Advanced Object Detection in Bio-Medical X-Ray Images for Anomaly Detection and Recognition

Garv Modwel, Amity University, India

 https://orcid.org/0000-0001-6370-1670

Anu Mehra, Amity University, India

Nitin Rakesh, Sharda University, India

K. K. Mishra, MNNIT Allhabad, India

ABSTRACT

The human vision system is mimicked in the format of videos and images in the area of computer vision. As humans can process their memories, likewise video and images can be processed and perceptive with the help of computer vision technology. There is a broad range of fields that have great speculation and concepts building in the area of application of computer vision, which includes automobile, biomedical, space research, etc. The case study in this manuscript enlightens one about the innovation and future scope possibilities that can start a new era in the biomedical image-processing sector. A pre-surgical investigation can be perused with the help of the proposed technology that will enable the doctors to analyses the situations with deeper insight. There are different types of biomedical imaging such as magnetic resonance imaging (MRI), computerized tomographic (CT) scan, x-ray imaging. The focused arena of the proposed research is x-ray imaging in this subset. As it is always error-prone to do an eyeball check for a human when it comes to the detailing. The same applied to doctors. Subsequently, they need different equipment for related technologies. The methodology proposed in this manuscript analyses the details that may be missed by an expert doctor. The input to the algorithm is the image in the format of x-ray imaging; eventually, the output of the process is a label on the corresponding objects in the test image. The tool used in the process also mimics the human brain neuron system. The proposed method uses a convolutional neural network to decide on the labels on the objects for which it interprets the image. After some pre-processing the x-ray images, the neural network receives the input to achieve an efficient performance. The result analysis is done that gives a considerable performance in terms of confusion factor that is represented in terms of percentage. At the end of the narration of the manuscript, future possibilities are being traces out to the limelight to conduct further research.

KEYWORDS

Bio-Medical Application, Convolutional Neural Network, Image Processing, Image Segmentation, Object Detection, Object Recognition, X-Ray Imaging

INTRODUCTION AND LITERATURE REVIEW

Medical technology is growing with improved technological advancement, driving to an advanced level of treatment and diagnostic in the twenty-first century. Day to day our food behavior is changing. The human body meets an enormous amount of chemical substances. This is resulting in different categories of infections, diseases, and syndromes. The current era demands to develop treatments that are efficient enough. However, efficiency is not only the key when it comes to the time constraints for the treatment. When technology fits properly into the medical sector, it can lead to an optimal solution for medical treatment.

In ancient days' technology was not much developed to detect different happenings in the human body with the improvement of technology now a day we are capable enough to reproduce an imagery system in different patterns. There are different types of imagery system that are MRI imaging, CT scan, X-Ray imaging, ultra-sonic imaging. The proposed methodology deals with the development of object detection techniques in X-Ray imaging.

X-ray is the light spectrums that has wavelength ranging from 0.01 to 10 nanometer and corresponding frequency range varies from $3*10^{16}$ to $3*10^{19}$ HZ. The property of the light is such that it passes through less dense material and obstructed through highly dense materials. The dense material from which the X-Ray is being reflected is captured in the sensors to form the corresponding image of the object. This principle of an X-Ray is the feature that drives this research in this manuscript. When patience is subject to the X-ray bone reflects the light that is sensed by the sensor. In the same way if the body contains a metallic part or any tumor that is, growing inside the body also shows the same characteristics as these are also dense objects. The denser the object the more it will reflect the X-Ray beam. The objects that are reflecting the X-Ray is represented as a light spot in the X-Ray imaging. Object through which the light passes is represented as a dark spot.

CT scan is also a 3-dimensional extension of X-Ray imaging. The patience is allowed to lay down on a bed and the complete body is scanned 360 degrees to generate the 3-dimensional image. These three-dimensional images are generated through digital geometry processing which combines the angle at which the images are taken and the corresponding images to form a 3-dimensional imaging.

When it comes to X-Ray image, it is very difficult to detect small objects with an eyeball check. Sometimes doctors need to attain many patients so they have very little time for a particular patience. The focus of the research in this manuscript is to detect important anomaly in X-Ray imaging. This will result in less time consumption in the detection of metallic objects and tumors in the human body. This will also enable the doctors to cure and advise many patients compared to an eyeball check on the X-Ray imaging alone. A small mistake in the detection of these objects may cause a serious issue to the patience's body we cannot rely on the doctor eyeball check. In some cases, it is hard to detect minor abnormalities. As the computer, vision technique uses pixel variation technique it can go to the pixel level of the image and analyze it. This is very difficult for an experienced doctor even. Suppose there is more than one part of the body that is infected by, a metal then even a minor variation is being detected a highlighted. Subsequently, different doctors to do a root cause analysis of the anomaly can analyze these highlighted areas.

