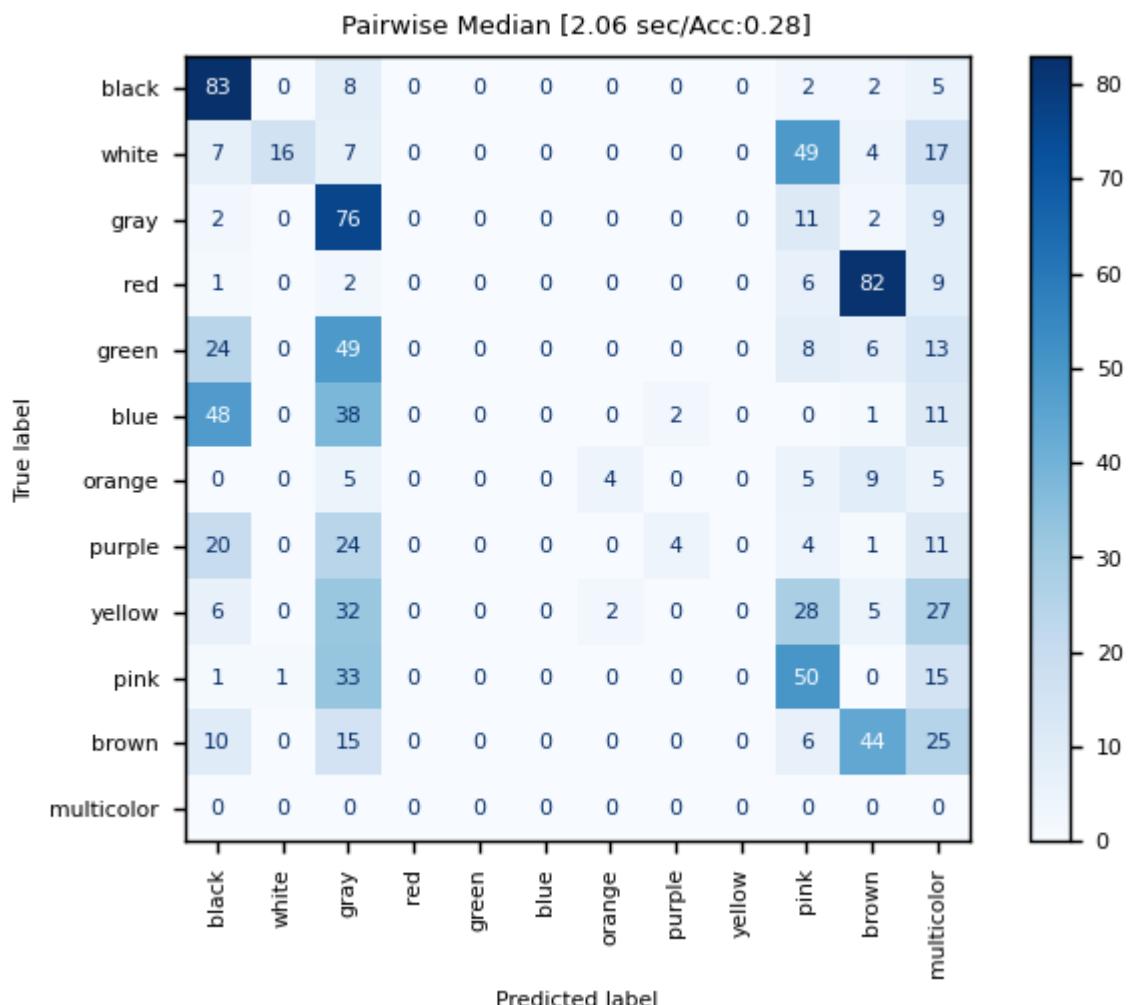
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Pairwise Median Result

Total Duration: 2.06 sec.

Multi-Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4109	0.8300	0.5497	100
white	0.9412	0.1600	0.2735	100
gray	0.2630	0.7600	0.3907	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.6667	0.1429	0.2353	28
purple	0.6667	0.0625	0.1143	64
yellow	0.0000	0.0000	0.0000	100
pink	0.2959	0.5000	0.3717	100
brown	0.2821	0.4400	0.3438	100
multicolor	0.0000	0.0000	0.0000	0
accuracy			0.2792	
macro avg	0.2939	0.2413	0.1899	992
weighted avg	0.2829	0.2792	0.2085	992



Multi-Color with RGB / Pairwise Median Confusion Matrix

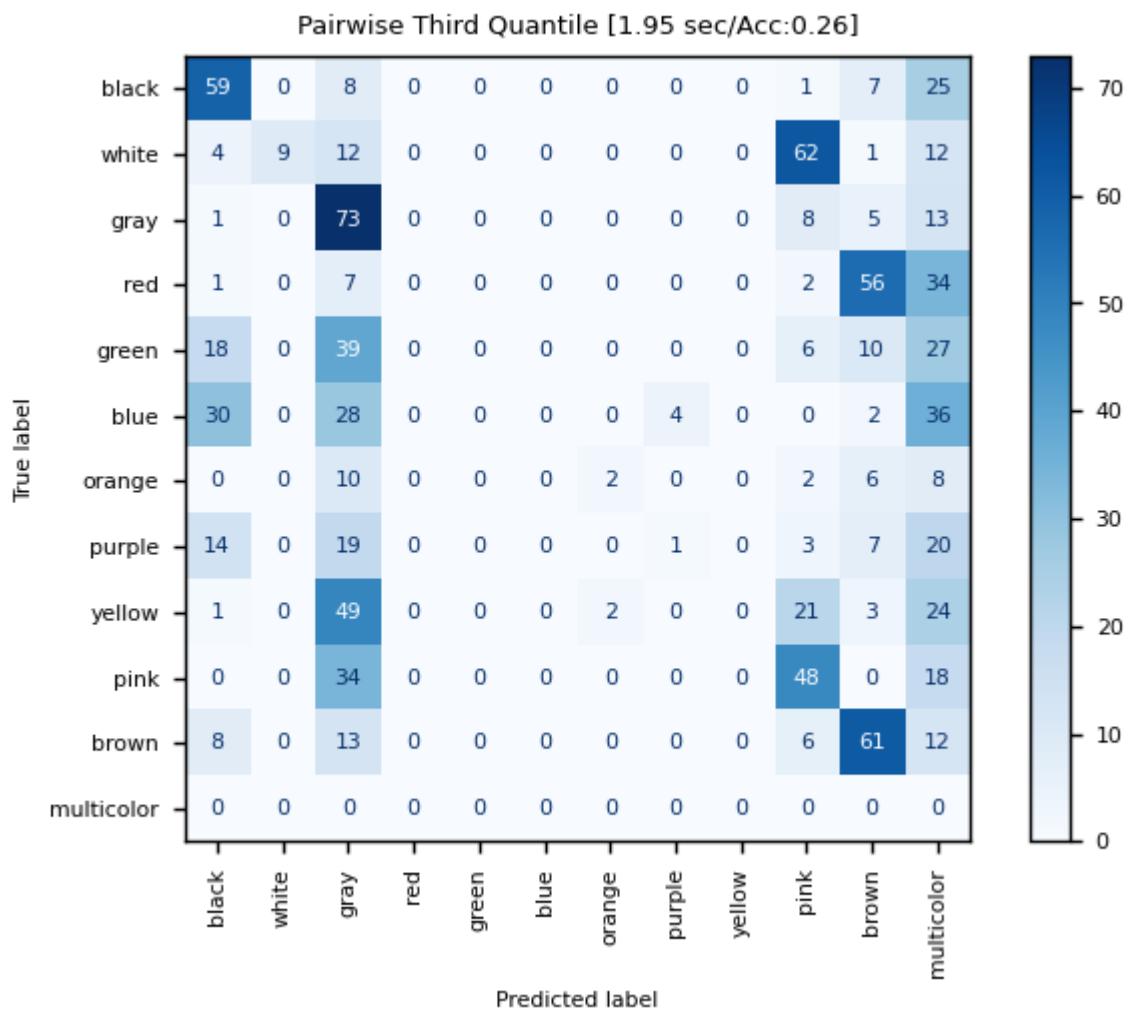
Manually Validated Dataset Improvement Run

Multi-Color with RGB / Pairwise Third Quantile Result

Total Duration: 1.95 sec.

Multi-Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4338	0.5900	0.5000	100
white	1.0000	0.0900	0.1651	100
gray	0.2500	0.7300	0.3724	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.5000	0.0714	0.1250	28
purple	0.2000	0.0156	0.0290	64
yellow	0.0000	0.0000	0.0000	100
pink	0.3019	0.4800	0.3707	100
brown	0.3861	0.6100	0.4729	100
multicolor	0.0000	0.0000	0.0000	0
accuracy			0.2550	
macro avg	0.2560	0.2156	0.1696	992
weighted avg	0.2661	0.2550	0.1950	992



Multi-Color with RGB / Pairwise Third Quantile Confusion Matrix

4.2. HSL Color Space

4.2.1. CNN AlexNet with Single Color

Manually Validated Dataset Improvement Run

CNN AlexNet Single Color with HSL

Total Duration: 678.59 sec.

Learning Rate: 0.000050

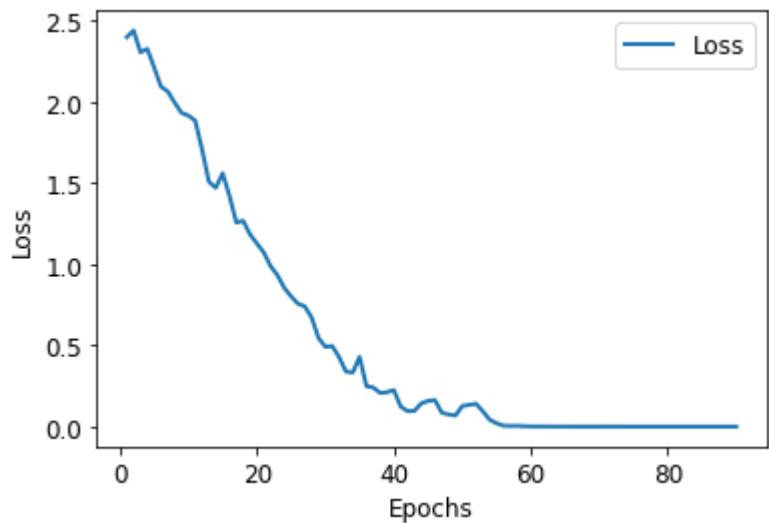
Dropout: 0.05

Batch Size: 16

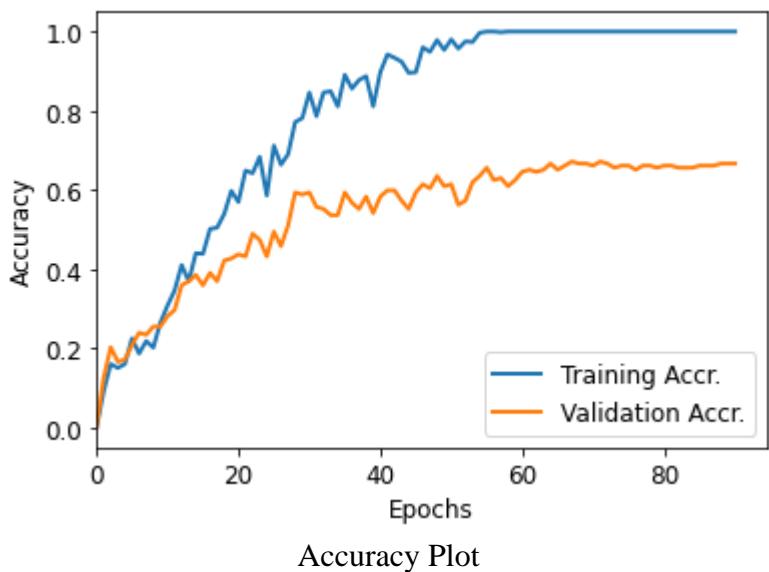
Epoch: 90

Best Performing Epoch: 71

Summary	
Method	Accuracy
Training Dataset	1.0000
Validation Dataset	0.6719
Test Dataset	0.6888



Loss Plot

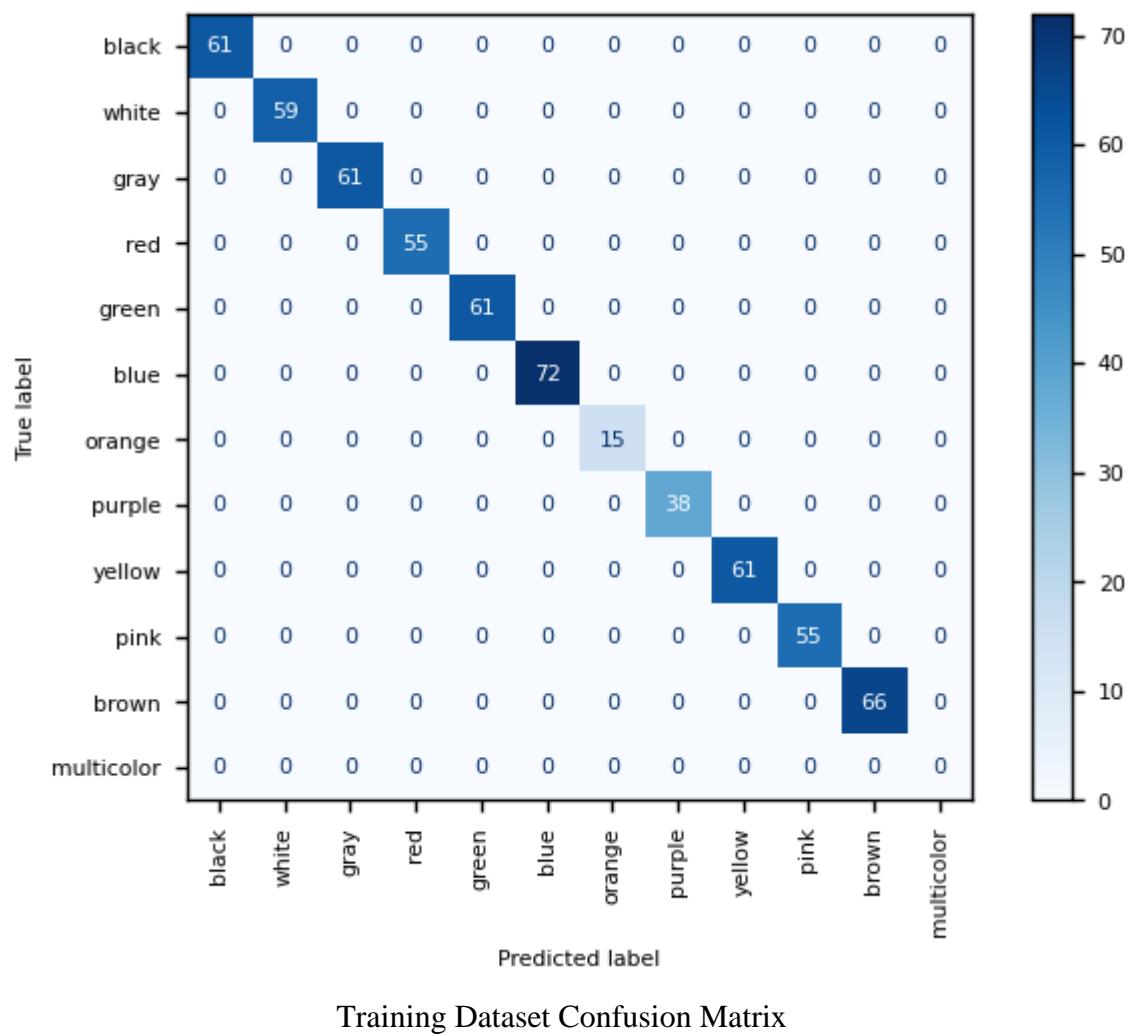


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	1.0000	1.0000	1.0000	61
white	1.0000	1.0000	1.0000	59
gray	1.0000	1.0000	1.0000	61
red	1.0000	1.0000	1.0000	55
green	1.0000	1.0000	1.0000	61
blue	1.0000	1.0000	1.0000	72
orange	1.0000	1.0000	1.0000	15
purple	1.0000	1.0000	1.0000	38
yellow	1.0000	1.0000	1.0000	61
pink	1.0000	1.0000	1.0000	55
brown	1.0000	1.0000	1.0000	66
multicolor	0.0000	0.0000	0.0000	0
micro avg	1.0000	1.0000	1.0000	604
macro avg	0.9167	0.9167	0.9167	604
weighted avg	1.0000	1.0000	1.0000	604

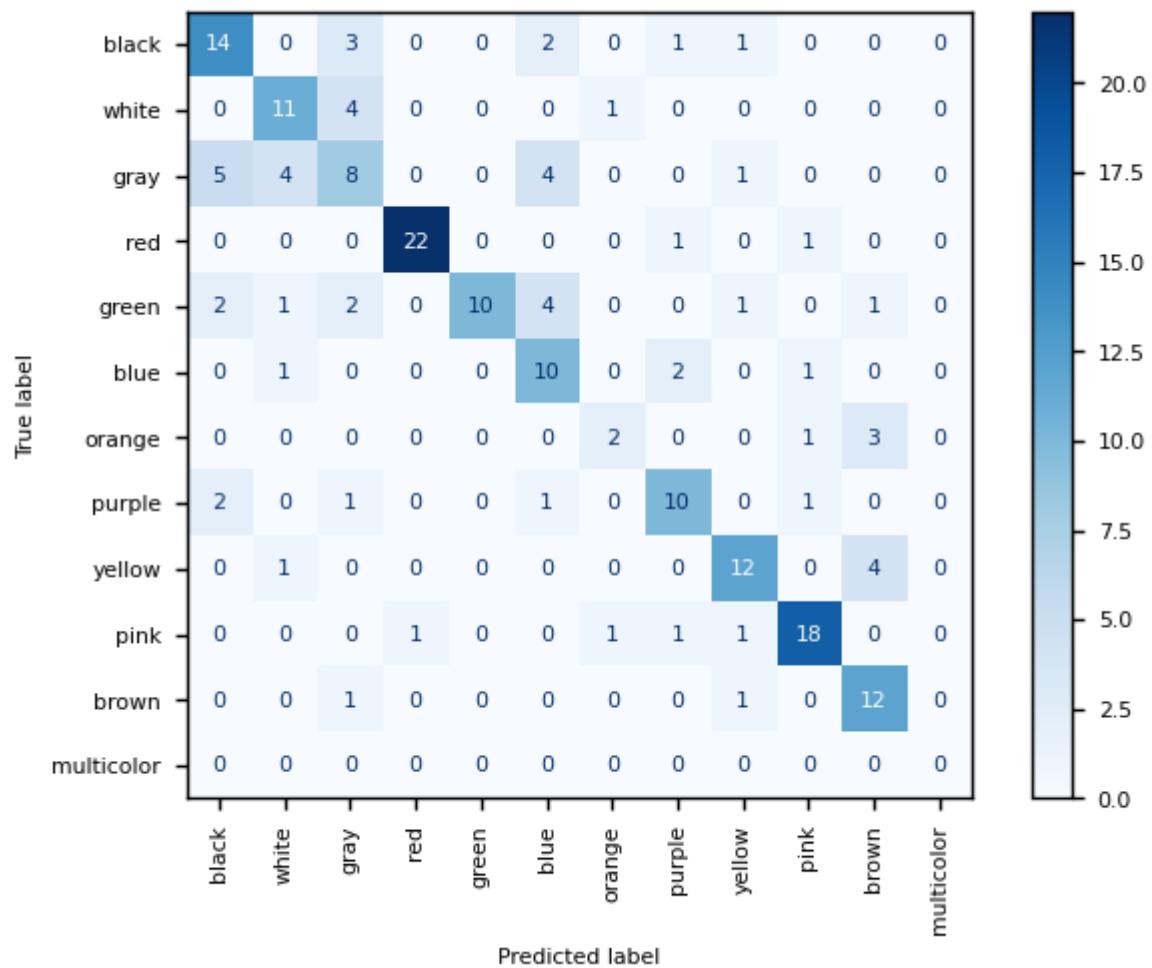


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.6087	0.6667	0.6364	21
white	0.6111	0.6875	0.6471	16
gray	0.4211	0.3636	0.3902	22
red	0.9565	0.9167	0.9362	24
green	1.0000	0.4762	0.6452	21
blue	0.4762	0.7143	0.5714	14
orange	0.5000	0.3333	0.4000	6
purple	0.6667	0.6667	0.6667	15
yellow	0.7059	0.7059	0.7059	17
pink	0.8182	0.8182	0.8182	22
brown	0.6000	0.8571	0.7059	14
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.6719	0.6719	0.6719	192
macro avg	0.6137	0.6005	0.5936	192
weighted avg	0.6971	0.6719	0.6698	192

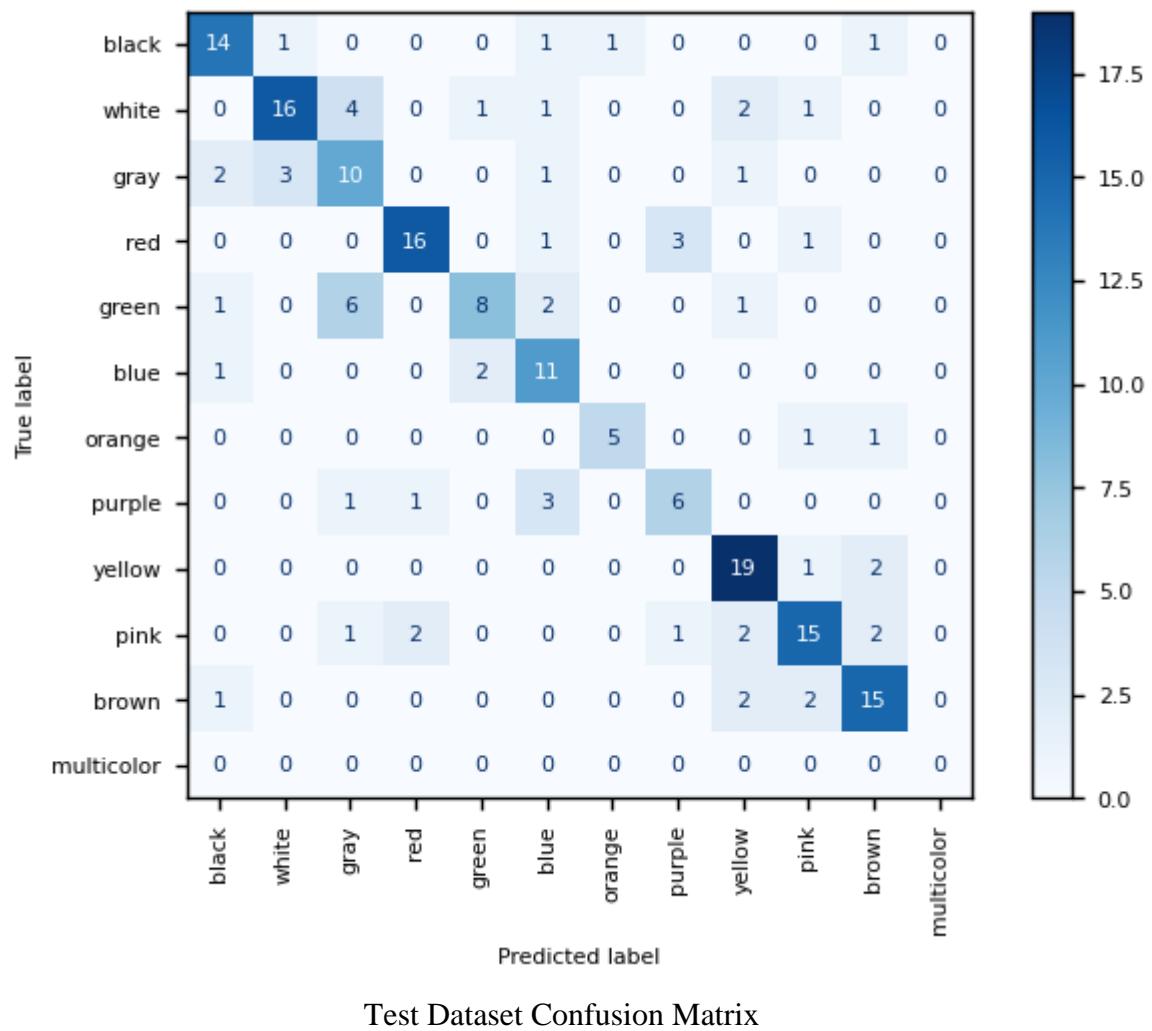


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7368	0.7778	0.7568	18
white	0.8000	0.6400	0.7111	25
gray	0.4545	0.5882	0.5128	17
red	0.8421	0.7619	0.8000	21
green	0.7273	0.4444	0.5517	18
blue	0.5500	0.7857	0.6471	14
orange	0.8333	0.7143	0.7692	7
purple	0.6000	0.5455	0.5714	11
yellow	0.7037	0.8636	0.7755	22
pink	0.7143	0.6522	0.6818	23
brown	0.7143	0.7500	0.7317	20
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.6888	0.6888	0.6888	196
macro avg	0.6397	0.6270	0.6258	196
weighted avg	0.7046	0.6888	0.6885	196



Test Dataset Confusion Matrix

blue (99.7%) blue (100.0%) black (100.0%) pink (71.4%) brown (98.1%) brown (100.0%)
 >> black >> blue >> black >> brown >> brown >> pink



black (88.0%) red (100.0%) blue (100.0%) green (100.0%) black (34.8%) pink (100.0%)
 >> black >> pink >> blue >> green >> black >> pink



gray (53.2%) red (100.0%) gray (73.0%) pink (100.0%) white (97.5%) black (100.0%)
 >> gray >> red >> white >> pink >> gray >> black



yellow (99.9%) white (93.2%) blue (100.0%) red (100.0%) red (99.4%) gray (93.4%)
 >> pink >> gray >> blue >> red >> red >> green



pink (100.0%) yellow (88.8%) gray (98.2%) pink (93.3%) brown (99.8%) brown (100.0%)
 >> pink >> brown >> gray >> pink >> brown >> brown



green (100.0%) gray (97.8%) brown (100.0%) black (100.0%) brown (100.0%) black (100.0%)
 >> green >> gray >> brown >> black >> brown >> black



Sample Images from Test Dataset

black (100.0%) brown (95.7%) pink (71.4%)
 >> brown >> black >> brown white (97.5%) gray (99.7%) gray (73.0%)
 >> gray >> pink >> white



yellow (99.9%) white (100.0%) black (99.9%) blue (64.7%) gray (99.9%) pink (99.7%)
 >> pink >> black >> gray >> green >> green >> red



red (100.0%) brown (100.0%) yellow (89.6%) green (94.8%) brown (99.7%) brown (100.0%)
 >> pink >> orange >> gray >> blue >> yellow >> pink



purple (97.8%) gray (70.2%) purple (92.5%) gray (87.4%) black (91.1%) black (100.0%)
 >> red >> green >> red >> white >> blue >> green



gray (99.9%) white (93.2%) black (100.0%) white (99.8%) blue (99.4%) orange (99.1%)
 >> green >> gray >> gray >> gray >> red >> black



gray (99.6%) gray (80.7%) purple (99.2%) yellow (100.0%) pink (100.0%) yellow (87.1%)
 >> green >> green >> red >> white >> yellow >> pink



Sample Images that cannot be accurately predicted in the Test Dataset

4.2.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Pairwise Argmin	0.64 sec.	0.3609
Pairwise Mean	0.53 sec.	0.3317
Array Mean Argmin	0.30 sec.	0.1008
Pairwise First Quantile	2.00 sec.	0.3609
Pairwise Median	2.03 sec.	0.3599
Pairwise Third Quantile	1.85 sec.	0.3024

* Quantization based methods are available only in RGB color space

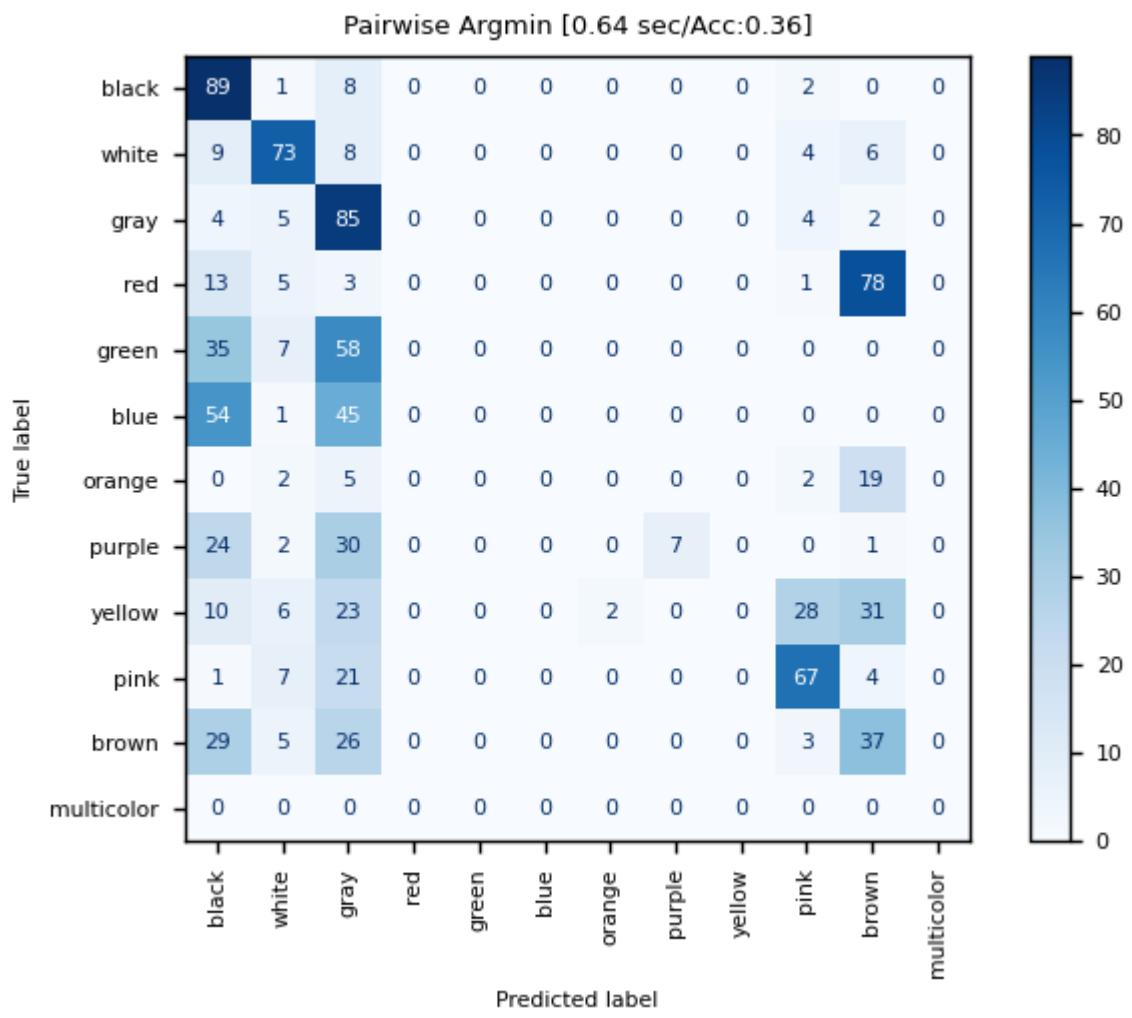
Manually Validated Dataset Improvement Run

Single Color with HSL / Pairwise Argmin Result

Total Duration: 0.64 sec.

Single Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.3321	0.8900	0.4837	100
white	0.6404	0.7300	0.6822	100
gray	0.2724	0.8500	0.4126	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.1094	0.1972	64
yellow	0.0000	0.0000	0.0000	100
pink	0.6036	0.6700	0.6351	100
brown	0.2079	0.3700	0.2662	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3609	0.3609	0.3609	992
macro avg	0.2547	0.3016	0.2231	992
weighted avg	0.2718	0.3609	0.2627	992



Single Color with HSL / Pairwise Argmin Confusion Matrix

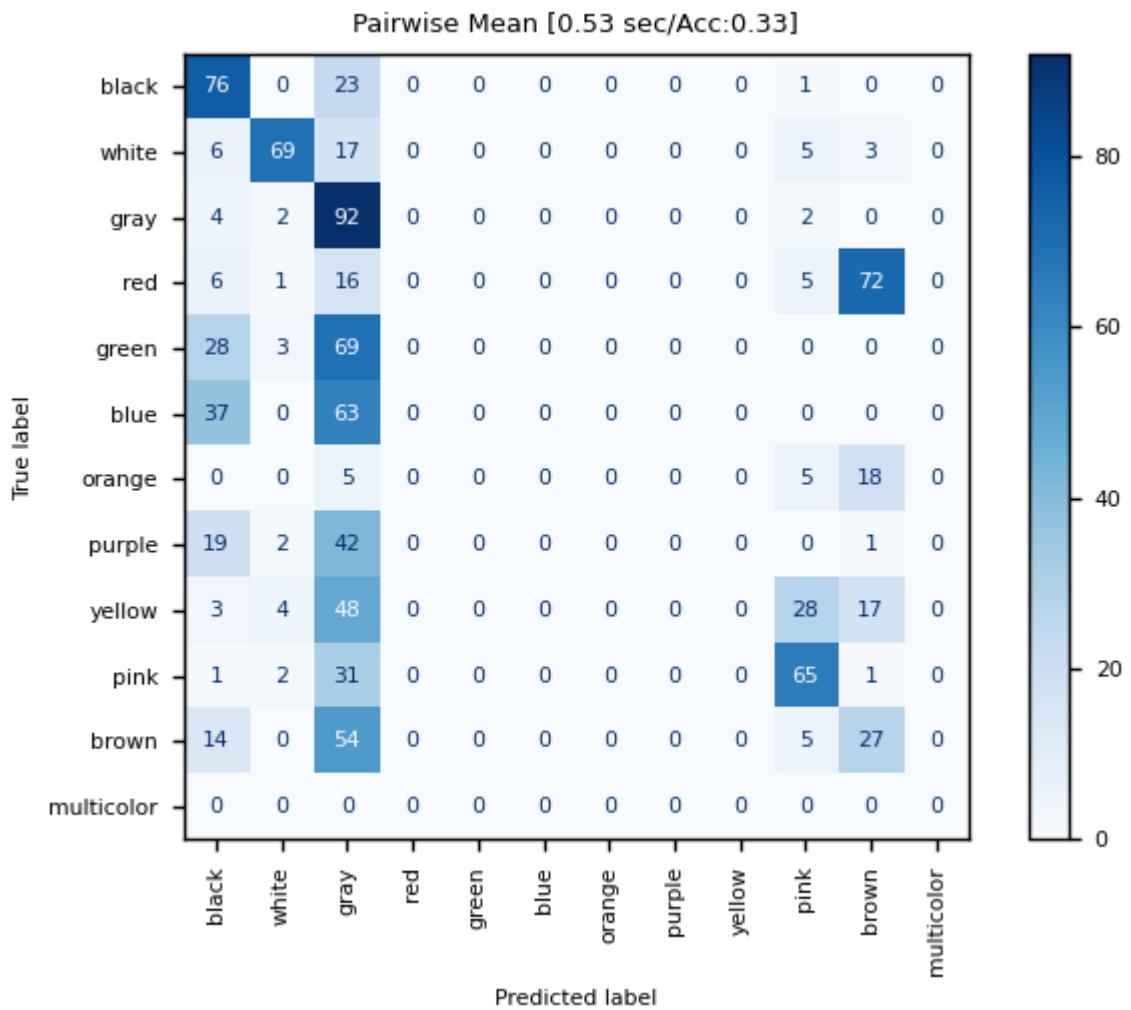
Manually Validated Dataset Improvement Run

Single Color with HSL / Pairwise Mean Result

Total Duration: 0.53 sec.

Single Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.3918	0.7600	0.5170	100
white	0.8313	0.6900	0.7541	100
gray	0.2000	0.9200	0.3286	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5603	0.6500	0.6019	100
brown	0.1942	0.2700	0.2259	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3317	0.3317	0.3317	992
macro avg	0.1815	0.2742	0.2023	992
weighted avg	0.2195	0.3317	0.2447	992



Single Color with HSL / Pairwise Mean Confusion Matrix

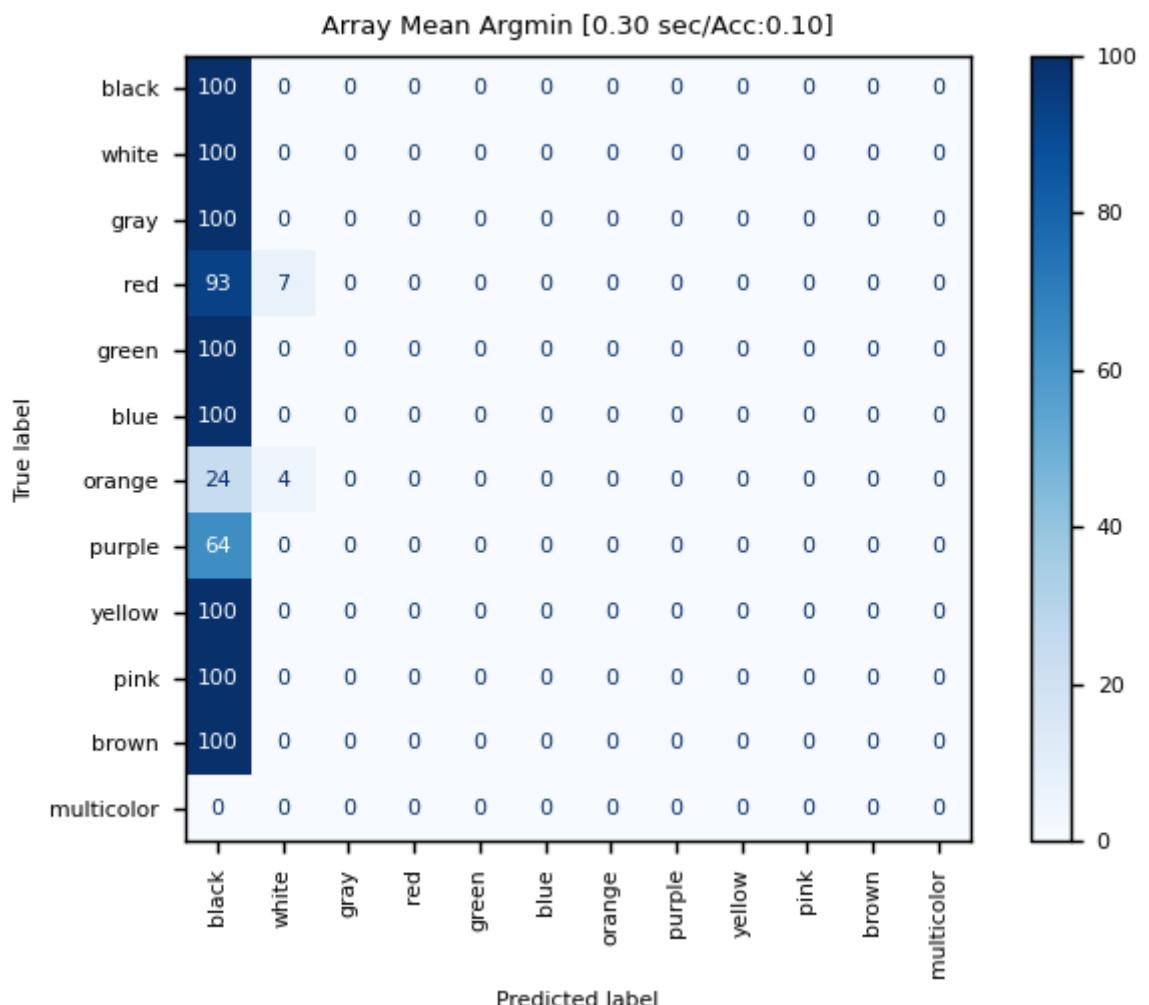
Manually Validated Dataset Improvement Run

Single Color with HSL / Array Mean Argmin Result

Total Duration: 0.30 sec.

