Current Trends in Biotechnology and Pharmacy
Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4


# Integrating AI with Liquid Biopsy for Cancer Detection: A Review of Current Advances and Future Prospects

**Sandhya Sindhu, and Sitalakshmi Thyagarajan***

Department of Bioinformatics, SIMATS Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Thandalam, Chennai-602105, Tamil Nadu, India

*Corresponding author: sitalakshmithyagarajan.sse@saveetha.com

**Abstract**

Liquid biopsy is an innovative and non-invasive technique which helps to analyses the circulating tumour-derived materials, like circulating tumour DNA, Circulating Tumour Cells, and exosomes from the blood samples to detect and monitor cancer stages. Cancer causes of death worldwide, due to delayed diagnosis often results in poor treatment outcomes. This approach offers a faster, safer and alternative method for traditional tissue biopsies method. By understanding the Deep Learning and Machine Learning Algorithms, AI enhances the sensitivity and specificity of liquid biopsy technologies. This review highlights the recent advancements techniques where AI-driven models such as DeepCNA, DELFI have demonstrated remarkable accuracy in early cancer detection, tumour classification, treatment guidance and post-treatment monitoring. In spite of successes, the challenges remain unexplored in the field of data quality issues, ethical concerns, and regulatory hurdles. Innovations in this field also allow for at-home cancer screening and even pre-symptomatic cancer detection. As we stand at the intersection of computational science and clinical oncology, the synergy between AI and liquid biopsy represent one of the most promising frontiers in precision medicine. This review highlights the comprehensive overview of current technologies, challenges, and future directions, underscores the transformative potential of AI-powered liquid biopsy in the fight against cancer.

**Keywords**: Machine learning, Liquid Biopsy, Cancer detection, AI in medicine

## Introduction

Cancer diseases are defined as an abnormal cell growth which invade the tissues and other parts of the body. If the spread is not controlled, it can result in death. The early detection of cancer is crucial for several reasons which include the early detection significantly improves the chances of successful treatment and survival (1). Colorectal cancer (CRC), analysing and early detection have shown an effective in reducing the mortality rates. Similarly, prostate cancer, analysing and early detection are done by using Digital Rectal Examination (DRE) and serum prostate-specific antigen (PSA) determination were considered to be useful for patients with at least a 10-year life expectancy. In the case of pancreatic cancer, early detection is vital for improved survival rates. For breast cancer, early detection done through appropriate screening programs which made significant contributions to improving outcomes (2).

Early detection of disease leads to less aggressive treatments and reduced treatment morbidity. Detecting cancer early is important for enhancing patient outcomes, as it enables more effective treatments with fewer side effects and improved long-term survival rates. Nonetheless, traditional cancer diagnostic techniques like tissue biopsy, tumor biomarker detection, and instrumental examinations are effective while having limitations that prompted the search for less invasive and more comprehensive alternatives. Tissue biopsy, cytological examination and human papillomavirus testing for cervical cancer, which have limitations in sensitivity and reproducibility and also painful and often unsuitable for long-

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

term study of tumor dynamics (3). It also fails to analysis the whole genomic landscape of tumors, especially in complex cancers (Fig. 1).

The limitations of traditional cancer detection methods give rise to liquid biopsy (LB) which is non-invasive technique that analyses cancer-derived materials in biofluids mainly in blood (4). Since blood circulates throughout the body and contacts most tumors, LB primarily involves in blood sampling; however, saliva, pleural effusions, urine, and cerebrospinal fluid can also be analysed. This approach offers an enhanced sensitivity that allows for repeated sampling during treatment and more convenient and less invasive than tissue biopsy. This primarily involves in detection of Circulating Tumor Cells, circulating tumor DNA and other biomarkers like microRNAs and exosomes (5). The research studies highlight the potential of liquid biopsy in early detection of tumours.

The Artificial Intelligence and Machine learning play a vital role in analysing the accuracy and effectiveness of liquid biopsy for cancer detection. Artificial Intelligence and Machine learning model process the complex data from blood-based biomarkers namely circulating tumor DNA (ctDNA), RNA, CTCs (circulating tumor cells) and exosomes to identify the patterns and molecular changes that are associated with cancer (6, 7). Machine learning were trained on dataset to recognize the difference between healthy and cancerous profile thus enabling earlier and more precise diagnosis. When paired with liquid biopsy, AI also detect cancer even in its early stages, monitor disease progression, and help predict treatment. Research studies reveal that minimizing human error of biological data quickly, AI significantly improves the specificity of liquid biopsy tests (8).

The review highlights relationship of Artificial Intelligence and liquid biopsy in cancer diagnosis and also give a complete overview of how AI technologies were applied to analyze the liquid biopsy data for detection, classification, and analysing of cancer. This review also discusses the types of biomarkers used in liquid biopsies, current AI models and their diagnostic capabilities, and the

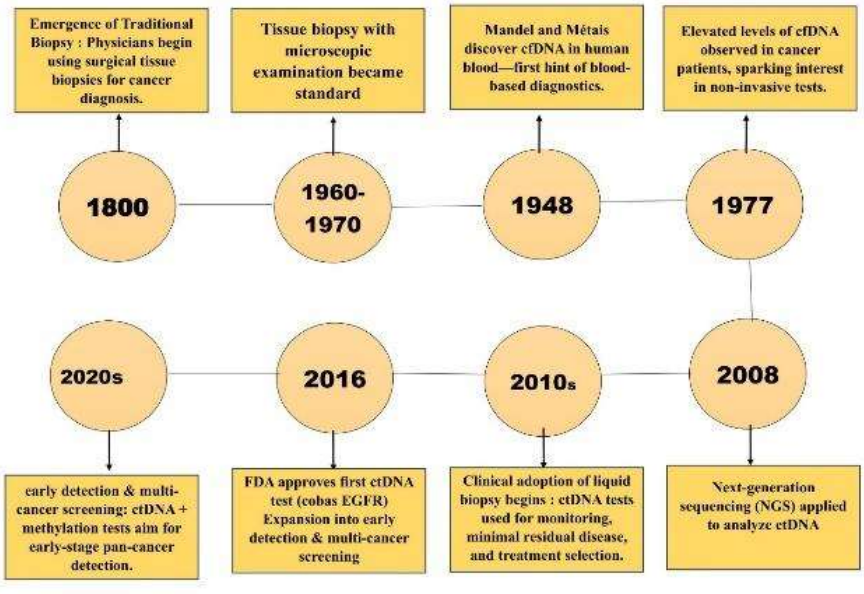


**Fig 1**: Schematic Representation for Time-line Cancer Dynamics

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

advantage and limitations of integrating AI into clinical workflows and also highlights recent research developments, real-world applications, and future prospects of combining AI with non-invasive liquid biopsy techniques to transform cancer diagnosis and personalized treatment strategies.