When body cells start growing at a faster rate compared to their usual metabolic performance the density of cells increases. This is the main reason behind the tumor. The early stage of cancer starts with tumors. If somehow we able to detect there is a tumor present in a patience body at this stage it can be detected and can be cured. About X-Ray as tumors are very dense rays cannot pass through. This results in the detection of a tumor. These tumors are less dense compared to bones in the human body. The denser the object the brighter the pixels of the x-ray image will be. If a cell is a tumor, it will be very difficult to detect it in X-Ray with an eyeball check. Therefore, if we subject these images to a computer vision system it can analyze a pixel level. At the pixel level, it will be easier to analyze these variations. If we can categories, the pixels according to the tumor levels in the patience body we will be able to detect which kind of tumor currently exists in the patience body.

This can be a future scope of this research for the researchers who are working on the field. Once the tumor detected not it comes to the doctor's decision on how to deal with such types of tumors.

When it comes to stones detection and metallic parts in the human body, it becomes easier for the computer vision system to detect. Usually, stones are less dense compared to metallic components. Along with that, metallic components have reflection capabilities compared to other structures in the human body. When a particular threshold pixel value is achieved in the X-Ray imaging that can be classified as a stone and detected it like a stone. At a higher level when the pixel is very bright then it can be detected as a metallic component. In metallic components there also different shapes. If the detected object is detected in the shape of a bullet, then it marks it as a bullet is detected otherwise it will be treated as a normal metallic object. This operation (shape detection) is done with morphological image processing.

In some cases, it is found that the above-mentioned objects can be present in a distributed format in the whole body. It is hard for doctors to see every single part of the body and trace out the components. In such a case, if the task is given to a computer vision system, the vision system can do detection of all possible abnormalities. Considering the abnormalities once, the object is detected. Subsequently, it will start recognizing and categorizing the detected components.

Using these detection techniques if we will be able to propose a suitable localization technique for the objects then there will be a big hope of future advancement in the sector. A robotic operator can be used in such a scenario that will conduct the safest surgery possible (The author of Nikolic M. (2016), Lu L.(2017) and Khosravan N.(2018) discussed in their research) . With the help of the detection technique, it will detect the object. Followed by a recognition technique that will classify the substance present in the body. Once the robotic operator is instructed to start the operation, it will take the safest operation possible in the current scenario. The proposed algorithm is able to detect both moving object as well as stationary objects. As the algorithm can scan a video frame by frame, it is not dependent of the movement of the object. At each fame of the capture video, the object will be detected and the proposed algorithm takes the current frame as the input. This is the reason it is independent of the motion of the object.

This manuscript describes how an X-ray image can be processed through different stages and by application of convolutional neural network how object is detected in the human body. After detection, the corresponding objects are being recognized. The recognized labels are then given on the corresponding detected objects. In this way before going to any type of surgery doctors can build a proper strategy for their operational procedure. By doing so the doctors will have a clear understanding of the root cause of the problem they will face while dealing with such situation. This can save a lot of life while a patience is undergoing a surgery.

Literature Review

There is an amend amount of development and active area of research going on in the biomedical image processing field. The case study is done on MRI, CT and X-ray imaging in a more focused way. Combining the diversity of the survey the manuscript takes its inspiration to develop a novel technique in doing object detection and recognition on X-Ray images. The following are some related work in the area that inspires the case study in the manuscript.

Unsupervised Subject Detection via Remote PPG (Method 1)

This is research conducted on the area of biomedical treatment based on different advanced thermal imaging technology. The technology involves scanning the human body and let the doctor know the different thermal report of the patients. The technique used is an unsupervised learning method rather than a supervised learning method. The proposed method in this research is voxel-pulse-spectral methodology which has the three different stages 1) creates hierarchical voxels across the video for temporally parallel pulse extraction; 2) builds a similarity matrix for hierarchical pulse signals based

on their intrinsic properties, and 3) utilizes incremental sparse matrix decomposition with hierarchical fusion to robustly identify and combine the voxels that correspond to single/multiple subjects.

Optimizing the Level Set Algorithm for Detecting Object Edges in MR and CT Images (Method 2)

This is a case study related to MR and CT image scan. It is usually difficult to detect different organs in the human body especially their boundary areas. To analyze particular organ doctors, need to detect areas occupied by different organs that helps them to detect problems and different organs. The proposed method in this case study is a genetic algorithm. The proposed method uses the training of the genetic algorithm, then testing on different sets of images related to MR and CT scans. Four quadratic majors are taken into consideration for the determination of the boundary of the object. In short, this method helps to detect different objects in a human body through MR and CT scans.

Real-Time Detection of Foreign Objects Using X-Ray Imaging for Dry Food Manufacturing Line

The proposed method applies X-Ray images to detect foreign objects in packaged food with irregular texture pattern. The application implements image enhancement by reducing the texture intensity of the food items. As mentioned in the research the foreign objects cannot be known in advance so it is classified based on one class implementation. Which means either it is a desired food item or it is a foreign element present in the food package. For real-time, operation, a max and min difference of mask operation is performed. The algorithm helps to detect different types of an element like glass, ceramic and metals from the X-Ray image. The detection operation is performed in a 180ms period. The clock frequency of the computer used in this experiment was 2.4 GHz. This method is used to improve food quality while doing the packaging. First, the background image is removed for faster processing. The background image is removed by using a process called thresholding. The resultant image is processed through an algorithm called the positive response of zero means. Following this process feature extraction of the objects is done. The features are used for the process of classification. As the foreign objects can be of various types and size it is classified as a foreign object if it fulfils some predefined criteria. In the experiment 170 images are being classified which includes 3 different types of food packaging. 12 different types of foreign objects are being detected per X-Ray image. Moreover, a tabular result is being constructed for analysis.