Single Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.1019	1.0000	0.1850	100
white	0.0000	0.0000	0.0000	100
gray	0.0000	0.0000	0.0000	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.0000	0.0000	0.0000	100
brown	0.0000	0.0000	0.0000	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.1008	0.1008	0.1008	992
macro avg	0.0085	0.0833	0.0154	992
weighted avg	0.0103	0.1008	0.0187	992



Single Color with HSL / Array Mean Argmin Confusion Matrix

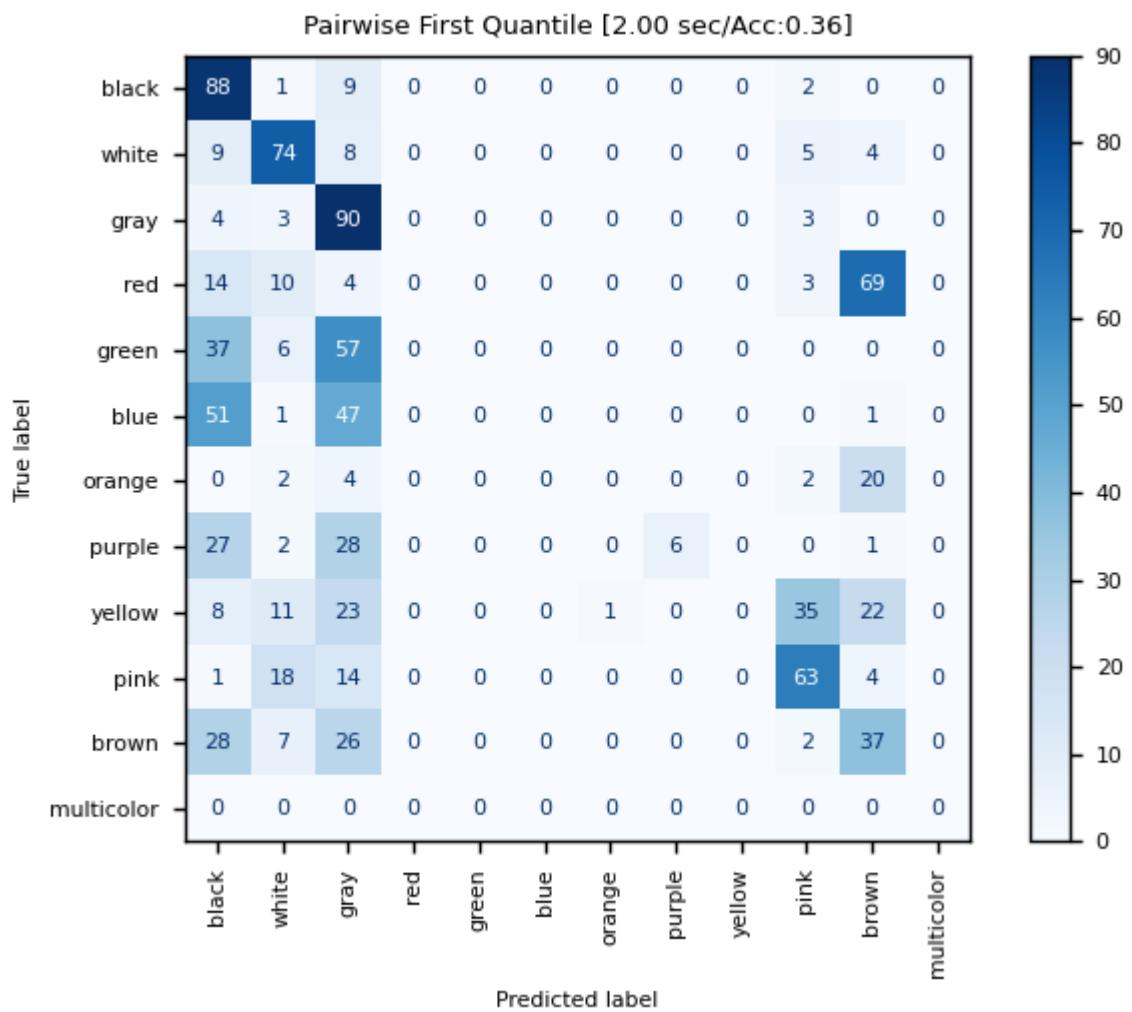
Manually Validated Dataset Improvement Run

Single Color with HSL / Pairwise First Quantile Result

Total Duration: 2.00 sec.

Single Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.3296	0.8800	0.4796	100
white	0.5481	0.7400	0.6298	100
gray	0.2903	0.9000	0.4390	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0938	0.1714	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5478	0.6300	0.5860	100
brown	0.2342	0.3700	0.2868	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3609	0.3609	0.3609	992
macro avg	0.2458	0.3011	0.2161	992
weighted avg	0.2611	0.3609	0.2551	992



Single Color with HSL / Pairwise First Quantile Confusion Matrix

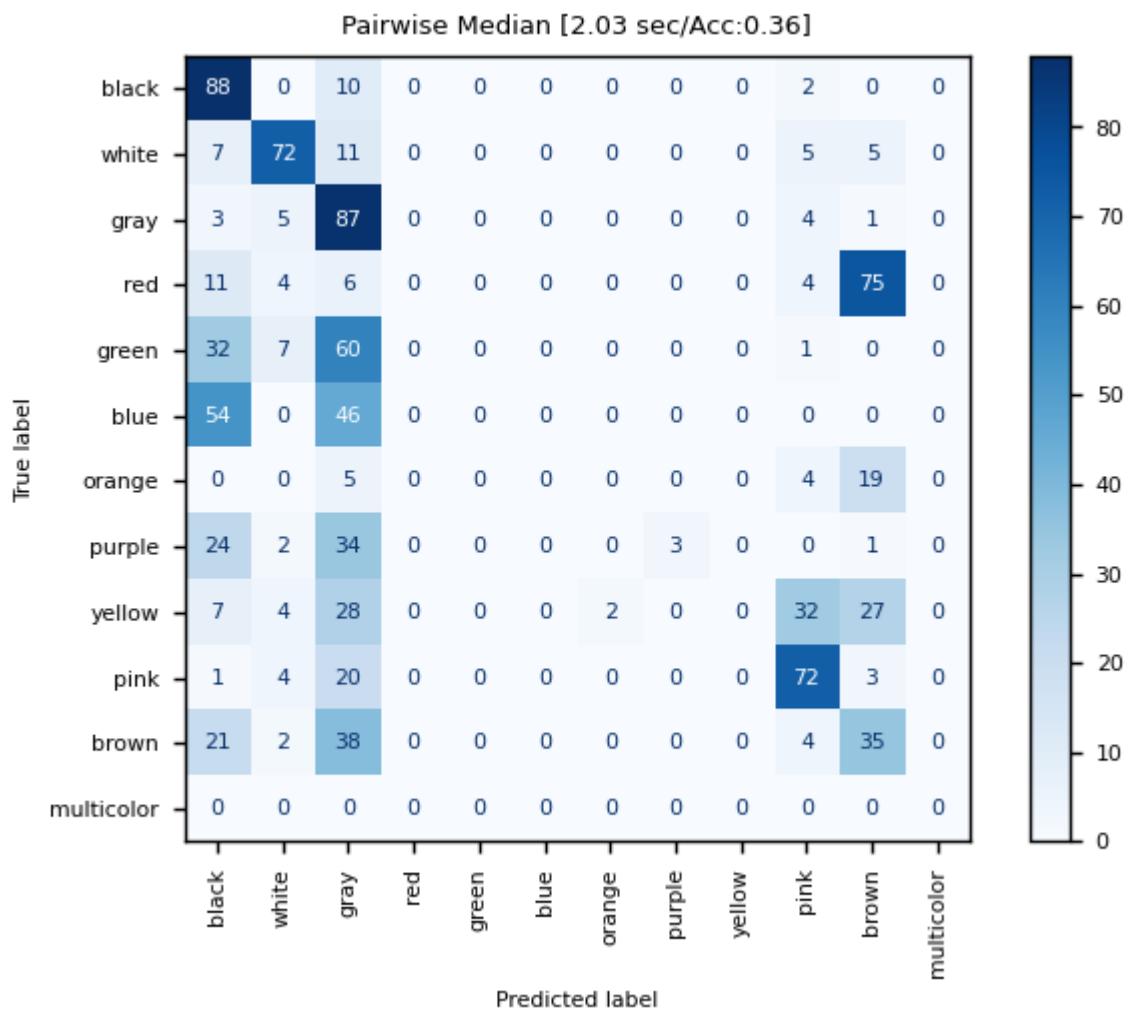
Manually Validated Dataset Improvement Run

Single Color with HSL / Pairwise Median Result

Total Duration: 2.03 sec.

Single Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.3548	0.8800	0.5057	100
white	0.7200	0.7200	0.7200	100
gray	0.2522	0.8700	0.3910	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0469	0.0896	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5625	0.7200	0.6316	100
brown	0.2108	0.3500	0.2632	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3599	0.3599	0.3599	992
macro avg	0.2584	0.2989	0.2168	992
weighted avg	0.2762	0.3599	0.2590	992



Single Color with HSL / Pairwise Median Confusion Matrix

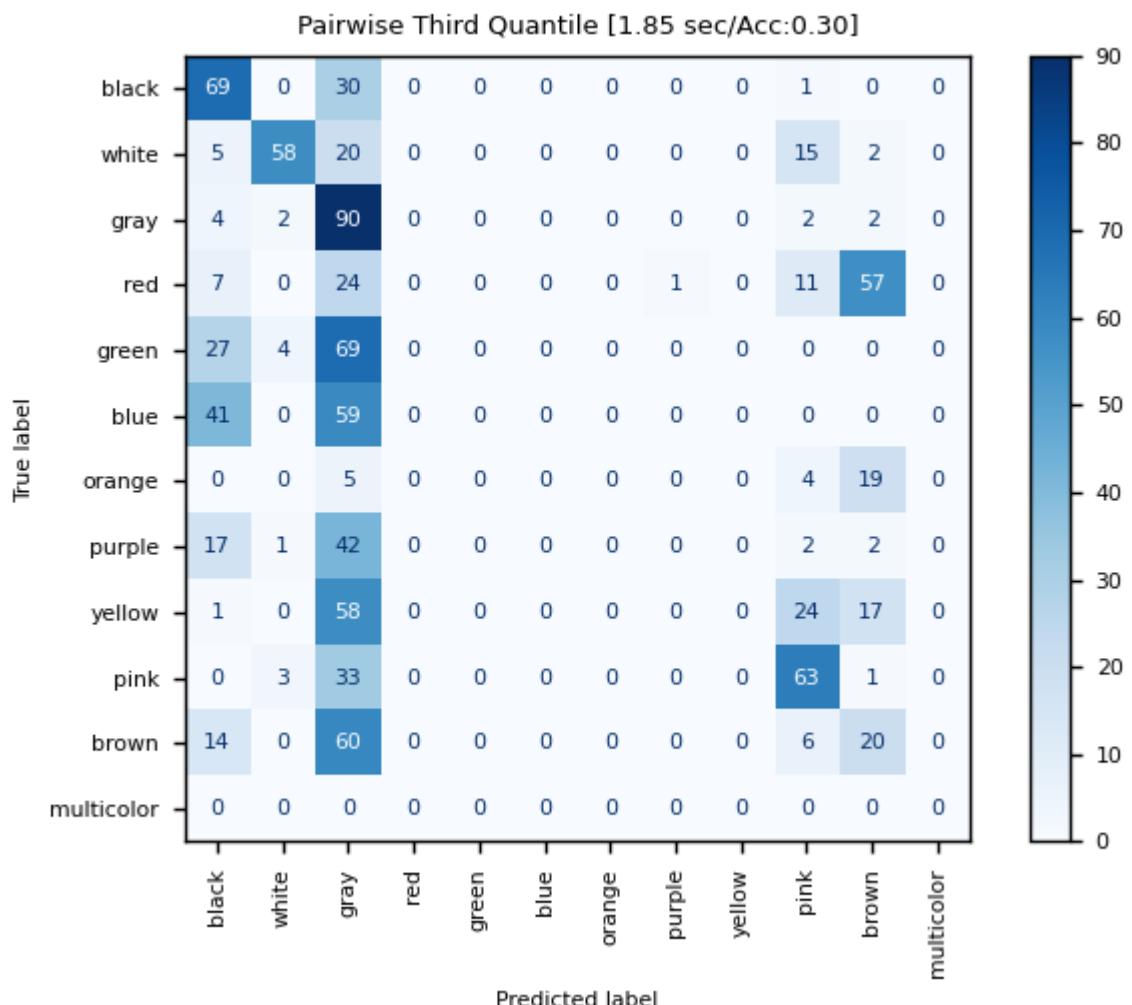
Manually Validated Dataset Improvement Run

Single Color with HSL / Pairwise Third Quantile Result

Total Duration: 1.85 sec.

Single Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.3730	0.6900	0.4842	100
white	0.8529	0.5800	0.6905	100
gray	0.1837	0.9000	0.3051	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.4922	0.6300	0.5526	100
brown	0.1667	0.2000	0.1818	100
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3024	0.3024	0.3024	992
macro avg	0.1724	0.2500	0.1845	992
weighted avg	0.2085	0.3024	0.2232	992



Single Color with HSL / Pairwise Third Quantile Confusion Matrix

4.2.3. CNN AlexNet with Multi-color

Manually Validated Dataset Improvement Run

CNN AlexNet Multi-Color with HSL

Total Duration: 727.56 sec.

Learning Rate: 0.000050

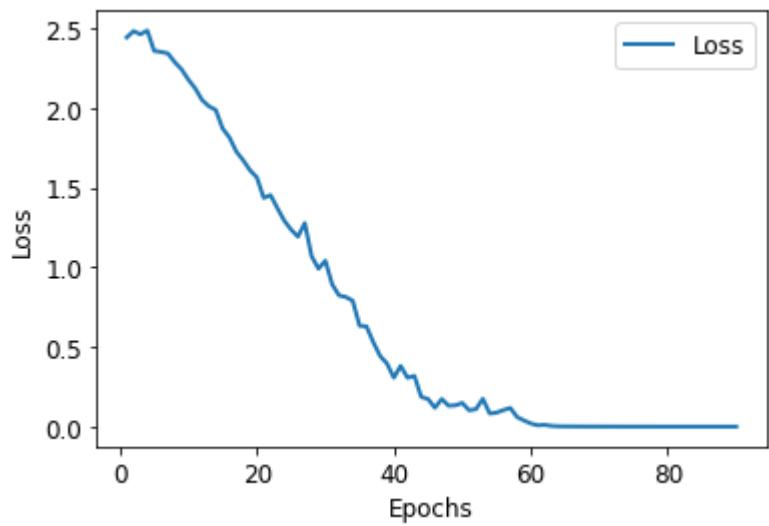
Dropout: 0.05

Batch Size: 16

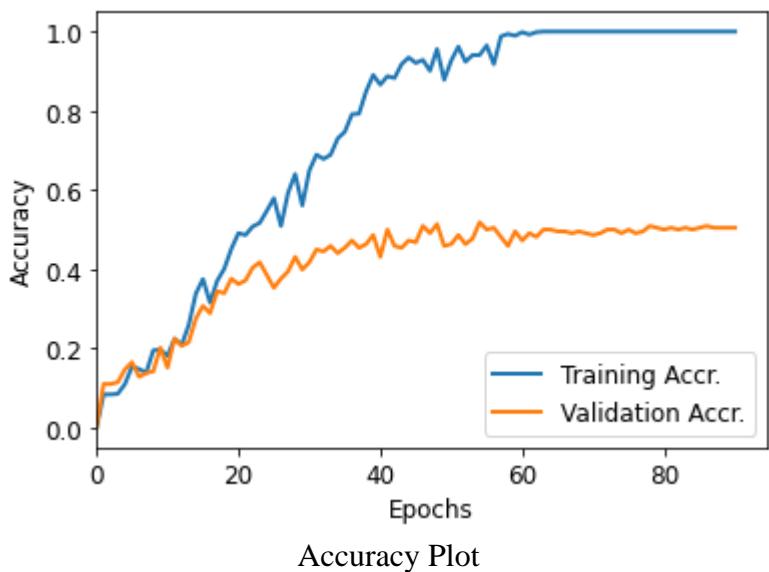
Epoch: 90

Best Performing Epoch: 54

Summary	
Method	Accuracy
Training Dataset	0.9405
Validation Dataset	0.5183
Test Dataset	0.5275



Loss Plot

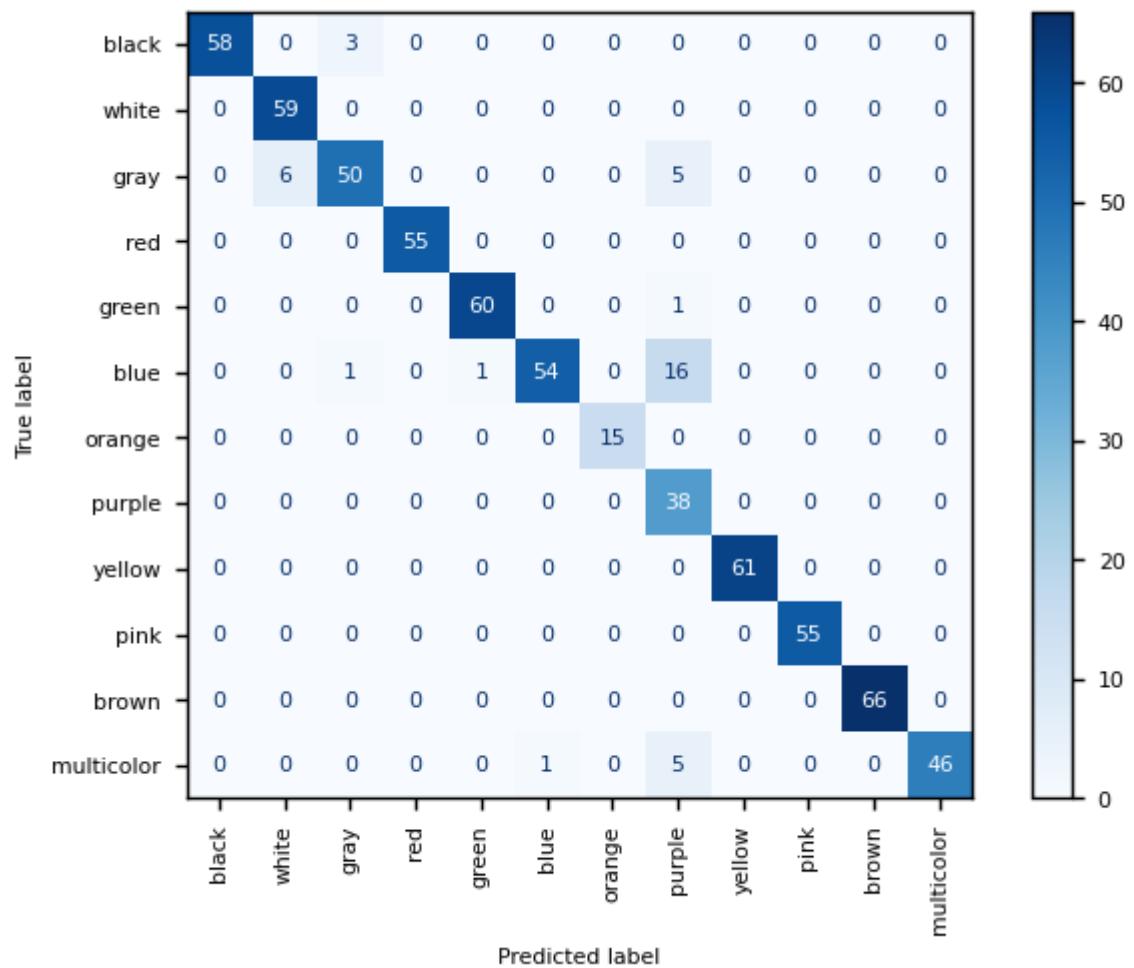


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	1.0000	0.9508	0.9748	61
white	0.9077	1.0000	0.9516	59
gray	0.9259	0.8197	0.8696	61
red	1.0000	1.0000	1.0000	55
green	0.9836	0.9836	0.9836	61
blue	0.9818	0.7500	0.8504	72
orange	1.0000	1.0000	1.0000	15
purple	0.5846	1.0000	0.7379	38
yellow	1.0000	1.0000	1.0000	61
pink	1.0000	1.0000	1.0000	55
brown	1.0000	1.0000	1.0000	66
multicolor	1.0000	0.8846	0.9388	52
accuracy			0.9405	
macro avg	0.9486	0.9491	0.9422	656
weighted avg	0.9572	0.9405	0.9432	656

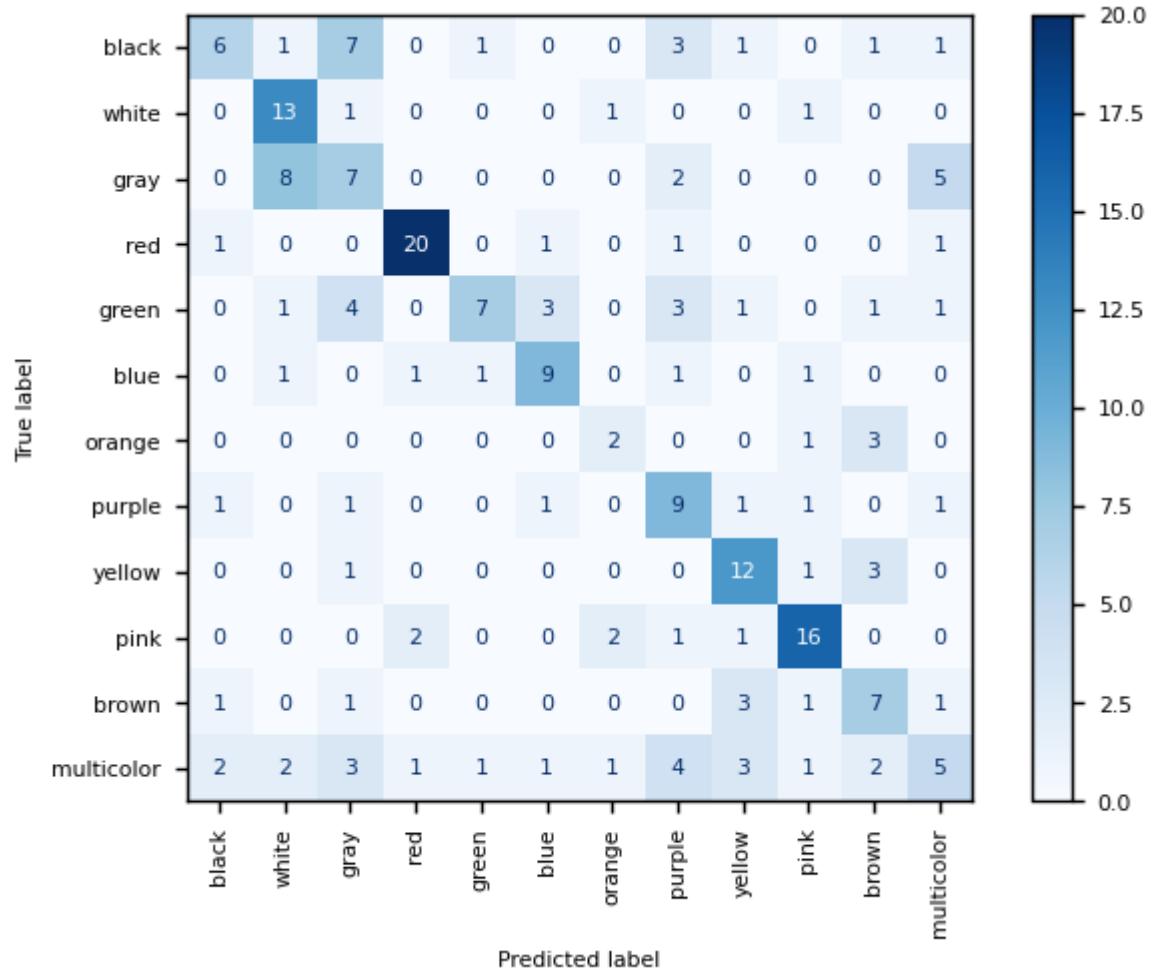


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.5455	0.2857	0.3750	21
white	0.5000	0.8125	0.6190	16
gray	0.2800	0.3182	0.2979	22
red	0.8333	0.8333	0.8333	24
green	0.7000	0.3333	0.4516	21
blue	0.6000	0.6429	0.6207	14
orange	0.3333	0.3333	0.3333	6
purple	0.3750	0.6000	0.4615	15
yellow	0.5455	0.7059	0.6154	17
pink	0.6957	0.7273	0.7111	22
brown	0.4118	0.5000	0.4516	14
multicolor	0.3333	0.1923	0.2439	26
accuracy			0.5183	
macro avg	0.5128	0.5237	0.5012	218
weighted avg	0.5291	0.5183	0.5055	218

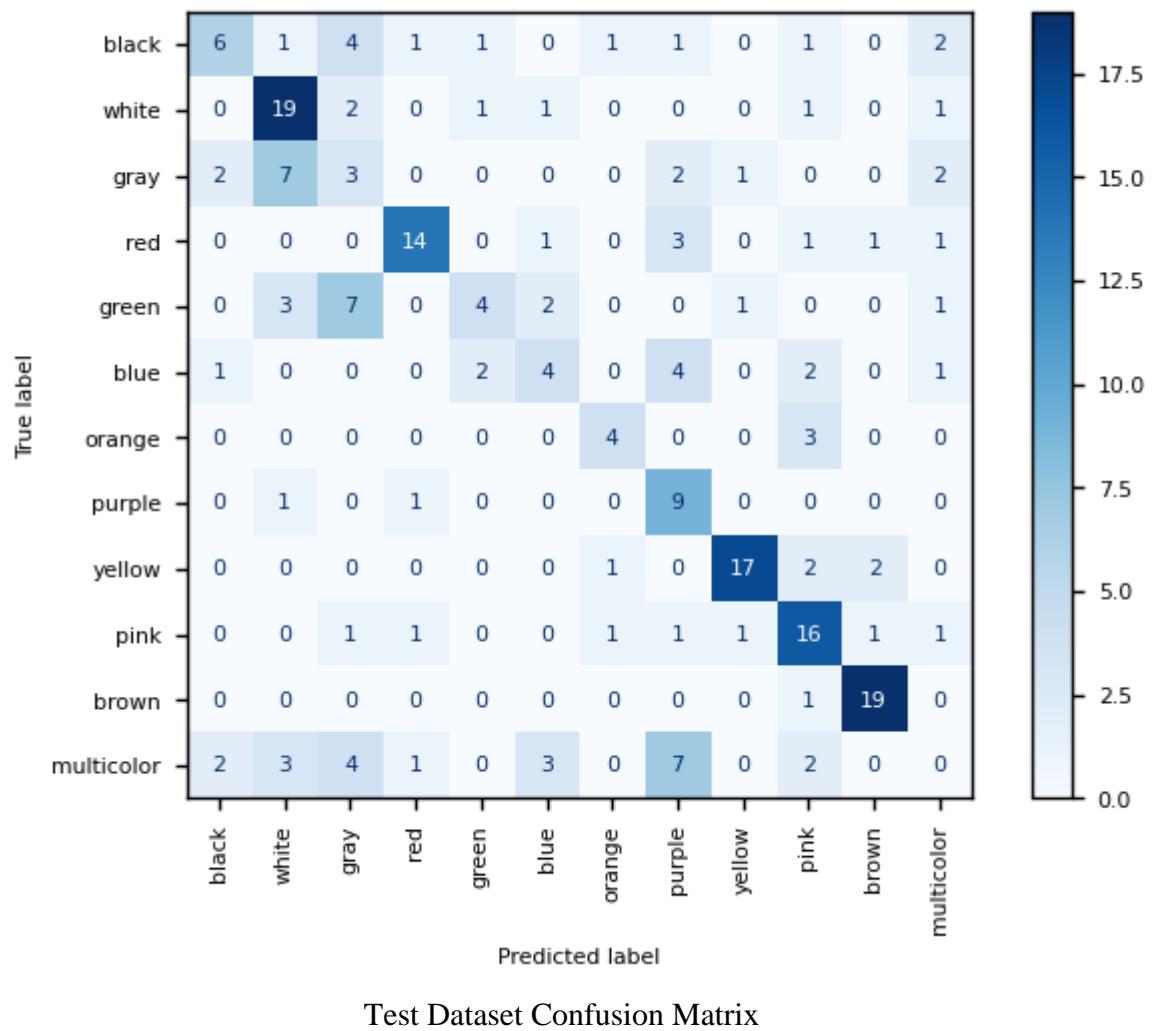


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.5455	0.3333	0.4138	18
white	0.5588	0.7600	0.6441	25
gray	0.1429	0.1765	0.1579	17
red	0.7778	0.6667	0.7179	21
green	0.5000	0.2222	0.3077	18
blue	0.3636	0.2857	0.3200	14
orange	0.5714	0.5714	0.5714	7
purple	0.3333	0.8182	0.4737	11
yellow	0.8500	0.7727	0.8095	22
pink	0.5517	0.6957	0.6154	23
brown	0.8261	0.9500	0.8837	20
multicolor	0.0000	0.0000	0.0000	22
accuracy			0.5275	
macro avg	0.5018	0.5210	0.4929	218
weighted avg	0.5148	0.5275	0.5054	218



Test Dataset Confusion Matrix

purple (83.8%) gray (88.9%) multicolor (39.8%) brown (100.0%) brown (100.0%) white (99.4%)
 >> gray >> multicolor >> pink >> brown >> brown >> white



gray (94.1%) black (100.0%) black (98.6%) multicolor (95.6%) multicolor (61.9%) gray (97.3%)
 >> multicolor >> black >> black >> gray >> black >> gray



brown (99.2%) white (99.8%) purple (93.8%) green (91.5%) purple (47.7%) brown (99.7%)
 >> brown >> white >> blue >> blue >> black >> brown



white (92.0%) black (90.3%) multicolor (46.4%) brown (100.0%) white (100.0%) brown (100.0%)
 >> white >> multicolor >> black >> brown >> white >> brown



blue (84.9%) blue (85.2%) green (64.2%) brown (98.5%) pink (99.9%) gray (66.7%)
 >> multicolor >> blue >> blue >> brown >> white >> gray



gray (93.3%) green (68.9%) black (99.9%) white (93.2%) brown (94.8%) purple (92.6%)
 >> green >> green >> black >> multicolor >> brown >> multicolor

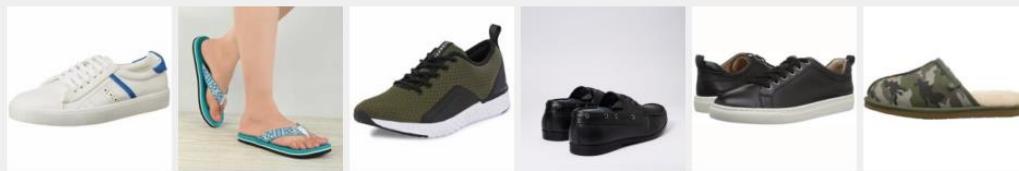


Sample Images from Test Dataset

white (93.2%) orange (65.9%) white (100.0%) purple (95.3%) gray (99.4%) gray (75.2%)
 >> multicolor >> black >> black >> multicolor >> black >> green



white (70.4%) white (98.4%) gray (93.3%) gray (94.9%) multicolor (61.9%) gray (81.5%)
 >> multicolor >> multicolor >> green >> black >> black >> green



pink (82.8%) purple (83.6%) black (90.3%) purple (51.3%) multicolor (95.6%) brown (73.9%)
 >> multicolor >> blue >> multicolor >> pink >> gray >> yellow



white (99.9%) purple (83.8%) multicolor (64.3%) purple (96.3%) purple (68.0%) purple (92.6%)
 >> gray >> gray >> blue >> pink >> blue >> multicolor



pink (86.2%) gray (79.4%) gray (88.9%) red (99.0%) pink (77.4%) pink (99.9%)
 >> blue >> green >> multicolor >> multicolor >> blue >> multicolor



black (75.7%) multicolor (51.8%) pink (99.9%) purple (87.7%) gray (94.6%) yellow (60.4%)
 >> gray >> red >> white >> gray >> multicolor >> pink



Sample Images that cannot be accurately predicted in the Test Dataset

4.2.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Pairwise Argmin	0.78 sec.	0.2546
Pairwise Mean	0.53 sec.	0.2701
Array Mean Argmin	0.46 sec.	0.0962
Pairwise First Quantile	2.36 sec.	0.3022
Pairwise Median	2.25 sec.	0.3068
Pairwise Third Quantile	2.09 sec.	0.2619

* Quantization based methods are available only in RGB color space

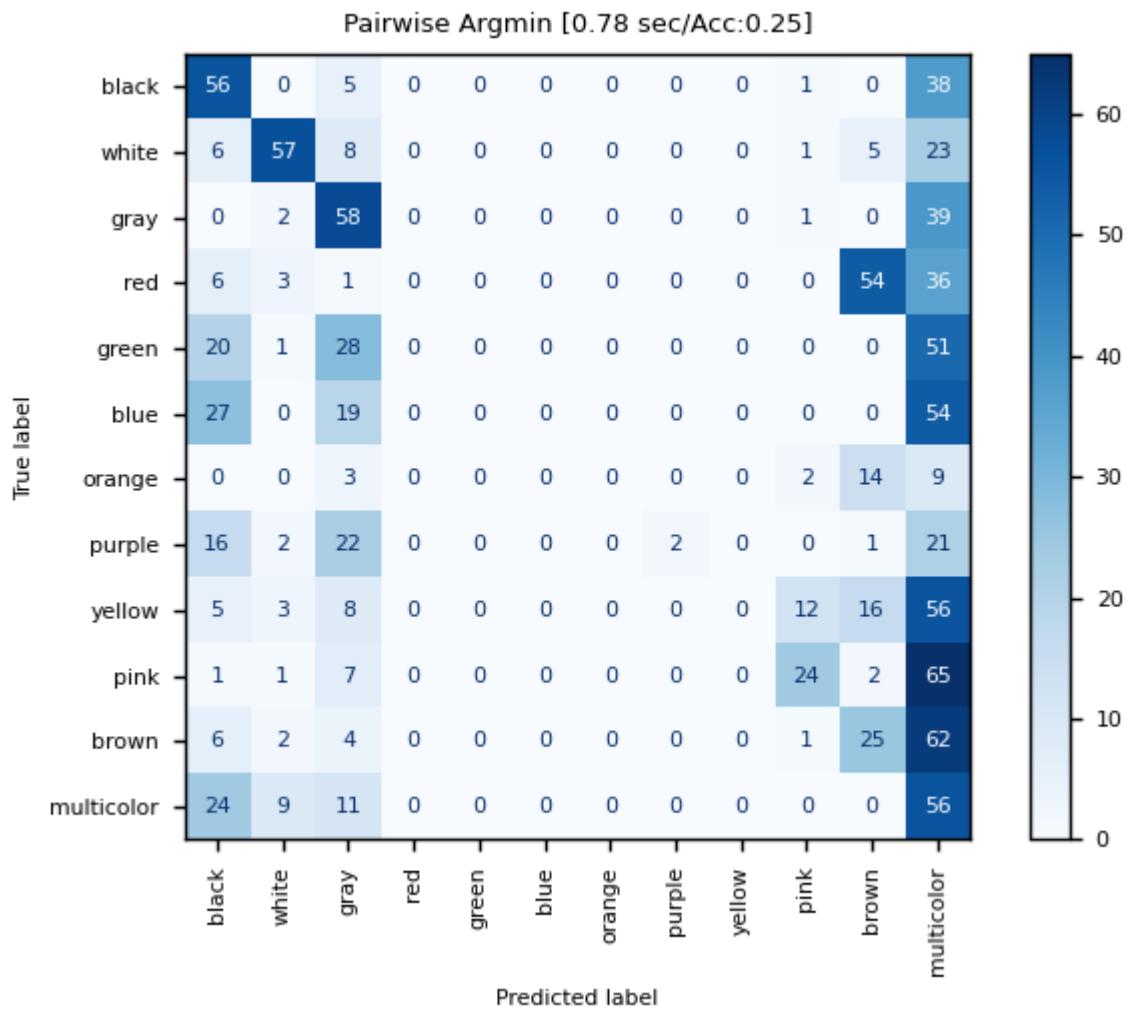
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Pairwise Argmin Result

Total Duration: 0.78 sec.