### *Introduction to Liquid Biopsy*

Tumour biopsy, an important standard for cancer subtyping, offers a brief view of the disease and can be quite challenging to acquire in many situations. Incomplete sampling and vague results in immune histochemical tumour analysis lead to inconsistencies in patient diagnoses, ultimately causing ineffective treatment. Standard monitoring makes the patient to bear the pain from repeated biopsy procedures throughout their therapy. To achieve this, advancements in cancer research and biomedical technology have led to the creation of innovative diagnostic tools that enhance traditional biopsies by utilizing sampling of bodily fluids. This capability to gather data on non-hematologic cancers is one of the most thrilling developments which has propelled significant progress (9).

The real-time insights of liquid biopsy in patient's status, enhance the use of cellular and molecular therapies which depends on regular assessments of vital biomarkers and considered to be byproducts or entities expelled from cells in the primary tumor and thought to significantly contribute to metastasis (10). The biomarkers in bodily fluids including blood and act as genetic messengers and offer a different pathway of cancer progression (Fig. 2).

This approach allows for continuous monitoring of disease progression and early detection of relapse, often before clinical symptoms appear, thereby improving patient prognosis and survival rates (11). Liquid Biopsies detects the minimal residual disease at levels not detectable by imaging technologies in therapy. As research advances, the combination of liquid biopsy in clinical practice gives a promising results in enhancing the effectiveness of cancer diagnostics and treatment (12).

Liquid biopsy testing follows a stepwise pre-analytical pipeline. First, a blood sample is drawn into special collection tubes that stabilize nucleic acids and prevent cell lysis. The sample is then processed and

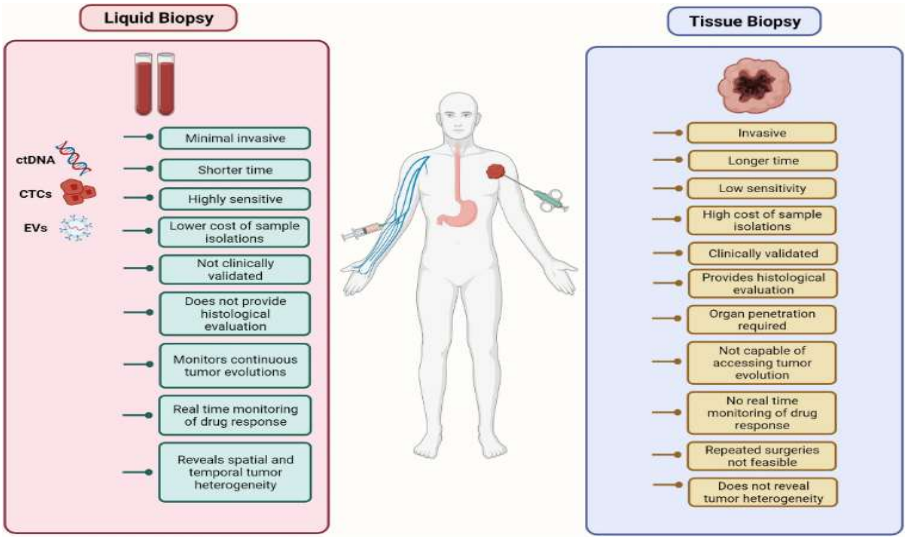


**Fig 2:** Comparison of traditional tissue biopsy and liquid biopsy. The schematic illustrates the advantages that liquid biopsies have gained over traditional invasive surgical methods Adapted from *SN Lone et al* (2022) (12).

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

typically centrifuged to separate plasma and the plasma is used for analyte extraction after plasma preservation, cell-free nucleic acids are isolated in parallel (cfDNA and cfRNA), quantified and subjected to quality control. Finally, the purified material is analyzed with digital PCR or next-generation sequencing to detect tumor-specific genomic or epigenetic alterations (13). Each step (collection, processing, extraction, and analysis) must be standardized to ensure the integrity of the low-abundance biomarkers.

**Introduction to Blood-Based Biomarkers**

Liquid biopsies targets on several key tumor-derived components in blood, mostly on circulating tumor cells and tumor-derived exosomes. ctDNA in the blood are released from tumor cells and CTCs are intact cancer cells which released from primary or metastatic tumors into the bloodstream; and exosomes are small membrane-bound vesicles of endosomal origin that are released into bodily fluids. Each of these analytes carries distinct molecular information about the tumor (e.g. ctDNA bears specific mutations, CTCs contain whole-cell genomic content, exosomes carry proteins and RNAs from their parent cells) (14, 15). These circulating biomarkers will be discussed in detail in the next section, examining their biology, isolation methods, and roles in cancer diagnostics.

***Circulating Tumor DNA* (ctDNA)**

Circulating tumor DNA and fragments of cell-free DNA are released by cancer cells into the bloodstream (Fig. 3). ctDNA undergoes apoptosis, necrosis, or active secretion (e.g. via exosomes), and carries tumor-specific genetic and epigenetic alterations (16). These fragments are typically short (on the order of 70–200 base pairs) but can span up to several kilobases. It is found in blood plasma and represents a small subcomponent of the total cell-free DNA which also seen in body fluids like blood without being enclosed in cells circulating in the body. Because ctDNA is derived from all tumor sites, it provides a composite ("liquid biopsy") view of tumor heterogeneity that exceeds the information from a single tissue biopsy (17). Notably, ctDNA is quickly cleared from the circulation (half-life ≈16 minutes to 2

