

STEP C1 — Class Distribution (Object & Track Level)

1. Insight Detail per Sub-step / Output

□ Output C1.1 — Object-Level Class Distribution (BBox Count)

Fakta Penting

- Kelas **car** memiliki **2.207.116 bbox (76,45%)**, mendominasi distribusi object-level.
- Kelas **pedestrian** berada di urutan kedua dengan **381.441 bbox (13,21%)**.
- Seluruh kelas lain masing-masing berada di bawah **6%**, bahkan sebagian besar < **1%**.
- Distribusi bbox menunjukkan pola **long-tail yang sangat tajam**.

Insight

- Dataset sangat tidak seimbang pada level deteksi.
- Sebagian besar sinyal visual berasal dari satu kelas utama.

Ini menunjukkan

- Data dashcam secara alami merekam dominasi kendaraan pribadi di jalan raya.
- Proses pengumpulan data mencerminkan kondisi lalu lintas umum, bukan skenario terkontrol.

Resiko

- Dominasi satu kelas dapat menyebabkan model belajar bias terhadap kelas tersebut dan mengabaikan kelas minoritas.
-

□ Output C1.2 — Track-Level Class Distribution (Unique Tracks)

Fakta Penting

- Kelas **car** memiliki **82.891 track (73,47%)**.
- Kelas **pedestrian** menyumbang **20.738 track (18,38%)**.

- Beberapa kelas memiliki jumlah track yang sangat kecil, seperti **train (24 track)** dan **trailer (97 track)**.

Insight

- Ketimpangan kelas tetap terjadi pada level tracking, tidak hanya pada level deteksi.
- Beberapa kelas sangat jarang muncul sebagai entitas temporal independen.

Ini menunjukkan

- Fenomena long-tail bersifat **spatio-temporal**, bukan hanya spasial.
- Kejadian dunia nyata tertentu (misalnya kereta api) memang jarang tertangkap oleh dashcam kendaraan.

Resiko

- Model tracking berpotensi tidak mampu membangun representasi temporal yang stabil untuk kelas dengan jumlah track sangat terbatas.
-

□ Output C1.3 — Object vs Track Distribution Comparison

Fakta Penting

- Nilai **bbox_per_track** bervariasi antar kelas:
 - car: ~26,6
 - pedestrian: ~18,4
 - train: ~66,3
- Kelas dengan track sedikit cenderung memiliki durasi track lebih panjang.

Insight

- Distribusi kelas pada level bbox tidak identik dengan level track.
- Ada kelas yang jarang muncul tetapi bertahan lama ketika muncul.

Ini menunjukkan

- Karakteristik kelas dipengaruhi oleh sifat objek di dunia nyata (misalnya kereta atau bus yang bergerak lambat dan konsisten).
- Mengukur ketidakseimbangan hanya dari jumlah bbox tidak cukup.

Resiko

- Jika hanya satu level distribusi digunakan, kesulitan tracking kelas tertentu dapat disalahestimasi.

□ Output C1.4 — Visual: BBox vs Track Count (Top Classes)

Fakta Penting

- Visual memperlihatkan gap skala yang ekstrem antara kelas **car** dan kelas lain.
- Perbedaan bbox dan track antar kelas terlihat jelas secara visual.

Insight

- Ketimpangan kelas bersifat struktural dan konsisten di berbagai metrik.
- Visualisasi menegaskan bahwa long-tail bukan artefak statistik.

Ini menunjukkan

- Dataset merepresentasikan kondisi lalu lintas nyata yang tidak seimbang.
- Evaluasi berbasis agregat berisiko menutupi performa buruk pada kelas langka.

Resiko

- Model dengan metrik global dapat tampak “baik” meskipun gagal pada kelas minoritas.
-

□ Output C1.5 — Minority Class Identification (Track-Level)

Fakta Penting

- Threshold minoritas (p20 track_count) adalah **225 track**.
- Kelas minoritas teridentifikasi:
 - **train**
 - **trailer**
 - **other person**
- Kelas-kelas ini juga memiliki proporsi bbox yang sangat kecil.

Insight

- Sebagian kelas memiliki keterbatasan data baik secara spasial maupun temporal.
- Kelas minoritas membentuk ekor panjang distribusi tracking.

Ini menunjukkan

- Beberapa skenario lalu lintas nyata sangat jarang terjadi pada data dashcam kendaraan pribadi.
- Informasi temporal untuk kelas tersebut sangat terbatas.

Resiko

- Model berpotensi gagal mengenali atau melacak kelas minoritas secara konsisten.
-

2 □ Insight Kesimpulan (Naratif — Terhubung & Konteks Dunia Nyata)

Pada sub-step C1.1, distribusi bbox menunjukkan dominasi ekstrem kelas car, yang selaras dengan realitas bahwa kamera dashcam kendaraan paling sering merekam mobil lain di jalan. Situasi ini mencerminkan kondisi lalu lintas sehari-hari, di mana kendaraan pribadi merupakan objek utama dalam field of view.

Sub-step C1.2 memperlihatkan bahwa dominasi ini tetap bertahan pada level tracking, menandakan bahwa ketimpangan kelas bukan hanya akibat frekuensi kemunculan sesaat, tetapi juga berlanjut secara temporal. Hal ini konsisten dengan pengalaman dunia nyata, di mana interaksi visual dengan kendaraan lain jauh lebih sering dibanding objek lain seperti kereta atau trailer.

Melalui sub-step C1.3, terlihat bahwa beberapa kelas yang jarang muncul justru memiliki durasi kemunculan yang panjang. Fenomena ini masuk akal dalam konteks lalu lintas, misalnya kereta api yang jarang terlihat oleh dashcam, tetapi tetap berada dalam frame dalam waktu lama ketika muncul.

Visualisasi pada sub-step C1.4 mempertegas bahwa ketimpangan ini bersifat struktural, bukan kebetulan. Sementara itu, sub-step C1.5 menegaskan keberadaan kelas minoritas yang sangat terbatas datanya, mencerminkan skenario lalu lintas yang jarang ditemui oleh kendaraan pribadi. Secara keseluruhan, STEP C1 menunjukkan bahwa kesulitan dataset ini berakar pada ketimpangan alami dunia nyata yang terekam oleh dashcam.

