RESEARCH ARTICLE · BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

Bias Mitigation in Ensemble-Based Meat Freshness Classification Using Grad-CAM

A Grad-CAM-guided attention framework that redirects lightweight ensemble models from packaging cues to genuine meat freshness indicators

Sercan Külcü & Duygu Balpetek Külcü

Giresun University — Computer Engineering & Food Engineering

Vol. 14 (2026) 74-82 | DOI: 10.17694/bajece.1817907

OVERVIEW

The Study at a Glance

Baseline CNNs can reach near-perfect accuracy on meat freshness classification while quietly relying on packaging trays instead of meat texture and colour. This study integrates a Grad-CAM-guided attention module into a lightweight hybrid ensemble to fix that — without sacrificing speed.

99.78%

Held-out test accuracy

99.66%

5-fold CV accuracy (± 0.23%)

4.3M

Model parameters (16.5 MB)

825 FPS

Single-GPU inference speed

INDEX TERMS

- Bias mitigation
- Computer vision
- Ensemble learning
- Food safety
- Grad-CAM
- Meat freshness

The Reliability Paradox

Spoiled meat threatens health and drives major economic loss, so fast automated inspection matters. But a model that scores 98%+ accuracy is not automatically trustworthy — it may be right for the wrong reasons.

Background bias

Grad-CAM analysis reveals baseline CNNs often focus on packaging trays instead of meat colour, texture, and moisture.

Safety-critical stakes

Misclassifying spoiled meat as fresh risks real health consequences and regulatory violations.

Edge-device constraints

Real-time, on-site inspection in low-resource supply chains needs small, fast, GPU-light models.

Four Contributions

This study explicitly identifies and corrects background-induced visual bias, rather than optimising accuracy alone.

1. **Grad-CAM-guided bias mitigation framework**

   Explicitly identifies and corrects spurious attention to background elements in meat freshness classification.

2. **Lightweight MiniCAM attention module**

   Integrated into MobileNetV2 to redirect model focus toward meat-specific visual cues — colour, texture, moisture.

3. **Hybrid ensemble fusion**

   Fuses deep embeddings from MobileNetV2+MiniCAM and Xception with SVM and XGBoost, combined via TTA and grid-optimised weighted ensembling.

4. **Comprehensive explainability analysis**

   Grad-CAM confirms the proposed model eliminates background bias and aligns attention with domain-relevant freshness indicators.

Meat Freshness Image Dataset

DATASET SUMMARY

2,266 high-resolution RGB images of beef samples, publicly hosted on Kaggle, collected from diverse market environments across varied lighting, orientation, and background.

Fresh **853 imgs**

Bright red, glossy, intact marbling

Semi-Fresh **789 imgs**

Duller tone, reduced sheen, early texture loss

Spoiled **624 imgs**

Discoloration, slime, visible microbial growth

PREPROCESSING & SPLIT STRATEGY

Original split	**1,815 train / 451 test**
Held-out val split	**10% stratified from train**
Resize	**416 × 416 px (stretch interp.)**
EXIF handling	**Auto-orientation, metadata stripped**
Provider augmentation	**None applied**

Baseline: YOLO11s-cls

A lightweight classification model derived from the YOLO11 architecture — a CSPNet-based backbone with C3k2 modules, an SPPF module, and a global-average-pooling classification head.

5.4M
Parameters

12.1
GFLOPs

98.89%
Top-1 accuracy

100%
Top-5 accuracy

BEST-EPOCH SNAPSHOT (EPOCH 19)

Train loss	0.0447 %
Validation loss	0.1178 %
Fresh F1	1.0000
Semi-Fresh F1	0.9840
Spoiled F1	0.9785
Weighted F1	0.9889

2.1 BASELINE MODEL — GRAD-CAM FINDING

High Accuracy, Wrong Reasons

Despite 98.89% Top-1 accuracy, Grad-CAM visualisation of the baseline model reveals a critical interpretability flaw.

Fig. 7 — Spoiled sample, 100.0% confidence

attention hotspot on the TRAY, not the meat

The model's decision relied predominantly on the packaging tray rather than meat-specific degradation cues — colour, texture, moisture.

WHY THIS MATTERS

- Confusion matrix still looks strong (5 spoiled → semi-fresh errors only)
- But reliance on non-meat cues fails to generalise across trays, backgrounds, and packaging styles
- A model must reason about color, texture, and moisture — not props — to be trustworthy in real deployments

2.2 PROPOSED MODEL

MiniCAM-Guided Hybrid Ensemble

Input Image

224×224 px
beef sample

→

Dual Feature Path

MobileNetV2+MiniCAM
+ frozen Xception

→

Fuse & Classify

Concatenated features →
SVM + XGBoost

→

Weighted Ensemble

TTA + grid-optimised
fusion weights

MiniCAM applies channel attention (global-avg-pool → MLP → sigmoid gating) and spatial attention (7×7 conv over avg/max-pooled maps) to redirect MobileNetV2 toward meat-specific regions (Eq. 1–2).

