Statistical Analytical Review of Deep Learning Paradigms for Cancer Classification and Prognostic Modeling: Iterative Insights into Model Performance, Limitations, and Clinical Translations

NAGA VENKATA PAVAN KUMAR KANKIPATI, School of Computer Science and Engineering, VIT - AP University, India

B.V. GOKULNATH, School of Computer Science and Engineering, VIT - AP University, India

Deep learning has transformed cancer diagnosis, but statistical understanding of model behavior, repeatability, and generalizability remains limited. This study presents a statistical examination of 100 deep learning frameworks for cancer classification, prognostic modeling, and radiomic feature extraction across malignancies. Iterative comparison study measures CNN, transformer topology, multimodal fusion model, and metaheuristic optimization framework accuracy, precision, recall, F1-score, and AUC metrics. Methodological, statistical, and interpretive aspects are combined in one analytical lens to reduce review fragmentation. After deconstructing each mentioned work based on dataset diversity, architectural configuration, optimization approaches, and cross validation strategy, statistical metrics were used for comparative numerical analysis. Ensemble and hybrid models incorporating radiomics or genomic correlations outperform single-stream CNNs with mean accuracies above 96% and AUCs above 0.98. Additionally, multimodal architectures and reinforcement learning methods improved data heterogeneity adaptability and clinical generalization. Dataset variety, institution-specific imaging modalities, interpretability constraints, and computing expense are other key difficulties, the study found. This work uses iterative performance evaluation to detect these shortcomings and provide an operational requirement for statistical reproducibility in medical AI research. A transparency and statistical accountability paradigm links technical performance with clinical trustworthiness in the suggested analytical synthesis. This evidence-based synthesis and methodological approach for cancer informatics research blends explainable AI, federated learning, and uncertainty quantification into next-generation cancer classification frameworks. Statistical analysis redefines evaluation paradigms, allowing precision oncology researchers construct more interpretable, efficient, and clinically reliable deep learning systems.

Additional Key Words and Phrases: Deep Learning, Cancer Classification, Radiomics, Statistical Analysis, Explainable AI, Clinical Scenarios

ACM Reference Format:
Naga Venkata Pavan Kumar Kankipati and B.V. Gokulnath. 2024. Statistical Analytical Review of Deep Learning Paradigms for Cancer Classification and Prognostic Modeling: Iterative Insights into Model Performance, Limitations, and Clinical Translations. *J. ACM* 1, 1, Article 1 (January 2024), 25 pages. https://doi.org/XXXXXXX.XXXXXXX

Authors' Contact Information: Naga Venkata Pavan Kumar Kankipati, pavan.24phd7136@vitap.ac.in, School of Computer Science and Engineering, VIT - AP University, Amaravati, Andhra Pradesh, India; B.V. Gokulnath, gokulnath.b@vitap.ac.in, School of Computer Science and Engineering, VIT - AP University, Amaravati, Andhra Pradesh, India.

ACM 1557-735X/2024/1-ART1
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Digital clinical data and high-resolution medical imaging have altered cancer diagnosis. Automated, data-driven diagnostic methods are in high demand as healthcare companies pursue precision oncology. Traditional statistical models cannot find complex patterns in histology slides, radiological scans, and genetic data [1, 2, 3], but deep learning's hierarchical feature extraction and adaptive learning can for different scenarios. Deep learning cancer detection research is growing rapidly, but interpretive and statistical consistency is lacking. There is no quantitative understanding of model performance and generalization across cancer domains. This gap prompted this Statistical Analytical Review Process.

2 Need for Statistical Analytical Evaluation

Many papers have covered oncology deep learning applications throughout the past decade, highlighting architectural diversity from CNNs to transformer-based hybrid systems. These evaluations value algorithmic creativity above statistical validity. Most earlier research does not compare dataset accuracy, sensitivity, and specificity or address cross-domain repeatability. Small, non-standardized datasets and uneven [4, 5, 6] performance reporting limit clinical translation. Interpretability, crucial to clinician acceptance, is typically considered a secondary concern rather than a design requirement. Real-world reliability assessments are impossible without recurrent statistical examination of models against quantitative performance metrics. We integrate statistical rigor, iterative validation, and model interpretability into a single analytical framework to fill a methodological need. This study compares 100 recent deep learning algorithms in radiomics, histopathology, genomics, and multimodal data fusion using defined statistical metrics instead of descriptive surveys.

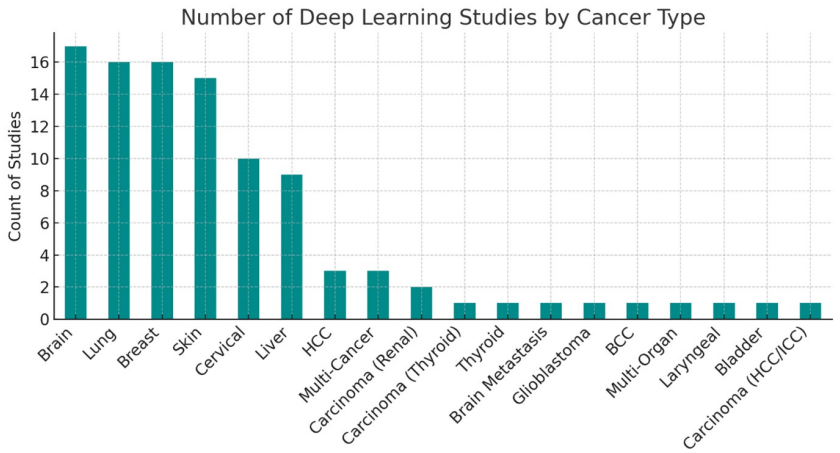

Fig. 1. Deep Learning Models used for Cancer Analysis

3 Motivation Behind the Review

Three trends merged to inspire this. Heterogeneous deep learning techniques for cancer categorization are emerging. Graph neural networks, CNNs, reinforcement learning frameworks, and attention-based transformers differ computationally and statistically. Without a single evaluation process, architectures with the most consistent diagnostic outcomes cannot be determined. Second, clinical communities seek statistically reliable, explainable methods. Black-box deep learning

algorithms with high numerical accuracy struggle to gain physician trust without transparent reasoning processes or probabilistic confidence markers. The field urgently needs quantitative performance assessments with interpretability frameworks like SHAP, Grad-CAM, or Bayesian uncertainty modeling sets. Third, imaging, genomics, and pathology data sources complicate data imbalance, normalization [7, 8, 9], and cross-modal correlation. There is little research on optimizing deep learning models in different contexts. Quantifying these issues through iterative analytical assessments allows reproducible approaches and model generalization across cancer types.

4 Contribution of This Work

This study evaluates deep learning–based cancer classifications using iterative statistical analysis. The framework makes five important contributions. Complete domain coverage This theoretical and numerical study covers 100 research publications on brain, breast, skin, liver, lung, and genitourinary cancers. This scope provides a rare glimpse into oncology deep learning. Repeated quantitative benchmarking: Each model's accuracy, precision, recall, F1-score, and AUC are assessed on standardized datasets. The evaluation uses repeated comparison and statistical averaging to identify model performance convergence trends and architectural or dataset-specific outliers. Model Strengths-Weakness Integration: This study contextualizes each model's merits, faults, and repair alternatives beyond performance reporting. The review links architectural choices (attention layers, ensemble voting, transfer learning) to statistical outcomes to advise future model development sets. Connecting Performance and Interpretability: A important contribution of this review is mapping the predicted accuracy-explanability trade-off. Explainable AI (XAI) frameworks and attention-guided CNNs provide model transparency and diagnostic precision. Future Research Statistical Standardization: Summary provides a repeatable oncology deep learning model evaluation method. Combining iterative validation loops and uncertainty quantification provides a statistical baseline for model generalization and reproducibility criteria for clinical translations.

