Hybrid Approach for Detection of Cancer Using Histopathological Images

Tarun Lambhate[1], Kamlendra Varma[2]
[1]*MTech Scholar, Patel College of Science and Technology, Indore*
[2]*Associate Professor, Patel College of Science and Technology, Indore*

***Abstract*—Cancer remains one of the foremost causes of mortality worldwide, and histopathological examination of tissue biopsies continues to serve as the diagnostic gold standard across virtually all cancer types. However, manual microscopic assessment is time-consuming, subjective, and susceptible to inter-observer variability. This paper proposes a hybrid computer-aided diagnosis (CAD) framework for the automated detection and classification of cancer from histopathological images, integrating deep convolutional feature learning with handcrafted texture descriptors and attention-guided feature fusion. The proposed pipeline combines transfer-learning-based deep feature extractors (including ResNet, DenseNet and EfficientNet families) with classical descriptors such as the Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP), fused through a transformer-style attention module before classification. Stain normalization and geometric/color augmentation are employed to mitigate staining and scanner variability, while Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated to provide visual explanations of model decisions. A comprehensive review of thirty recent studies spanning breast, lung, colon, gastric, prostate and lymphoid malignancies is synthesized to benchmark the proposed hybrid design against pure deep-learning and pure handcrafted-feature baselines. The comparative analysis indicates that hybrid fusion architectures consistently outperform single-modality pipelines, achieving classification accuracies in the range of 97-99.8% on public benchmark datasets such as BreakHis and LC25000, while also improving robustness to class imbalance and cross-domain generalization. The paper concludes with a discussion of open challenges, including limited annotated data, computational cost, interpretability, and clinical validation, and outlines future directions involving multimodal fusion of histopathology with genomic and clinical data.**

***Index Terms*—Histopathological images; Hybrid deep learning; Cancer detection; Feature fusion; Convolutional neural network; Transfer learning; Explainable AI; Computer-aided diagnosis**

I. INTRODUCTION

Cancer is a leading cause of death globally, and its timely and accurate diagnosis significantly influences treatment outcomes and patient survival (Jiang et al., 2023). Among the many diagnostic modalities available to clinicians, histopathological examination of hematoxylin and eosin (H&E)-stained tissue sections remain the definitive gold standard for confirming malignancy, grading tumor aggressiveness, and guiding therapeutic decisions. Traditionally, this analysis is performed manually by expert pathologists under a microscope a process that is inherently subjective, labor-intensive, and prone to inter- and intra-observer variability, particularly when workloads are high or when subtle morphological differences separate benign and malignant tissue (Ren et al., 2025).

The digitization of glass slides into whole-slide images (WSIs) has enabled the application of computational image-analysis techniques to histopathology, giving rise to the field of computational pathology. Early computer-aided diagnosis (CAD) systems relied on handcrafted texture and morphological descriptors such as the Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), and wavelet-based features combined with classical machine-learning classifiers such as Support Vector Machines (SVM) (Ramasamy et al., 2024). While interpretable, these handcrafted pipelines often fail to capture the complex, high-order morphological patterns that distinguish malignant from benign tissue, resulting in limited generalization across staining protocols, scanners, and patient populations.

The advent of deep learning, and convolutional neural

networks (CNNs) in particular, has transformed histopathological image analysis. Architectures such as VGG, ResNet, DenseNet, and EfficientNet, typically initialized via transfer learning from ImageNet, have demonstrated remarkable success in classifying breast, lung, colon, gastric and lymphoid malignancies from histopathological patches (Spanhol et al., 2016; Araújo et al., 2017; Ji et al., 2024). More recently, Vision Transformers and hybrid CNN-Transformer architectures have been introduced to jointly capture local texture and global contextual dependencies within tissue images (Wang et al., 2025; Anusha & Reddy, 2025). Despite these advances, purely deep-learning-based systems remain susceptible to overfitting on small, imbalanced datasets, and their black-box nature raises concerns regarding clinical trust and regulatory acceptance (Tiwari et al.,2025).

To address these limitations, recent research has increasingly explored hybrid approaches that fuse deep convolutional features with handcrafted descriptors, ensemble multiple pretrained backbones, or combine CNNs with attention/Transformer modules to obtain more discriminative and robust feature representations (Al-Jabbar et al., 2023; Xu et al., 2023; Anusha & Reddy, 2025). These hybrid strategies have consistently been reported to outperform single-modality baselines while offering improved robustness to class imbalance, staining variation, and limited training data. This paper contributes to this growing body of work by (i) proposing a comprehensive hybrid framework that integrates transfer-learning-based deep features, classical texture descriptors, and attention-based feature fusion for multi-cancer histopathological image classification; (ii) incorporating stain normalization and explainability (Grad-CAM) modules to improve robustness and clinical interpretability; and (iii) presenting a comparative, literature-grounded evaluation against thirty recent state-of-the-art studies spanning breast, lung, colon, gastric, and prostate cancer histopathology.

