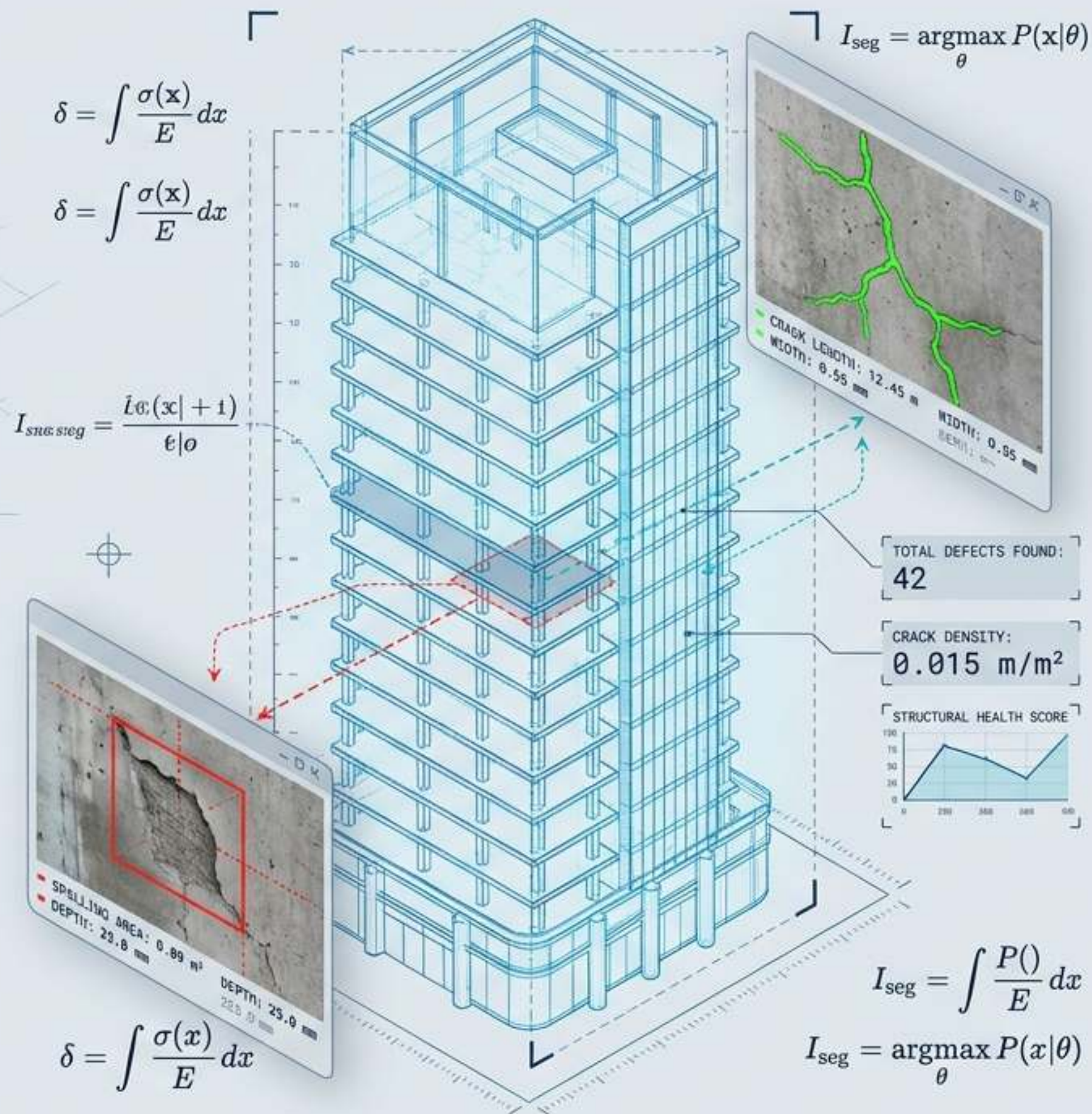


FROM PIXELS TO PHYSICS

ピクセルから物理量へ：次世代インフラ点検のパラダイムシフト

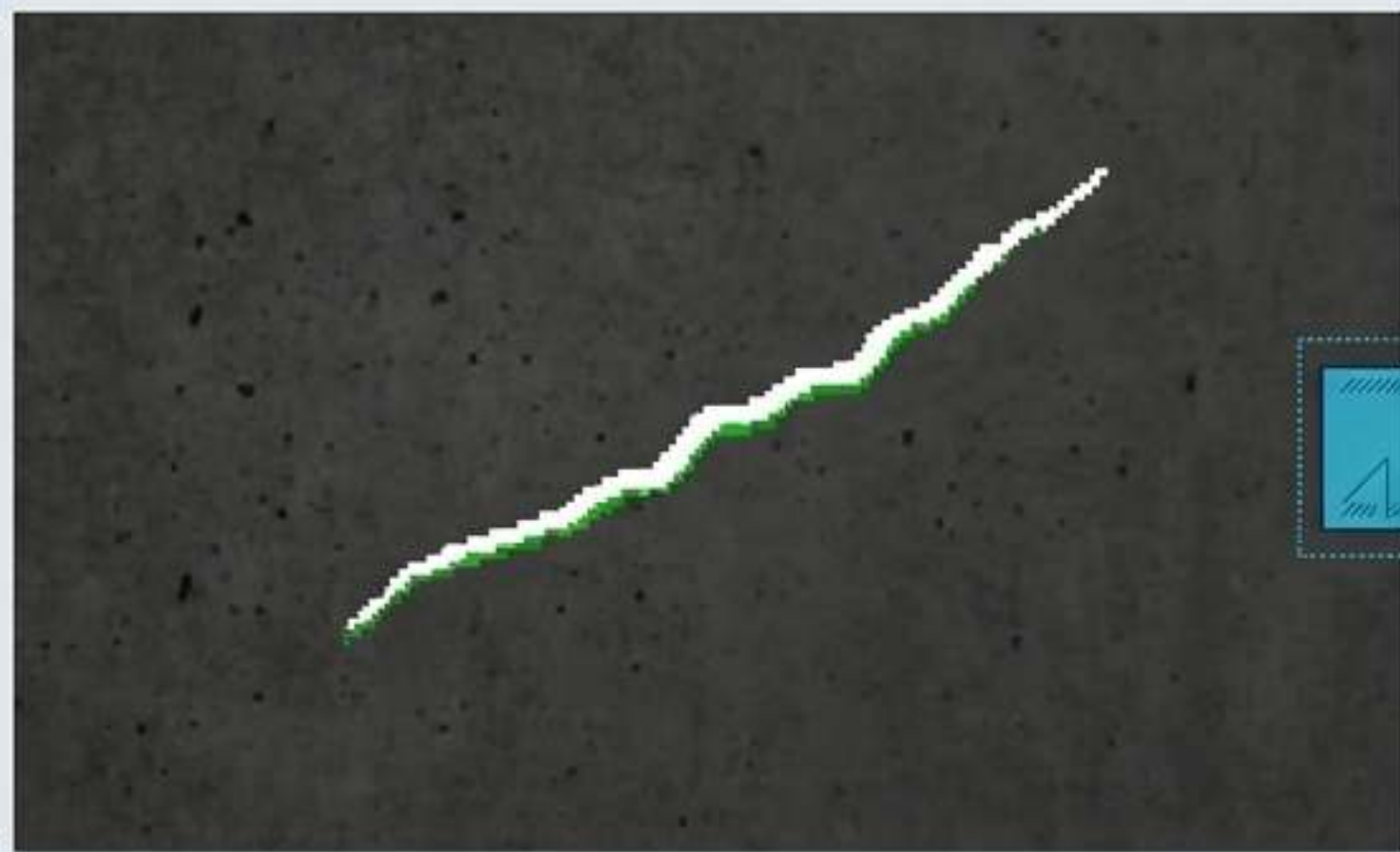
CUBIT-InSegデータセットが実現する、完全自動・物理的定量化・デジタルツイン統合型の外壁評価アプローチ



