Survey on Cancer Detection Techniques

S.J. Subhashini[3], K.Nagalakshmi[2], M.Ilakkiya[1]
Associate Professor [3, 2], PG (M.E) Scholar[1]
Department of CSE, KLN College of Information Technology, Pottapalayam, Sivagangai District, Tamil Nadu, India.

***Abstract*—This paper proposed to analyze the earlier cancer detection method to easily and efficiently identify the cancer tumours such as lung cancer, skin cancer, breast cancer and blood cancer. These are using the image processing techniques and machine learning techniques. Periodic clinical checkups and self-tests help in early detection and thereby significantly increase the chances of survival. In this work, an automated system is to achieving error-free detection of breast cancer using mammogram. A number of methods have been developed on classification, clustering, and probabilistic techniques for detection of CaP (prostate cancer). X-ray mammography is the main test used for screening and early diagnosis, and its analysis and processing are the keys to improving breast cancer prognosis. The proposed system analyzes the techniques to overcome the disadvantages for the existing methods.**

I. INTRODUCTION

Big Data is set of data which is so complex and voluminous that application of data processing software is inadequate to deal with them. it challenges include data analysis, data storage, visualization, capturing data, updating, search, sharing, transfer, querying information privacy and data source. it can be contrasted with small data, another evolving term that's often used to describe data whose volume and format can be easily used for self-service analytics. A commonly quoted axiom is that big data is for machines; small data is for people.

There are a number of concepts associated with big data: originally there were 3 concepts volume, variety, and velocity. Other concepts later attributed with big data are veracity and value.

1. Volume

The quantity of generated and stored data. The size of the data determines the value and potential insight and whether it can be considered big data or not.

2. Variety

This concept is the type of nature. It helps people who analyze to effectively use the resulting insight. Big data draw from text, images, audio and video; plus it completes missing pieces through the data fusion.

3. Velocity

In this context, the speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development. Big data is often available in real-time.

4. Variability

Inconsistency of the data set can champer processes to handle and manage it.

5. Veracity

The data quality of captured data can vary greatly, affecting the accurate analysis.

Big Data tend to refer to the use of predictive analytic, user behavior analytic and advanced data analytic methods that extract value from data, and seldom to a particular size of the data set. Data set grow rapidly - in part because they are increasingly gathered from cheap and numerous information-sensing Internet of things devices such as mobile devices, aerial (remote sensing), software log, camera, microphone, Radio-Frequency identification (RFID) readers and wireless sensor networks. The world's technology per-capita capacity to store information has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 Exabyte (2.5×10^{18}) of data are generated. Based on an Idiopathic Dilated Cardiomyopathy (IDC) report prediction, the global data volume will grow exponentially from 4.4 petabyte to 44 petabyte between 2013 to 2020. By 2025, IDC predicts there will be 163 petabytes of data. One question for large enterprise is determining who should own big-data initiatives that affect the entire organization. The growing maturity of the concept more starkly delineates the difference between big data and Business Intelligence. business Intelligence uses descriptive statistics with data with high information density to measure things, detect trends, etc. Big data use inductive statistics and concepts from nonlinear system identification to infer laws (regressions, nonlinear relationships, and causal effects) from large sets of data with low information density to reveal relationships and dependencies, or to perform predictions of outcomes and behaviors.

Applications:

- Government
- International development
- Manufacturing
- Healthcare
- Education
- Media

- Internet of Things (IOT)
- Information Technology

The *digital image processing* (DIP) has been employed in a number of areas, particularly for feature extraction and to obtain patterns of digital images. Recognition of characters are a novel problem, and although currently there are widely-available digital image processing algorithms and implementations that is able to detect characters from images, selection of an appropriate technique that can straightforwardly acclimatize to diverse type of images, that are very specific or complex is very important. This paper presents a brief overview of digital image processing techniques such as image restorations, image enhancements, and feature extraction, a framework for processing images and aims at presenting an adaptable digital image processing method for recognition of characters in digital images. This processing is under the concept of big data, because the collection of data or data collected to the different sources its managing the big data and using image processing techniques for the cancer detection. This processing technique involved preprocessing, feature extraction, image acquisition, line segmentation to detect the cancer tumors in the human. These techniques are providing the result of high accuracy and efficiency of cancer detection.