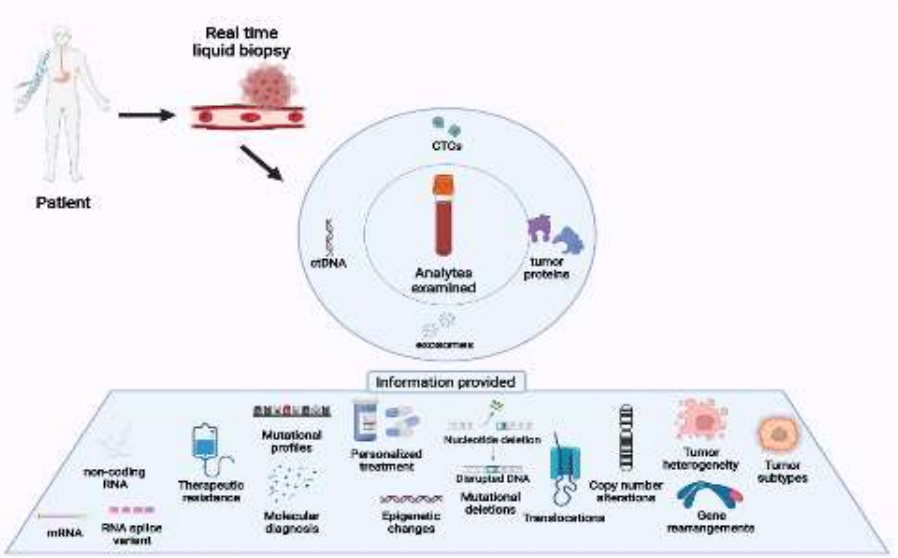


**Fig 3**: Real-time liquid biopsy identifies circulating tumor biomarkers for molecular profiling. Adapted from *SN Lone et al* (2022) (12).

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

hours, meaning its levels reflect real-time tumor dynamics. In healthy individuals, total plasma cfDNA is low (~0–100 ng/mL) but can exceed 1000 ng/mL in cancer patients; ctDNA includes a small fraction of this (often <1% variant allele frequency), which necessitates highly sensitive detection. Because ctDNA originates from tumor genomes, it carries somatic mutations (a cancer-specific DNA change that occurred in a tumor cell during a person's life), methylation changes (abnormal patterns of epigenetic modification), copy-number alterations (gains or losses of large DNA segments) and other cancer-specific features (18). This makes ctDNA a powerful tumor biomarker: detecting a known driver mutation or methylation signature in plasma directly indicates the presence of cancer cells. Thus, distinguishing true tumor-derived ctDNA from background cfDNA is critical (19). The ctDNA testing workflow comprises carefully controlled pre-analytical steps, sensitive molecular assays, and rigorous data analysis. Each phase is designed to preserve the fragile ctDNA fragments and maximize assay sensitivity. The main steps include:

***Blood Sample Collection and Plasma Preparation***

Blood is drawn into anticoagulant tube to prevent clotting and DNA release. Plasma (not serum) should be used, since clotting (as in serum) releases leukocyte DNA that dilutes ctDNA (20). Tubes must be mixed by gentle inversion (8–10 times) and kept at room temperature or refrigerated. If using EDTA tubes, plasma should be separated within 4-6 hours of collection to minimize genomic DNA contamination from lysed blood cells. During transport and handling, agitation and temperature fluctuations should be minimized to avoid hemolysis (21). Plasma is prepared by a two-step centrifugation: first a low-speed spin (e.g. 800-1,600g for 5 min) to pellet cells, and a high-speed spin (≈14,000-16,000g for 5 min) to eradicate residual debris. The cell-free plasma is collected in fresh tubes. Plasma should be processed immediately or stored at 4 °C short-term (≤3 h) and then frozen. For long-term storage, plasma aliquots are kept at –80 °C as rapid processing and cold storage inactivate nucleases that degrade ctDNA.

**ctDNA Extraction and Quantification**

cfDNA (including ctDNA) is extracted from the plasma using high-efficiency kits (e.g. silica-membrane or magnetic-bead based cfDNA isolation kits). ctDNA is a small tumor-derived component of total cfDNA, which also includes DNA from normal cells. Because ctDNA concentration is very low (often pg per mL of plasma), extraction yield is maximized by using large plasma volumes (several mL) or multiple tubes. The purified cfDNA is then quantified (e.g. by fluorometry) to estimate total cfDNA yield (22, 23).

***Mutation Detection Technologies in ctDNA Analysis***

Detection of tumor-derived mutations in ctDNA relies on two main approaches: PCR-based assays and next-generation sequencing (NGS). Data obtained from PCR-based assays and next-generation sequencing (NGS) are processed through specialized bioinformatic pipelines including alignment, error correction, and variant calling—AI models are employed to further interpret the results (24). To enhance the interpretation of this complex data, artificial intelligence (AI) and machine learning (ML) models are increasingly employed. These models integrate genomic, epigenomic, and clinical features to improves ctDNA detection, which aids in early cancer detection, and monitor treatment responses. The incorporation of AI into ctDNA analysis signifies an important improvement in precision oncology, allowing more accurate and personalized cancer diagnostics and prognostics (25).

***Circulating Tumor Cells (CTCs):***

Research study of tumour-derived mutations in ctDNA rely on two main approaches which include, PCR-based assays and next-generation sequencing (NGS). The data obtained from these are processed through specialised bioinformatic

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

pipelines, including alignment and error correction (26). The AI models were employed to interpret the result that enhance the interpretation of the complex data. These models integrate genomic, epigenomic, and clinical features to improve ctDNA detection that aid in early cancer detection and treatment responses (27).

Cancer metastasis are preliminary cause for death worldwide and remains one of the important challenges in curing cancer. Most of the patients with metastatic disease take systemic drugs for survival and improve signs, but their mortality rate is very high. Metastasis involves in intravasation, extravasation, migration and regeneration, in which cancer cells detach itself and invade the tissues using the bloodstream as a transport system.

First observed in the 19th century. CTCs are now understood to be central actors in the metastatic cascade (28). By acting as the “seeds” of secondary tumors, CTCs underlie the spread of cancer from the original site to other organs. In modern oncology, CTCs are studied not only for their direct role in metastasis but also reviewed for their biomarkers. Because a simple blood sample can capture CTCs, their molecular characterization offers “liquid biopsy” of the tumor’s genotype and phenotype (29). The number of CTCs in blood identifies the disease stage in patient and also highlighting both their significance in cancer progression and their potential clinical utility.