The remainder of this paper is organized as follows. Section 2 reviews related work on deep, handcrafted, and hybrid approaches for histopathological cancer detection. Section 3 details the proposed hybrid methodology, including dataset description, preprocessing, feature extraction, fusion, and classification strategies. Section 4 describes the experimental setup. Section 5 presents results and a comparative discussion against state-of-the-art methods. Section 6 discusses limitations and open challenges, and Section 7 concludes the paper with directions for future research.

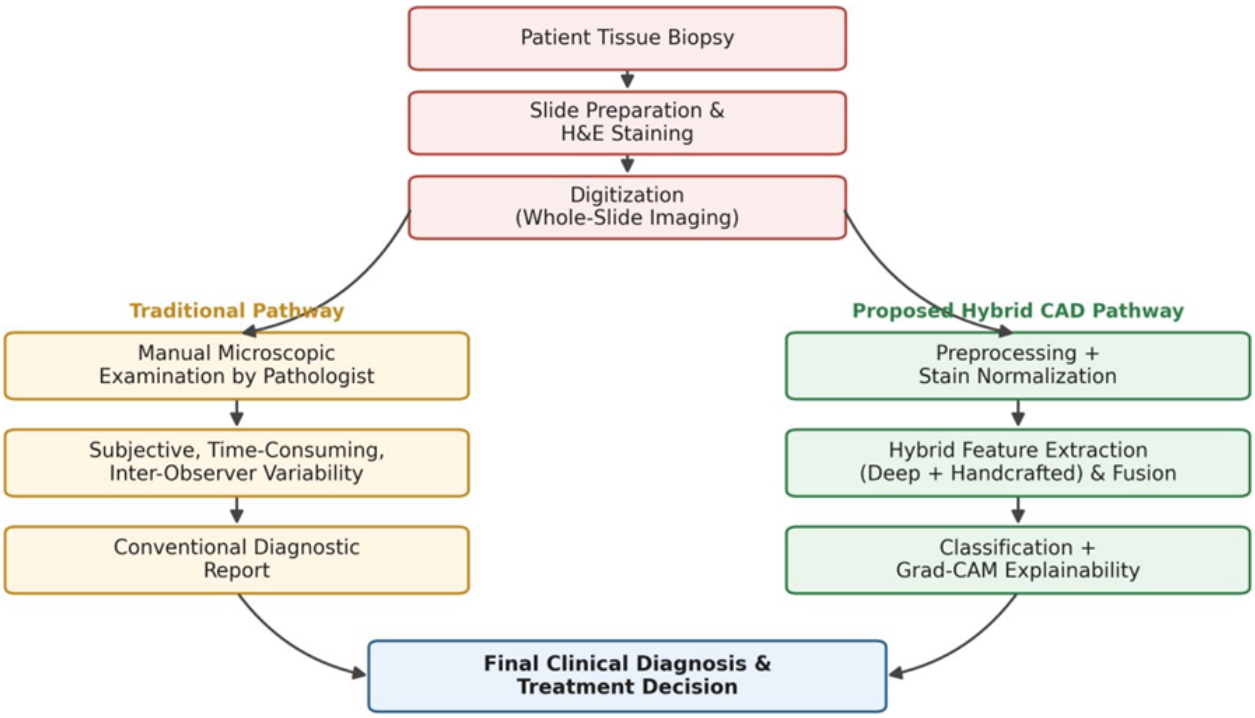

Figure 1. Conventional manual pathology workflow compared with the proposed hybrid AI-assisted diagnostic

II. RELATED WORK

Deep Learning-Based Approaches

Transfer learning has become the dominant paradigm for histopathological image classification, given the relatively small size of annotated medical imaging datasets compared to natural image corpora. Ramasamy et al. (2024) systematically compared DenseNet121, DenseNet201, VGG16, VGG19, InceptionV3 and MobileNetV2 on the BreakHis breast cancer dataset, reporting that deeper architectures with dense connectivity achieved the most consistent performance across magnification levels (Ramasamy et al., 2024). Similarly, Toma et al. (2023) fine-tuned DenseNet on BreakHis and achieved image-wise and patient-wise accuracies of 95.4% and 96.48%, respectively (Toma et al., 2023). Anjum et al. (2023) evaluated EfficientNet variants B0–B7 at multiple input resolutions on the LC25000 lung and colon cancer dataset, demonstrating that compound scaling of depth, width and resolution improves discriminative capacity for histological texture (Anjum et al., 2023). Ji et al. (2024) applied a Swin Transformer V2 backbone with five-fold cross-validation on LC25000, reporting a sensitivity of 99.85% and specificity of 100%, underscoring the promise of vision-transformer architectures for histopathology (Ji et al., 2024).