II. LITERATURE SURVEY

The most common cause of death for women's Breast cancer. It's second leading in worldwide. The Breast cancer tested by the screening and diagnosis. This processing and analysis is improved the breast cancer using X-ray mammography in early stages very complex for primary prevention. Internist and radiologist's diagnosis supported by the computer aided diagnosis scheme. The proposed diagnosis system is digital mammogram; it's based on texture segmentation to detect the early stage tumours. The proposed system introduced new algorithm took the more images from the digital database for screening mammography for cancer research and diagnosis. It was found to be absolutely suitable to distinguish masses and micro calcifications from the background tissue using morphological operators, extract through machine learning techniques and a intensity-based segmentation by clustering algorithm. [1]

In New Zealand, Northern America, Europe and India most of the people can be affected by the Prostate Cancer is shortly known as CaP. Its detected by probabilistic, clustering and classification techniques their pros and cons basic gaps of conventional methods need to be addressed in CaP diagnosis and detection. [2]

One of the most dangerous disease is lung cancer in the world. It evaluations of computer aided techniques and main aim of this paper to analyzed one of the best technique and also found their advantages and limitations and to improve the model of the best model. Machine learning and image processing are done by CAD cells can be diagnosis accurately and also CT scan is one of the best technique for imaging but it's difficult to interpret and identified by the doctors. All the techniques are to produce higher and lower accuracy but not nearby 100%. The main focus of this paper to increased the accuracy nearby 100. [3]

Breast cancer and lung cancer have the second mortality rate in women, the proposed automated system for achieved best cancer error free detection using the mammogram 1 women can be affected the breast cancer out of 8 women as per clinical statistics, self test and periodic clinical checkups help early detection of the breast cancer, the cause technique of invasive detection was a tumor may be cracked and to scatter an adjoining areas. Sparse auto encoder and stacked sparse auto encoder is using conventional neural network in deep learning techniques, these techniques can be analyzed the performance of existing techniques. [4]

The image processing technique can detect the lung cancer and recognize the information about the lung tissue. Auto detection detects the small nodules in early stage lung cancer is difficult one so we can introduce a ridge detection algorithm, Intermediate nodules (non cancerous) are diagnosed. The proposed algorithm to allow the curative resection of early stage malignant nodules and avoided the mortality and morbidity of surgery for benign modules. Some traditional image segmentation technique compared with a ridge detection algorithm, both results are satisfactory the diagnosis. [5]

The proposed system to introduce the cancer detection method, to improve the surface enhanced Raman spectroscopy in vitro and vivo from the basic principle of Raman spectroscopy (RS) to discover the cancer, elevated to the nanoparticle increases the specificity and sensitivity. The main focus of this paper to introduce new technologies of nanoparticle using the RS method for cancer detection, SERS to cooperate with different diagnostic imaging tools. [6]

The proposed system to developed hysteresis loop analysis method for to diagnose breast cancer which is validates the diet diagnostic approach using by the silicone phantom test, in this hysteresis loops are constructed the breast surface by using the reference point of measured displacements and mass will be calculated using by normalized restoring force and to estimate the breast surface using regression analysis and hypothesis test. The displacement of mechanical actuation is captured using the DIET system for 4 silicone breast phantoms. The main focus of this proposed system is sensitivity analysis of shown displacement reconstruction errors from very small motions can be eliminated using a threshold; This method generalized the diet system to human clinical data and justifies the clinical trials. [7]