現在のAIは「色を塗る」だけで、「測定」ができない

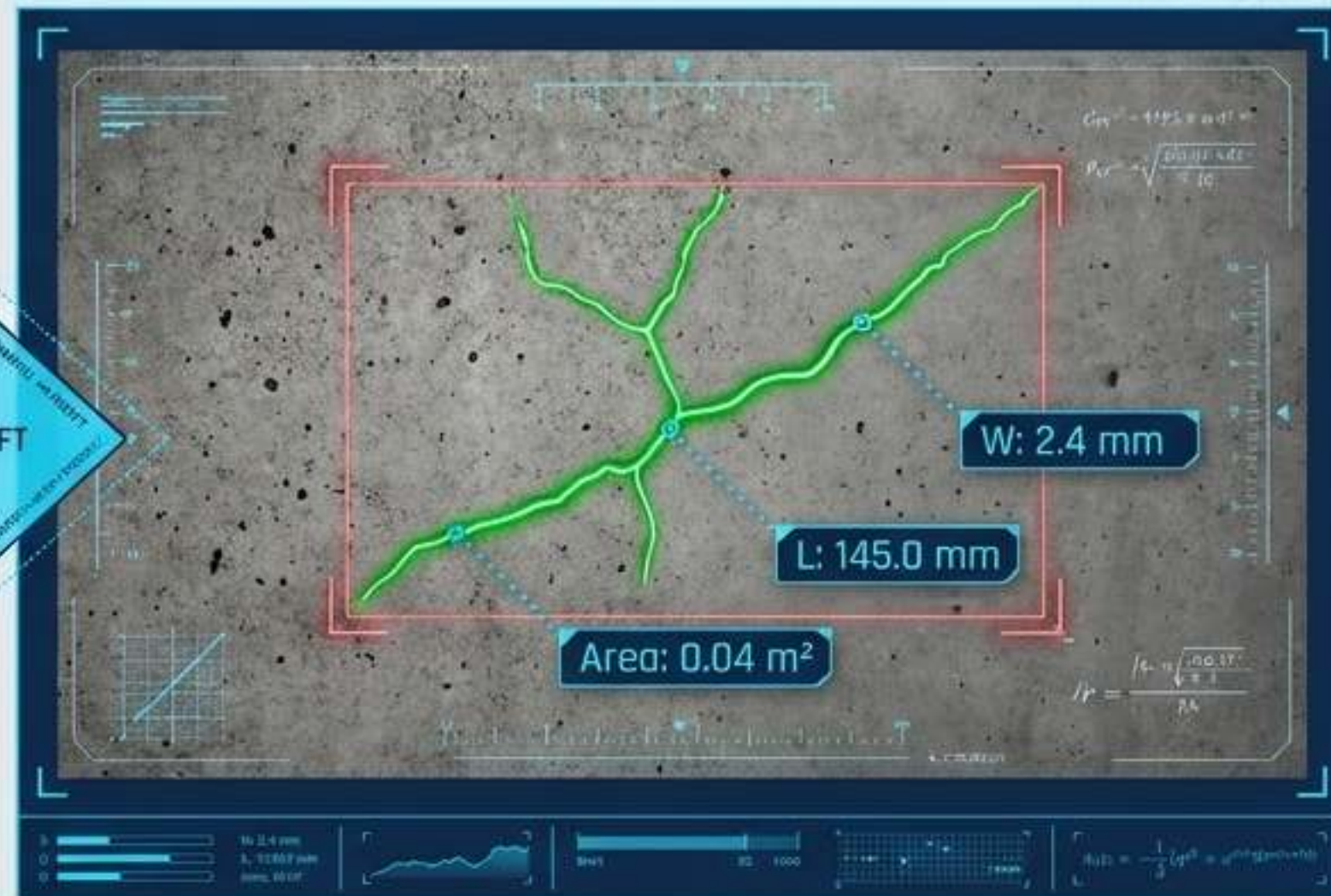
既存のインフラ点検AIは、画像のひび割れをピクセル単位で識別（Semantic Segmentation）できても、各欠陥を独立したインスタンスとして分離し、現実の寸法を計算することができません。実用的な保守計画に必要なのは、ピクセルの数ではなく、標準規格に準拠した「物理的な重症度」の定量化です。

AIの視点：単なるピクセルの集合（寸法不明）



PARADIGM SHIFT

エンジニアの視点：物理的スケールとインスタンス構造



既存のデータセットが抱える構造的なギャップ

データセット	Modality	Resolution	Annotation Type	Physical Scale
UAV75	UAV	低 (512×512)	Semantic Mask	非対応
GAPs-10m	地上車両	高 (5030×11505)	Semantic Mask	非対応
Crack-Seg	地上車両/混在	低 (416×416)	Semantic Mask	非対応
CUBIT-InSeg (Ours)	UAVベース	超高解像度 (4800×3200)	完全な Instance Level	対応 (等距離撮影)