3 □ Penjelasan Masalah yang Terjadi (Rangkuman Risiko & Hipotesis)

- Dataset memiliki ketimpangan kelas yang sangat kuat pada level deteksi dan tracking.
 - Terdapat kelas dengan jumlah track dan bbox yang sangat terbatas.
 - Long-tail bersifat struktural dan berasal dari distribusi kejadian dunia nyata, bukan kesalahan anotasi.
-

4 □ Kaitan dengan Preprocessing dan Modeling

Kaitan dengan Preprocessing

- Distribusi kelas perlu diperhitungkan sebelum melakukan sampling frame atau sequence.
- Statistik object-level dan track-level harus dijaga konsistensinya saat filtering atau batching.

Kaitan dengan Modeling

- Model perlu dievaluasi secara class-aware, bukan hanya menggunakan metrik agregat.
- Ketimpangan bbox vs track mengindikasikan perlunya perhatian khusus pada aspek temporal saat melatih model tracking.
- Performa pada kelas minoritas perlu dianalisis terpisah untuk menghindari bias evaluasi.

FASE 4 — MENCARI INSIGHT

STEP C2 — Object Density per Frame (Crowdedness)

1. Insight Detail per Sub-step / Output

□ Output C2.1 — Object Count per Frame (Construction)

Fakta Penting

- Total **274.149 frame berlabel** dianalisis.
- Perhitungan hanya dilakukan pada **frame yang memiliki label** (window-level).
- Nilai `num_objects` minimum adalah **1 objek per frame**.

Insight

- Analisis kepadatan objek benar-benar merepresentasikan area yang dianotasi, bukan keseluruhan durasi video.
- Tidak ada asumsi implisit bahwa frame tanpa label bersifat kosong atau negatif.

Ini menunjukkan

- Dataset BDD100K pada split ini menggunakan **window annotation strategy**, sehingga kepadatan objek hanya terdefinisi pada segmen tertentu dari video.
- Interpretasi crowdedness harus selalu dikaitkan dengan window berlabel, bukan full scene driving.

Resiko

- Jika frame tanpa label diasumsikan sebagai frame kosong, estimasi kepadatan objek akan bias ke arah sparse.
-

□ Output C2.2 — Global Object Density Statistics

Fakta Penting

- Rata-rata objek per frame adalah **10,53**.
- Median berada di **10 objek**, menunjukkan distribusi relatif simetris di sekitar nilai tengah.

- Nilai ekstrem:
 - $p90 = 19$
 - $p95 = 22$
 - $p99 = 29$
 - maksimum = **62 objek per frame**
- Deviasi standar cukup besar (~**6,09**).

Insight

- Sebagian besar frame berada pada tingkat kepadatan menengah.
- Namun terdapat **ekor distribusi yang panjang** menuju frame-frame dengan kepadatan sangat tinggi.

Ini menunjukkan

- Scene lalu lintas pada dataset ini sering kali moderat, tetapi secara berkala terdapat kondisi padat (misalnya persimpangan, kemacetan).
- Variasi kepadatan mencerminkan dinamika lalu lintas dunia nyata yang tidak homogen.

Resiko

- Model yang hanya dioptimalkan pada kepadatan rata-rata berpotensi gagal pada scene padat yang jarang tetapi krusial.
-

□ Output C2.3 — Crowded Frame Thresholds (Quantile)

Fakta Penting

- Threshold deskriptif kepadatan objek:
 - $q90 = 19 \text{ objek/frame}$
 - $q95 = 22 \text{ objek/frame}$
 - $q99 = 29 \text{ objek/frame}$

Insight

- Definisi “crowded frame” sangat bergantung pada quantile yang dipilih.
- Perbedaan kecil pada threshold menghasilkan perubahan signifikan pada proporsi frame yang dianggap padat.

Ini menunjukkan

- Crowdedness bersifat relatif terhadap distribusi dataset, bukan nilai absolut tetap.
- Tidak ada satu ambang universal yang mewakili seluruh kondisi lalu lintas.

Resiko

- Penggunaan threshold tetap tanpa mempertimbangkan distribusi data dapat menyebabkan over- atau under-estimation terhadap scene padat.
-

□ Output C2.4 — Histogram: Objects per Frame

Fakta Penting

- Histogram menunjukkan puncak distribusi di sekitar **8–12 objek per frame**.
- Frekuensi menurun secara gradual menuju nilai tinggi, membentuk **right-skewed distribution**.
- Frame dengan objek sangat banyak jumlahnya jauh lebih sedikit.

Insight

- Majoritas frame memiliki kepadatan moderat.
- Scene ekstrem padat merupakan minoritas namun tetap signifikan.

Ini menunjukkan

- Dataset merepresentasikan kondisi berkendara normal dengan sesekali menghadapi situasi lalu lintas padat.
- Variasi kepadatan adalah karakter inheren dari data dashcam.

Resiko

- Model dapat cenderung “nyaman” pada kepadatan menengah dan kurang terlatih menghadapi kondisi ekstrem.
-

□ Output C2.5 — ECDF: Objects per Frame

Fakta Penting

- Sekitar **90% frame** memiliki ≤ 19 objek.
- Sekitar **99% frame** memiliki ≤ 29 objek.
- Ekor distribusi terlihat jelas pada bagian atas ECDF.

Insight

- Frame dengan kepadatan sangat tinggi memang jarang, tetapi bukan outlier tunggal.
- Ada proporsi non-trivial frame yang berada di wilayah kepadatan tinggi.

Ini menunjukkan

- Scene crowded adalah bagian dari distribusi alami dataset, bukan noise atau kesalahan anotasi.
- Model akan tetap menemui scene semacam ini saat inference di dunia nyata.

Resiko

- Jika ekor distribusi diabaikan, performa model pada situasi lalu lintas padat dapat menurun drastis.
-

□ Output C2.6 — Crowded vs Sparse Frame Split (Descriptive)

Fakta Penting

- Proporsi frame di atas threshold:
 - $\geq q90 (\geq 19 \text{ objek})$: **10,38%**
 - $\geq q95 (\geq 22 \text{ objek})$: **5,39%**
 - $\geq q99 (\geq 29 \text{ objek})$: **1,07%**
- Frame sangat padat (≥ 29 objek) jumlahnya kecil tetapi konsisten.

Insight

- Scene padat bukan kejadian langka, tetapi juga bukan mayoritas.
- Ada spektrum kepadatan yang jelas dari sparse ke crowded.

Ini menunjukkan

- Dataset mencakup berbagai kondisi lalu lintas, dari lengang hingga sangat padat.
- Tingkat kesulitan inference akan sangat bervariasi antar frame.