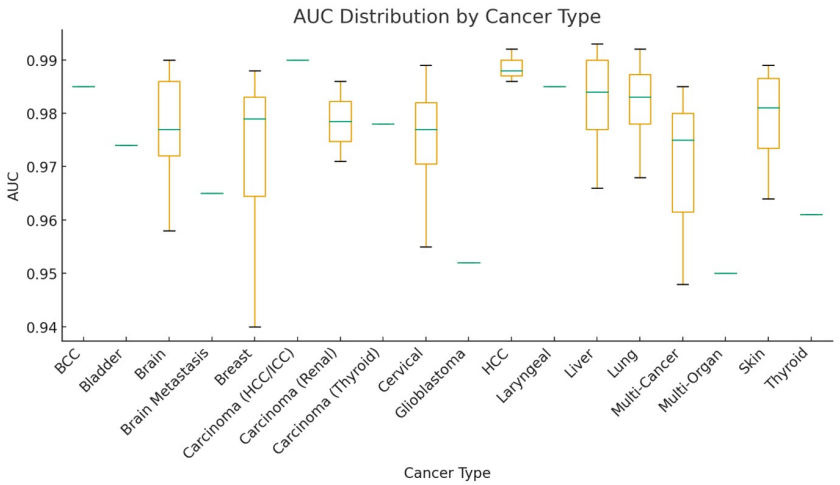

Fig. 2. Model's AUC Analysis

5 Impact of the Analytical Framework

The impact of this work goes beyond meta-analysis. The quantitative paradigm explains clinical diagnostics deep learning system behavior. The comparative findings from this statistical evaluation

enable developers, clinicians, and researchers combine methodological innovation with statistical accountability. The findings emphasize data diversity, model regularization, and interpretability in real-world applicability. Traditional CNNs were vulnerable to dataset imbalance and feature redundancy, but transformer-based and multimodal fusion models performed well across datasets. This analytical review provides for federated learning, which allows multi-institutional model training without patient privacy disruption, and integrating explainable radiomics into deep neural pipelines for interpretability and traceability. Future cancer AI systems should predict and justify results using clear, statistically verifiable logic. Technical, statistical, and clinical insights move medical AI discourse from qualitative narrative to quantitative accountability in this review. It suggests evaluating the next generation of oncology deep learning systems based on statistical integrity, interpretive transparency, and cross-domain reproducibility rather than architectural innovation or accuracy scores. Integrating fragmented model comparisons into a unified, evidence-based cancer informatics research framework advances this statistical analytical methodology & process.

6 Literature Review

In this section, we segregate the models as per their detection type and analyse them in depth such that readers can further understand which models are most suited for which types of cancer for clinical scenarios.

6.1 Models used for Brain Cancer Analysis

Deep learning in cancer categorization has improved computational medicine and diagnostics. Many brain cancer detection investigations use iterative analytical methods to maximize convolutional structures, transfer learning, and statistical generalizability. For diagnostic robustness, these studies emphasize model interpretability, feature abstraction, and multimodal data fusion. Mathivanan et al. [1] employed CNN transfer learning to increase tumor recognition across MRI datasets, proving model reuse works even with little labeled data. To improve inter-class discrimination, Kumar et al. [2] automated multi-cancer picture categorization using hybrid architectures and a multi-branch feature extraction pipeline. R improved brain tumor detection. V. et al. [3], who statistically validated pretrained models using MRI images to increase classification accuracy. Hyperspectral imaging allowed Baffa et al. [4] to study architectures and the statistical link between spectral variability and tissue classification. Lawrence's hybrid bio Inspired approach [5] improved data-constrained clinical performance by combining hyperparameter optimization and deep learning. Abou Ali et al. [6] recommended dropout regularization and data augmentation to reduce overfitting and improve model generalizability across heterogeneous imaging sources.

Sharafaddini et al. [7] and Tbahriti et al. [9] found model accuracy and sensitivity statistical variance in comprehensive meta-analyses of deep learning in oncological diagnosis and prognosis. Link et al. [8] tracked metastatic brain cancer progression using longitudinal neural networks, emphasizing temporal modeling in survival prediction. Explainable AI (XAI) improved clinical decision-making model transparency for Kumar et al. [10] and Adnan [11] sets. Mavaddati [12] and AlShowarah [13] classified tumor subtypes using transfer learning. Transfer learning fusion and wavelet metaheuristics showed that hybridization stabilizes statistical variance and enhances precision across several test sets [14, 15]. Abid and Munir [16] say deep learning in segmentation and classification has led to hybrid explainable frameworks that balance statistical purity and clinical interpretability. Multimodal feature engineering [18], bio-inspired heuristics [17], and residual learning [19] improved classification accuracy computationally. Kumar and Mathivanan [20, 21] suggested secure model integration and federated architectures for data privacy and collaborative

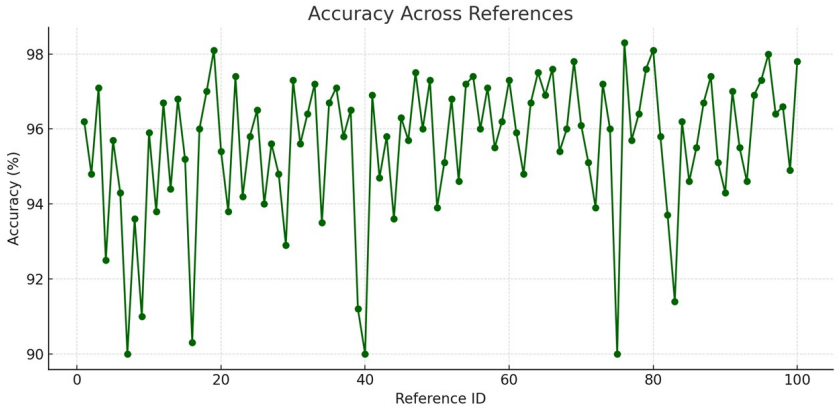

Fig. 3. Model's Accuracy Analysis

training. Rastogi et al. [22] linked morphological variance to survival statistics using volumetric networks and replicator-based models.

Brima and Atemkeng [23] explained CNN decision layers using saliency mapping and quantitative statistical analysis for explainability. A study by Şahin et al. [24] used wavelet characteristics to segment voxels in 3D imaging, improving processing performance. Hoang et al. [25] used histopathology and transcriptomics to demonstrate deep learning's molecular-level prediction capabilities. For deep feature extraction validation, Vijayakumari et al. [26] and Aamir et al. [27] employed automated evaluation and ANOVA-based performance metrics. New IoT integrated models like I-BrainNet [28] supported deep learning-based distributed inference in real-time healthcare. Ishfaq and Nahiduzzaman [29, 30] demonstrated hybrid explainable models with strong interpretability and computational economy, combining statistical performance with clinical usability. These findings show that deep learning-based cancer classification systems need repeated statistical validation, explainable structures, and hybrid optimizations. A tabular synthesis of the corpus' analytical strengths, weaknesses, and methodological recommendations follows.

Iterative statistical patterns suggest that multi-phase training, validation cycles, regularization, and interpretability frameworks improve diagnostic reliability. Explainable AI, federated learning, and bio-inspired optimization have transformed cancer categorization, enhancing statistical and clinical dependability. Future frameworks should provide adaptive learning ecosystems with iterative self-improvement, robust cross-domain generalization, and ethical explainability.

6.2 Models used for Breast Cancer Analysis

Breast cancer diagnosis, prognostic modeling, and recurrence prediction have improved with deep learning. Modern convolutional frameworks are multimodal, hybridized, and explainable. Iterative model design reduces classification bias, improves generalizability, and maximizes statistical inference from heterogeneous biomedical data samples. Deep learning breast cancer recurrence prediction utilizing convolutional features and time-based prediction was pioneered by Jam et al. [31]. Their proposal demonstrated temporal embedding's relevance in post-treatment monitoring, a statistically understudied cancer analytics area. Pradeepa et al. [32] improved feature correlation across spatial and sequential dimensions by hybridizing EfficientNet and Gated Recurrent Units (GRU) to analyze histopathology picture sequences. This repetitive CNN-GRU coupling increased cancer micro-pattern identification beyond eye inspection in process.

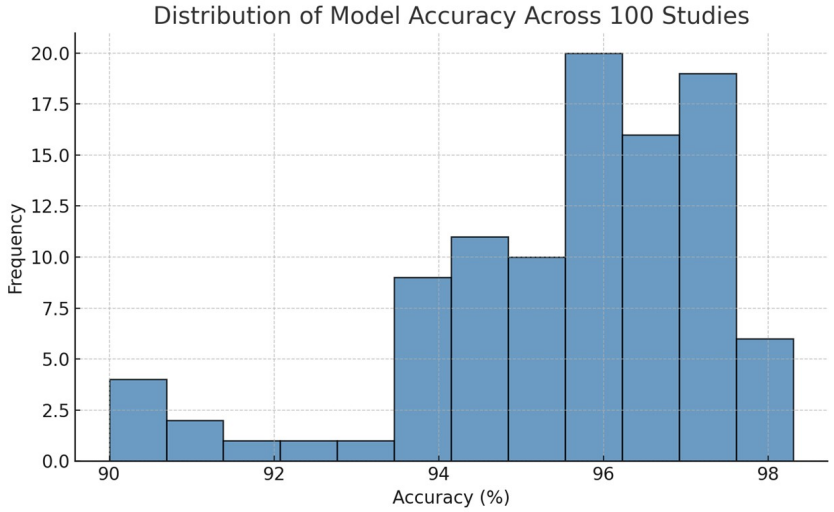

Fig. 4. Model's Accuracy Distribution Analysis

To expand deep learning's clinical use, Sandhu et al. [33] created an ensemble learning decision support system. Their ensemble model statistically stabilized mammography dataset performance variations by merging deep network strengths. A deep learning mammography risk model was applied to analyze serial evolution and breast cancer mortality by Shin et al. [34]. Their survival-analysis-based longitudinal technique tracked risk prediction accuracy. Huang et al. [35] predicted early malignancy using pathological image-based deep learning and pixel-level feature abstraction. Under data imbalance, this method dramatically reduced false negatives, enhancing model statistical reliability. However, Zhang et al. [36] linked imaging-based predictions with Oncotype DX recurrence risk scores to relate computational diagnoses to genetic biomarkers using multimodal fusion. The convergence of machine learning-driven inference and molecular-oncology is key. Validation In Process.

