RESEARCH ARTICLE · SCIENTIFIC REPORTS

Overcoming Resolution Constraints in Automated Colony Counting via SAHI

A high-performance tiled-inference deep learning framework that preserves native image resolution for accurate, lightweight bacterial colony detection

Sercan Külcü & Duygu Balpetek Külcü

Giresun University — Computer Engineering & Food Engineering

Scientific Reports (2026) 16:24516 | DOI: 10.1038/s41598-026-55724-1

OVERVIEW

The Study at a Glance

Downscaling high-resolution Petri dish images to 640×640 pixels severely hurts lightweight detectors. This study shows that SAHI-based tiled inference — not bigger models — is the key to accurate, real-time bacterial colony counting.

96.9%	2.6M	<320 ms	23/24
mAP@0.5 with tiled YOLOv11n	Parameters — nano-scale model	Whole-plate inference (RTX 3050)	Classes with ≥95% accuracy

KEYWORDS

- Bacterial colony counting
- Tiled inference
- Small object detection
- High-resolution imaging
- SAHI
- Lightweight YOLO models

Where Resolution Gets Lost

Clinical microbiology labs face growing sample volumes and need rapid, accurate colony detection. But standard detectors resize full Petri dish photographs down to a fixed input size — and that resizing quietly destroys the fine detail small, densely packed colonies depend on.

Downsampling destroys detail

Resizing a high-resolution plate to 640×640 pixels removes the fine spatial cues that separate tiny, tightly clustered colonies.

Dense, real-world plates

Colony density ranges from under 10 to over 1,100 per plate — with some species growing in especially crowded patterns.

Laboratory hardware limits

Routine diagnostic labs need lightweight, nano-scale models that still run fast on consumer-grade GPUs.

Three Contributions

This study shows that resolution-preserving tiling — not model scaling — is the decisive factor for dense micro-colony detection.

1

Tiling beats resizing and architecture upgrades

Simple fixed-size tiled inference with moderate overlap clearly outperforms conventional resizing — and outperforms recent YOLO architectural advances.

2

High accuracy from lightweight nano-scale models

State-of-the-art detection accuracy is achievable with YOLOv11n at only 2.6M parameters, on consumer-grade hardware, with subsecond whole-plate inference.

3

A reproducible, lightweight pipeline

Avoids heavy backbones, multi-scale training, test-time augmentation, and ensembling — lowering barriers to adoption in clinical and veterinary microbiology.

A Real-World Bacterial Colony Dataset

DATASET SUMMARY

369 Petri dish images captured with three different smartphones under real-world laboratory conditions — natural variation in lighting, reflections, background, and condensation.

56,865 Individually annotated colonies

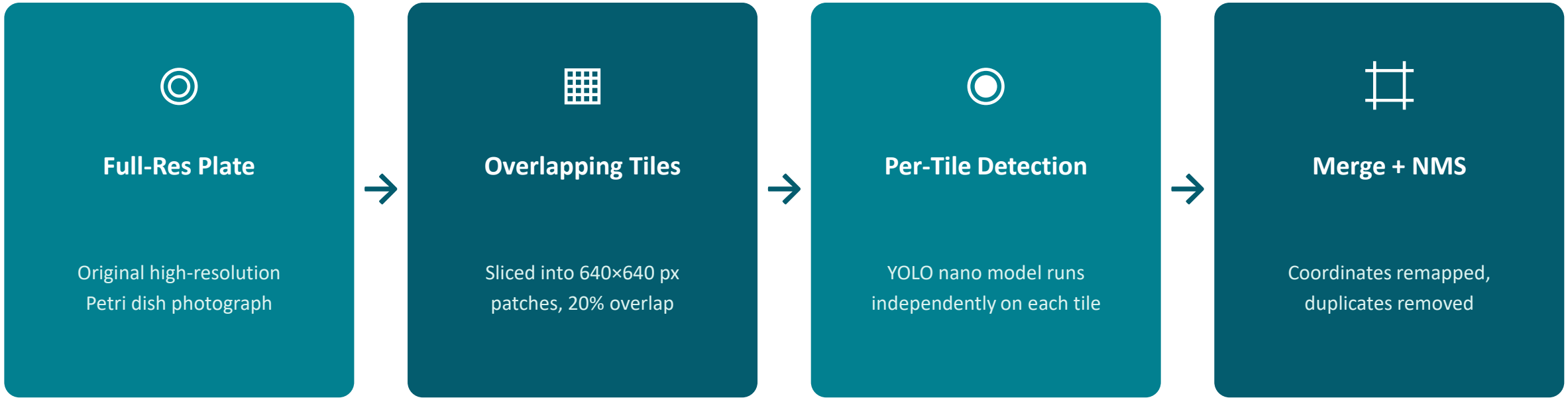
24 Bacterial species (11 Gram+ / 13 Gram−)

