

Driver Monitoring System — Performance Analysis Report

This comprehensive report evaluates the Driver Monitoring System (DMS) across multiple performance dimensions including face detection, gaze estimation, drowsiness detection, and distraction classification.

Report Configuration: - Dataset: ds_dms_eval_2024_v2 - Model Version: DMS-Core-v3.2.1 - Evaluation Date: 2024-12-15

Executive Summary

This section provides high-level metrics suitable for external stakeholders.

Overall system accuracy across all detection tasks.

Metric	Value
Overall System Accuracy	96.7 %

99th percentile inference latency on target hardware.

Metric	Value
P99 Latency	23.4 ms

Summary of key performance indicators for executive review.

Category	Metric	Value	Target	Status
Detection	Face Detection	98.0%	>95%	PASS
Detection	Eye Detection	96.5%	>95%	PASS
Estimation	Gaze Accuracy	4.2°	<5°	PASS
Estimation	Pose Accuracy	3.1°	<5°	PASS
Classification	Drowsiness	94.5%	>90%	PASS
Classification	Distraction	91.3%	>90%	PASS
Performance	Latency P99	23.4ms	<30ms	PASS

1. Face Detection Performance

1.1 Detection Accuracy

Face detection precision under standard lighting conditions.

Metric	Value
Face Detection Precision	98.2 %

Face detection recall ensuring no missed detections.

Metric	Value
Face Detection Recall	97.8 %

Detection performance breakdown by lighting condition (day, night, mixed).

Lighting Condition	Precision	Recall	F1 Score	Samples
Daylight	98.5	98.2	98.3	45000
Night (IR)	97.2	96.8	97	38000
Dusk/Dawn	96.8	96.1	96.4	22000
Tunnel Exit	94.5	93.8	94.1	8000
Direct Sunlight	95.2	94.6	94.9	12000

Analysis Notes: - Night performance shows 3% degradation compared to day-time - IR illumination compensates effectively for low-light conditions - Mixed lighting (tunnel exits) remains the most challenging scenario

1.2 Pose Estimation

Yaw angle estimation accuracy across the operational range.

Yaw Range	MAE (degrees)	Std Dev	Samples
-90° to -60°	5.8	2.1	8500
-60° to -30°	3.2	1.4	22000
-30° to 0°	2.1	0.9	45000
0° to 30°	2	0.8	48000
30° to 60°	3.4	1.5	21000
60° to 90°	6.1	2.3	7500

Pitch angle estimation accuracy for head tilt detection.

Pitch Range	MAE (degrees)	Std Dev
-45° to -30°	4.2	1.8
-30° to -15°	2.8	1.2
-15° to 15°	1.9	0.7

Pitch Range	MAE (degrees)	Std Dev
15° to 30°	2.6	1.1
30° to 45°	4.5	1.9

Distribution of pose estimation errors across the test dataset.



Figure 1: Visualization: Pose Error Dist

2. Gaze Estimation

2.1 Gaze Vector Accuracy

Mean angular error of gaze direction estimation.

Metric	Value
Gaze Angular Error	4.2 degrees

Gaze accuracy breakdown by dashboard region (instrument cluster, center console, mirrors).

Gaze Region	Accuracy (%)	Samples
Road Ahead	96.2	85000
Instrument Cluster	93.8	25000
Center Console	94.5	18000
Left Mirror	98.1	8000
Right Mirror	97.8	7500
Rear View Mirror	95.2	6500

Heatmap showing gaze estimation accuracy across the driver field of view.

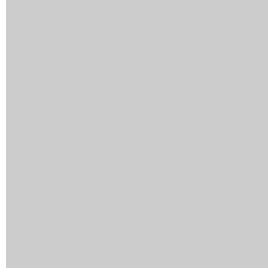


Figure 2: Visualization: Gaze Heatmap

2.2 Gaze Zone Classification

Confusion matrix for gaze zone classification (road, left mirror, right mirror, instrument, phone, other).

Actual Zone	→ Road	→ Instrument	→ Console	→ L Mirror	→ R Mirror	→ Phone
Road	96.2	2.1	0.8	0.3	0.3	0.3
Instrument	3.8	93.8	1.5	0.2	0.2	0.5
Console	1.2	2.3	94.5	0.1	0.1	1.8
L Mirror	0.8	0.3	0.1	98.1	0.5	0.2
R Mirror	0.9	0.4	0.1	0.6	97.8	0.2
Phone	0.5	1.2	3.5	0.1	0.1	94.6

Key Observations: - Road vs. instrument cluster confusion is the primary error mode - Left/right mirror discrimination is highly accurate (>98%) - “Phone” detection specifically important for distraction alerts

3. Drowsiness Detection

3.1 Eye State Analysis

PERCLOS (Percentage of Eye Closure) detection accuracy.