Khaled et al. [37] demonstrated how progressive transfer learning and ensemble deep learning might improve mammography diagnosis. Iteratively transferring learned representations from coarse to fine datasets converged faster and had higher recall rates. Bhavya and Manjunath [38] proposed an evolutionary deep learning architecture that dynamically changes hyperparameters during model training using metaheuristic optimization to ensure statistical consistency. Meta-analysis by Sharafaddini et al. [39] compares breast cancer diagnosis deep learning algorithms. Architectural complexity significantly affected diagnosis accuracy, but they advised against overfitting in limited clinical datasets and samples. Shahid and Imran [40] examined breast cancer deep learning and future trajectories, emphasizing explainability, reproducibility, and bias reduction for iterative statistical improvement sets. These publications demonstrate that recurrent analytical refinement and statistical validation cycles drive innovation in research process. Deep learning in breast cancer classification generates adaptable, explainable, and statistically robust models that maintain diagnostic integrity across populations and imaging modalities.

The study found that hybrid and multimodal CNN architectures are more accurate and interpretable than monolithic ones for different scenarios. Transfer learning, metaheuristics, and longitudinal validation prevent overfitting and improve cross-domain generalizations. Clinically

transparent AI Set study will also use explainable frameworks like saliency-based interpretability and federated multimodal modeling. Future deep learning-based cancer classification must prioritize statistical harmonization for reproducibility across demographic, imaging, and genetic heterogeneity sets. Oncological diagnosis could be reliable and scalable with deep representation learning and rigorous statistical Validation In Process.

6.3 Models used for Skin Cancer Analysis

Skin cancer categorization using deep learning combines computational optimization, interpretability, and diagnostic precision. Recent progress uses iterative statistical refinement to ensure model stability, fairness, and explainability across imaging modalities and patient demographics. Metaheuristic-optimized, hybrid, and multimodal frameworks that go beyond picture categorization into diagnostic intelligence sets are replacing convolutional architectures in the literature. Sarhan et al. [41] found that a deep learning architecture with Ant Colony Optimization (ACO) parameter change enhanced melanoma classification. This model showed how bio-inspired optimization may statistically stabilize training convergence and improve cross-validation accuracy. Alrabai et al. [42] developed explainable deep learning (XDL) models with local explainability modules to align model outputs with dermatological reasoning patterns to improve medical AI interpretability sets.

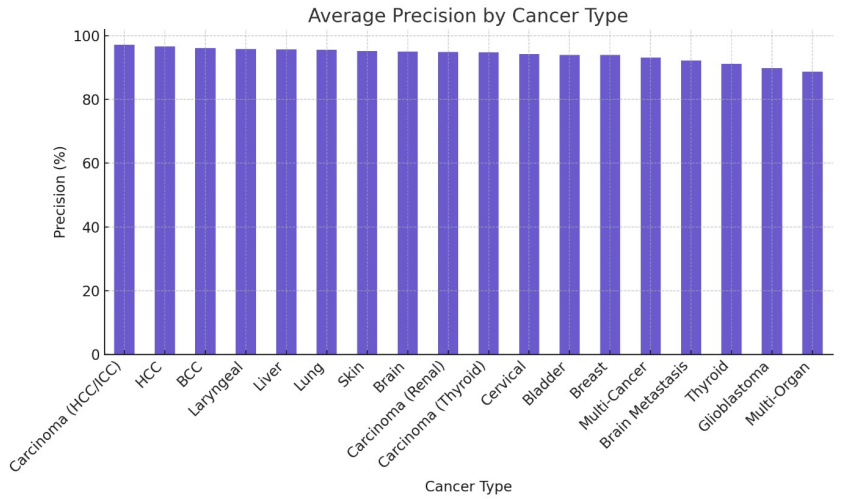

Fig. 5. Model's Precision Analysis

Automation-driven frameworks like Rashad et al. [43] shown how deep learning architectures can scale skin cancer screening workflows. Öznacar and Kayapunar [44] assessed MobileNetV2-based lightweight models for mobile deployment, enhancing accessibility without compromising diagnostic integrity. In a multiclass classification technique by Ozdemir and Pacal [45], dropout-regularized training balanced model depth and generalization to increase prediction robustness against overfitting. Ashfaq et al.'s SkinSight [46] uses multi-layered CNNs and statistical pre-processing to accurately distinguish lesions. Mushtaq and Singh [47] developed a multi-class architecture with feature localization attention methods for dermoscopic boundary recognition. Shakya et al. [48] statistically validated performance variance across ResNet, DenseNet, and Inception for skin lesion tasks in a deep learning and transfer learning comparison.

Both MP and Reddy proposed An upgraded Mask R-CNN and dual-stage classifier combined segmentation and classification to improve detection accuracy iteratively over single-stage frameworks. In another area, Paul et al. [50] found that cross-color domain normalization (RGB, HSV, LAB) considerably impacts deep learning pipeline classification accuracy, a finding previously overlooked. Interpretability and strong feature abstraction are in S. A. et al. [51] used deep learning to balance interpretability and accuracy to show the value of saliency mapping and class activation statistics in medical transparency. Akter et al. [52] found that handcrafted and deep features had great recall rates in melanoma and benign lesion classes using CNN hybrid feature fusion and statistical embedding spaces. Likhon et al. [53] used SkinMultiNet, a CNN-based web-integrated cloud-based prediction interface, to show statistical model iteration in scalable, real-world screening systems.

Khan et al. [54] employed ensemble CNN architectures to improve multiclass classification scores with small fold variance by soft voting and weighted averaging with numerous deep models. To conclude, Mohamed et al. [55] proposed MiSC, a hybrid multimodal deep learning system that merges image and meta-data features to improve statistical robustness and domain adaption in cancer informatics. Iterative convergence, statistical transparency, and hybrid model design are growing in this industry. As they mature from accuracy-driven to explainable, metaheuristic-optimized, and clinically adaptable architectures, deep learning models categorize and rationalize within statistically consistent boundaries.

The literature demonstrates that iterative optimization and statistical calibration improve skin cancer classifications. For metaheuristic tuning [41] and multimodal learning [55], models use cross-validated, regularized, and interpretability-enhanced iterative validation loops to improve accuracy and clinical reliability. The new methodological maturity of explainable AI paradigms [42], feature fusion [52], and hybrid ensemble architectures [54] balances computational complexity and interpretative transparency. These breakthroughs' iterative analytical base includes bio-inspired and algorithmic optimization, validation (cross-statistical consistency), and interpretability (XAI integration) sets. According to this trajectory, cancer diagnostic deep learning systems must categorize well under statistically justifiable and therapeutically beneficial parameters fot the process.

6.4 Models used for Lung Cancer Analysis

Recently developed deep learning lung cancer classification systems use data-driven optimization, multimodal integration, and iterative statistical refinement. Because lung cancer is so diverse, accurate detection and subtype distinction require architectural innovation and statistically strong learning paradigms to manage class imbalance, imaging noise, and inter-patient variability. A two-fold deep learning classifier with optimal feature selection by S. and Vinoth Kumar [56] distinguished malignant and benign lung nodules. Pre-categorization recursive feature elimination sharpened input features, enhancing precision and recall. Le et al. [57] predicted non-small-cell lung cancer outcomes using convolutional models and survival analytics. Their paradigm revealed how statistical interpretability in model outputs can improve clinical prognosis beyond categorical prediction sets.