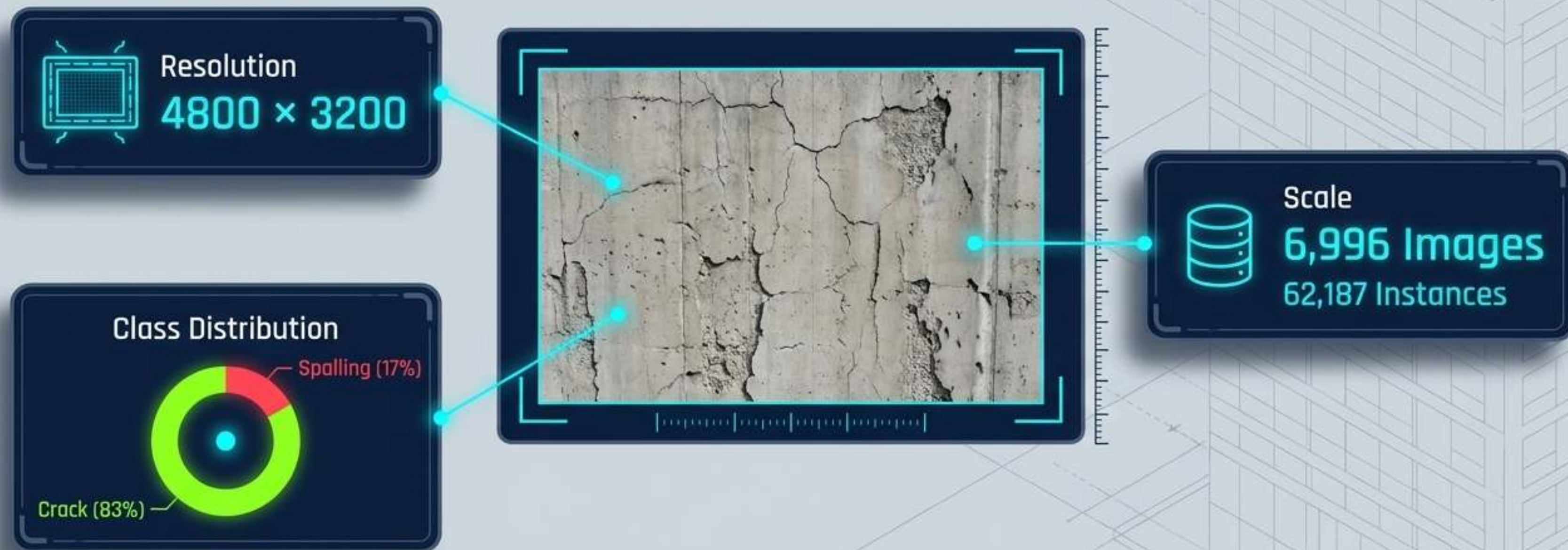
解像度不足による
微小クラックの欠落

地上視点のため高
層ビル点検に不適格

唯一、物理的な重症度
評価に直結する設計

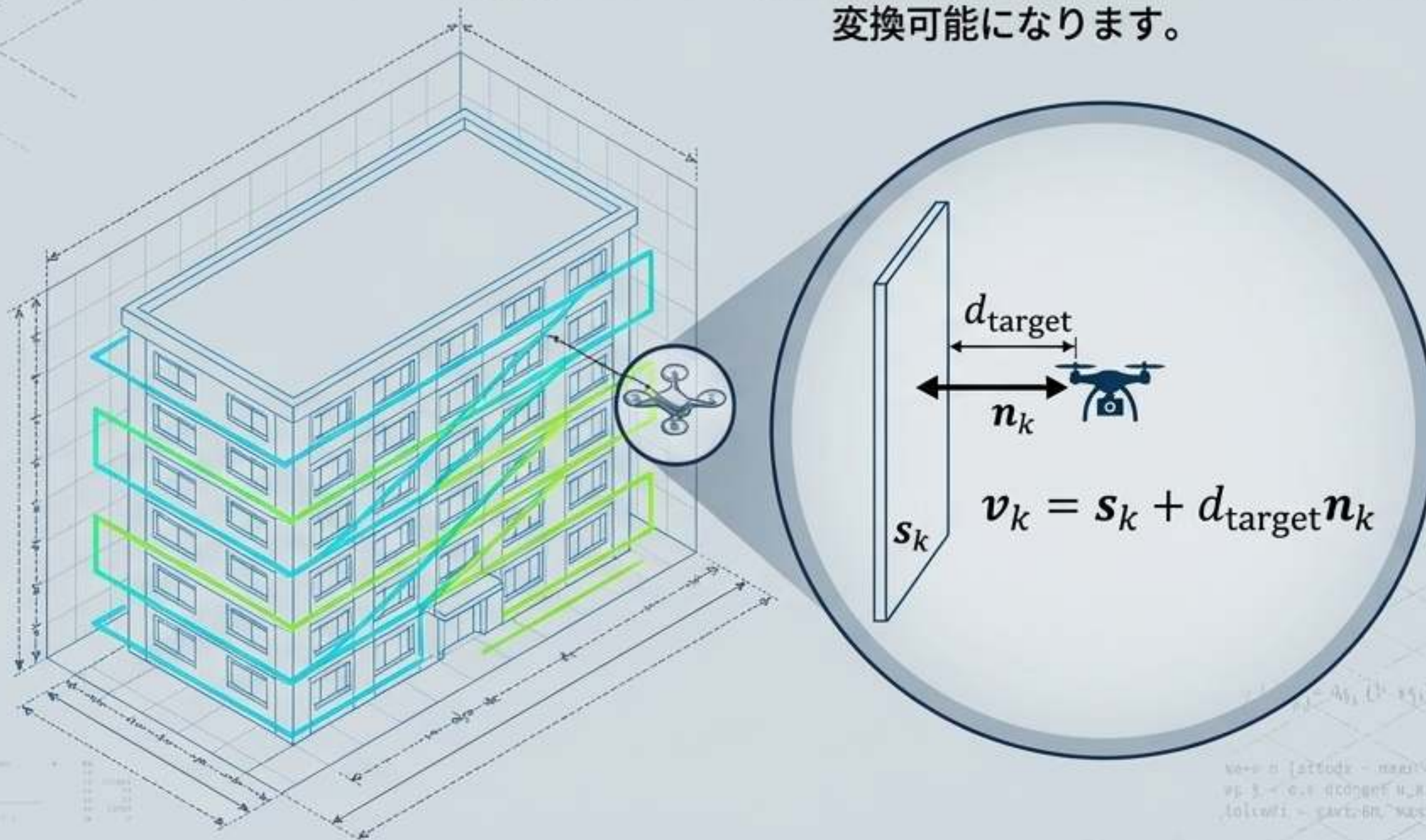
CUBIT-InSeg : コンピュータビジョンと土木工学の架け橋

実際の高層ビル点検環境を再現した、世界最大級のUAVベース外壁欠陥インスタンスセグメンテーション・データセット。ひび割れ (Crack) と剥離 (Spalling) という安全上極めて重要な2つの欠陥を、極限の解像度で捉えます。



等距離撮影モデル：GSDを固定し、スケールを保証する

ランダムな手動飛行ではなく、3D再構築に基づき壁面から一定距離（ d_{target} ）を保つ自律飛行戦略を採用。これにより、すべての画像でGSDが一定となり、AIが弾き出したピクセル数が直接ミリメートル（mm）へと変換可能になります。



The Key to Physical Measurement



GSD (Ground Sample Distance: 地上画素寸法)