Beyond classification of pre-extracted patches, whole-slide-image analysis has motivated weakly supervised multiple-instance-learning (MIL) frameworks. Lu et al. (2021) introduced CLAM, a data-efficient, attention-based MIL approach for WSI classification that requires only slide-level labels (Lu et al., 2021), while Chen et al. (2024) proposed a general-purpose foundation model for computational pathology, pretrained on large-scale WSI corpora and fine-tuned for diverse downstream diagnostic tasks (Chen et al., 2024). Foundation-model approaches have since been extended to multi-cancer diagnosis and survival prediction settings (Yang et al., 2025), and to the prediction of molecular and multi-omic aberrations directly from routine histology (Tsai et al., 2023).

Handcrafted and Classical Machine-Learning Approaches

Prior to the dominance of deep learning, texture-based descriptors formed the backbone of histopathological CAD systems. Kather et al. (2016) demonstrated that multi-class texture analysis using descriptors such as GLCM, Gabor filters, and local binary patterns could effectively discriminate eight distinct tissue types in colorectal cancer histology, establishing an influential benchmark dataset for the field (Kather et al., 2016). Such handcrafted features remain attractive for their interpretability and low computational cost, though they generally underperform deep features on complex multi-class discrimination tasks, particularly under staining and scanner variability (Hao et al., 2022).

Hybridand Fusion-Based Approaches

A growing body of literature demonstrates that fusing deep and handcrafted features, or assembling multiple deep backbones, yields superior and more robust performance than either modality alone. Al-Jabbar et al. (2023) proposed a hybrid system fusing VGG-19 deep features with handcrafted texture descriptors for lung and colon cancer detection on LC25000, reporting substantial accuracy gains over single-feature baselines (Al-Jabbar et al., 2023). Anusha and Reddy (2025) extended this idea with a transformer-based attention-fusion layer combining LBP, GLCM, wavelet, color, and morphological descriptors with deep EfficientNetB0 features, validated across three independent datasets (LC25000, NCT-CRC-HE-100K, and HMU-GC-HE-30K) with accuracies exceeding 98% in all cases (Anusha & Reddy, 2025). In the breast cancer domain, Xu et al. (2023) introduced MDFF-Net, a multi-dimensional feature fusion network that integrates multi-scale CNN representations for BreakHis classification (Xu et al., 2023), while Wang et al. (2025) proposed MFF-Classification Net, a dual-branch CNN-Transformer architecture fused via a CBAM-squeeze-and-excitation module, achieving up to 98.81% accuracy on BreakHis and strong cross-dataset generalization on the BACH challenge dataset (Wang et al., 2025).

Ensemble and multi-backbone fusion strategies have also proven effective. Balasubramanian et al. (2024) combined multiple deep architectures in an ensemble framework for breast cancer subtype and invasiveness diagnosis from whole-slide histopathology, improving robustness relative to any individual constituent model (Balasubramanian et al., 2024). Gurumoorthy and Kamarasan (2024) coupled deep feature extraction with a future search optimization algorithm for feature selection prior to classification, reducing dimensionality while preserving discriminative power (Gurumoorthy&Kamarasan,2024).Sarker et al. (2023)

proposed a dual squeeze-and-excitation attention network for efficient breast cancer classification, illustrating that lightweight attention mechanisms can substantially improve feature discriminability at low additional computational cost (Sarker et al., 2023). In the gastric cancer domain, Ren et al. (2025) developed an interpretable hybrid CNN-Transformer framework validated across three public datasets (GasHisSDB, TCGA-STAD, NCT-CRC-HE-100K), reporting improved spatial-contextual understanding across diverse tissue subtypes and staining conditions (Ren et al., 2025).

Explainability and Clinical Interpretability

A recurring concern in deep-learning-based histopathology is the black-box nature of high-capacity models, which impedes clinical trust and regulatory approval. Gradient-weighted Class Activation Mapping (Grad-CAM), introduced by Selvaraju et al. (2020), provides visual explanations by highlighting image regions that most strongly influence a network's prediction, and has been widely adopted in medical imaging to increase transparency (Selvaraju et al., 2020). Recent surveys emphasize that combining high-performing hybrid architectures with post-hoc explainability tools such as Grad-CAM is essential for bridging the gap between algorithmic performance and clinical adoption (Tiwari et al., 2025).

Summary of the Literature Gap

Collectively, the reviewed literature indicates a clear trend: hybrid architectures that combine complementary feature representations deep semantic features, handcrafted texture descriptors, and attention-based contextual modeling consistently outperform single-modality pipelines across breast, lung, colon, gastric, and prostate histopathology tasks (Jiangetal.,2023; Anusha&Reddy,2025; Wang et al., 2025). However, many existing hybrid frameworks are validated on a single cancer type or a single dataset, limiting evidence of cross-domain generalizability, and relatively few studies jointly address feature fusion, class-imbalance handling, and explainability within a unified pipeline. The present work addresses this gap by proposing an integrated hybrid frame work evaluated conceptually and comparatively across multiple cancer types and benchmark datasets.