Metric	Value
PERCLOS Accuracy	94.5 %

Blink detection metrics including blink rate and duration analysis.

Metric	Value	Unit	Alert Threshold
Blink Detection Rate	97.2%		>95%
Mean Blink Duration	215	ms	>400ms
PERCLOS (1 min)	8.3%		>15%
Blink Rate	14.2	blinks/min	<8 or >25

ROC curve for drowsiness detection at various thresholds.

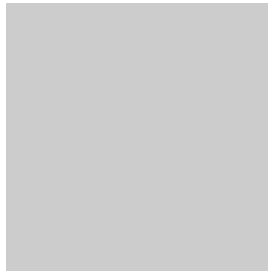


Figure 3: Visualization: Drowsiness Roc

3.2 Yawn Detection

Yawn detection F1 score.

Metric	Value
Yawn Detection F1	0.923

Temporal analysis of yawn detection including onset and duration accuracy.

Detection Phase	Accuracy (%)	Latency (ms)
Onset Detection	94.2	85
Peak Detection	97.8	120
End Detection	91.5	95
Duration Estimation	88.3	0

4. Distraction Classification

4.1 Activity Recognition

Classification accuracy for distracted driving activities.

Activity	Precision	Recall	F1
Phone Usage	95.2	93.3	94.2
Texting	92.1	89.5	90.8
Eating/Drinking	81.2	75.6	78.3
Reaching Back	88.5	85.2	86.8
Talking to Passenger	79.8	82.1	80.9
Adjusting Controls	86.3	84.7	85.5

Confusion matrix visualization for distraction activity classification.

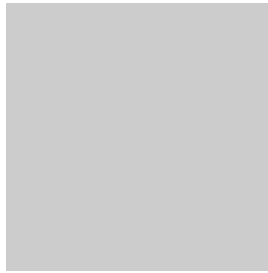


Figure 4: Visualization: Distraction Confusion

Critical Findings: - Phone usage detection achieves 94.2% accuracy - Eating/drinking distinction remains challenging (78.3%) - Talking to passenger vs. phone call differentiation needs improvement

4.2 Attention Score

Composite attention score combining all distraction indicators.

Metric	Value
Attention Score Accuracy	91.3 %

Alert threshold calibration across attention score ranges.

Attention Level	Score Range	System Action	Occurrence (%)
Fully Attentive	90-100	None	78.5
Mild Distraction	70-89	Visual Alert	15.2
Moderate Distraction	50-69	Audio Alert	4.8
Severe Distraction	30-49	Haptic + Audio	1.2
Critical	0-29	Emergency Protocol	0.3

5. System Performance

5.1 Latency Analysis

Latency breakdown by processing stage (detection, landmarks, classification).

Processing Stage	Mean (ms)	P50 (ms)	P99 (ms)
Image Preprocessing	1.2	1.1	2.1
Face Detection	4.5	4.2	6.8
Landmark Extraction	2.8	2.6	4.2
Pose Estimation	3.2	3	5.1
Gaze Estimation	2.1	2	3.5
Classification	1.8	1.7	2.9
Post-processing	0.8	0.7	1.2
Total	16.4	15.3	23.4

Histogram of end-to-end processing latency.



Figure 5: Visualization: Latency Histogram

5.2 Resource Utilization

CPU/GPU utilization and memory footprint during inference.

Resource	Mean Usage	Peak Usage	Limit
CPU	45%	62%	70%
GPU	38%	55%	80%
Memory	512 MB	680 MB	1 GB
NPU	72%	85%	100%

6. Failure Analysis (Internal Only)

This section contains detailed failure analysis for internal engineering review.

6.1 Error Categorization

Breakdown of detection failures by root cause category.