The metastatic journey of CTCs begins when tumor cells acquire the ability of detachment from primary mass and intravasate into the circulation (30, 31). This requires overcoming strong cell–cell and cell–matrix adhesions, typically through processes such as from epithelial to mesenchymal transition process wherein epithelial cells change into mesenchymal cells, attaining an enhanced migratory and proteolytic degradation of the extracellular matrix. Invasion is aided by tumor-associated blood vessels and stromal cells. Thus, cancer cells invade through surrounding stroma, enters the vessel wall, and to lymphatic system, becoming CTCs. The fast-moving bloodstream exerts shear forces and can induce apoptosis in free tumor cells and simultaneously, the immune system patrols the blood and may recognize CTCs as foreign (7, 32). To survive, CTCs employ various strategies which may include the arrests in capillaries and form microemboli with platelets or clotting factors, which can shield them from shear stress and immunity. Intriguingly, CTCs often display increased plasticity and altered biomechanics (e.g. increased deformability) that help them squeeze through narrow capillaries and extravasate. Ultimately, only a small fraction of CTCs survives to exit the vasculature – extravasating into distant tissues and potentially founding metastatic.

***Exosomes***

Studies revealed that exosomes are promising biomarkers for liquid biopsy in cancer monitoring which contain biomolecules that modifies the tumor microenvironment and promote cancer progression. Exosomes are a specialized subset of extracellular vesicles which plays an important role in intercellular regions and release a wide range of cell types, including cancer cells (33, 34). These nano-sized vesicles typically from 50 -150 nm and carry a diverse cargo of biomolecules such as proteins, lipids, and nucleic acids. This molecular content not only identifies the pathological state of their cell of origin but also has a good potential to reprogram recipient cells, thereby influencing various biological processes (35).

In cancer treatment exosomes have gained increasing attention as a key facilitator of disease progression. They deeply involved in modulating the immune response, reshaping the tumour microenvironment and also promoting metastasis. One of the notable features of exosomes in cancer biology is their responsiveness to environmental cause, particularly hypoxic conditions within the tumor microenvironment.

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

Under such stress, these cancer cells alter both the quantity and composition of exosomes and result in tumor aggressiveness and metastatic potential. Moreover, exosomes contribute to therapy resistance. They can shield tumor cells from the cytotoxic effects of chemotherapeutic agents and enable the horizontal transfer of drug resistance traits to surrounding cells (36, 37).

***DNA sequencing and data conversion***

Liquid biopsy samples (typically blood plasma) yield cell-free DNA (cfDNA) fragments (~150–170 bp) shed from various cells, including tumors. This cfDNA is extracted and subjected to molecular assays. One approach is PCR-based detection: for example, droplet digital PCR (ddPCR) partitions cfDNA into thousands of droplets and amplifies target loci, yielding absolute counts of mutant vs. wild-type sequences. ddPCR (or BEAMing) can detect known hotspot mutations at very low allele fractions, often below 0.1%, due to its high sensitivity. However, PCR-based assays require prior knowledge of target variants and do not produce sequence reads; their outputs are quantitative (droplet counts or fluorescence signals) that are interpreted with specialized software (38). In practice, ddPCR complements sequencing assays by confirming specific variants, but it cannot discover new mutations or profile many genes simultaneously. For broader genomic profiling, next-generation sequencing (NGS) of cfDNA is used. In targeted NGS panels, cfDNA is converted to a sequencing library by attaching adapters (and often sample indices and Unique Molecular Identifiers, UMIs). Amplicon-based panels use multiplex PCR to enrich dozens or hundreds of regions; hybrid-capture panels use biotinylated probes to pull down selected exonic regions. After library amplification and enrichment, sequencing (typically on Illumina or Ion platforms) generates millions of short reads. The sequencer outputs raw data (Illumina's BCL files or Ion's raw signals) which are changed into FASTQ files comprising of nucleotide sequences. These FASTQ files are the starting point for computational analysis (39).

Whole-genome sequencing (WGS) of cfDNA extends this approach genome-wide. Here, library preparation uses random fragmentation (often negligible additional PCR), and sequencing covers the entire genome. WGS yields very large FASTQ datasets; for example, low-pass WGS (~0.1–1× coverage) can be used to detect copy number variations (CNVs), while deep WGS (>30×) is needed for reliable single-nucleotide variant (SNV) calling. WGS of cfDNA has the advantage of unbiased variant discovery (including SNVs, indels, structural rearrangements and methylation patterns), but at great cost and data volume. Because ctDNA can be rare, WGS often has lower sensitivity for low-frequency variants than targeted panels unless sequenced extremely deeply. In all sequencing approaches, quality control is critical: library yields, insert sizes, and sequencing quality metrics (e.g. Phred scores) are checked (40, 41).

Once sequencing data has been converted into variants and other genomic features, it can be integrated into AI and ML-based analysis. These structured data can then feed into machine learning models. Recent studies apply random forests or deep neural networks to such features to classify cancer vs. normal samples, predict tumor type, or monitor treatment response. In practice, critical pre-processing steps (filtering false positives, normalizing features) ensure clean input for AI.

**Machine Learning, Artificial Intelligence and Deep Learning in cancer detection**

Artificial intelligence (AI) is a new technology that simulates and extends human intelligence in machines that are programmed to illustrate the human actions. It has designed to think about computer intelligence and also analyse to do anything from a computer program for detection of disease

Machine learning involves in algorithms that can make predictions or decisions based upon data. In context, ML algorithms identify high-risk populations, and enrich patient populations to benefit from advanced imaging tests. Deep learning, uses

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

neural networks obtained by the human brain's architecture to analyze complex medical images and enhance diagnostic precision (42). Machine learning (ML) is developed as a powerful tool for enhancing cancer monitoring using liquid biopsy. These AI-based approaches can integrate high-dimensional features from various biomarkers to improve sensitivity and accuracy in cancer diagnosis and monitoring.

In liquid biopsy analysis, ML algorithms have been applied to combine multiple biomarkers into composite biomarkers for clinical analysis. This multianalyte approach has driven improvements in sensitivity and accuracy for cancer detection, particularly in early-stage cancers where biomarker concentrations are typically low (43). For instance, the LIQUORICE algorithm has been developed to observe tumor DNA based on specific chromatin, demonstrating the potential of ML in exploiting epigenetic deregulation for cancer detection(36, 44).