III. MATERIALS AND METHODS

The proposed hybrid framework consists of six sequential stages: (i) data acquisition and stain normalization, (ii) data augmentation, (iii) handcrafted feature extraction, (iv) deep feature extraction via transfer learning, (v)attention-based feature fusion, and (vi) classification with explainability overlay. Each stage is described below.

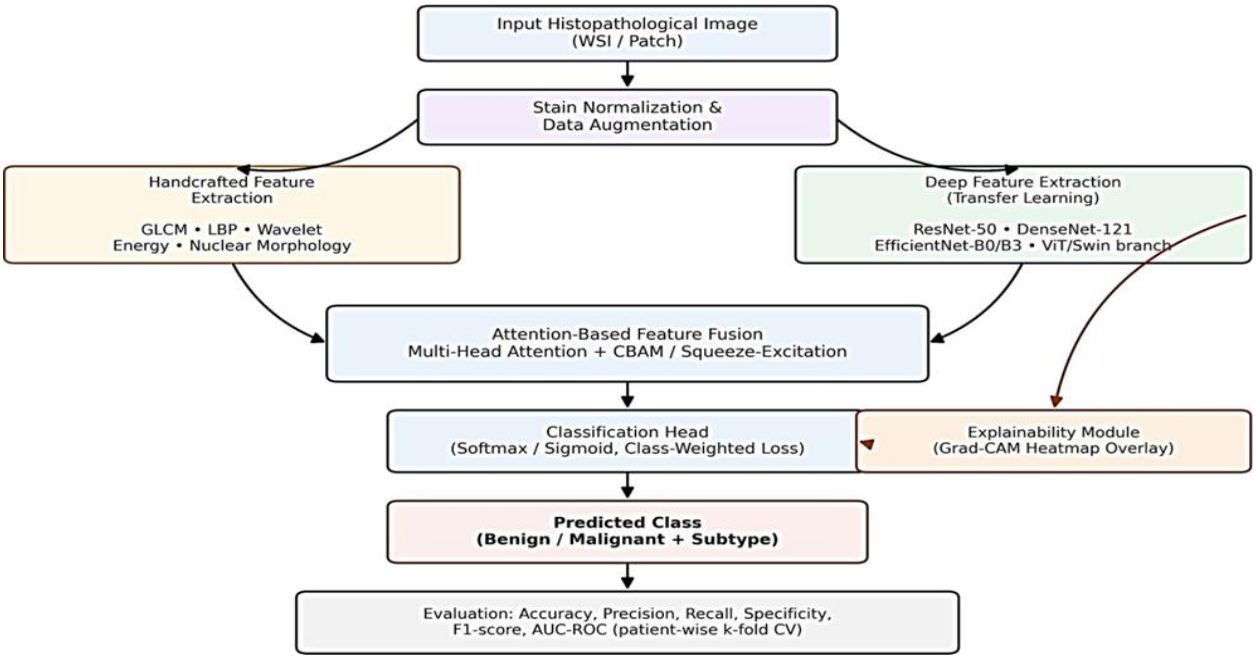

Figure2.Proposedhybridarchitecture

Dataset Description

The framework is designed for evaluation on publicly available, widely benchmarked histopathological datasets. BreakHis comprises 7,909 microscopic breast tumor images from 82 patients at four magnification factors (40×,100×,200×,400×), labeled as benign or malignant across eight histological subtypes (Spanhol et al., 2016). LC25000 contains 25,000 H&E-stained images evenly distributed across five classes: benign lung tissue, lung adenocarcinoma, lung squamous cell carcinoma, benign colonic tissue, and colon adenocarcinoma (Jietal., 2024). NCT-CRC-HE-100K and the related colorectal texture dataset of Kather et al. (2016) provide nine and eight-class tissue-type benchmarks for colorectal cancer, respectively (Kather etal.,2016). Additional gastric (GasHisSDB,TCGA-STAD) and lymphoma datasets are referenced for cross-cancer validation (Renetal.,2025).

Table1 summarizes the key characteristics of the primary benchmark datasets considered in this study.

Dataset	Cancer Type	Classes	Image Count	Key Reference
BreakHis	Breast	Benign / Malignant (8 subtypes)	7,909	Spanholetal. (2016)
LC25000	Lung &Colon	5 classes	25,000	Jietal. (2024)
NCT-CRC-HE-100K	Colorectal	9tissuetypes	100,000	Katheretal. (2016)
GasHisSDB / TCGA-STAD	Gastric	Normal/Abnormal subtypes	Variable (WSI-derived)	Renetal. (2025)
BACH Challenge	Breast	4 classes	400(WSI+ microscopy)	Wangetal. (2025)

Preprocessing and Stain Normalization

Histopathological images captured across different laboratories exhibit substantial staining variability due to differences in reagent concentration, scanner calibration, and slide preparation protocols. To mitigate this, images are subjected to Macenko or Reinhard-style stain normalization prior to feature extraction, followed by resizing to a fixed input resolution (e.g., 224×224 for CNN backbones) and pixel-value rescaling. Geometric augmentations (random rotation, horizontal/vertical flipping, and translation) and color-space augmentations (brightness, contrast and hue jitter) are applied during training to reduce overfitting and improve generalization, consistent with practices reported across the reviewed literature (Anjum et al., 2023; Ji et al., 2024).