2.2 TRAINING CONFIGURATION

Fusion and Training Parameters

MODEL & OPTIMISATION

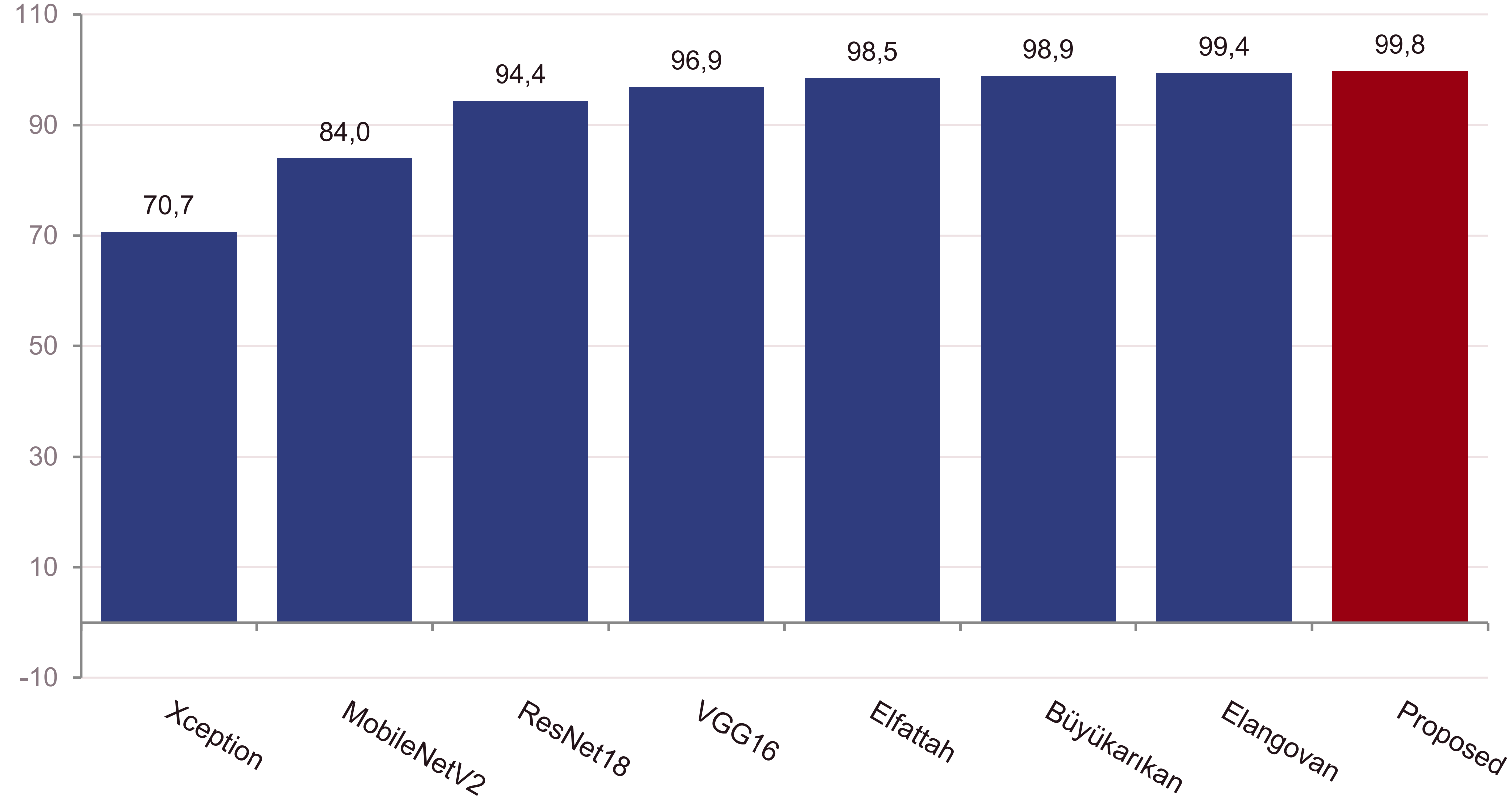
Backbone	**MobileNetV2 + MiniCAM**
Feature extractor	**Xception (frozen)**
Image resolution	**224 × 224**
Optimizer	**AdamW**
Learning rate	**3e-4**
Batch size	**16**
Early stopping	**10 epochs**
Loss function	**Cross-entropy**

ENSEMBLE FUSION

Test-time augmentation	**8 transforms (flip×4 rot.)**
SVM kernel	**Linear**
XGBoost params	**max_depth=7, n_est.=300**
Fusion strategy	**Grid-search weighted avg**
Search range w_1	**0.4 – 0.6**
Search range w_2, w_3	**0.15 – 0.3**
Optimal weights	**w_1=0.5, w_2=0.3, w_3=0.2**

Outperforming State-of-the-Art Models

Accuracy across representative baselines (Table V)



KEY TAKEAWAY

99.78%

accuracy, precision & recall — best across every SOTA baseline compared

+1–2 pts

margin over recent lightweight/edge models (Elangovan, Büyükarıkan)

This gain is practically significant: misclassifying spoiled meat risks health and regulatory consequences.