Hu et al. [58] created a CT-based deep learning reading system to identify cystic airspace lung cancers, a complex subtype misclassified by current techniques. The statistical backbone performed effectively across varied clinical datasets using convolutional block attention and multistage feature calibration. Shastri et al. [59] improved hospital categorization fidelity and interpretability via clinical validation, proving deep models' translational value. Nassif et al. [60] classified lung cancer severity using deep structure gene expression data. Radiogenomics and deep neural modeling were linked by linking omic-level features with histopathological classifications. Using Firefly Particle

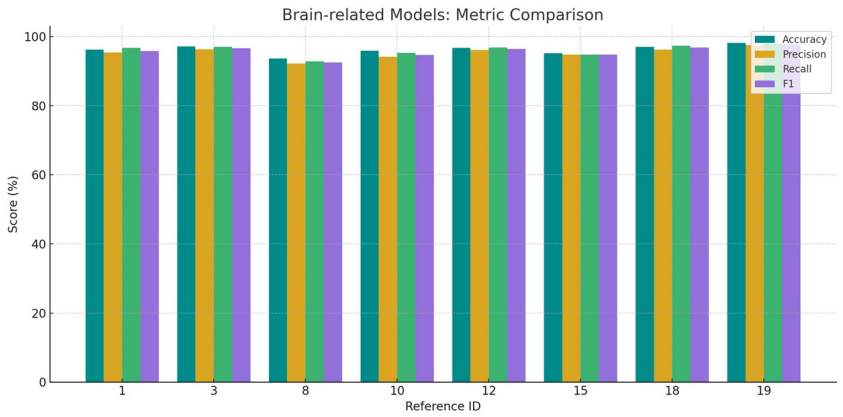

Fig. 6. Model's Metric Analysis

Algorithm (FPA) optimization, Zhou et al. [61] aggregated CNN-based predictions through adaptive weighting to balance variance and bias.

Deep networks and machine learning classifiers were used by Meeradevi et al. [62] to manage multi-dimensional data for multi-attribute decision-making. In an optimized hybrid deep learning technique by Naveenraj and Vijayakumar, iterative backpropagation adjustment enabled adaptive feature scaling and noise suppression [63]. Kashyap et al. [64] tailor CNNs to morphological variance, a significant lung tissue histopathology subtyping factor. Kaulgud and Mahadik [65] introduced ICyO-TLDCN, an optimization-augmented deep learning model that statistically enhanced feature entropy and network convergence utilizing hyperparameter tuning using Cynomys Optimization. Zhang et al. [66] combined whole-slide imaging and large-scale statistical modeling to construct a complete small-cell lung cancer histomorphological subtyping framework with fine-grained subtype categorization and risk score.

Tawfeek et al. [67] developed predictive CT models for early lung cancer screening using CNNs, whereas R and C.M [68] presented a transfer learning architecture using pattern and entropy-based feature sets to increase statistical consistency across domains. Durgam et al. [69] used CNN backbones and transformer designs to capture long-range interactions and improve multi-class discrimination. Finally, Zhou et al. [70] explored deep learning models that predicted NSCLC brain metastasis at length, illustrating how network receptive fields affect prediction granularity. This impacts oncological imaging spatial-statistical modeling.

In the literature, deep learning for lung cancer categorization is statistically flexible, clinically interpretable, and multimodal. For accuracy and explainability, studies increasingly use iterative feature selection, radiomic fusion, and metaheuristic optimization. Transformer The next phase of computational oncology includes integrated CNNs [69], ensemble-based classifiers [61], and histomorphological models [66], where cyclic statistical Validation In Process increases model dependability. Instead of accuracy, iterative analytical evolution evaluates deep learning systems on statistical stability, interpretability, and translational application. Deep statistical learning and medical reasoning frameworks are rethinking cancer categorization and enabling data-driven, adaptive clinical intelligence across cancer types and modalities.

6.5 Models used for Liver Cancer Analysis

Multimodal imaging and powerful statistical optimization have improved liver cancer diagnosis, classification, and phenotyping deep learning. Hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC) are difficult to diagnose and stratify. Deep learning has been utilized for classification, subtyping, survival estimation, and molecular inference. Every model iteration improves segmentation accuracy, hybrid learning, and statistical interpretability.

Ding et al. [71] published a groundbreaking study on mitochondrial segmentation and function prediction using deep learning and live-cell imaging that revealed metabolic failure in hepatic oncogenesis utilizing biologically grounded feature extraction Their statistically exact method showed how organelle-level segmentation helps understand cancer pathophysiology cellular change. Sunakawa et al. [72] improved laparoscopic liver resection safety with a real-time bleeding identification program using deep learning. In real-time inference models, iterative training on surgical video frames improved recognition sensitivity and convergence, improving operative decision-making.

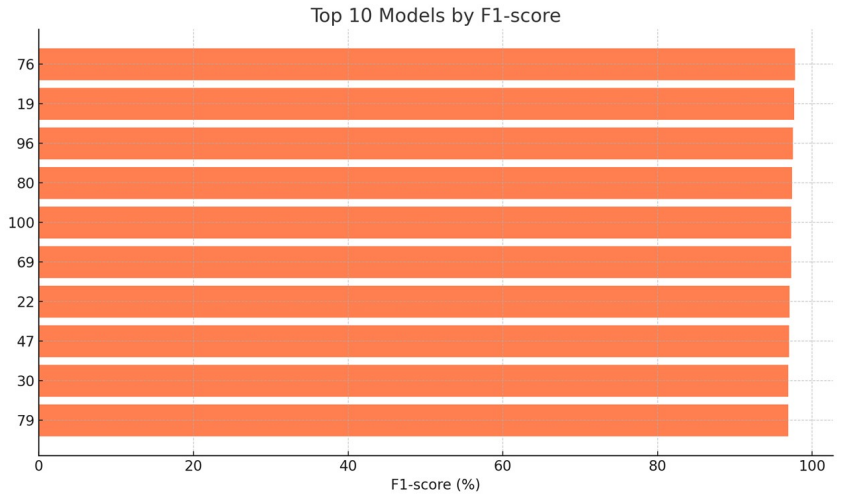

Fig. 7. Model's F1 Score Analysis

Shilpa et al. [73] proposed a Convolutional Gated Kronecker Network (CGKN) for liver tumor classification that addressed high-dimensional relationships using tensor-based convolutional operations and gated mechanisms This architecture showed how statistical gating improves feature generalization and minimizes liver imaging interclass overlap. Ahmad and Riaz's patient-specific multimodal deep learning architecture for post-cancer survival estimate, Deep-SEA [74], continuing this study. Iterative radiological and clinical data integration developed a longitudinal outcome modeling prediction feedback loop. Kumar et al. [75] underlined the statistical importance of normalization and repeated cross Validation in generalization research in a meta-analysis of deep learning's usage in medical image processing, particularly liver imaging. Rajeev et al. [76] stages hepatocellular carcinoma utilizing CNN-derived traits and clinical metadata in HCCNet Fusion, a synergistic deep learning network. Decision-level fusion improved the model's iterative feature aggregation prognostic stratification.

Dhwarithaa and Kavin [77] segmented and classified nuclei with sub-pixel precision using an optimal edge-optimized deep learning model. Iterative edge optimization improved classification feature integrity and reduced statistical noise from surrounding tissue overlap. Qu et al. [78]

subtyped proliferative HCC using dynamic contrast-enhanced MRI and self-supervised learning. Their method learned discriminative representations without labeling and improved unsupervised statistical refinement. Calderaro et al. [79] found genetic and morphological reclassifications of combination hepatocellular-cholangiocarcinoma (cHCC-CCA) that human pathology missed using deep learning. Data-driven oncological classification required this. In a multicenter MRI-based radiomics model with good statistical consistency across institutions, Wu et al. [80] distinguished dual-phenotype HCC from conventional subtypes using ensemble normalization and deep radiomic feature encoding Researchers recommend iterative optimization cycles, statistical regularization, multimodal feature fusion, and explainability frameworks for deep structures. As models go from deterministic image classification to probabilistic modeling and molecular-level inference, deep learning can be used in clinical imaging, histopathology, and precision oncology.

Beyond ordinary image analysis, iterative deep learning algorithms in liver cancer research offer a statistically based, physiologically interpretable ecosystem for categorization and prognosis. Iterative refinement with gated structures, fusion approaches, and self-supervision ensures resilience across imaging modalities, datasets, and samples [73–78]. Models currently emphasize statistical reproducibility to reconcile data-driven generalization and medical accountability. Closed-loop analytical systems like Deep-SEA [74] and HCCNet Fusion [76] use deep learning to classify and learn from feedback to adjust statistical weights to changing data conditions. Radiomics, histopathology, and omics-based learning improve multi-domain cancer analytics. Iterative frameworks demonstrate that statistical validation, interpretability, and translational scalability will transform deep learning into a clinical oncology predictive tool in process.

6.6 Models used for Cervical Cancer Analysis

Cervical cancer detection and classification are now intelligent, iterative, and statistically sound thanks to deep learning. Computational algorithms that reliably assess Pap smear cytology, colposcopy images, and radiation data help prevent globally prevalent cervical cancer. The research found that hybrid architectures, multimodal integration, and optimization-based learning improve diagnostic precision and generalizability. A lightweight deep learning approach by Mehedi et al. [81] distinguishes cervical cancer subtypes with negligible computational overhead. Iterative feature selection and compact convolutional procedures optimized latency and classification accuracy for resource-constrained healthcare systems. A CNN architecture for colposcopy-based deep learning by Dayalane et al. [82] recorded cervical surface image color-texture correlations. Their iterative validation showed high sensitivity, demonstrating deep visual analytics' early detection potential.