The application of machine learning (ML) in cancer detection has grown rapidly, leveraging computational power to analyze complex biomedical data and improve diagnostic accuracy. Various ML algorithms are used depending upon availability of the input data like imaging, genomics, or clinical features. These algorithms fall into different categories based on their learning approach and architecture.

***Supervised learning algorithms*** are widely employed in cancer classification where labelled datasets are available. Convolutional Neural Networks (CNNs) are particularly effective on image-based cancer detection, such as identifying tumors in radiology or pathology images. Support Vector Machines (SVM), Random Forests, and Decision Trees are also commonly used in classifying cancer types using structured clinical or genomic data. Gradient-boosted trees, including algorithms like XGBoost and LightGBM, are known for accuracy to handle heterogeneous data. Penalized logistic regression, which includes Lasso (L1), Ridge (L2), and Elastic Net variants, is especially useful for gene expression profiles due to its regularization capabilities (44).

***Unsupervised learning algorithms*** are applied when labels are not available, aiming to discover hidden patterns or subgroups within the data. K-means clustering and hierarchical clustering are often used to identify molecular subtypes of cancer. Autoencoders, are the type of neural network, that used for dimensional reduction in complex datasets. Principal Component Analysis (PCA) are also another technique used to reduce the genomic or proteomic data while preserving variance.

***Ensemble methods*** allow the predictions from multiple models to progress the overall performance. The various techniques like bagging (e.g., Random Forests), boosting (e.g., AdaBoost, XGBoost), and stacked ensembles are frequently employed in cancer prediction tasks to reduce overfitting and improve generalization. Stacked ensembles, in particular, leverage the strengths of different models by combining them through a meta-learner (45).

***Deep learning architectures*** beyond CNNs, namely Deep Belief Networks (DBNs), Long Short-Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), and Transformers, are also being explored. These models are especially useful in processing the sequential data, like time-series clinical data or genomic sequences, offering new avenues for cancer prognosis and biomarker discovery.

***Reinforcement learning*** is an emerging area in cancer research, primarily focused on optimizing personalized treatment plans and decision-making processes, such as the timing and method of biopsies. While still in early stages, its potential for dynamic, adaptive systems makes it a promising addition to cancer informatics.

The integration of diverse machine learning algorithms in cancer detection allows for tailored approaches based on data type and clinical goals. This growing toolkit is enabling earlier and more accurate diagnosis, paving the way for precision oncology (43).

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

***AI-powered platforms to detect cancer***

Several real-world AI-powered platforms are already advancing the clinical utility of liquid biopsy. For instance, DeepCNA is an AI model used to identify copy number alterations in tumor DNA, while DELFI analyses the fragmentation patterns of cfDNA to distinguish between healthy and cancerous cells. These models exemplify how AI is actively transforming the landscape of cancer diagnostics through liquid biopsy.

**DeepCNA**

Copy number alterations (CNAs) are significant genomic events involving the amplification or deletion of DNA segments ranging from kilobases to metabases. These alterations are widely observed in cancer genomes and are considered critical in cancer biology due to their ability to trigger oncogenes or deactivate tumor suppressor genes. CNAs invade a genome in cancer than any other genetic alteration and thus hold strong potential as biomarkers. Recent studies have verified the predictive power of CNAs across multiple cancer types, with CNA-derived models effectively defining risk groups in over 90% of analyzed cancers. Their importance is further reinforced by discoveries in various cancer, where CNAs provide a high-resolution map of genomic imbalances that inform disease progression and patient outcomes (46).

Historically, CNAs have been detected using karyotyping and FISH. Towards the turn of the century more advanced platforms like comparative genomic hybridization (CGH), SNP arrays, and PCR-based methods such as multiplex ligation-dependent probe amplification (MLPA) have provided greater resolution in identifying CNAs. MLPA, in particular, is mostly used in clinical diagnostics because of its speed, cost-effective, and their ability to detect known CNAs. Next-Generation Sequencing (NGS), helps to detect CNAs from both DNA and RNA, either in bulk or at the single-cell level, making liquid biopsy as an informative approach for non-invasive cancer detection. To harness this information computationally, the DeepCNA model was introduced as a Convolutional Neural Network (CNN)-based approach specifically designed for cancer analysis using CNA data and chromatin 3D structure (46).

DeepCNA is a model that processes CNA data through a series of novel steps, allowing it to be analyzed using machine learning. CNNs, are the foundation of DeepCNA, are a class of deep learning models known for their efficiency in recognizing features in image-like data structures. The major advantage of CNNs is their ability to identify the significant features from raw input data without the handcrafted features. This has made them popular not just in computer vision but also in biomedical data analysis. For instance, in brain tumor detection, CNNs have shown high accuracy in classifying MRI images into different tumor categories, including glioma, meningioma, and pituitary tumors The use of data augmentation techniques had further upgraded the performance of CNN models in brain tumor detection, addressing the challenge of limited datasets CNN architectures have been widely adopted for image classification, object detection, facial recognition, and recently, for genomics-based cancer classification tasks. CNN are used to make use of 2D input-data structures which uses the small number of parameters, which simplifies the training process and speeds up the network. Before training, CNA data undergoes a three-step preprocessing pipeline.

- First, all CNA values are clipped within the empirical range of (0, 10) to regulate data distribution and eliminate outliers that may hinder the learning process.
- Second, zero-padding is applied to adjust the length of each CNA vector so it fits the required input size of the neural network.
- Third, the padded data for 2D CNN is reshaped into a 176 × 176 × 1 structure, similar to grayscale images, making it suitable for spatial pattern recognition by convolutional layers.

In practical application, DeepCNA has been trained on approximately 15,000

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

samples of CNA data sourced from COSMIC, along with Hi-C data from two human cell lines. It achieved approximately 60% accuracy in distinguishing among 25 different cancer types. This performance demonstrates the model's effectiveness in leveraging high-dimensional CNA data, especially when enhanced by chromatin 3D context.