Colorectal cancer screening program is modalities based on evidence to its including visualization and biomarkers development of reliable, sensitive, patient friendly, minimal invasive technology are one of the ultimate goals of this paper, Circulating tumour cells are the best understood source of blood-based colorectal cancer biomarkers but their detection is more useful for assessing responses to treatment for metastatic disease rather than in screening. Analysis of VOMs represents a promising new screening tool for the detection of colorectal cancer. VOMs are gaseous carbon-based chemicals, representing the molecular basis for sense of smell. The successful uptake of prostate-specific antigen testing for prostate cancer, despite conflicting evidence regarding its efficiency, can, in part, be attributed to the fact that it is relatively non-invasive and can be administered at the primary point of patient care. [8]

The incidence of HPV-related head and neck squamous cell carcinoma (HPV- HNSCC) is increasing. Oral samples are easy and non-invasive to collect, but the diagnostic accuracy of oral HPV detection methods for classifying HPV-positive HNSCC tumors have not been exploring well. In this method have been meeting the four types of criteria problem.

1. Tumour HPV or p16 testing.
2. A publication date within the last 10 years.
3. HPV detection in oral rinse or oral swab samples.
4. At least 15 HNSCC cases.

This test may have utility for screening in some settings and subgroups, but is not currently and commercially available. There also an interested in developing biomarkers based on miRNA expression profiles in HNSCC. There was substantial variation across the studies included in our meta-analysis with regard to methods of oral and tumor HPV testing, as well as patient characteristics such as geographic location. [9]

Cancer molecular genetics to improve the lung cancer detection technology. AFB has shown to increase detection sensitivity of recurrent or new intraepithelial neoplasias and invasive carcinomas when added to WLB (from 25% for WLB alone to 75% when AFB is used in conjunction with WLB) in the postoperative surveillance of patients who underwent curative resection for NSCLC. AFB combined with CT of the thorax in patients with radiographically suspicious and occult lung cancer has shown to be an effective lung cancer staging and tumor extension assessment modality with impact on therapeutic strategy choice. [10]

Skin cancer diagnosis in early stages possible to highly advanced medical technology to involve the preprocessing for skin cancer and the classification technique are used. Dermatologists are obligated to perform a surgery to define the tumor as menacing or less threatening. Malignant melanoma is the most communal type of skin cancer which is incurable if found in its later stages, but can be treated in its early stages. Skin lesion classification from a dermatoscopic image after acquiring the image consists of four phases: Image Preprocessing, Image foreground and background Segmentation, appropriate Feature extraction or selection and finally decision-making using Classification.[11]

Biomarker of cancer is discovered by circulating tumour cells (CTC) to analyses the uterine cervical cancer patients samples of blood were hemolyzed and infected with the virus and then labeled with fluorescent anti-CD45 and anti-pan cytokeratin antibodies. GFP (+)/CD45 (_) cells.6 of 23 Patients are detected by the CTC cervical cancers (26.0%). [12]

Automated classification of the breast cancer to using the histopathological images one of the challenging task of accurate detection of the tumor sub system, This process facilitated by the machine learning techniques and established the method for computer aided diagnosis (CAD). [13]

Cancer death reducing by the early stage detection method to detect the level of mutations and proteins in cell free DNA called cancer SEEK to clinically detect cancers of the ovary stomach, liver, colorectum, pancreas, esophagus, breast or Cancer SEEK of the lung to localize the cancer number of anatomic sites in small, a median of 83% of the patients. Cosmic dataset used to detect the larger fraction of tumour. One of the most important attributes of a screening test is the ability to detect cancers at relatively early stages. [14]

Efficient and robust cell detection serve as critical prerequisite for many subsequent biomedical image analysis methods an computer- aided diagnosis (CAD). It remains a challenging task due to touching cells inhomogeneous background noise, and large variations in cell sizes and shapes. Scanned images available in the dataset further for demand efficient processing algorithm. Efficient cell detection of proposed convolution neural network for the regression model, report of the new system to produce Accuracy of detection and reducing running time.[15]

III. CONCLUSION

Cancer Detection is a very important research area which is motivated by many numbers of applications. The purpose of this project presented an outline of established tactics for Cancer detection. The detection involves all types of cancers to increasing the performance of an early detection cancer method. The difficulties of these techniques to diagnosing skin cancer, which involves the basic preprocessing but problem is faced during selection of a classifier gives best diagnose. To improve the classification techniques, algorithm for early detection method. The adenoma detection rate was inversely associated with the risks of interval colorectal cancer, advanced-stage interval cancer, and fatal interval cancer.