Multi-Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.3353	0.5600	0.4195	100
white	0.7125	0.5700	0.6333	100
gray	0.3333	0.5800	0.4234	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0312	0.0606	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5714	0.2400	0.3380	100
brown	0.2137	0.2500	0.2304	100
multicolor	0.1098	0.5600	0.1836	100
accuracy			0.2546	
macro avg	0.2730	0.2326	0.1907	1092
weighted avg	0.2670	0.2546	0.2076	1092



Multi-Color with HSL / Pairwise Argmin Confusion Matrix

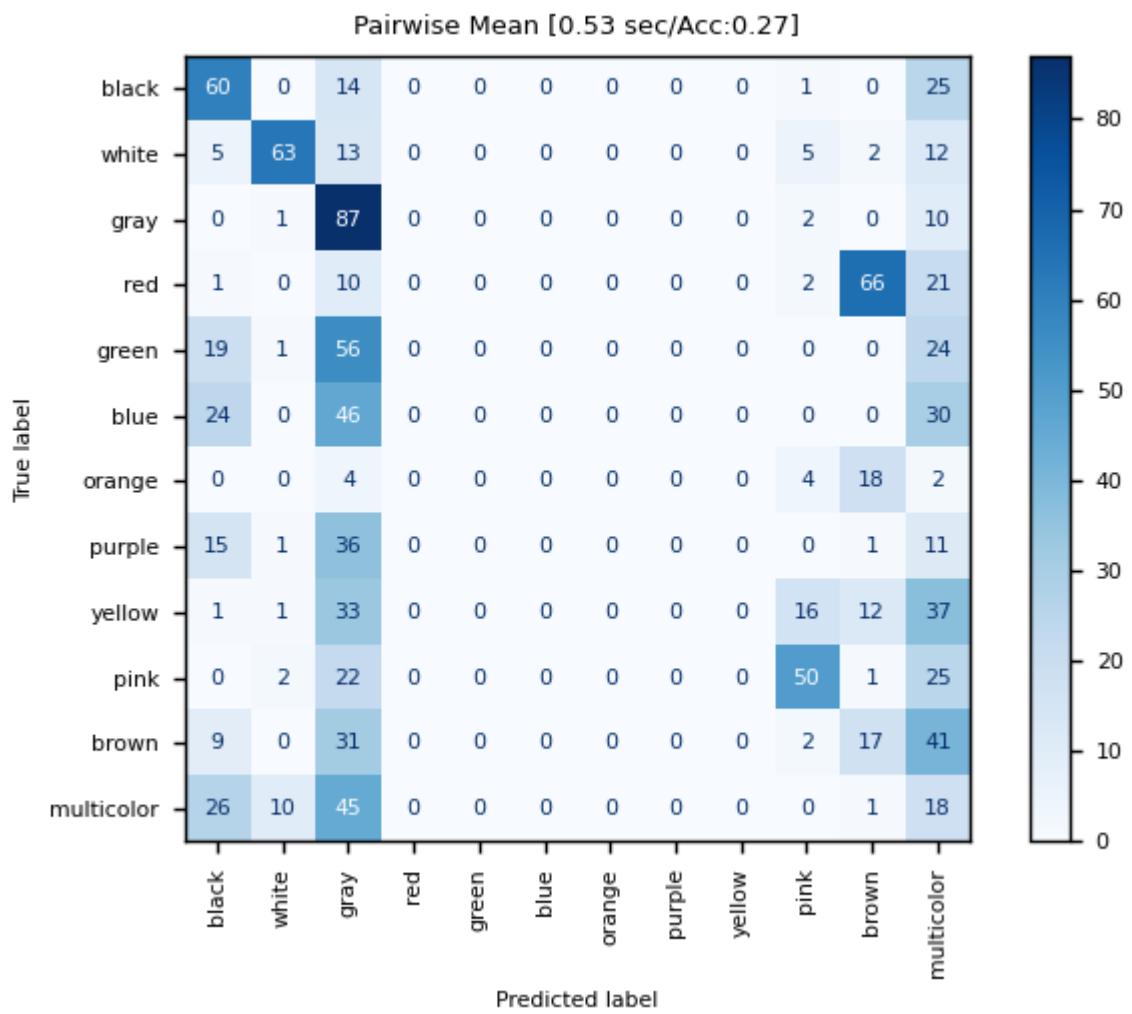
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Pairwise Mean Result

Total Duration: 0.53 sec.

Multi-Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.3750	0.6000	0.4615	100
white	0.7975	0.6300	0.7039	100
gray	0.2191	0.8700	0.3501	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.6098	0.5000	0.5495	100
brown	0.1441	0.1700	0.1560	100
multicolor	0.0703	0.1800	0.1011	100
accuracy			0.2701	
macro avg	0.1846	0.2458	0.1935	1092
weighted avg	0.2029	0.2701	0.2126	1092



Multi-Color with HSL / Pairwise Mean Confusion Matrix

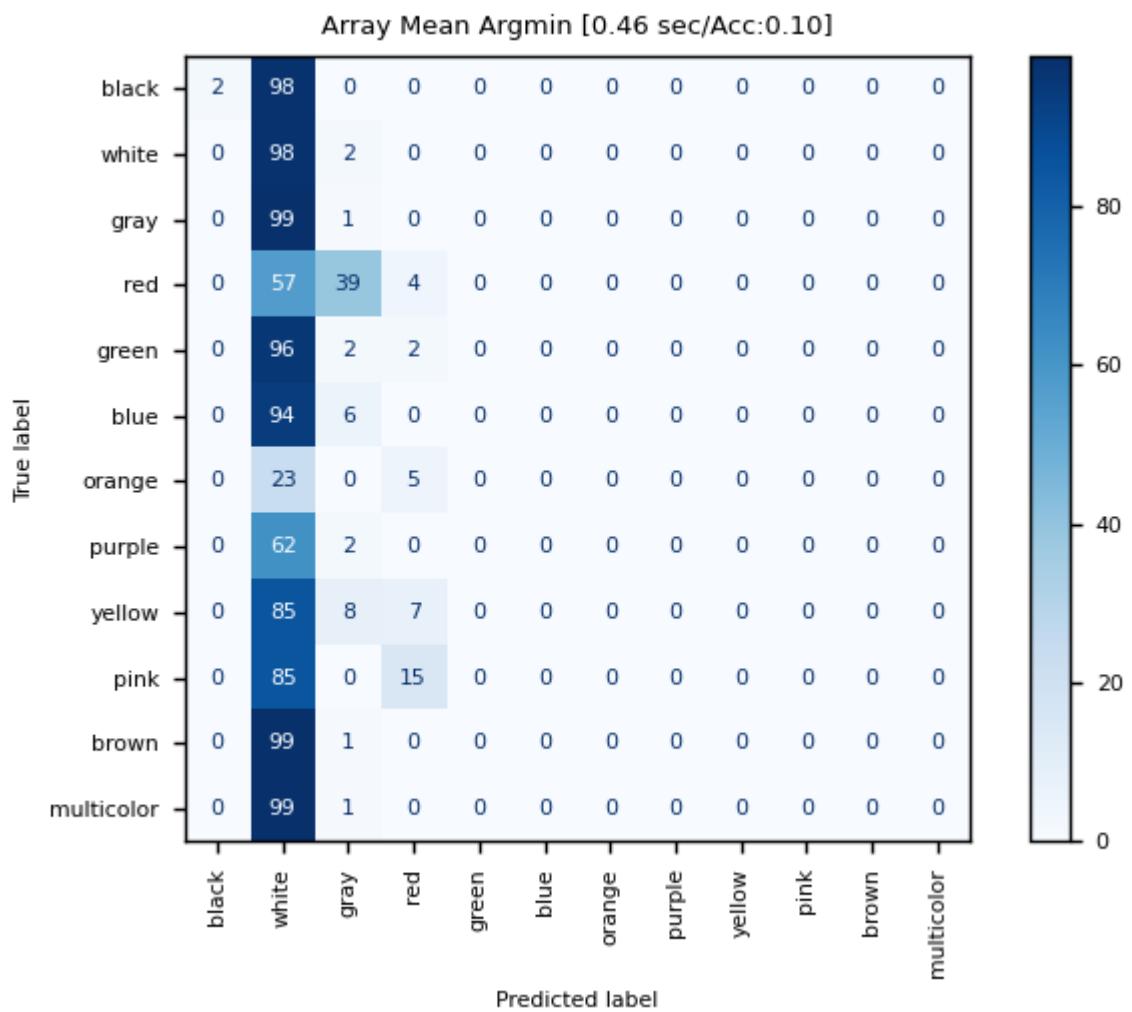
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Array Mean Argmin Result

Total Duration: 0.46 sec.

Multi-Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	1.0000	0.0200	0.0392	100
white	0.0985	0.9800	0.1790	100
gray	0.0161	0.0100	0.0123	100
red	0.1212	0.0400	0.0602	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.0000	0.0000	0.0000	100
brown	0.0000	0.0000	0.0000	100
multicolor	0.0000	0.0000	0.0000	100
accuracy			0.0962	
macro avg	0.1030	0.0875	0.0242	1092
weighted avg	0.1132	0.0962	0.0266	1092



Multi-Color with HSL / Array Mean Argmin Confusion Matrix

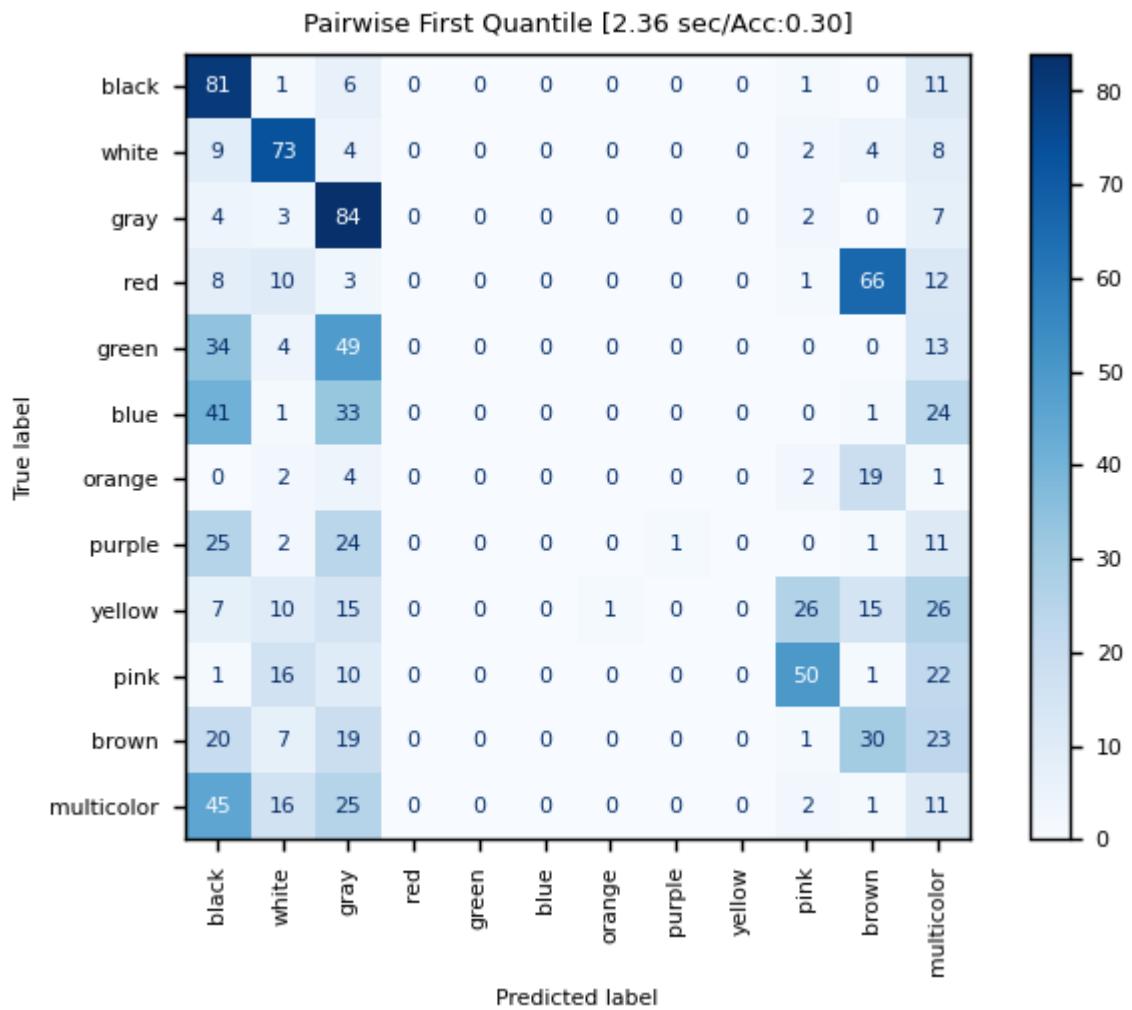
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Pairwise First Quantile Result

Total Duration: 2.36 sec.

Multi-Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.2945	0.8100	0.4320	100
white	0.5034	0.7300	0.5959	100
gray	0.3043	0.8400	0.4468	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0156	0.0308	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5747	0.5000	0.5348	100
brown	0.2174	0.3000	0.2521	100
multicolor	0.0651	0.1100	0.0818	100
accuracy			0.3022	
macro avg	0.2466	0.2755	0.1978	1092
weighted avg	0.2381	0.3022	0.2164	1092



Multi-Color with HSL / Pairwise First Quantile Confusion Matrix

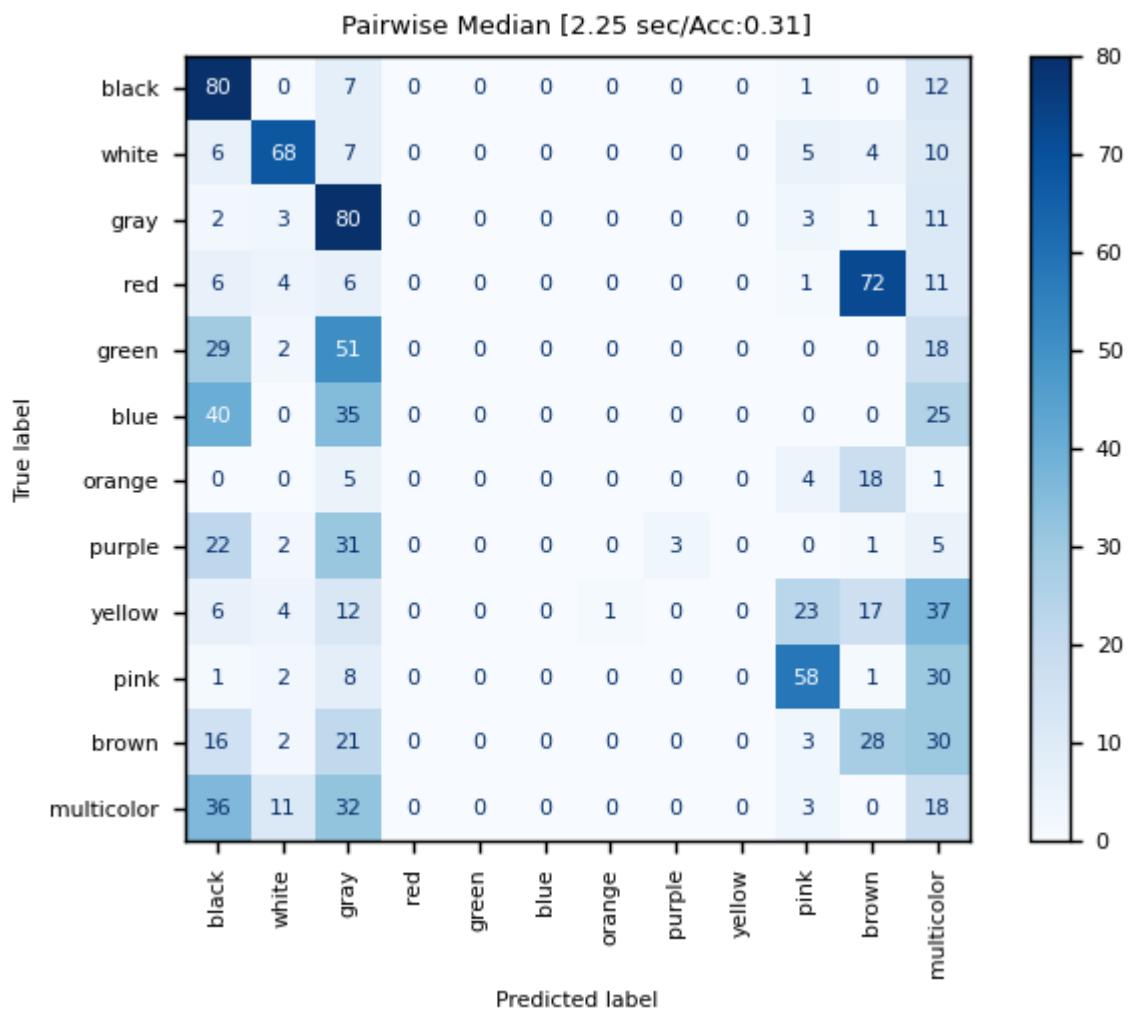
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Pairwise Median Result

Total Duration: 2.25 sec.

Multi-Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.3279	0.8000	0.4651	100
white	0.6939	0.6800	0.6869	100
gray	0.2712	0.8000	0.4051	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	1.0000	0.0469	0.0896	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5743	0.5800	0.5771	100
brown	0.1972	0.2800	0.2314	100
multicolor	0.0865	0.1800	0.1169	100
accuracy			0.3068	
macro avg	0.2626	0.2806	0.2143	1092
weighted avg	0.2556	0.3068	0.2326	1092



Multi-Color with HSL / Pairwise Median Confusion Matrix

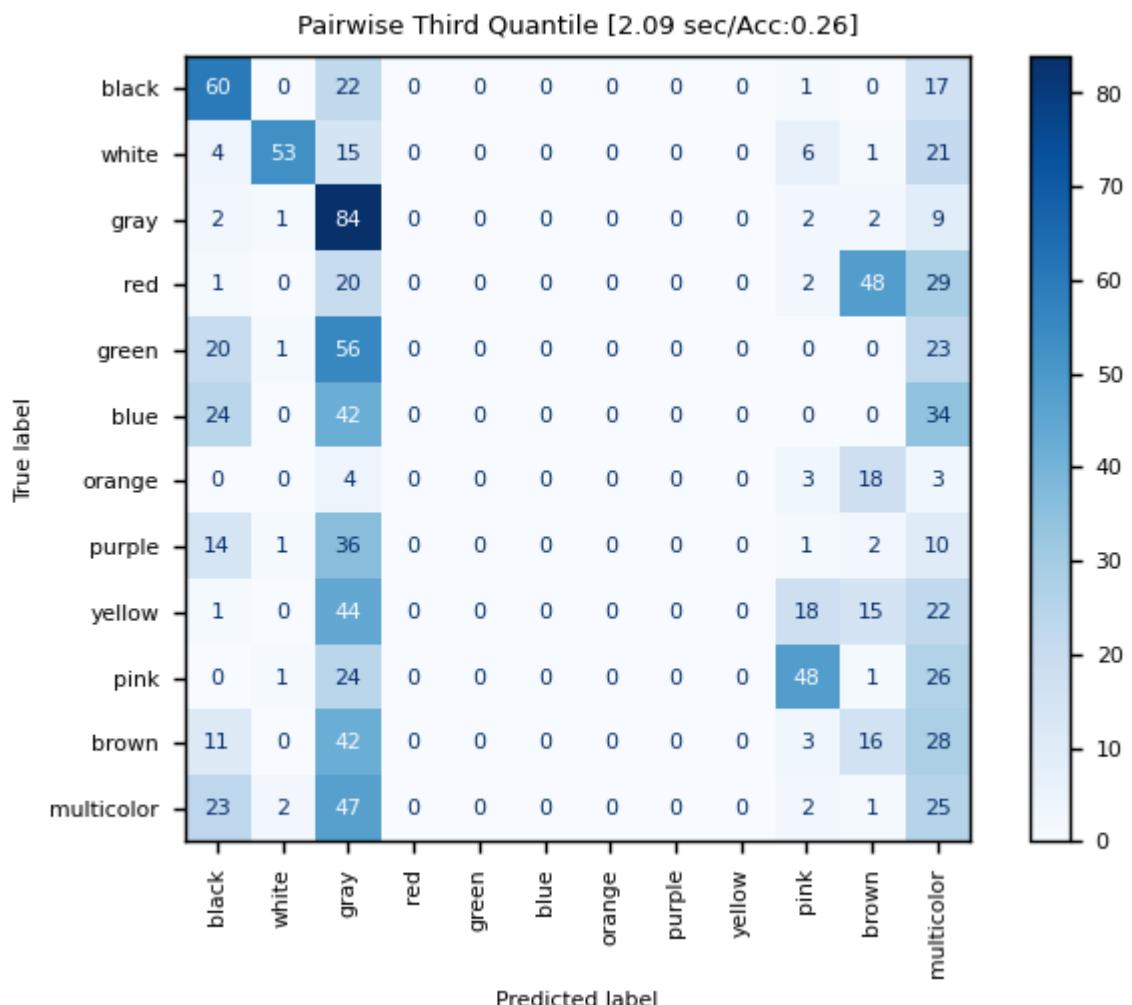
Manually Validated Dataset Improvement Run

Multi-Color with HSL / Pairwise Third Quantile Result

Total Duration: 2.09 sec.

Multi-Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.3750	0.6000	0.4615	100
white	0.8983	0.5300	0.6667	100
gray	0.1927	0.8400	0.3134	100
red	0.0000	0.0000	0.0000	100
green	0.0000	0.0000	0.0000	100
blue	0.0000	0.0000	0.0000	100
orange	0.0000	0.0000	0.0000	28
purple	0.0000	0.0000	0.0000	64
yellow	0.0000	0.0000	0.0000	100
pink	0.5581	0.4800	0.5161	100
brown	0.1538	0.1600	0.1569	100
multicolor	0.1012	0.2500	0.1441	100
accuracy			0.2619	
macro avg	0.1899	0.2383	0.1882	1092
weighted avg	0.2087	0.2619	0.2068	1092



Multi-Color with HSL / Pairwise Third Quantile Confusion Matrix

REFINED DATASET IMPROVEMENT RUN

5.1. RGB Color Space

5.1.1. CNN AlexNet with Single Color

Refined Dataset Improvement Run

CNN AlexNet Single Color with RGB

Total Duration: 2128.86 sec.

Learning Rate: 0.000070

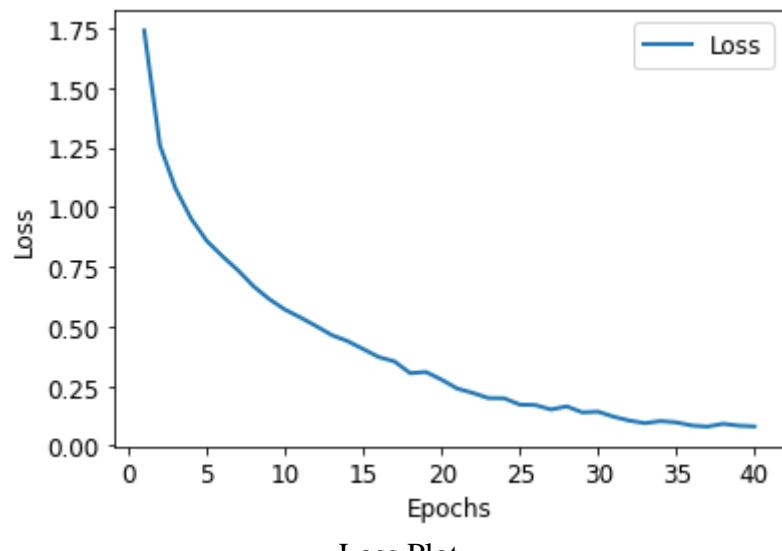
Dropout: 0.10

Batch Size: 32

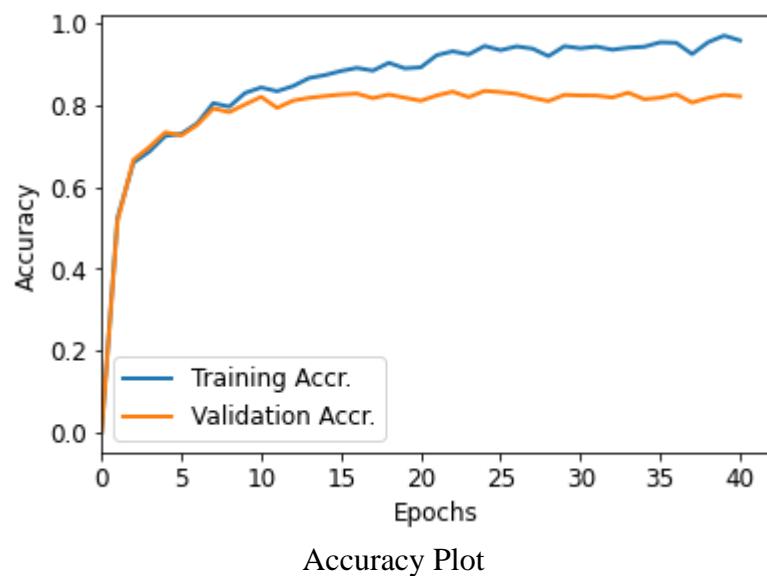
Epoch: 40

Best Performing Epoch: 24

Summary	
Method	Accuracy
Training Dataset	0.9435
Validation Dataset	0.8339
Test Dataset	0.8222



Loss Plot

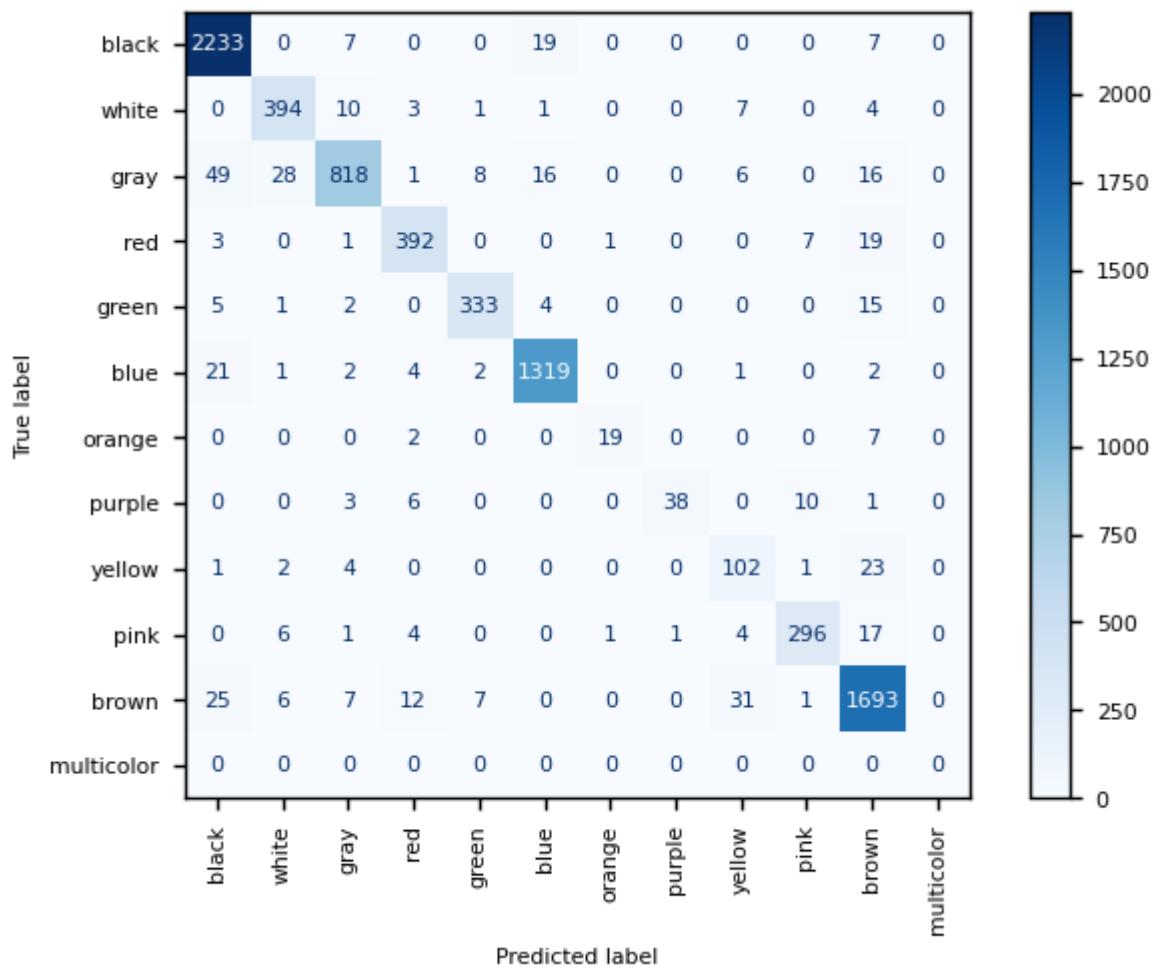


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9555	0.9854	0.9702	2266
white	0.8995	0.9381	0.9184	420
gray	0.9567	0.8684	0.9104	942
red	0.9245	0.9267	0.9256	423
green	0.9487	0.9250	0.9367	360
blue	0.9706	0.9756	0.9731	1352
orange	0.9048	0.6786	0.7755	28
purple	0.9744	0.6552	0.7835	58
yellow	0.6755	0.7669	0.7183	133
pink	0.9397	0.8970	0.9178	330
brown	0.9385	0.9501	0.9442	1782
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.9435	0.9435	0.9435	8094
macro avg	0.8407	0.7972	0.8145	8094
weighted avg	0.9443	0.9435	0.9432	8094

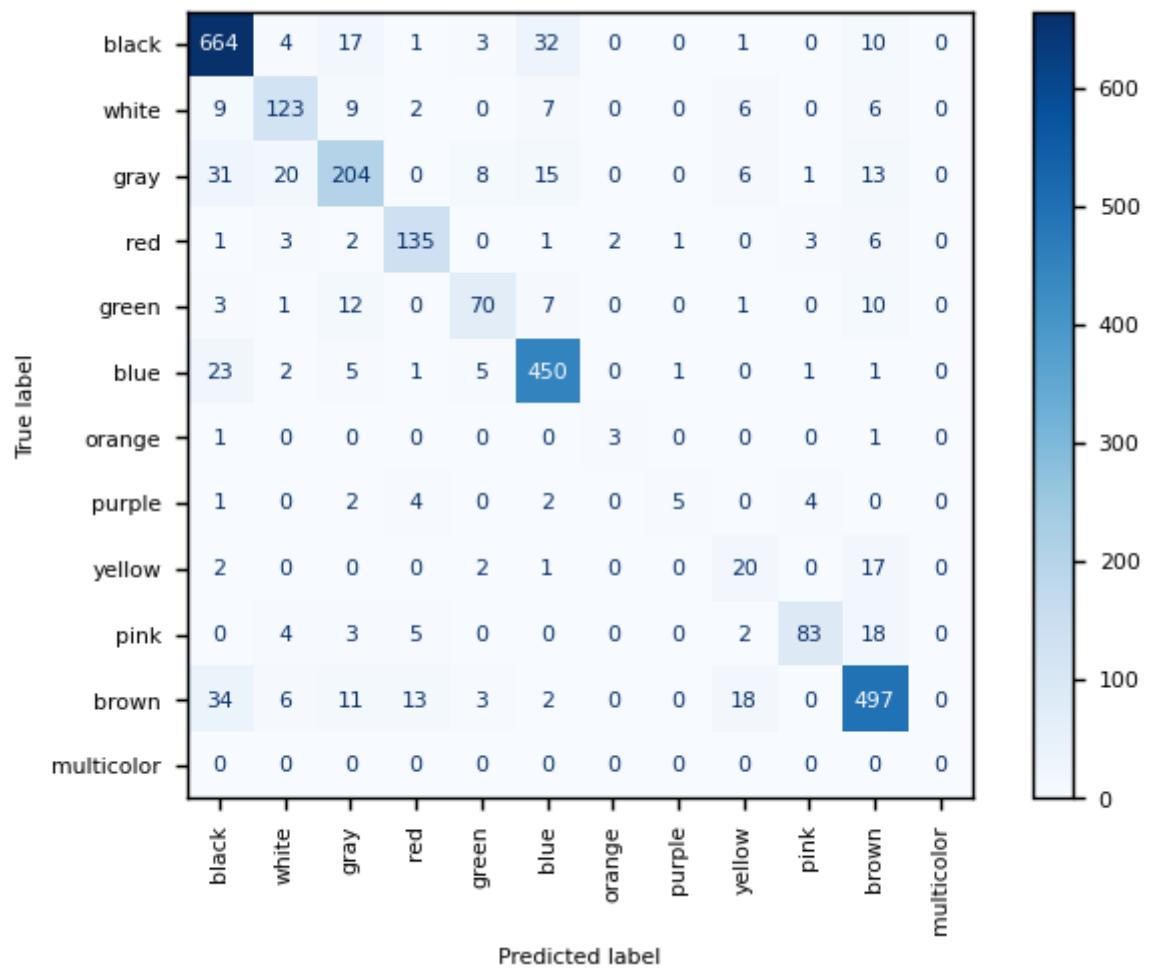


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8635	0.9071	0.8847	732
white	0.7546	0.7593	0.7569	162
gray	0.7698	0.6846	0.7247	298
red	0.8385	0.8766	0.8571	154
green	0.7692	0.6731	0.7179	104
blue	0.8704	0.9202	0.8946	489
orange	0.6000	0.6000	0.6000	5
purple	0.7143	0.2778	0.4000	18
yellow	0.3704	0.4762	0.4167	42
pink	0.9022	0.7217	0.8019	115
brown	0.8584	0.8510	0.8547	584
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.8339	0.8339	0.8339	2703
macro avg	0.6926	0.6456	0.6591	2703
weighted avg	0.8342	0.8339	0.8322	2703

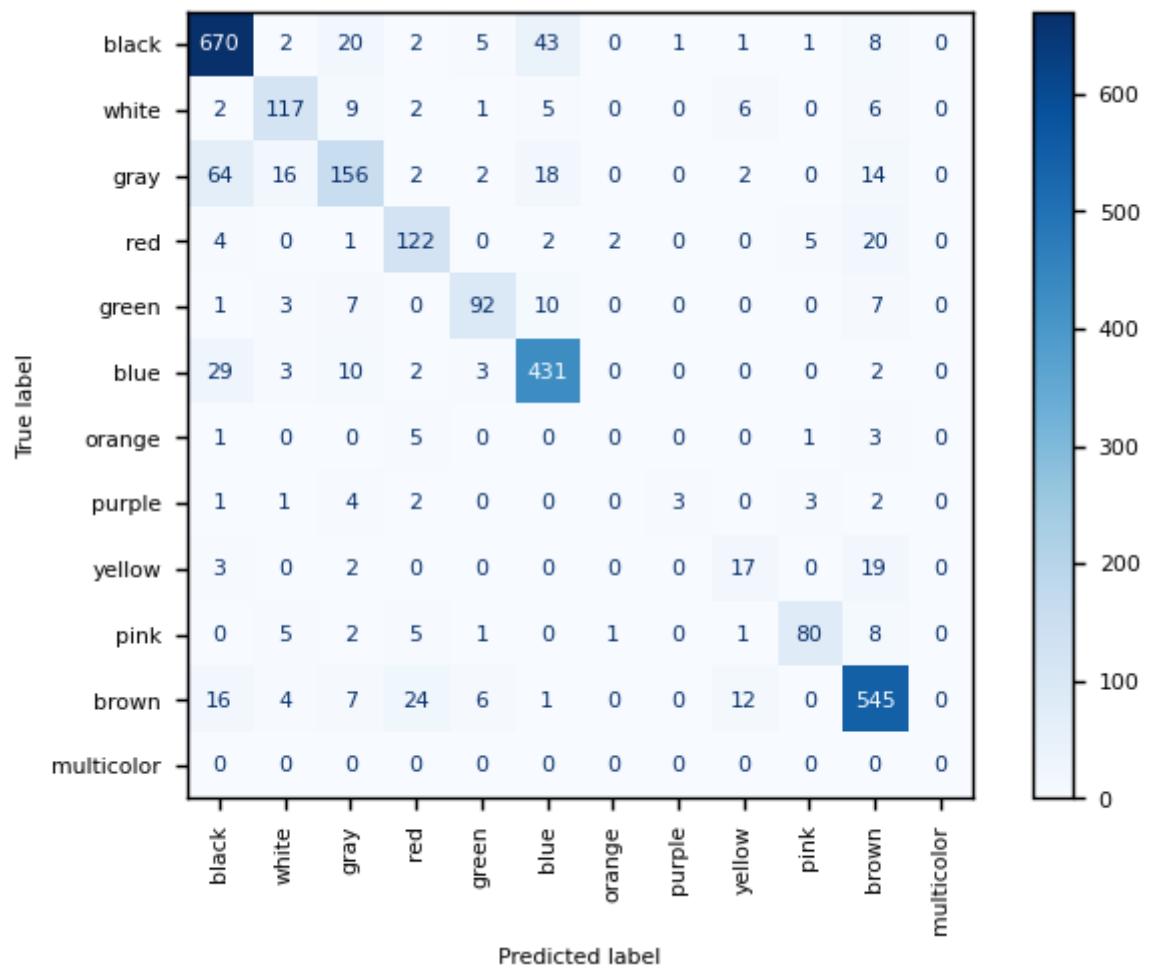


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8470	0.8898	0.8679	753
white	0.7748	0.7905	0.7826	148
gray	0.7156	0.5693	0.6341	274
red	0.7349	0.7821	0.7578	156
green	0.8364	0.7667	0.8000	120
blue	0.8451	0.8979	0.8707	480
orange	0.0000	0.0000	0.0000	10
purple	0.7500	0.1875	0.3000	16
yellow	0.4359	0.4146	0.4250	41
pink	0.8889	0.7767	0.8290	103
brown	0.8596	0.8862	0.8727	615
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.8222	0.8222	0.8222	2716
macro avg	0.6407	0.5801	0.5950	2716
weighted avg	0.8171	0.8222	0.8172	2716



Test Dataset Confusion Matrix

gray (100.0%) brown (98.8%) brown (99.6%) brown (54.3%) blue (100.0%) blue (99.9%)
 >> white >> brown >> brown >> yellow >> blue >> blue



gray (33.6%) pink (97.9%) black (100.0%) green (67.7%) red (99.9%) pink (97.3%)
 >> blue >> pink >> black >> gray >> red >> pink



black (99.3%) brown (66.7%) red (96.9%) brown (100.0%) red (85.4%) black (93.9%)
 >> black >> brown >> brown >> red >> blue



black (100.0%) red (100.0%) gray (48.8%) white (99.7%) blue (99.9%) red (66.0%)
 >> black >> red >> gray >> white >> blue >> brown



yellow (50.4%) black (100.0%) blue (93.5%) black (85.1%) brown (99.5%) brown (100.0%)
 >> yellow >> black >> green >> black >> brown >> brown



black (99.9%) blue (100.0%) brown (96.9%) black (100.0%) brown (96.2%) black (100.0%)
 >> black >> blue >> brown >> black >> brown >> black



Sample Images from Test Dataset

gray (33.6%)
 >> blue blue (87.0%)
 >> gray black (98.4%)
 >> red black (93.9%)
 >> blue green (67.7%)
 >> gray brown (54.3%)
 >> yellow



blue (46.2%)
 >> green gray (100.0%)
 >> white black (96.5%)
 >> gray pink (87.5%)
 >> red black (100.0%) blue (87.6%)
 >> gray >> gray



red (65.8%)
 >> brown white (70.4%)
 >> gray gray (95.9%)
 >> black white (97.8%)
 >> brown red (55.6%)
 >> orange brown (50.7%)
 >> gray



white (83.2%)
 >> gray black (66.8%)
 >> yellow brown (49.7%)
 >> white green (56.7%)
 >> blue gray (95.6%)
 >> black brown (77.5%)
 >> yellow



gray (79.9%)
 >> blue brown (90.1%)
 >> black blue (93.5%)
 >> green blue (51.8%)
 >> black yellow (45.1%)
 >> gray black (86.9%)
 >> blue



black (74.5%)
 >> gray white (66.3%)
 >> gray brown (41.9%)
 >> white black (98.6%)
 >> gray red (97.7%)
 >> white red (66.0%)
 >> brown



Sample Images that cannot be accurately predicted in the Test Dataset

5.1.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Quantization Fastoctree	35.31 sec.	0.4315
Quantization Mmcq	538.97 sec.	0.4143
Quantization Libimagequant	228.15 sec.	0.3602
Pairwise Argmin	15.96 sec.	0.4308
Pairwise Mean	8.33 sec.	0.3985
Array Mean Argmin	7.15 sec.	0.1456
Pairwise First Quantile	37.92 sec.	0.4200
Pairwise Median	37.71 sec.	0.4271
Pairwise Third Quantile	35.72 sec.	0.3727

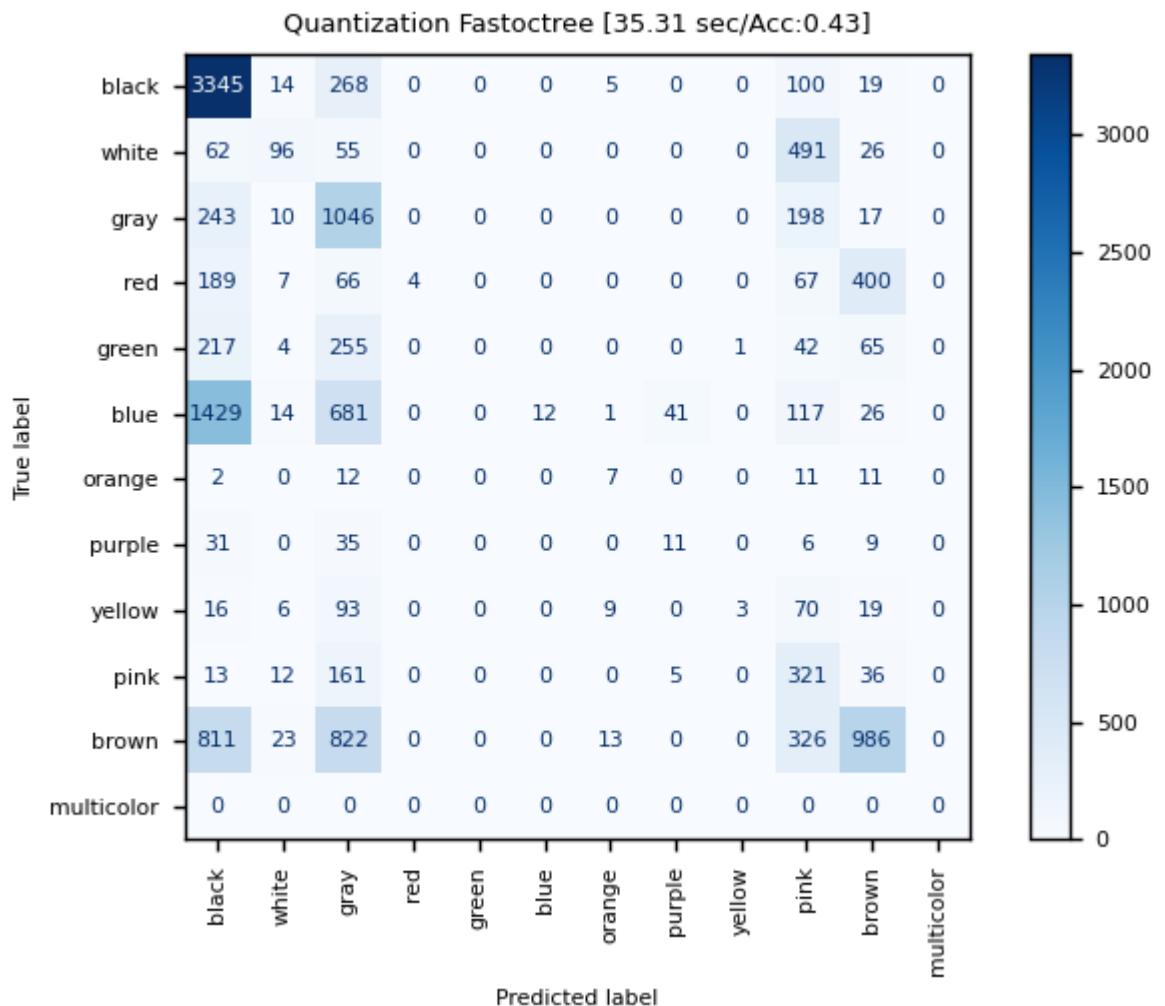
Refined Dataset Improvement Run

Single Color with RGB / Quantization Fastoctree Result

Total Duration: 35.31 sec.