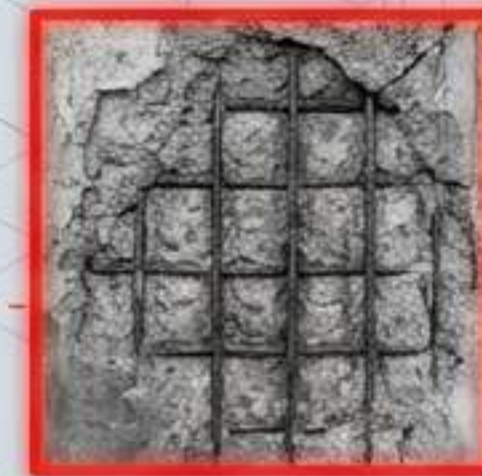
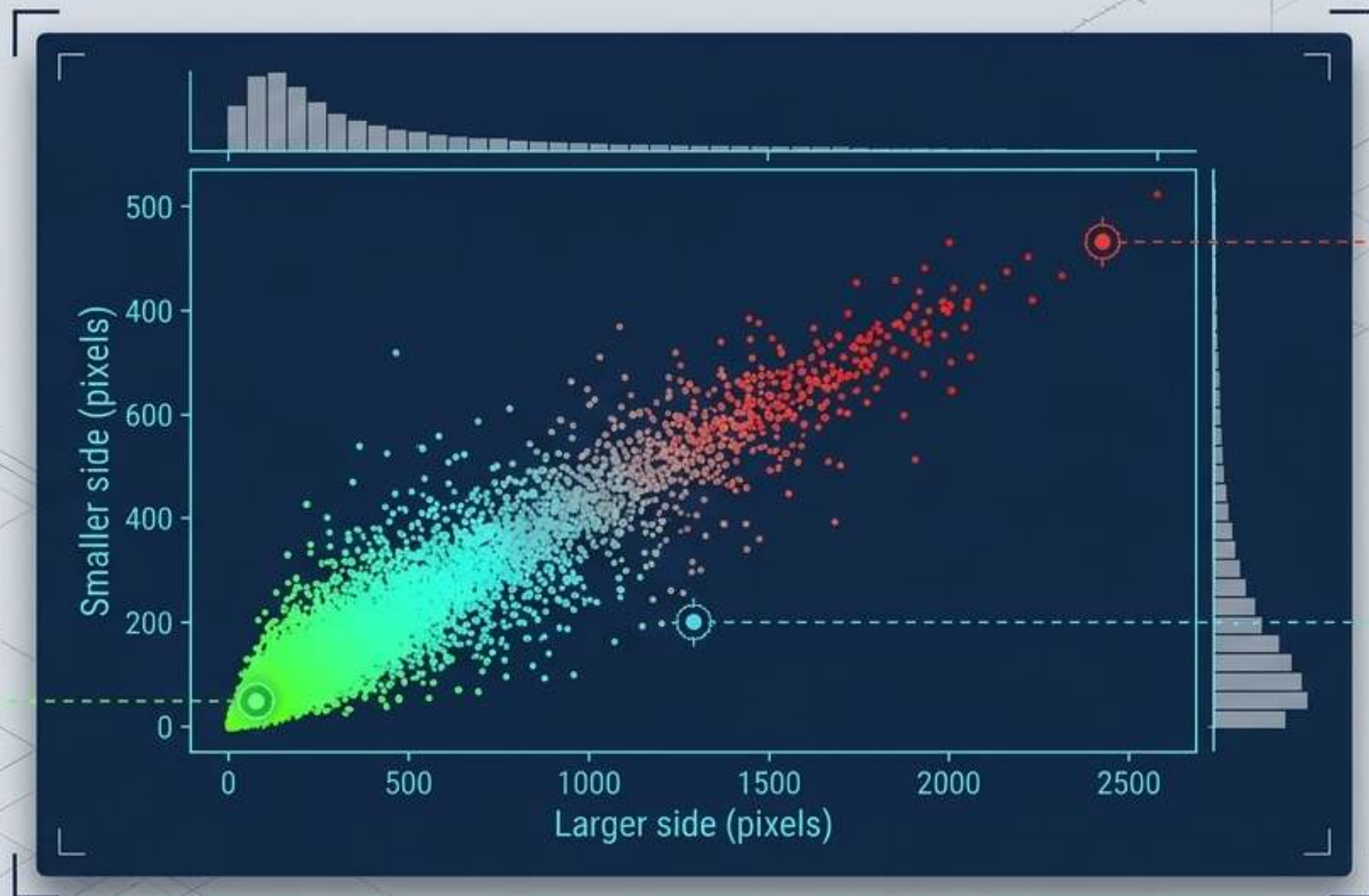
d_{target} の固定により、画像上の1ピクセルが現実世界の絶対寸法（mm）にロックされる。

ミクロの亀裂から、巨大な剥離まで

ターゲットのサイズ分布は極端なロングテールを描きます。数ピクセル幅の細いひび割れから、画像全体を覆うような巨大な剥離まで、CUBIT-InSegは現実のインフラ環境に存在する強烈なスケール変化とノイズ（影、汚れ、低コントラスト）を網羅しています。



Micro-Crack
(< 5 pixels width)



Massive Spalling
(> 4000 pixels)



Medium Spalling
+ Noise

80+ SOTAモデル検証：End-to-End vs. Single-Stage

17のモデルファミリー、80以上の最新アーキテクチャを単一のプロトコルで学習・評価。複雑なコンテキストを捉えるTransformerか、展開可能性に優れるYOLO派生型か。インフラ点検に最適なパラダイムを決定づけます。

End-to-End Models (CNN & Transformer)



