Food Watch

A CV app for Allergen Detection

Computer Vision Final Project

2026

GIURISSICH ANDREA
PINTORE MATTEO

Starting point

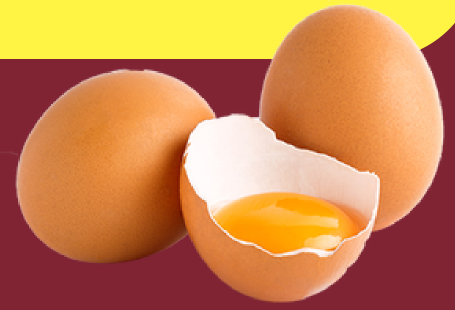
ALLERGEN30: DETECTING FOOD ITEMS WITH POSSIBLE ALLERGENS USING DEEP LEARNING-BASED COMPUTER VISION[1]

- By Mishra, M., Sarkar, T., Choudhury, T. et al. 2022
- Food allergies impose a significant health concern on the community
- Presents a new food allergen dataset
- Develops different CV allergen recognition models

The Allergens

The paper pinpoints 6 main allergens

Ovomucoid



Lactose

Histamine

Gluten

Salicylate

Caffeine

The Dataset

Allergen30

Datset Overview

14.5K 416*416 **FOOD IMAGES**
DIVIDED IN
30 CLASSES AND
HAND LABELED

Train-Test Split

TRAIN SET **70%**

VAL SET **20%**

TEST SET **10%**

Augmentation

RANDOM **SHEAR** OF **±15° HORIZONTALLY** AND **±15° VERTICALLY**

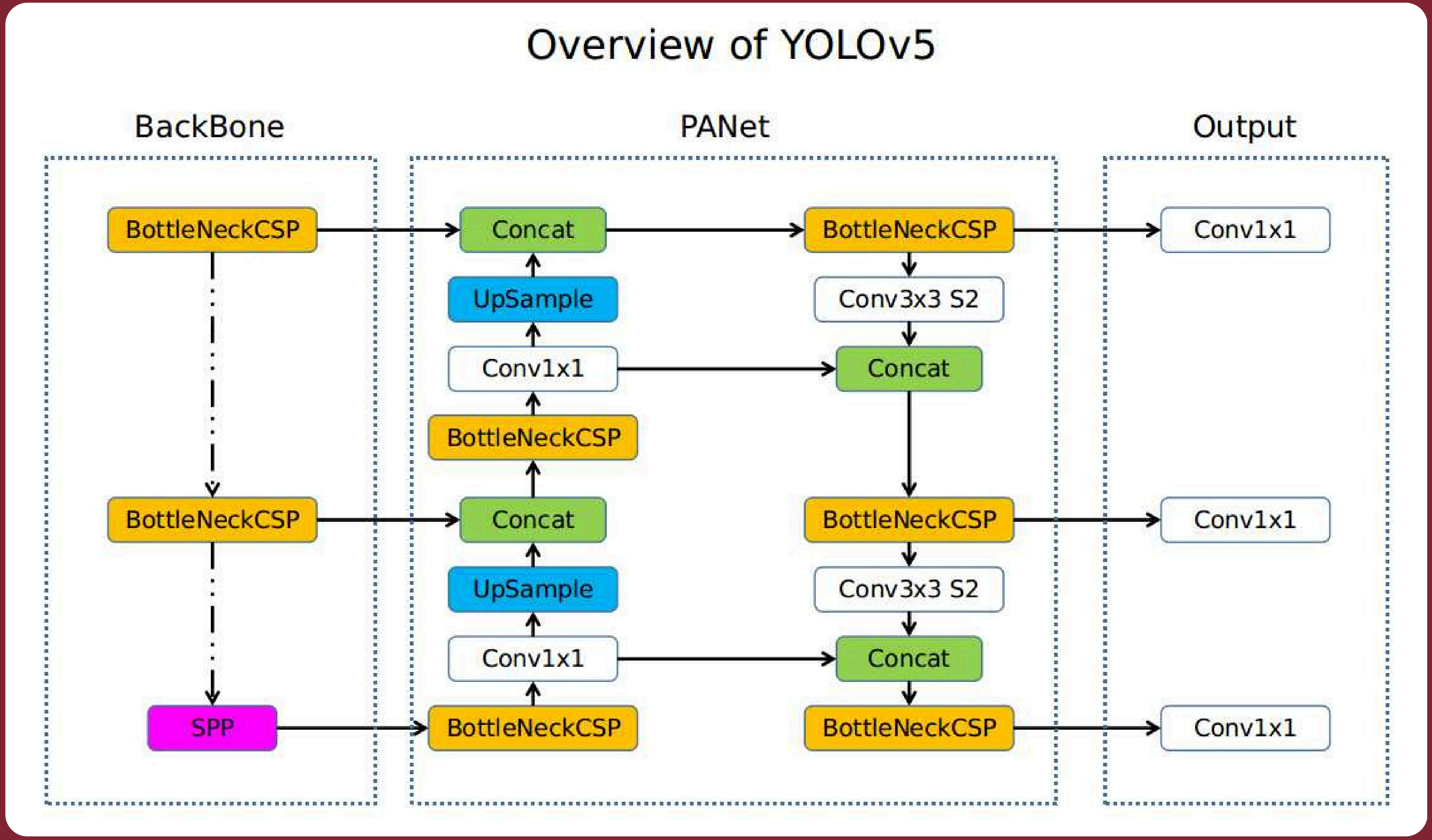
RANDOM **90° ROTATIONS** OF THE **BOUNDING BOXES**