CNNs proved to be an effective tool in cancer detection across various imaging modalities. Their ability to extract features and learn to classify simultaneously makes them particularly suited for medical image analysis. However, challenges remain, including the requirement for large datasets, computational resources, and model interpretability. Future research should focus on improving model interpretability, combining clinical records with imaging data, and exploring transfer learning and synthetic data acquisition to address data limitations and enhance model performance (47, 48).

**DNA Evaluation of Fragments for Early Interception**

Cell-free DNA fragmentomics refers to the analysis of genomic and epigenomic signals encoded in the size, coverage, and nucleotide context of circulating DNA fragments. DNA Evaluation of Fragments for Early Interception is a paradigm to detect cancer-specific fragment patterns. In DELFI, the genome is divided into large bins (e.g. 5-Mb windows) and millions of cfDNA fragments are counted by size and genomic position. Fragmentation signals – including the relative abundance of short (e.g. 100–150 bp) versus long fragments, nucleosome-protected coverage, and sequence motifs at fragment ends – are then fed to a machine learning model to classify cancer vs. non-cancer. tudies showed that a DELFI classifier could distinguish cancer patients across multiple tumor types at high specificity: for seven cancer types (lung, breast, colorectal, ovarian, pancreatic, gastric and bile-duct), Cristiano *et al.* reported sensitivities from 57–99% at 98% specificity (AUC ≈0.94) using genome-wide fragmentomic features. The model also inferred tissue-of-origin in ~75% of cases. Thus, DELFI demonstrated the potential of genome-scale cfDNA patterns as a pan-cancer early detection tool.

In practical pipelines, cfDNA sequencing data are pre-processed through standard NGS workflows. Raw reads are trimmed for adapters and aligned to the human reference genome. PCR duplicates are removed and properly paired-end reads are converted to fragment coordinates. Fragments are then binned across the genome, commonly in non-overlapping 5-Mb windows (~80,000 fragments per bin). Regions with low mappability or extreme GC content are usually excluded or normalized. The end result is a matrix of fragment counts and coverage in each window, stratified by fragment length or GC class (49, 50).

DELFI and related fragmentomic tests use supervised learning to classify samples, which means the model is trained using labelled data, samples that are already known to be from either healthy individuals or cancer patients. The original multi-cancer DELFI model employed gradient-boosted trees (XGBoost) on features including fragment size/coverage, arm-level CNV, and mitochondrial DNA counts . Reaseachers also used penalized logistic regression on the fragmentomic bins. Newer studies often use stacked ensembles: for example, Shi et al. (2024) built a stacked model combining three fragmentomic feature sets. Deep learning is also being explored: the EMIT transformer model can learn cfDNA end-motif patterns in a self-supervised way, potentially bypassing laborious feature engineering (50). Model training and validation follow standard procedures. Data are split into training and test sets, often with external validation cohorts. Velculescu's group reported 10-fold cross-validation in which data is split into 10 parts and the model is trained and tested multiple times—to make their model more reliable. In a large lung cancer study, Mazzone et al. trained their model on data from 576 people and then tested it on an independent group of 382 people (51).

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

Certain applications that made DELFI effective, In lung cancer, DELFI was first applied prominently to lung cancer screening.In a prospective "LUCAS" study (365 at-risk individuals) and validation in 46 lung cancer vs. 385 controls, combining cfDNA fragmentation with clinical factors achieved 94% sensitivity for lung cancer at 80% specificity (91% for Stage I/II). uilding on this, Mazzone *et al.* (2024) enrolled 958 *screening-eligible* individuals (576 train / 382 test) and trained a low-coverage WGS fragmentome assay followed by LDCT. The model demonstrated 84% overall sensitivity in the held-out validation. Importantly, incorporating DELFI into lung screening pathways could significantly increase early detection: the authors' decision analysis predicted that offering this blood test could prevent thousands of lung cancer deaths by raising effective screening rates. Liver cancer (Hepatocellular Carcinoma) causes high cancer mortality, especially in persons with viral hepatitis (51). A machine learning model (penalized logistic regression) attained 88% sensitivity and 98% specificity for detection in an average-risk cohort, and 85% sensitivity at 80% specificity in a cirrhosis surveillance cohort. Early-stage HCC was also detected with high sensitivity. These results greatly surpass traditional AFP screening (sens 47–84%) and ultrasound (51).

**Challenges and Future Prospects**

Several strategies are emerging to address the challenges that are faced in cancer treatment. Multimodal integration is a powerful trend which combines the liquid biopsy with other modalities that can boost signal. Research studies by Medina et al. combined cfDNA fragmentation features (DELFI) with two protein markers (CA-125, HE4) to screen for ovarian cancer. This multianalyte model achieved >99% specificity and sensitivities of 72–100% for stages I–IV (versus ~28–63% for each protein alone). Similarly, Yang et al. reported a prospective gallbladder cancer study ("GBCseeker") that fuses cfDNA signatures, imaging features, and clinical data; this multimodal model effectively distinguished cancer from benign disease across multiple centres. In general, combining cfDNA with proteomic, transcriptomic, or radiologic data allows AI models to cross-validate findings and reduce false calls.

Groups (e.g. the NCI's Liquid Biopsy Consortium) of uniting institutions standardize the assays and pool insights. Federated ML enlarge this by enabling joint model development without data exchange. Such approaches can mitigate biases and improve robustness across populations. Large-scale studies are ongoing which aims to validate new liquid biopsy assays and address known obstacles. Other early-detection trials (e.g. GRAIL's Galleri, PATHFINDER, DETECT-A) and cancer screening programs are beginning to incorporate ML-based cfDNA analysis. Success in these trials will build confidence and inform regulatory approval. AI researchers are applying cutting-edge methods to cfDNA wherein, the transformer-based models and self-supervised pretraining on genomic data are being explored. Explainable AI tools are also under development to trace predictions back to genomic features, which may improve clinician acceptance. Specialized AI frameworks for cfDNA are maturing to demonstrate the value of deep learning on genomic patterns. Similarly, the DELFI approach to fragment omics has seen extensions: Sangphukieo et al. (2024) introduced "xDELFI," an exome-based fragmentomic model using just 0.08× WGS coverage which achieved AUC ≈0.90 for cancer detection, nearly matching the full-genome DELFI (AUC ≈0.92), showing the targeted sequencing can retain high accuracy at lower cost. The DELFI framework has also expanded into monitoring: e.g. a recent Nature Communications study developed "DELFI-TF" to quantify tumor fraction longitudinally during treatment in colorectal cancer (not cited here). The DELFI concept has even been extended to include protein markers (as above) and to tissue-of-origin prediction. Together, these innovations suggest AI-augmented fragmentomics can

become both more sensitive and more accessible. This AI integration in liquid biopsy is progressing rapidly, yet has to overcome the hurdles in data, validation, and regulation.