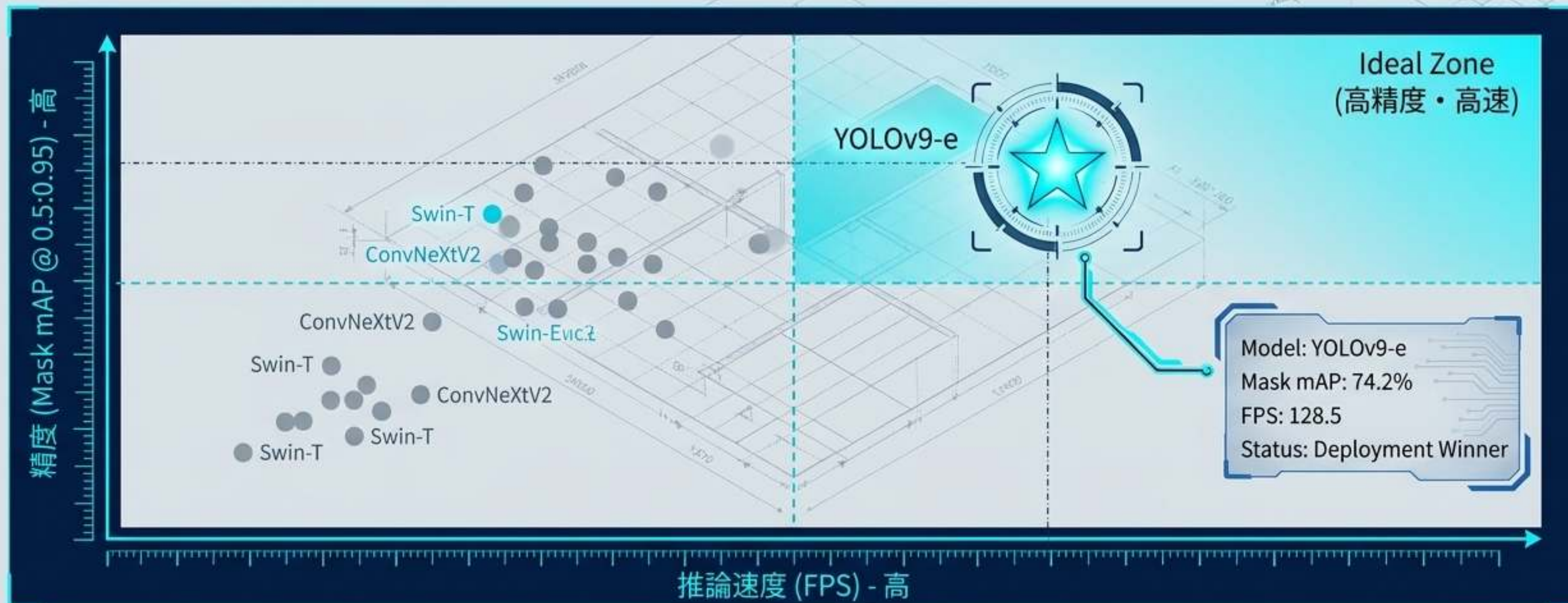
VS

Single-Stage Real-Time Models (YOLO variants)



展開可能性のランドスケープ：YOLOv9-eの圧倒的優位

複雑なEnd-to-Endモデル（ConvNeXtV2-Largeなど）と比較して、YOLOv9-eは計算量（FLOPs）を抑えつつ、最も厳しいIoU=0.5:0.95の評価基準においてBox mAP・Mask mAPともに最高精度（74.2%）を記録しました。実運用における最適解です。



ピクセルをミリメートルに変換する数学的アプローチ

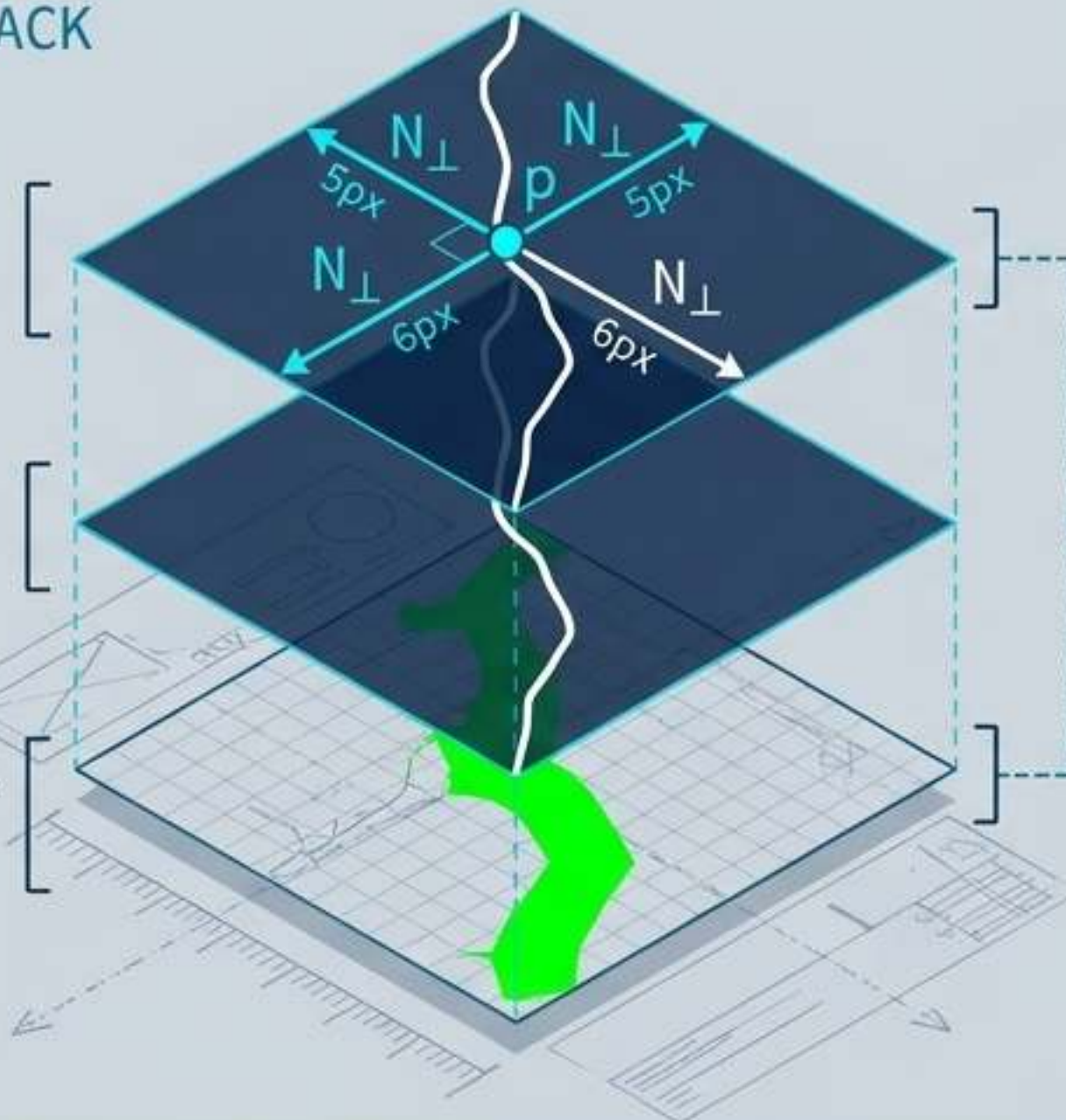
セグメンテーションマスクの骨格 (Skeleton) を抽出し、法線ベクトルを計算。エッジまでの連続ピクセル数に、撮影時にロックしたGSD (地上画素寸法) を乗算することで、2Dピクセルから現実世界の正確な「ひび割れ幅」と「長さ」を自動算出します。

