

# CVEDA Dataset Audit

Dataset root dataset

## ***Run metadata***

Features run : 40

Checks : completeness, bbox\_sanity, summary

# Executive summary

Images indexed 11704

Check	Status
completeness	ok
bbox_sanity	ok
summary	ok

## ***Absolute Coordinate Patterns***

Feature absolute\_coordinate\_patterns results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	absolute_coordinate_patterns
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Annotation Confidence***

Feature annotation\_confidence results.

### ***Key findings***

- Images inspected 11704

Metric	Value
feature	annotation_confidence
n_images	11704
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Annotation Lifespan Drift***

Feature annotation\_lifespan\_drift results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	annotation_lifespan_drift
status	no_timestamps
message	No timestamp metadata found

### ***Recommendations***

No automated recommendations generated for this feature

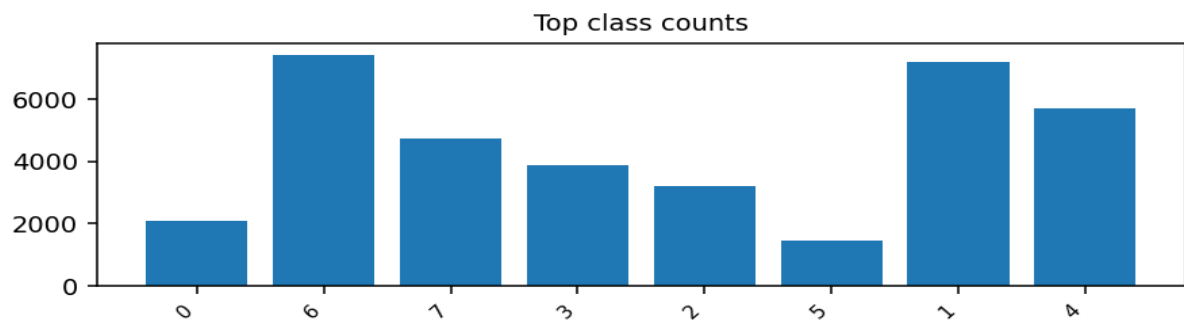
## Annotation Rarity Buckets

Feature annotation\_rarity\_buckets results.

### Key findings

- Top classes 6:7404, 1:7184, 4:5712, 7:4720, 3:3880, 2:3196

Metric	Value
feature	annotation_rarity_buckets
status	ok



### Recommendations

No automated recommendations generated for this feature

## ***Background Relevance***

Feature background\_relevance results.

### ***Key findings***

- Images inspected 11704

Metric	Value
feature	background_relevance
n_images	11704
mean_annotation_fraction	0.45111477078606815
median_annotation_fraction	0.2757026627218935
low_coverage_count	188
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## Bbox Border Alignment

Feature bbox\_border\_alignment results.

### Key findings

- No automatic takeaways generated for this feature

Metric	Value
feature	bbox_border_alignment
n_boxes	35624
border_touch_fraction	0.3124859645183023
status	ok

### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



frame\_1021\_jpg.rf.0cec4cd85f4b1c85b7729  
71da5e12e28.jpg



frame\_1021\_jpg.rf.0cec4cd85f4b1c85b7729  
71da5e12e28.jpg

### Recommendations

- Many boxes touch image borders. Consider re annotating clipped objects.



## **Camera Diversity**

Feature camera\_diversity results.

### **Key findings**

- No automatic takeaways generated for this feature

Metric	Value
feature	camera_diversity
scanned	500
status	ok

### **Recommendations**

No automated recommendations generated for this feature

## ***Color Cast Detection***

Feature color\_cast\_detection results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	color_cast_detection
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## Compression Anomaly

Feature compression\_anomaly results.

### Key findings

- No automatic takeaways generated for this feature

Metric	Value
feature	compression_anomaly
sampled	200
status	ok

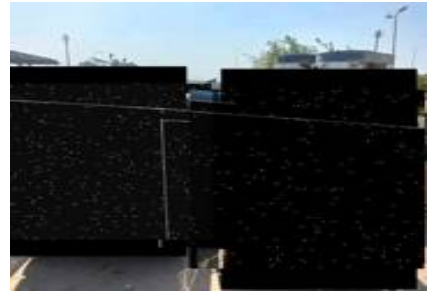
### Sample evidence



1559032\_0\_jpg.rf.ec1ee786392dda6e7f589b6a5a4d0d28.jpg



1559032\_0\_jpg.rf.ec1ee786392dda6e7f589b6a5a4d0d28.jpg



1559032\_0\_jpg.rf.d35291bc891e0fe9bb408c0b381e9d18.jpg



1559032\_0\_jpg.rf.d35291bc891e0fe9bb408c0b381e9d18.jpg



1559032\_0\_jpg.rf.e8e280b00d5a156d4032976b511c02f8.jpg



1559032\_0\_jpg.rf.e8e280b00d5a156d4032976b511c02f8.jpg

### Recommendations

No automated recommendations generated for this feature

## ***Containment Detection***

Feature containment\_detection results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	containment_detection
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Duplicate Annotation Ids***

