

Deep learning-based models of molecular phenotypes for predicting the overall survival in cancer

ABSTRACT

Aims: The aim of the study is to justify the need of deep learning predictive model in obtaining molecular phenotypes of overall cancer survival.

Study design: The study is based on the secondary qualitative data analysis through usage of systematic review.

Methodology: A qualitative study has been conducted to analyse the necessity of deep learning. It also includes the need for deep learning models to obtain the imaging of the cancer cells. In the study, a detailed discussion on deep learning has been made. The analysis of the primary sources has been obtained by evaluating the quality of the resources in the study. The study also comprises of a thematic analysis that enlightens the benefits of deep learning. The study is based on the analysis of 14 primary research-based articles out of 112 quantitative articles and structuring of a systematic review from the collected data.

Result: The morphological and physiological changes that occur in the cancerous cells have been clearly evaluated in the research. The result signifies the prediction can be made by implementing deep learning in terms of cancer survival. Advancements in terms of technology in the medical field can thus be improved with the help of the deep learning process. It states the advancements of the deep learning models that are helpful in predicting the model of cancer to determine survival rate.

Conclusion: Deep learning is a process that is considered to be a subset of artificial intelligence. Deep learning programmes are meant to be performed for complex learning models. Although there is difference in the concept of deep learning and image processing still artificial intelligence brings both together so as to ensure better performance in image processing. The need for deep learning models has become invasive, and it helps to build a strong ground for cancer survival.

Keywords: Mutagenesis, Cancer, Deep Learning, Gene expression, Convolutional Neural Network Models

1. INTRODUCTION

Deep learning models have now become an important part of increasing the prognosis rate in cancer patients. These are considered to be novel frameworks that implement computational algorithms in biology. The idea of deep learning, as well as machine learning, completely differs from each other. However, both of them can be put together to perform in collaboration so that better results can be obtained. As per the view of Liu and Lang (2019), it has been manifested that deep learning is a process that has become a hotspot for the machine learning processes. It is thus necessary to influence modern techniques in the medical field as it would help to tend a better treatment process. The main objective of the research is to analyse deep learning based models related to molecular phenotypes of cancer and its implication in overall survival in cancer. The research also has the objective of identifying the cellular signals that help in image processing.

As illustrated by Rathore *et al.* (2019), tumour cells that take the shape of malignancy require a critical treatment method to get reduced. It is indeed a true statement that all cancers are tumours, whereas all tumours are not cancer. The statement states that the initial stage of every cancer is a tumour. It shows that the treatment must be done at the very beginning stage. However, no such diagnosis has been evaluated that can remove the cancerous cells completely without giving any radiation to the body. Computational approaches have become an important domain of microscopic images of tissue specimens that might also deal with cancer. The study aims to determine the characteristics of deep learning predictive models in testing the overall survival rate of cancer patients. The objective of the study is, thus, to introduce the present advancements of deep learning predictive models in testing the hypothesis of overall survival rate in cancer in terms of molecular phenotypes. The objective of the study also deals with the features of molecular anatomical changes that occur in the cancerous cell. In cancerous cells, the molecular structure needs to be determined as it shows several structural changes and chromosomal changes. The point mutations are found in the chromosomes that are hard to be featured and distinguished without proper images. There the need for a deep learning model arises so that it can give a definite structure of the modifications of cancerous cells. However, in the case of survival rate, it has done wonders because it would even define the initial stage of the cancer that is, the tumor that would help to start the treatment at an immediate basis.

Background

Cancer is mainly caused due to malfunctioning of the cellular signalling process owing to which there is uncontrolled growth of cells. This uncontrolled growth is facilitated by metastases which lead to spreading of the infection to other parts of the body thus causing rapid spread of the infection throughout the system. Neuro-endocrine metastatic tumours are slow growing on the contrary AML is a fast growing cancer. Thus the growth of the cells depends on the cell type and the type of mutation along with the environment in which the tumour is present. As per the view of Karpiński (2019), approximately 9.6 million of people have died from cancer due to lack of proper treatment methods. As per the view of Chen *et al.* (2020), several decades have passed, and there is not much progress in the determination of cancerous stages. The patients are always treated at the intermediate or last stage and never in the initial stage due to less visible reports. A number of limited information can be found near the pathologists; however, the number is insufficient. On the other hand, in the complicated cancerous stages, such as in the case of brain cancer or even in hepatic cancer, a clear image is being required to identify the channels of that particular cancer. However, there is no such diagnosis method except deep learning that can build up a clear image of the happenings in the cancerous stages. As illustrated by Cheerla and Gevaert (2019), it is necessary to evaluate the deep learning of predictive models because the present invitations fail to produce proper and suggestive information about cancer. A vast mode of multimodal data is required to identify the stages of the patients. However, these data are necessary to build a strong perception about the outcomes of survival rate in the very last or intermediate stage. Surprisingly, no such evidence has been identified to date that can show the betterment of the outcomes of cancer patients and that also without any side effects of being susceptible or any other diseases. As demonstrated by Suarez-Ibarrola *et al.* (2020) the advancement in every field of the world is occurring with AI or artificial intelligence. It is thus necessary to implement the same in the medical field to get rid of the typically harsh radiation for treating cancer. In the medical field, computational technology stimulates the intellectual process of typical cognitive functioning. It is thus necessary to build a strong base that can provide adequate information about the stages of the cancer that is recurring inside the patient's body. Additionally, adopting computational programming in terms of artificial intelligence in this field would help intelligent entities. The initial stage of tumour is required to be treated so that the malignancy can be stopped. In order to analyse the beginning stage of the tumour, microscopic images of the cross-sectional areas of human tissue is needed. As narrated by Kather *et al.* (2019) deep learning can be used to extract the information about the complex stages of cancer. This model helps to develop the predictions of the outcome directly from the histological images of the cancer cells or tissues. It is thus considered to be the most important and innovative learning process in the medical field that helps in the prediction of the survival rate of cancerous patient. In contrast to this statement, it is justified that the initiation of the treatment in the beginning stage might help to determine the survival rate as the faster the treatment, the more will be the survival rates of cancer patients. Kather *et al.* (2019) have also cited that there are cancer stages that show complexities with increasing life span. However, it also hampers the immune

system and makes the human body ineligible to deal with other common diseases. It is necessary to identify the stage with a proper contrasting image so that the recurrence of cancer cells can be identified. It has been noticed that in many cases even after giving radiation to the body there is no improvement of the diseases. It might be because the origin of the cancerous cells in the body is unknown and cannot be determined by any other technique.

