Project G13: Toxic Plant Classification

- **Start with non-DL baseline**: Random Forest on handcrafted features (RGB histograms 64-bin, GLCM contrast/energy/homogeneity, Canny edge density)—if accuracy >80%, justify the CNN necessity.
- For the CNN choice, you can **start with EfficientNet-B0** (5M params, ∼20MB) not B7 (66M params)—B7 maybe an overkill for 10K images.
- DeepTextSpotter/ASTER references are confusing—**clarify**: do Kaggle images have text labels on leaves? If not, remove OCR?

Some actions before Milestone-2

- **Dataset verification**: Download Kaggle dataset, verify 10 classes × ∼1000 images, confirm toxic/non-toxic labels exist separately from species labels, check for class imbalance.
- **Create validation split**: Use 70/15/15 train/val/test (stratified by both species and toxicity).
- **Random Forest baseline**: Extract color histograms + GLCM texture + shape features (circularity, aspect ratio), train RF (100 trees), report binary accuracy (toxic/non-toxic) and 10-class accuracy on val set.
- **Define task priority**: Is primary goal binary toxicity classification (safer), 10-class species ID (more informative), or both? Metrics and loss functions depend on this choice.
- **Recommended: Setup tracking**: W&B or MLflow with fixed seeds (random 42, torch 42, numpy 42), log augmentation params (rotation ±30°, flip, ColorJitter hue=0.1), track training time.

Ablations

- **Attention (CBAM) on/off**: Add CBAM before final classifier—hypothesis is +2–4% accuracy by focusing on leaf texture/veins vs background, testing SCAM-Herb findings.
- **Data augmentation strength**: Compare light (rotation/flip only) vs heavy (rotation/flip/ColorJitter/Cutout)—expect +5–10% test accuracy with heavy augmentation given iNaturalist background diversity.
- **Pretrained weights**: ImageNet vs random init—expect +10–15% accuracy boost from pretraining on 10K images, validating transfer learning despite plant-specific features.

Risks & mitigations

- **Risk**: Model overfits to background (iNaturalist images vary in setting) rather than plant features—**Mitigation**: Apply Cutout or MixUp, use Grad-CAM to verify attention on leaves not background.
- **Risk**: Toxic/non-toxic pairs are visually similar (poison ivy 3 leaflets vs virginia creeper 5 leaflets)—**Mitigation**: Focus on high precision for toxic class (minimize false negatives), report confusion matrix for similar pairs.

Open questions

- Are you building one multi-task model (shared backbone, two output heads: binary + 10-class) or two separate models—which is primary evaluation metric?
- Why reference “multimodal approach (NDVI imagery + CNN)” in lit review—does your Kaggle dataset include multispectral channels beyond RGB, or is this a misunderstanding?
- Will you collect any real-world test images (phone photos with poor lighting, partial occlusion) to validate beyond iNaturalist distribution, or rely solely on Kaggle holdout?