Failure Category	Count	Percentage	Severity
Sunglasses (reflective)	1842	23.1	High
Extreme Pose	1456	18.3	Medium
Motion Blur	1123	14.1	Medium
Occlusion (hand)	987	12.4	Medium
Low Contrast	856	10.7	Low
IR Saturation	723	9.1	High
Multiple Faces	512	6.4	Low
Other	471	5.9	Low

Engineering Notes: - Sunglasses remain the top failure mode (23% of all failures) - Consider adding polarized lens detection as a feature flag - IR reflection issues in certain sunglasses types identified

6.2 Edge Cases

Catalog of identified edge cases with reproduction steps.

Case ID	Description	Frequency	Status
EC-001	Polarized sunglasses with IR reflection	Rare	Under Investigation
EC-002	Face tattoos affecting landmark detection	Very Rare	Known Issue
EC-003	Extreme backlight during sunset	Occasional	Mitigated
EC-004	Medical eye patches	Very Rare	Documented
EC-005	Face masks (post-COVID)	Common	Resolved

Visual examples of challenging edge cases.

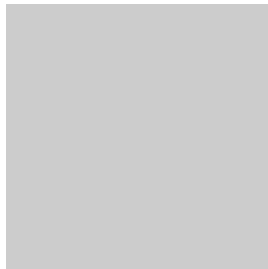


Figure 6: Visualization: Edge Cases Examples

6.3 Debug Information

Internal feature statistics for model debugging.

Feature	Mean	Std	Min	Max
face_confidence	0.92	0.08	0.51	0.99
eye_aspect_ratio	0.28	0.05	0.12	0.45
mouth_aspect_ratio	0.15	0.12	0.02	0.68
head_pose_yaw	-2.3	18.5	-85.2	82.1
gaze_pitch	-8.5	12.3	-45	35.2

7. Experimental Results (Draft)

These results are from experimental features not yet validated for production.

7.1 Multi-Occupant Detection

Preliminary accuracy for multi-occupant detection.

Metric	Value
Multi-Occupant Detection	87.2 %

Detection accuracy by seating position (driver, front passenger, rear).

Seating Position	Detection Rate	Notes
Driver	99.2	Primary target
Front Passenger	94.5	Good coverage
Rear Left	78.3	Limited FOV
Rear Center	65.2	Occluded
Rear Right	76.8	Limited FOV

7.2 Emotion Recognition (WIP)

Preliminary emotion recognition results - work in progress.

Emotion	Accuracy	Status
Neutral	82.5	WIP
Happy	75.2	WIP
Angry	68.3	WIP
Surprised	71.8	WIP

Emotion	Accuracy	Status
Drowsy	88.2	Validated

8. Recommendations

Based on the analysis, we recommend the following improvements:

1. **Sunglasses Handling:** Implement adaptive IR intensity for sunglasses detection
2. **Gaze Calibration:** Add per-driver gaze calibration for improved accuracy
3. **Latency Optimization:** Target 15ms reduction through model pruning
4. **Edge Case Coverage:** Expand training data for identified failure modes

Appendix A: Methodology

A.1 Dataset Description

The evaluation dataset (ds_dms_eval_2024_v2) contains: - 150,000 annotated frames - 500 unique subjects - 12 lighting conditions - 8 vehicle types

A.2 Evaluation Protocol

All metrics were computed using: - 5-fold cross-validation - Stratified sampling by subject - Held-out test set (20% of data)

Appendix B: Detailed Tables (Verbose)

B.1 Per-Subject Performance

Performance metrics broken down by individual test subjects.

Subject ID	Age Group	Gender	Accuracy	Sessions
S0000	18-25	M	90	3
S0001	26-35	F	91	4
S0002	36-45	M	92	5
S0003	46-55	F	93	6
S0004	56+	M	94	7
S0005	18-25	F	95	3
S0006	26-35	M	96	4
S0007	36-45	F	97	5

Subject ID	Age Group	Gender	Accuracy	Sessions
S0008	46-55	M	98	6
S0009	56+	F	99	7
S0010	18-25	M	90	3
S0011	26-35	F	91	4
S0012	36-45	M	92	5
S0013	46-55	F	93	6
S0014	56+	M	94	7
S0015	18-25	F	95	3
S0016	26-35	M	96	4
S0017	36-45	F	97	5
S0018	46-55	M	98	6
S0019	56+	F	99	7
S0020	18-25	M	90	3
S0021	26-35	F	91	4
S0022	36-45	M	92	5
S0023	46-55	F	93	6
S0024	56+	M	94	7
S0025	18-25	F	95	3
S0026	26-35	M	96	4
S0027	36-45	F	97	5
S0028	46-55	M	98	6
S0029	56+	F	99	7
S0030	18-25	M	90	3
S0031	26-35	F	91	4
S0032	36-45	M	92	5
S0033	46-55	F	93	6
S0034	56+	M	94	7
S0035	18-25	F	95	3
S0036	26-35	M	96	4
S0037	36-45	F	97	5
S0038	46-55	M	98	6
S0039	56+	F	99	7
S0040	18-25	M	90	3
S0041	26-35	F	91	4
S0042	36-45	M	92	5
S0043	46-55	F	93	6
S0044	56+	M	94	7
S0045	18-25	F	95	3
S0046	26-35	M	96	4
S0047	36-45	F	97	5
S0048	46-55	M	98	6
S0049	56+	F	99	7
S0050	18-25	M	90	3
S0051	26-35	F	91	4