2. METHODS AND MATERIALS

Database

Databases are being used to access the data required to conduct any research. In this study the use of PubMed and Google Scholar has been used to collect the primary quantitative and secondary qualitative data respectively. The collected data are analysed properly in the study in terms of resources to estimate the probable outcome of the study. The databases that have been chosen in this study to collect data have helped to extract the data in an organized way. To conduct any design and modelling techniques, databases show significance as important sources can be collected from the databases to generate the idea of preparing any model [*Refer to appendix 1*].

Inclusion and exclusion criteria

An inclusion and exclusion criterion of this study is based on the fact that there is no article comprising a different language rather than English is included and thus it is excluded. The journal articles that are being selected in this study are not more than 10 years older. That means the articles are being collected from the range of 2011 to 2021. To conduct the thematic analysis, 14 primary quantitative articles have been chosen and qualitative articles are excluded. The keywords that deal with the research topic are only included. The study also includes the features and characteristics of deep learning while the features of other techniques have been excluded.

Data collection

The collection of the data is based on the extraction of secondary sources. However, each secondary source consists of primary quantitative research that has been conducted in this study. As described by Pal (2017) data analysis is a way of obtaining real data by analysing the existing data. The variables and the factors in a systematic review can be determined by using secondary sources in terms of journal articles. In this study, secondary sources have been collected to build themes based on the chosen topic. Data collection has been made by using several databases such as PubMed and Google scholar. It has been taken into account that statistical data is present in every chosen article so that comparative analysis can be drawn. Pal (2017) has also manifested that quantitative data collection mainly focuses on surveys and measurement of the pre-existing data sources.

3. RESULTS

The findings of the study deal with the secondary article that comprises primary quantitative study. In this study a quality review and an axial coding table has been made. The thematic analysis of the study is based on the themes created in terms of axial coding. A Boolean table has also been made that shows the keywords for the study [Refer to appendix 2]. **Figure 1 illustrates the Prisma model for the selected articles.**

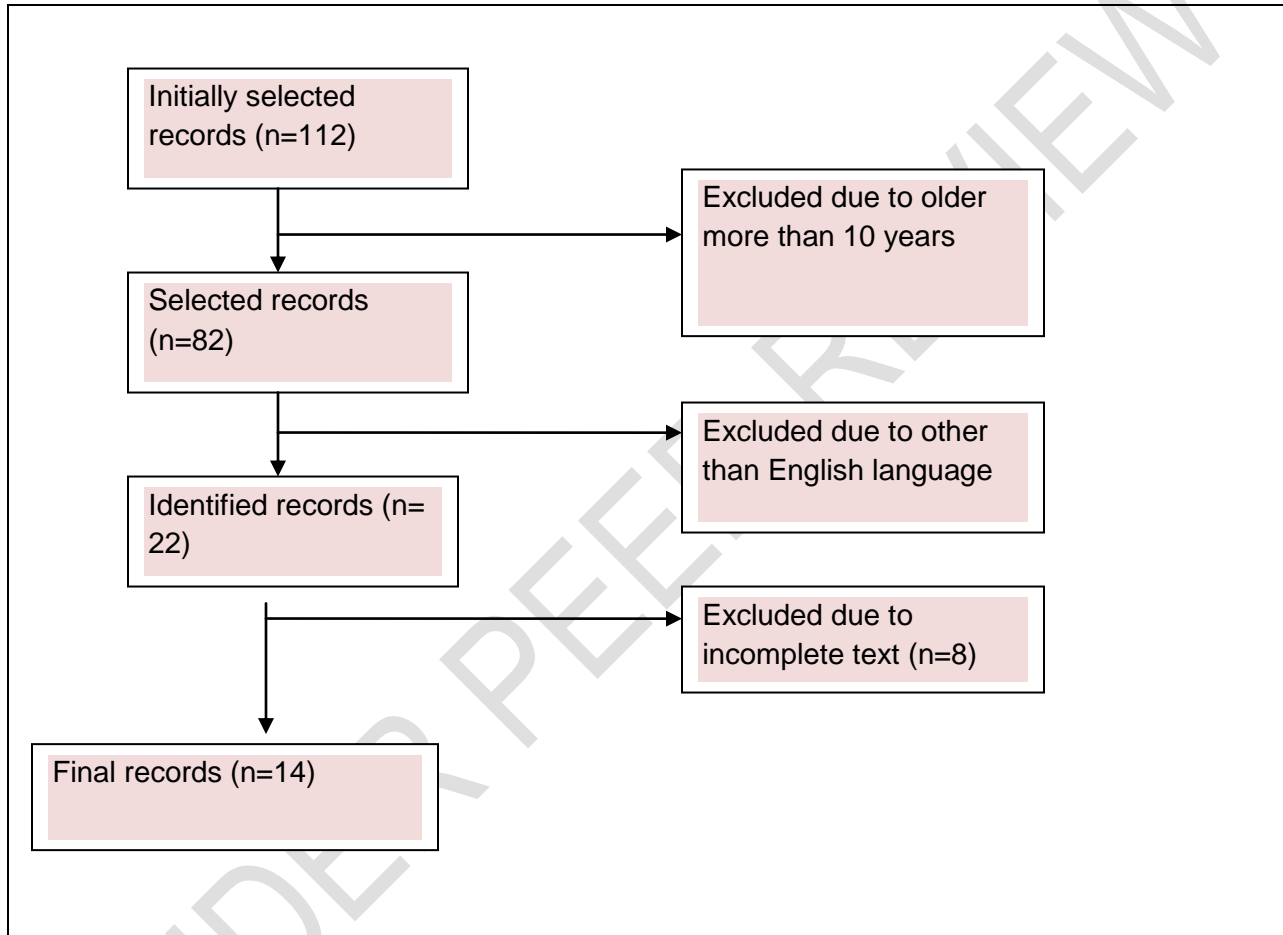


Figure 1: PRISMA diagram

Table 1: Quality review

| Authors | Study design | Accuracy in addressing the issue | Accuracy of outcomes | Proper identification of factors | Sufficiency of subjects to meet the objectives | Confounding of the factors have been done |
|---------|--------------|----------------------------------|----------------------|----------------------------------|--|---|
| | | | | | | |