Resiko

- Tanpa kesadaran distribusi ini, pipeline dapat mengabaikan sekitar 1–5% frame paling sulit.
-

2 □ Insight Kesimpulan (Naratif — Terhubung & Konteks Dunia Nyata)

Sub-step C2.1 menegaskan bahwa analisis kepadatan objek pada dataset BDD100K hanya relevan pada window berlabel, yang mencerminkan pendekatan anotasi selektif terhadap segmen

video tertentu. Hal ini sejalan dengan praktik dunia nyata, di mana anotator fokus pada bagian video yang dianggap informatif atau penting.

Pada sub-step C2.2, statistik global menunjukkan bahwa rata-rata frame dashcam berisi sekitar sepuluh objek, mencerminkan kondisi lalu lintas normal di jalan perkotaan atau jalan raya. Namun, keberadaan ekor distribusi yang panjang mengindikasikan adanya situasi lalu lintas padat seperti persimpangan besar atau kemacetan, yang meskipun jarang, tetap menjadi bagian penting dari pengalaman berkendara nyata.

Sub-step C2.3 dan C2.6 bersama-sama memperlihatkan bahwa crowdedness bersifat relatif dan berlapis. Dalam konteks dunia nyata, perbedaan antara lalu lintas lancar dan macet tidak bersifat biner, melainkan kontinu, dan hal ini tercermin jelas dalam distribusi kuantil kepadatan objek.

Visualisasi pada sub-step C2.4 dan C2.5 memperkuat pemahaman bahwa sebagian kecil frame membawa beban kesulitan yang jauh lebih tinggi. Frame-frame ini kemungkinan merepresentasikan kondisi ekstrem seperti jam sibuk atau area dengan banyak pengguna jalan, yang sering kali menjadi titik kegagalan sistem persepsi kendaraan otomotif.

3 Penjelasan Masalah yang Terjadi (Rangkuman Risiko & Hipotesis)

- Terdapat variasi kepadatan objek yang sangat lebar antar frame.
 - Sebagian kecil frame memiliki kepadatan ekstrem yang jauh di atas rata-rata.
 - Scene padat bersifat inheren terhadap data dashcam dan bukan anomali.
-

4 Kaitan dengan Preprocessing dan Modeling

Kaitan dengan Preprocessing

- Distribusi kepadatan objek perlu dipertimbangkan saat menyusun batching atau sampling frame.
- Frame dengan kepadatan ekstrem sebaiknya tidak tereliminasi hanya karena jumlahnya kecil.

Kaitan dengan Modeling

- Model deteksi dan tracking perlu dievaluasi secara eksplisit pada frame dengan kepadatan tinggi.
- Mekanisme pasca-proses (misalnya NMS) akan menghadapi tekanan terbesar pada frame-frame crowded.

- Evaluasi performa sebaiknya mencakup analisis terpisah antara frame sparse dan crowded.

FASE 4 — MENCARI INSIGHT

STEP C3 — Bounding Box Size Profiling (Small / Medium / Large)

1 Insight Detail per Sub-step / Output

□ Output C3.1 — Bounding Box Area & Ratio Construction (box2d.*)

Fakta Penting

- Kolom bbox yang digunakan adalah `box2d.x1`, `box2d.y1`, `box2d.x2`, `box2d.y2`.
- Kolom kelas yang dipakai adalah `category`.
- Total baris yang terbentuk pada `size_df` adalah **2.886.916 rows** (sinkron dengan total anotasi object valid pada joinable subset).
- Contoh nilai `bbox_area_ratio` yang muncul pada preview berada di rentang kecil (mis. ~0.000396 sampai ~0.018331 pada beberapa contoh baris yang terlihat).

Insight

- Schema bbox sudah terkonfirmasi dan konsisten untuk seluruh analisis size profiling.
- Area ratio digunakan sebagai metrik utama sehingga size profiling bersifat **scale-aware** terhadap resolusi frame.

Ini menunjukkan

- Dataset memakai format BDD100K yang umum (box2d.*) dan siap dianalisis secara kuantitatif tanpa perlu rekonstruksi nested object.
- Menggunakan ratio membuat analisis tidak “terkunci” pada pixel absolut dan lebih kompatibel untuk keputusan modeling.

Resiko

- -
-

□ Output C3.2 — Global BBox Size Distribution (Area Ratio Quantiles)

Fakta Penting

- Jumlah sampel: **2.886.916** bbox.
- Statistik area ratio (bbox_area/frame_area):
 - mean ≈ **0.013128**
 - median (50%) ≈ **0.0023556**
 - 75% ≈ **0.008262**
 - 90% ≈ **0.031782**
 - 95% ≈ **0.067111**
 - 99% ≈ **0.175113**
 - min ≈ **2.265e-05**
 - max ≈ **0.835328**
- Distribusi sangat skewed: median jauh di bawah mean (indikasi ekor kanan panjang).

Insight

- Mayoritas bbox berukuran **sangat kecil relatif terhadap frame** (median ~0.235% dari luas frame).
- Ada ekor panjang menuju bbox besar (hingga menutupi ~83% frame), namun itu adalah minoritas.

Ini menunjukkan

- Scene dashcam cenderung berisi banyak objek jauh (small-object dominated), dengan sebagian kecil kasus objek sangat dekat kamera (large boxes).
- “Extreme tiny bbox” nyata terjadi (min area ratio sangat kecil), yang biasanya menjadi bottleneck kualitas deteksi.

Resiko

- Jika model/training pipeline tidak robust terhadap small objects, performa mAP dapat turun signifikan meski metrik agregat terlihat baik (karena dominasi jumlah bbox kecil).

□ Output C3.3 — Size Bucket Distribution per Class (Quantile-based Small/Medium/Large)

Catatan: bucket dibentuk secara **quantile-based** (batas small/medium/large berdasarkan distribusi area ratio global), bukan threshold pixel absolut.

Fakta Penting

- Distribusi bucket per kelas (persen small/medium/large yang terlihat):
 - **pedestrian:** ~42.80% / 40.20% / 17.00%
 - **rider:** ~39.54% / 32.27% / 28.20%
 - **bicycle:** ~30.71% / 37.60% / 31.70%
 - **car:** ~32.39% / 31.89% / 35.72%

- **truck:** ~23.24% / 31.15% / 45.62%
- **bus:** ~24.13% / 30.23% / 45.64%
- **other vehicle:** ~15.69% / 33.57% / 50.74%
- **trailer:** ~15.57% / 29.22% / 55.20%
- **train:** ~7.29% / 18.84% / 73.87%
- **other person:** ~36.78% / 46.82% / 16.39%
- Dari sisi hitungan, kelas dominan seperti **car** memiliki jumlah bbox sangat besar di semua bucket (small/medium/large).