Sample train image with shear and 180° box rotation

Paper’s models

MODEL	PRECISION	RECALL	F1	MAP
YOLOv5s	0.801	0.681	0.7361	0.747
YOLOv5m	0.861	0.679	0.7592	0.749
YOLOv5l	0.845	0.707	0.7698	0.766
YOLOR	0.830	0.741	0.7829	0.811

The Model
YOLOv5



BACKBONE

Feature extraction with **BottleNeckCSP** and **Spatial Pyramid Pooling** (SPP) to capture context at multiple scales

PANet NECK

Feature fusion in two flows:

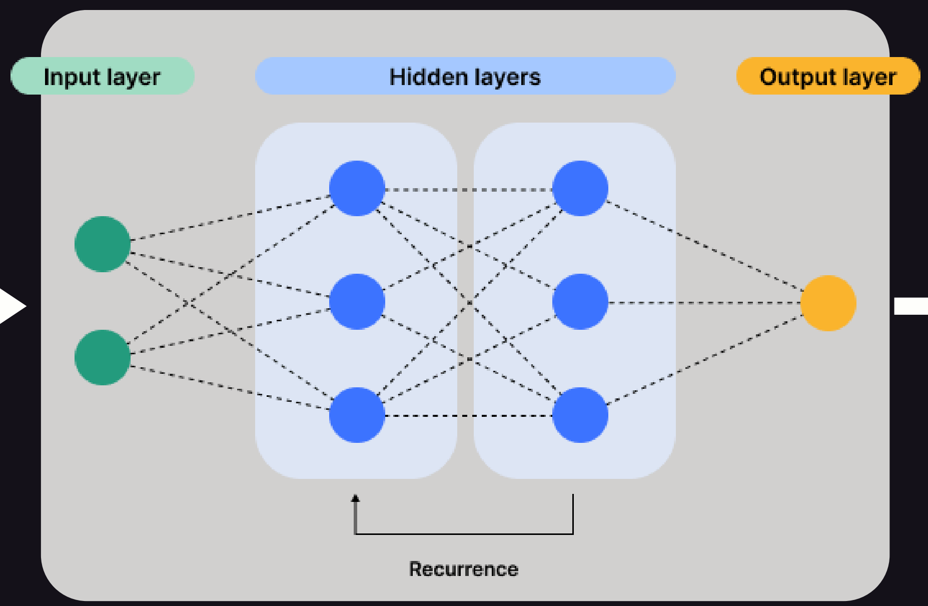
- one upsamples semantic features then concatenates shallow details.
- the other downsamples and concatenates

