🧬 Interpretation of `enhanced_invasion_metrics.csv`

Each **row** = one `(dataset, sample, timepoint)`
Each **column** = a **quantitative feature** describing invasion and cell behavior at that specific timepoint.

🧱 1. Identification Columns

Column	Meaning
dataset	Source group (e.g., `train_S1`, `test_S2`). Represents condition or treatment.
sample	Biological replicate or specific experimental run within that dataset.
timepoint	Temporal frame of invasion (e.g., t0, t1, … t12). Used to track progression.

📏 2. Radial Spread Metrics

Describe **how far cells have migrated** from the invasion center.

Column	Meaning	Interpretation per sample
mean_radius	Average distance of cells from the invasion center.	Larger → cells on average migrated farther. Indicates invasion extent.
median_radius	Midpoint of radial distances (less sensitive to outliers).	Shows the “typical” cell distance.
max_radius	Maximum cell distance from center.	Indicates the most invasive cell in the population.
min_radius	Minimum distance (usually near the core).	Helps assess how compact or hollow the core region is.
sd_radius	Standard deviation of distances.	High = variable invasion (some cells far, some near).
p25_radius, p75_radius	25th and 75th percentiles of distances.	Spread between them = heterogeneity in migration.

p90_radius, p95_radius	90th and 95th percentiles.	Capture the extreme invasive front beyond the bulk population.

🧠 **Together:** These describe the **spatial extent** and **variability** of invasion.
For each sample, increasing radii and SD over time show progressive invasion and dispersion.

🦸 3. Leader Cell Metrics

Quantify **cells at the invasive front** that guide migration.

Column	Meaning	Interpretation per sample
n_leader_cells	Number of cells beyond 90th percentile distance (top 10%).	More leaders → stronger invasive front.
mean_leader_radius	Average distance of leader cells.	Indicates how far the most aggressive subset has moved.

🧠 **Together:** Track **frontline invasion dynamics**.
Rising leader-cell numbers or distance suggest increasing aggressiveness or collective invasion.

📊 4. Distribution Shape Metrics

Describe **how uniform or skewed** the spatial distribution is.

Column	Meaning	Interpretation per sample
radius_skewness	Asymmetry of radius distribution.	Positive → long tail of outliers (few very invasive cells).
radius_kurtosis	Peakedness of distribution.	High → many cells near mean (tight cluster). Low → flatter, more dispersed population.

🧠 **Together:** These quantify **population structure** — whether invasion is homogeneous or driven by a few outliers.

🌐 5. Spatial Clustering Metrics

Characterize **how cells are spatially arranged**.

Column	Meaning	Interpretation per sample
dispersion_index	Ratio of variance to mean in spatial distribution.	>1 → clustered; ≈1 → random; <1 → uniform spacing.
mean_nn_dist	Mean nearest-neighbor distance.	Larger → cells spread apart (less cohesive).
median_nn_dist	Median of nearest-neighbor distances.	Similar to mean but more robust to outliers.
sd_nn_dist	Variation in neighbor spacing.	High = irregular, uneven cell spacing.

🧠 **Together:** Reveal **cell-cell interaction** and invasion mode — cohesive vs dispersed migration.

🧫 6. Morphological Metrics

Describe **individual cell shapes and volumes**.

Column	Meaning	Interpretation per sample
mean_volume	Average cell volume.	Larger → cell swelling or active spreading.
sd_volume	Volume variability.	Reflects heterogeneity in cell states.
mean_compactness	Compactness = (surface area)² / volume.	High = rounder, less spread cells; low = elongated or irregular.
sd_compactness	Variation in compactness.	Indicates morphological diversity.
mean_extent	Extent = cell bounding box occupancy ratio.	High = well-filled shape (less protrusive).
sd_extent	Variation in extent.	Measures irregularity in shape and spread.

🧠 **Together:** Show **cell morphology evolution** — e.g., compact clusters early → elongated invasive cells later.

🧍‍♀️ 7. Population and Density Metrics

Summarize **overall population size and invasion volume**.

Column	Meaning	Interpretation per sample
cell_count	Total number of segmented cells at that timepoint.	Indicates proliferation or cell loss.
invasion_volume	Total 3D volume occupied by all cells.	Reflects the physical extent of invasion.
cell_density	cell_count / invasion_volume.	High = dense cluster; low = dispersed invasion.

🧠 **Together:** Capture **macro-level expansion** — how the colony's overall footprint changes over time.

🧩 Summary: What It Explains per Sample

Each **sample** represents one invasion experiment under a specific condition (dataset).
Across **timepoints**, these 26 metrics describe **how that invasion evolves**, in terms of:

Aspect	What It Tells You
Spatial Spread (radius metrics)	How far and variably cells migrate from the center.
Leader Cells	How many and how active the invasive front cells are.
Distribution Shape	Whether invasion is dominated by few leaders or broad migration.
Spatial Clustering	Whether cells move collectively (clustered) or individually (dispersed).
Morphology	Changes in cell shape associated with invasion (compact → elongated).
Population/Density	How cell number and occupied volume evolve (growth vs spread).

🧠 Interpreting Time-Series Behavior per Sample

- **Early timepoints:**
    - Low mean/max radius, high density → compact spheroid.
    - Few leader cells → invasion not yet initiated.
- **Mid timepoints:**
    - Rising SD and percentiles → spread begins.
    - Leader cells appear → invasion front forms.
    - Compactness decreases, volume increases.
- **Late timepoints:**
    - High mean/max radius and dispersion index → full invasion spread.
    - Lower density → cells scattered in matrix.
    - Morphological diversity peaks (high SDs).

🧬 In Short

Each row in your CSV captures the **quantitative snapshot** of invasion at a given timepoint for a specific sample — how far, how fast, how coordinated, and how morphologically variable the cells have become as invasion progresses.