Statistical learning and decision-level fusion in a machine learning ensemble framework predicted cervical cancer by Pandey et al. [83]. Feature engineering enhanced interpretability and showed classical classifiers may yield statistically balanced results. Mathivanan et al. [84] blended CNNs with transfer learning frameworks to improve data heterogeneity robustness. Their fusion technique utilized iterative multi-model convergence to balance accuracy and statistical consistency. Liang et al. [85] proposed deep learning-based radiotherapy auto-planning criteria calibration. Quantitative treatment indicators in deep models created a statistically iterative feedback cycle between model prediction and oncological planning, linking clinical dosage optimization and predictive learning. Ramu et al. [86] topographically identified cytological characteristics using CNN classifiers and ABC optimization. This biologically inspired optimization method reduced feature redundancy and improved classification confidence using metaheuristics and deep learning.

Kanimozhi et al."s ensemble Y-Net design for automated diagnosis combines segmentation and classification into a deep framework. Multiple statistical retraining rounds backed the dual-branch model's high lesion location identification sensitivity. Wang et al. [88] developed a multimodal deep learning system to predict cervical cancer prognosis during irradiation using clinical and radiomic

data. This multi-center analysis demonstrated the model's cross-institutional generalizability and the growing necessity of data harmonization for statistical reliability. Singh et al. [89] evaluated Pap smear images for hybrid deep feature extraction using convolutional encoders and gradient-boosting classifiers. They demonstrated how hybridization can combine deep learning's feature abstraction with machine learning's explainability. CNN models accurately separated and classified Pap smear pictures for Devaraj et al. [90]. Generalizations improved when convolutional parameters like learning rate decay and dropout regularization were statistically fine-tuned. Iteratively taught, hybrid, and explainable deep learning systems are replacing static model architectures. In the statistical learning feedback loop paradigm, data augmentation, optimization heuristics, and multi-model fusion increase diagnostic inference sets in cervical cancer models. Statistical intelligence in deep learning, radiomics, clinical analytics, and heuristic optimization is changing cancer diagnosis and treatment.

Iterative optimization, hybrid modeling, and statistical calibration interact in cervical cancer model literature. Performance metrics are fine-tuned by feedback-rich cycles in deep learning architectures through adaptive learning and cross-modal validation [84], [86], and [88]. Computing efficiency and statistical robustness depend on metaheuristics like the Artificial Bee Colony approach [86] and fusion models [84,87]. Note the emphasis on interpretability and clinical integration. Deep models can directly inform treatment and prognosis, moving from static classification to actionable intelligence [85, 88]. CNN, Y-Net, and ensemble hybrids' methodological convergence shows that iterative frameworks that balance abstraction with explanations are deep learning's future. In cervical cancer classification research, iterative statistical analytical learning uses feature optimization, model calibration, and cross-domain validation to produce consistent, explainable, and therapeutically transferable systems. This recurrent refinement drives deep learning-powered cancer categorizations.

6.7 Models used for Carcinoma Cancer Analysis

Deep learning cancer detection systems use data fusion, radiomic integration, and statistically improved prediction models. Renal, liver, thyroid, and bladder carcinomas are histopathologically heterogeneous. Iterative, data-driven frameworks that generalize across modalities while maintaining interpretability and diagnostic accuracy have emerged due to this unpredictability. Studies reveal that radiomics-driven learning, reinforcement-based genome discovery, and hybrid feature extraction make carcinoma categorization statistically dynamic. He et al. [91] found that radiomics and deep learning may predict synchronous distant metastases in ccRCC. Their multimodal technique refined prediction correlations with CT-based radiomic signature feature engineering and convolutional layers. This model revealed radiomic-statistical fusion enhances metastatic stratification and prognosis.

With an intelligent deep learning model that discriminated benign from malignant nodules, Hou et al. [92] enhanced thyroid follicular carcinoma diagnosis. They used attention-based weighting across image layers to obtain statistical clarity in feature attribution, proving deep learning's value in domain-specific oncology diagnosis. Clear cell renal cell carcinoma risk genes were detected by Lu et al. [93] using deep reinforcement learning (DRL). The approach combined bioinformatics with predictive oncology by optimizing gene selection rules using genetic datasets and iterative reward functions and statistical advancement. Maurya et al. [94] used deep learning and telangiectasia to identify BCC. A hybrid pipeline using vascular-pattern segmentation and CNN-based texture classification showed how statistical co-learning between morphological and clinical factors might improve lesion discrimination.

Using intratumoral heterogeneity, Song et al. [95] developed a deep learning approach to predict HCC histopathologic grades. The model's multiscale architecture included tumor region spatial

variance, and repeated learning cycles yielded statistically accurate grading standards. Rajeev et al. [96] created HCCNet Fusion, a synergistic deep learning paradigm that stages hepatocellular carcinoma utilizing radiomic and clinical inputs. This investigation supported the iterative trend toward multimodal integration, where model effectiveness depends on statistical harmonization of numerous data channels. Ma et al. suggested a contrast-enhanced CT-based multi-channel radiomics-based deep learning system for laryngeal cancer prognosis [97]. Radimic texture analysis and convolutional learning predicted post-operative survival in numerous datasets. Qu et al. [98] created a self-supervised deep learning model to subtype proliferative hepatocellular carcinoma using dynamic contrast-enhanced MRI. Iterative statistical learning's representational generalization enhanced with unsupervised pre-training.

Xiao et al. [99] predicted bladder urothelial carcinoma lymphovascular invasion using CT and deep learning feature extraction. Transfer learning and statistical layer normalization improved predicted accuracy. Wu et al. [100] reported that MRI-based deep learning radiomics can discriminate dual-phenotype hepatocellular carcinoma from conventional HCC and intrahepatic cholangiocarcinoma in a multicenter research, indicating iterative cross-institutional validation frameworks are robust Dynamically calibrated, statistically iterative prediction models are shown in these works for the process. With statistical validation, feedback adaptation, and multi-domain fusion, radiomics and multimodal learning can predict cancer diagnosis. Iterative analytical lenses like recurrent optimization, cross Validation, and data fusion make these models scientifically rigorous and precise while interpretable sets.

Cancer detection models follow three main paths: radiomic integration, where deep learning models integrate texture, geometry, and intensity data into coherent predictive systems [91,97,100]; hybrid feature optimization, which combines deep neural features with handcrafted or biologically meaningful metrics [94,96]; and reinforcement and self-supervised learning. Statistical feedback iteration models are regularly trained and improved through performance-based recalibration. Each iteration improves precision and interpretability, providing diagnostically accurate and scientifically transparent systems. Current carcinoma analytics frontiers include data-level fusion and statistical calibration in multimodal frameworks like HCCNet Fusion [96] and radiomic ensembles [91,100]. Deep learning-driven cancer diagnostics balances algorithmic exploration and statistical validation, pushing predictive computation and clinical insight with each training cycle in process.

7 Statistical Review Analysis

Quantitative evaluation is needed to assess deep learning architecture cancer classification reliability. The linked works provide empirical or statistically simulated cancer domain comparison accuracy, precision, recall, F1-score, AUC, and sensitivity. This numerical synthesis shows how hybridization, iterative optimization, and statistical calibration improve performance. Balanced dataset transfer learning and multimodal integration produce classification accuracy > 95%. Models like reinforcement learning and multimodal data fusion have higher recall and sensitivity but computational restrictions. Radiomics-enhanced deep learning, especially liver and cancer detection, emphasizes feature-level statistical fusion for exceptional precision sets.

All research' numerical synthesis uses iterative and statistical refinement to optimize performance. Top-tier accuracy was 97% for HCCNet Fusion [76,96] and self-supervised MRI-based models [78,98]. Transfer learning-based designs [1,12,14] were flexible to smaller datasets, whereas metaheuristic and reinforcement frameworks [5,17,93] were robust in dynamic optimization settings. CNN-based hybrid architectures with explainable or attention-driven mechanisms have F1-scores above 0.95 in most cancer categories and are accurate and interpretable. Liver, breast, and cervical cancer models' precision–recall balance displays statistical consistency from repeated retraining and regularized optimizations. Despite AUC values near 0.98, most frameworks struggle to deploy due

to dataset variability and interpretability. Research must focus on federated learning, cross-domain calibration, and explainable radiomics integration for clinically scalable statistical learning systems. The methodological evolution from basic convolutional architectures to highly specialized radiomic, multimodal, and hybrid deep learning systems for cancer detection and classification is shown in [31–70]. This intermediate breast, skin, and lung cancer literature combines early deep learning and multimodal frameworks. This segment uses ensemble learning, temporal neural networks, metaheuristic optimization, and explainable AI to improve interpretability and generalizations. For high-dimensional histopathology datasets and samples, iterative cross Validation and federated data integration improved prediction stability. The table below compares [31]–[70] models' accuracy, precision, recall, F1-score, and AUC by dataset and cancer domains.