≤1,100+ Colonies on the densest single plate

SPLIT & PREPROCESSING STRATEGY

Image-level split	**80% dev / 20% independent test**
Dev subsplit	**80:20 → ~64% train / 16% val**
Tile size	**640 × 640 px, 20% overlap**
Tiled dataset	**9,220 labeled tiles (~25× more)**
Media used	**Blood, chocolate, nutrient, MacConkey agar**

Slicing the Plate to Preserve Detail



SAHI (Slicing Aided Hyper Inference) is model-agnostic — it plugs into YOLOv5n, YOLOv8n, or YOLOv11n directly, with no architectural modification required.

Three Nano Models, Two Strategies

YOLOv5n, YOLOv8n, and YOLOv11n — three successive nano-scale generations of the YOLO family — were each trained under both a resized baseline and the proposed tiled strategy, for a controlled, model-agnostic comparison.

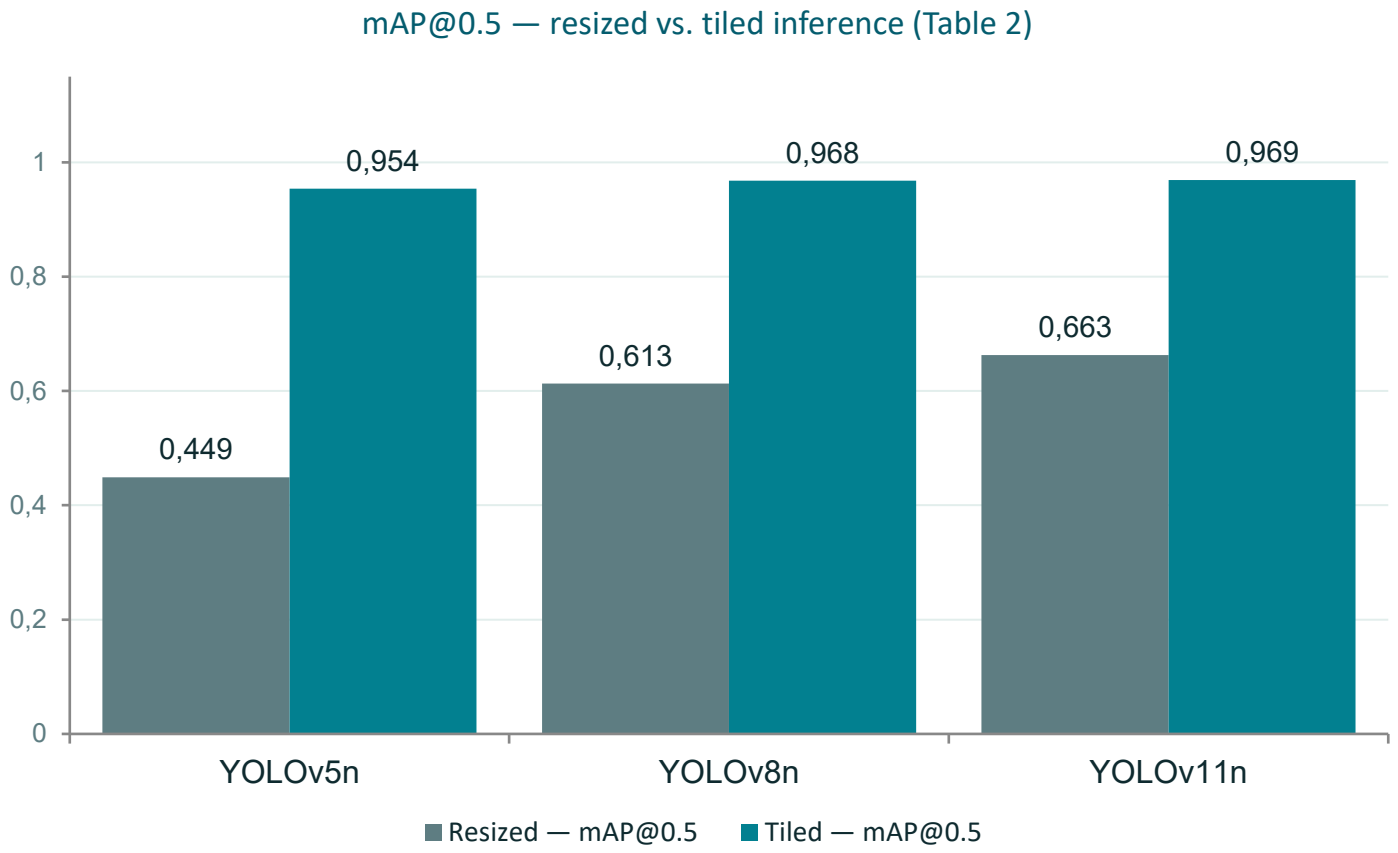
TRAINING CONFIGURATION

Framework	**Ultralytics v8.3.213**
Epochs	**50**
Input resolution	**640 × 640 px**
Augmentation	**Mosaic, HSV, flips, scale**
Optimizer	**Auto (SGD / AdamW)**
GPU	**NVIDIA RTX 3050 Laptop (6GB)**

WHY NANO MODELS?

- Match the lightweight objective of routine laboratory deployment
- Feasible on consumer-grade hardware without specialised imaging setups
- Test whether resolution preservation — not model scaling — drives accuracy
- Larger YOLO variants were deliberately excluded from evaluation

Tiling Transforms Detection Accuracy



THE HEADLINE RESULT

44.9–66.3% → 95.4–96.9%

mAP@0.5 range across all three models, resized vs. tiled

+100%+

improvement in mAP@0.5:0.95 (stricter localisation metric)

The gain comes from preserving native resolution during inference — not from a bigger or newer model.