Insight

- Kelas “manusia” (pedestrian, rider) condong ke **small-object** lebih kuat dibanding kendaraan besar.
- Kelas kendaraan besar (bus, truck, trailer, train) condong ke **large-object** (proporsi large tinggi).
- Ada “profil ukuran khas” per class, bukan sekadar variasi acak.

Ini menunjukkan

- Pola ukuran konsisten dengan karakter objek di dunia nyata: manusia sering muncul kecil (jarak jauh), sedangkan kendaraan besar cenderung menempati area besar ketika dekat kamera.
- Ketidakseimbangan bukan hanya antar class, tapi juga **antar class-size regime** (mis. pedestrian lebih banyak di small regime).

Resiko

- Jika training/evaluasi tidak memperhatikan perbedaan size profile per class, maka kegagalan pada sub-regime (mis. pedestrian kecil) bisa tersamarkan oleh performa kelas/size lain.

□ Output C3.4 — Histogram: BBox Area Ratio

Fakta Penting

- Histogram menunjukkan konsentrasi besar bbox berada sangat dekat nol (area ratio kecil).
- Distribusi memiliki ekor panjang sampai ~0.8+ (sesuai max pada statistik).

Insight

- Dataset sangat **small-object heavy** secara global.
- Ekor panjang mengindikasikan adanya campuran kondisi: objek jauh vs objek sangat dekat.

Ini menunjukkan

- Skala objek dalam data dashcam tidak homogen: model akan menghadapi perubahan skala ekstrem antar frame/video.
- Ini tipikal untuk scene berkendara: ada kendaraan jauh di horizon dan ada kendaraan yang memotong dekat.

Resiko

- Jika preprocessing/training tidak mengakomodasi rentang skala yang lebar, model bisa “overfit” ke rentang skala yang paling sering.
-

□ Output C3.5 — Boxplot: BBox Size by Class (Top Classes)

Fakta Penting

- Boxplot memperlihatkan perbedaan distribusi area ratio antar kelas top (bicycle, bus, car, motorcycle, other vehicle, pedestrian, rider, truck).
- Pedestrian/rider tampak memiliki pusat distribusi lebih rendah dibanding bus/truck/other vehicle.
- Outlier besar muncul di berbagai kelas (hingga ~0.6–0.8 pada beberapa kelas).

Insight

- Per kelas, ada **pergeseran lokasi** (median/kuartil) yang jelas: manusia lebih kecil, kendaraan besar lebih besar.
- Outlier besar lintas kelas mengindikasikan variasi jarak dan perspektif kamera.

Ini menunjukkan

- Ukuran bbox bukan hanya “noise”, tapi sinyal yang terstruktur mengikuti jenis objek dan kondisi jalan (jarak, lane position, occlusion).
- Dataset mengandung skenario ekstrem (objek sangat dekat) yang bisa berpengaruh ke training stabilitas.

Resiko

- Evaluasi global tanpa stratifikasi ukuran dapat menyesatkan: performa bagus di objek besar belum tentu berarti bagus di objek kecil yang lebih menantang.
-

2 Insight Kesimpulan (Naratif — 1 paragraf per sub-step + konteks dunia nyata)

Pada C3.1, format anotasi bbox terbukti mengikuti pola khas BDD100K (box2d.*) sehingga perhitungan ukuran objek bisa dilakukan secara konsisten dan repeatable. Ini penting dalam dataset real-world dashcam karena schema sering berasal dari pipeline anotasi yang kompleks dan nested; ketika schema sudah terkunci, interpretasi hasil EDA menjadi jauh lebih dapat dipercaya.

Pada C3.2, statistik kuantil memperlihatkan dominasi objek kecil yang sangat kuat (median area ratio ~0.00236), yang selaras dengan kenyataan video dashcam: banyak objek muncul jauh di depan kendaraan (kendaraan lain, pejalan kaki di trotoar, pengendara sepeda di sisi jalan) dan hanya sesekali objek berada sangat dekat kamera (misalnya kendaraan memotong jalur atau berhenti di depan). Extreme tiny bbox yang muncul bukan hal aneh di jalan raya—itu merepresentasikan objek yang jauh, kadang hanya beberapa puluh piksel, yang memang menjadi sumber utama kegagalan deteksi.

Pada C3.3, bucket per kelas menunjukkan bahwa ukuran objek sangat bergantung pada jenisnya: pedestrian dan rider lebih sering berada pada rezim kecil, sedangkan bus/truck/trailer/train lebih dominan pada rezim besar. Ini konsisten dengan dunia nyata: manusia cenderung berada di pinggir jalan atau jauh dari ego-vehicle, sementara kendaraan besar ketika muncul biasanya cukup dekat atau cukup besar secara fisik sehingga mengambil area lebih besar. Artinya, “kesulitan” tidak merata—kelas tertentu secara intrinsik lebih sering berada pada kondisi sulit (small regime).

Pada C3.4, histogram menguatkan bahwa keseluruhan dataset lebih berat ke objek kecil dengan ekor panjang menuju objek besar. Dalam konteks berkendara, ini menggambarkan variasi skala yang ekstrem akibat perubahan jarak, kecepatan relatif, dan komposisi scene (misalnya jalan bebas hambatan vs persimpangan padat).

Pada C3.5, boxplot memperlihatkan perbedaan distribusi ukuran antar kelas dan adanya outlier besar lintas kelas. Ini cocok dengan fenomena real-world: objek bisa menjadi sangat besar ketika dekat kamera, termasuk sepeda motor atau mobil saat berada di lane yang sama dan jarak dekat. Dengan demikian, model perlu menghadapi dua dunia sekaligus: banyak objek kecil yang sulit dan sedikit objek besar yang menuntut kestabilan prediksi pada skala besar.

3 Penjelasan Masalah yang Terjadi (Rangkuman Risiko & Hipotesis)

- Dataset sangat **didominasi objek kecil** (small-object heavy) sehingga ada risiko bottleneck performa deteksi pada skala kecil.

- Distribusi ukuran memiliki **rentang ekstrem** (hingga area ratio ~0.835), menunjukkan variasi skala besar yang dapat mempengaruhi stabilitas training.
 - Setiap kelas memiliki **profil ukuran berbeda**, sehingga risiko kegagalan dapat terkonsentrasi pada kelas tertentu (misalnya pedestrian/rider di small regime).
-

4 Kaitan dengan Preprocessing dan Modeling

Kaitan dengan Preprocessing

- Analisis ukuran berbasis **area ratio** dapat digunakan untuk membuat **stratifikasi evaluasi** (misalnya laporan performa per bucket size) tanpa mengubah data.
- Perlu memastikan pipeline batching/sampling tidak “menghilangkan” frame/objek kecil secara tidak sengaja (misalnya akibat resizing agresif atau filter area minimum).