Subject ID	Age Group	Gender	Accuracy	Sessions
S0052	36-45	M	92	5
S0053	46-55	F	93	6
S0054	56+	M	94	7
S0055	18-25	F	95	3
S0056	26-35	M	96	4
S0057	36-45	F	97	5
S0058	46-55	M	98	6
S0059	56+	F	99	7
S0060	18-25	M	90	3
S0061	26-35	F	91	4
S0062	36-45	M	92	5
S0063	46-55	F	93	6
S0064	56+	M	94	7
S0065	18-25	F	95	3
S0066	26-35	M	96	4
S0067	36-45	F	97	5
S0068	46-55	M	98	6
S0069	56+	F	99	7
S0070	18-25	M	90	3
S0071	26-35	F	91	4
S0072	36-45	M	92	5
S0073	46-55	F	93	6
S0074	56+	M	94	7
S0075	18-25	F	95	3
S0076	26-35	M	96	4
S0077	36-45	F	97	5
S0078	46-55	M	98	6
S0079	56+	F	99	7
S0080	18-25	M	90	3
S0081	26-35	F	91	4
S0082	36-45	M	92	5
S0083	46-55	F	93	6
S0084	56+	M	94	7
S0085	18-25	F	95	3
S0086	26-35	M	96	4
S0087	36-45	F	97	5
S0088	46-55	M	98	6
S0089	56+	F	99	7
S0090	18-25	M	90	3
S0091	26-35	F	91	4
S0092	36-45	M	92	5
S0093	46-55	F	93	6
S0094	56+	M	94	7
S0095	18-25	F	95	3

Subject ID	Age Group	Gender	Accuracy	Sessions
S0096	26-35	M	96	4
S0097	36-45	F	97	5
S0098	46-55	M	98	6
S0099	56+	F	99	7

B.2 Per-Frame Analysis

Frame-by-frame analysis with confidence scores.

Frame	Confidence	Yaw	Pitch	Gaze X	Gaze Y
F000000	0.85	-30	-15	-20	-10
F000001	0.86	-29	-14	-19	-9
F000002	0.87	-28	-13	-18	-8
F000003	0.88	-27	-12	-17	-7
F000004	0.89	-26	-11	-16	-6
F000005	0.9	-25	-10	-15	-5
F000006	0.91	-24	-9	-14	-4
F000007	0.92	-23	-8	-13	-3
F000008	0.93	-22	-7	-12	-2
F000009	0.94	-21	-6	-11	-1
F000010	0.95	-20	-5	-10	0
F000011	0.96	-19	-4	-9	1
F000012	0.97	-18	-3	-8	2
F000013	0.98	-17	-2	-7	3
F000014	0.99	-16	-1	-6	4
F000015	0.85	-15	0	-5	5
F000016	0.86	-14	1	-4	6
F000017	0.87	-13	2	-3	7
F000018	0.88	-12	3	-2	8
F000019	0.89	-11	4	-1	9
F000020	0.9	-10	5	0	-10
F000021	0.91	-9	6	1	-9
F000022	0.92	-8	7	2	-8
F000023	0.93	-7	8	3	-7
F000024	0.94	-6	9	4	-6
F000025	0.95	-5	10	5	-5
F000026	0.96	-4	11	6	-4
F000027	0.97	-3	12	7	-3
F000028	0.98	-2	13	8	-2
F000029	0.99	-1	14	9	-1
F000030	0.85	0	-15	10	0
F000031	0.86	1	-14	11	1
F000032	0.87	2	-13	12	2