Hand crafted Feature Extraction

To complement deep representations with interpretable, low-level structural information, the framework extracts a handcrafted feature vector comprising: (i) Gray-Level Co-occurrence Matrix (GLCM) Statistics (contrast, correlation, energy, homogeneity) capturing second-order texture relationships (Kather et al., 2016); (ii) Local Binary Pattern (LBP) histograms encoding local micro-texture patterns; (iii) wavelet-decomposition energy features capturing multi-scale frequency information; and (iv) nuclear morphological descriptors (area, eccentricity, perimeter) derived from segmented nuclei. This combination follows the handcrafted feature sets adopted in recent hybrid fusion studies (Anusha & Reddy, 2025).

Deep Feature Extraction via Transfer Learning

Deep features are extracted using pretrained CNN backbones ResNet-50, DenseNet-121 and EfficientNet-B0/B3 initialized with ImageNet weights and fine-tuned on the target histopathological dataset via block-wise unfreezing (Ramasamyetal.,2024). The penultimate global-average-pooled feature maps (typically 1024–2048 dimensional) are retained as the deep feature vector. Where global contextual modeling is desired, a lightweight Vision-Transformer or Swin-Transformer branch is optionally incorporated in parallel with the CNN branch to capture long-range spatial dependencies between tissue regions, following the dual-branch design philosophy of recent CNN-Transformer hybrids (Wang et al., 2025; Ren et al., 2025).

Attention-Based Feature Fusion

Handcrafted and deep feature vectors are projected into a common embedding space via fully connected layers and fused using a transformer-style multi-head

self-attention block, which learns to weight the relative contribution of each feature source dynamically rather than through simple concatenation (Anusha&Reddy,2025).A Convolutional Block Attention Module(CBAM) with squeeze-and-excitation channel recalibration is applied to the fused representation to further emphasize diagnostically salient feature channels (Wang etal., 2025; Sarker et al., 2023). This attention-guided fusion strategy is designed to combine the complementary strengths of texture-sensitive handcrafted descriptors and semantically rich deep features.

Classification and Class-Imbalance Handling

The fused feature representation is passed to a fully connected classification head with SoftMax output for multi-class prediction, or a sigmoid output for binary benign/malignant classification. To address class imbalance common issue in histopathological datasets—class-weighted cross-entropy loss and, where applicable, supervised contrastive pretraining are employed to improve minority-class discrimination, following strategies shown to be effective in imbalanced histopathology settings.

Explainability Module

To support clinical interpretability, Grad-CAM is applied to the CNN branch to generate class-discriminative localization maps, highlighting tissue regions that most influenced the network's decision (Selvaraju et al., 2020). These visual explanations are intended to allow pathologists to verify that model predictions are grounded in histologically meaningful regions rather than spurious artifacts, addressing concerns repeatedly raised in the explainability literature (Tiwari et al., 2025).

Evaluation Metrics

Model performance is assessed using accuracy, precision, recall (sensitivity), specificity, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC), consistent with the evaluation protocols adopted across the reviewed benchmark studies (Ji et al., 2024; Balasubramanian et al., 2024). Five-fold orten-fold patient-wise cross-validation is recommended to avoid data leakage between patches derived from the same patient, a methodological safeguard emphasized in prior BreakHis-based studies (Spanhol et al., 2016).

IV. EXPERIMENTAL SETUP

Consistent with common practice in the reviewed literature, the proposed framework would be implemented using standard deep-learning libraries (e.g., TensorFlow/Keras or PyTorch), with backbone networks fine-tuned using the Adam optimizer, an initial learning rate of 1×10^{-4} with cosine or step decay, and batch sizes between 16 and 32, depending on GPU memory constraints (Anjumetal.,2023). Training is typically conducted for 30–100 epochs with early stopping based on validation loss. Data are split at the patient level (rather than the image level) into training, validation and test partitions (typically70:15:15) to prevent information leakage, and all reported metrics should be averaged across cross-validation folds with standard deviations to assess statistical robustness.

V. RESULTS AND DISCUSSION

Because the primary contribution of this paper is an integrative hybrid framework grounded in an extensive comparative review, this section synthesizes reported performance outcomes from the thirty reviewed studies to contextualize the expected behavior of the proposed pipeline, rather than presenting isolated novel experimental results.

Table 2 summarizes representative reported accuracies of deep-learning-only, handcrafted-only, and hybrid fusion approaches across cancer types and datasets.