Kaitan dengan Modeling

- Karena dominasi small objects, pemilihan konfigurasi model/training sebaiknya memperhatikan kemampuan menangkap detail skala kecil (misalnya strategi multi-scale atau resolusi input yang memadai).
- Profil ukuran per kelas mengindikasikan perlunya evaluasi **class × size regime**, terutama untuk kelas yang cenderung kecil (pedestrian, rider).
- Ekor ukuran besar menunjukkan model juga harus stabil pada objek dekat kamera; training perlu robust pada variasi skala ekstrem.

FASE 4 — MENCARI INSIGHT

STEP C4 — Aspect Ratio & Shape Bias

1. Insight Detail per Sub-step / Output

□ Output C4.1 — Aspect Ratio Construction (w / h) (PDF hlm. 1)

Fakta Penting

- Aspect ratio dihitung sebagai **w / h** menggunakan kolom `box2d.x1`, `box2d.y1`, `box2d.x2`, `box2d.y2`.
- Total bbox dianalisis: **2.886.916**.
- **Tidak ada** nilai NaN dan **tidak ada** aspect ratio ≤ 0 .
- Contoh nilai awal menunjukkan variasi lebar–tinggi yang nyata (mis. car ~ 2.7 , pedestrian ~ 0.9).

Insight

- Perhitungan aspect ratio bersih dan stabil secara numerik.
- Dataset siap dianalisis untuk bias bentuk tanpa isu teknis dasar.

Ini menunjukkan

- Pipeline anotasi bbox konsisten.
- Variasi bentuk yang muncul adalah karakteristik data, bukan artefak error perhitungan.

Resiko

- -
-

□ Output C4.2 — Global Aspect Ratio Distribution (PDF hlm. 2)

Fakta Penting

- Statistik global aspect ratio:
 - mean $\approx 1,30$
 - median (50%) $\approx 1,19$

- $p90 \approx 2,14$
- $p95 \approx 2,68$
- $p99 \approx 4,18$
- $\text{min} \approx 0,027$
- $\text{max} \approx 79,34$
- Distribusi memiliki **ekor kanan sangat panjang**.

Insight

- Mayoritas bbox relatif mendekati bentuk “persegi-horizontal ringan”.
- Terdapat **shape ekstrem** dengan aspect ratio sangat besar (bbox sangat lebar & tipis).

Ini menunjukkan

- Scene dashcam mengandung objek dengan variasi bentuk ekstrem, misalnya objek memanjang secara horizontal (kendaraan dekat kamera, occlusion parsial, atau artifact perspektif).
- Distribusi bentuk tidak simetris dan tidak sempit.

Resiko

- Model yang tidak robust terhadap aspect ratio ekstrem berpotensi gagal pada sebagian kecil tetapi signifikan bbox.
-

□ Output C4.3 — Aspect Ratio Distribution per Class (*PDF hlm. 3*)

Fakta Penting

- Nilai median (50%) per class (perkiraan):
 - **pedestrian:** ~0,42
 - **rider:** ~0,52
 - **bicycle:** ~0,82
 - **motorcycle:** ~0,82
 - **car:** ~1,28
 - **truck:** ~1,15
 - **bus:** ~1,22
- Pedestrian dan rider konsisten **tinggi & kurus ($w/h < 1$)**.
- Kendaraan besar (car, bus, truck) dominan **lebih lebar daripada tinggi ($w/h > 1$)**.
- Outlier ekstrem muncul di hampir semua kelas (max car ~29, truck ~15, bus ~11).

Insight

- Setiap kelas memiliki **profil bentuk khas**.
- Bias bentuk sangat kuat dan konsisten antar kelas.

Ini menunjukkan

- Bentuk bbox merefleksikan geometri objek dunia nyata (manusia vs kendaraan).
- Distribusi bentuk bukan acak, tetapi terstruktur oleh kelas.

Resiko

- Jika model tidak sensitif terhadap perbedaan bentuk antar kelas, performa bisa timpang (baik di kendaraan, buruk di manusia atau sebaliknya).
-

□ Output C4.4 — Global Aspect Ratio Density (*PDF hlm. 4*)

Fakta Penting

- Kurva density menunjukkan puncak utama sekitar **aspect ratio ~1**.
- Ekor kanan panjang terlihat jelas hingga nilai sangat besar.

Insight

- Distribusi global didominasi bbox “normal”, tetapi terdapat minoritas bbox ekstrem.
- Extreme shape bukan noise tunggal, melainkan bagian dari distribusi.

Ini menunjukkan

- Dalam konteks lalu lintas nyata, mayoritas objek berada pada orientasi wajar, namun kondisi tertentu (jarak dekat, sudut kamera, occlusion) memicu bentuk ekstrem.

Resiko

- Model dapat overfit ke bentuk umum dan mengabaikan bentuk ekstrem yang jarang.
-

□ Output C4.5 — Boxplot: Aspect Ratio by Class (*PDF hlm. 5*)

Fakta Penting

- Boxplot memperlihatkan:
 - Pedestrian & rider: box sempit, median rendah.
 - Car, bus, truck: median lebih tinggi dan IQR lebih lebar.
 - Banyak outlier tinggi terutama pada **car** dan **truck**.

Insight

- Variabilitas bentuk antar kelas sangat berbeda.
- Kendaraan menunjukkan rentang bentuk yang jauh lebih lebar dibanding manusia.

Ini menunjukkan

- Objek kendaraan dalam dashcam mengalami variasi perspektif yang jauh lebih besar (lane, jarak, sudut).
- Manusia relatif konsisten bentuknya (tinggi–kurus).

Resiko

- Tanpa perlakuan tepat, model bisa bias ke bentuk dominan dan kurang adaptif pada variasi besar kendaraan.
-

□ Output C4.6 — Extreme Aspect Ratio Outliers (*PDF hlm. 6*)

Fakta Penting

- Threshold ekstrem:
 - $p01 \approx \mathbf{0,269}$
 - $p99 \approx \mathbf{4,177}$
- Bbox di luar threshold:
 - $p01: \mathbf{28.870 \, bbox} (\sim 1,00\%)$
 - $p99: \mathbf{29.035 \, bbox} (\sim 1,01\%)$
- Total bbox ekstrem ~2% dari data.

Insight

- Shape ekstrem mencakup proporsi kecil tetapi nyata.
- Extreme aspect ratio bukan kejadian sangat langka.

Ini menunjukkan

- Dalam deployment dunia nyata, model akan tetap menemui bbox dengan bentuk ekstrem secara reguler.