Feature duplicate\_annotation\_ids results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	duplicate_annotation_ids
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

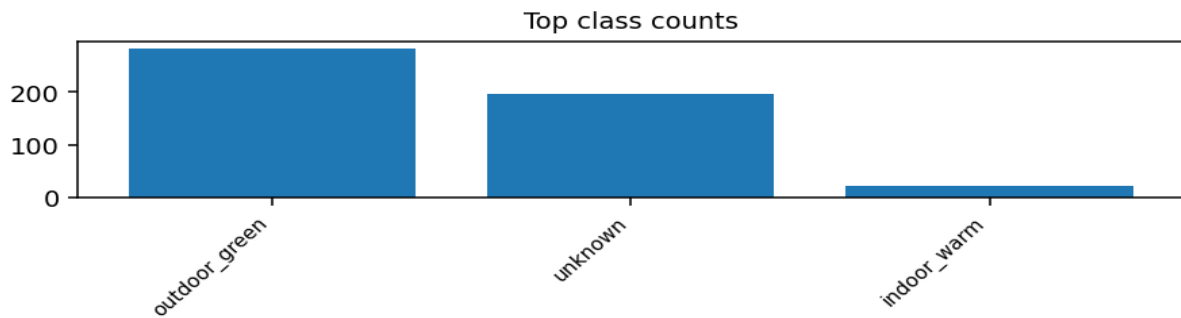
## Environment Diversity

Feature environment\_diversity results.

### Key findings

- Top classes outdoor\_green:282, unknown:196, indoor\_warm:22

Metric	Value
feature	environment_diversity
scanned	500
status	ok



### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg

### Recommendations

No automated recommendations generated for this feature

## ***Geographic Clustering***

Feature geographic\_clustering results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	geographic_clustering
status	no_gps

### ***Recommendations***

No automated recommendations generated for this feature

**Heirarchy Validation**

Feature heirarchy\_validation results.

**Key findings**

- No automatic takeaways generated for this feature

Metric	Value
status	no-runner
note	No runner run_heirarchy_validation or run callable found in cveda.feature

**Recommendations**

No automated recommendations generated for this feature



***Illumination Diversity***

Feature illumination\_diversity results.

***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	illumination_diversity
scanned	500
status	ok

***Sample evidence***



frame\_1254\_jpg.rf.2bc9f4c8b964bd2ebf9a3861101567b3.jpg



frame\_1254\_jpg.rf.2bc9f4c8b964bd2ebf9a3861101567b3.jpg



frame\_201\_jpg.rf.e96dfb8b469d60fd9df6ebb5afce10dd.jpg



frame\_201\_jpg.rf.e96dfb8b469d60fd9df6ebb5afce10dd.jpg



IMG\_5310\_JPG.rf.998009a095670560181a3c1cbcce5998.jpg



IMG\_5310\_JPG.rf.998009a095670560181a3c1cbcce5998.jpg

***Recommendations***

No automated recommendations generated for this feature

## ***Inter Annotator Disagreement***

Feature inter\_annotator\_disagreement results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	inter_annotator_disagreement
available	False
message	No annotator identifiers found in annotations; cannot compute disagreement
status	no_annotators

### ***Recommendations***

No automated recommendations generated for this feature

## ***Metadata Drift***

Feature metadata\_drift results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	metadata_drift
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## Mislabel Candidate Generator

Feature mislabel\_candidate\_generator results.

### Key findings

- No automatic takeaways generated for this feature

Metric	Value
feature	mislabel_candidate_generator
status	ok

### Sample evidence



frame\_1582\_jpg.rf.14058eadab1c087534dc9e5e80246430.jpg



frame\_1582\_jpg.rf.14058eadab1c087534dc9e5e80246430.jpg



frame\_1582\_jpg.rf.14058eadab1c087534dc9e5e80246430.jpg



frame\_1582\_jpg.rf.14058eadab1c087534dc9e5e80246430.jpg



img\_129\_jpg.rf.95d834561252eb48c4abcf7bca04a3af.jpg



img\_129\_jpg.rf.95d834561252eb48c4abcf7bca04a3af.jpg

### Recommendations

No automated recommendations generated for this feature

## ***Multiresolution Similarity***

Feature multiresolution\_similarity results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	multiresolution_similarity
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Mutual Exclusion***