**Conclusion**

This review paper helps us to explore the field of liquid biopsy and also highlighting its potential as a non-invasive and dynamic tool for cancer detection. By focusing on biomarkers, this review emphasizes their critical role in providing molecular insights into tumor biology without surgical procedures.

This review paper also discussed how raw biological data from these biomarkers was converted into computationally analysable formats using advanced molecular techniques like PCR, NGS, and WGS. These methods also serve as crucial mediators by translating complex biological signals into digital data that can be processed and interpreted through computational models.

The AI, particularly Machine Learning (ML) and Deep Learning (DL) models, have improved the accuracy of liquid biopsy analysis. The model enables the identification of genomic alterations which are often missed by conventional methods, thereby helping in early cancer detection and personalised treatment strategies.

Despite these advancements, the paper has acknowledged several ongoing challenges, including variability in sample collection, the need for standardized protocols, and the requirement for robust clinical validation. This review gives various information on several ongoing challenges, including variability in sample collection, the need for standardized protocols, and the requirement for clinical validation. However, the future prospects remain optimistic. With continuous innovation in AI and molecular diagnostics, liquid biopsy has made its importance as an integral component of precision oncology.

**References**

1. Jhaveri FM, Brawer MK, Klein EA (2003) The role of prostate specific antigen and its variants in prostate cancer screening. In: Prostate Cancer Screening. Humana Press, New Jersey, pp 79–90
2. Milosevic M, Jankovic D, Milenkovic A, Stojanov D (2018) Early diagnosis and detection of breast cancer. Technol Health Care 26:729–759
3. Kaur S, Baine MJ, Jain M, Sasson AR, Batra SK (2012) Early diagnosis of pancreatic cancer: challenges and new developments. Biomark Med 6:597–612
4. Bao Y, Zhang D, Guo H, Ma W (2024) Beyond blood: Advancing the frontiers of liquid biopsy in oncology and personalized medicine. Cancer Sci 115:1060–1072
5. Singh S, Podder PS, Russo MJ, Henry C, Cinti S (2022) Correction: Tailored point-of-care biosensors for liquid biopsy in the field of oncology. Lab Chip 22:5008
6. Ge Q, Zhang Z-Y, Li S-N, Ma J-Q, Zhao Z (2024) Liquid biopsy: Comprehensive overview of circulating tumor DNA (Review). Oncol Lett 28:548
7. Wen X, Pu H, Liu Q, Guo Z, Luo D (2022) Circulating tumor DNA-A novel biomarker of tumor progression and its favorable detection techniques. Cancers (Basel) 14:6025
8. Li L, Sun Y (2024) Circulating tumor DNA methylation detection as biomarker and its application in tumor liquid biopsy: advances and challenges. MedComm 5:e766
9. Alimirzaie S, Bagherzadeh M, Akbari MR (2019) Liquid biopsy in breast cancer: A comprehensive review. Clin Genet 95:643–660
10. Vaidyanathan R, Soon RH, Zhang P, Jiang K, Lim CT (2018) Cancer diagnosis: from tumor to liquid biopsy and beyond. Lab Chip 19:11–34
11. Wu H-J, Chu P-Y (2022) Current and developing liquid biopsy techniques for breast cancer. Cancers (Basel) 14:2052
12. Lone SN, Nisar S, Masoodi T, et al (2022) Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments. Mol Cancer 21:79
13. Sheriff S, Saba M, Patel R, et al (2025) A scoping review of factors influencing the implementation of liquid biopsy for cancer care. J Exp Clin Cancer Res 44:50

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

14. Ma L, Guo H, Zhao Y, Liu Z, Wang C, Bu J, Sun T, Wei J (2024) Liquid biopsy in cancer current: status, challenges and future prospects. Signal Transduct Target Ther 9:336
15. Lee J-S, Cho EH, Kim B, et al (2024) Clinical practice guideline for blood-based circulating tumor DNA assays. Ann Lab Med 44:195–209
16. Čeri A, Somborac-Bačura A, Fabijanec M, Hulina-Tomašković A, Matusina M, Detel D, Verbanac D, Barišić K (2024) Establishment of liquid biopsy procedure for the analysis of circulating cell free DNA, exosomes, RNA and proteins in colorectal cancer and adenoma patients. Sci Rep 14:26925
17. Maass KK, Schad PS, Finster AME, Puranachot P, Rosing F, Wedig T, Schwarz N, Stumpf N, Pfister SM, Pajtler KW (2021) From sampling to sequencing: A liquid biopsy pre-analytic workflow to maximize multi-layer genomic information from a single tube. Cancers (Basel) 13:3002
18. Kim H, Park KU (2023) Clinical circulating tumor DNA testing for precision oncology. Cancer Res Treat 55:351–366
19. Sánchez-Herrero E, Serna-Blasco R, Robado de Lope L, González-Rumayor V, Romero A, Provencio M (2022) Circulating tumor DNA as a cancer biomarker: An overview of biological features and factors that may impact on ctDNA analysis. Front Oncol 12:943253
20. Moding EJ, Nabet BY, Alizadeh AA, Diehn M (2021) Detecting liquid remnants of solid tumors: Circulating tumor DNA minimal residual disease. Cancer Discov 11:2968–2986
21. Walter N, Groth J, Zu Zwerger B von U (2025) Evaluation of an innovative multi-cancer early detection test: high sensitivity and specificity in differentiating cancer, inflammatory conditions, and healthy individuals. Front Oncol 15:1520869
22. Bartolomucci A, Nobrega M, Ferrier T, Dickinson K, Kaorey N, Nadeau A, Castillo A, Burnier JV (2025) Circulating tumor DNA to monitor treatment response in solid tumors and advance precision oncology. NPJ Precis Oncol 9:84
23. Kurma K, Alix-Panabières C (2023) Mechanobiology and survival strategies of circulating tumor cells: a process towards the invasive and metastatic phenotype. Front Cell Dev Biol 11:1188499
24. Krebs MG, Hou J-M, Ward TH, Blackhall FH, Dive C (2010) Circulating tumour cells: their utility in cancer management and predicting outcomes. Ther Adv Med Oncol 2:351–365
25. Salu P, Reindl KM (2024) Advancements in circulating tumor cell research: Bridging biology and clinical applications. Cancers (Basel).
https://doi.org/10.3390/cancers16061213
26. Zavyalova MV, Denisov EV, Tashireva LA, Savelieva OE, Kaigorodova EV, Krakhmal NV, Perelmuter VM (2019) Intravasation as a key step in cancer metastasis. Biochemistry (Mosc) 84:762–772
27. Lu Y-J, Wang P, Peng J, Wang X, Zhu Y-W, Shen N (2017) Meta-analysis reveals the prognostic value of circulating tumour cells detected in the peripheral blood in patients with non-metastatic colorectal cancer. Sci Rep 7:905
28. Ruivo CF, Adem B, Silva M, Melo SA (2017) The biology of cancer exosomes: Insights and new perspectives. Cancer Res 77:6480–6488
29. Yang S, Sun Y, Liu W, Zhang Y, Sun G, Xiang B, Yang J (2024) Exosomes in glioma: Unraveling their roles in progression, diagnosis, and therapy. Cancers (Basel) 16:823
30. Karampoga A, Tzaferi K, Koutsakis C, Kyriakopoulou K, Karamanos NK (2022) Exosomes and the extracellular matrix: a dynamic interplay in cancer progression. Int J Dev Biol 66:97–102
31. Wang W, Han Y, Jo HA, Lee J, Song YS (2020) Non-coding RNAs shuttled via exosomes reshape the hypoxic tumor microenvironment. J Hematol Oncol 13:67
32. Fan M, Weng Y, Liu Y, Lu Y, Xu L, Ye J, Lin D, Qiu S, Feng S (2024) Surface-enhanced Raman spectroscopy-based optical biosensor for liquid biopsy: Toward precision medicine. Laser Photon Rev.