Resiko

- Jika shape ekstrem tidak terakomodasi, error akan terkonsentrasi pada subset kecil namun berulang.
-

2 Insight Kesimpulan (Naratif — Terhubung & Konteks Dunia Nyata)

Pada sub-step C4.1, perhitungan aspect ratio menunjukkan bahwa data bbox BDD100K bersih secara numerik dan siap dianalisis. Ini penting karena dalam dataset dunia nyata, error pada anotasi sering kali mencemari analisis bentuk, namun pada kasus ini variasi yang muncul murni berasal dari data.

Sub-step C4.2 mengungkap bahwa meskipun mayoritas bbox memiliki bentuk mendekati seimbang ($w/h \approx 1$), terdapat ekor panjang menuju bentuk ekstrem. Dalam konteks dashcam, ini masuk akal karena kendaraan yang sangat dekat kamera atau terpotong frame dapat menghasilkan bbox yang sangat lebar atau sangat tipis.

Pada C4.3 dan C4.5, perbedaan pola bentuk antar kelas menjadi sangat jelas. Pedestrian dan rider secara konsisten memiliki bbox tinggi–kurus, sementara kendaraan lebih lebar dan variatif. Ini mencerminkan realitas lalu lintas: manusia berdiri tegak dan relatif kecil, sedangkan kendaraan memiliki orientasi dan ukuran yang lebih beragam tergantung posisi dan jarak.

C4.4 dan C4.6 memperkuat bahwa shape ekstrem adalah bagian alami dari distribusi, bukan outlier yang bisa diabaikan. Dalam sistem persepsi kendaraan, kondisi ekstrem seperti kendaraan sangat dekat, sudut kamera tajam, atau occlusion parsial memang sering menjadi titik kegagalan, sehingga temuan ini sangat relevan secara praktis.

3 Penjelasan Masalah yang Terjadi (Rangkuman Risiko & Hipotesis)

- Distribusi aspect ratio sangat bervariasi dengan ekor panjang.
 - Terdapat sekitar **2% bbox dengan bentuk ekstrem**.
 - Setiap kelas memiliki profil bentuk berbeda yang berpotensi menimbulkan bias performa model.
 - Risiko kegagalan tidak merata dan dapat terkonsentrasi pada subset bbox tertentu.
-

4 Kaitan dengan Preprocessing dan Modeling

Kaitan dengan Preprocessing

- Analisis shape dapat digunakan untuk **stratifikasi evaluasi** (normal vs extreme shape).
- Perlu memastikan preprocessing (resize, crop) tidak merusak bbox dengan aspect ratio ekstrem.

Kaitan dengan Modeling

- Model deteksi/tracking perlu dievaluasi lintas variasi aspect ratio, bukan hanya agregat.
- Bias bentuk per kelas menunjukkan pentingnya representasi yang adaptif terhadap orientasi dan bentuk objek.
- Extreme shape cases sebaiknya dianalisis terpisah dalam evaluasi error.

FASE 4 — MENCARI INSIGHT

STEP C5 — Temporal Difficulty (Track Length & Fragmentation)

1. Insight Detail per Sub-step / Output

□ Output C5.1 — Track Length Construction (*PDF hlm. 1*)

Fakta Penting

- Total **unique tracks** yang dianalisis: **112.819 track**.
- Track length didefinisikan sebagai jumlah **frame unik** per `(videoName, track_id)`.
- Track bersifat **lokal per video** dan hanya mencakup **window frame yang berlabel** (bukan full video).
- Contoh track menunjukkan variasi panjang yang besar (misalnya 6 frame hingga 72 frame dalam satu video).

Insight

- Dataset menyediakan jumlah track yang cukup besar untuk analisis tracking-level.
- Track length yang teramati mencerminkan durasi kemunculan objek dalam window anotasi, bukan durasi objek di dunia nyata secara penuh.

Ini menunjukkan

- Kesulitan tracking yang terukur di sini adalah **difficulty relatif terhadap anotasi window**, bukan durasi objek sebenarnya di sepanjang video.
- Analisis temporal harus diinterpretasikan dalam konteks windowing (hasil Block B).

Resiko

- Jika interpretasi track length disamakan dengan durasi objek sesungguhnya, maka kesimpulan tentang stabilitas tracking bisa bias.
-

□ Output C5.2 — Global Track Length Statistics (*PDF hlm. 2*)

Fakta Penting

- Statistik utama track length:
 - mean \approx **25,6 frame**
 - median (50%) = **15 frame**
 - $25\% \leq$ **7 frame**
 - $75\% \leq$ **30 frame**
 - $90\% \leq$ **62 frame**
 - $95\% \leq$ **93 frame**
 - $99\% \leq$ **167 frame**
 - max = **263 frame**
- Distribusi **right-skewed** (mean > median).

Insight

- Mayoritas track bersifat **pendek hingga menengah**, dengan hanya sebagian kecil track panjang.
- Track sangat panjang memang ada, tetapi merupakan minoritas.

Ini menunjukkan

- Dalam konteks dashcam, banyak objek hanya terlihat sebentar (misalnya objek yang melintas cepat, tertutup, atau keluar frame).
- Track panjang kemungkinan berasal dari objek yang berada lama di sekitar ego-vehicle (misalnya kendaraan di depan).

Resiko

- Model tracking dapat terlihat “cukup baik” secara rata-rata, tetapi sebenarnya gagal mempertahankan identitas pada track panjang yang lebih menantang.

□ Output C5.3 — Short Track Analysis (*PDF hlm. 3*)

Fakta Penting

- Proporsi track pendek:
 - ≤ 1 frame: **2,63%**
 - ≤ 2 frame: **6,81%**
 - ≤ 3 frame: **11,22%**
 - ≤ 5 frame: **19,84%**
 - ≤ 10 frame: **38,01%**
- Hampir **40%** track muncul ≤ 10 frame.

Insight

- Track pendek merupakan bagian signifikan dari dataset.
- Hampir seperlima track sangat singkat (≤ 5 frame).

Ini menunjukkan

- Banyak objek sulit dilacak secara temporal karena kemunculannya singkat.
- Kondisi ini realistik pada scene jalan raya: occlusion, objek keluar–masuk frame cepat, atau anotasi window terbatas.

Resiko

- Tracking ID-switch atau track loss mudah terjadi dan sulit dideteksi jika evaluasi terlalu bergantung pada track pendek.
 - Model bisa “terlihat stabil” karena banyak track memang singkat.
-

□ Output C5.4 — Histogram Track Length (*PDF hlm. 4*)

Fakta Penting

- Histogram menunjukkan konsentrasi besar track pada panjang pendek ($< 20\text{--}30$ frame).
- Ekor distribusi memanjang hingga > 200 frame.