| | | | | | | |
|-------------------------------|----------------------|--------|--------|-----|-----|-----|
| Munir <i>et al.</i> (2019) | Primary quantitative | High | High | Yes | Yes | Yes |
| Yala <i>al.</i> (2019) | Primary quantitative | High | High | Yes | Yes | Yes |
| Schmauch <i>et al.</i> (2020) | Primary quantitative | Medium | High | Yes | Yes | Yes |
| Smith <i>et al.</i> (2020) | Primary quantitative | High | High | Yes | Yes | Yes |
| Ardila <i>et al.</i> (2019) | Primary quantitative | High | High | Yes | Yes | Yes |
| Metovic <i>et al.</i> (2018) | Primary quantitative | Medium | Medium | Yes | Yes | Yes |
| Brunet <i>al.</i> (2020) | Primary quantitative | Medium | High | Yes | Yes | Yes |
| Zhou <i>et al.</i> (2020) | Primary quantitative | High | High | Yes | Yes | Yes |
| Coudray <i>et al.</i> (2018) | Primary quantitative | High | Medium | Yes | Yes | Yes |

| | | | | | | |
|----------------------------------|----------------------|------|------|-----|-----|-----|
| | | | | | | |
| Avanzo <i>et al.</i> (2020) | Primary quantitative | High | High | Yes | Yes | Yes |
| Lee and Fujita (2020) | Primary quantitative | High | High | Yes | Yes | Yes |
| Xu <i>et al.</i> (2019) | Primary quantitative | High | High | Yes | Yes | Yes |
| Mobadersany <i>et al.</i> (2018) | Primary quantitative | High | High | Yes | Yes | Yes |
| Slavik <i>et al.</i> (2014) | Primary quantitative | High | High | Yes | Yes | Yes |
| Indolia <i>et al.</i> (2018) | Primary quantitative | High | High | Yes | Yes | Yes |
| Kim <i>et al.</i> (2021), | Primary quantitative | High | High | Yes | Yes | Yes |
| Lambin <i>et al.</i> (2017), | Primary quantitative | High | High | Yes | Yes | Yes |
| Chaunzwa <i>et</i> | Primary | High | High | Yes | Yes | Yes |

| | | | | | | |
|-------------------|--------------|--|--|--|--|--|
| <i>al. (2021)</i> | quantitative | | | | | |
|-------------------|--------------|--|--|--|--|--|

Table 2: Axial coding

| Authors | Codes | Sub-theme |
|---|--|---|
| <p>Munir <i>et al.</i> (2019)</p> <p>Yala <i>al.</i> (2019)</p> <p>Kim <i>et al.</i> (2021),</p> | <p>Deep learning analysis</p> <p>Diagnosis</p> | <p>Analysis of diagnosis with the help of deep learning</p> <p>Uses of deep learning</p> <p>Diagnosis based on deep learning</p> |
| <p>Schmauch <i>et al.</i> (2020)</p> <p>Smith <i>et al.</i> (2020)</p> | <p>Gene expression</p> <p>Phenotype</p> | <p>Understanding of phenotype prediction from gene expression with the help of deep learning</p> <p>Nature of phenotype of the cancer cells</p> <p>Gene expression or transcriptomics</p> |
| <p>Ardila <i>et al.</i> (2019)</p> <p>Metovic <i>et al.</i> (2018)</p> <p>Chaunzwa <i>et al.</i> (2021)</p> | <p>Screening</p> <p>Cancer</p> | <p>Evaluating the role of deep learning in the screening of cancer stages</p> <p>Screening of cancer cells</p> <p>Screening of stages in lung carcinoma</p> |
| | | |

| | | |
|--|--|---|
| | | |
| <p>Brunet <i>et al.</i> (2020)</p> <p>Zhou <i>et al.</i> (2020)</p> <p>Indolia <i>et al.</i> (2018)</p> | <p>Deep learning</p> <p>Activators</p> | <p>Analysing the activators of cancer with the help of introduction deep learning</p> <p>Identification of the activators of cancer cells</p> <p>Role of deep learning in the determination of activators</p> |
| | | |
| <p>Coudrayet <i>al.</i> (2018)</p> <p>Avanzoet <i>al.</i> (2020)</p> <p>Lambin <i>et al.</i> (2017),</p> | <p>Mutagenesis</p> <p>Cancerous cells</p> | <p>Identification of the mutagenesis occurring in the cancer cells with the help of deep learning</p> <p>Stages of cancerous cell</p> |
| <p>Lee and Fujita (2020)</p> <p>Xu <i>et al.</i> (2019)</p> | <p>Medical imaging</p> | <p>Justification of the uses of deep learning in obtaining medical images</p> <p>Obtaining a medical images</p> <p>Role of deep learning in giving a clear overview of the image</p> |
| <p>Mobadersanyet <i>al.</i> (2018)</p> <p>Slaviket <i>al.</i> (2014)</p> | <p>Histopathology</p> <p>Deep learning</p> | <p>Understanding of the predictions made in the basis of histopathological observation of cancer cell by deep learning</p> |

Critical analysis of the cancer diagnosis using deep learning

The diagnosis of cancer has become a worldwide concern in present days. It has been observed that there is no such advanced procedure that can determine the initial stage of the cancer that is a tumour. The diagnosis feature thus needed to be improved so that a proper outcome can be achieved in terms of survival rate. As described by Munir *et al.* (2019) raw images of cancer stages might not give a proper solution for the analysis of the stages. On the other hand, it has also been noted that the other processes of obtaining images of cancer cells in terms of pre-processing might create immense noise. It is thus necessary to reduce the noise along with increasing contrast of the images so that they can be analyzed properly. On the contrary, Yala *et al.* (2019) have elaborated that models have always supported risk-adjusted screening. It means the screening with higher radiation might create disruptions in other somatic cells of the body rather than the cancerous cells. **Figure 2 shows that the results obtained by Yala *et al.* (2019) with deep learning models for full field mammograms are better at identifying breast cancer risks over other TC or Tyrer-Cuzick tests is more significant.** Somatic cells might overcome the issues caused to them however; the radiation might also cause mutations in the chromosomal structure of the somatic cells. In this aspect it can be stated that there must be the implementation of drug-related prices so that the patient gets painless treatment. Munir *et al.* (2019) have also discussed that to identify the mutation in the cancerous cells and to resist such mutation by conducting different research requires a handful of proper resources. The resources such as images of cancerous cells and tissues can be obtained from the deep learning model.