OUTPUT

Three heads to make predictions on different scales simultaneously

FOOD WATCH

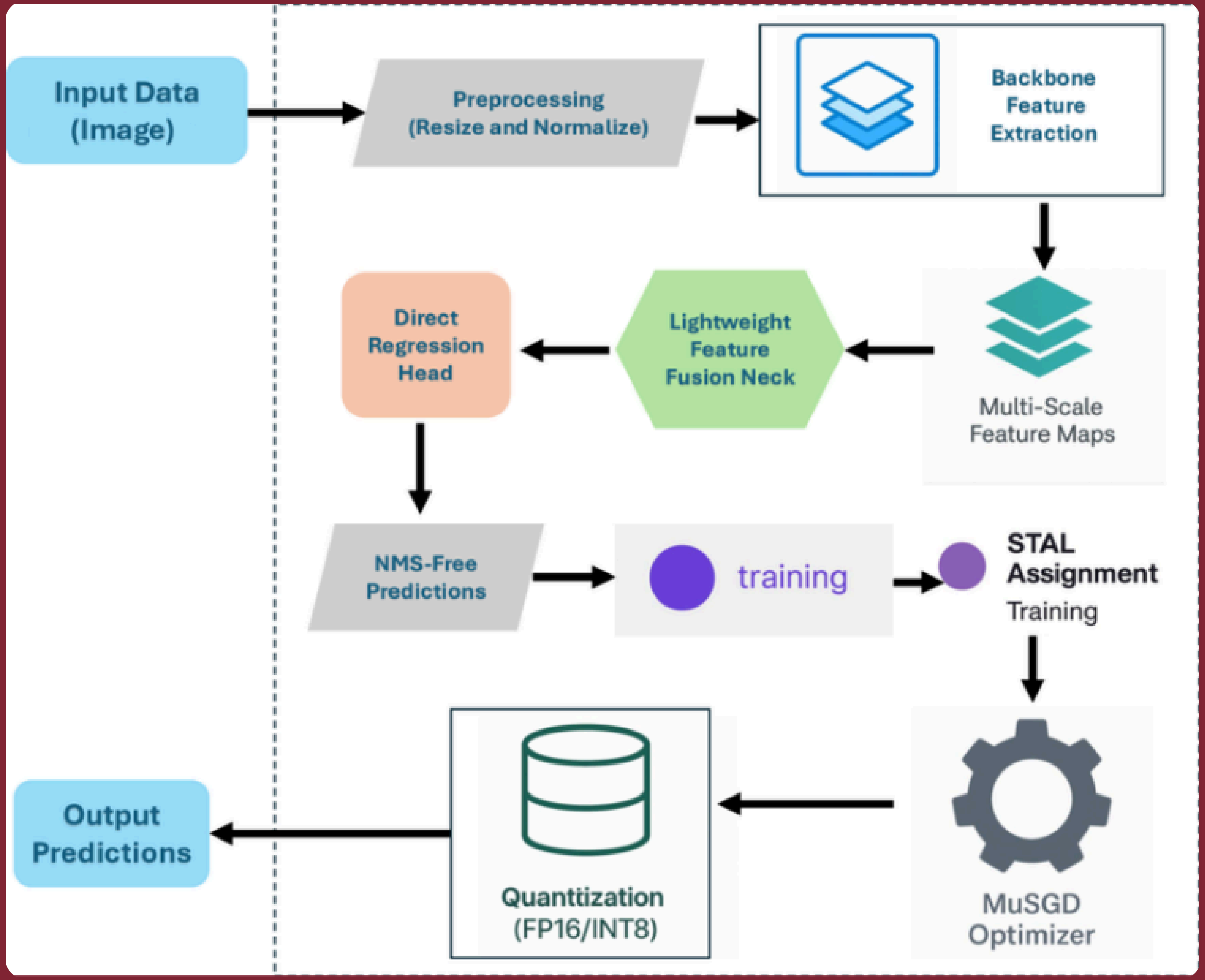
Our Goal



SALICYLATE

- Develop a fast object detection model for mobile use
- Beat the previous models' performances