Frame	Confidence	Yaw	Pitch	Gaze X	Gaze Y
F000033	0.88	3	-12	13	3
F000034	0.89	4	-11	14	4
F000035	0.9	5	-10	15	5
F000036	0.91	6	-9	16	6
F000037	0.92	7	-8	17	7
F000038	0.93	8	-7	18	8
F000039	0.94	9	-6	19	9
F000040	0.95	10	-5	-20	-10
F000041	0.96	11	-4	-19	-9
F000042	0.97	12	-3	-18	-8
F000043	0.98	13	-2	-17	-7
F000044	0.99	14	-1	-16	-6
F000045	0.85	15	0	-15	-5
F000046	0.86	16	1	-14	-4
F000047	0.87	17	2	-13	-3
F000048	0.88	18	3	-12	-2
F000049	0.89	19	4	-11	-1
F000050	0.9	20	5	-10	0
F000051	0.91	21	6	-9	1
F000052	0.92	22	7	-8	2
F000053	0.93	23	8	-7	3
F000054	0.94	24	9	-6	4
F000055	0.95	25	10	-5	5
F000056	0.96	26	11	-4	6
F000057	0.97	27	12	-3	7
F000058	0.98	28	13	-2	8
F000059	0.99	29	14	-1	9
F000060	0.85	-30	-15	0	-10
F000061	0.86	-29	-14	1	-9
F000062	0.87	-28	-13	2	-8
F000063	0.88	-27	-12	3	-7
F000064	0.89	-26	-11	4	-6
F000065	0.9	-25	-10	5	-5
F000066	0.91	-24	-9	6	-4
F000067	0.92	-23	-8	7	-3
F000068	0.93	-22	-7	8	-2
F000069	0.94	-21	-6	9	-1
F000070	0.95	-20	-5	10	0
F000071	0.96	-19	-4	11	1
F000072	0.97	-18	-3	12	2
F000073	0.98	-17	-2	13	3
F000074	0.99	-16	-1	14	4
F000075	0.85	-15	0	15	5
F000076	0.86	-14	1	16	6

Frame	Confidence	Yaw	Pitch	Gaze X	Gaze Y
F000077	0.87	-13	2	17	7
F000078	0.88	-12	3	18	8
F000079	0.89	-11	4	19	9
F000080	0.9	-10	5	-20	-10
F000081	0.91	-9	6	-19	-9
F000082	0.92	-8	7	-18	-8
F000083	0.93	-7	8	-17	-7
F000084	0.94	-6	9	-16	-6
F000085	0.95	-5	10	-15	-5
F000086	0.96	-4	11	-14	-4
F000087	0.97	-3	12	-13	-3
F000088	0.98	-2	13	-12	-2
F000089	0.99	-1	14	-11	-1
F000090	0.85	0	-15	-10	0
F000091	0.86	1	-14	-9	1
F000092	0.87	2	-13	-8	2
F000093	0.88	3	-12	-7	3
F000094	0.89	4	-11	-6	4
F000095	0.9	5	-10	-5	5
F000096	0.91	6	-9	-4	6
F000097	0.92	7	-8	-3	7
F000098	0.93	8	-7	-2	8
F000099	0.94	9	-6	-1	9
F000100	0.95	10	-5	0	-10
F000101	0.96	11	-4	1	-9
F000102	0.97	12	-3	2	-8
F000103	0.98	13	-2	3	-7
F000104	0.99	14	-1	4	-6
F000105	0.85	15	0	5	-5
F000106	0.86	16	1	6	-4
F000107	0.87	17	2	7	-3
F000108	0.88	18	3	8	-2
F000109	0.89	19	4	9	-1
F000110	0.9	20	5	10	0
F000111	0.91	21	6	11	1
F000112	0.92	22	7	12	2
F000113	0.93	23	8	13	3
F000114	0.94	24	9	14	4
F000115	0.95	25	10	15	5
F000116	0.96	26	11	16	6
F000117	0.97	27	12	17	7
F000118	0.98	28	13	18	8
F000119	0.99	29	14	19	9
F000120	0.85	-30	-15	-20	-10