Single Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.5261	0.8918	0.6618	3751
white	0.5161	0.1315	0.2096	730
gray	0.2994	0.6909	0.4177	1514
red	1.0000	0.0055	0.0109	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0052	0.0103	2321
orange	0.2000	0.1628	0.1795	43
purple	0.1930	0.1196	0.1477	92
yellow	0.7500	0.0139	0.0273	216
pink	0.1835	0.5858	0.2795	548
brown	0.6109	0.3308	0.4292	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4315	0.4315	0.4315	13513
macro avg	0.4399	0.2448	0.1978	13513
weighted avg	0.5896	0.4315	0.3522	13513



Single Color with RGB / Quantization Fastoctree Confusion Matrix

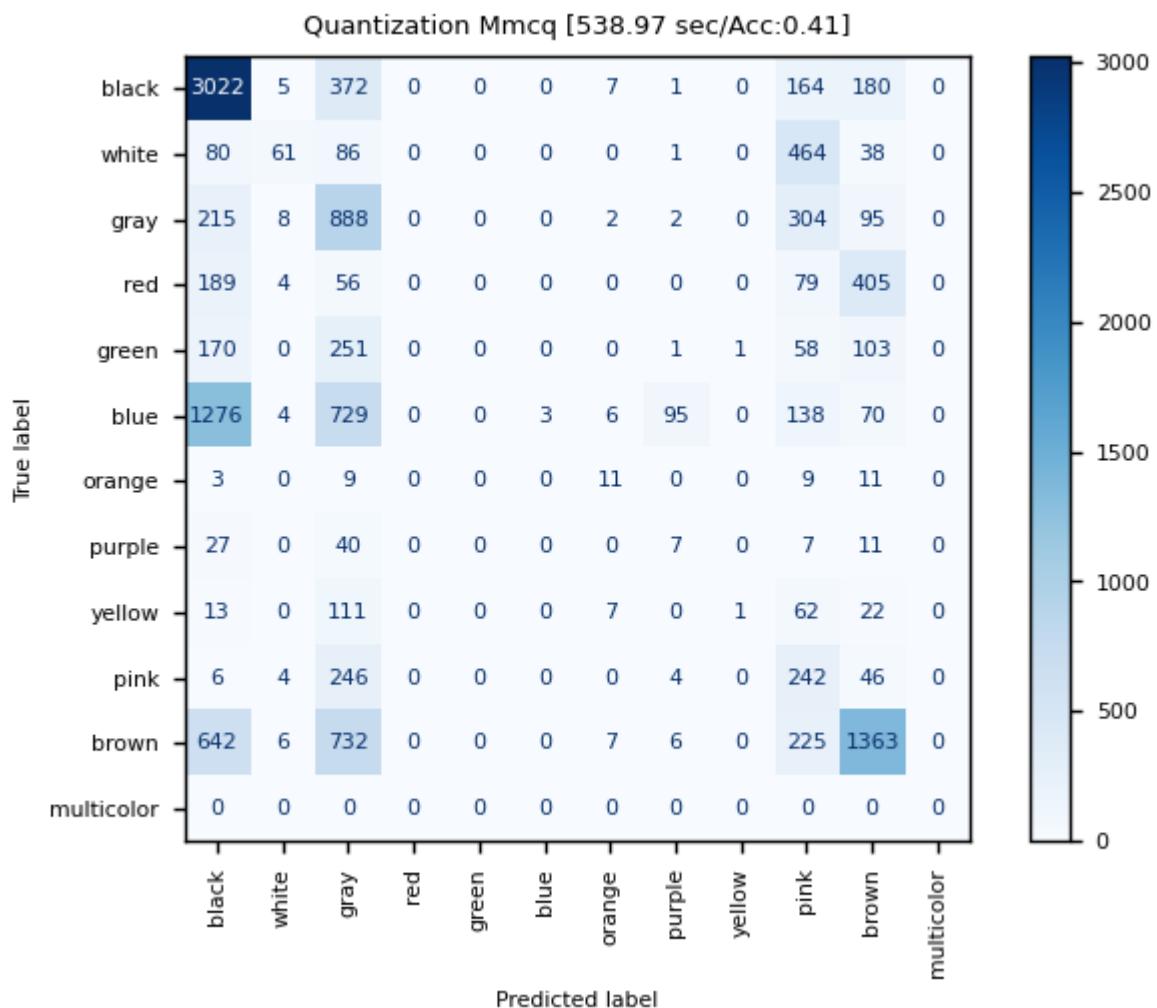
Refined Dataset Improvement Run

Single Color with RGB / Quantization Mmcq Result

Total Duration: 538.97 sec.

Single Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.5355	0.8057	0.6434	3751
white	0.6630	0.0836	0.1484	730
gray	0.2523	0.5865	0.3528	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0013	0.0026	2321
orange	0.2750	0.2558	0.2651	43
purple	0.0598	0.0761	0.0670	92
yellow	0.5000	0.0046	0.0092	216
pink	0.1381	0.4416	0.2104	548
brown	0.5815	0.4572	0.5119	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4143	0.4143	0.4143	13513
macro avg	0.3338	0.2260	0.1842	13513
weighted avg	0.5277	0.4143	0.3495	13513



Single Color with RGB / Quantization Mmcq Confusion Matrix

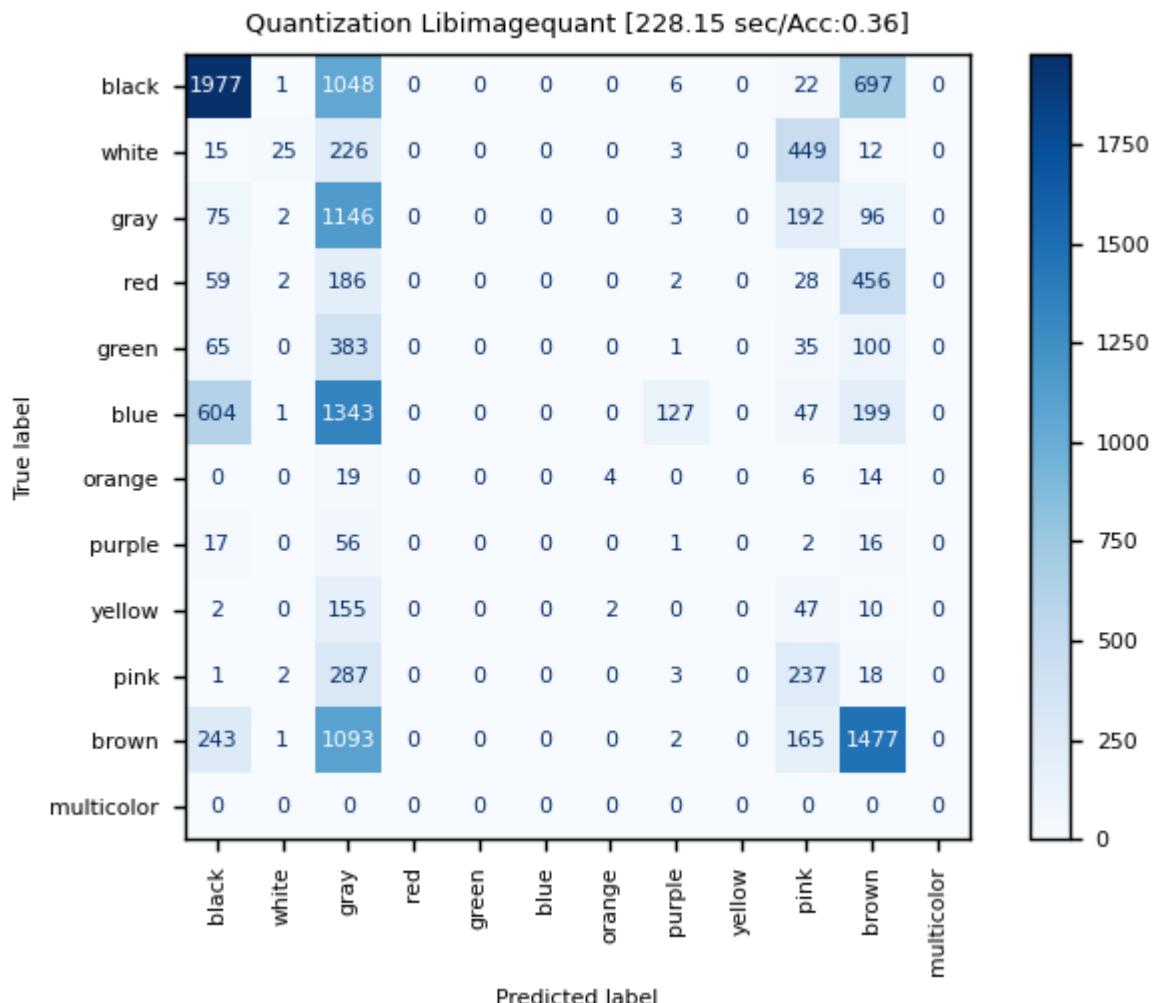
Refined Dataset Improvement Run

Single Color with RGB / Quantization Libimagequant Result

Total Duration: 228.15 sec.

Single Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.6465	0.5271	0.5807	3751
white	0.7353	0.0342	0.0654	730
gray	0.1929	0.7569	0.3074	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.6667	0.0930	0.1633	43
purple	0.0068	0.0109	0.0083	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1927	0.4325	0.2666	548
brown	0.4772	0.4955	0.4862	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3602	0.3602	0.3602	13513
macro avg	0.2432	0.1958	0.1565	13513
weighted avg	0.3560	0.3602	0.3178	13513



Single Color with RGB / Quantization Libimagequant Confusion Matrix

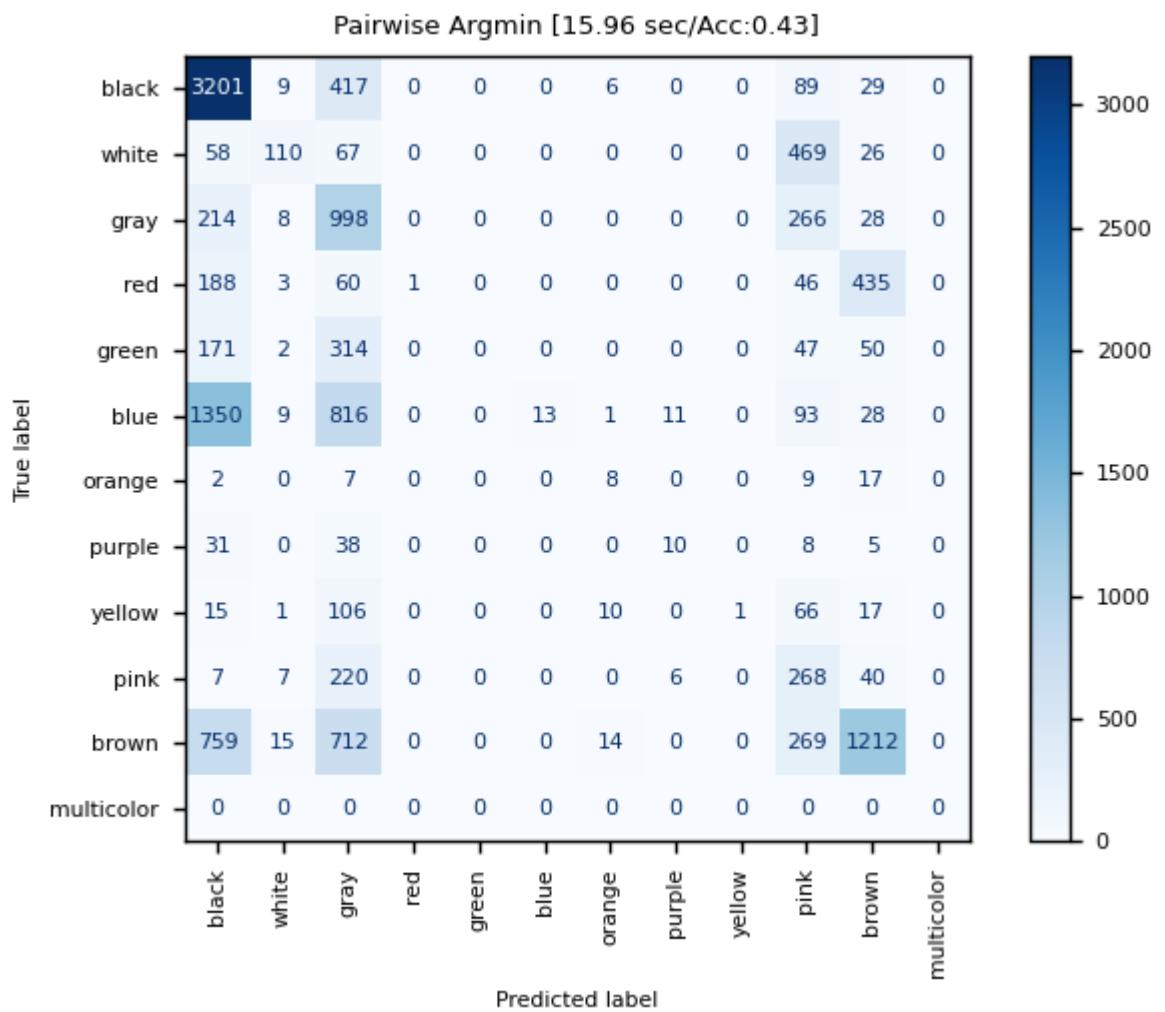
Refined Dataset Improvement Run

Single Color with RGB / Pairwise Argmin Result

Total Duration: 15.96 sec.

Single Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5339	0.8534	0.6568	3751
white	0.6707	0.1507	0.2461	730
gray	0.2658	0.6592	0.3788	1514
red	1.0000	0.0014	0.0027	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0056	0.0111	2321
orange	0.2051	0.1860	0.1951	43
purple	0.3704	0.1087	0.1681	92
yellow	1.0000	0.0046	0.0092	216
pink	0.1644	0.4891	0.2461	548
brown	0.6423	0.4066	0.4979	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4308	0.4308	0.4308	13513
macro avg	0.4877	0.2388	0.2010	13513
weighted avg	0.6077	0.4308	0.3619	13513



Single Color with RGB / Pairwise Argmin Confusion Matrix

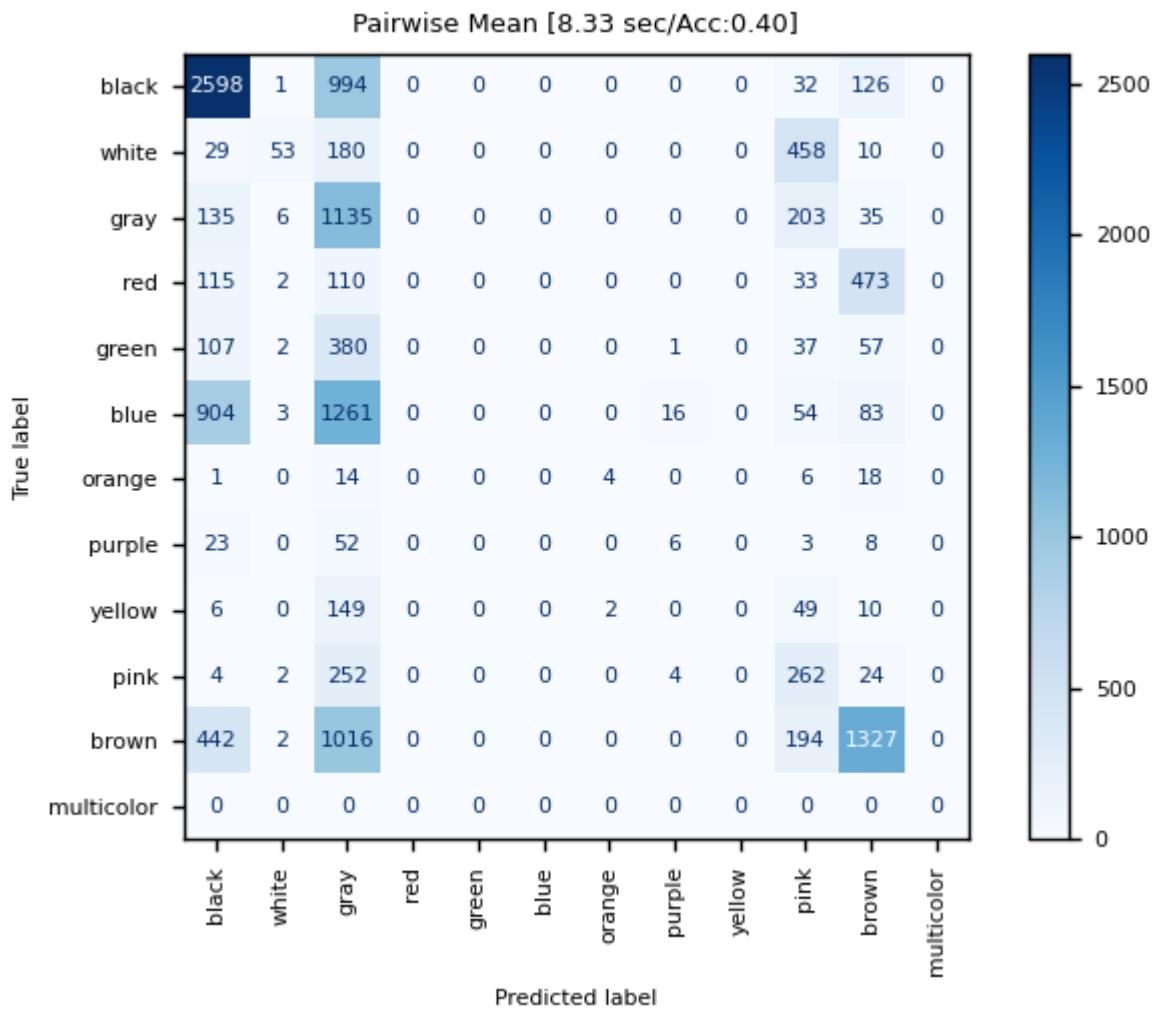
Refined Dataset Improvement Run

Single Color with RGB / Pairwise Mean Result

Total Duration: 8.33 sec.

Single Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.5953	0.6926	0.6403	3751
white	0.7465	0.0726	0.1323	730
gray	0.2048	0.7497	0.3217	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.6667	0.0930	0.1633	43
purple	0.2222	0.0652	0.1008	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1968	0.4781	0.2789	548
brown	0.6112	0.4452	0.5151	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3985	0.3985	0.3985	13513
macro avg	0.2703	0.2164	0.1794	13513
weighted avg	0.3750	0.3985	0.3471	13513



Single Color with RGB / Pairwise Mean Confusion Matrix

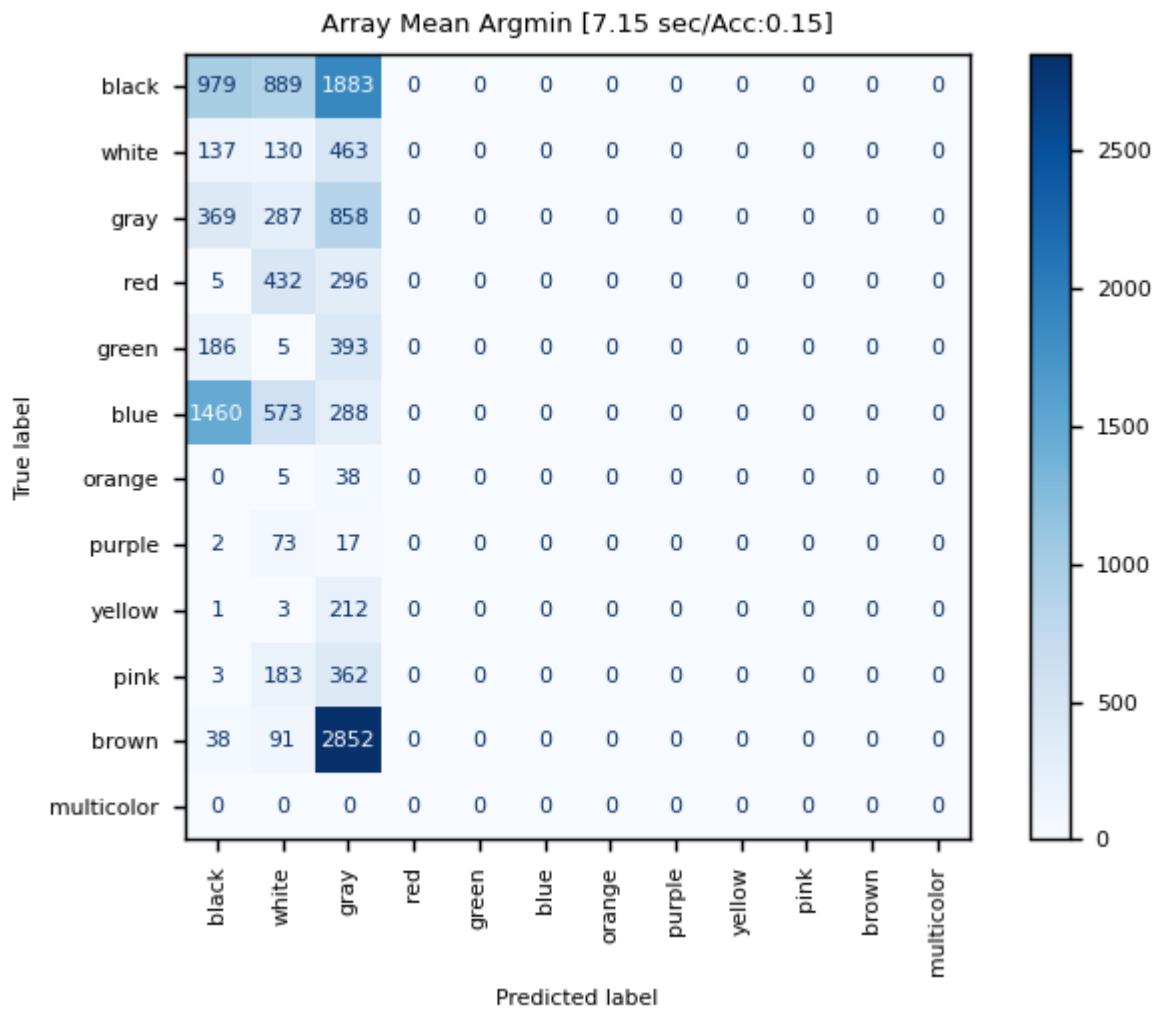
Refined Dataset Improvement Run

Single Color with RGB / Array Mean Argmin Result

Total Duration: 7.15 sec.

Single Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.3079	0.2610	0.2825	3751
white	0.0487	0.1781	0.0764	730
gray	0.1120	0.5667	0.1870	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.0000	0.0000	0.0000	548
brown	0.0000	0.0000	0.0000	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.1456	0.1456	0.1456	13513
macro avg	0.0390	0.0838	0.0455	13513
weighted avg	0.1006	0.1456	0.1035	13513



Single Color with RGB / Array Mean Argmin Confusion Matrix

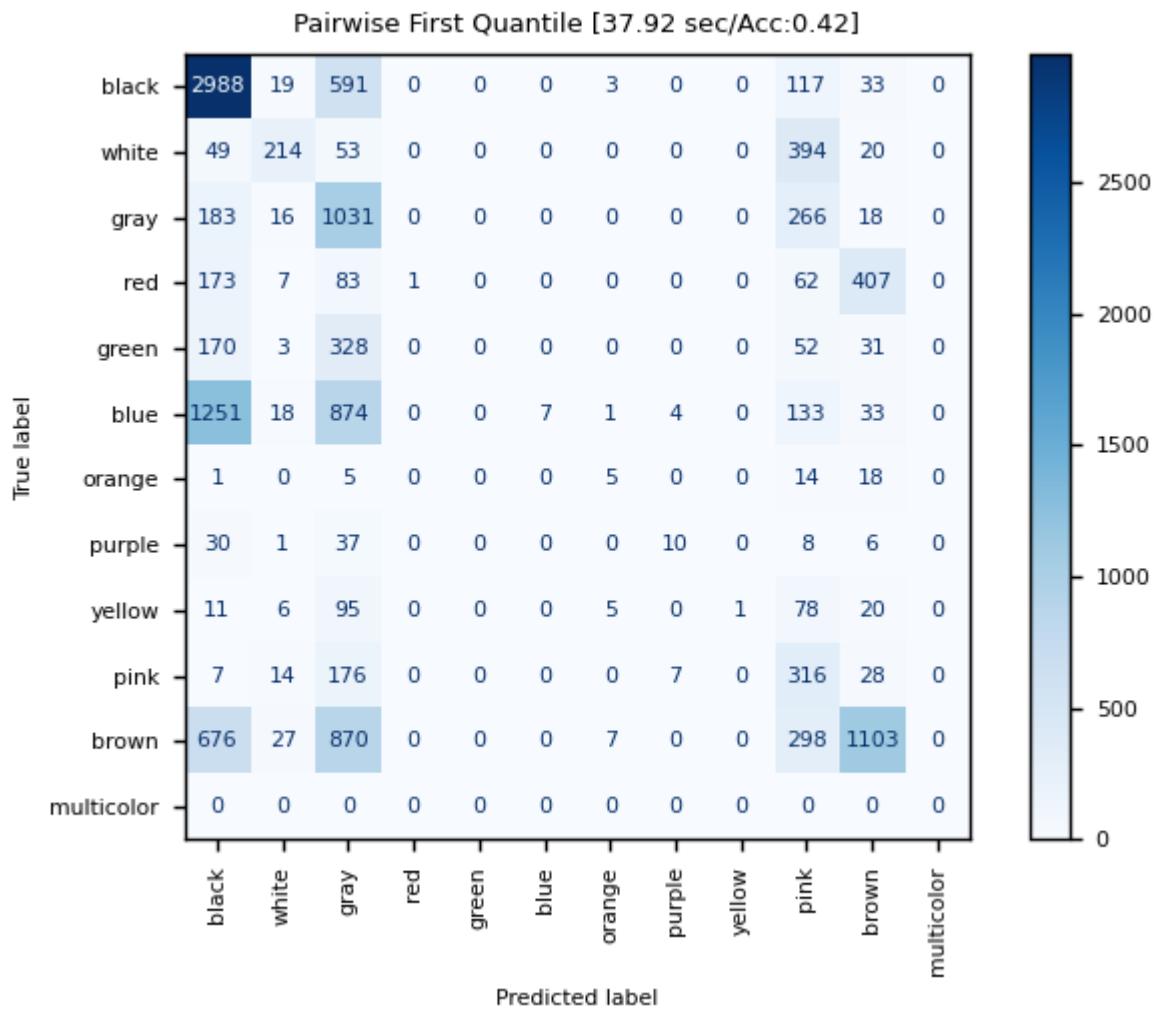
Refined Dataset Improvement Run

Single Color with RGB / Pairwise First Quantile Result

Total Duration: 37.92 sec.

Single Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5394	0.7966	0.6433	3751
white	0.6585	0.2932	0.4057	730
gray	0.2489	0.6810	0.3645	1514
red	1.0000	0.0014	0.0027	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0030	0.0060	2321
orange	0.2381	0.1163	0.1562	43
purple	0.4762	0.1087	0.1770	92
yellow	1.0000	0.0046	0.0092	216
pink	0.1818	0.5766	0.2765	548
brown	0.6424	0.3700	0.4696	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4200	0.4200	0.4200	13513
macro avg	0.4988	0.2459	0.2092	13513
weighted avg	0.6083	0.4200	0.3591	13513



Single Color with RGB / Pairwise First Quantile Confusion Matrix

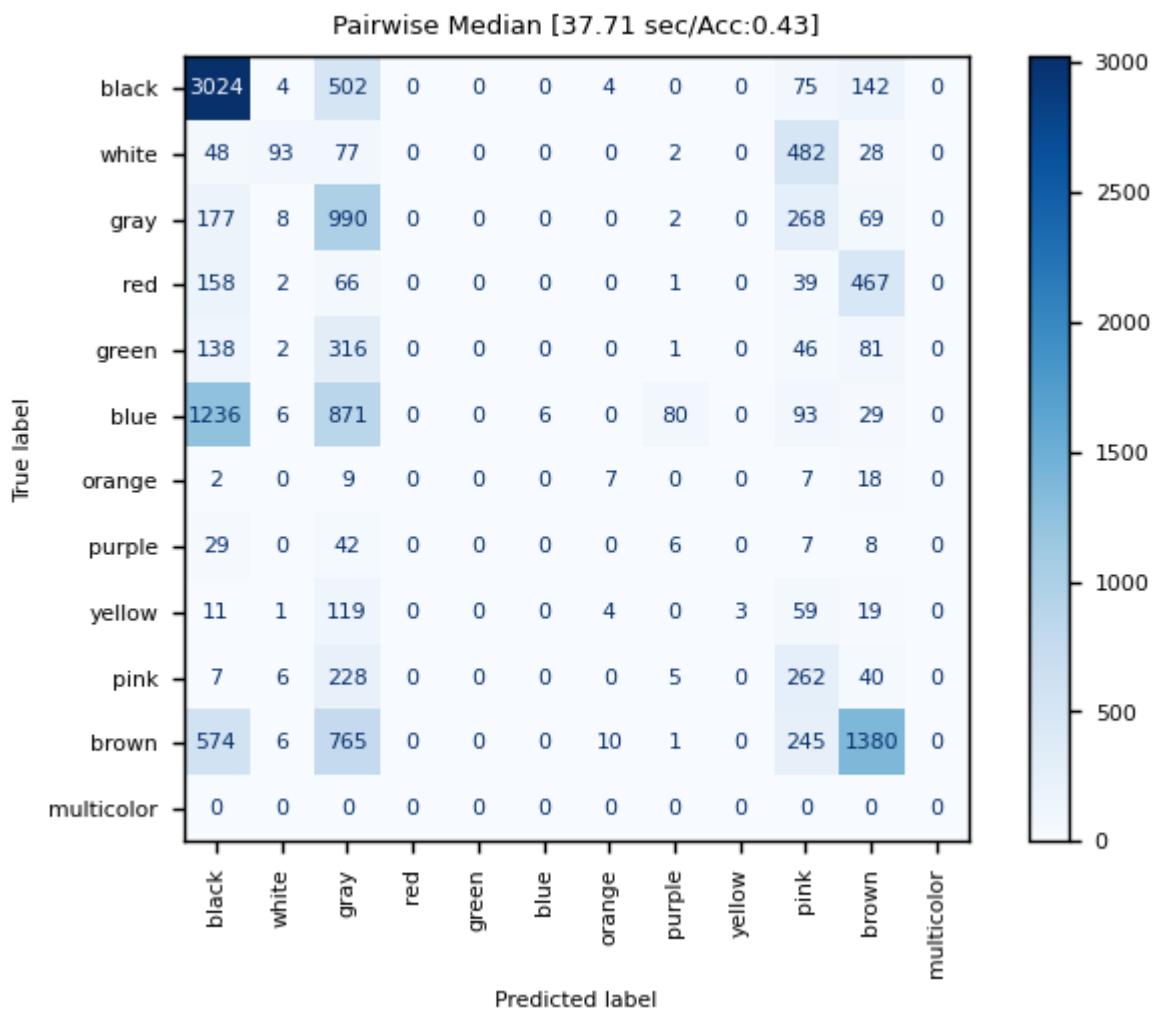
Refined Dataset Improvement Run

Single Color with RGB / Pairwise Median Result

Total Duration: 37.71 sec.

Single Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.5596	0.8062	0.6606	3751
white	0.7266	0.1274	0.2168	730
gray	0.2484	0.6539	0.3601	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0026	0.0052	2321
orange	0.2800	0.1628	0.2059	43
purple	0.0612	0.0652	0.0632	92
yellow	1.0000	0.0139	0.0274	216
pink	0.1655	0.4781	0.2459	548
brown	0.6050	0.4629	0.5245	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4271	0.4271	0.4271	13513
macro avg	0.3872	0.2311	0.1925	13513
weighted avg	0.5516	0.4271	0.3635	13513



Single Color with RGB / Pairwise Median Confusion Matrix

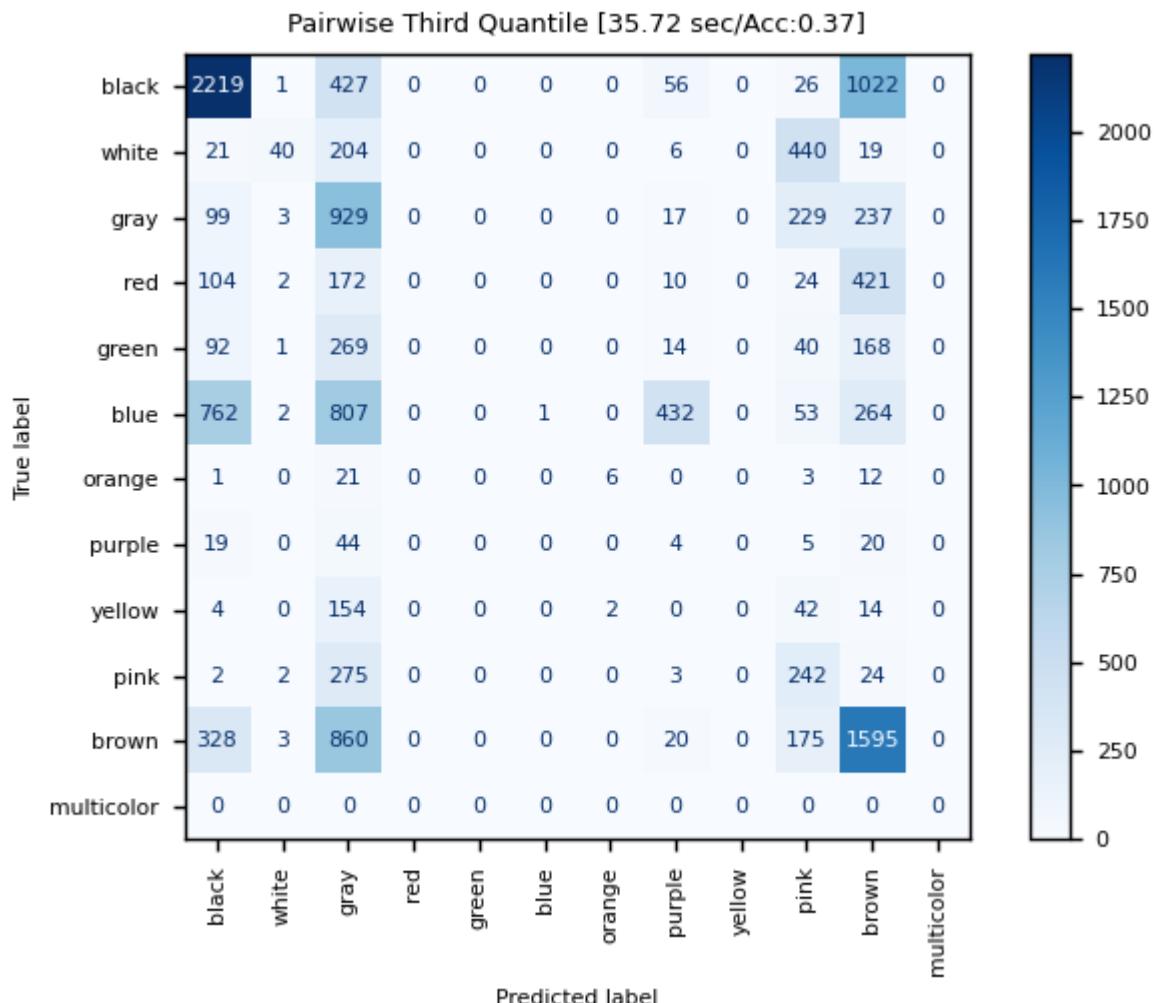
Refined Dataset Improvement Run

Single Color with RGB / Pairwise Third Quantile Result

Total Duration: 35.72 sec.