The Model
YOLO26

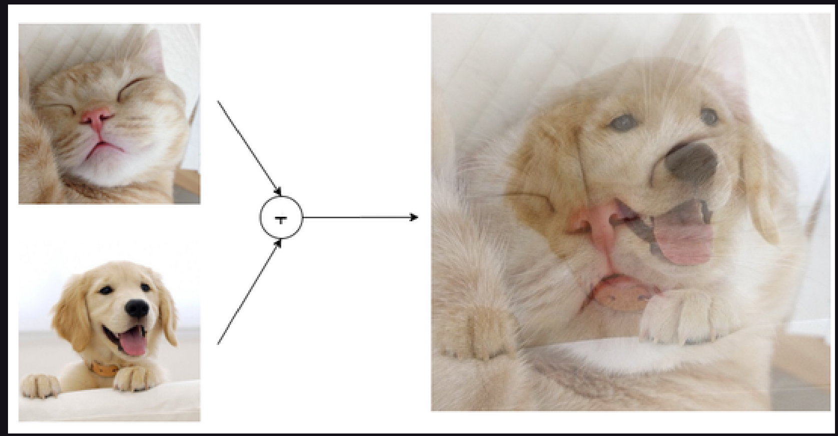


- Using **lightweight backbone** and feature fusion neck enables feature extraction with minimal CPU usage
- It removes **Distribution Focal Loss** (DFL) for faster CPU inference.
- Adopts a native end-to-end **NMS-free** design based on one to one matching
- **Quantization** enables faster computation with simpler types like INT8
- Aims to **directly predict the final box** avoiding complex post processing steps

FOOD WATCH

Training
YOLO26

- **100 Epochs, 32 Batch size** with **Patience of 25** to enable **Early Stopping**
- **AdamW** optimizer with a **Starting Learning Rate of 0.001** and a **Cosine Scheduler** for smooth convergence
- **Label Smoothing** to prevent overfitting
- To be more robust on smaller objects we combined **Mixup** and **Mosaic**



MIXUP

MOSAIC

Evaluation

3 models

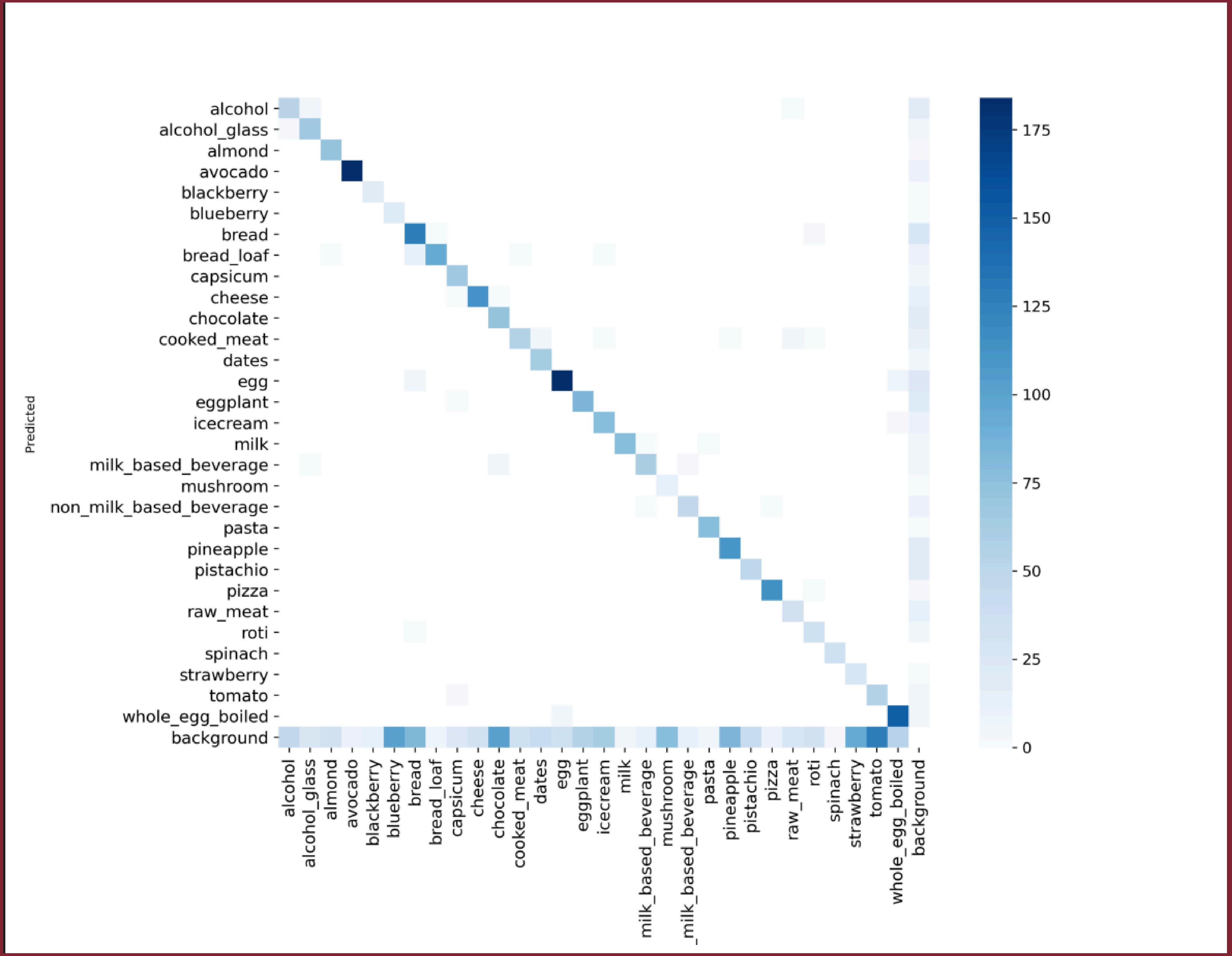
MODEL	PRECISION	RECALL	F1	MAP
YOLO26n	0.823	0.636	0.717	0.730
YOLO26s	0.835	0.664	0.740	0.769
YOLO26m	0.838	0.640	0.726	0.746
YOLOv5s	0.801	0.681	0.736	0.747