Insight

- Distribusi panjang track tidak homogen.
- Track panjang adalah kasus khusus yang relatif jarang.

Ini menunjukkan

- Kesulitan tracking bersifat **heterogen**: sebagian besar mudah (karena pendek), sebagian kecil sangat menantang (karena panjang).

Resiko

- Tanpa analisis stratifikasi berdasarkan track length, kegagalan pada track panjang bisa tertutup oleh dominasi track pendek.
-

□ Output C5.5 — Track Fragmentation Statistics (*PDF hlm. 5*)

Fakta Penting

- Statistik fragmentasi (num_segments):

- mean \approx **1,38 segmen per track**
- median = **1 segmen**
- $90\% \leq$ **2 segmen**
- $95\% \leq$ **3 segmen**
- $99\% \leq$ **6 segmen**
- max = **29 segmen**
- Mayoritas track **tidak terfragmentasi** (1 segmen).

Insight

- Secara umum, track bersifat cukup kontinu dalam window anotasi.
- Namun terdapat ekor panjang track yang sangat terfragmentasi.

Ini menunjukkan

- Untuk sebagian besar objek, anotasi tracking relatif stabil.
- Kasus fragmentasi tinggi kemungkinan berkaitan dengan occlusion berat, objek kecil, atau objek keluar-masuk frame berulang.

Resiko

- Track dengan fragmentasi tinggi berpotensi menyebabkan ID-switch dan kesalahan asosiasi serius pada sistem tracking.
-

□ Output C5.6 — Histogram Fragmentation (*PDF hlm. 6*)

Fakta Penting

- Distribusi `num_segments` sangat terpusat di 1–2 segmen.
- Track dengan segmen banyak (>5) sangat jarang.

Insight

- Fragmentasi ekstrem adalah minoritas, tetapi nyata.
- Tracking difficulty tinggi terkonsentrasi pada subset kecil track.

Ini menunjukkan

- Dalam praktik dunia nyata, sebagian besar objek mudah diikuti secara temporal, tetapi sebagian kecil sangat sulit dan menjadi sumber error dominan.

Resiko

- Evaluasi rata-rata dapat mengabaikan kegagalan kritis pada track yang kompleks.

2 □ Insight Kesimpulan (Naratif — Terhubung & Konteks Dunia Nyata)

Hasil STEP C5 menunjukkan bahwa secara temporal, dataset BDD100K memiliki karakter campuran antara track pendek yang sangat dominan dan sejumlah kecil track panjang yang berpotensi sulit. Median track length hanya 15 frame, menandakan bahwa dalam window anotasi, banyak objek hanya muncul sebentar. Ini sangat sesuai dengan realitas dashcam, di mana objek sering terhalang, melintas cepat, atau keluar dari field of view.

Proporsi track pendek yang besar (sekitar 40% \leq 10 frame) mengindikasikan bahwa banyak kasus tracking sebenarnya “trivial” karena durasi yang singkat, sementara tantangan sesungguhnya berada pada subset track yang lebih panjang. Track panjang ini biasanya merepresentasikan kendaraan di depan ego-vehicle atau objek yang bergerak searah dan relatif stabil.

Dari sisi fragmentasi, mayoritas track bersifat kontinu, namun adanya track dengan fragmentasi tinggi (hingga 29 segmen) menegaskan keberadaan skenario dunia nyata yang sulit, seperti occlusion berulang, objek kecil, atau interaksi kompleks antar objek. Dengan demikian, kesulitan tracking dalam dataset ini tidak merata, melainkan terkonsentrasi pada kasus-kasus tertentu yang realistik dan kritis.

3 □ Penjelasan Masalah yang Terjadi (Rangkuman Risiko & Hipotesis)

- Mayoritas track pendek → risiko **overestimasi performa tracking**.
 - Track panjang yang jarang namun penting → berpotensi menjadi sumber utama kegagalan ID consistency.
 - Fragmentasi tinggi pada sebagian kecil track → indikasi kondisi sulit (occlusion, small object, crowded scene).
 - Kesulitan tracking bersifat **long-tail problem**, bukan masalah merata.
-

4 □ Kaitan dengan Preprocessing dan Modeling

Kaitan dengan Preprocessing

- Analisis temporal dapat digunakan untuk **stratifikasi evaluasi** berdasarkan track length dan fragmentasi.
- Window-based nature perlu selalu diingat agar tidak salah tafsir durasi track.

Kaitan dengan Modeling

- Evaluasi tracking sebaiknya tidak hanya berbasis agregat, tetapi juga fokus pada **track panjang dan terfragmentasi**.
- Baseline tracker perlu diuji secara eksplisit pada subset track sulit (panjang & fragmented).
- Informasi ini relevan untuk menentukan apakah pendekatan tracking sederhana cukup atau perlu mekanisme asosiasi temporal yang lebih kuat.

Matriks 1 — Kumpulan Insight Informatif (Bukan Problem)

