



# Damaged vs Undamaged Egg

An AI image classifier that flags cracked egg shells

Muhammad Usama · Candidate No: 2345

LRN International GCSE Artificial Intelligence (7923) — Paper 2



# The Problem

- The system looks at a photo of an egg and decides whether the shell is damaged or not.
- A hairline crack lets bacteria in, but writing fixed rules for 'a crack' is very hard - cracks vary in size, angle and shadow.
- So the model learns what damage looks like from example photos instead.

# Tool: Edge Impulse



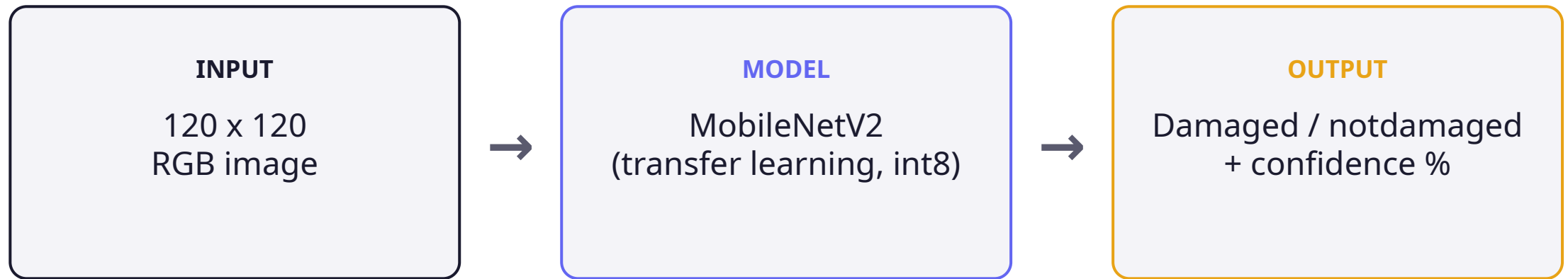
- A browser-based platform for building image models with no coding.
- I chose it because I can upload my own dataset and train in the browser.
- It exports to WebAssembly, so the check can run on the device at the point of inspection.

# The Data



- *Dataset source and licence:* \_\_\_\_\_
- Two labels: damaged and notdamaged.
- Removed unclear photos, then split 80/20 into training and test data.

# How It Works



The photo is resized to 120x120, packed into 14,400 RGB values, and the model returns two scores gated at a 0.60 threshold.



# Training and Testing

- The model was trained on the training images.
- It was then tested on separate images it had never seen.
- Test accuracy: 79%
- Testing on unseen data proves it learned what damage looks like, not just memorised photos.

# Improving the Model



## BEFORE

- Missed faint cracks and low-contrast damage
- e.g. hairline cracks, shadowed or dirty shells

## AFTER

- Added more varied damaged examples
- Retrained the model
- Higher confidence on the same hard eggs

*Test → find failures → add examples → retrain → compare.*



# Ethics and Users

- Privacy: the model runs locally in the browser, so no photo of stock leaves the device.
- Limitation: only two labels, so a dirty or oddly shaped but sound egg can still be pushed into 'damaged'.
- Bias: trained on mostly one shell colour, so brown or speckled eggs may score worse; the fix is a more varied dataset.



# Reflection

- Hardest part: finding photos where the damage is genuinely visible rather than implied by the caption.
- What I learned: a wrong 'undamaged' costs far more than a wrong 'damaged', so the threshold matters as much as the accuracy.
- Next improvement: add a third label for dirty-but-sound eggs so those stop being called damage.