Feature mutual\_exclusion results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	mutual_exclusion
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

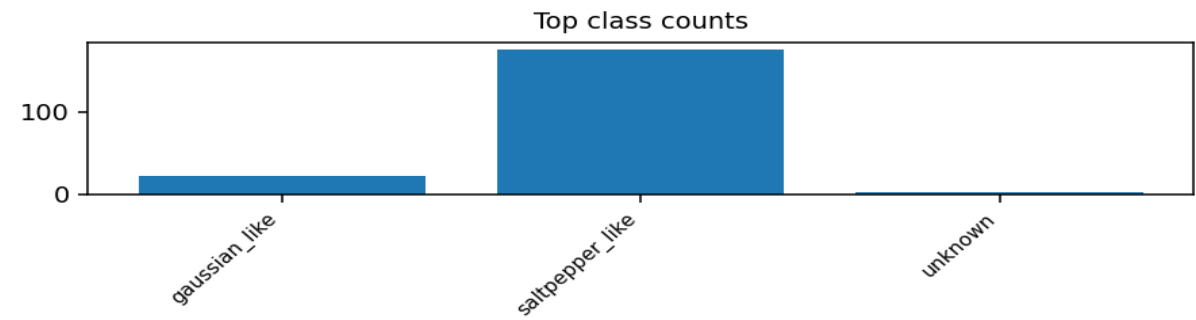
Noise Type Classifier

Feature noise\_type\_classifier results.

Key findings

- Top classes saltpepper\_like:176, gaussian\_like:22, unknown:2

Metric	Value
feature	noise_type_classifier
processed	200
status	ok



Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg

Recommendations

No automated recommendations generated for this feature

## ***Object Alignment Patterns***

Feature object\_alignment\_patterns results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	object_alignment_patterns
status	ok

### ***Recommendations***

No automated recommendations generated for this feature



## ***Object Distance Distribution***

Feature object\_distance\_distribution results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	object_distance_distribution
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## Objects Touching Edges

Feature objects\_touching\_edges results.

### Key findings

- No automatic takeaways generated for this feature

Metric	Value
feature	objects_touching_edges
n_boxes	35624
touching_boxes	11132
touch_fraction_global	0.3124859645183023
status	ok

### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



frame\_1021\_jpg.rf.0cec4cd85f4b1c85b7729  
71da5e12e28.jpg



frame\_1021\_jpg.rf.0cec4cd85f4b1c85b7729  
71da5e12e28.jpg

### Recommendations

No automated recommendations generated for this feature

## ***Occlusion Estimator***

Feature occlusion\_estimator results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	occlusion_estimator
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Polygon Smoothness***

Feature polygon\_smoothness results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	polygon_smoothness
n_polygons	0
status	no_polygons

### ***Recommendations***

No automated recommendations generated for this feature

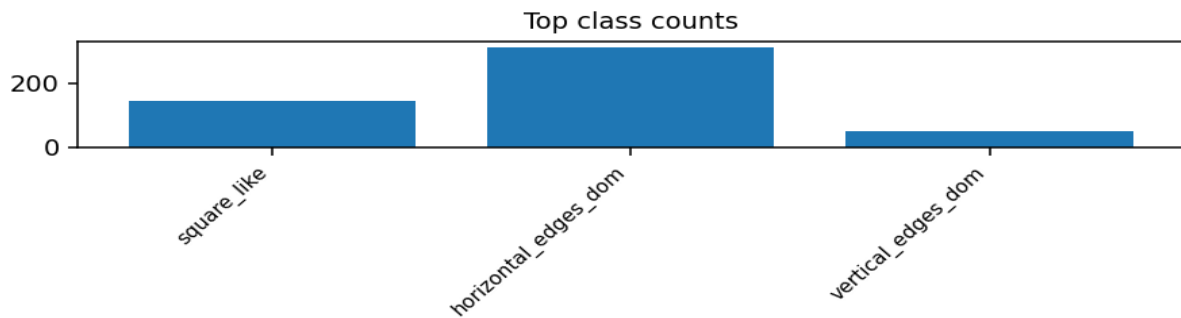
## Pose Diversity

Feature pose\_diversity results.

### Key findings

- Top classes horizontal\_edges\_dom:312, square\_like:144, vertical\_edges\_dom:48

Metric	Value
feature	pose_diversity
processed	504
status	ok



### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg

### Recommendations

No automated recommendations generated for this feature

Random Crop Inspection

Feature random\_crop\_inspection results.