Area	Insight Informatif (Disatukan)	Bukti ringkas dari C1–C5
Class distribution (object-level)	Distribusi bbox sangat didominasi car ; kelas lain mengikuti pola long-tail.	C1: car ~76% bbox; pedestrian ~13%; sisanya jauh lebih kecil.
Class distribution (track-level)	Distribusi track juga long-tail; car dan pedestrian mendominasi track.	C1: car ~73% track; pedestrian ~18%.
Bbox vs track behavior	Distribusi bbox ≠ distribusi track; beberapa kelas jarang tetapi saat muncul cenderung bertahan lama.	C1: variasi bbox_per_track antar kelas (mis. train tinggi).
Frame crowdedness (global)	Mayoritas frame berada pada kepadatan menengah , dengan ekor panjang ke scene padat.	C2: mean ~10.5 obj/frame; p90 ~19; p99 ~29; max ~62.
Crowded frame frequency	Scene padat adalah minoritas namun nyata dan terukur (berlapis berdasarkan quantile).	C2: $\geq q90 \sim 10\%$; $\geq q95 \sim 5\%$; $\geq q99 \sim 1\%$.
Bbox size (global)	Dataset small-object heavy ; median bbox area ratio sangat kecil, namun ada ekor panjang sampai bbox besar.	C3: median area_ratio ~0.00236; p99 ~0.175; max ~0.835.
Bbox size by class	Tiap kelas punya “profil ukuran” khas: manusia cenderung kecil; kendaraan besar cenderung large.	C3: pedestrian small tinggi (~43%); bus/truck large tinggi (~46%); train dominan large (~74%).
Aspect ratio (global)	Aspect ratio mayoritas di sekitar ~1, tapi memiliki ekor kanan panjang (shape ekstrem ada).	C4: median ~1.19; p99 ~4.18; max sangat besar (~79).
Shape bias by class	Pedestrian/rider cenderung tinggi-kurus ($w/h < 1$); kendaraan lebih lebar ($w/h > 1$).	C4: median pedestrian ~0.42; rider ~0.52; car ~1.28, bus ~1.22.
Extreme shape prevalence	Extreme shape bukan outlier tunggal; ~2% bbox berada di ekstrem (p01/p99).	C4: p01 ~0.269 & p99 ~4.177 → ~1% masing-masing.
Track length (global)	Mayoritas track pendek-menengah, dengan ekor panjang sampai ratusan frame (dalam window label).	C5: median 15; mean ~25.6; p99 167; max 263.
Short-track dominance	Track sangat pendek cukup besar proporsinya.	C5: ≤ 10 frame ~38%; ≤ 5 frame ~19.8%.
Fragmentation (global)	Mayoritas track tidak terfragmentasi (1 segmen), tetapi ada ekor kasus kompleks.	C5: median segmen=1; mean ~1.38; p99 ≤6; max 29 segmen.
Heterogeneous difficulty	Kesulitan dataset bersifat “campuran”: banyak kasus relatif mudah + subset kecil kasus sangat sulit (crowded, tiny, fragmented).	Konsisten dari C2 (tail), C3 (tiny), C5 (long-tail frag).

Area	Insight Informatif (Disatukan)	Bukti ringkas dari C1–C5
Konteks dashcam	Pola ini konsisten dengan dunia nyata: objek jauh banyak (small), kondisi macet sesekali (crowded), occlusion/masuk-keluar frame memicu track pendek/frag.	Konsistensi interpretasi C2–C5 dalam domain dashcam.

Matriks 2 — Kumpulan Masalah (Risk Register Block C)

Observasi (Problem)	Resiko (Jika tidak ditangani)	Hipotesis (Penyebab)	Saran penanganan (≥ 1)
Class imbalance sangat tajam (bbox & track long-tail).	Model bias ke kelas dominan; kelas minoritas underfit; metrik agregat menyesatkan.	Distribusi lalu lintas nyata memang didominasi car; pengambilan data tidak seimbang.	(1) Evaluasi per-class wajib. (2) Reweighting loss / sampling berbasis class. (3) Two-stage training / curriculum.
Minority class sangat kecil di track-level (mis. train, trailer, other person di bawah threshold).	Generalisasi buruk pada kelas langka; performa tracking kelas ini tidak stabil.	Kejadian nyata jarang; sedikit contoh temporal untuk belajar asosiasi.	(1) Laporkan metrik terpisah kelas minoritas. (2) Oversampling sequence yang memuat kelas minoritas. (3) Jika perlu, merge/ignore class sangat langka tergantung objective.
Crowded tail nyata (p99 ~29 obj/frame, max 62).	NMS/association makin sulit; false positive/ID switch meningkat pada scene padat.	Scene persimpangan/macet; banyak overlap antar objek.	(1) Buat evaluasi khusus untuk frame di atas q95/q99. (2) Tuning post-process (NMS/association) berbasis kondisi crowded. (3) Data augmentation yang mensimulasikan crowding.
Small-object heavy (median area_ratio sangat kecil; extreme tiny ada).	mAP turun pada objek kecil; recall rendah; tracking putus-putus untuk objek kecil.	Objek jauh dominan pada dashcam; resolusi input efektif kecil untuk tiny bbox.	(1) Uji multi-scale / resolusi input lebih tinggi. (2) Augmentasi crop/scale untuk memperkaya small regime. (3) Evaluasi stratifikasi size (small/med/large).
Profil ukuran berbeda per class (pedestrian lebih small; kendaraan besar lebih large).	Model tampak bagus di vehicle tetapi gagal di pedestrian/rider kecil; bias kelas-size.	Geometri dunia nyata: manusia jauh & kecil, kendaraan punya skala lebih variatif.	(1) Analisis error per class×size. (2) Sampling yang memastikan representasi pedestrian kecil. (3) Loss yang menekankan small-object.
Aspect ratio memiliki ekor ekstrem (max sangat tinggi; ~2% extreme).	Anchor/priors mismatch (jika anchor-based); regresi bbox tidak stabil pada shape ekstrem; false localization.	Occlusion/perspektif/objek terpotong frame menghasilkan bbox tipis/lebar.	(1) Evaluasi terpisah untuk extreme AR (p01/p99). (2) Jika anchor-based, pertimbangkan anchor tuning. (3) Augmentasi yang mempertahankan extreme AR.

Observasi (Problem)	Resiko (Jika tidak ditangani)	Hipotesis (Penyebab)	Saran penanganan (≥ 1)
Shape bias kuat antar kelas (pedestrian kurus, kendaraan lebar).	Model bisa bias terhadap bentuk dominan; beberapa kelas sulit jika representasi bentuk tidak adaptif.	Bentuk objek inherent; sudut pandang dashcam memperkuat perbedaan.	(1) Pastikan model/heads robust terhadap variasi AR. (2) Per-class calibration / analysis.
Track pendek dominan (≤ 10 frame $\sim 38\%$).	Evaluasi tracking bisa bias “terlihat bagus” karena track pendek mudah; model lemah pada long-term association.	Window labeling membatasi durasi; objek masuk-keluar frame cepat; occlusion.	(1) Stratifikasi evaluasi tracking berdasarkan track length. (2) Fokus analisis pada track panjang. (3) Sampling sequence dengan long tracks untuk training.
Fragmentation tail ada (max 29 segmen).	ID-switch/association error tinggi pada subset sulit; performa tracking tidak stabil di kondisi kompleks.	Occlusion berat, crowded, objek kecil, anotasi window memotong track.	(1) Analisis subset track dengan segmen tinggi. (2) Robust association (motion+appearance) bila tersedia. (3) Evaluasi dan debugging visual di Block D5.
Heterogeneous difficulty (mix kasus mudah + subset sulit).	Model overfit pada mayoritas kondisi “mudah”, gagal di edge cases yang menentukan kualitas nyata.	Distribusi dunia nyata long-tail; sampling training tidak menekankan tail.	(1) Buat evaluation suite berbasis strata: crowded/tiny/long/fragmented. (2) Hard-example mining. (3) Perlu audit visual terarah (Block D2/D5).