Single Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.6078	0.5916	0.5996	3751
white	0.7407	0.0548	0.1020	730
gray	0.2232	0.6136	0.3273	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0004	0.0009	2321
orange	0.7500	0.1395	0.2353	43
purple	0.0071	0.0435	0.0122	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1892	0.4416	0.2649	548
brown	0.4202	0.5351	0.4707	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3727	0.3727	0.3727	13513
macro avg	0.3282	0.2017	0.1677	13513
weighted avg	0.5083	0.3727	0.3242	13513



Single Color with RGB / Pairwise Third Quantile Confusion Matrix

5.1.3. CNN AlexNet with Multi-color

Refined Dataset Improvement Run

CNN AlexNet Multi-Color with RGB

Total Duration: 2281.19 sec.

Learning Rate: 0.000070

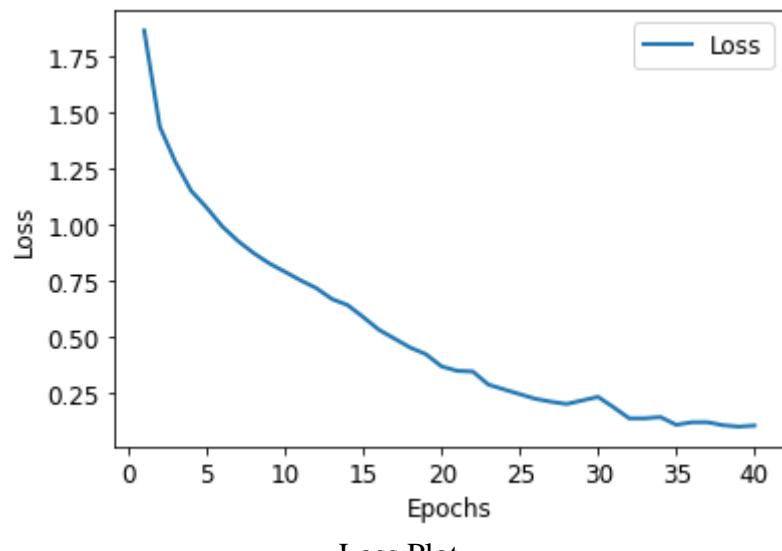
Dropout: 0.10

Batch Size: 32

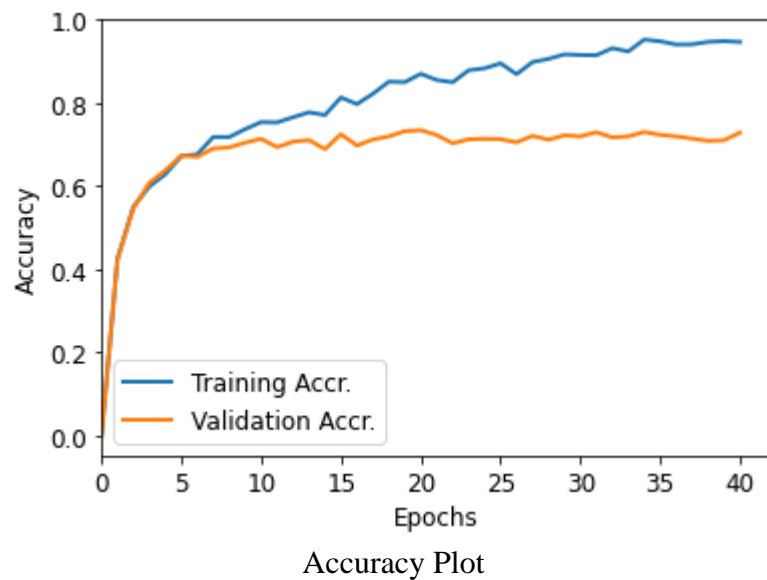
Epoch: 40

Best Performing Epoch: 20

Summary	
Method	Accuracy
Training Dataset	0.8705
Validation Dataset	0.7348
Test Dataset	0.7241



Loss Plot

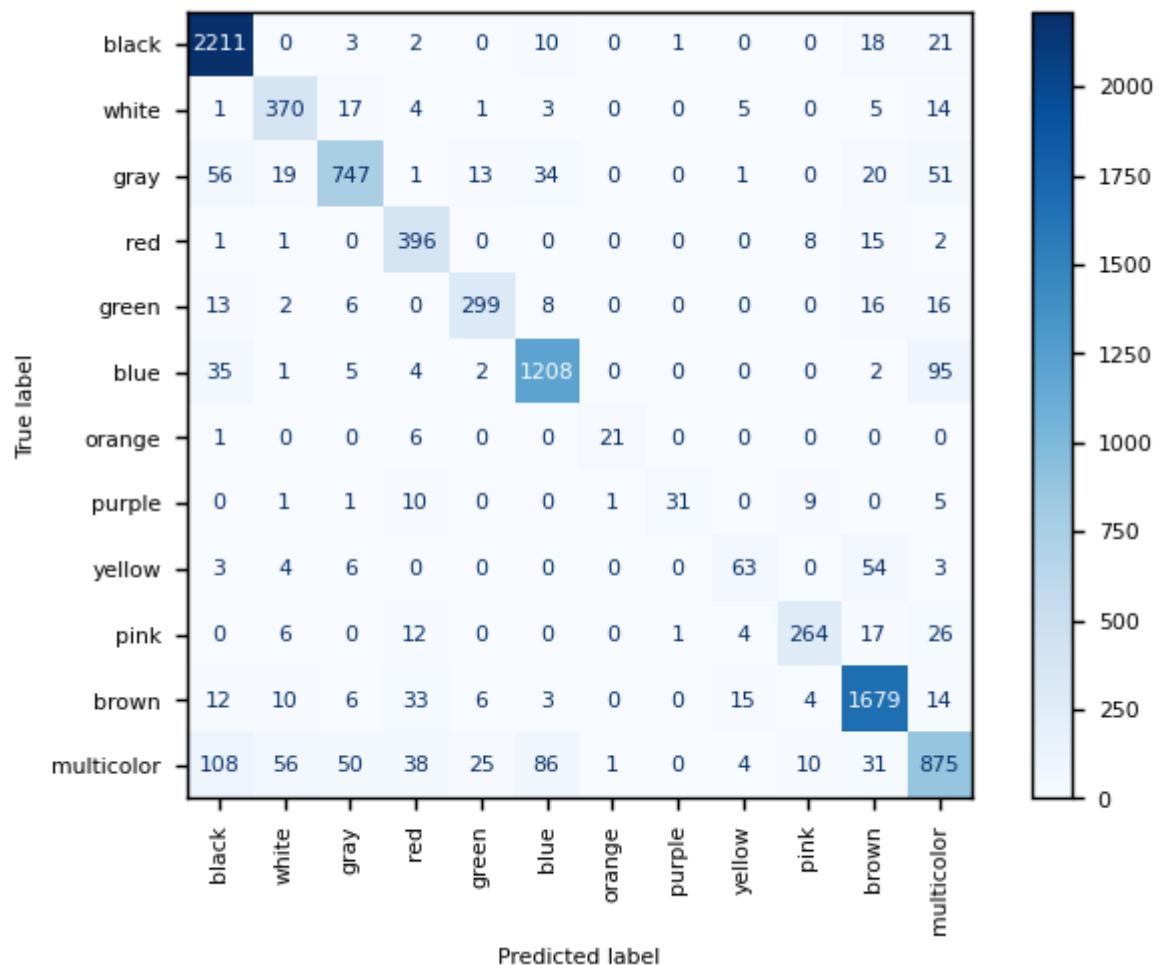


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.9058	0.9757	0.9395	2266
white	0.7872	0.8810	0.8315	420
gray	0.8882	0.7930	0.8379	942
red	0.7826	0.9362	0.8525	423
green	0.8642	0.8306	0.8470	360
blue	0.8935	0.8935	0.8935	1352
orange	0.9130	0.7500	0.8235	28
purple	0.9394	0.5345	0.6813	58
yellow	0.6848	0.4737	0.5600	133
pink	0.8949	0.8000	0.8448	330
brown	0.9041	0.9422	0.9228	1782
multicolor	0.7799	0.6815	0.7273	1284
accuracy			0.8705	
macro avg	0.8531	0.7910	0.8135	9378
weighted avg	0.8689	0.8705	0.8675	9378

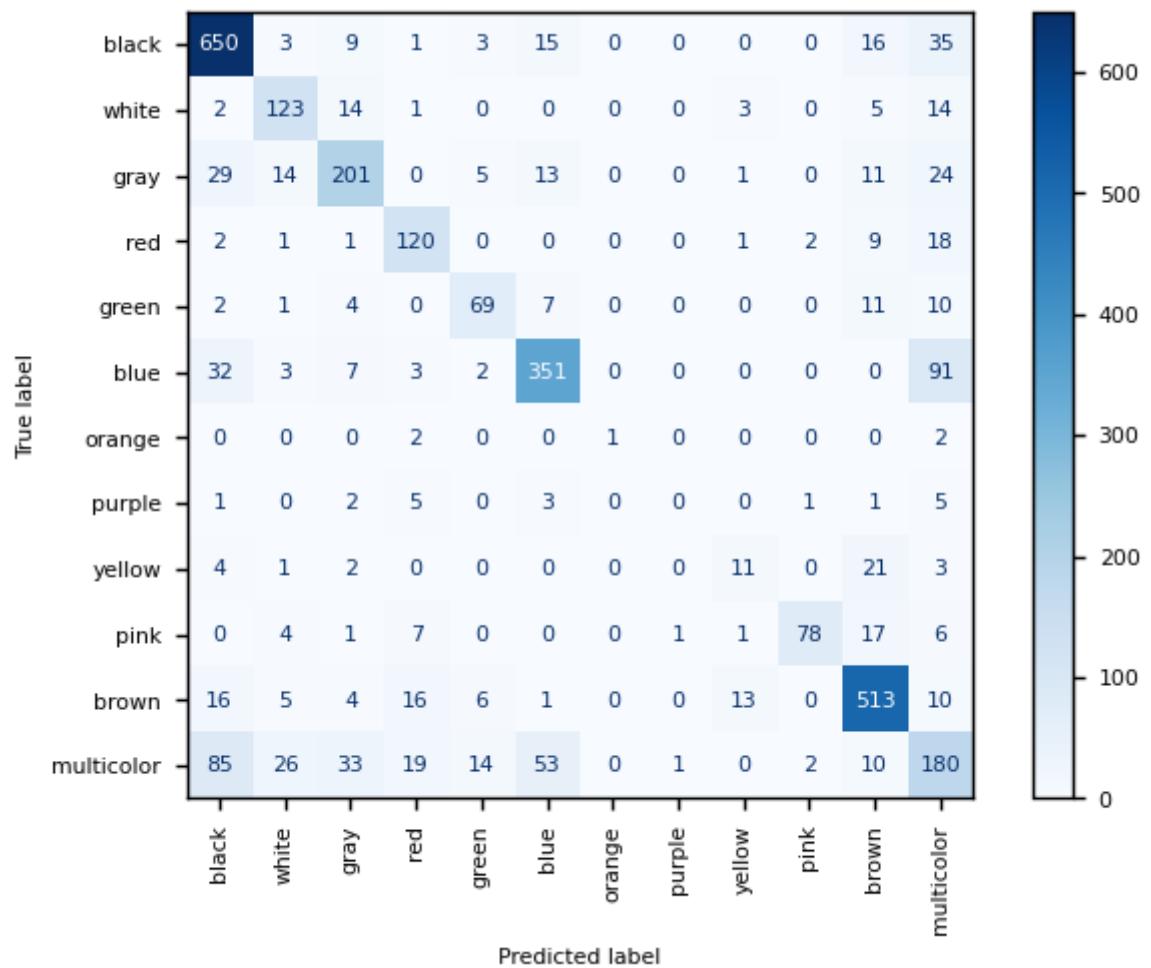


Training Dataset Confusion Matrix

Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7898	0.8880	0.8360	732
white	0.6796	0.7593	0.7172	162
gray	0.7230	0.6745	0.6979	298
red	0.6897	0.7792	0.7317	154
green	0.6970	0.6635	0.6798	104
blue	0.7923	0.7178	0.7532	489
orange	1.0000	0.2000	0.3333	5
purple	0.0000	0.0000	0.0000	18
yellow	0.3667	0.2619	0.3056	42
pink	0.9398	0.6783	0.7879	115
brown	0.8355	0.8784	0.8564	584
multicolor	0.4523	0.4255	0.4385	423
accuracy			0.7348	
macro avg	0.6638	0.5772	0.5948	3126
weighted avg	0.7286	0.7348	0.7289	3126

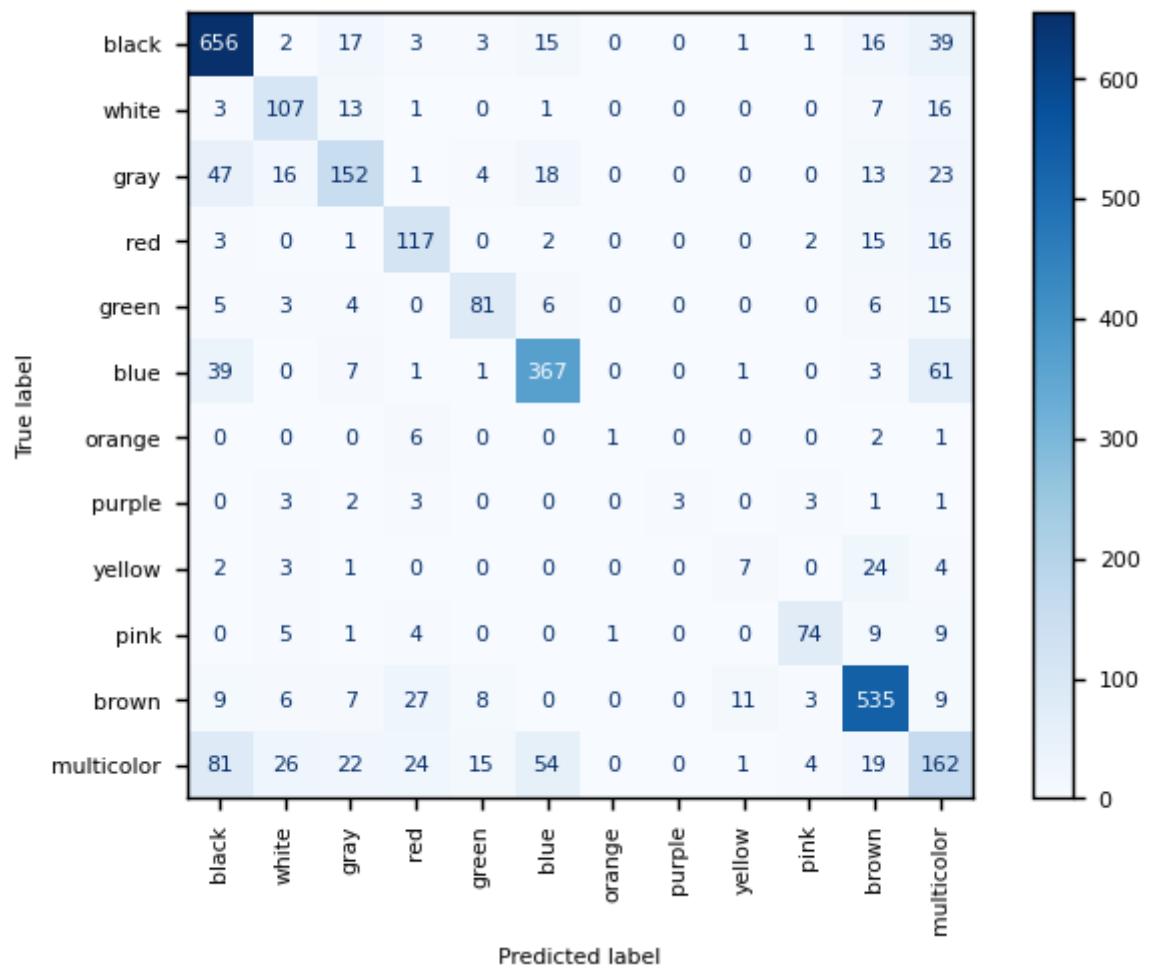


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7763	0.8712	0.8210	753
white	0.6257	0.7230	0.6708	148
gray	0.6696	0.5547	0.6068	274
red	0.6257	0.7500	0.6822	156
green	0.7232	0.6750	0.6983	120
blue	0.7927	0.7646	0.7784	480
orange	0.5000	0.1000	0.1667	10
purple	1.0000	0.1875	0.3158	16
yellow	0.3333	0.1707	0.2258	41
pink	0.8506	0.7184	0.7789	103
brown	0.8231	0.8699	0.8458	615
multicolor	0.4551	0.3971	0.4241	408
accuracy			0.7241	
macro avg	0.6813	0.5652	0.5846	3124
weighted avg	0.7169	0.7241	0.7161	3124



Test Dataset Confusion Matrix

gray (66.6%) black (99.8%) blue (100.0%) brown (73.8%) white (99.3%)
 >> blue >> black >> blue >> yellow >> white



multicolor (49.4%) black (97.0%) brown (73.7%) brown (99.9%) multicolor (61.0%) black (83.9%)
 >> multicolor >> black >> brown >> brown >> gray >> black



black (100.0%) gray (92.7%) gray (57.5%) black (98.6%) black (99.9%) brown (99.8%)
 >> black >> gray >> black >> multicolor >> black >> brown



black (90.7%) multicolor (99.4%) black (100.0%) blue (93.0%) multicolor (60.1%) brown (99.9%)
 >> blue >> black >> black >> blue >> brown >> brown



gray (40.0%) red (96.6%) white (99.5%) multicolor (39.6%) black (100.0%) blue (90.0%)
 >> white >> red >> white >> multicolor >> black >> blue



blue (91.9%) brown (100.0%) brown (56.3%) black (100.0%) red (70.2%) multicolor (36.0%)
 >> green >> brown >> yellow >> black >> red >> green



Sample Images from Test Dataset

multicolor (55.3%) gray (75.2%) gray (64.9%) multicolor (61.0%) black (73.0%) gray (57.5%)
 >> black >> multicolor >> blue >> gray >> multicolor >> black



multicolor (51.7%) white (72.7%) multicolor (95.3%) brown (73.8%) multicolor (60.1%) black (98.6%)
 >> gray >> multicolor >> red >> yellow >> brown >> multicolor



brown (71.1%) white (98.2%) multicolor (90.7%) (81.8%) red (97.6%) multicolor (99.4%)
 >> black >> brown >> black >> brown >> brown >> black



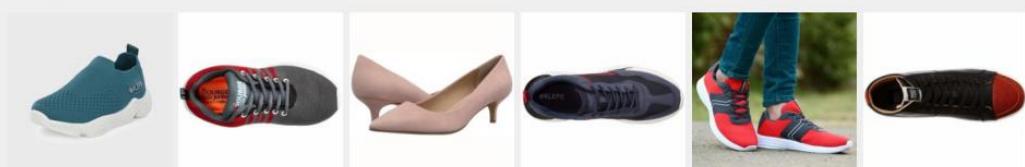
black (96.7%) yellow (65.4%) blue (49.6%) multicolor (98.9%) black (95.2%) pink (89.3%)
 >> gray >> brown >> gray >> gray >> red >> red



yellow (48.0%) blue (69.6%) brown (49.6%) brown (95.9%) gray (66.6%) black (95.9%)
 >> brown >> multicolor >> yellow >> gray >> blue >> blue



multicolor (61.1%) blue (65.0%) pink (77.8%) blue (82.2%) red (86.8%) multicolor (90.8%)
 >> green >> multicolor >> red >> multicolor >> orange >> black



Sample Images that cannot be accurately predicted in the Test Dataset

5.1.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Quantization Fastoctree	41.71 sec.	0.3641
Quantization Mmcq	630.72 sec.	0.3448
Quantization Libimagequant	264.62 sec.	0.2914
Pairwise Argmin	34.44 sec.	0.2859
Pairwise Mean	12.97 sec.	0.3140
Array Mean Argmin	14.91 sec.	0.1259
Pairwise First Quantile	54.61 sec.	0.3388
Pairwise Median	54.74 sec.	0.3492
Pairwise Third Quantile	52.13 sec.	0.3246

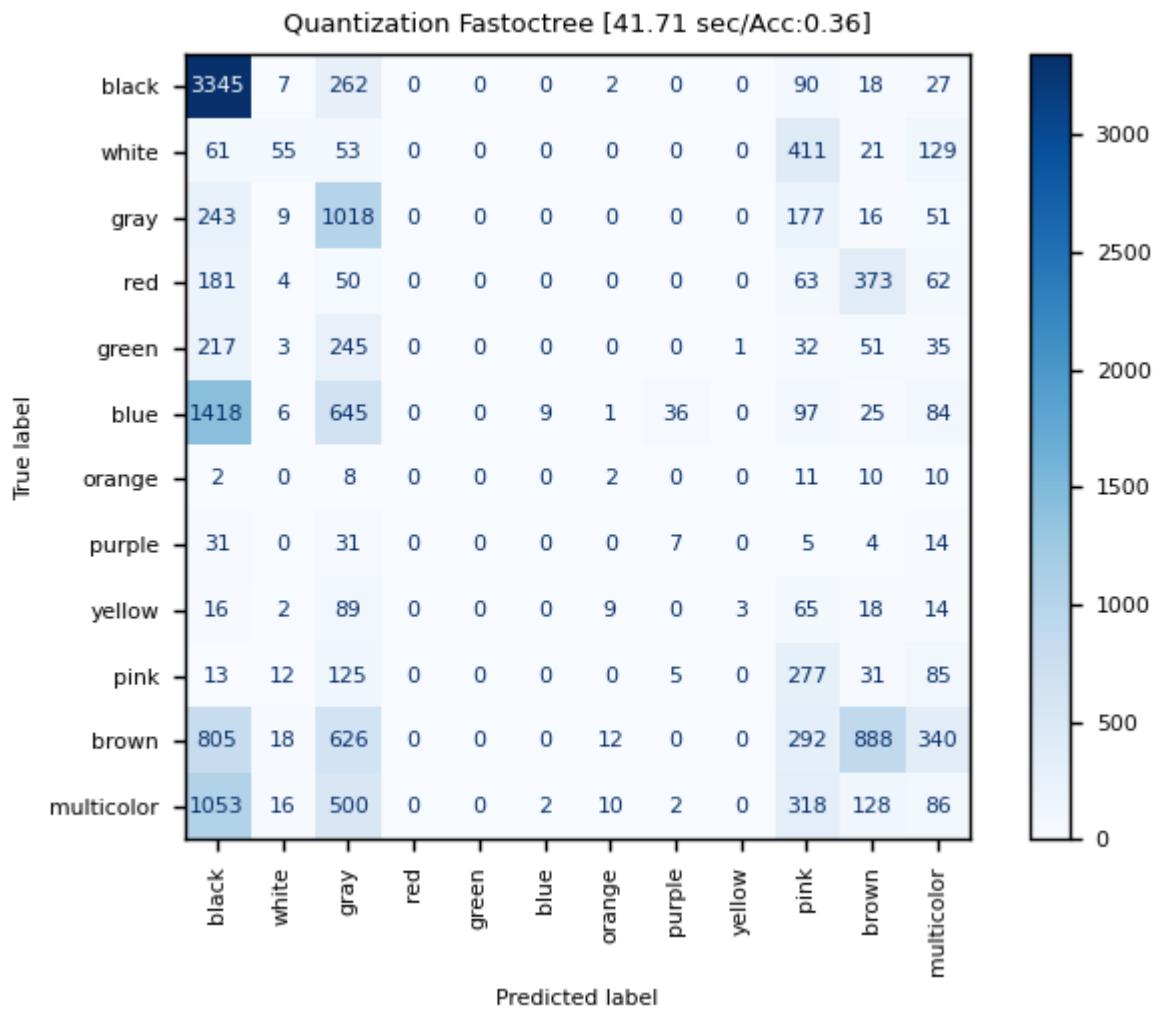
Refined Dataset Improvement Run

Multi-Color with RGB / Quantization Fastoctree Result

Total Duration: 41.71 sec.

Multi-Color with RGB / Quantization Fastoctree Classification Report

	Precision	Recall	F1 Score	Support
black	0.4529	0.8918	0.6008	3751
white	0.4167	0.0753	0.1276	730
gray	0.2788	0.6724	0.3941	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.8182	0.0039	0.0077	2321
orange	0.0556	0.0465	0.0506	43
purple	0.1400	0.0761	0.0986	92
yellow	0.7500	0.0139	0.0273	216
pink	0.1507	0.5055	0.2322	548
brown	0.5610	0.2979	0.3891	2981
multicolor	0.0918	0.0407	0.0564	2115
accuracy			0.3641	
macro avg	0.3096	0.2187	0.1654	15628
weighted avg	0.4127	0.3641	0.2806	15628



Multi-Color with RGB / Quantization Fastoctree Confusion Matrix

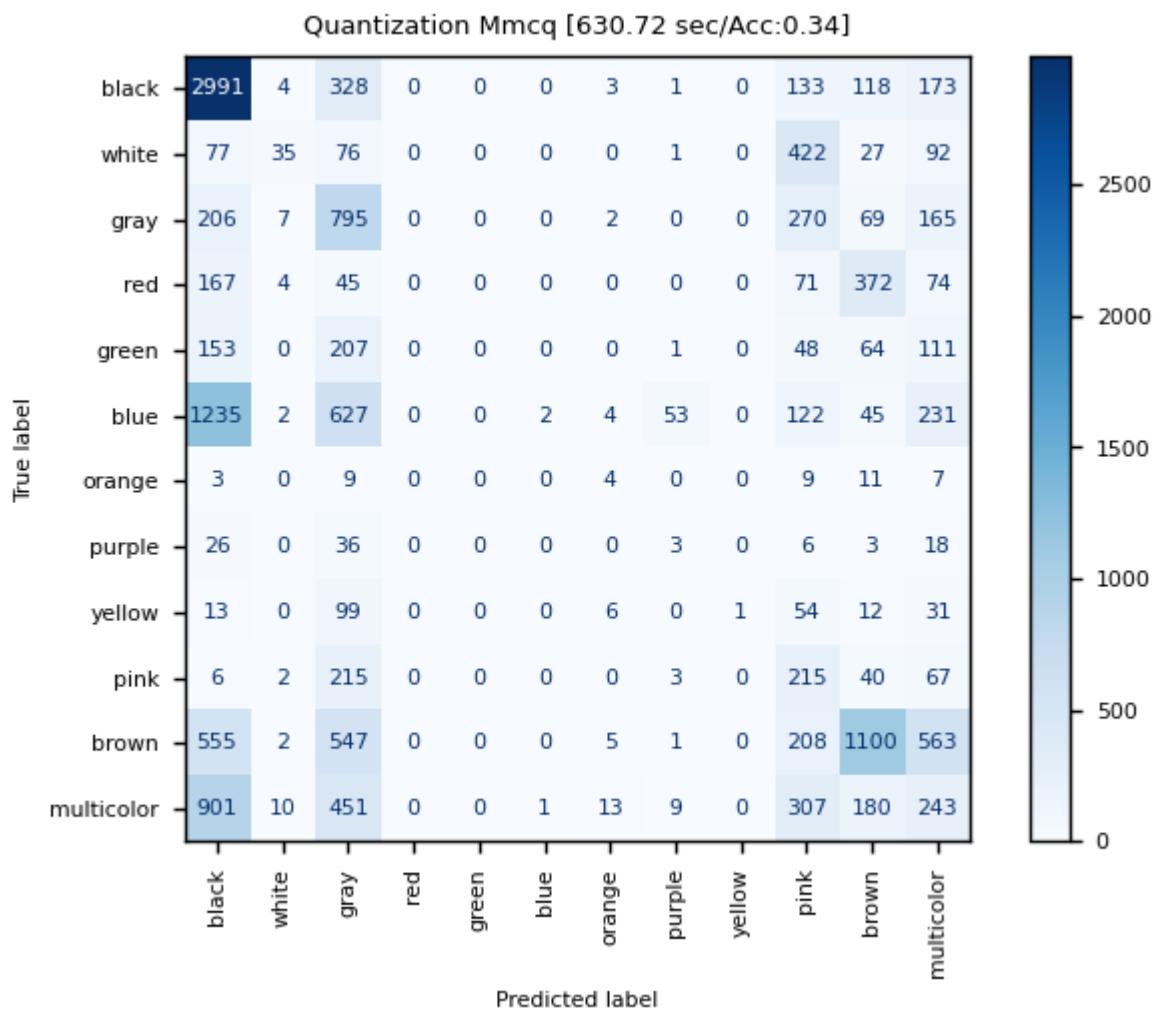
Refined Dataset Improvement Run

Multi-Color with RGB / Quantization Mmcq Result

Total Duration: 630.72 sec.

Multi-Color with RGB / Quantization Mmcq Classification Report

	Precision	Recall	F1 Score	Support
black	0.4723	0.7974	0.5932	3751
white	0.5303	0.0479	0.0879	730
gray	0.2314	0.5251	0.3213	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.6667	0.0009	0.0017	2321
orange	0.1081	0.0930	0.1000	43
purple	0.0417	0.0326	0.0366	92
yellow	1.0000	0.0046	0.0092	216
pink	0.1153	0.3923	0.1782	548
brown	0.5390	0.3690	0.4381	2981
multicolor	0.1369	0.1149	0.1249	2115
accuracy			0.3448	
macro avg	0.3201	0.1981	0.1576	15628
weighted avg	0.3993	0.3448	0.2852	15628



Multi-Color with RGB / Quantization Mmcq Confusion Matrix

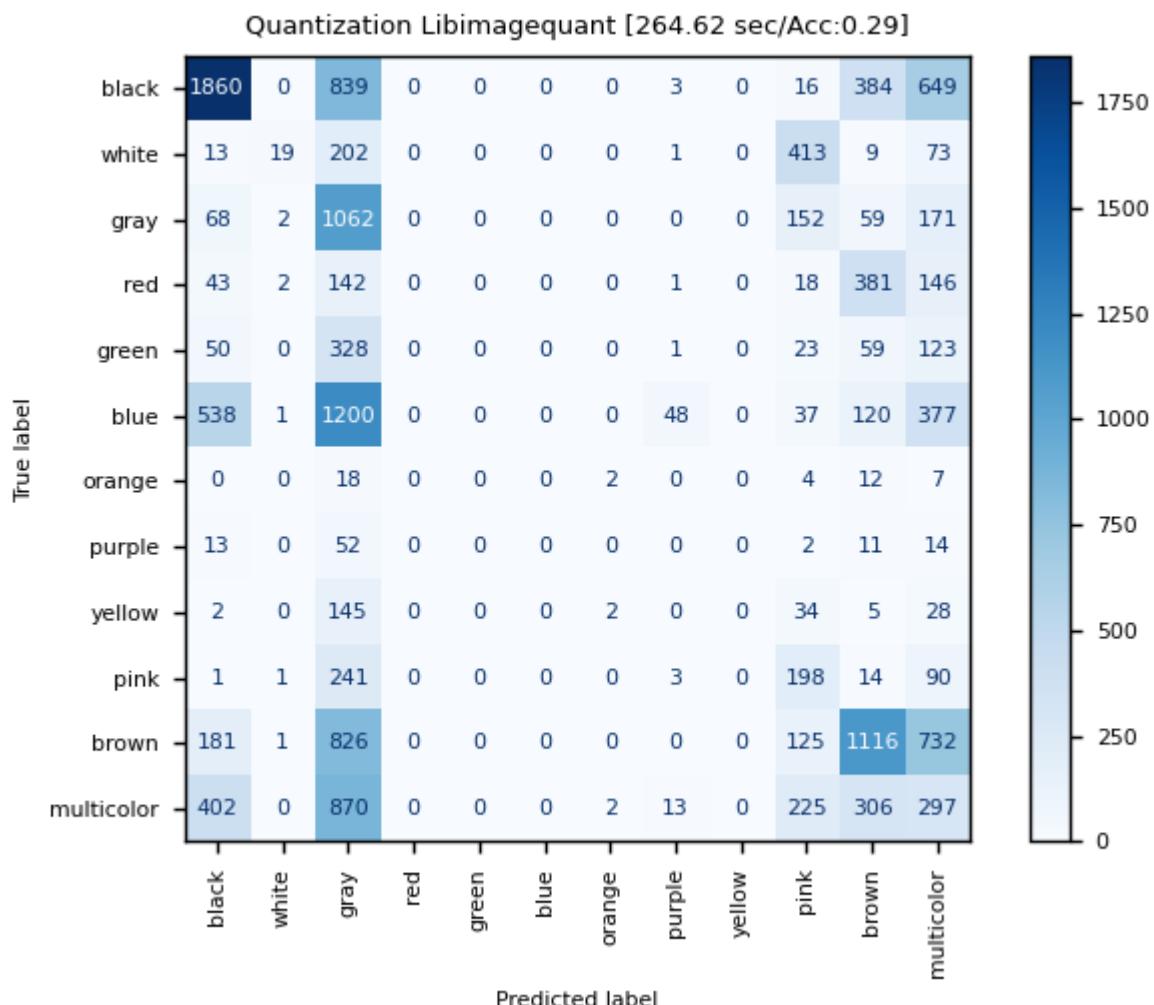
Refined Dataset Improvement Run

Multi-Color with RGB / Quantization Libimagequant Result

Total Duration: 264.62 sec.

Multi-Color with RGB / Quantization Libimagequant Classification Report

	Precision	Recall	F1 Score	Support
black	0.5866	0.4959	0.5374	3751
white	0.7308	0.0260	0.0503	730
gray	0.1792	0.7015	0.2855	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.3333	0.0465	0.0816	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1588	0.3613	0.2206	548
brown	0.4507	0.3744	0.4090	2981
multicolor	0.1097	0.1404	0.1232	2115
accuracy			0.2914	
macro avg	0.2124	0.1788	0.1423	15628
weighted avg	0.2996	0.2914	0.2616	15628



Multi-Color with RGB / Quantization Libimagequant Confusion Matrix

Refined Dataset Improvement Run

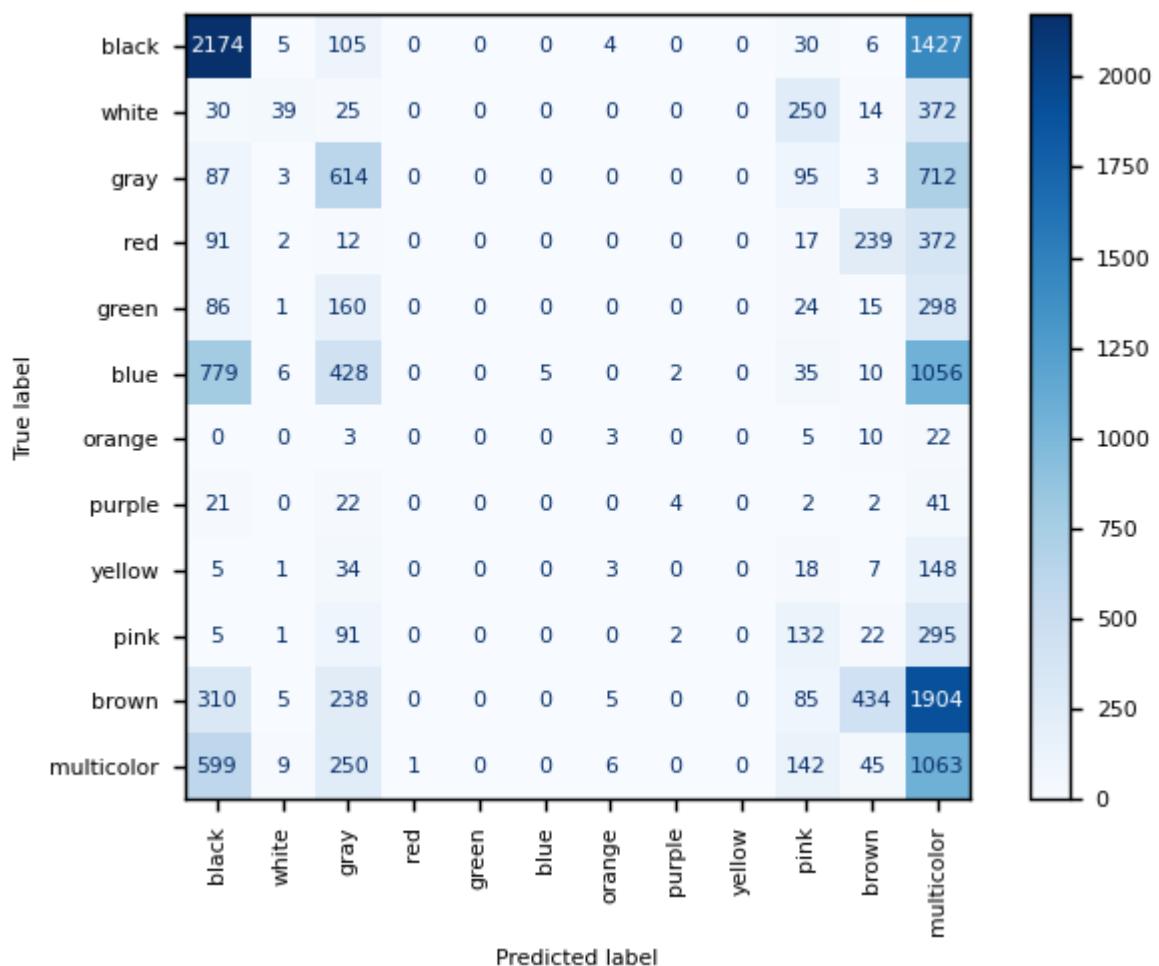
Multi-Color with RGB / Pairwise Argmin Result

Total Duration: 34.44 sec.