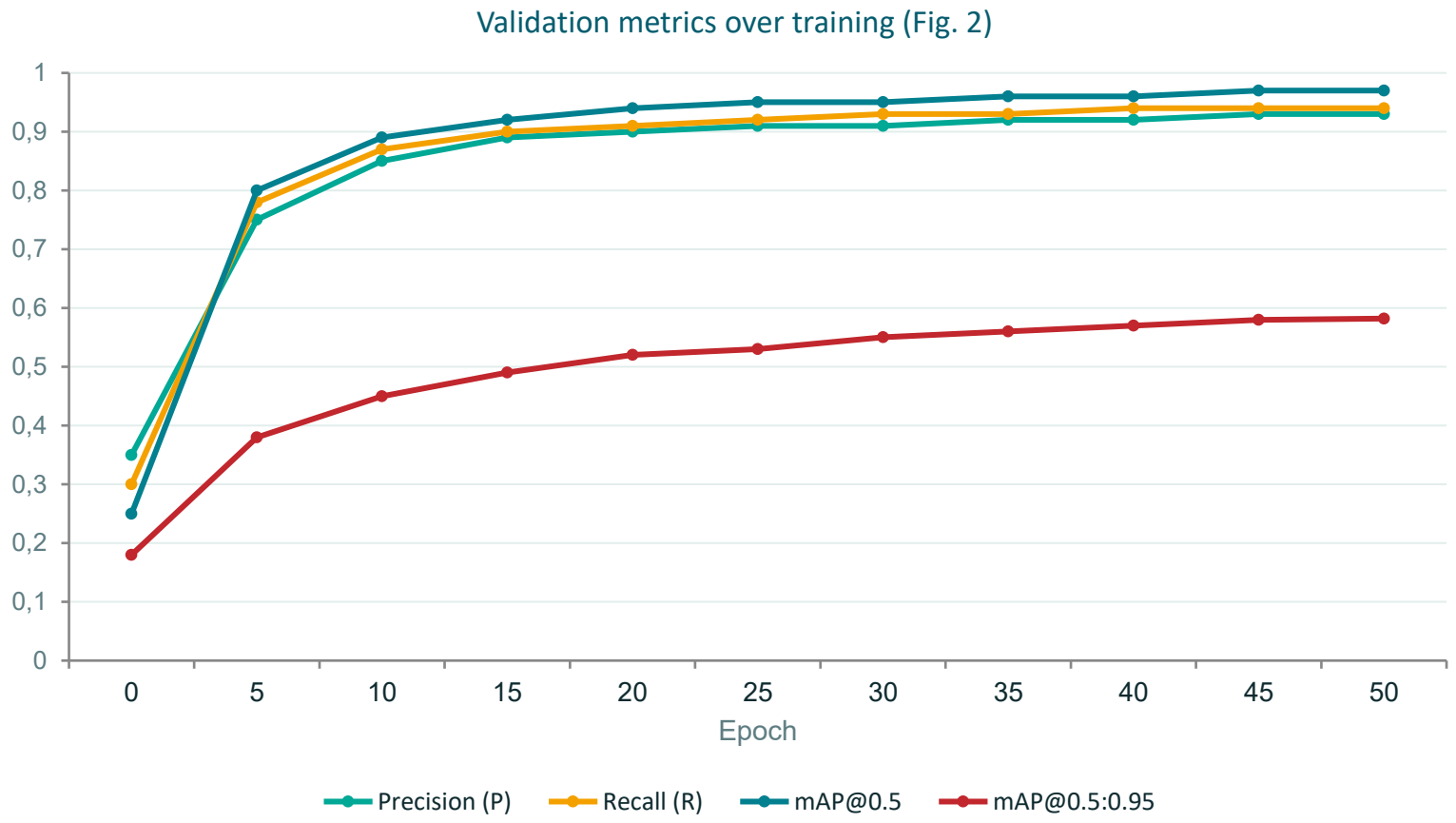
A Worthwhile Trade-Off

Tiling multiplies training time roughly fourfold — expected, since the tiled dataset has ~25× more samples — but whole-plate inference stays firmly practical for routine lab use.

Model	Train (resized)	Train (tiled)	Infer (resized)	Infer (tiled)
YOLOv5n	~1h 15min	~4h 50min	42 ms	285 ms
YOLOv8n	~1h 20min	~5h 10min	58 ms	298 ms
YOLOv11n	~1h 10min	~4h 45min	62 ms	312 ms

Whole-plate inference stays under 320 ms on a laptop RTX 3050 GPU — no TensorRT or quantisation required.

YOLOv11n: Steady Gains Over 50 Epochs



READING THE CURVES

0.21 → 0.582

mAP@0.5:0.95 rises steadily with no plateau — more epochs or mild augmentation could push it further

Precision ≈ Recall

by epoch 50 (≈0.93/0.94) — a well-balanced detector with few false positives or missed colonies

Consistent Separation Across 24 Species

23 / 24

Bacterial classes reaching
≥ 0.95 diagonal accuracy

< 0.01

Typical inter-class
confusion rate

0.16

Highest background
false-positive rate
(Streptococcus agalactiae)

The normalised confusion matrix (Fig. 3) shows the proposed tiled YOLOv11n model reliably distinguishes morphologically similar veterinary pathogens under varying imaging conditions. Diagonal values exceed 0.95 in 23 of 24 classes; nearly all inter-class confusions remain below 0.01. The only notable off-diagonal activity is between a few classes and the background — not between bacterial species themselves, which is the harder and more clinically important distinction.

Missed Colonies, Recovered

On a high-density blood agar plate with 26 closely packed colonies (Fig. 4), resize-based inference lost fine spatial detail — tiled inference recovered it completely.

RESIZED (640×640)