REFERENCES

[1]. Guzmán-Cabrera, R., Guzmán- Sepúlveda, J. R., Torres-Cisneros, M., May-Arrioja, D. A., Ruiz-Pinales, J., Ibarra-Manzano, O. G.,

... & Parada, A. G. (2013). Digital image processing technique for breast cancer detection. *International Journal of Thermophysics*, *34*(8-9), 1519- 1531.

[2]. Garg, G., & Juneja, M. (2018). A Survey on Computer-Aided Detection Techniques of Prostate Cancer. In *Progress in Advanced Computing and Intelligent Engineering* (pp. 115-125). Springer, Singapore.

[3]. Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, (2018). Lung Cancer Detection using CT Scan Images. *Procedia Computer Science*, *125*, 107-114.

[4]. Selvathi, D., & Poornila, A. A. (2018). Deep Learning Techniques for Breast Cancer Detection Using Medical Image Analysis. In *Biologically Rationalized Computing Techniques For Image Processing Applications* (pp. 159-186). Springer, Cham.

[5]. Wang, W., & Wu, S. (2006, June). A study on Lung cancer detection by Image processing. In *Communications, Circuits and Systems Proceedings, 2006 International Conference on* (Vol. 1, pp. 371-374). IEEE.

[6]. Ravanshad, R., Karimi Zadeh, A., Amani, A. M., Mousavi, S. M., Hashemi, S. A., Savar Dashtaki, A., ... & Zare, B. (2018). Application of nanoparticles in cancer detection by Raman scattering based techniques. *Nano Reviews & Experiments*, *9*(1), 1373551.

[7]. Zhou, C., Chase, J. G., Ismail, H., Signal, M. K., Haggers, M., Rodgers, G. W., & Pretty, C. (2018). Silicone phantom validation of breast cancer tumor detection using nominal stiffness identification in digital imaging elasto-tomography (DIET). *Biomedical Signal Processing and Control*, *39*, 435-447.

[8]. Ryan, E. J., & Creagh, E. M. (2018). Emerging methods in colorectal cancer screening. *BJS*, *105*(2).

[9]. Gipson, B. J., Robbins, H. A., Fakhry, C., & D'Souza, G. (2018). Sensitivity and specificity of oral HPV detection for HPV-positive head and neck cancer. *Oral Oncology*, *77*, 52-56.

[10]. Nakajima, T., & Yasufuku, K. (2018). Early Lung Cancer: Methods for Detection. In *Interventions in Pulmonary Medicine* (pp. 245-256). Springer, Cham.

[11]. Arora, G., Dubey, A. K., & Jaffery, (2018). Classifiers for the Detection of Skin Cancer. In *Smart Computing and Informatics* (pp. 351-360). Springer, Singapore.

[12]. Takakura, M., Matsumoto, T., Nakamura, M., Mizumoto, Y., Myojyo, S., Yamazaki, R., ... & Iizuka, T. (2018). Detection of circulating tumor cells in cervical cancer using a conditionally replicative adenovirus targeting telomerase-positive cells.*Cancer science*.

[13]. Motlagh, N. H., Jannesary, M., Aboulkheyr, H., Khosravi, P., Elemento, O., Totonchi, M., & Hajirasouliha, I. (2018). Breast Cancer Histopathological Image Classification: A Deep Learning Approach.

[14]. Yuanpu Xie, Fuyong Xing, Xiaoshuang Shi, Xiangfei Kong, Hai Su, Lin Yang, Effcient and Robust Cell Detection: A Structured Regression Approach, *Medical Image Analysis* (2017),doi: 10.1016/j.media.2017.07.003

[15]. Joshua Cohen, J. D., Li, L., Wang, Y., Thoburn, C., Afsari, B., Danilova, L.,& Hruban, R. H. (2018). Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*, eaar3247.