Multi-Color with RGB / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5192	0.5796	0.5477	3751
white	0.5417	0.0534	0.0973	730
gray	0.3098	0.4055	0.3513	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0022	0.0043	2321
orange	0.1429	0.0698	0.0937	43
purple	0.5000	0.0435	0.0800	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1581	0.2409	0.1909	548
brown	0.5378	0.1456	0.2291	2981
multicolor	0.1379	0.5026	0.2164	2115
accuracy			0.2859	
macro avg	0.3206	0.1703	0.1509	15628
weighted avg	0.4586	0.2859	0.2511	15628

Pairwise Argmin [34.44 sec/Acc:0.29]



Multi-Color with RGB / Pairwise Argmin Confusion Matrix

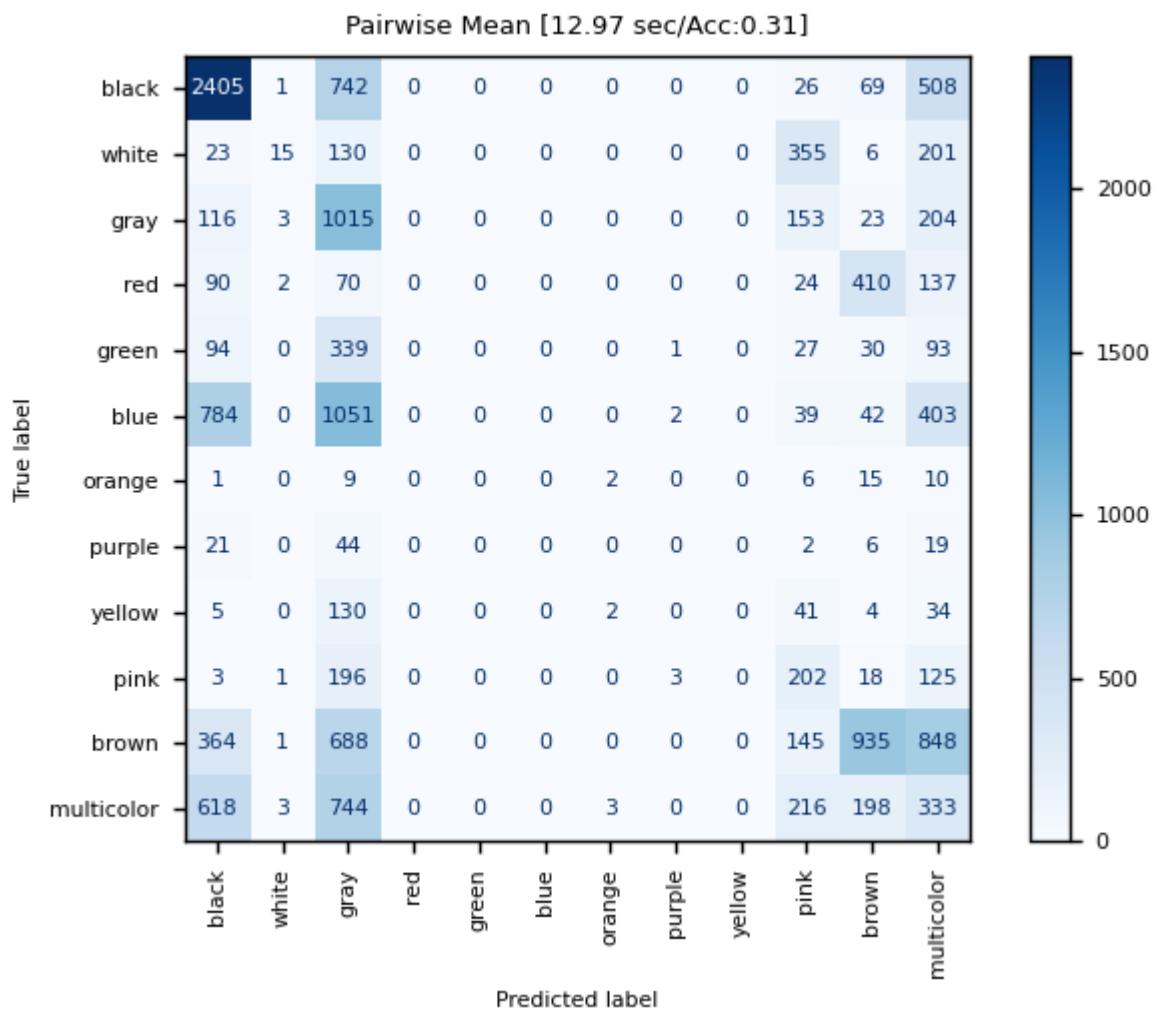
Refined Dataset Improvement Run

Multi-Color with RGB / Pairwise Mean Result

Total Duration: 12.97 sec.

Multi-Color with RGB / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.5316	0.6412	0.5813	3751
white	0.5769	0.0205	0.0397	730
gray	0.1968	0.6704	0.3043	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.2857	0.0465	0.0800	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1634	0.3686	0.2265	548
brown	0.5325	0.3137	0.3948	2981
multicolor	0.1142	0.1574	0.1324	2115
accuracy			0.3140	
macro avg	0.2001	0.1849	0.1466	15628
weighted avg	0.2972	0.3140	0.2722	15628



Multi-Color with RGB / Pairwise Mean Confusion Matrix

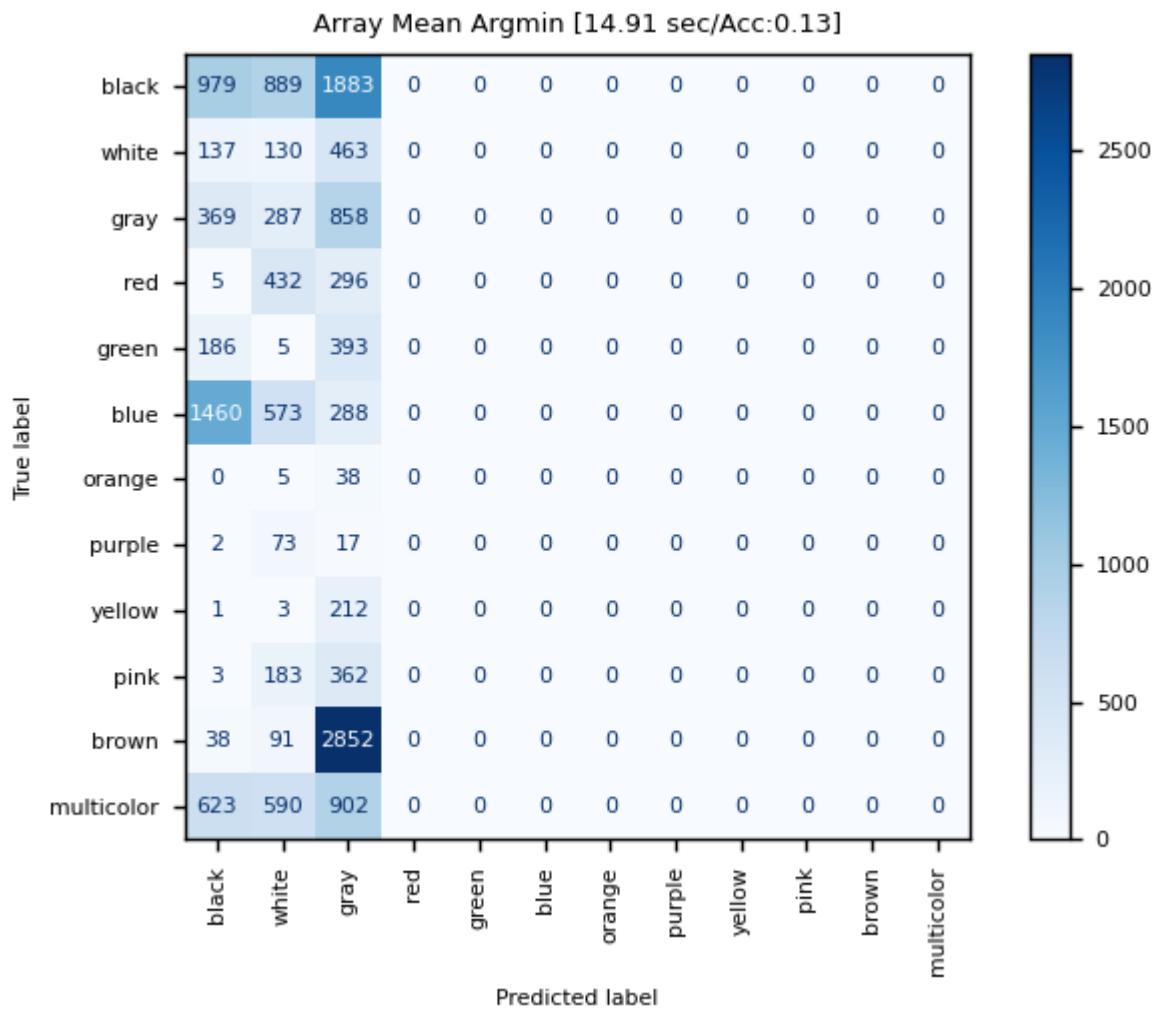
Refined Dataset Improvement Run

Multi-Color with RGB / Array Mean Argmin Result

Total Duration: 14.91 sec.

Multi-Color with RGB / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2574	0.2610	0.2592	3751
white	0.0399	0.1781	0.0651	730
gray	0.1002	0.5667	0.1703	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.0000	0.0000	0.0000	548
brown	0.0000	0.0000	0.0000	2981
multicolor	0.0000	0.0000	0.0000	2115
accuracy			0.1259	
macro avg	0.0331	0.0838	0.0412	15628
weighted avg	0.0734	0.1259	0.0818	15628



Multi-Color with RGB / Array Mean Argmin Confusion Matrix

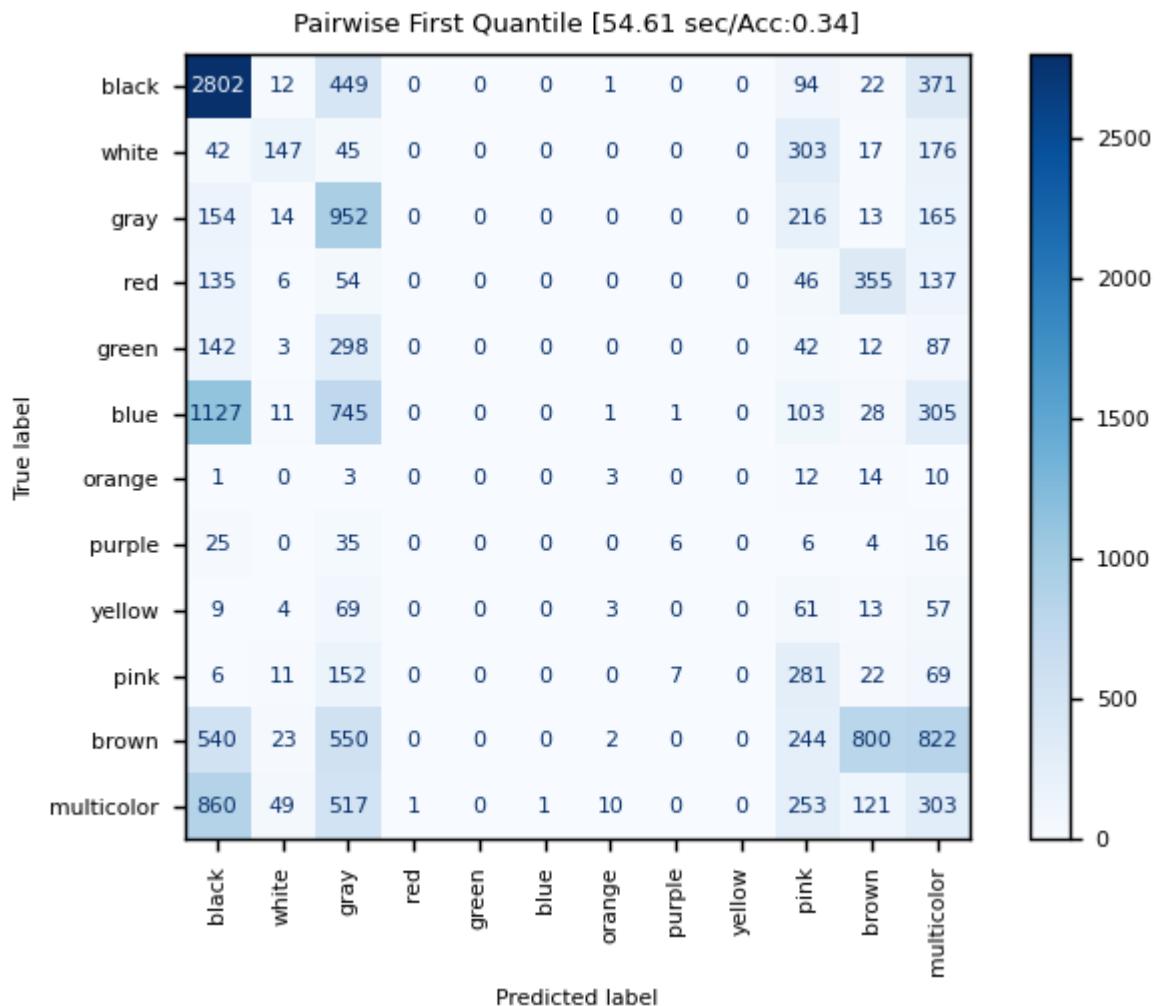
Refined Dataset Improvement Run

Multi-Color with RGB / Pairwise First Quantile Result

Total Duration: 54.61 sec.

Multi-Color with RGB / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4795	0.7470	0.5841	3751
white	0.5250	0.2014	0.2911	730
gray	0.2461	0.6288	0.3537	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.1500	0.0698	0.0952	43
purple	0.4286	0.0652	0.1132	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1692	0.5128	0.2544	548
brown	0.5630	0.2684	0.3635	2981
multicolor	0.1203	0.1433	0.1308	2115
accuracy			0.3388	
macro avg	0.2235	0.2197	0.1822	15628
weighted avg	0.2960	0.3388	0.2849	15628



Multi-Color with RGB / Pairwise First Quantile Confusion Matrix

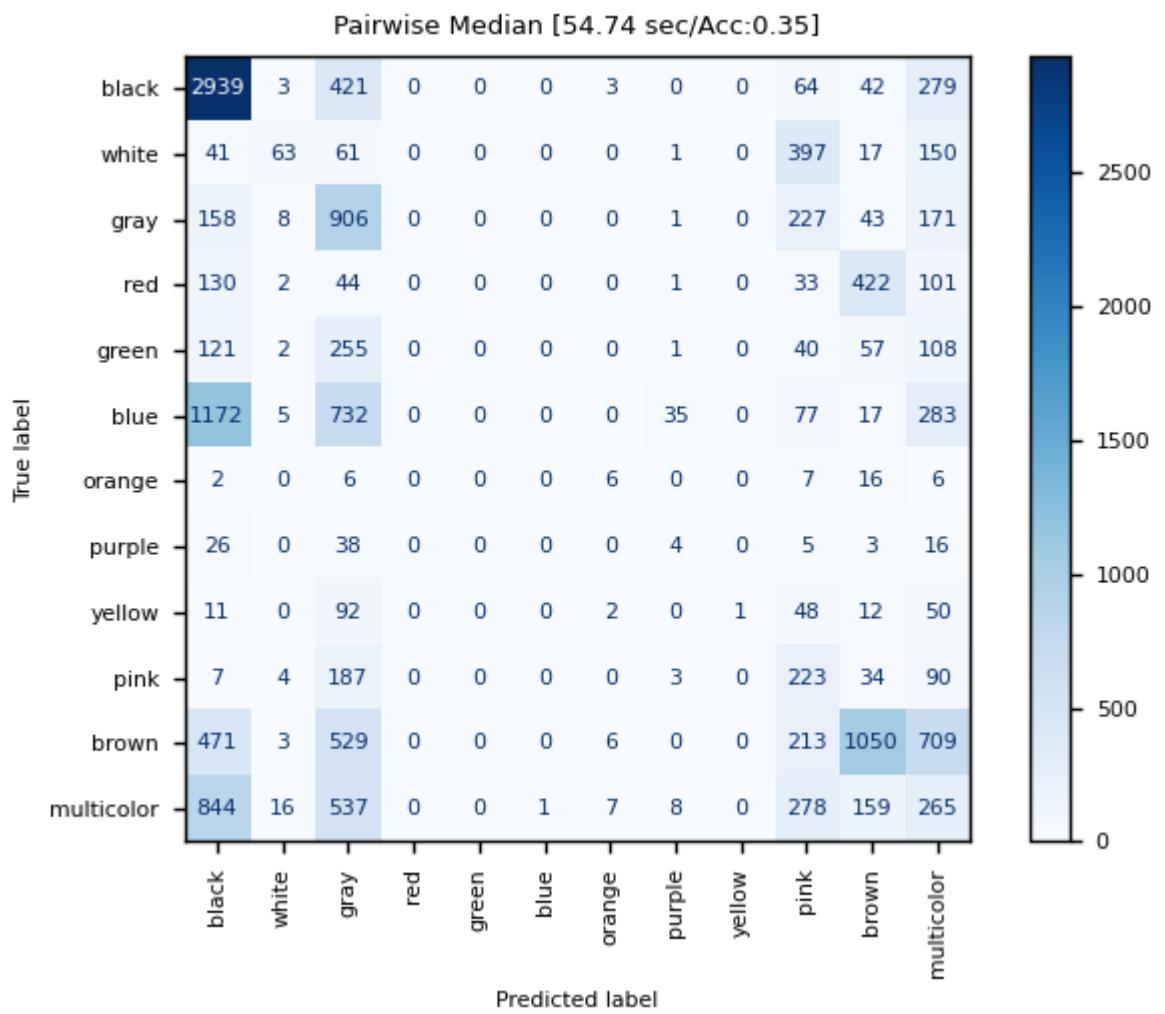
Refined Dataset Improvement Run

Multi-Color with RGB / Pairwise Median Result

Total Duration: 54.74 sec.

Multi-Color with RGB / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4963	0.7835	0.6077	3751
white	0.5943	0.0863	0.1507	730
gray	0.2379	0.5984	0.3405	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.2500	0.1395	0.1791	43
purple	0.0741	0.0435	0.0548	92
yellow	1.0000	0.0046	0.0092	216
pink	0.1383	0.4069	0.2065	548
brown	0.5609	0.3522	0.4327	2981
multicolor	0.1189	0.1253	0.1220	2115
accuracy			0.3492	
macro avg	0.2892	0.2117	0.1753	15628
weighted avg	0.3128	0.3492	0.2931	15628



Multi-Color with RGB / Pairwise Median Confusion Matrix

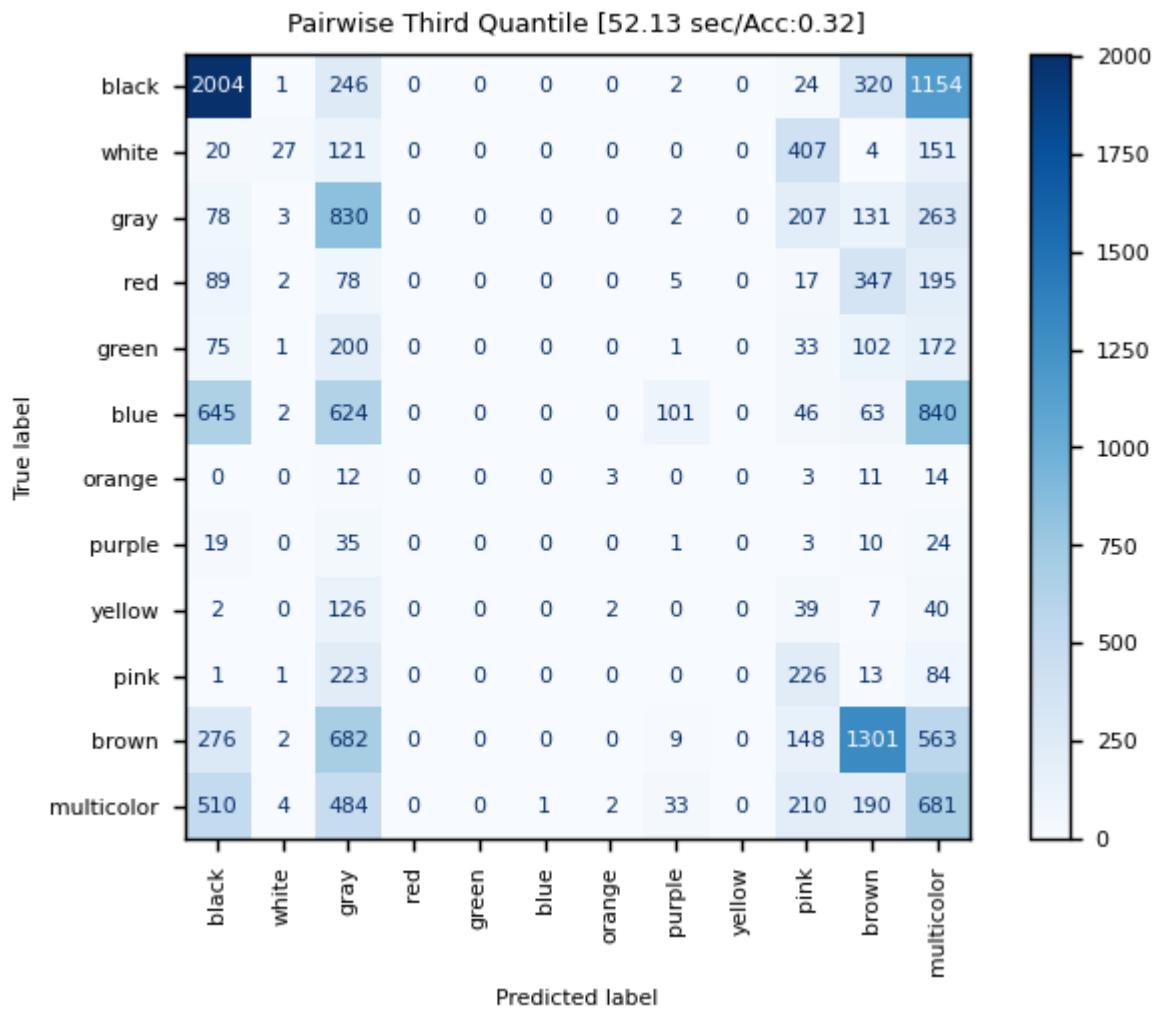
Refined Dataset Improvement Run

Multi-Color with RGB / Pairwise Third Quantile Result

Total Duration: 52.13 sec.

Multi-Color with RGB / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5389	0.5343	0.5365	3751
white	0.6279	0.0370	0.0699	730
gray	0.2267	0.5482	0.3208	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.4286	0.0698	0.1200	43
purple	0.0065	0.0109	0.0081	92
yellow	0.0000	0.0000	0.0000	216
pink	0.1658	0.4124	0.2365	548
brown	0.5206	0.4364	0.4748	2981
multicolor	0.1629	0.3220	0.2163	2115
accuracy			0.3246	
macro avg	0.2232	0.1976	0.1652	15628
weighted avg	0.3090	0.3246	0.2916	15628



Multi-Color with RGB / Pairwise Third Quantile Confusion Matrix

5.2. HSL Color Space

5.2.1. CNN AlexNet with Single Color

Refined Dataset Improvement Run

CNN AlexNet Single Color with HSL

Total Duration: 3643.50 sec.

Learning Rate: 0.000070

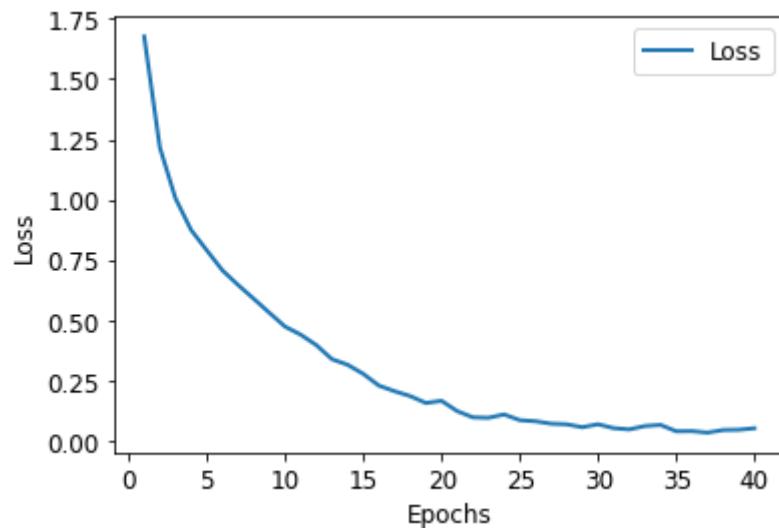
Dropout: 0.10

Batch Size: 32

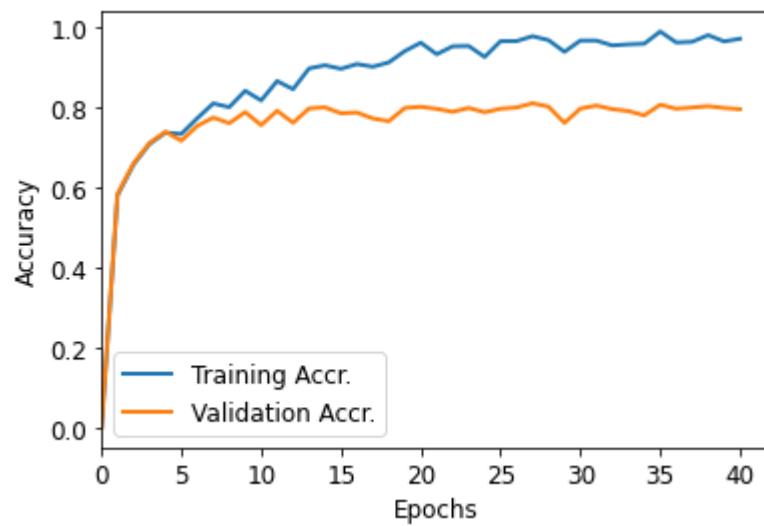
Epoch: 40

Best Performing Epoch: 27

Summary	
Method	Accuracy
Training Dataset	0.9785
Validation Dataset	0.8110
Test Dataset	0.8085



Loss Plot

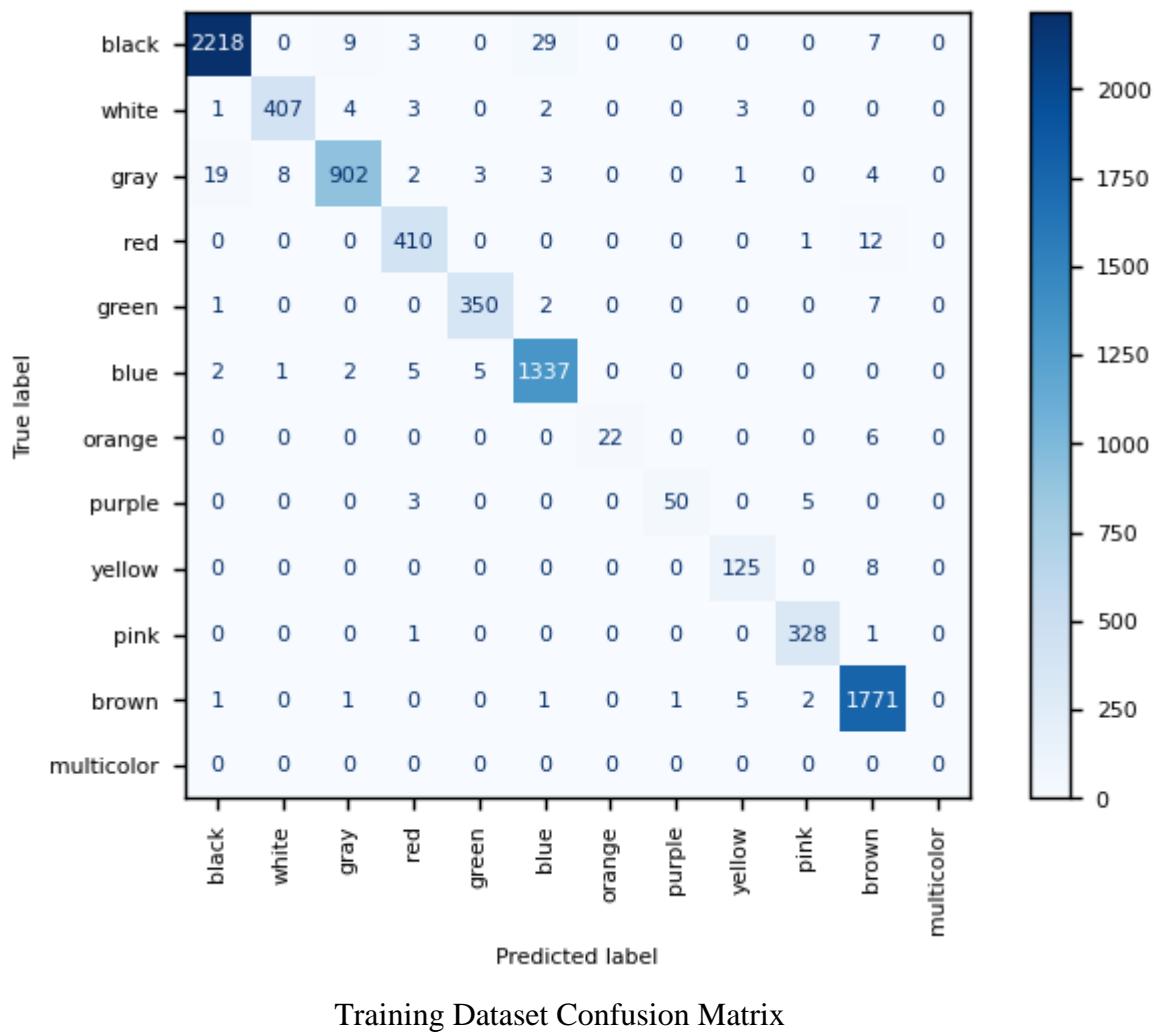


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

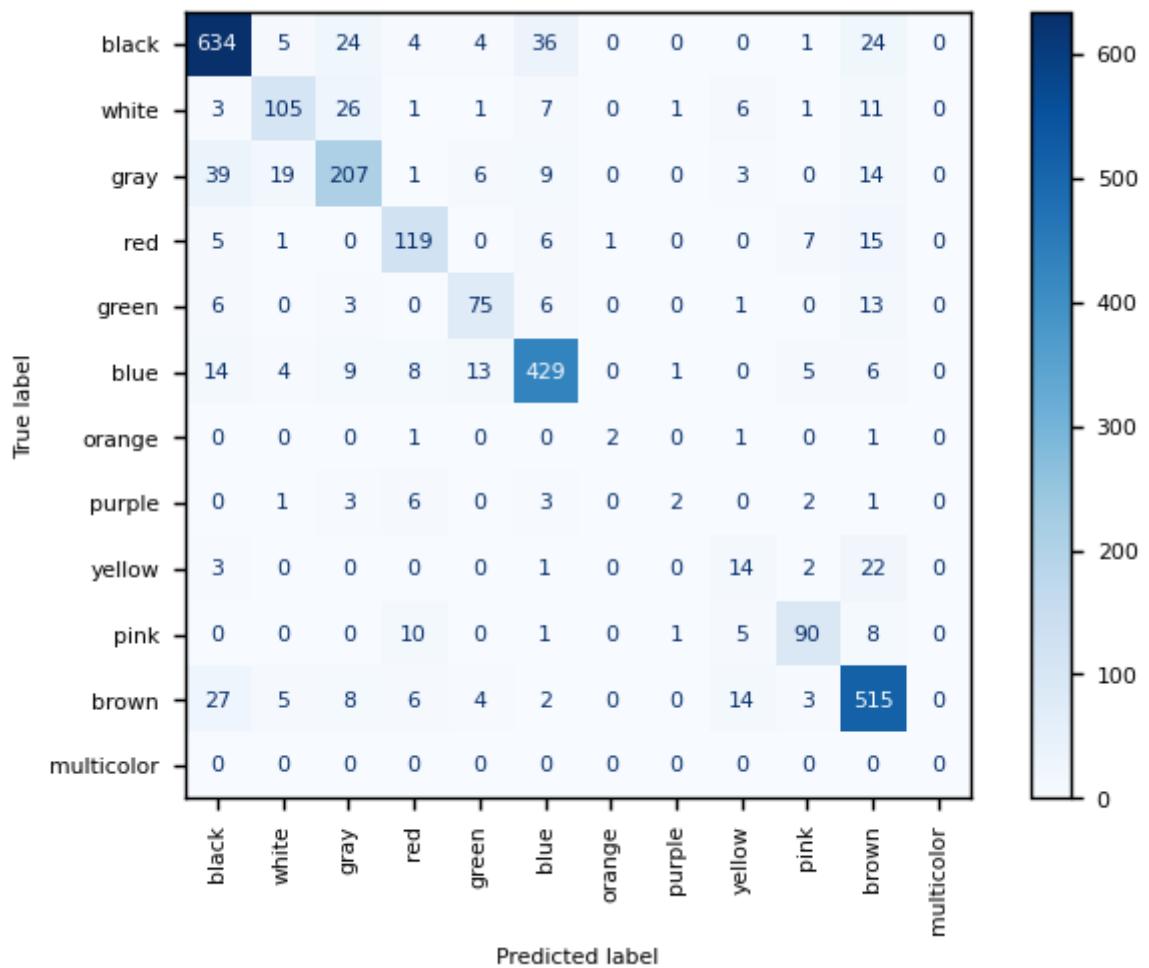
	Precision	Recall	F1 Score	Support
black	0.9893	0.9788	0.9840	2266
white	0.9784	0.9690	0.9737	420
gray	0.9826	0.9575	0.9699	942
red	0.9602	0.9693	0.9647	423
green	0.9777	0.9722	0.9749	360
blue	0.9731	0.9889	0.9809	1352
orange	1.0000	0.7857	0.8800	28
purple	0.9804	0.8621	0.9174	58
yellow	0.9328	0.9398	0.9363	133
pink	0.9762	0.9939	0.9850	330
brown	0.9752	0.9938	0.9844	1782
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.9785	0.9785	0.9785	8094
macro avg	0.8938	0.8676	0.8793	8094
weighted avg	0.9786	0.9785	0.9784	8094



Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8673	0.8661	0.8667	732
white	0.7500	0.6481	0.6954	162
gray	0.7393	0.6946	0.7163	298
red	0.7628	0.7727	0.7677	154
green	0.7282	0.7212	0.7246	104
blue	0.8580	0.8773	0.8675	489
orange	0.6667	0.4000	0.5000	5
purple	0.4000	0.1111	0.1739	18
yellow	0.3182	0.3333	0.3256	42
pink	0.8108	0.7826	0.7965	115
brown	0.8175	0.8818	0.8484	584
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.8110	0.8110	0.8110	2703
macro avg	0.6432	0.5907	0.6069	2703
weighted avg	0.8080	0.8110	0.8083	2703

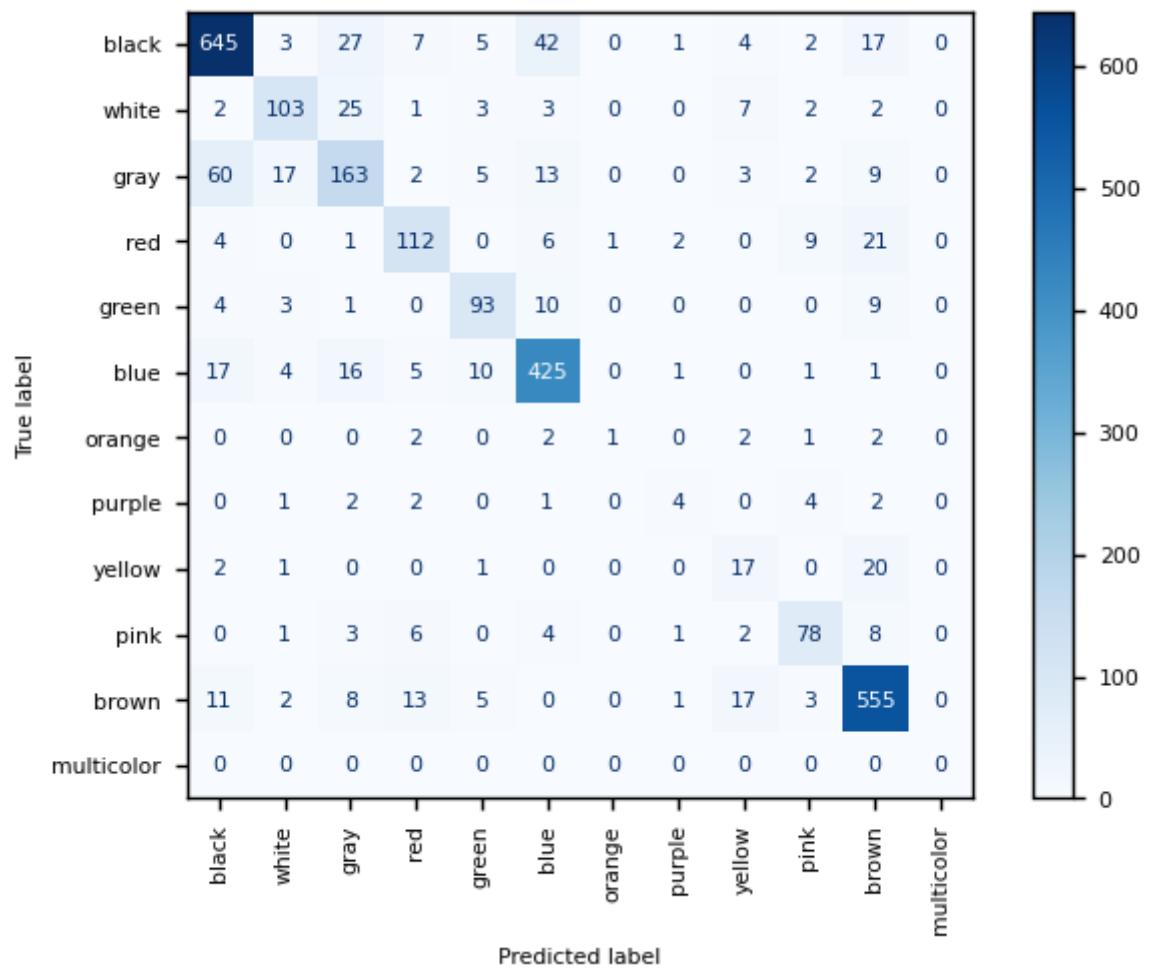


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.8658	0.8566	0.8611	753
white	0.7630	0.6959	0.7279	148
gray	0.6626	0.5949	0.6269	274
red	0.7467	0.7179	0.7320	156
green	0.7623	0.7750	0.7686	120
blue	0.8399	0.8854	0.8621	480
orange	0.5000	0.1000	0.1667	10
purple	0.4000	0.2500	0.3077	16
yellow	0.3269	0.4146	0.3656	41
pink	0.7647	0.7573	0.7610	103
brown	0.8591	0.9024	0.8803	615
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.8085	0.8085	0.8085	2716
macro avg	0.6242	0.5792	0.5883	2716
weighted avg	0.8061	0.8085	0.8061	2716