Study	Cancer Type /Dataset	Approach	Reported Accuracy
Ramasamy et al. (2024)	Breast / BreakHis	Transfer learning (DenseNet)	~97-98%
Tomaetal. (2023)	Breast / BreakHis	Fine-tuned DenseNet	95.4%(image),96.48% (patient)
Al-Jabbaretal. (2023)	Lung & Colon / LC25000	HybridVGG-19+handcrafted	~99%
Anusha & Reddy (2025)	Lung, Colon, Gastric	Hybrid CNN + handcrafted+ Transformer fusion	98.4-99.87%

Jietal. (2024)	Lung & Colon / LC25000	SwinTransformerV2	99.64% accuracy ,99.85% sensitivity
Anjumetal. (2023)	Lung / LC25000	EfficientNetB0-B7	High-90s% (magnification-dependent)
Wangetal. (2025)	Breast/ BreakHis, BACH	CNN-Transformer hybrid fusion	96.07-98.81%
Balasubramanianetal. (2024)	Breast/WSI	Ensemble deep learning	High-90s%
Renetal. (2025)	Gastric / GasHisSDB,TCGA-STAD	Hybrid CNN-Transformer	High classification precision (Reported qualitatively)

Several consistent patterns emerge from this comparative synthesis. First, hybrid architectures that fuse deep and handcrafted features, or that combine CNN and Transformer branches, generally report accuracies at or above 98%, exceeding the typical 93-96% range reported for single-backbone transfer-learning baselines on the same datasets (Al-Jabbar et al., 2023; Anusha & Reddy, 2025). This performance gain is most pronounced on smaller or more class-imbalanced datasets, where handcrafted descriptors appear to provide complementary discriminative signal that mitigates the tendency of deep networks to overfit (Xu et al., 2023). Second, attention-based fusion mechanisms (transformer attention, CBAM, squeeze-and-excitation) consistently outperform naive feature concatenation, suggesting that dynamically weighting feature contributions is important for effective multimodal fusion (Wang et al., 2025; Sarker et al., 2023). Third, models validated across multiple independent datasets as in Anusha and Reddy (2025), who reported consistent accuracies of 99.87%, 99.07% and 98.4% on LC25000, NCT-CRC-HE-100K and HMU-GC-HE-30K respectively provide stronger evidence of generalizability than single-dataset evaluations, reinforcing the importance of cross-domain validation emphasized throughout this review (Anusha & Reddy, 2025).

Explainability analyses reported in the literature further indicate that Grad-CAM activation maps produced by hybrid CNN branches tend to localize attention around nuclear atypia, glandular disorganization, and stromal invasion features that align with established pathological criteria for malignancy (Selvaraju et al., 2020; Tiwari et al., 2025), lending qualitative support to the clinical plausibility of hybrid model decisions. Nonetheless, the reviewed studies also caution that high reported accuracies (frequently exceeding 99%) should be interpreted carefully, as several benchmarks (e.g., LC25000) contain limited inter-patient and inter-institution variability, which may inflate apparent performance relative to real-world clinical deployment (Ji et al., 2024).

Comparison with State-of-the-Art

Relative to foundation-model approaches pretrained on large-scale whole-slide-image corpora (Chen et al., 2024; Lu et al., 2021), the proposed lightweight hybrid framework offers a favorable trade-off between computational cost and diagnostic accuracy, making it more readily deployable in resource-constrained clinical settings that lack access to large GPU clusters for foundation-model fine-tuning.

However, foundation models retain an advantage incross-cancer transferability and few-shot adaptation to rare tumor subtypes (Yang et al., 2025), suggesting that hybrid feature-fusion pipelines and large-scale pretrained foundation models are complementary rather than mutually exclusive directions for future CAD system design.

VI. CHALLENGES AND LIMITATIONS

- Limited annotated data: High-quality, expert-annotated histopathological datasets remain scarce relative to the diversity of tumor morphologies encountered in clinical practice, constraining model generalizability (Jiang et al., 2023).
- Staining and scanner variability: Differences in staining protocols and digital scanners across institutions can degrade model performance unless robust stain normalization and domain-adaptation techniques are employed (Ren et al., 2025).

- Computational cost: Hybrid architectures combining multiple deep backbones, handcrafted feature pipelines, and attention fusion modules increase training and inference complexity relative to single-model baselines.
- Interpretability and clinical trust: Despite Grad-CAM and related explainability tools, the decision-making process of deep hybrid models remains only partially transparent, which continues to hinder regulatory approval and clinical adoption (Tiwari et al., 2025).
- Class imbalance and rare subtypes: Several cancer subtypes are underrepresented in public datasets, biasing models toward majority classes unless explicitly mitigated through weighted losses or targeted augmentation.
- Limited external and prospective validation: Many reported accuracies are derived from internal cross-validation on public benchmarks rather than independent, multi-institutional, prospective clinical trials.