FOOD WATCH

Evaluation

Confusion matrix

- Mushrooms are problematic due to intra-class variance



The App Logic

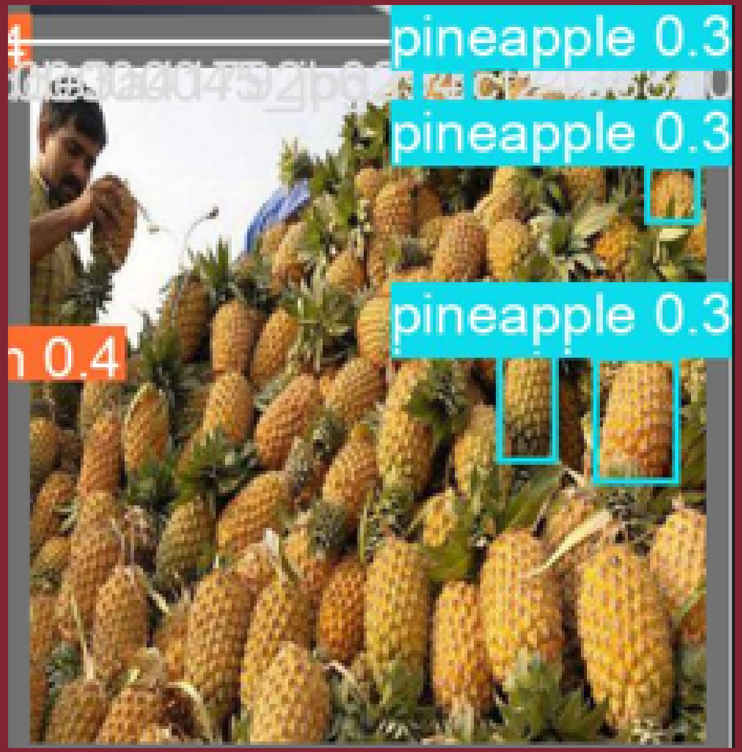
INITIALLY THE SYSTEM DETECTED LOW CONFIDENCE PREDICTIONS FOR VARIOUS BACKGROUND TEXTURES

- We couldn’t rely on a global threshold because it would cut out smaller objects
- We implemented a post processing layer directly in Kotlin
- We applied a Hybrid Filtering composed of **Confidence Thresholds** and **Aspect Ratio Physics**

Conclusions

And further developments

- Our **YOLO26s** performance is **slightly better** than the paper's results
- We suggest using the small model for mobile production



- Smaller models tend to have low recall on images with multiple instances

EXPAND THE DATASET

- Real scenario images
- Align with EU regulations on allergens[2]

Thank you for your attention!

Questions?

[1] Mishra, M., Sarkar, T., Choudhury, T. et al. Allergen30: Detecting Food Items with Possible Allergens Using Deep Learning-Based Computer Vision. Food Anal. Methods 15, 3045–3078 (2022); doi: 10.1007/s12161-022-02353-9

[2] EFSA NDA Panel (EFSA Panel on Dietetic Products, Nutrition and Allergies), 2014. Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes. EFSA Journal 2014; doi:10.2903/j.efsa.2014.3894