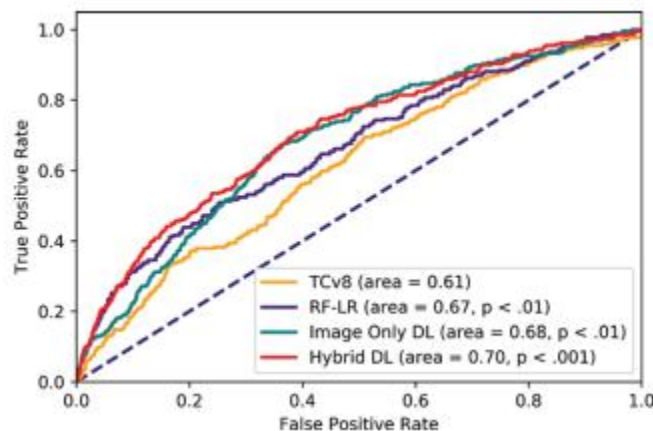


Figure 2. Curve showing true and false-positive rates of the models used in the diagnosis procedure of breast cancer by analysing the singles released in the deep learning process

(Source: Yala *et al.* 2019)

It can be stated that DL on improvement can help in working as a better machine learning approach that would help to deal with clear images of the cancer as a diagnosis procedure. **Figure 2** presents a curve

for false positive rate in the diagnosis of breast cancer. As per the view of Kim *et al.* (2021), deep learning-based computer-aided diagnosis can help to detect the breast masses and reduce the rate of giving false-positive rates. This, in turn, can be stated that the DL gives a proper accuracy of cancerous growth rate. Figure 2 displays that application of deep learning is very useful to increase the accuracy of the results by reducing the false-positive rates considering some non-malicious activities due to improper growth rate. Hence, it can be stated that DL on improvement can help to obtain a much clearer image along with significant value of the histological pathways. The more significant value suggests that there is an accuracy of occurring higher chances of cancer.

Critical analysis of the deep representation of learning on phenotype prediction from gene expression

Determination of the prediction that can confidently give rise to the gene expression in the body due to cancer can be done by deep learning process. As illustrated by Smith *et al.* (2020) that there are various traditional as well as scientific processes that can analyse the expression mechanism in the patients suffering from cancer. It has been analyzed that machine learning methods often provide a clear cut idea about the physiological and morphological changes that are occurring in the cells due to cancer. Smith *et al.* (2020) have also discussed the fact that genes are rarely active and act in isolation. It might be because along with the isolation procedure the physiological changes also happen in the genes in terms of its regulation.

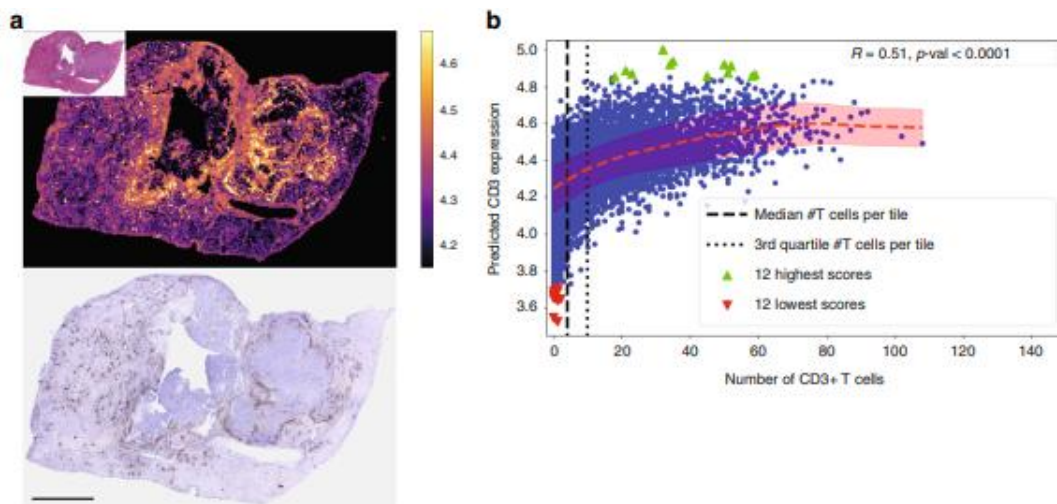


Figure 3. Virtual specialization of the expression of CD3 and CD20 genes that are responsible for giving rise to cancer

(Source: Schmauch *et al.* 2020)

The above-mentioned figure as opined by Schmauch *et al.* (2020) justifies the presence and expression of CD3 and CD20 genes that are being percent during the formation of cancer. In a word, it can be stated that these genes are responsible for giving rise to cancer. **Figure 3** signifies that there must be an expression of genes that can show the growth of cancerous cells. As per the image the details given are hemotoxylin and eosin stained RNA-Seq data for 28 different types of cancer. This kind of image in neural network analysis allows transcriptome representation from images as well as improving predictive analysis for cancer occurrence. CD3 and CD20 genes produce the protein that acts as a marker for the lymphomas. Additionally, it is necessary to include deep learning in terms of evaluation of such genes because these genes are not often found in the traditional machine learning diagnosis pattern. Lee *et al.* (2017) have cited that these two genes are present in the B and T lineage lymphomas. However, there are no such standards immunohistochemistry tests that can help to find out the staining pattern of the CD3 and CD20 genes as markers. As these two genes are not visible in traditional diagnosis methods, hence staining them without advanced imaging system becomes difficult. (et al. (2020) stated that deep learning models need to be evaluated in terms of RNA sequencing so that expression of these cytokines can be observed.

Critical understanding of screening of lung cancer with three dimensional deep learning and radiomics

Avanzo *et al.* (2020) have demonstrated the fact that the imaging of the cancerous cells has become a crucial challenge in today's world. However, no such evidence has been found that can help in the formation of better images of cancerous cells. On the other hand, being smaller in size the lung nodules do not get properly imaged on the verge of the detection of lung cancer. Additionally, it has been noted that the accuracy of detection of tumour lesions can be improved significantly by implementing the AI interfaces. Avanzo *et al.* (2020) have also demonstrated the idea of histological results that are being obtained from biopsy. However, radiomics features by describing the intensity, shape along with heterogeneity of the tumour. As per the view of Coudray *et al.* (2018), it has been noted that the American Cancer Society and the Cancer statistic Centre has reported over 150,000 patients suffering from lung carcinoma. It has been noted that there are many diagnosis methods that can be used to determine lung cancer yet deep learning is the most accurate one that determines the histopathology of lung cancer by obtaining its proper image.

