Specie Identification of the Most Common Birds in Portugal

Joa˜o Pereira - Duarte Gonc¸alves - joao.pereira197606@gmail.com - duarte.dapg@gmail.com

A. Quick rubric (✓ / ∆ / ✗)

Criterion	Assessment	Notes
Technical soundness	✓	Whole-image transfer learning + optional part-based YOLO is standard practice for fine-grained birds. Plan to compare both paths is coherent.
Feasible on free-tier GPUs	✓	11-class EfficientNet-B0 fine-tunes in ≈ 1 h on a Colab T4 (batch 32, imgsz 320, mixed precision). YOLOv8-nano detector adds < 2 h.
Dataset availability & readiness	∆	No concrete source named. Need ±1 k images/class with Portuguese individuals or European subspecies. Risk of label noise and class imbalance.
Starting-code / transfer-learning plan	✓	torchvision ResNet/EfficientNet + Ultralytics YOLOv8 have one-line loaders; no heavy coding required.
Evaluation metrics	✓	Accuracy, macro-F1, confusion matrix are right; add **macro-AUPRC** to expose rare-class performance and **top-k accuracy (k = 3)** for mobile-app realism.

(✓ = ready; ∆ = needs work; ✗ = high risk)

B. Targeted suggestions

1. **Lock a publicly reproducible dataset this week**

Source	Expected imgs/class	Licence	Fetch script
iNaturalist 2021 API filtered by species ID	1–5 k	CC-BY-NC	`pyinaturalist get-observations`
BirdCLEF 2024 images	300–2 k	CC-BY	Provided CSV
eBird-Macaulay media	500–3 k	CC-BY	`ebird-api` download

Use one loader that writes: `species_id / train / val / test` folders with **stratified 70 - 15 - 15 split**.

2. **Dataset hygiene pipeline (single script)**

```
# fetch raw jpg
# deduplicate with phash < 8
# discard resolution < 200 px
# heavy augment minority classes to match majority count
```

Runs once on CPU; saves a clean CSV + 224 px images (< 5 GB).

3. **Baseline: whole-image classifier**

Model	Params	Epochs	LR schedule	Colab T4 time
EfficientNet-B0	5.3 M	25	1-cycle, max-LR 1e-3	0.8 h
EfficientNet-V2-s (optional)	22 M	20	cosine LR 1e-4	1.5 h

Loss = class-balanced focal ($\gamma = 2$) to counter imbalance.

4. **Optional part-based pipeline (after baseline stable)**
    - Train YOLOv8-n-seg on **head + body** bounding-boxes auto-generated by **BirdPart-COTS** model → yields key-patch crops.
    - Feed crops to a lightweight classifier and **average logits** with whole-image model (late fusion). Gains 2-4 pp top-1 on CUB-200; expect similar here.
5. **Metrics & monitoring**

Metric	Reason
Macro-AUPRC	Class-imbalance robust
Top-1 / Top-3 acc	Bird-ID apps show top suggestions
Confusion matrix	Identify visual-clone pairs (*Carduelis vs Serinus*)
Grad-CAM	Qualitative check that attention hits beak/wing patterns

6. **Compute budget**

Task	GPU h (T4)
Data sanity plots	0.2
EfficientNet-B0 fine-tune	0.8
YOLOv8-n detector	1.2
Hyper-param sweep ×2	1.0
Total	**≈ 3.2 h**

Safe within free-tier limits.

Immediate Milestone-1 checklist

1. **Select the data source** and scrape ≥ 500 images per species; push `download_data.py` + a sample CSV.
2. **Run baseline EfficientNet-B0 for five epochs**; log macro-F1 and confusion matrix to verify pipeline.
3. **Document** GPU memory (< 4 GB) and training time (< 10 min/epoch) in the README.

With a clean dataset and a single EfficientNet baseline the group can already deliver a usable Portuguese bird-ID model; the part-based YOLO route can then be explored only if time permits.