Research Papers Summary

Title	Dataset name and URL	Dataset description (samples, classes, images per class or split)	Methods name	Accuracy of the model	Research Questions	Pros and Cons	Citation
Transfer Learning from One Cancer to Another via Deep Learning Domain Adaptation	Multi Cancer Dataset (Kaggle). URL: doi:10.34740/KAGGLE/DSV/9537604	4 cancer types: breast, kidney, lung, and colon. 10,000 samples per type (5,000 normal, 5,000 malignant). Split: 70/15/15.	Domain Adversarial Neural Network (DANN) with ResNet-50 backbone. XAI: Integrated Gradients.	95.56% (DANN adapted to unlabeled lung data).	Can a classifier distinguish malignant from normal cases across organs? Do domain-adaptation strategies improve robustness to shift?	**Pros:** Mitigates label scarcity; leverages biologically meaningful features. **Cons:** Poor transfer between CT and histopathology.	Justin Cheung, et al. (2025).
Cancer Detection and Classification Using CNN Model	Multi cancer image dataset (Kaggle). URL: www.kaggle.com/datasets/obulisainaren/multi-cancer	8 cancer types (e.g., lymphoma, oral, brain). 2,000 images per class for training, 500 for testing (Total: 65,000).	CNN architecture with 3 convolutional and 3 pooling layers; Softmax output.	Above 90% for all models. Kidney cancer achieved 100% accuracy.	Can manual examination errors be alleviated using automated CNNs? How does CNN compare to traditional ML?	**Pros:** High accuracy; reduces human error (fatigue/bias). **Cons:** "Black box" nature; high computational resource needs.	Percy Okae, et al. (2024).
Enhanced and Interpretable Prediction of Multiple Cancer Types Using a Stacking Ensemble Approach with SHAP Analysis	Lung Cancer (Data World), BCWD (Kaggle), Cervical Cancer (Kaggle). Dataset Links: www.kaggle.com/datasets/mysarah	Lung: 309 samples; Breast: 569 samples; Cervical: 835 samples. SMOTE used for balancing. Split: 70:30.	Stacking Ensemble using 12 base learners (LDA, QDA, etc.) and meta-models. XAI: SHAP analysis.	99.28% average accuracy across three cancer datasets.	Can stacking ensemble models improve accuracy using lifestyle/clinical data? How does SHAP enhance interpretability?	**Pros:** Outperforms base learners; provides global/local interpretability. **Cons:** Limited to 3 cancer types; demographic homogeneity.	Shahid Mohammad Ganie, et al. (2025).

	madbhat/lung-cancer						
Browser-Based Multi-Cancer Classification Framework Using Depthwise Separable Convolutions for Precision Diagnostics	Composite dataset from 8 public sources (LC25000, SIPaKMeD, BreakHis, etc.).	Over 130,000 images spanning 26 cancer types. 4,000 train / 500 test images per class.	Fine-tuned Xception architecture; deployed via TensorFlow.js. XAI: Grad-CAM.	Top-1 Accuracy: 99.85%; Top-5 Accuracy: 100%.	Can a privacy-preserving framework perform real-time browser inference? Is Xception more efficient than VGG16/ResNet?	**Pros:** Privacy-preserving; zero infrastructure cost. **Cons:** Artificial balancing; Grad-CAM attribution is coarse.	Divine Sebukpor, et al. (2025).
Evaluation of Vision Transformers for Multi-Organ Tumor Classification Using MRI and CT Imaging	Combined Brain MRI, Lung CT, and Kidney CT datasets.	Brain (15,000 MRIs), Lung (1,054 CTs), Kidney (11,929 CTs). Total: ~28,000 images.	Vision Transformer (ViT), Swin Transformer (Swin-t), and MaxViT Transformer.	Swin Transformer achieved 99.43% accuracy on the combined dataset.	Can ViTs generalize across multi-organ datasets and modalities (MRI/CT)? Benefits of combining datasets?	**Pros:** High accuracy across modalities; low computational footprint for Swin-t. **Cons:** Lack of intra-organ multimodal analysis.	Óscar A. Martín and Javier Sánchez. (2025).