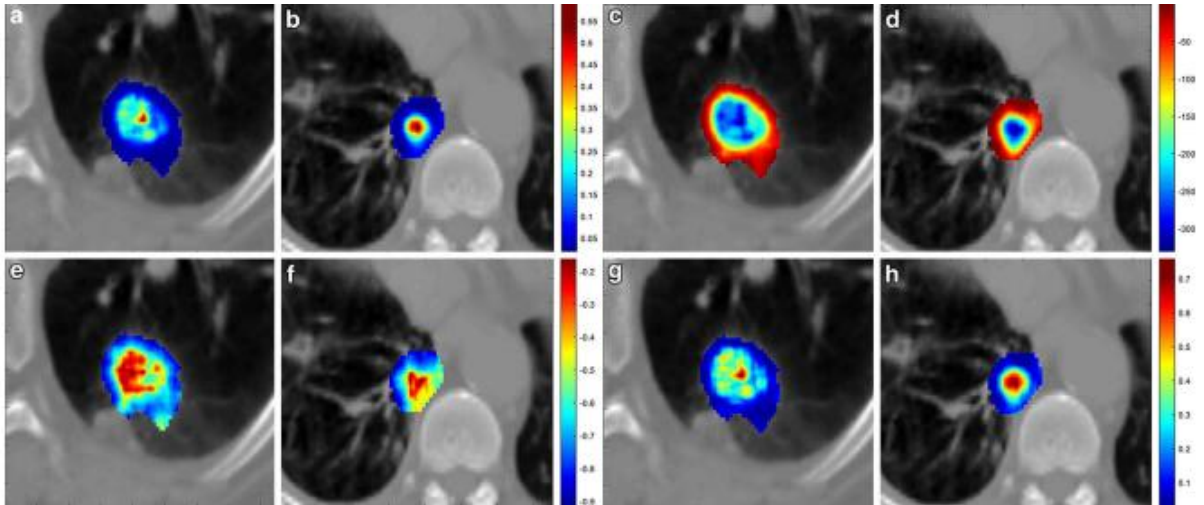


Figure 4. Image obtained by radiomics to uncover the tumoral patterns and calculated from CT scan to analyse the masses of the cancer cells

(Source: Avanzo *et al.* 2020)

Avanzo *et al.* (2020) have elaborated that chromosomal rearrangements can lead to gene fusion. Therefore, radiomics helps to determine the rearrangement of anaplastic lymphoma kinase or ALK that potentially makes connotative as a multi-target. **Figure 4 demonstrates images of radiomics test to analyze the fragments and cancer affected areas inside the lungs. Every tumour has its own pattern that is necessary to be identified with the help of diagnostics tests. Such images allows prediction of lung carcinoma. The differentiating pattern might fail with the naked eye in terms of observation. As per the view of Lambin *et al.* (2017), it has been noted that that the process of radiomics deals with computer-aided diagnosis. The reason behind including it into the deep learning methods is also related to the computer-aided process.**

Critical justification of activators of cancer cells that is synthetic by deep learning

As illustrated by Brunet *et al.* (2020) it has been noted that convolution neural channels are the important method to recognize the choice of the image. It has been noted that in such imaging a protein activator has been identified. CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) models are thus used to identify chemokines receptor CXCR4 modulators that consist of high accuracy.

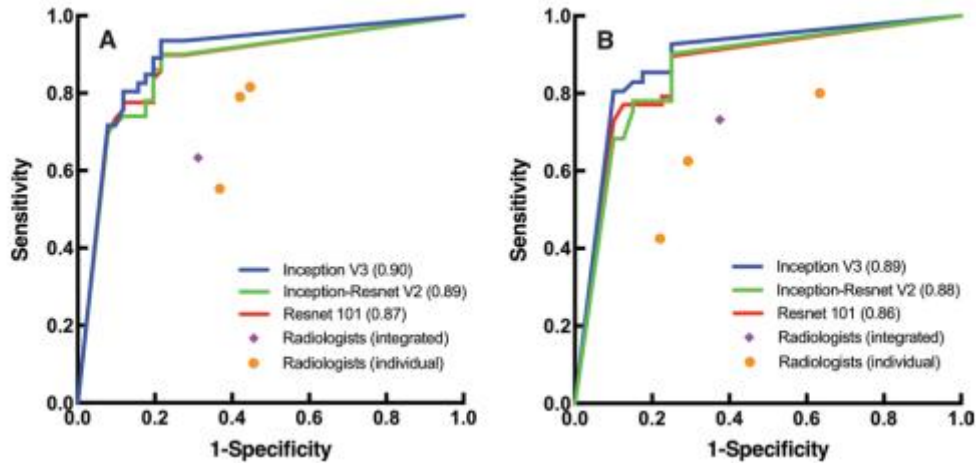


Figure 5. Receiver operating the curves by deep learning convolution models to process an image of signals received by the cancerous cells

(Source: Zhou *et al.* 2020)

Figure 5 highlights that there must be a curve that can show high accuracy in the growth rate of the cancerous cells. Here the images are to be processed based on the signals from the cancer cells. As demonstrated by Indolia *et al.* (2018), it is noticed that CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) helps to overcome the limitations of traditional learning of diagnosis. Additionally, it has been noted that CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) helps to give the proper images of the receptors or activators of the cancerous cells with specificity in the growth rate or in the curve. The above mentioned figure shows that the initial point of the activators can be observed at 0.89, that means it gets stable in that place of sensitivity. However, these growth rates can only be determined by the markers that are required to be implemented by the deep learning methods. The reason behind using convolution models is because it shows a clear description of the neural networks that are necessary in the deep learning models. In deep learning the CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) can help to feature the input with proper essence with spatial features. Brunet *et al.* (2020) has also cited that there are channels that consist of G-protein coupled receptors and are observed with the help of deep learning princesses. On the other hand, Zhou *et al.* (2020) have argued on the fact that deep learning also has exhibited excellence in the performances of imaging the quality image in breast cancer. Additionally, in the present-day convolution neural networks or CNN (CONVOLUTIONAL NEURAL NETWORK MODELS)s act as the most important deep learning process in generating medical images. However, identifying CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) images is indeed difficult. It needs the most innovative approach so that the images can be analysed properly. Brunet *et al.* (2020) has discussed the receptors that are present in the cancerous cells. It is thus necessary to determine such receptors with a proper imaging system to get the idea of molecular morphology and their survival rates.

Critical analysis of the mutagenesis occurring in small cell lung cancer by observing histopathological images via deep learning

There are several changes that occur in the case of morphological changes of cancerous cells. As discussed by Ardila *et al.* (2019) it has been noted that low doses of computed tomography could not properly establish the changes of the cancerous cells. In contrast to this statement, it can be stated that accuracy of the cancerous cell and mutagenesis can only be achieved by implementing deep learning methods. Deep learning method, on the other hand, by analysing the initial stage helps the patients to survive with proper treatment methods. Ardila *et al.* (2019), have also cited that the deep learning model provides an additional approach of enlightening the world by detecting the mutagen of the cancerous cells.