Near-Perfect, Class-by-Class

Confusion Matrix — Proposed Model (validation split)

	FRESH	SEMI-FRESH	SPOILED
FRESH	0.994	0.006	0.000
SEMI-FRESH	0.000	1.000	0.000
SPOILED	0.000	0.000	1.000

WHAT CHANGED VS. BASELINE

- Baseline: 5 spoiled samples misclassified as semi-fresh
- Proposed: only 1 fresh sample (0.6%) misclassified as semi-fresh
- Zero confusion remains between semi-fresh and spoiled — the safety-critical boundary
- Weighted precision, recall, and F1 all reach ≈0.998–1.000

From Packaging Trays to Meat Cues

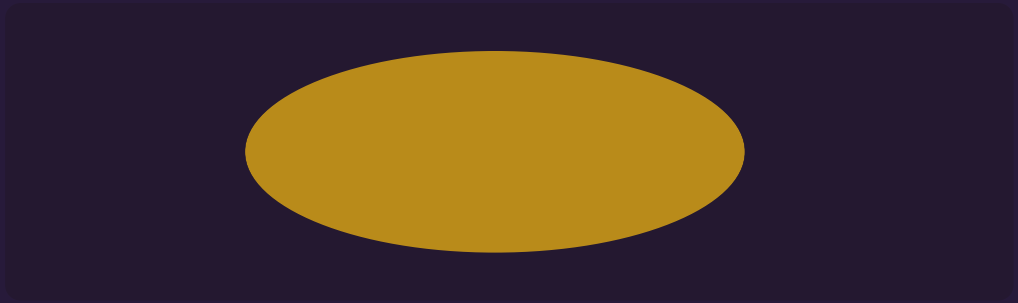
Grad-CAM visualisations across all three classes confirm that the MiniCAM-guided model attends to semantically valid regions.

BASELINE (YOLO11s-cls)

hotspot: tray corner (background)

Packaging tray focus

PROPOSED (MiniCAM ensemble)



hotspot: meat surface — colour, texture, moisture

Meat surface focus

Fast, Small, and Stable Across Folds

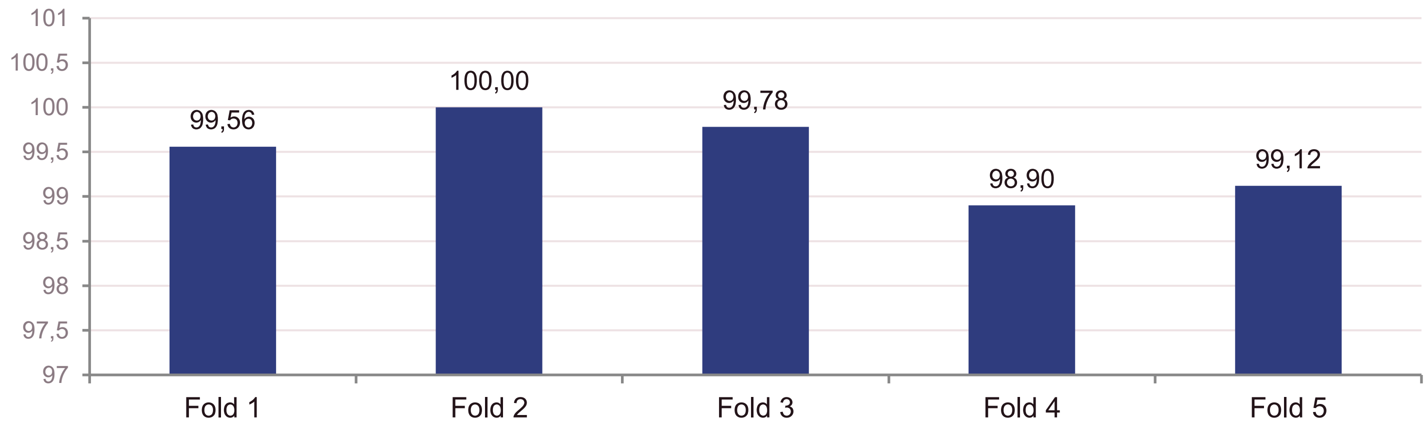
4.3M
Parameters

16.5 MB
Model size

825.1
FPS (single GPU)

99.66±0.23%
5-fold CV accuracy

5-Fold Cross-Validation Accuracy (%)



Optimal ensemble weights ($w_1 \approx 0.5$, $w_2 \approx 0.3$, $w_3 \approx 0.2$) stayed consistent in 4 of 5 folds — the fusion strategy is stable and not overly sensitive to data partitioning.

Trustworthy, Deployable Freshness Assessment

By integrating a Grad-CAM-guided MiniCAM attention module into a lightweight MobileNetV2 + Xception ensemble, this framework reaches 99.78% test accuracy (99.66% ± 0.23% under 5-fold CV) while eliminating background-induced bias — redirecting model attention from packaging trays to genuine meat freshness cues, at 825 FPS on just 4.3M parameters.

Bias-mitigated | **Explainable** | **Edge-deployable** | **Real-time** | **High-accuracy**

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