SeleneX Mini-Prototype — Model Card

Model Overview

Purpose. Assist clinicians by predicting benign vs. malignant ovarian tumors from ultrasound images and basic biomarkers (Age, CA-125, BRCA).
Intended Use. Research/educational prototype; not a medical device.
Out of Scope. Autonomous diagnosis, triage, or treatment decisions.

Data

Image Modality. MMOTU 2D ovarian ultrasound images (JPG).
Labels. Case-level benign/malignant derived from provided class IDs.
Tabular Modality. Synthetic biomarkers aligned to image IDs: Age (28–78), CA-125 (5–35 for benign; 35–600 malignant, skewed higher), BRCA (0/1; 10–15% positive with higher prevalence among malignant).
Splits. Train 70%, Val 15%, Test 15% with effort to maintain case-level consistency.

Model Architectures

- **Image-only:** Lightweight CNN (3 conv blocks, GAP, dense, sigmoid).
- **Tabular-only:** MLP with two hidden layers, sigmoid output.
- **Fused:** Parallel encoders (CNN + MLP), concatenation, dense, sigmoid.

Training & Tuning

Objective. Binary cross-entropy; monitor ROC-AUC on validation for early stopping.
Regularization. Data augmentation (rot/shift/flip/brightness), optional dropout/L2, class weights for imbalance.
Optimization. Adam, batch size 32, epochs with early stopping.

Evaluation

Report the following on the held-out test set:

- **Threshold-agnostic:** ROC-AUC.
- **Thresholded (default 0.5 or tuned):** Accuracy, Precision, Recall, F1.
- **Diagnostics:** ROC curves; confusion matrices per model.

Placeholder Metrics (replace with your results):

Model	Acc	Prec (Malig)	Rec (Malig)	F1 (Malig)	ROC-AUC
Image-only	–	–	–	–	–
Tabular-only	–	–	–	–	–
Fused	–	–	–	–	–

Intended Users & Context

Radiology/ML researchers exploring multimodal diagnostic pipelines. Requires expert oversight; outputs are decision support, not determinations.

Ethical Considerations

Biases. Synthetic biomarkers may not match real distributions; demographic generalization unknown.
Harms. False negatives may delay treatment; false positives may drive unnecessary anxiety or tests.

Limitations

- Small dataset; potential overfitting.
- Ultrasound variability across devices/sites not fully represented.
- Synthetic tabular features with a hard CA-125 cutoff.
- No clinical covariates beyond Age/CA-125/BRCA.

Disclaimer

This prototype is for research/education only and is **not** FDA/CE approved. Do not use for clinical diagnosis or patient management.