Key findings

- No automatic takeaways generated for this feature

Metric	Value
feature	random_crop_inspection
n_images_sampled	200
mean_crop_kills_fraction	0.16699999999999993
status	ok

Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg

Recommendations

No automated recommendations generated for this feature

## ***Relative Size Ratios***

Feature relative\_size\_ratios results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	relative_size_ratios
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Repeated Background Detector***

Feature repeated\_background\_detector results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	repeated_background_detector
status	ok

### ***Recommendations***

No automated recommendations generated for this feature



## ***Repetitive Layouts***

Feature repetitive\_layouts results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	repetitive_layouts
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Seasonal Coverage***

Feature seasonal\_coverage results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	seasonal_coverage
status	no_timestamps

### ***Recommendations***

No automated recommendations generated for this feature

## Shadow Intensity

Feature shadow\_intensity results.

### Key findings

- Sampled images 20

Metric	Value
feature	shadow_intensity
sampled_images	20
mean_shadow_fraction	0.07631286982248521
status	ok

### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg

### Recommendations

No automated recommendations generated for this feature

**Spatial Culture Inference**

Feature spatial\_culture\_inference results.

**Key findings**

- No automatic takeaways generated for this feature

Metric	Value
feature	spatial_culture_inference
status	ok

**Sample evidence**



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



COCO\_train2014\_000000115502\_jpg.rf.0cc  
3e5369d622bc49739a1145f40dd29.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg



Datacluster-Auto--28-\_jpg.rf.c56f6719fbc095  
9a4b2e123d32a90c65.jpg

**Recommendations**

No automated recommendations generated for this feature

## ***Texture Smoothness***

Feature texture\_smoothness results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	texture_smoothness
n_sampled	100
mean_laplacian_variance	0.03446014943690364
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Time Of Day Mapping***

Feature time\_of\_day\_mapping results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	time_of_day_mapping
status	no_timestamps

### ***Recommendations***

No automated recommendations generated for this feature

## ***Train Ready Scoring***

Feature train\_ready\_scoring results.

### ***Key findings***

- Mean score 0.451

Metric	Value
feature	train_ready_scoring
mean_score	0.45135621340298915
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

## ***Triplet Cooccurrence***

Feature triplet\_cooccurrence results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	triplet_cooccurrence
status	ok

### ***Recommendations***

No automated recommendations generated for this feature



## ***Unsafe Formats***

Feature unsafe\_formats results.

### ***Key findings***

- No automatic takeaways generated for this feature

Metric	Value
feature	unsafe_formats
scanned	1000
status	ok

### ***Recommendations***

No automated recommendations generated for this feature

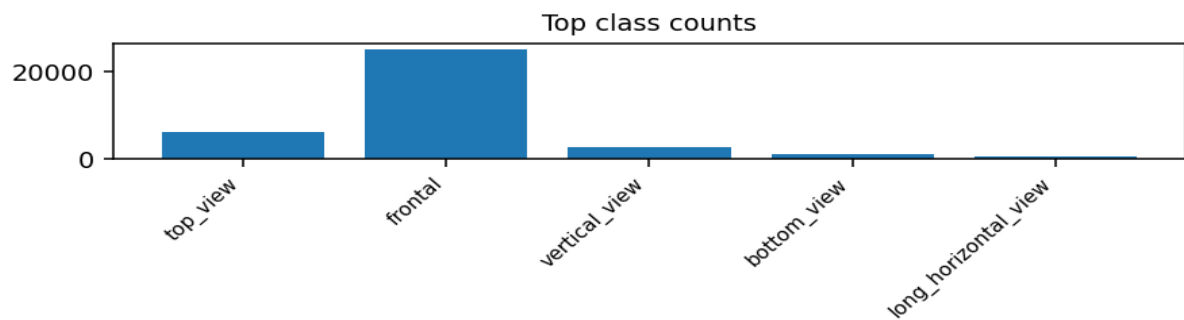
## Viewpoint Variety

Feature viewpoint\_variety results.

### Key findings

- Top classes frontal:25044, top\_view:6216, vertical\_view:2812, bottom\_view:1048, long\_horizontal\_view:504

Metric	Value
feature	viewpoint_variety
status	ok



### Sample evidence



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



20210528\_15\_39\_07\_000\_kSBbVSEI9VV0  
elUslY8iz5Ua7N03\_F\_3264\_2448\_jpg.rf.b6  
d73b1952967995b2882c3038b3d004.jpg



frame\_1135\_jpg.rf.23c830c4a130537c860a  
d92df3f3da6e.jpg

### Recommendations

No automated recommendations generated for this feature

## Notes and next steps

This report was generated by CVEDA. Review feature sections to see detailed evidence and recommendations.