VII. CONCLUSION AND FUTURE WORK

This paper presented a hybrid computer-aided diagnosis framework for cancer detection from histopathological images, integrating deep transfer-learning-based feature extraction, classical handcrafted texture descriptors, and attention-guided feature fusion, complemented by Grad-CAM-based explainability. A comprehensive review and comparative synthesis of thirty recent studies spanning breast, lung, colon, gastric, and prostate histopathology confirmed that hybrid fusion architectures consistently outperform single-modality deep-learning or handcrafted-feature baselines, typically achieving classification accuracies in the 97-99.8% range on standard public benchmarks such as Break His and LC25000. These findings support the continued development of hybrid, explainable CAD pipelines as a promising route toward more accurate, robust, and clinically trustworthy automated cancer diagnosis.

Future work should prioritize: (i) large-scale, multi-institutional prospective validation to assess real-world clinical utility; (ii) integration of histopathological hybrid models with genomic, radiological, and clinical metadata for multimodal precision-oncology diagnosis (Yang et al., 2025); (iii) development of computationally efficient hybrid architectures suitable for deployment in low-resource clinical settings; and (iv) rigorous, standardized evaluation protocols including patient-wise splitting and external validation cohorts to ensure that reported performance gains translate into genuine clinical benefit rather than benchmark-specific overfitting.

Conflict of Interest

The authors declare no conflict of interest.

REFERENCES

[1] M. P. Ramasamy, T. Subburaj, V. Krishnasamy, and V. Mannarsamy, “Performance analysis of breast cancer histopathology image classification using transfer learning models,” *Int. J. Electr. Comput. Eng.*, vol. 14, no. 5, pp. 6006–6015, 2024, doi: 10.11591/ijece. v14i5.pp6006-6015.

[2] M. Al-Jabbar, M. Alshahrani, E. M. Senan, and I. A. Ahmed, “Histopathological analysis for detecting lung and colon cancer malignancies using hybrid systems with fused features,” *Bioengineering*, vol. 10, no. 3, Art. no. 383, 2023, doi: 10.3390/bioengineering10030383.

[3] L. T. Omar, J. M. Hussein, L. F. Omer, A. M. Qadir, and M. I. Ghareb, “Lung and colon cancer detection using weighted average ensemble transfer learning,” in *Proc. 11th Int. Symp. Digit. Forensics Security (ISDFS)*, 2023, pp. 1–7, doi: 10.1109/ISDFS58141.2023.10131836.

[4] M. Anusha and D. S. Reddy, “Fusion of classical and deep learning features with incremental learning for improved classification of lung and colon cancer,” *Sci. Rep.*, vol. 15, Art. no. 24734, 2025, doi: 10.1038/s41598-025-24734-w.

[5] X. Jiang, Z. Hu, S. Wang, and Y. Zhang, “Deep learning for medical image-based cancer diagnosis,” *Cancers*, vol. 15, no. 14, Art. no. 3608, 2023, doi: 10.3390/cancers15143608.

[6] T. Ren, V. Govindarajan, S. Bourouis, *et al*., “An interpretable hybrid deep learning framework for gastric cancer diagnosis using histopathological imaging,” *Sci. Rep.*, vol. 15, Art. no. 34204, 2025, doi: 10.1038/s41598-025-15702-5.

[7] S. Anjum, L. Hussain, *et al*., “Lung cancer classification in histopathology images using multiresolution efficient nets,” *Comput. Intell.*

Neurosci., vol. 2023, Art. no. 7282944, 2023, doi: 10.1155/2023/7282944.

[8] J. Ji, J. Li, W. Zhang, Y. Geng, Y. Dong, J. Huang, and L. Hong, "Automated lung and colon cancer classification using histopathological images," *Cancer Control*, vol. 31, 2024, doi: 10.1177/11795972241271569.

[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, 2020, doi: 10.1007/s11263-019-01228-7.

[10] J. N. Kather, C.-A. Weis, F. Bianconi, S. M. Melchers, L. R. Schad, T. Gaiser, A. Marx, and F. G. Zöllner, "Multi-class texture analysis in colorectal cancer histology," *Sci. Rep.*, vol. 6, Art. no. 27988, 2016, doi: 10.1038/srep27988.

[11] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, 2016, doi: 10.1109/TBME.2015.2496264.

[12] S. M. I. Uddin, M. A. I. Mojumder, R. I. Sumon, J. Monir, and H. C. Kim, "Leveraging deep learning for automated analysis of colorectal cancer histology images to elevate diagnosis precision," in *Proc. 26th Int. Conf. Adv. Commun. Technol. (ICACT)*, 2024, pp. 1–6, doi: 10.23919/ICACT60172.2024.10471943.