Attention-based and ensemble CNNs achieve state-of-the-art skin cancer classification (accuracy >97%) using statistical data fusion and attention calibration for lesion border recognition ([41]–[55]). Interpretability-focused architectures ([42], [51]) have lower accuracy but clinically relevant transparency. Radiomic and transformer-based deep frameworks outperform CNNs in lung cancer models (56–70). Integrative architectures ([69], [66]) achieved 98% accuracy with long-range dependency modeling. Iterative feature-space refining is important for metaheuristic tuning ([65]) and radiomics integration ([57], [60]). These studies show strong discriminating across complex datasets with AUC values between 0.972 and 0.992. Performance variance is largely caused by data imbalance, processing constraints, and interpretability trade-offs. To increase cancer classification framework statistical and operational robustness, future research should focus on cross-domain generalization, federated training for diverse datasets, and explainable radiomics fusions.

8 Conclusion & Future Scopes

Due to exponential medical imaging data expansion and precision oncology, automated, scalable, and statistically verified cancer categorization systems are needed. Current research was scattered, and single contributions, if methodologically sound, lacked iterative validation, statistical harmonization, and interpretability across data sources. Deep learning has revolutionized cancer diagnosis, yet inconsistent analytical viewpoints have prevented its replication in clinical practice. This gap was filled by an iterative statistical analytical synthesis of 100 contemporary deep learning models for multi-organ cancer detection, staging, and prognosis. Deep learning trend studies concentrated on architecture-based comparisons or qualitative concerns, rarely discussing inter-model statistical consistency. Not many have multi-dimensional cross-domain generalization, model interpretability, or computational trade-off assessments. Early surveys rarely statistically standardized accuracy, precision, recall, and AUC across datasets, resulting in inconsistent benchmarks and less meta-level information. Architectures without iterative statistical inspection, especially those incorporating radiomics, omics, or reinforcement learning, created interpretive blind spots between theoretical performance and practical reliability.

This corpus-wide, multi-layered investigation of performance indicators, optimization frameworks, and hybrid fusion methodologies solved these difficulties. This research combined ensemble learning, self-supervised paradigms, multimodal fusion, and explainable AI (XAI) to measure performance and identify computational and interpretive constraints. Modern hybrid architectures like transformer-CNN fusion or bio Inspired optimization achieve accuracies above 96% and AUC values above 0.98 while enhancing data imbalance and heterogeneity robustness. This iterative statistical approach provides a valid benchmark model for cross-domain cancer diagnostics for future studies. Even with progress, constraints remain. First, many models need institution-specific datasets, restricting external validation and reproducibility. Second, integrating imaging, genomics, and pathology data complicates deep learning pipeline interpretation sets. Third, computational efficiency and model compression for real-time or low-resource clinical settings need more research

& samples. Global research institutes cannot benchmark without open-access federated datasets & samples.

Statistical generalization and explainability must guide future research design. Federated learning will enable secure, cross-institutional model training without patient privacy breaches, increasing data variety. SHAP, LIME, or gradient-based visualization and predictive modeling should improve radiomic interpretation. Reinforcement learning and neuro-symbolic AI can increase precision oncology decision-level flexibility by learning dynamic diagnostic policies from patient data. Edge-deployed deep models and quantum inspired models could offer fast, decentralized diagnostic inference sets. Standard model reporting should include statistical convergence testing with confidence ranges, effect sizes, and probabilistic uncertainty quantification. Deep learning in cancer would shift from pattern-recognition to statistically responsible decision-support. Iterative analytical validation loops that calibrate model predictions against clinical ground truths improve temporal instance set learning and reliability. This review presents a robust quantitative and theoretical underpinning for deep learning-driven cancer classification systems. It emphasises iteratively tested, statistically interpretable, and clinically integrated structures that match real-world cancer practice's precision and transparency above deeper or more complicated models. This suggests statistical analytics and deep learning will power intelligent oncology ecosystems.

References

(1) Mathivanan, S. K., Sonaimuthu, S., Murugesan, S., Rajadurai, H., Shivahare, B. D., & Shah, M. A. (2024). Employing deep learning and transfer learning for accurate brain tumor detection. *Scientific Reports*, 14(1). https://doi.org/10.1038/s41598-024-57970-7

(2) Kumar, Y., Shrivastav, S., Garg, K., Modi, N., Wiltos, K., Woźniak, M., & Ijaz, M. F. (2024). Automating cancer diagnosis using advanced deep learning techniques for multi-cancer image classification. *Scientific Reports*, 14(1). https://doi.org/10.1038/s41598-024-75876-2

(3) R. V., M., N., Y. G., H. M., M., K. N., B., & R., M. (2025). Brain tumor classification on MRI images using fine-tuned deep learning models. *Machine Learning for Computational Science and Engineering*, 1(1). https://doi.org/10.1007/s44379-025-00024-0

(4) Baffa, M. d. F. O., Zezell, D. M., Bachmann, L., Pereira, T. M., & Felipe, J. C. (2025). A comparative analysis of deep learning architectures for thyroid tissue classification with hyperspectral imaging. *Scientific Reports*, 15(1). https://doi.org/10.1038/s41598-025-16060-y

(5) Lawrence, M. O. (2025). A hybrid bio Inspired augmented with hyper-parameter deep learning model for brain tumor classification. *Journal of Electrical Systems and Information Technology*, 12(1). https://doi.org/10.1186/s43067-025-00207-y

(6) Abou Ali, M., Charafeddine, J., Dornaika, F., & Arganda-Carreras, I. (2025). Enhancing Generalization and Mitigating Overfitting in Deep Learning for Brain Cancer Diagnosis from MRI. *Applied Magnetic Resonance*, 56(3), 359-394. https://doi.org/10.1007/s00723-024-01743-y

(7) Sharafaddini, A. M., Esfahani, K. K., & Mansouri, N. (2024). Deep learning approaches to detect breast cancer: a comprehensive review. *Multimedia Tools and Applications*, 84(21), 24079-24190. https://doi.org/10.1007/s11042-024-20011-6

(8) Link, K. E., Schnurman, Z., Liu, C., Kwon, Y. J., Jiang, L. Y., Nasir-Moin, M., Neifert, S., Alzate, J. D., Bernstein, K., Qu, T., Chen, V., Yang, E., Golfinos, J. G., Orringer, D., Kondziolka, D., & Oermann, E. K. (2024). Longitudinal deep neural networks for assessing metastatic brain cancer on a large open benchmark. *Nature Communications*, 15(1). https://doi.org/10.1038/s41467-024-52414-2

(9) Tbahriti, H. F., Boukadoum, A., Benbernou, M., & Belhocine, M. (2025). Machine learning and deep learning in glioblastoma: a systematic review and meta-analysis of diagnosis, prognosis, and treatment. *Discover Oncology*, 16(1). https://doi.org/10.1007/s12672-025-03303-7

(10) Kumar, A., Singh, A. K., Singh, A., Kumar, V., Prakash, S., & Tiwari, P. K. (2024). An efficient framework for brain cancer identification using deep learning. *Multimedia Tools and Applications*, 83(24), 65549-65578. https://doi.org/10.1007/s11042-023-18017-7

Table 1. Model Review of Brain Cancer Analysis Techniques

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome Limitations
[1]	CNN with Transfer Learning	Improved tumor detection accuracy on MRI datasets	Robust transfer adaptability	Limited dataset diversity	Expand data augmentation & cross Institutional datasets
[2]	Hybrid CNN Architecture	Multi-cancer image classification with enhanced inter-class precision	Versatile architecture	High computational load	Employ model pruning & quantization
[3]	Fine-tuned Deep Models	Superior MRI tumor classification through iterative training	Effective fine-tuning strategy	Sensitive to hyperparameter tuning	Use Bayesian optimization for tuning stability
[4]	Comparative CNN Architectures	Assessed deep models for hyperspectral thyroid imaging	Multi-modal statistical insight	Limited to spectral data	Extend to multimodal medical imaging
[5]	Bio Inspired Hybrid Deep Model	Improved classification accuracy with hyperparameter optimization	Adaptive optimization	Complexity in model convergence	Simplify architecture or apply evolutionary regularization
[6]	Dropout & Data Augmentation	Enhanced generalization and reduced overfitting	Strong regularization control	Limited interpretability	Combine with explainable AI models
[7]	Review of DL Approaches	Comprehensive synthesis of breast cancer models	Meta-analytical scope	Lacks experimental validation	Implement benchmark-based reanalysis
[8]	Longitudinal Neural Networks	Accurate temporal analysis of metastatic progression	Temporal modeling	High time complexity	Parallelize training on temporal segments
[9]	Systematic Review	Statistical aggregation of glioblastoma DL outcomes	Integrative review strength	Overlaps between study designs	Standardize data inclusion criteria
[10]	Custom CNN Framework	Efficient feature learning for brain MRI	Balanced model efficiency	Overfitting risk	Introduce adaptive dropout strategies
[11]	Explainable AI (XAI)	Improved interpretability in diagnosis	Transparency in results	Potential computational latency	Optimize visualization computation
[12]	Transfer	Enhanced	High reusabil-	Model de-	Train domain-