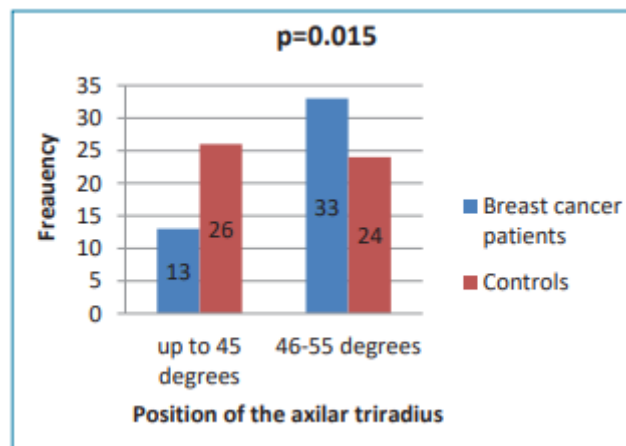


Figure 6. Accuracy in determining the position of axillary triradius in the palm because of dermatoglyphics

(Source: Metovic *et al.* 2018)

Metovic *et al.* (2018) have cited that the positioning of the dermatoglyphics are the determining properties of breast cancers in women. It can be stated that the positioning of dermatoglyphics has been incurred by a deep learning process. In some cases the changes in the hand structure such as the dermatoglyphics might give an idea of such diseases. **Figure 6 demonstrates the occurrence of leverages in females of age 45-55 is much more than that of the controls ones. The controls are those who have been with proper precautions after the diagnosis test with deep learning. As illustrated by Chaunzwa *et al.* (2021), it has been noted that tumor histology is an important predictor that can show the growth of cancers in terms of symptoms. However, in the case of breast cancers or even lung cancers, the symptoms can be seen as dermatoglyphics. That leads to the occurrence of leverages in the palm. These leverages or dermatoglyphics can be observed with the help of deep learning. Chaunzwa *et al.* (2021) has also cited that the minimum biopsy fails to analyse the smaller cell lung cancers, it is thus necessary to implement**

computer-aided diagnosis. The symptoms such as dermatoglyphics are also captured in computer-aided techniques such as deep learning.

Critical evaluation of deep learning in obtaining medical images

Medical imaging is an important tool in the determination of the images. As per the view of Lee and Fujita (2020) it has been noted that there are several aspects to obtaining a medical image for the determination of the diseases. On the other hand, it has been noted that conventional machine learning spark or deep learning has become more advanced in comparison to other magnetic resonance obtaining images or any other radiology method. Lee and Fujita (2020) have also discussed about the fact that the effectiveness of the description of the image depends on the domain expertise of CAD learning model.

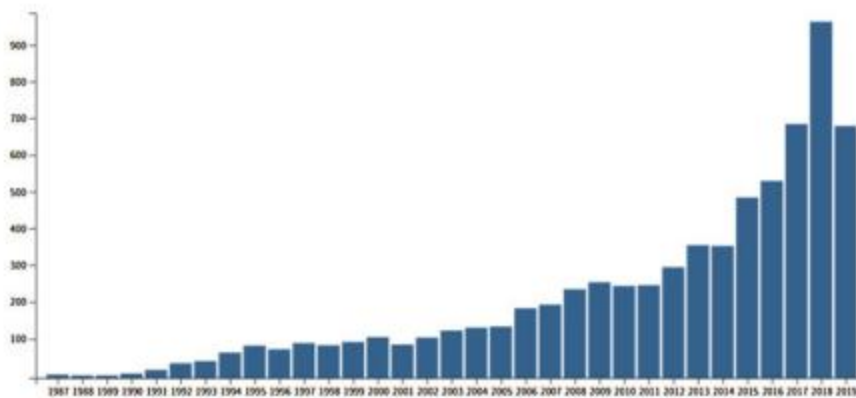


Figure 7. Data records suggests the result of search engine of machine or deep learning journals that shows the interest of people in the deep learning methods

(Source: Lee and Fujita, 2020)

Xu *et al.* (2019) have discussed the fact that as tumours are continuously enveloping it is necessary to improve the imaging effects too. To obtain a healthy image of the cancerous cell, it is necessary to implement advancements. **Figure 7** displays that application of deep learning methods for image processing studies are increasing faster in recent years. The advancement would be in terms of the improvement of the techniques of image obtaining of cancerous cells and tissues. However, there are many other diagnostic methods that can also create issues in the determination of the mutagenesis in the cancerous cells. The records are based on the search engines of the patients. Then advanced technologies in the medical field have been emerging to a greater extent and it has generated the trust and interest of the patients. The deep learning process along with artificial intelligence can be brought into collaboration and it can help to generate better numbers of images with proper accuracy.

Critical understanding of the predictions that are made by observing outcomes from histology using genomic convolution

The predictions are always made on the basis of the images obtained in the diagnostic methods. However, to understand the initial stage of cancer it is necessary to determine the formation of tumours so that the treatment can proceed with proper pathways. As per the view of Mobadersany *et al.* (2018), it has been noted that cancer histology reflects the molecular progression in the body along with changes in the cell structure.

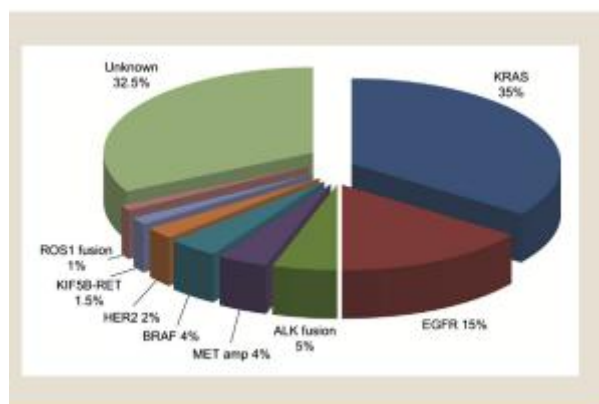


Figure 8. Analysis of oncogenes drivers in the cancerous cells that gives rise to cancer

(Source: Slavik *et al.* 2014)

Figure 8 depicts the activators of cancerous cells such as RAS, ROS1 and many more that are responsible for giving rise to cancers. However, the effectiveness of the drugs can only suppress the activation of such factors and not completely remove the issues created by these factors. Slavik *et al.* (2014) has also cited that there are 32.5% of the activators that are responsible for giving rise to cancers. Identification of the cancerous factor with the help of histology has become an important tool. On the contrary, Slavik *et al.* (2014) have evaluated that where drugs are available, the predictions can be made on the basis of survival rate. However, the effectiveness of drugs and other radiations can only be achieved by determining the cancer genomics at the very initial stage. The genomic evolution of cancerous cells is indeed difficult to analyse. Because it has been observed that the mutations that are found in the cancerous cells are never be able to repair. Mobadersany *et al.* (2018) have illustrated that the histological structure of the tissues are also responsible for determination of the cancerous cells. However, a tumor formation might give an overview of turning into malignant tumours and cancer rapidly. On the other hand, as cancerous cells are immortal they might show several variations without getting damaged even after getting radiation.