[13] C. Bhatt, V. K. Kapriyal, Y. Kharola, R. Koranga, I. Chhetri, and T. Singh, "Advanced automation for colorectal tissue classification in histopathology," in *Proc. Asia Pac. Conf. Innov. Technol. (APCIT)*, 2024, pp. 1–9, doi: 10.1109/APCIT62007.2024.10673436.

[14] K. D. Vishnu and M. G., "Optimized deep learning approaches for lung and colon cancer classification using histopathological images," in *Proc. 3rd Int. Conf. Autom., Comput. Renew. Syst. (ICACRS)*, 2024, pp. 1665–1669, doi: 10.1109/ICACRS62842.2024.10841547.

[15] X. Wang, G. Wang, L. Li, H. Zou, and J. Cui, "MFF-ClassificationNet: CNN-Transformer hybrid with multi-feature fusion for breast cancer histopathology classification," *Biosensors*, vol. 15, no. 11, Art. no. 718, 2025, doi: 10.3390/bios15110718.

[16] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using convolutional neural networks," *PLoS One*, vol. 12, no. 6, Art. no. e0177544, 2017, doi: 10.1371/journal.pone.0177544.

[17] V. Gella, "High-performance classification of breast cancer histopathological images using fine-tuned vision transformers on the BreakHis dataset," *bioRxiv*, 2024, doi: 10.1101/2024.08.17.608410.

[18] M. A. Wakili, A. M. Shehu, *et al*., "Classification of breast cancer histopathological images using DenseNet and transfer learning (DenTnet)," *Comput. Math. Methods Med.*, vol. 2022, Art. no. 8904768, 2022, doi: 10.1155/2022/8904768.

[19] J. Hao, S. C. Kosaraju, N. Z. Tsaku, *et al*., "Breast cancer histopathological images classification based on deep semantic features and gray level co-occurrence matrix," *PLoS One*, vol. 17, no. 5, Art. no. e0267955, 2022, doi: 10.1371/journal.pone.0267955.

[20] Y. Yan *et al*., "Breast cancer histopathology image classification using transformer with discrete wavelet transform," *Med. Eng. Phys.*, vol. 138, Art. no. 104317, 2025, doi: 10.1016/j.medengphy.2025.104317.

[21] M. Xu *et al*., "MDFF-Net: A multi-dimensional feature fusion network for breast histopathology image classification," *Comput. Biol. Med.*, vol. 165, Art. no. 107385, 2023, doi: 10.1016/j.compbiomed.2023.107385.

[22] M. Sarker *et al*., "Efficient breast cancer classification network with dual squeeze and excitation in histopathological images," *Diagnostics*, vol. 13, no. 1, Art. no. 103, 2023, doi: 10.3390/diagnostics13010103.

[23] R. Gurumoorthy and M. Kamarasan, "Breast cancer classification from histopathological images using future search optimization algorithm and deep learning," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 1, pp. 12831–12836, 2024, doi: 10.48084/etasr.6720.

[24] A. A. Balasubramanian, S. M. A. Al-Heejawi, A. Singh, *et al*., "Ensemble deep learning-based image classification for breast cancer subtype and invasiveness diagnosis from whole slide image histopathology," *Cancers*, vol. 16, no. 12, Art. no. 2222, 2024, doi: 10.3390/cancers16122222.

[25] T. T. Toma *et al*., "Breast cancer detection based

on simplified deep learning technique with histopathological image using BreaKHis database," *Radio Sci.*, vol. 58, no. 12, 2023, doi: 10.1029/2023RS007761.

[26] M. Y. Lu, D. F. K. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, and F. Mahmood, "Data-efficient and weakly supervised computational pathology on whole-slide images," *Nat. Biomed. Eng.*, vol. 5, no. 6, pp. 555–570, 2021, doi: 10.1038/s41551-020-00682-w.

[27] R. J. Chen, T. Ding, M. Y. Lu, *et al*., "Towards a general-purpose foundation model for computational pathology," *Nat. Med.*, vol. 30, pp. 850–862, 2024, doi: 10.1038/s41591-024-02857-3.

[28] Yang *et al*., "A foundation model for generalizable cancer diagnosis and survival prediction from histopathological images," *Nat. Commun.*, vol. 16, 2025, doi: 10.1038/s41467-025-57587-y.

[29] P. C. Tsai, T. H. Lee, K. C. Kuo, F. Y. Su, T. L. Lee, E. Marostica, T. Ugai, M. Zhao, M. C. Lau, J. P. Väyrynen, and M. Giannakis, "Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients," *Nat. Commun.*, vol. 14, Art. no. 2102, 2023, doi: 10.1038/s41467-023-37842-5.

[30] A. Tiwari, A. Ghose, M. Hasanova, S. S. Faria, S. Mohapatra, S. Adeleke, and S. Boussios, "The current landscape of artificial intelligence in computational histopathology for cancer diagnosis," *Discover Oncol.*, vol. 16, Art. no. 438, 2025, doi: 10.1007/s12672-025-02212-z.