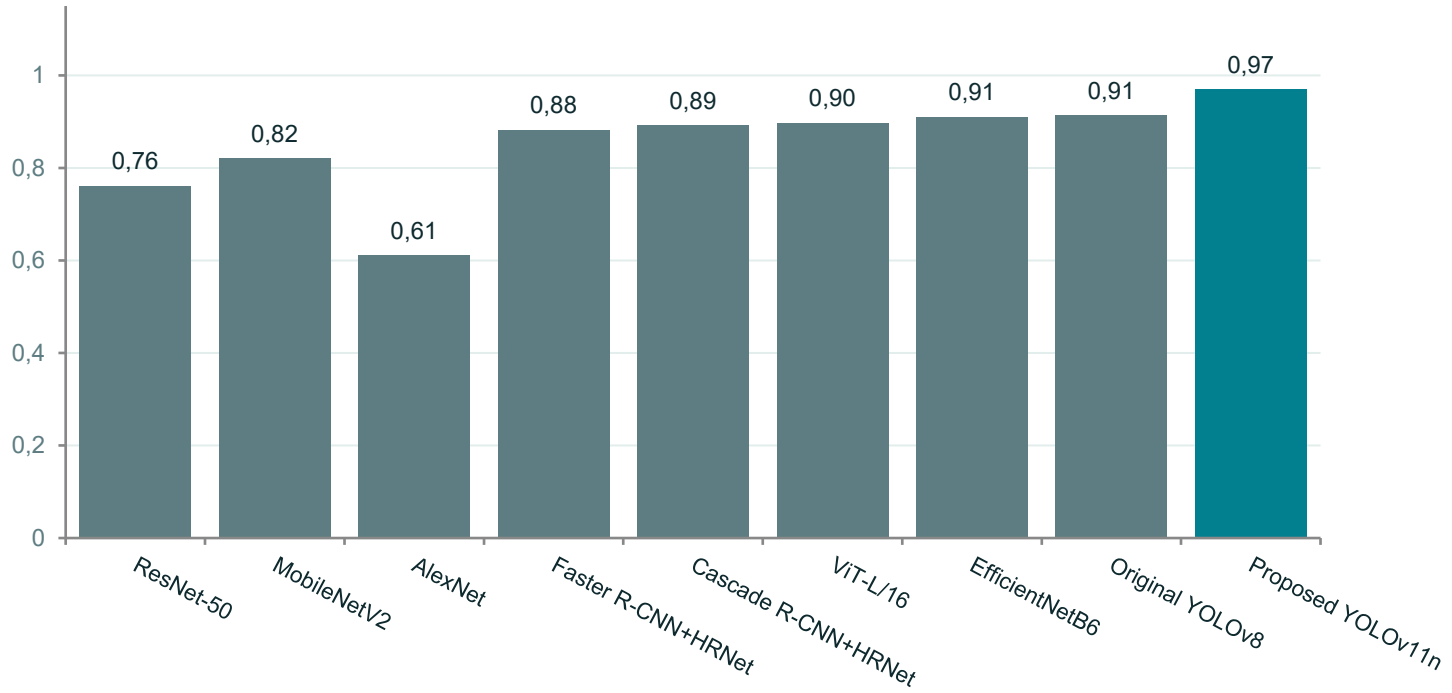
- YOLOv8n-resized: misses 1 colony, 5 misclassifications
- YOLOv11n-resized: misses 2 colonies, 3 misclassifications
- Fine spatial detail lost during downsampling

TILED (20% overlap + NMS)

- YOLOv8n-tiled: zero missed detections
- YOLOv11n-tiled: zero missed detections
- Only 1 duplicate false positive per model at tile borders

Small Model, Best-in-Class Accuracy

mAP@0.5 vs. published detectors (Table 4)



Proposed YOLOv11n reaches 0.969 mAP@0.5 with only 2.6M parameters — versus 27–304M in earlier heavy two-stage or transformer methods.

OVERLAP-RATIO SENSITIVITY

10% → 25% overlap changes mAP@0.5:0.95 by only 0.002 — 20% was chosen as the balanced setting.

5-FOLD CROSS-VALIDATION

0.967 ± 0.004

tiled mAP@0.5, vs. 0.663 ± 0.009 resized — stable across every fold

4. CONCLUSION

Resolution Preservation Over Model Scaling

A simple tiled-inference pipeline with 20% overlap eliminates the resolution loss that cripples lightweight YOLO models on high-density Petri dish images — reaching 96.9% mAP@0.5 and 58.2% mAP@0.5:0.95 with YOLOv11n's 2.6M parameters, and under 320 ms whole-plate inference on a laptop GPU.

Resolution-preserving | **Lightweight** | **Model-agnostic** | **Reproducible** | **Real-time**

Sercan Külcü & Duygu Balpetek Külcü · Giresun University

sercan.kulcu@giresun.edu.tr

Thank You

Questions & Discussion

Sercan Külcü & Duygu Balpetek Külcü · Giresun University

sercan.kulcu@giresun.edu.tr