4. DISCUSSION

The entire study deals with the fact that there are many ways to generate imaging process. However, the only applicable and innovative process is the deep learning process. As per the view of Zhou *et al.* (2020) it has been noted that deep learning helps in building two or even three dimensional images of cancer cells. It is thus expected that in the diagnostic process, this model can be a promising approach. In the topic it has been mentioned that deep learning is the method can determine the survival rate. It might be because it deals with analysis of complex images of neural networks. It also determines the stage of the cancer. The results states that axillary triradius can be determined by using deep learning process and it is the symptom of breast cancer. Hence, the treatment of breast cancer can be initiated at an early stage. On the other hand, it has been noticed that there are protein activators that get expressed along with the raising of cancer. As stated by Brunet *et al.* (2020) it has been noticed CXCR4 modulators are such protein kinase that gets expressed in the presence of cancerous cells. The size of protein activators and their receptors on the cancerous cells are indeed smaller. Indolia *et al.* (2018), have cited that the CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) is also used with the help of deep learning process that gives a clear network about the images of tumor cells. However, those can be determined only by analysing the accurate imaging of the cancerous cells. As per the view of Avanzo *et al.* (2020) rearrangement of the chromosomes also helps to determine the presence of cancer. From the findings it has been evaluated that oncogenes driver can be predicted through deep learning model. In corresponds to the aim, it can be stated that the characteristics of the deep learning method suggest that it is an advanced process to predict the cancerous model. In respect to the objective of the study also helps to understand the modification and advancements of the deep learning process in the medical field. From the results, it can thus be determined that deep learning model has become the most important topic for research for predicting cancerous model as the graphical data of the search engine suggest the advantages. One of its important features is the determination of histological images. The prediction of histological models thus can be made using deep learning method. However, the rearrangements can only be evaluated by sharpening of the image obtained by deep learning. As the cancerous cells are immortal, it is thus necessary to identify them at the very initial stage of their development. As per the view of Ardila *et al.* (2019) deep learning models are an essential part that can even detect the smallest mutation in the cancerous cells. It is thus necessary to evaluate the approaches of cancer detection in terms of imaging so that the proper diagnosis method can be reached. In cancerous tissues or cells, the main thing that is to be determined is the molecular and morphological changes occurring in the cell. Chaunzwa *et al.* (2021), it is necessary to predict the size and nature of the tumor in order to get a beneficial analysis of the cancerous tissues and cells. The results of the research thus evaluate the need of deep learning that can be acted with the artificial intelligence and gives an innovative result in image. In corresponds to the objectives, the finding suggests that it has several recent advancements of the deep learning process that are needed to determine the molecular morphology of the cancerous predictive model. The survival rate thus can be determined by analyzing the initial impact of the cancerous cell.

5. CONCLUSION

The study concludes the benefits of the deep learning and imaging process that help to determine the changes in the cancerous cells. The need for deep learning can thus be evaluated from the study. The imaging process thus obtained in the research needs to be brought into the future world so that an advanced learning process in the medical field can be induced. However, along with deep learning, radiomics can also be introduced as an improved method for the deep learning process. In a conclusive term it can be stated that the deep learning model can contribute in the early determination of the cancer stages. The results suggests that in very cancerous stages it can be proved to be helpful as it determines the stages of the tissue consisting. It has been evaluated in the study that the imaging process can be generated properly by obtaining the gene expression and activity of protein kinase. In short it can be concluded that there must be incorporation of the improved deep learning model so that the survival rate of the cancerous patients can be estimated. The initial stage can thus be analysed and treated with proper diagnosis as well. It can thus be concluded that when there is a need of opting for cancerous diagnosis method one must opt for the deep learning event. In a nutshell it is concluded that deep learning has become the most important model in the field of cancer detection and it is the only thing that can bring advancement.

ETHICAL CONSIDERATION

The technical consideration of this systematic review can be justified as the maintenance of proper and authentic data. The validation of every research depends on the ethical issues that are being considered properly. Very sources that are being selected from the databases are acknowledged in this study. The extracted data are being analyzed properly and it has been observed that every source is authentic and did not cause any harm to the moral or social culture of humans. It has been kept into account that there is not any wrong symbol or sign being used in this study that can give rise to ethical challenges of the study. For ensuring the ethical concern, throughout the research, every author's names have been acknowledged. It has been kept into account that no single belief of any community is harmed in the study. The data that is included in the study as resources are also not the copyright of anything.

FUTURE SCOPE

The study opens up pathways of using the improved technique of deep learning in the medical field. It has been noted that the study enlightens all the benefits of deep learning that are being evaluated in different clinics and path labs for obtaining accurate and contrasting images of cancerous cells. The necessary implementations that are required to be made for getting prior images of cancerous cells are also being evaluated in the study. The study would also enlighten the necessity of opting for a diagnostic method at the very beginning stage of the cancer. It has been noted that many people considered tumours not to be

cancerous. However, it is indeed a wrong perception and its need to be modified properly the study will also open up the broad ideas of deep learning that are necessary for getting proper images of cancerous cells. The developmental stages of a cancerous cell in terms of molecular morphology have also been determined in the study.

REFERENCES

Additional resources

Cheerla, A. and Gevaert, O., 2019. Deep learning with multimodal representation for pancancer prognosis prediction. *Bioinformatics*, 35(14), pp.i446-i454.

Chen, M., Zhang, B., Topatana, W., Cao, J., Zhu, H., Juengpanich, S., Mao, Q., Yu, H. and Cai, X., 2020. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *NPJ precision oncology*, 4(1), pp.1-7.

Karpiński, T. M. (2019). Role of oral microbiota in cancer development. *Microorganisms*, 7(1), 20.

Kather, J.N., Krisam, J., Charoentong, P., Luedde, T., Herpel, E., Weis, C.A., Gaiser, T., Marx, A., Valous, N.A., Ferber, D. and Jansen, L., 2019. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1), p.e1002730.

Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. *applied sciences*, 9(20), 4396.

Pal, A., 2017. Quantitative Data Analysis and Representation. *International Journal of Engineering Science and Computing*, 7(3), pp.4853-4856.

Rathore, S., Iftikhar, M.A. and Mourelatos, Z., 2019. Prediction of overall survival and molecular markers in gliomas via analysis of digital pathology images using deep learning. *arXiv preprint arXiv:1909.09124*.