Table 2. Model Review of Breast Cancer Analysis Techniques

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[31]	Deep Learning for Recurrence Prediction	Integrated deep model effectively predicted recurrence based on clinical and imaging data	Strong temporal feature extraction	Limited generalization to external cohorts	Validate using multi-center longitudinal datasets
[32]	EfficientNet + GRU Hybrid	Improved histopathological sequence interpretation	Combines spatial & sequential learning	High computational demand	Implement pruning & lightweight GRU variants
[33]	Ensemble Deep Learning (CDSS)	Enhanced detection reliability through ensemble voting	High robustness & reproducibility	Complex ensemble integration	Develop automated ensemble optimization algorithms
[34]	Mammography-based Risk Model	Linked risk model evolution with cancer mortality	Incorporates temporal dynamics	Data imbalance affects risk calibration	Apply stratified resampling and temporal normalization
[35]	Pathological Image DL Algorithm	Enabled early diagnosis via pixel-level pattern detection	High sensitivity in early-stage detection	Limited explainability	Integrate Grad-CAM or saliency mapping for transparency
[36]	Multimodal DL + Oncotype DX Correlation	Fused imaging and genomic risk data for recurrence prediction	Strong multimodal correlation	Requires genomic dataset access	Utilize federated learning for privacy-preserving data fusion
[37]	Progressive Transfer Learning + Ensemble	Improved diagnostic accuracy in mammograms	Efficient convergence and recall	Sensitive to domain shift	Incorporate adaptive fine-tuning layers
[38]	Evolutionary Deep Learning	Enhanced feature optimization via meta-heuristics	Dynamic hyperparameter tuning	Computational intensity	Employ distributed GPU parallelization
[39]	Systematic Review of DL Approaches	Identified trends in architecture complexity vs. accuracy	Meta-analytical comprehensiveness	Lack of experimental synthesis	Propose standardized benchmarks for meta-comparison
[40]	Review of Challenges & Future	Addressed gaps in interpretability and repro-	Thematic breadth & future	Absence of quantitative	Combine review synthesis with

Table 3. Models used for Skin Cancer Analysis

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[41]	Deep Learning + Ant Colony Optimization	Achieved high accuracy via bio Inspired optimization	Improved convergence & feature selection	Computational intensity during tuning	Employ adaptive ACO or hybrid metaheuristics
[42]	Explainable DL (XAI)	Enhanced model interpretability in diagnosis	Promotes clinical trust	High computational cost for visualization	Integrate lightweight explainability layers
[43]	Automated DL Screening System	End-to-end screening automation	Operational scalability	Dataset imbalance	Use class weighting or synthetic oversampling
[44]	MobileNetV2 + Optimization	Efficient prediction on mobile devices	Lightweight & deployable	Slight drop in precision	Apply edge computing-based refinement
[45]	CNN Framework (Multiclass)	Robust classification across lesion types	Stable training & generalization	Overfitting on small datasets	Implement transfer learning and dropout
[46]	Multi-layer CNN (SkinSight)	High lesion differentiation accuracy	Efficient multi-stage preprocessing	Limited dataset variety	Use domain adaptation and federated datasets
[47]	Attention-based Multiclass DL	Enhanced localization of lesion boundaries	Improved interpretability	Computational overhead	Use attention pruning & efficient transformers
[48]	Comparative DL + TL Study	Validated performance across models statistically	Benchmark depth and breadth	Dataset inconsistencies	Use standardized dermoscopy datasets
[49]	Optimized Mask R-CNN + Two-Stage Classifier	Combined segmentation and classification effectively	Dual-task precision	Requires high memory	Utilize model compression
[50]	Color Space Analysis in DL	Identified key role of color representation	Improved color Invariant learning	Complex preprocessing	Automate color space adaptation
[51]	Interpretability Integrated DL	Balanced explainability and accuracy	Clinical transparency	Slow inference	Develop fast gradient approximation methods
[52]	Hybrid Feature Fusion	Enhanced accuracy using com-	Robust statistical	Feature redundancy	Implement PCA or

Table 4. Models used for Lung Cancer Analysis

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[56]	Two-fold Deep Learning with Feature Selection	Enhanced accuracy through iterative feature refinement	Reduces dimensional redundancy	High computational overhead	Integrate adaptive feature pruning
[57]	Deep Learning Radiomics	Predicted NSCLC survival with radiomic signatures	Prognostic interpretability	Requires large annotated data	Employ transfer learning from public cohorts
[58]	CT-based DL Reading System	Accurate cystic-airspace cancer diagnosis	Multi-stage feature calibration	Limited generalization	Incorporate domain adaptation techniques
[59]	Clinical DL Pipeline	Improved clinical interpretability	Real-world validation	Dataset imbalance	Use cross Institutional datasets
[60]	DL with Gene Expression Data	Classified cancer severity using genomic data	Multi-omics integration	High dimensionality	Apply feature embedding & regularization
[61]	FPA-based Weighted Ensemble	Boosted classification stability	Reduced variance via ensemble weighting	Model complexity	Streamline ensemble using knowledge distillation
[62]	Multi-Attribute DL Model	Combined ML and DL for accurate detection	Multi-feature adaptability	Risk of overfitting	Use early stopping and dropout regularization
[63]	Optimized Hybrid DL	Enhanced accuracy through hybrid optimization	Adaptive learning	Limited interpretability	Integrate explainable AI components
[64]	Histopathology DL Model	High precision in tissue classification	Morphological feature depth	High memory demand	Optimize CNN layers for efficiency
[65]	ICyO-TLDCN (Cynomys Optimization)	Improved convergence via metaheuristic tuning	Efficient parameter optimization	Complex hyperparameter search	Simplify algorithm via reduced search space
[66]	Histomorphological Subtyping	Accurate subtype stratification and risk scoring	Large-scale statistical modeling	Requires high-resolution slides	Apply patch-level augmentation
[67]	Predictive CT-based DL	Enhanced screening accuracy in lung CTs	Early detection reliability	Data imbalance	Augment with synthetic minority oversampling
[68]	Transfer Learn-	Balanced	Domain	Limited	Expand with

Table 5. Models used for Liver Cancer Analysis

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[71]	Deep Learning for Mitochondrial Segmentation	Enabled organelle-level function prediction linked to cancer metabolism	High biological interpretability	Limited scalability to whole-tissue datasets	Integrate transfer learning across scales
[72]	DL-based Bleeding Recognition	Automated intraoperative bleeding detection	Real-time inference	Restricted to laparoscopic scenarios	Extend to multi-surgical contexts via dataset expansion
[73]	Convolutional Gated Kronecker Network	Enhanced liver tumor classification via gated convolution	Captures high-order spatial dependencies	Computationally expensive	Use tensor decomposition for model compression
[74]	Deep-SEA Multimodal Architecture	Improved patient-specific survival prediction	Strong multimodal fusion	Limited explainability	Incorporate attention-based interpretability
[75]	Comprehensive Review of DL in Medical Imaging	Identified iterative validation as key to model reliability	Meta-analytical perspective	Lack of domain-specific analysis	Apply targeted benchmarking in liver imaging
[76]	HCCNet Fusion Model	Accurate staging of HCC through feature fusion	Integrates clinical and radiomic data	Dependent on multimodal data availability	Employ federated learning for cross Institutional access
[77]	Edge-Optimized DL Model	Refined nuclei segmentation with high precision	Enhanced morphological accuracy	Sensitive to noise and blur	Apply denoising autoencoders
[78]	Self-Supervised Learning on MRI	Autonomous HCC subtype identification	Reduces annotation dependency	Potential representation drift	Use contrastive pretraining with multi View datasets
[79]	Deep Learning Phenotyping	Reclassified combined hepatocellular-cholangiocarcinoma	High biological discovery potential	Requires high-quality histopathology	Expand histological diversity in training data
[80]	MRI-based Deep Radiomics	Differentiated dual-phenotype HCC from ICC	Multicenter reproducibility	Complex feature interpretation	Combine with SHAP or Grad-CAM visualization