ANATOMY OF A CRACK QUANTIFICATION

3. 法線ベクトルの
ピクセルカウント

2. 骨格線 (Skeleton)
の抽出

1. 生のAIセグメンテーション
マスク



$$W = N_{\perp} \times \text{GSD}$$

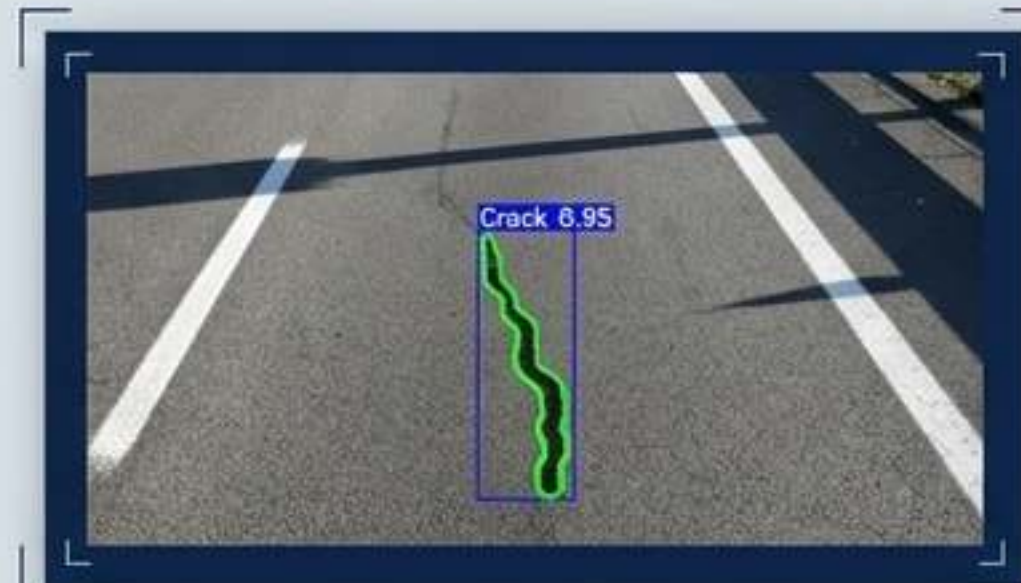
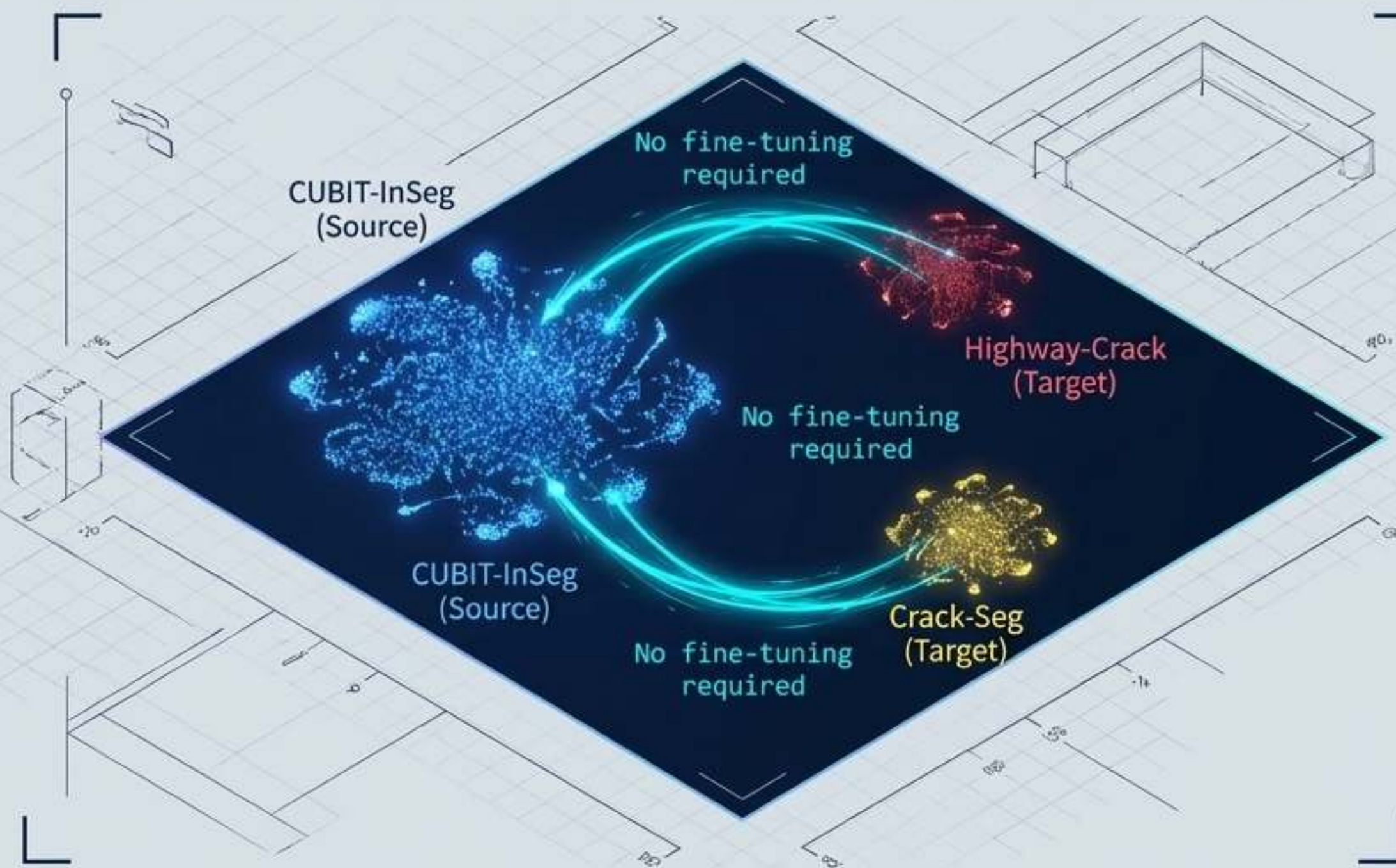
-> Crack Width (mm)

$$L = \sum \|p_{j+1} - p_j\| \times \text{GSD}$$

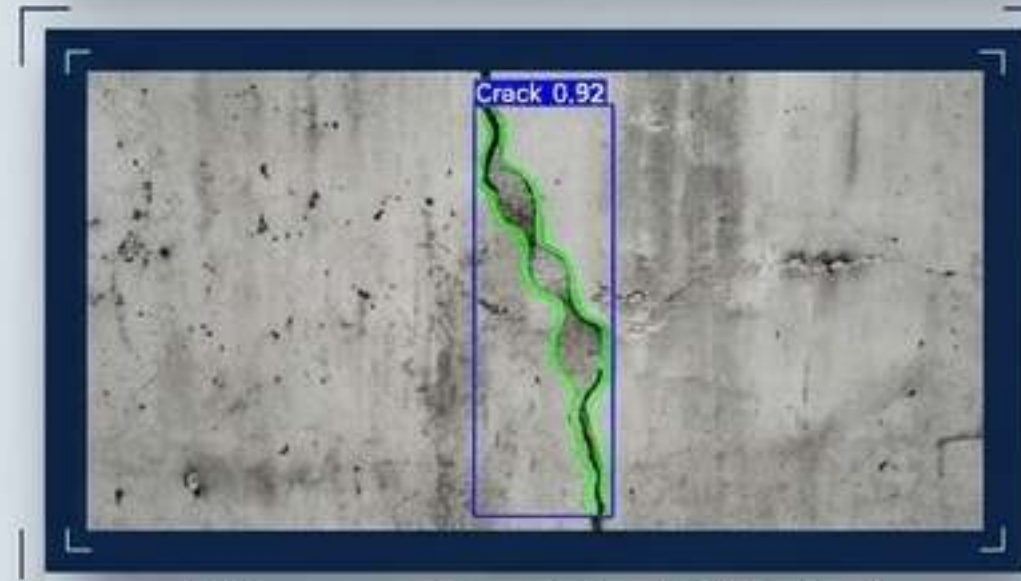
-> Crack Length (mm)