Test Dataset Confusion Matrix

red (100.0%) brown (100.0%) red (100.0%) gray (95.6%) white (100.0%) black (100.0%)
 >> red >> brown >> red >> gray >> white >> black



blue (98.5%) black (69.5%) black (99.5%) brown (99.3%) brown (100.0%) black (100.0%)
 >> blue >> gray >> black >> pink >> brown >> black



green (99.2%) black (99.9%) brown (100.0%) black (100.0%) yellow (53.1%) blue (54.4%)
 >> green >> black >> brown >> black >> brown >> red



white (99.7%) white (99.9%) black (100.0%) black (100.0%) brown (79.1%) black (100.0%)
 >> white >> white >> black >> black >> yellow >> black



gray (99.9%) blue (96.7%) blue (100.0%) blue (100.0%) black (100.0%) black (60.4%)
 >> gray >> blue >> blue >> blue >> black >> black



brown (100.0%) black (100.0%) gray (76.3%) black (100.0%) gray (100.0%) black (99.8%)
 >> brown >> black >> gray >> black >> gray >> black



Sample Images from Test Dataset

gray (97.9%)
 >> white
 black (97.8%)
 >> gray
 brown (79.1%)
 >> yellow
 blue (73.2%)
 >> pink
 gray (81.9%)
 >> black
 green (63.0%)
 >> blue



blue (99.6%)
 >> green
 green (50.7%)
 >> gray
 yellow (88.8%)
 >> brown
 pink (97.7%)
 >> blue
 black (93.9%)
 >> blue
 red (82.2%)
 >> brown



white (99.2%)
 >> gray
 white (96.5%)
 >> gray
 white (98.1%)
 >> black
 black (69.5%)
 >> gray
 yellow (70.5%)
 >> brown
 blue (97.5%)
 >> black



yellow (53.1%)
 >> brown
 pink (98.1%)
 >> red
 yellow (57.4%)
 >> white
 pink (76.2%)
 >> red
 brown (99.3%)
 >> pink
 brown (37.6%)
 >> white



blue (54.4%)
 >> red
 red (97.0%)
 >> brown
 pink (73.7%)
 >> red
 purple (53.2%)
 >> red
 brown (82.4%)
 >> red
 blue (99.5%)
 >> red



blue (92.5%)
 >> gray
 brown (66.9%)
 >> gray
 red (99.8%)
 >> gray
 black (93.8%)
 >> red
 blue (51.9%)
 >> orange
 yellow (49.5%)
 >> orange



Sample Images that cannot be accurately predicted in the Test Dataset

5.2.2. Baseline Methods with Single Color

Method	Duration	Accuracy
Pairwise Argmin	21.80 sec.	0.4518
Pairwise Mean	7.41 sec.	0.3959
Array Mean Argmin	12.73 sec.	0.2776
Pairwise First Quantile	30.55 sec.	0.4439
Pairwise Median	30.75 sec.	0.4400
Pairwise Third Quantile	28.28 sec.	0.3699

* Quantization based methods are available only in RGB color space

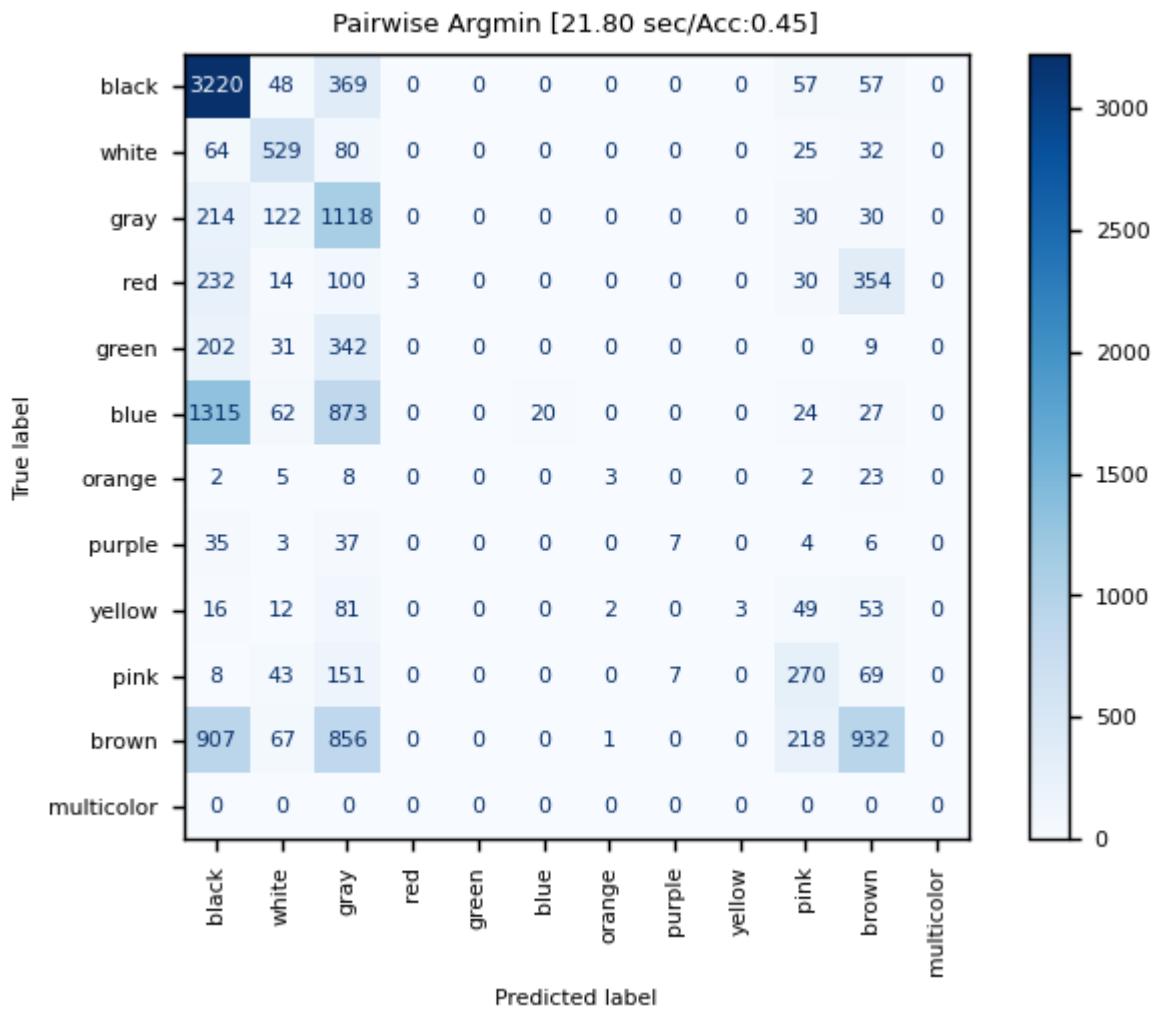
Refined Dataset Improvement Run

Single Color with HSL / Pairwise Argmin Result

Total Duration: 21.80 sec.

Single Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5181	0.8584	0.6462	3751
white	0.5652	0.7247	0.6351	730
gray	0.2785	0.7384	0.4044	1514
red	1.0000	0.0041	0.0082	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0086	0.0171	2321
orange	0.5000	0.0698	0.1224	43
purple	0.5000	0.0761	0.1321	92
yellow	1.0000	0.0139	0.0274	216
pink	0.3808	0.4927	0.4296	548
brown	0.5854	0.3126	0.4076	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4518	0.4518	0.4518	13513
macro avg	0.5273	0.2749	0.2358	13513
weighted avg	0.5971	0.4518	0.3714	13513



Single Color with HSL / Pairwise Argmin Confusion Matrix

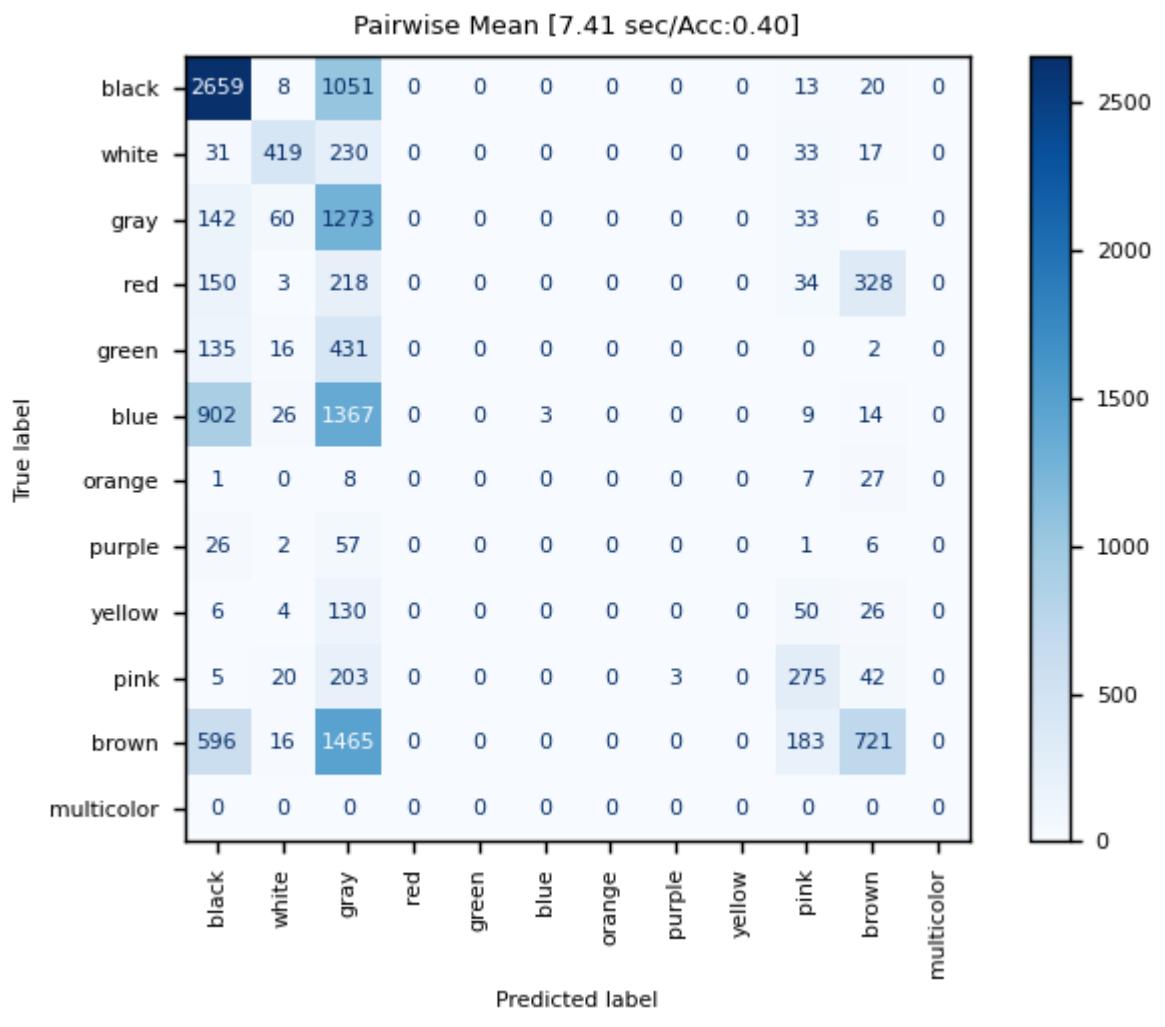
Refined Dataset Improvement Run

Single Color with HSL / Pairwise Mean Result

Total Duration: 7.41 sec.

Single Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.5715	0.7089	0.6328	3751
white	0.7300	0.5740	0.6426	730
gray	0.1979	0.8408	0.3204	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0013	0.0026	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.4310	0.5018	0.4637	548
brown	0.5964	0.2419	0.3442	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3959	0.3959	0.3959	13513
macro avg	0.2939	0.2391	0.2005	13513
weighted avg	0.5410	0.3959	0.3414	13513



Single Color with HSL / Pairwise Mean Confusion Matrix

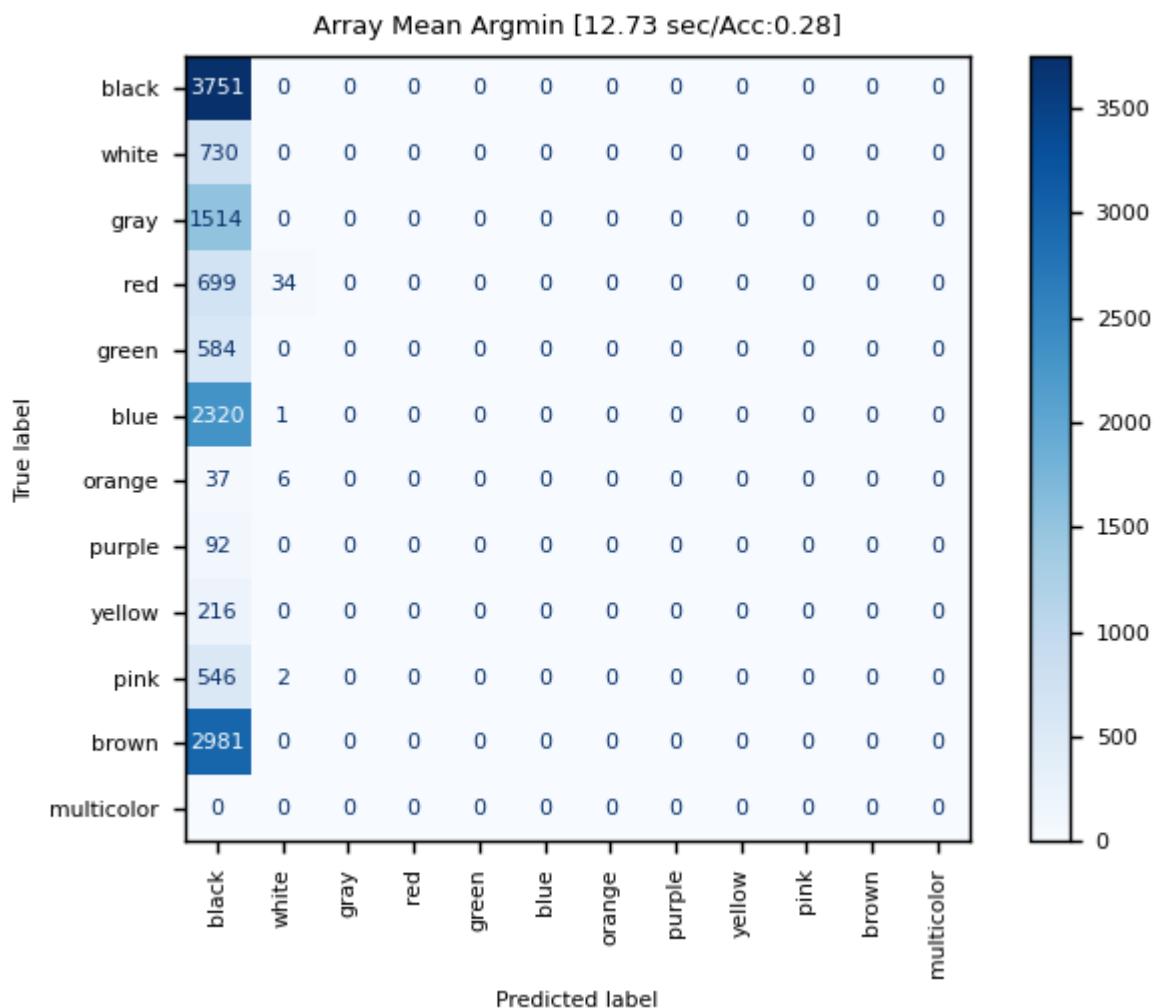
Refined Dataset Improvement Run

Single Color with HSL / Array Mean Argmin Result

Total Duration: 12.73 sec.

Single Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2785	1.0000	0.4356	3751
white	0.0000	0.0000	0.0000	730
gray	0.0000	0.0000	0.0000	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.0000	0.0000	0.0000	548
brown	0.0000	0.0000	0.0000	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.2776	0.2776	0.2776	13513
macro avg	0.0232	0.0833	0.0363	13513
weighted avg	0.0773	0.2776	0.1209	13513



Single Color with HSL / Array Mean Argmin Confusion Matrix

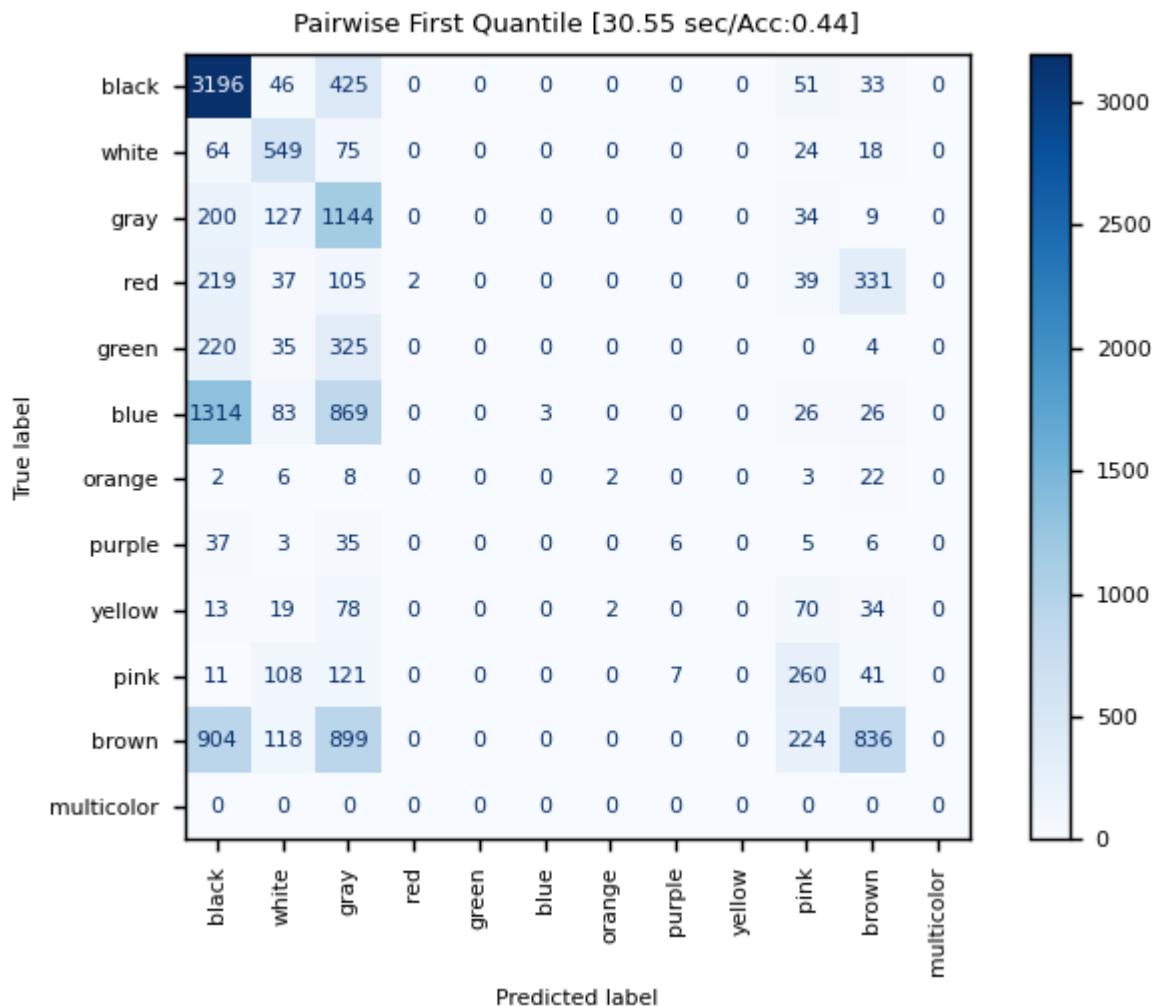
Refined Dataset Improvement Run

Single Color with HSL / Pairwise First Quantile Result

Total Duration: 30.55 sec.

Single Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5172	0.8520	0.6436	3751
white	0.4854	0.7521	0.5900	730
gray	0.2801	0.7556	0.4087	1514
red	1.0000	0.0027	0.0054	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0013	0.0026	2321
orange	0.5000	0.0465	0.0851	43
purple	0.4615	0.0652	0.1143	92
yellow	0.0000	0.0000	0.0000	216
pink	0.3533	0.4745	0.4050	548
brown	0.6147	0.2804	0.3852	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4439	0.4439	0.4439	13513
macro avg	0.4343	0.2692	0.2200	13513
weighted avg	0.5818	0.4439	0.3595	13513



Single Color with HSL / Pairwise First Quantile Confusion Matrix

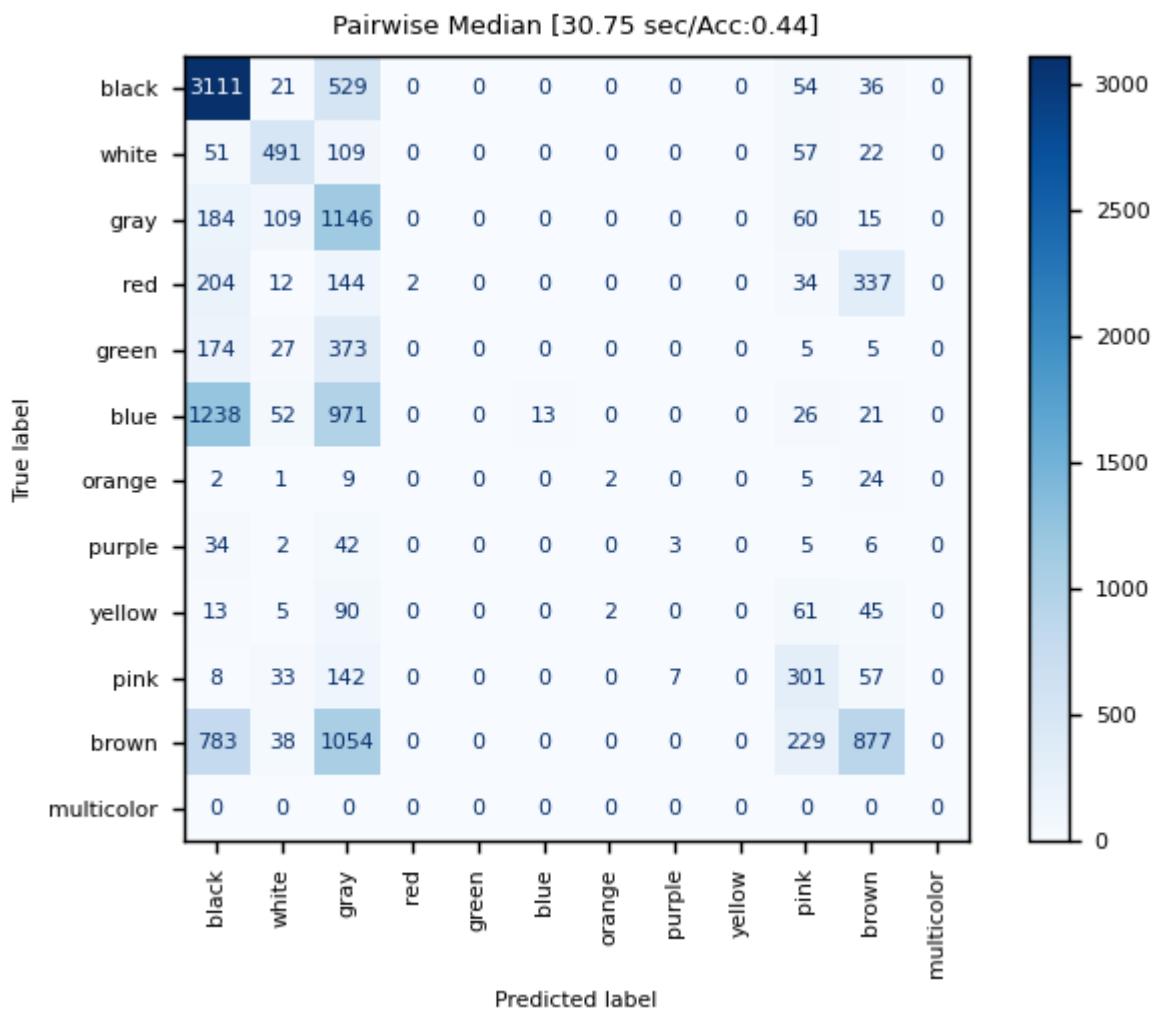
Refined Dataset Improvement Run

Single Color with HSL / Pairwise Median Result

Total Duration: 30.75 sec.

Single Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.5362	0.8294	0.6513	3751
white	0.6207	0.6726	0.6456	730
gray	0.2486	0.7569	0.3743	1514
red	1.0000	0.0027	0.0054	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0056	0.0111	2321
orange	0.5000	0.0465	0.0851	43
purple	0.3000	0.0326	0.0588	92
yellow	0.0000	0.0000	0.0000	216
pink	0.3596	0.5493	0.4347	548
brown	0.6069	0.2942	0.3963	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.4400	0.4400	0.4400	13513
macro avg	0.4310	0.2658	0.2219	13513
weighted avg	0.5883	0.4400	0.3655	13513



Single Color with HSL / Pairwise Median Confusion Matrix

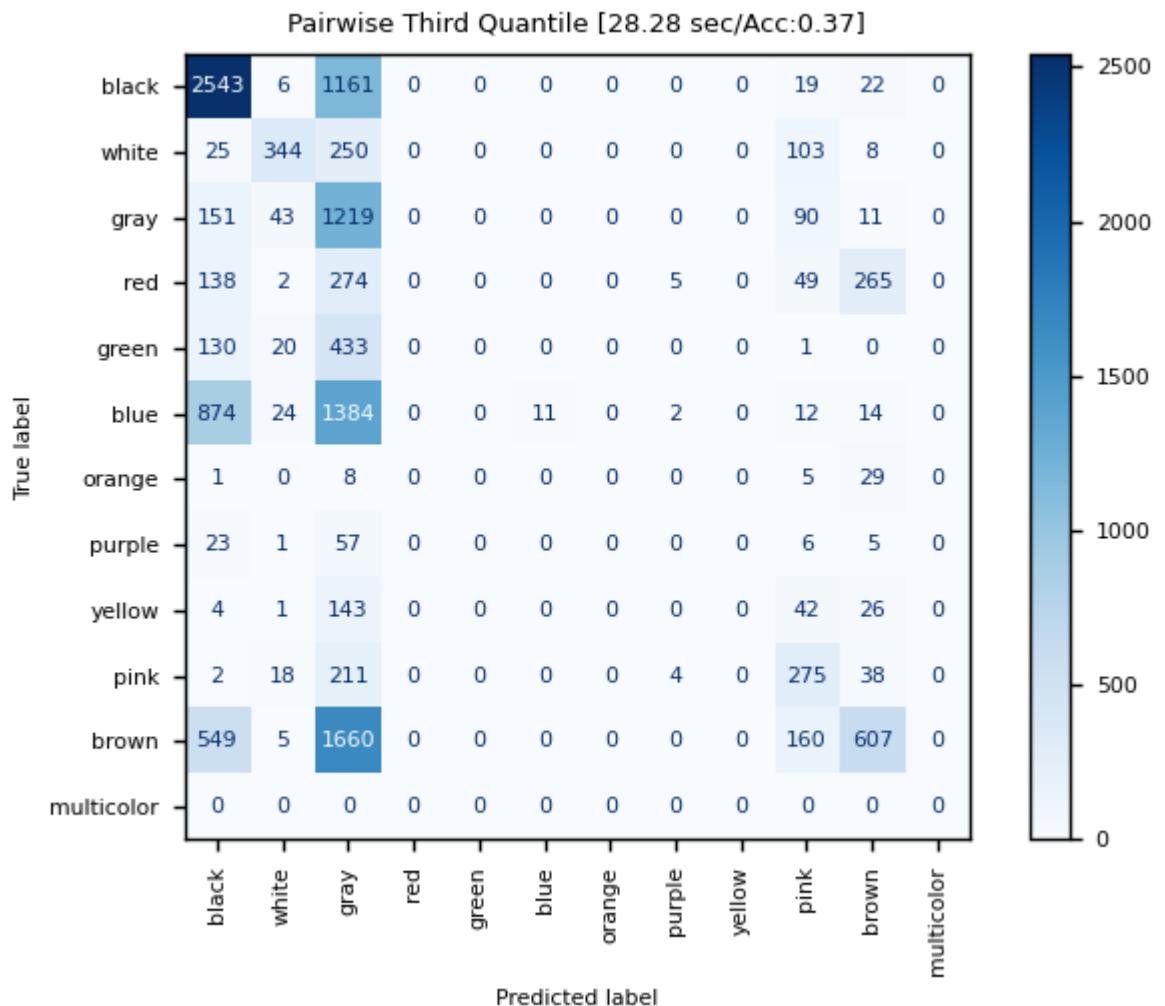
Refined Dataset Improvement Run

Single Color with HSL / Pairwise Third Quantile Result

Total Duration: 28.28 sec.

Single Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5727	0.6780	0.6209	3751
white	0.7414	0.4712	0.5762	730
gray	0.1793	0.8052	0.2932	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	1.0000	0.0047	0.0094	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.3609	0.5018	0.4198	548
brown	0.5922	0.2036	0.3030	2981
multicolor	0.0000	0.0000	0.0000	0
micro avg	0.3699	0.3699	0.3699	13513
macro avg	0.2872	0.2220	0.1852	13513
weighted avg	0.5362	0.3699	0.3218	13513



Single Color with HSL / Pairwise Third Quantile Confusion Matrix

5.2.3. CNN AlexNet with Multi-color

Refined Dataset Improvement Run

CNN AlexNet Multi-Color with HSL

Total Duration: 4071.40 sec.

Learning Rate: 0.000070

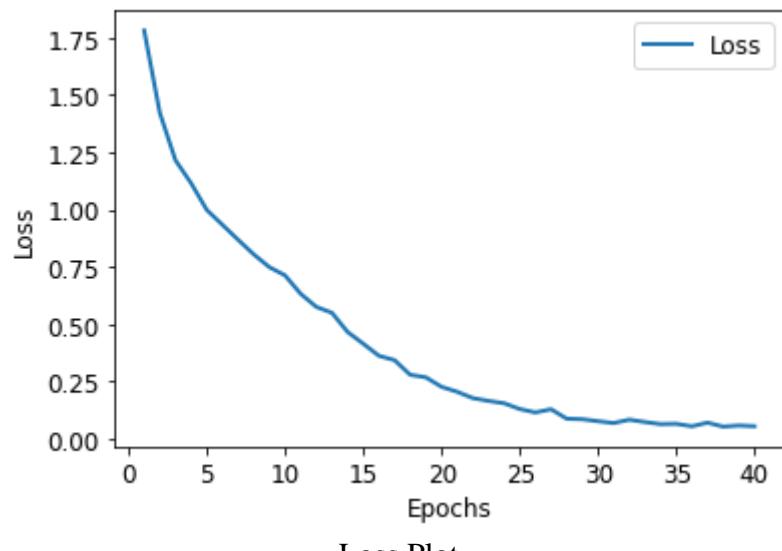
Dropout: 0.10

Batch Size: 32

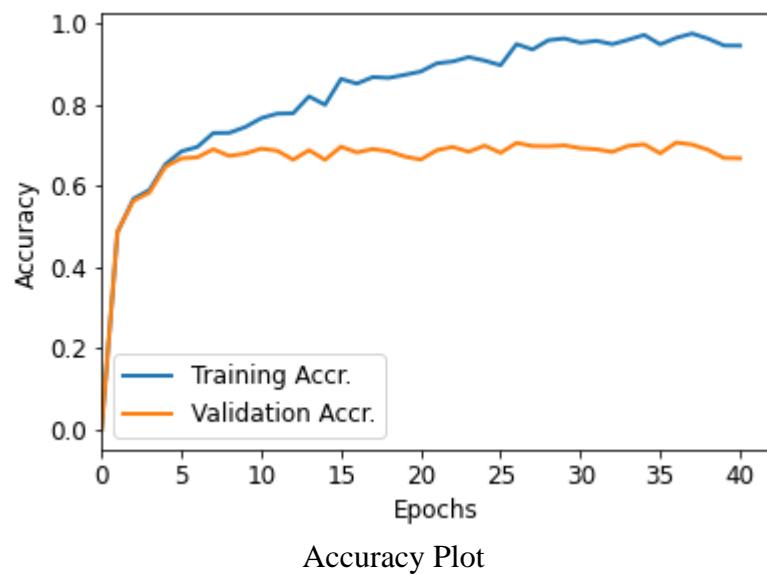
Epoch: 40

Best Performing Epoch: 36

Summary	
Method	Accuracy
Training Dataset	0.9643
Validation Dataset	0.7067
Test Dataset	0.7020



Loss Plot

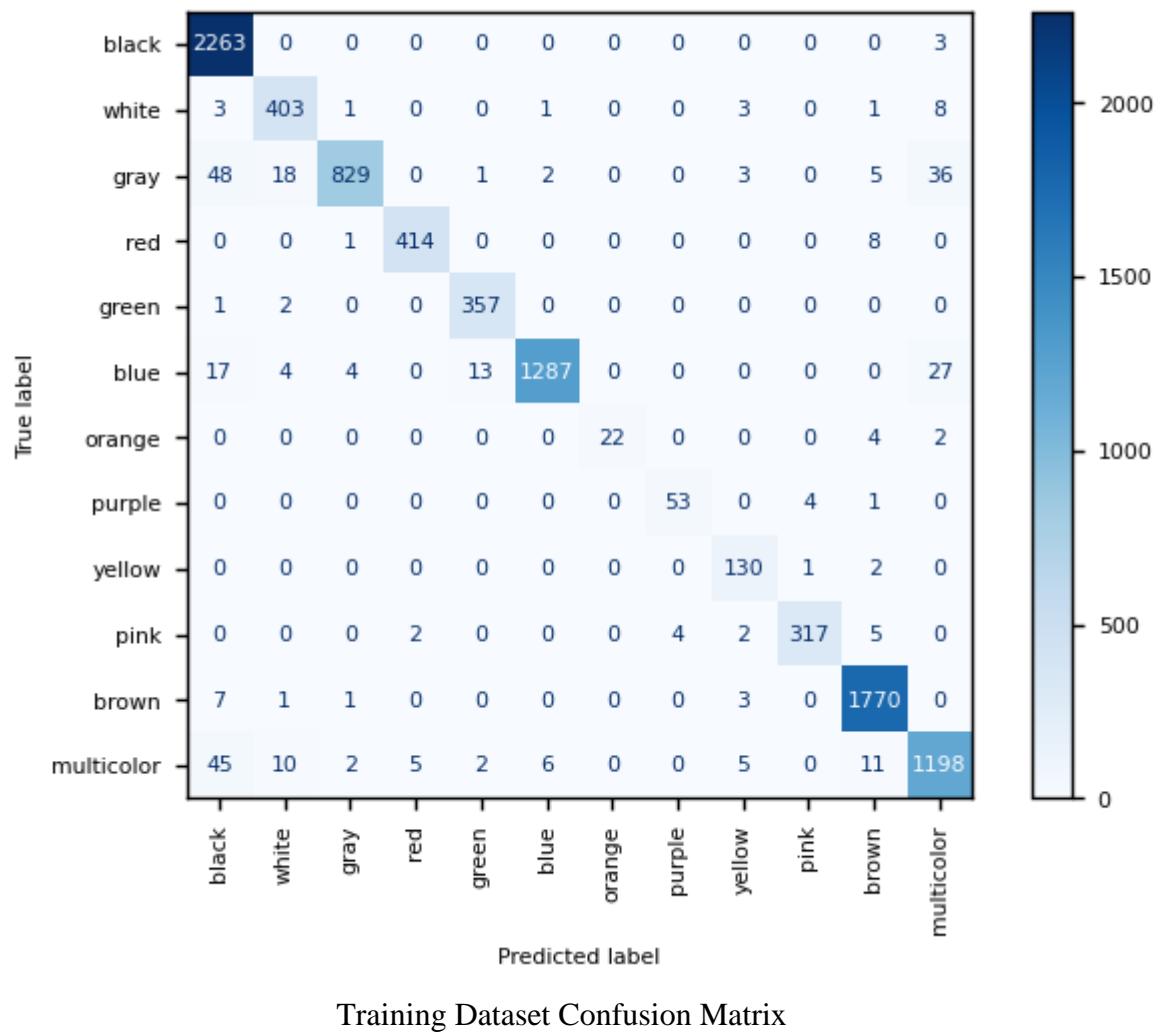


Accuracy Plot

Training Dataset Result

Training Dataset Classification Report

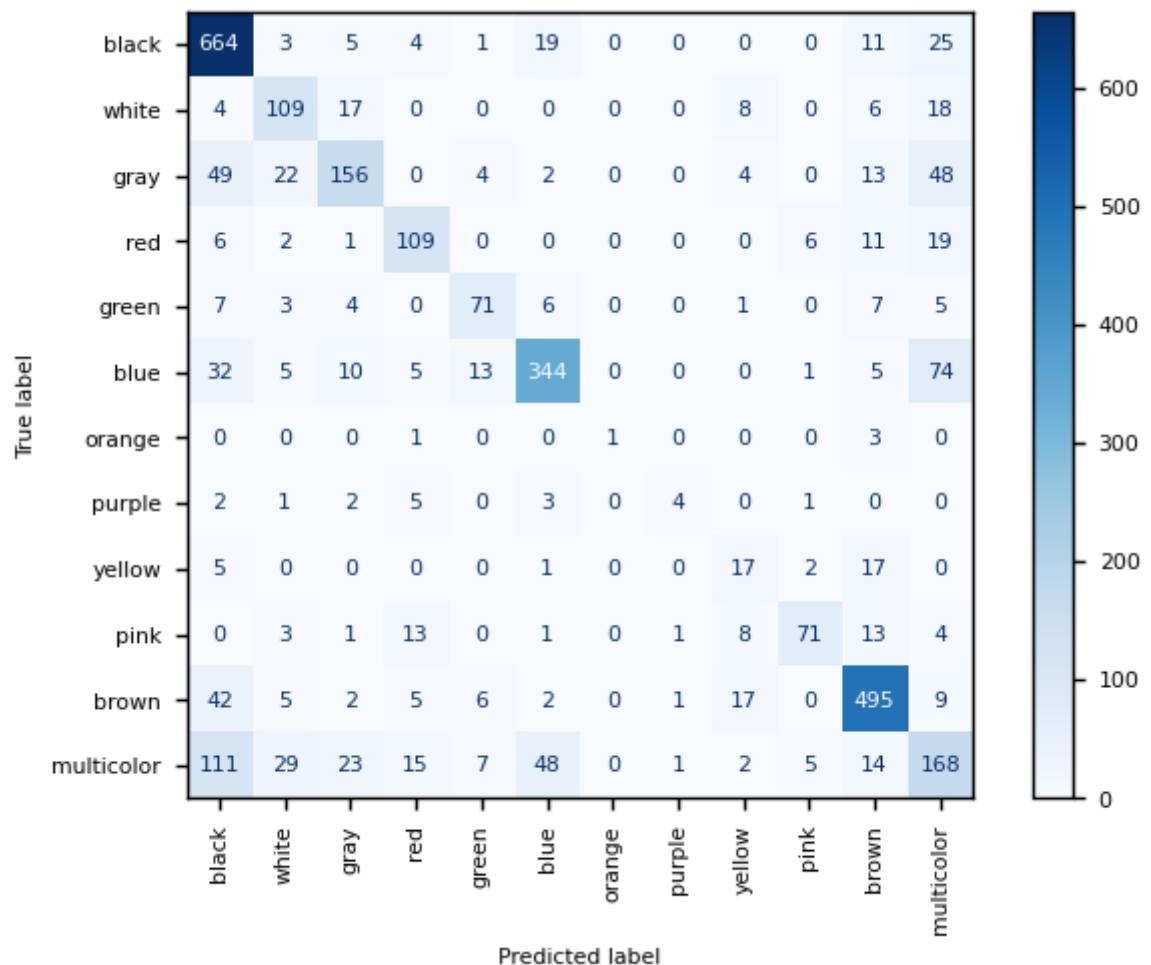
	Precision	Recall	F1 Score	Support
black	0.9492	0.9987	0.9733	2266
white	0.9201	0.9595	0.9394	420
gray	0.9893	0.8800	0.9315	942
red	0.9834	0.9787	0.9810	423
green	0.9571	0.9917	0.9741	360
blue	0.9931	0.9519	0.9721	1352
orange	1.0000	0.7857	0.8800	28
purple	0.9298	0.9138	0.9217	58
yellow	0.8904	0.9774	0.9319	133
pink	0.9845	0.9606	0.9724	330
brown	0.9795	0.9933	0.9863	1782
multicolor	0.9403	0.9330	0.9367	1284
accuracy			0.9643	
macro avg	0.9597	0.9437	0.9500	9378
weighted avg	0.9651	0.9643	0.9640	9378