Frame	Confidence	Yaw	Pitch	Gaze X	Gaze Y
F000121	0.86	-29	-14	-19	-9
F000122	0.87	-28	-13	-18	-8
F000123	0.88	-27	-12	-17	-7
F000124	0.89	-26	-11	-16	-6
F000125	0.9	-25	-10	-15	-5
F000126	0.91	-24	-9	-14	-4
F000127	0.92	-23	-8	-13	-3
F000128	0.93	-22	-7	-12	-2
F000129	0.94	-21	-6	-11	-1
F000130	0.95	-20	-5	-10	0
F000131	0.96	-19	-4	-9	1
F000132	0.97	-18	-3	-8	2
F000133	0.98	-17	-2	-7	3
F000134	0.99	-16	-1	-6	4
F000135	0.85	-15	0	-5	5
F000136	0.86	-14	1	-4	6
F000137	0.87	-13	2	-3	7
F000138	0.88	-12	3	-2	8
F000139	0.89	-11	4	-1	9
F000140	0.9	-10	5	0	-10
F000141	0.91	-9	6	1	-9
F000142	0.92	-8	7	2	-8
F000143	0.93	-7	8	3	-7
F000144	0.94	-6	9	4	-6
F000145	0.95	-5	10	5	-5
F000146	0.96	-4	11	6	-4
F000147	0.97	-3	12	7	-3
F000148	0.98	-2	13	8	-2
F000149	0.99	-1	14	9	-1
F000150	0.85	0	-15	10	0
F000151	0.86	1	-14	11	1
F000152	0.87	2	-13	12	2
F000153	0.88	3	-12	13	3
F000154	0.89	4	-11	14	4
F000155	0.9	5	-10	15	5
F000156	0.91	6	-9	16	6
F000157	0.92	7	-8	17	7
F000158	0.93	8	-7	18	8
F000159	0.94	9	-6	19	9
F000160	0.95	10	-5	-20	-10
F000161	0.96	11	-4	-19	-9
F000162	0.97	12	-3	-18	-8
F000163	0.98	13	-2	-17	-7
F000164	0.99	14	-1	-16	-6

Frame	Confidence	Yaw	Pitch	Gaze X	Gaze Y
F000165	0.85	15	0	-15	-5
F000166	0.86	16	1	-14	-4
F000167	0.87	17	2	-13	-3
F000168	0.88	18	3	-12	-2
F000169	0.89	19	4	-11	-1
F000170	0.9	20	5	-10	0
F000171	0.91	21	6	-9	1
F000172	0.92	22	7	-8	2
F000173	0.93	23	8	-7	3
F000174	0.94	24	9	-6	4
F000175	0.95	25	10	-5	5
F000176	0.96	26	11	-4	6
F000177	0.97	27	12	-3	7
F000178	0.98	28	13	-2	8
F000179	0.99	29	14	-1	9
F000180	0.85	-30	-15	0	-10
F000181	0.86	-29	-14	1	-9
F000182	0.87	-28	-13	2	-8
F000183	0.88	-27	-12	3	-7
F000184	0.89	-26	-11	4	-6
F000185	0.9	-25	-10	5	-5
F000186	0.91	-24	-9	6	-4
F000187	0.92	-23	-8	7	-3
F000188	0.93	-22	-7	8	-2
F000189	0.94	-21	-6	9	-1
F000190	0.95	-20	-5	10	0
F000191	0.96	-19	-4	11	1
F000192	0.97	-18	-3	12	2
F000193	0.98	-17	-2	13	3
F000194	0.99	-16	-1	14	4
F000195	0.85	-15	0	15	5
F000196	0.86	-14	1	16	6
F000197	0.87	-13	2	17	7
F000198	0.88	-12	3	18	8
F000199	0.89	-11	4	19	9

B.3 Raw Feature Distributions

Raw feature value distributions for model interpretability.

Bin	Face Conf	Eye AR	Mouth AR
0-10	50	30	20
10-20	70	45	30

Bin	Face Conf	Eye AR	Mouth AR
20-30	90	60	40
30-40	110	75	50
40-50	130	90	60
50-60	150	105	70
60-70	170	120	80
70-80	190	135	90
80-90	210	150	100
90-100	230	165	110

Appendix C: Dossier-Only Content

This section contains highly condensed summary for regulatory dossier.

Regulatory compliance summary table.

Regulatory Requirement	Standard	Test Result	Compliance
Driver Attention Detection	UNECE R79	PASS	Compliant
Drowsiness Warning	EU 2019/2144	PASS	Compliant
Distraction Warning	EU 2019/2144	PASS	Compliant
System Availability	ISO 26262	99.2%	ASIL-B
False Positive Rate	Internal	<2%	Compliant

Composite safety score for regulatory filing.

Metric	Value
Composite Safety Score	94.8 %