未知の領域への適応：ゼロショット転移の証明

CUBIT-InSegで訓練されたモデルは、再学習（ファインチューニング）なしで、視点やテクスチャの異なるハイウェイ路面（Highway-Crack）や異種コンクリート（Crack-Seg）のデータセットにも高い精度で適応します。これはモデルが本質的な欠陥の特徴を学習している証拠です。



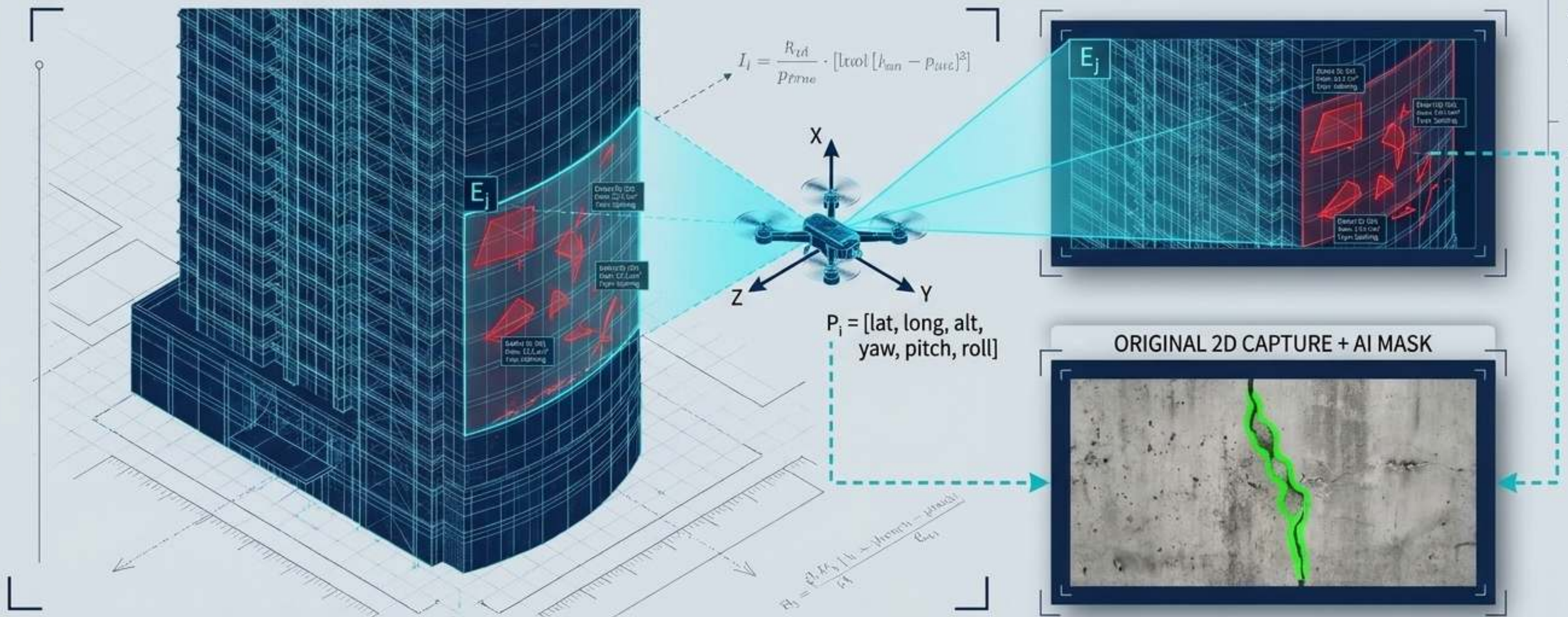
Inference on Highway-Crack (YOLOv9-e)



Inference on Crack-Seg (YOLOv9-e)

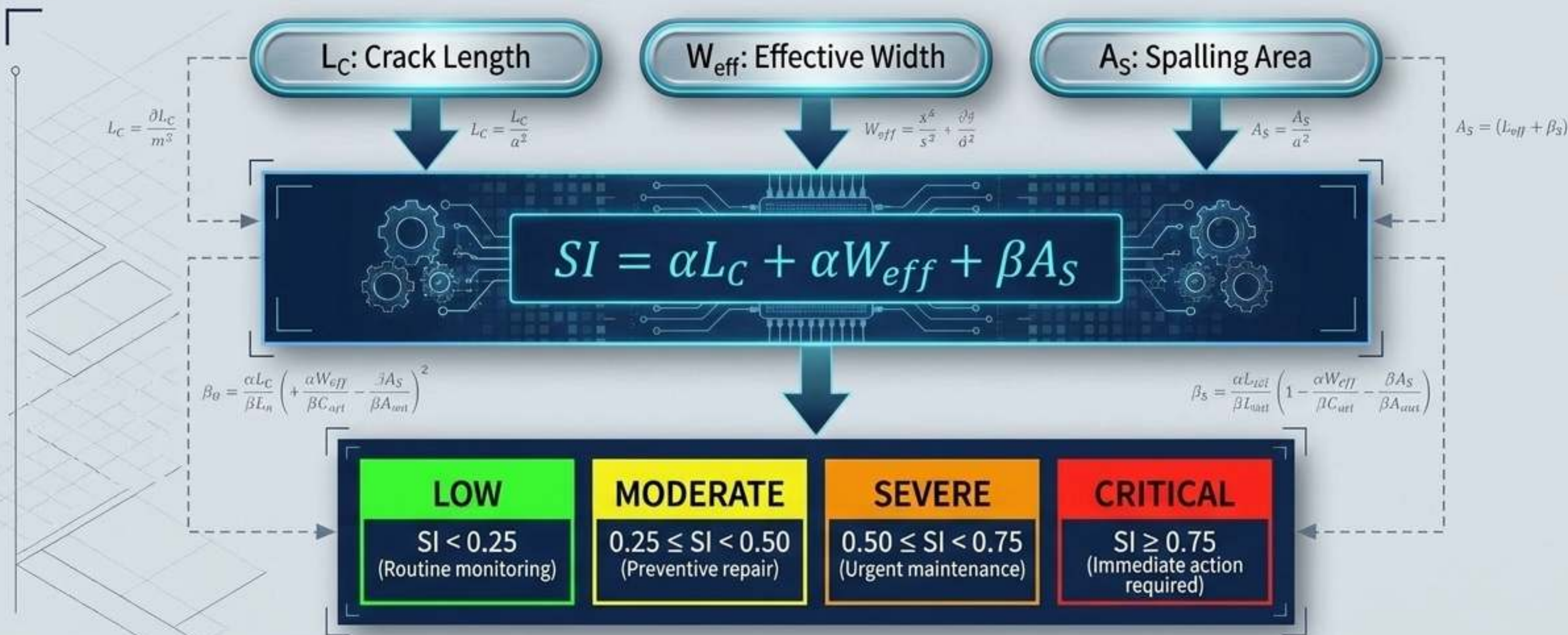
2D画像から3Dデジタルツインへの空間レジストレーション

個々の画像内で検出・定量化された欠陥は、ドローンの位置情報 (Pose) とカメラFOVを用いて、3D再構築されたデジタルツイン (DT) 上の対応する外壁コンポーネントへ正確にマッピングされます。これにより、画像ベースの管理からコンポーネントベースの管理へと移行します。