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.4

33. Abdul Rasool Hassan B, Mohammed AH, Hallit S, Malaeb D, Hosseini H (2025) Exploring the role of artificial intelligence in chemotherapy development, cancer diagnosis, and treatment: present achievements and future outlook. Front Oncol 15:1475893
34. Pak S, Park SG, Park J, Cho ST, Lee YG, Ahn H (2024) Applications of artificial intelligence in urologic oncology. Investig Clin Urol 65:202–216
35. Wei L, Niraula D, Gates EDH, Fu J, Luo Y, Nyflot MJ, Bowen SR, El Naqa IM, Cui S (2023) Artificial intelligence (AI) and machine learning (ML) in precision oncology: a review on enhancing discoverability through multiomics integration. Br J Radiol 96:20230211
36. Peneder P, Stütz AM, Surdez D, et al (2021) Multimodal analysis of cell-free DNA whole-genome sequencing for pediatric cancers with low mutational burden. Nat Commun 12:3230
37. Lyu N, Hassanzadeh-Barforoushi A, Rey Gomez LM, Zhang W, Wang Y (2024) SERS biosensors for liquid biopsy towards cancer diagnosis by detection of various circulating biomarkers: current progress and perspectives. Nano Converg 11:22
38. Lawal AO, Ogunniyi TJ, Oludele OI, Olorunfemi OA, Okesanya OJ, Ogaya JB, Manirambona E, Ahmed MM, Lucero-Prisno DE 3rd (2025) Innovative laboratory techniques shaping cancer diagnosis and treatment in developing countries. Discov Oncol 16:137
39. Al-Badri MA, Cross WCH, Barnes CP (2024) DeepCNA: an explainable deep learning method for cancer diagnosis and cancer-specific patterns of copy number aberrations. bioRxiv. https://doi.org/10.1101/2024.03.08.584160
40. Yuan Y, Shi Y, Su X, Zou X, Luo Q, Feng DD, Cai W, Han Z-G (2018) Cancer type prediction based on copy number aberration and chromatin 3D structure with convolutional neural networks. BMC Genomics 19:565
41. Guardado A, Aguirre-Gamboa R, Treviño V (2024) Systematic modeling of risk-associated copy number alterations in cancer. Int J Mol Sci. https://doi.org/10.3390/ijms251910455
42. Mostavi M, Chiu Y-C, Huang Y, Chen Y (2020) Convolutional neural network models for cancer type prediction based on gene expression. BMC Med Genomics 13:44
43. Taylor BS, Barretina J, Socci ND, Decarolis P, Ladanyi M, Meyerson M, Singer S, Sander C (2008) Functional copy-number alterations in cancer. PLoS One 3:e3179
44. Carey-Smith SL, Kotecha RS, Cheung LC, Malinge S (2024) Insights into the clinical, biological and therapeutic impact of copy number alteration in cancer. Int J Mol Sci 25:6815
45. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8:53
46. Gong B, Deveson IW, Mercer T, Johann DJ Jr, Jones W, Tong W, Xu J (2022) Ultra-deep sequencing data from a liquid biopsy proficiency study demonstrating analytic validity. Sci Data 9:170
47. Röner S, Burkard L, Speicher MR, Kircher M (2024) cfDNA UniFlow: a unified preprocessing pipeline for cell-free DNA data from liquid biopsies. Gigascience. https://doi.org/10.1093/gigascience/giae102
48. Garcia J, Kamps-Hughes N, Geiguer F, Couraud S, Sarver B, Payen L, Ionescu-Zanetti C (2021) Sensitivity, specificity, and accuracy of a liquid biopsy approach utilizing molecular amplification pools. Sci Rep 11:10761
49. Mathios D, Johansen JS, Cristiano S, et al (2021) Detection and characterization of lung cancer using cell-free DNA fragmentomes. Nat Commun 12:5060
50. Foda ZH, Annapragada AV, Boyapati K, et al (2023) Detecting liver cancer using cell-free DNA fragmentomes. Cancer Discov 13:616–631
51. Cristiano S, Leal A, Phallen J, et al (2019) Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 570:385–389