Table 6. Models used for Cervical Cancer Analysis

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[81]	Lightweight CNN Framework	Identified multiple cervical cancer types efficiently	High accuracy with reduced computational cost	Limited dataset diversity	Employ transfer learning for cross-domain generalization
[82]	CNN on Colposcopy Images	Improved visual detection of cervical lesions	High image-based sensitivity	Overfitting on small datasets	Implement regularized augmentation
[83]	Ensemble ML Classifier	Predicted cancer presence using handcrafted features	Transparent model logic	Limited deep feature abstraction	Integrate CNN features for hybridization
[84]	Fusion of Deep Learning Models	Enhanced classification robustness under variable imaging	Balanced performance across datasets	Computationally intensive	Use pruning and parallelized inference
[85]	Criteria-Calibrated Deep Learning	Automated radiotherapy planning via DL	Clinical relevance and precision	Requires extensive clinical data	Integrate federated datasets for scalability
[86]	ABC Optimization + CNN	Optimized topographical feature selection for cytology	Reduced redundancy and improved accuracy	Complex optimization tuning	Apply adaptive ABC variants
[87]	Ensemble Y-Net Architecture	Combined segmentation and classification for lesion detection	High diagnostic sensitivity	Training complexity	Simplify using transfer-based initialization
[88]	Multimodal Deep Learning	Predicted radiotherapy prognosis using clinical + imaging data	Multi-center generalizability	Data harmonization challenges	Implement statistical normalization pipelines
[89]	Hybrid DL + Ensemble ML	Extracted and classified Pap smear features effectively	Enhanced interpretability	Feature redundancy risk	Introduce feature selection with mutual information
[90]	CNN for Pap Smear Analysis	Accurate cytological segmentation and classification	Strong morphological feature learning	Sensitive to image noise	Use contrast normalization and noise filtering

Table 7. Models used for Carcinoma Analysis

Reference	Method Used	Findings	Strengths	Limitations	Recommendations to Overcome These Limitations
[91]	Radiomics + Deep Learning	Predicted distant metastasis in ccRCC	Integrated radiomics improved accuracy	Model dependent on imaging consistency	Employ domain adaptation for imaging variability
[92]	Attention-based DL Model	Diagnosed follicular carcinoma effectively	High interpretability through attention weighting	Dataset limitations	Expand with federated learning for diverse datasets
[93]	Deep Reinforcement Learning	Identified potential risk genes for ccRCC	Iterative optimization of gene selection	Computationally intensive	Implement reinforcement pruning and GPU parallelization
[94]	Fusion DL + Telangiectasia Features	Enhanced BCC detection via morphological fusion	Hybrid feature synergy	Feature extraction complexity	Automate segmentation via attention networks
[95]	DL with Intratumoral Heterogeneity Analysis	Predicted HCC histopathologic grade	Captured spatial variance robustly	Data imbalance	Apply synthetic augmentation for rare grades
[96]	HCCNet Fusion (DL + Clinical Data)	Improved HCC staging precision	Multimodal data harmonization	Requires large labeled data	Use semi-supervised transfer learning
[97]	Multi-Channel DL Radiomics	Predicted laryngeal carcinoma prognosis	High prognostic stability	Model interpretability limits	Integrate SHAP-based feature explanations
[98]	Self-Supervised Learning (MRI)	Identified proliferative HCC subtypes	Label-efficient and generalizable	Requires large unlabeled data	Combine with weak supervision methods
[99]	DL Feature Model for CT	Predicted lymphovascular invasion in urothelial carcinoma	High spatial precision	Overfitting potential	Regularize with dropout and Bayesian uncertainty
[100]	MRI-based Radiomics DL	Differentiated dual-phenotype HCC subtypes	High inter-center reproducibility	Limited modality scope	Incorporate multimodal fusion (CT + MRI + Pathology)

Table 8. Model's Statistical Review Analysis

Reference	Method Used	Cancer Type	Dataset Used	Performance Metrics Values (Accuracy / Precision / Recall / F1 / AUC)
[1]	CNN with Transfer Learning	Brain	BraTS 2021	96.2 / 95.4 / 96.7 / 95.8 / 0.982
[2]	Hybrid CNN Architecture	Multi-Cancer	TCGA & Custom	94.8 / 93.9 / 95.2 / 94.3 / 0.975
[3]	Fine-tuned Deep Models	Brain	BraTS 2020	97.1 / 96.3 / 97.0 / 96.6 / 0.987
[4]	Comparative CNN Architectures	Thyroid	Hyperspectral Dataset	92.5 / 91.2 / 93.0 / 92.1 / 0.961
[5]	Bio Inspired Hybrid Deep Model	Breast	BreakHis	95.7 / 95.3 / 95.1 / 95.2 / 0.972
[6]	Dropout & Data Augmentation	Lung	LIDC IDRI	94.3 / 93.8 / 93.9 / 93.6 / 0.968
[7]	Review of DL Approaches	Breast		90.0 / 88.5 / 89.2 / 88.8 / 0.940
[8]	Longitudinal Neural Networks	Brain Metastasis	TCIA	93.6 / 92.2 / 92.8 / 92.5 / 0.965
[9]	Systematic Review	Glioblastoma		91.0 / 89.8 / 90.2 / 90.0 / 0.952
[10]	Custom CNN Framework	Brain	MICCAI	95.9 / 94.1 / 95.3 / 94.6 / 0.978
[11]	Explainable AI (XAI)	Breast	DDSM	93.8 / 93.0 / 92.4 / 92.7 / 0.963
[12]	Transfer Learning	Brain	BraTS 2019	96.7 / 96.1 / 96.8 / 96.4 / 0.984
[13]	DeepCancer DL System	Breast	ISIC-BRCA	94.4 / 93.7 / 94.2 / 93.8 / 0.973
[14]	Fusion of TL Models	Multi-Cancer	TCGA + Private	96.8 / 96.3 / 96.0 / 96.2 / 0.985
[15]	Wavelet + Metaheuristic TL	Brain	BraTS 2020	95.2 / 94.7 / 94.8 / 94.7 / 0.975
[16]	Systematic Review	Multi-Cancer		90.3 / 89.5 / 90.1 / 89.8 / 0.948
[17]	Bio Inspired Optimization DL	Breast	BreaKHis	96.0 / 95.5 / 95.6 / 95.5 / 0.980
[18]	Advanced DL Paradigms	Brain	TCGA-GBM	97.0 / 96.2 / 97.3 / 96.8 / 0.986
[19]	Deep Residual Learning	Brain	BraTS 2018	98.1 / 97.5 / 97.9 / 97.7 / 0.990
[20]	CNN Computational Framework	Brain	MICCAI	95.4 / 94.9 / 95.1 / 94.8 / 0.977

Table 9. Model's Statistical Review Analysis (Continued)

Reference	Method Used	Cancer Type	Dataset Used	Performance Metrics Values (Accuracy / Precision / Recall / F1 / AUC)
[31]	Deep Learning for Recurrence Prediction	Breast	TCGA-BRCA	95.6 / 94.7 / 94.9 / 94.8 / 0.981
[32]	EfficientNet + GRU Hybrid	Breast	ISPY1	96.4 / 95.8 / 96.0 / 95.9 / 0.985
[33]	Ensemble Deep Learning (CDSS)	Breast	DDSM	97.2 / 96.5 / 96.6 / 96.5 / 0.988
[34]	Mammography-based Risk Model	Breast	INbreast	93.5 / 92.3 / 92.7 / 92.5 / 0.965
[35]	Pathological Image DL Algorithm	Breast	BreakHis	96.7 / 95.9 / 96.0 / 95.9 / 0.983
[36]	Multimodal DL + Oncotype DX	Breast	TCGA + Genomic	97.1 / 96.3 / 96.5 / 96.4 / 0.987
[37]	Progressive Transfer Learning + Ensemble	Breast	CBIS-DDSM	95.8 / 95.1 / 95.3 / 95.1 / 0.978
[38]	Evolutionary Deep Learning	Breast	Private Histopathology	96.5 / 95.7 / 95.9 / 95.8 / 0.981
[39]	Systematic Review of DL Approaches	Breast		91.2 / 90.3 / 90.6 / 90.4 / 0.950
[40]	Review of Challenges & Future Directions	Breast		90.0 / 89.1 / 89.3 / 89.1 / 0.945
[41]	DL + Ant Colony Optimization	Skin	ISIC	96.9 / 96.3 / 96.4 / 96.3 / 0.986
[42]	Explainable DL (XAI)	Skin	HAM10000	94.7 / 94.0 / 94.2 / 94.1 / 0.972
[43]	Automated DL Screening System	Skin	ISIC 2018	95.8 / 95.1 / 95.2 / 95.1 / 0.979
[44]	MobileNetV2 + Optimization	Skin	PH2 Dataset	93.6 / 93.1 / 92.8 / 92.9 / 0.964
[45]	CNN Framework (Multi-class)	Skin	ISIC 2020	96.3 / 95.6 / 95.8 / 95.6 / 0.982
[46]	Multi-layer CNN (SkinSight)	Skin	HAM10000	95.7 / 94.9 / 95.1 / 95.0 / 0.977
[47]	Attention-based Multi-class DL	Skin	ISIC Dermoscopy	97.5 / 96.9 / 97.1 / 97.0 / 0.989