ISO準拠の客観的リスク評価（Severity Index）

物理的に定量化された寸法データは、BS ISO 15686-7:2017等の国際基準に準拠したSeverity Index (SI) の数式に投入されます。人間の主観を排除し、ひび割れの広がりや剥離面積に基づき、4段階の明確な修繕優先度を自動で判定します。



フルオートメーション：収集から意思決定までを単一のチェーンに

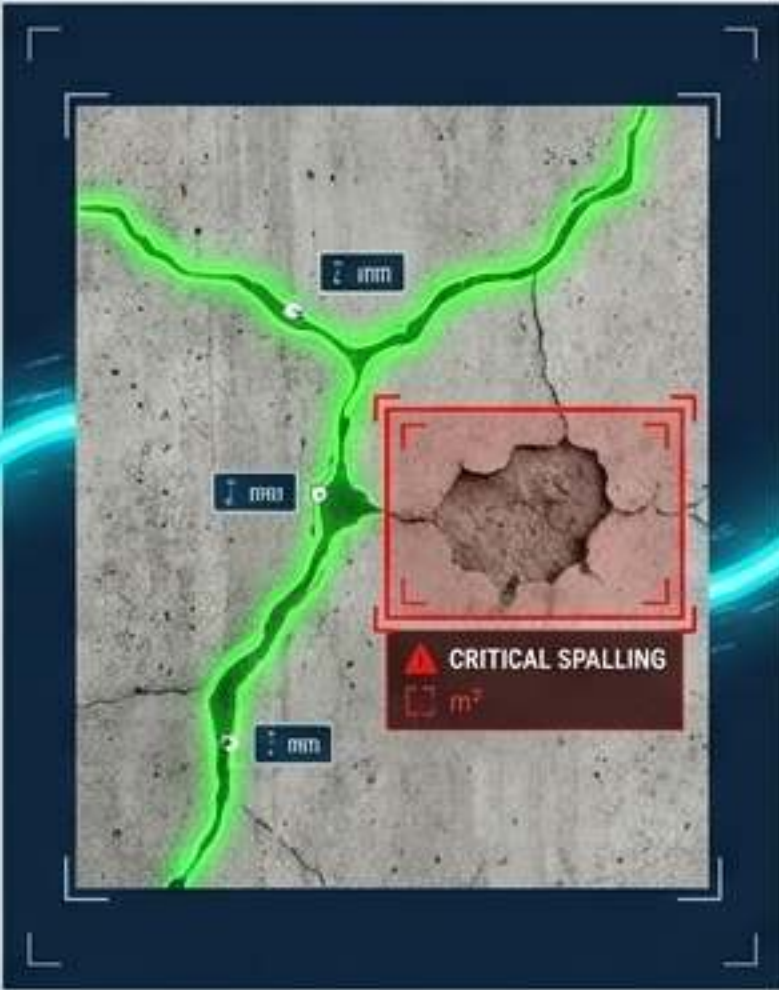
CUBIT-InSegが可能にしたのは、単なるAIの精度向上ではありません。幾何学的な飛行計画、インスタンス精度のAI、数学的な物理量変換を統合することで、「画像を撮る」だけで「ISO準拠の3D保守ダッシュボード」が完成する、完全に閉じたワークフローの実現です。

1. Data Collection



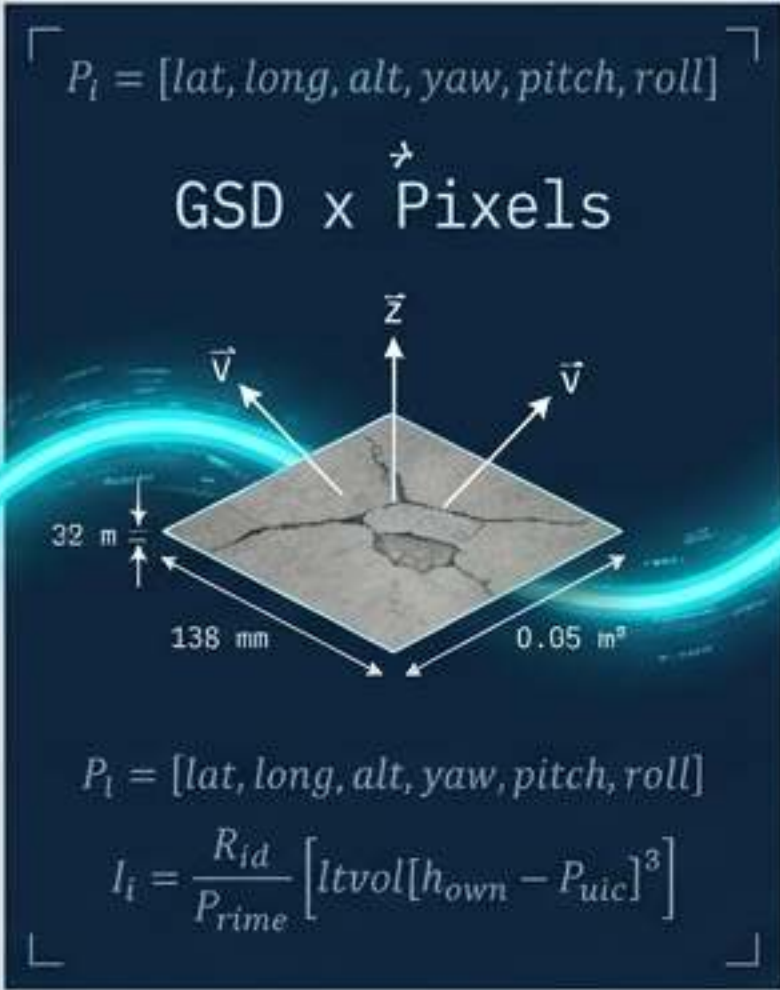
Equal-Distance Flight

2. Instance Segmentation



YOLOv9-e Pixel Masking

3. Physical Translation



Pixel to Metric (mm/m²)

4. Digital Twin Dashboard



ISO-Compliant
Actionable Insights

「異常を発見する」時代から、 「物理的リスクを定量化する」時代へ

属人的で危険な目視点検や、物理スケールを持たない2D AIの限界は過去のものとなりました。CUBIT-InSegを基盤とするこのワークフローは、ピクセルを物理量に変換し、インフラ維持管理におけるスケーラブルで定量的な意思決定の新たな基準を確立します。

AI-Driven Infrastructure Inspection
CUBIT-InSeg
Pixel-to-Physical Conversion
Scalable Quantitative Decision Making

Critical - Action Required

Engineering value delivered.

エンジニアリング価値の具現化