Suarez-Ibarrola, R., Hein, S., Reis, G., Gratzke, C. and Miernik, A., 2020. Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World journal of urology*, 38(10), pp.2329-2347.

Selected resources

Ardila, D., Kiraly, A.P., Bharadwaj, S., Choi, B., Reicher, J.J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G. and Naidich, D.P., 2019. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature medicine*, 25(6), pp.954-961.

Avanzo, M., Stancanello, J., Pirrone, G. and Sartor, G., 2020. Radiomics and deep learning in lung cancer. *Strahlentherapie und Onkologie*, pp.1-9.

Bruns, D., Gawehn, E., Kumar, K.S., Schneider, P., Baumgartner, M. and Schneider, G., 2020. Identification of synthetic activators of cancer cell migration by hybrid deep learning. *ChemBioChem*, 21(4), pp.500-507.

Chaunzwa, T.L., Hosny, A., Xu, Y., Shafer, A., Diao, N., Lanuti, M., Christiani, D.C., Mak, R.H. and Aerts, H.J., 2021. Deep learning classification of lung cancer histology using CT images. *Scientific reports*, 11(1), pp.1-12.

Coudray, N., Ocampo, P.S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A.L., Razavian, N. and Tsirigos, A., 2018. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10), pp.1559-1567.

India, S., Goswami, A.K., Mishra, S.P. and Asopa, P., 2018. Conceptual understanding of convolutional neural network-a deep learning approach. *Procedia computer science*, 132, pp.679-688.

Kim, S.Y., Choi, Y., Kim, E.K. *et al.* Deep learning-based computer-aided diagnosis in screening breast ultrasound to reduce false-positive diagnoses. *Sci Rep* 11, 395 (2021).

Lambin, P., Leijenaar, R.T., Deist, T.M., Peerlings, J., De Jong, E.E., Van Timmeren, J., Sanduleanu, S., Larue, R.T., Even, A.J., Jochems, A. and van Wijk, Y., 2017. Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*, 14(12), pp.749-762.

Lee, E. J., Kim, M., Kim, H. S., Kang, H. J., Kim, H. J., Min, S. K., & Lee, Y. K. (2017). CD3 and CD20 Immunohistochemical Staining Patterns of Bone Marrow-Infiltrating Malignant Lymphoma Cells. *Annals of Clinical & Laboratory Science*, 47(2), 136-143.

Lee, G. and Fujita, H. eds., 2020. *Deep learning in medical image analysis: challenges and applications* (Vol. 1213). Springer Nature.

Metovic, A., Musanovic, J., Alicelebic, S., Pepic, E., Sljuka, S. and Mulic, M., 2018. Predictive Analysis of Palmar Dermatoglyphics in Patients with Breast Cancer for Small Bosnian-Herzegovinian Population. *medical archives*, 72(5), p.357.

Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D.A., Barnholtz-Sloan, J.S., Vega, J.E.V., Brat, D.J. and Cooper, L.A., 2018. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proceedings of the National Academy of Sciences*, 115(13), pp.E2970-E2979.

Munir, K., Elahi, H., Ayub, A., Frezza, F. and Rizzi, A., 2019. Cancer diagnosis using deep learning: a bibliographic review. *Cancers*, 11(9), p.1235.

Schmauch, B., Romagnoni, A., Pronier, E., Saillard, C., Maillé, P., Calderaro, J., Kamoun, A., Sefta, M., Toldo, S., Zaslavskiy, M. and Clozel, T., 2020. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nature communications*, 11(1), pp.1-15.

Slavik, T., Asselah, F., Fakhruddin, N., El Khodary, A., Torjman, F., Anis, E., Quinn, M., Khankan, A. and Kerr, K.M., 2014. Diagnosis and predictive molecular analysis of non-small-Cell lung Cancer in the Africa-Middle East region: challenges and strategies for improvement. *Clinical lung cancer*, 15(6), pp.398-404.

Smith, A. M., Walsh, J. R., Long, J., Davis, C. B., Henstock, P., Hodge, M. R., Maciejewski, M., Mu, X. J., Ra, S., Zhao, S., Ziemek, D., & Fisher, C. K. (2020). Standard machine learning approaches outperform deep representation learning on phenotype prediction from transcriptomics data. *BMC bioinformatics*, 21(1), 119.

Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., Mak, R.H. and Aerts, H.J., 2019. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clinical Cancer Research*, 25(11), pp.3266-3275.

Yala, A., Lehman, C., Schuster, T., Portnoi, T. and Barzilay, R., 2019. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*, 292(1), pp.60-66.

Zhou, L.Q., Wu, X.L., Huang, S.Y., Wu, G.G., Ye, H.R., Wei, Q., Bao, L.Y., Deng, Y.B., Li, X.R., Cui, X.W. and Dietrich, C.F., 2020. Lymph node metastasis prediction from primary breast cancer US images using deep learning. *Radiology*, 294(1), pp.19-28.

UNDER PEE

APPENDICES

Appendix 1: Databases

PubMed.gov search results for "predictive analysis" in cancer. The search results page displays 90 results. The top two results are:

- Low lncRNA ZNF385D-AS2 expression and its prognostic significance in liver cancer.**
Zhang Z, Wang S, Liu Y, Meng Z, Chen F.
Oncol Rep. 2019 Sep;42(3):1110-1124. doi: 10.3892/or.2019.7238. Epub 2019 Jul 16.
PMID: 31322274 [Free PMC article.](#)
Cox **analysis** identified ZNF385D-AS2 low-expression as an independent prognostic variable (AUC=0.594) for overall survival in liver **cancer** patients. ...In conclusion, our results suggests that low expression of ZNF385D-AS2 is **predictive** of a poor prognosis of ...
- Diagnosis and predictive molecular analysis of non-small-cell lung cancer in the Africa-Middle East region: challenges and strategies for improvement.**
Slavik T, Asselah F, Fakhruddin N, El Khodary A, Torjman F, Anis E, Quinn M, Khankan A, Kerr KM.
Clin Lung Cancer. 2014 Nov;15(6):398-404. doi: 10.1016/j.clcc.2014.06.005. Epub 2014 Jun 21.
PMID: 25043310

RESULTS BY YEAR: 1984 to 2021. Text Availability: Abstract, Free full text.

Appendix 2: Boolean table

| Keywords | And/or | Keywords |
|---|---------------|------------------|
| Deep learning | And | Machine learning |
| Mutagenesis | And | Cancer cells |
| Molecular testing | And | Carcinoma |
| CT scan | And | Radiomics |
| CNN (CONVOLUTIONAL NEURAL NETWORK MODELS) | And | X-ray |
| Dermatoglyphics | And | Breast cancer |
| Gene expression | And | Genomics |