Validation Dataset Result

Validation Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7202	0.9071	0.8029	732
white	0.5989	0.6728	0.6337	162
gray	0.7059	0.5235	0.6012	298
red	0.6943	0.7078	0.7010	154
green	0.6961	0.6827	0.6893	104
blue	0.8075	0.7035	0.7519	489
orange	1.0000	0.2000	0.3333	5
purple	0.5714	0.2222	0.3200	18
yellow	0.2982	0.4048	0.3434	42
pink	0.8256	0.6174	0.7065	115
brown	0.8319	0.8476	0.8397	584
multicolor	0.4541	0.3972	0.4237	423
accuracy			0.7067	
macro avg	0.6837	0.5739	0.5956	3126
weighted avg	0.7068	0.7067	0.7004	3126

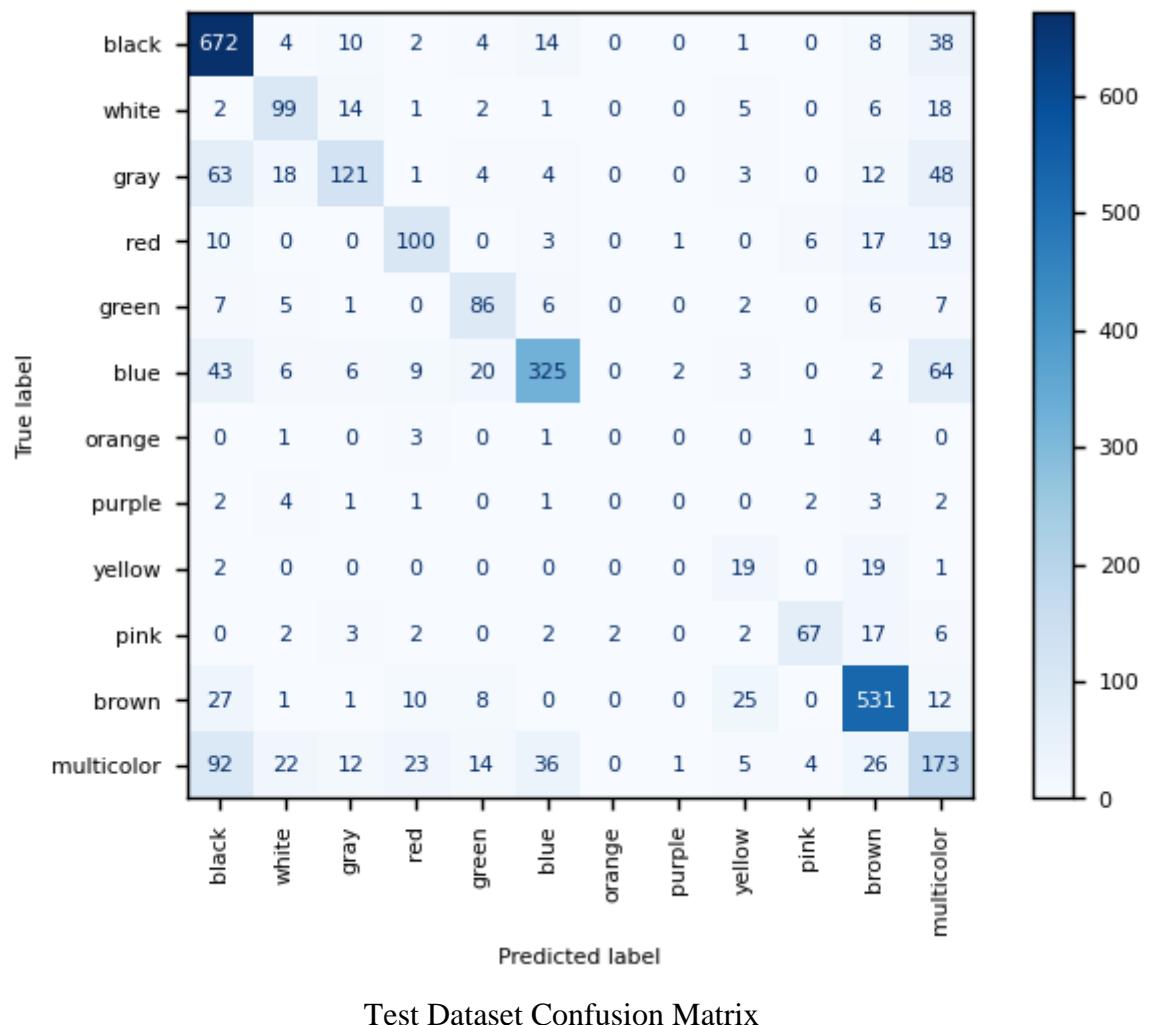


Validation Dataset Confusion Matrix

Test Dataset Result

Test Dataset Classification Report

	Precision	Recall	F1 Score	Support
black	0.7304	0.8924	0.8033	753
white	0.6111	0.6689	0.6387	148
gray	0.7160	0.4416	0.5463	274
red	0.6579	0.6410	0.6494	156
green	0.6232	0.7167	0.6667	120
blue	0.8270	0.6771	0.7446	480
orange	0.0000	0.0000	0.0000	10
purple	0.0000	0.0000	0.0000	16
yellow	0.2923	0.4634	0.3585	41
pink	0.8375	0.6505	0.7322	103
brown	0.8157	0.8634	0.8389	615
multicolor	0.4459	0.4240	0.4347	408
accuracy			0.7020	
macro avg	0.5464	0.5366	0.5344	3124
weighted avg	0.7019	0.7020	0.6950	3124



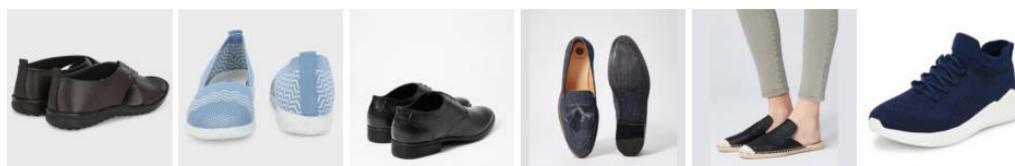
purple (84.7%) multicolor (100.0%) black (100.0%) black (100.0%) red (100.0%) white (99.5%)
 >> red >> black >> black >> black >> red >> white



black (96.8%) black (100.0%) multicolor (78.6%) black (99.8%) black (91.5%) brown (99.9%)
 >> gray >> black >> blue >> multicolor >> black >> brown



multicolor (97.1%) multicolor (94.2%) black (100.0%) black (100.0%) black (100.0%) multicolor (87.1%)
 >> brown >> blue >> black >> blue >> black >> multicolor



multicolor (96.9%) brown (100.0%) blue (100.0%) green (95.2%) brown (100.0%) brown (97.4%)
 >> multicolor >> brown >> blue >> green >> brown >> brown



white (99.8%) black (99.0%) gray (100.0%) black (100.0%) black (100.0%) black (100.0%)
 >> white >> gray >> gray >> black >> black >> black



blue (54.8%) red (92.5%) blue (100.0%) brown (100.0%) black (100.0%) brown (99.8%)
 >> blue >> red >> blue >> brown >> black >> black



Sample Images from Test Dataset

multicolor (98.6%) multicolor (91.9%) multicolor (97.8%) blue (97.7%)
 >> blue >> blue >> gray >> pink black (100.0%) multicolor (100.0%)
 >> red >> black



brown (99.8%) yellow (100.0%) blue (99.8%) black (99.0%) purple (87.3%) black (100.0%)
 >> black >> blue >> multicolor >> gray >> blue >> gray



multicolor (96.6%) multicolor (100.0%) white (47.7%) multicolor (94.2%) black (96.8%) black (100.0%)
 >> blue >> white >> gray >> blue >> gray >> blue



black (79.9%) blue (37.5%) brown (72.1%) black (100.0%) black (100.0%) blue (62.5%)
 >> brown >> gray >> white >> multicolor >> blue >> black



brown (99.9%) multicolor (99.9%) yellow (82.8%) blue (60.1%) multicolor (78.6%) brown (98.1%)
 >> yellow >> blue >> brown >> orange >> blue >> black



multicolor (100.0%) white (99.1%) multicolor (56.9%) brown (99.7%) black (99.8%) red (100.0%)
 >> black >> multicolor >> blue >> pink >> multicolor >> brown



Sample Images that cannot be accurately predicted in the Test Dataset

5.2.4. Baseline Methods with Multi-color

Method	Duration	Accuracy
Pairwise Argmin	25.60 sec.	0.3097
Pairwise Mean	8.42 sec.	0.2400
Array Mean Argmin	26.09 sec.	0.2400
Pairwise First Quantile	38.30 sec.	0.3723
Pairwise Median	41.53 sec.	0.3643
Pairwise Third Quantile	39.84 sec.	0.3018

* Quantization based methods are available only in RGB color space

Refined Dataset Improvement Run

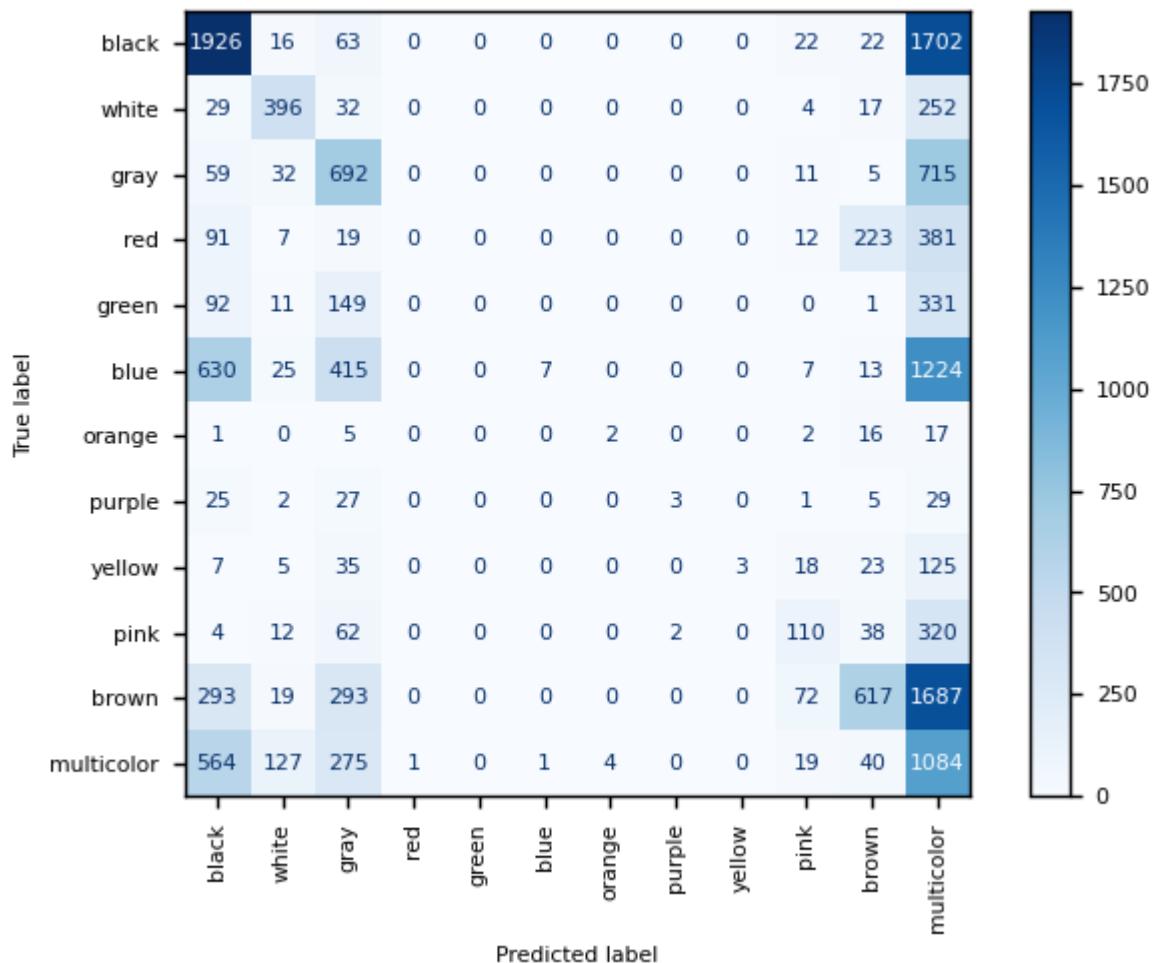
Multi-Color with HSL / Pairwise Argmin Result

Total Duration: 25.60 sec.

Multi-Color with HSL / Pairwise Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.5176	0.5135	0.5155	3751
white	0.6074	0.5425	0.5731	730
gray	0.3348	0.4571	0.3865	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.8750	0.0030	0.0060	2321
orange	0.3333	0.0465	0.0816	43
purple	0.6000	0.0326	0.0619	92
yellow	1.0000	0.0139	0.0274	216
pink	0.3957	0.2007	0.2663	548
brown	0.6049	0.2070	0.3084	2981
multicolor	0.1378	0.5125	0.2172	2115
accuracy			0.3097	
macro avg	0.4505	0.2108	0.2037	15628
weighted avg	0.4812	0.3097	0.2874	15628

Pairwise Argmin [25.60 sec/Acc:0.31]



Multi-Color with HSL / Pairwise Argmin Confusion Matrix

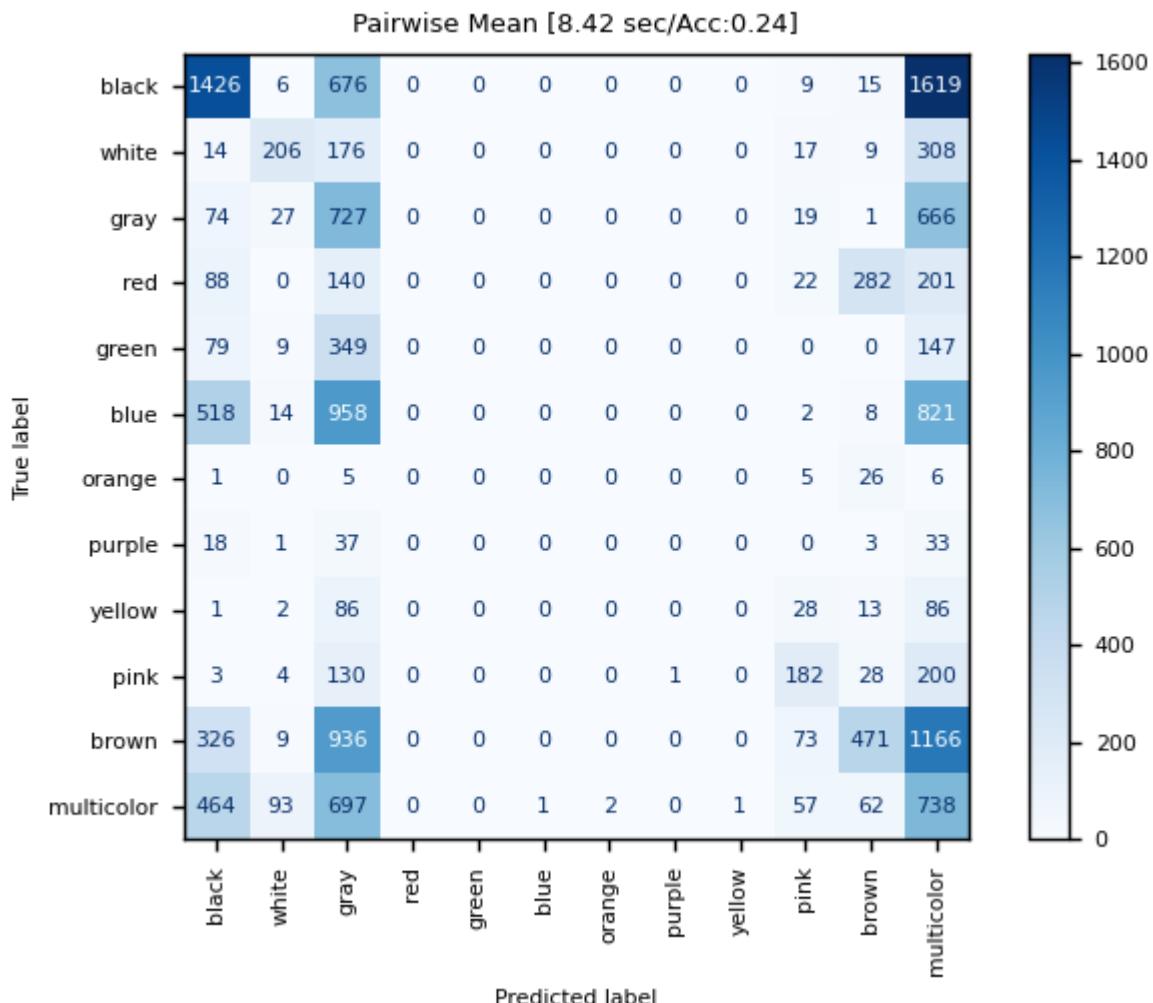
Refined Dataset Improvement Run

Multi-Color with HSL / Pairwise Mean Result

Total Duration: 8.42 sec.

Multi-Color with HSL / Pairwise Mean Classification Report

	Precision	Recall	F1 Score	Support
black	0.4734	0.3802	0.4217	3751
white	0.5553	0.2822	0.3742	730
gray	0.1479	0.4802	0.2261	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.4396	0.3321	0.3784	548
brown	0.5131	0.1580	0.2416	2981
multicolor	0.1232	0.3489	0.1821	2115
accuracy			0.2400	
macro avg	0.1877	0.1651	0.1520	15628
weighted avg	0.2838	0.2400	0.2246	15628



Multi-Color with HSL / Pairwise Mean Confusion Matrix

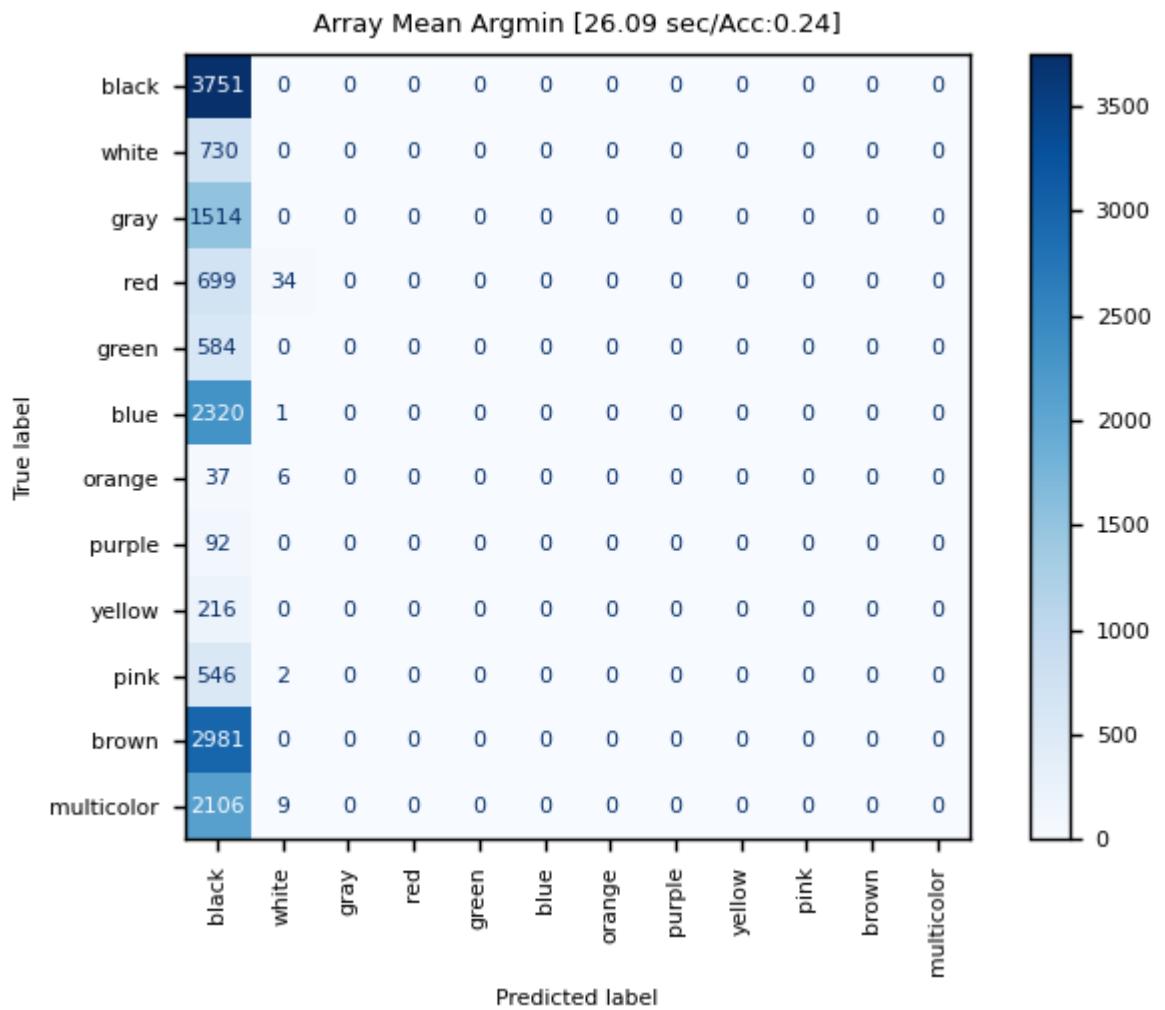
Refined Dataset Improvement Run

Multi-Color with HSL / Array Mean Argmin Result

Total Duration: 26.09 sec.

Multi-Color with HSL / Array Mean Argmin Classification Report

	Precision	Recall	F1 Score	Support
black	0.2408	1.0000	0.3882	3751
white	0.0000	0.0000	0.0000	730
gray	0.0000	0.0000	0.0000	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.0000	0.0000	0.0000	548
brown	0.0000	0.0000	0.0000	2981
multicolor	0.0000	0.0000	0.0000	2115
accuracy			0.2400	
macro avg	0.0201	0.0833	0.0323	15628
weighted avg	0.0578	0.2400	0.0932	15628



Multi-Color with HSL / Array Mean Argmin Confusion Matrix

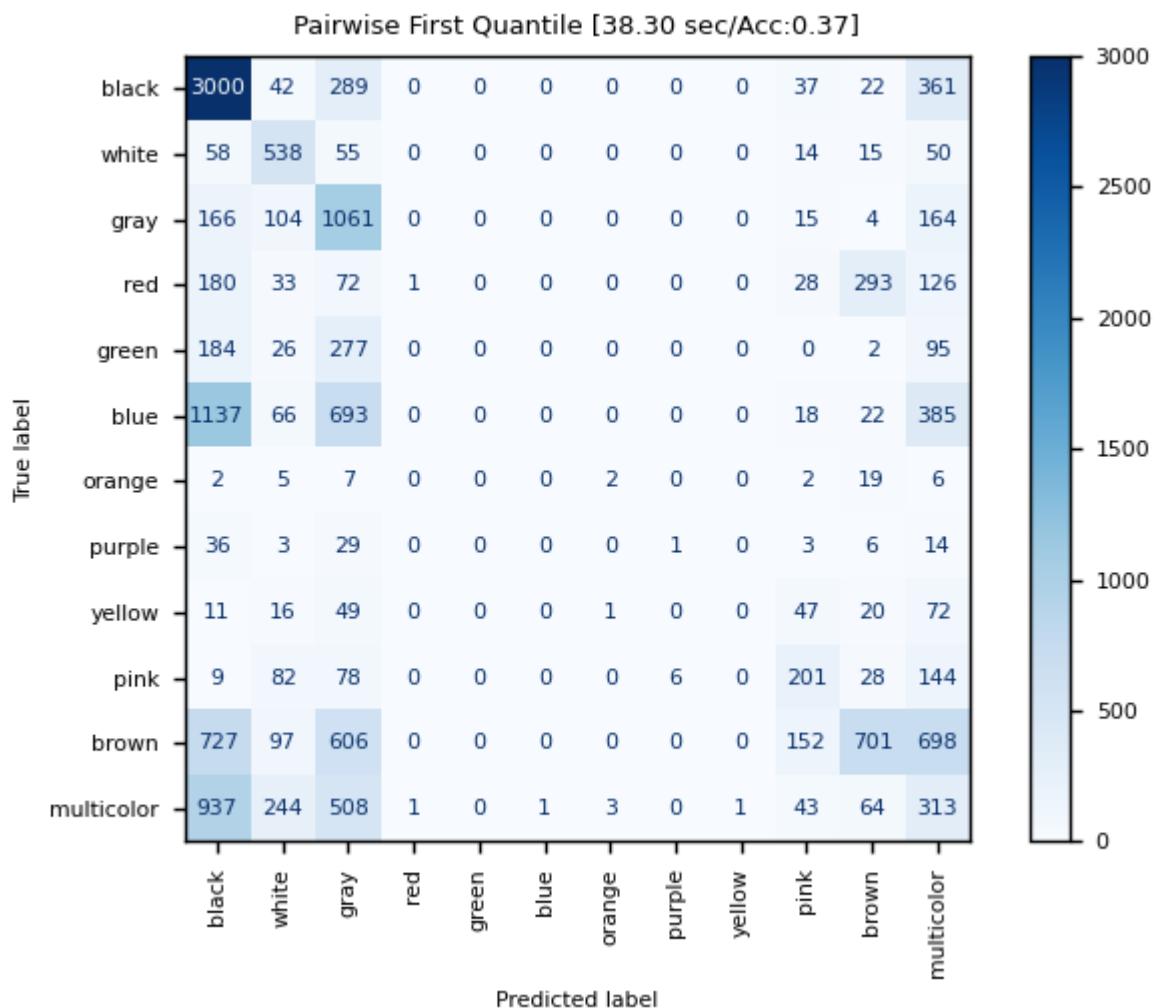
Refined Dataset Improvement Run

Multi-Color with HSL / Pairwise First Quantile Result

Total Duration: 38.30 sec.

Multi-Color with HSL / Pairwise First Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.4653	0.7998	0.5884	3751
white	0.4283	0.7370	0.5418	730
gray	0.2849	0.7008	0.4051	1514
red	0.5000	0.0014	0.0027	733
green	0.0000	0.0000	0.0000	584
blue	0.0000	0.0000	0.0000	2321
orange	0.3333	0.0465	0.0816	43
purple	0.1429	0.0109	0.0202	92
yellow	0.0000	0.0000	0.0000	216
pink	0.3589	0.3668	0.3628	548
brown	0.5861	0.2352	0.3356	2981
multicolor	0.1289	0.1480	0.1378	2115
accuracy			0.3723	
macro avg	0.2691	0.2539	0.2063	15628
weighted avg	0.3263	0.3723	0.3016	15628



Multi-Color with HSL / Pairwise First Quantile Confusion Matrix

Refined Dataset Improvement Run

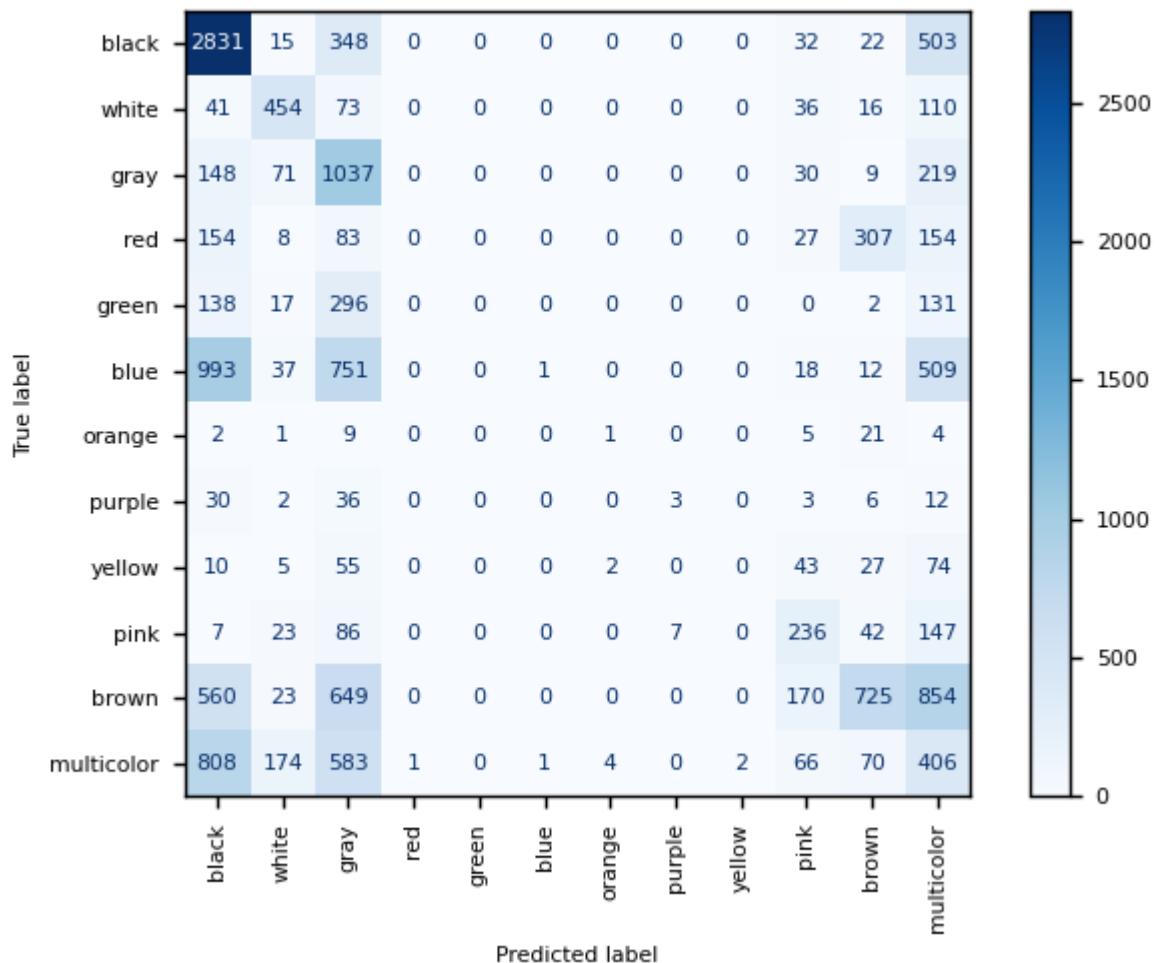
Multi-Color with HSL / Pairwise Median Result

Total Duration: 41.53 sec.

Multi-Color with HSL / Pairwise Median Classification Report

	Precision	Recall	F1 Score	Support
black	0.4948	0.7547	0.5977	3751
white	0.5470	0.6219	0.5821	730
gray	0.2589	0.6849	0.3757	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.5000	0.0004	0.0009	2321
orange	0.1429	0.0233	0.0400	43
purple	0.3000	0.0326	0.0588	92
yellow	0.0000	0.0000	0.0000	216
pink	0.3544	0.4307	0.3888	548
brown	0.5759	0.2432	0.3420	2981
multicolor	0.1300	0.1920	0.1550	2115
accuracy			0.3643	
macro avg	0.2753	0.2486	0.2117	15628
weighted avg	0.3857	0.3643	0.3075	15628

Pairwise Median [41.53 sec/Acc:0.36]



Multi-Color with HSL / Pairwise Median Confusion Matrix

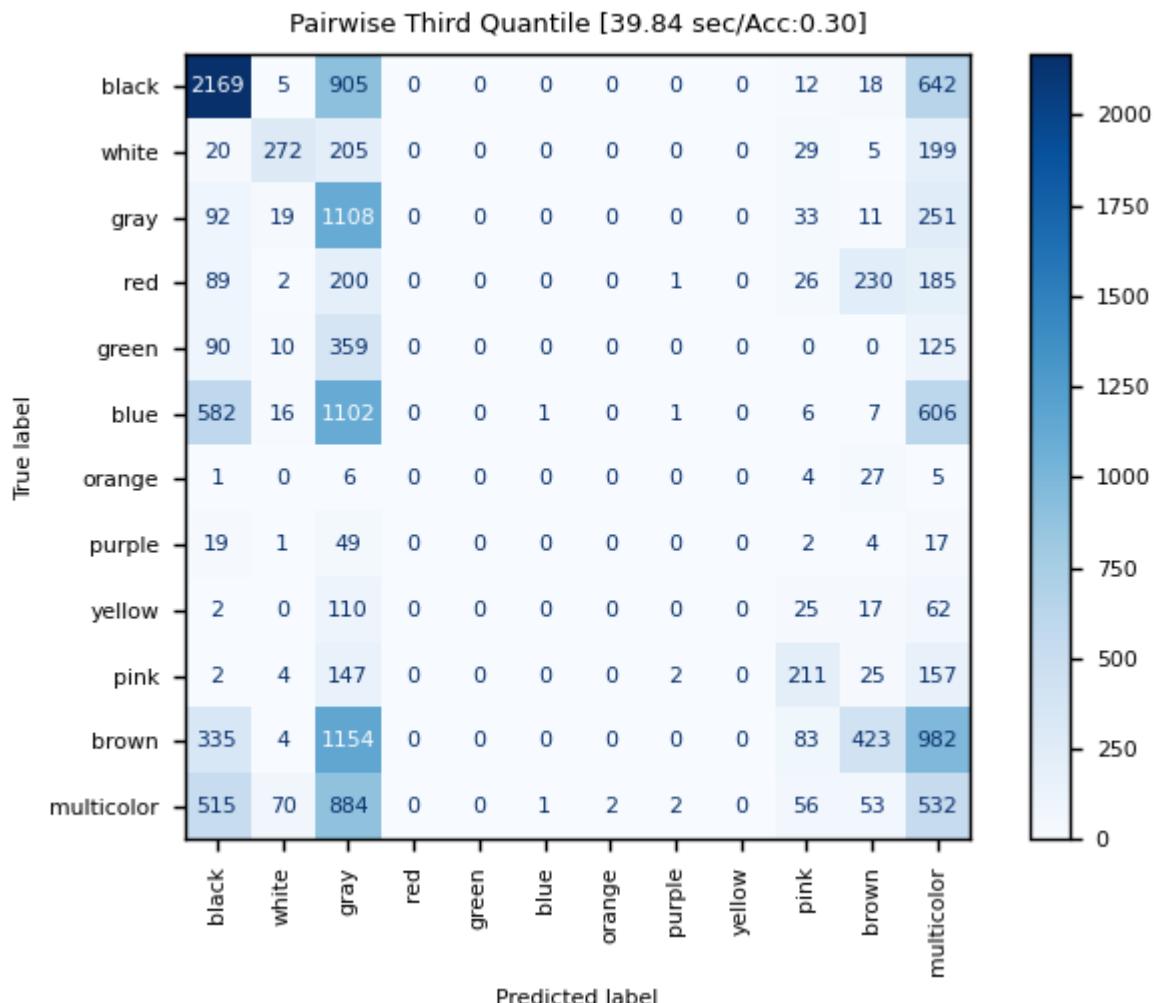
Refined Dataset Improvement Run

Multi-Color with HSL / Pairwise Third Quantile Result

Total Duration: 39.84 sec.

Multi-Color with HSL / Pairwise Third Quantile Classification Report

	Precision	Recall	F1 Score	Support
black	0.5539	0.5782	0.5658	3751
white	0.6749	0.3726	0.4801	730
gray	0.1779	0.7318	0.2862	1514
red	0.0000	0.0000	0.0000	733
green	0.0000	0.0000	0.0000	584
blue	0.5000	0.0004	0.0009	2321
orange	0.0000	0.0000	0.0000	43
purple	0.0000	0.0000	0.0000	92
yellow	0.0000	0.0000	0.0000	216
pink	0.4333	0.3850	0.4077	548
brown	0.5159	0.1419	0.2226	2981
multicolor	0.1414	0.2515	0.1810	2115
accuracy			0.3018	
macro avg	0.2498	0.2051	0.1787	15628
weighted avg	0.3887	0.3018	0.2673	15628



Multi-Color with HSL / Pairwise Third Quantile Confusion Matrix