The X-Ray Object Recognition Test (X-Ray ORT) – A Reliable and Valid Instrument for Measuring Visual Ability Needed in X-Ray Screening

In this research, the application area is an aviation security system where X-Ray imaging is used. When a bag is closely packed, it is very difficult to detect the object with the naked eye. In such a case, image-processing technology can be used to detect different objects. The technology implemented in this paper is the X-Ray Object Recognition Test (ORT). The algorithm will tell which objects are prohibited and which are permissible to carry. By doing basic image processing operations the objects are detected using the shape of the object for example, a knife has a different shape compared to a gun. As it is a real-time application, each image is scanned for around 4 seconds to get a detailed understanding of the content that is present in the bag. The result of the test is whether the bag is complexly packed or not. It also shows the list of identified items in the corresponding bag, the time lag for each result production is around 200ms and below.

BACKGROUND

To have a detain insight into the proposed method we need to have a clear understanding of the related concepts that are used in this manuscript. These related concepts are listed and described briefly in this section of the manuscript.

Image Thresolding

This method of image processing converts a greyscale image into a binary image. The binary image can contain only the darkest pixel or the brightest pixel. The darkest pixel represented as a logic zero and the brightest pixel is represented as a logic one. A transition from brightest to darkest or vice-versa is considered as an age in the image. For converting a greyscale to a binary image thresholding process is needed, whichever value of the pixel is more than threshold it is considered as 1 otherwise it is considered as 0. The binary images are generated to do image segmentation. Before doing image segmentation we usually do process morphological operation on the binary image. (The author of Mohammed M. (2018),Koniar D.(2017) and Litjens G.(2017) discussed in their research) There are different types of morphological operators present that are circular operator, square operator. The binary image is tuned to a suitable level by varying the parameters of the morphological operators. Figure 1 shows how the thresholding phenomenon is achieved in a greyscale image.

Image Segmentation

As the name suggests this leads to the segmentation of the image into different sections. Image segmentation helps in extract different possible objects present in the image. Once these objects are detected, we enclose a bounding box around these objects. The bounding box in an image represents

Figure 1. The thresholding process done on a greyscale image on the leftmost image, the mid part of the image is representing the histogram, the redline on the histogram represents the thresholding point, after thresholding the rightmost part represents the binary image that is generated which is ready for image segmentation
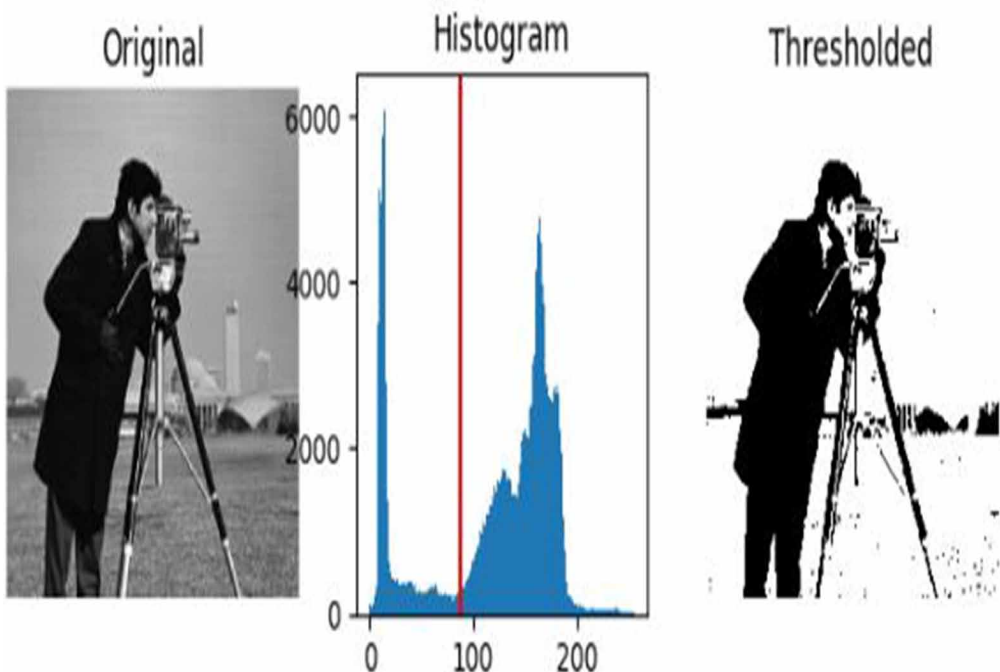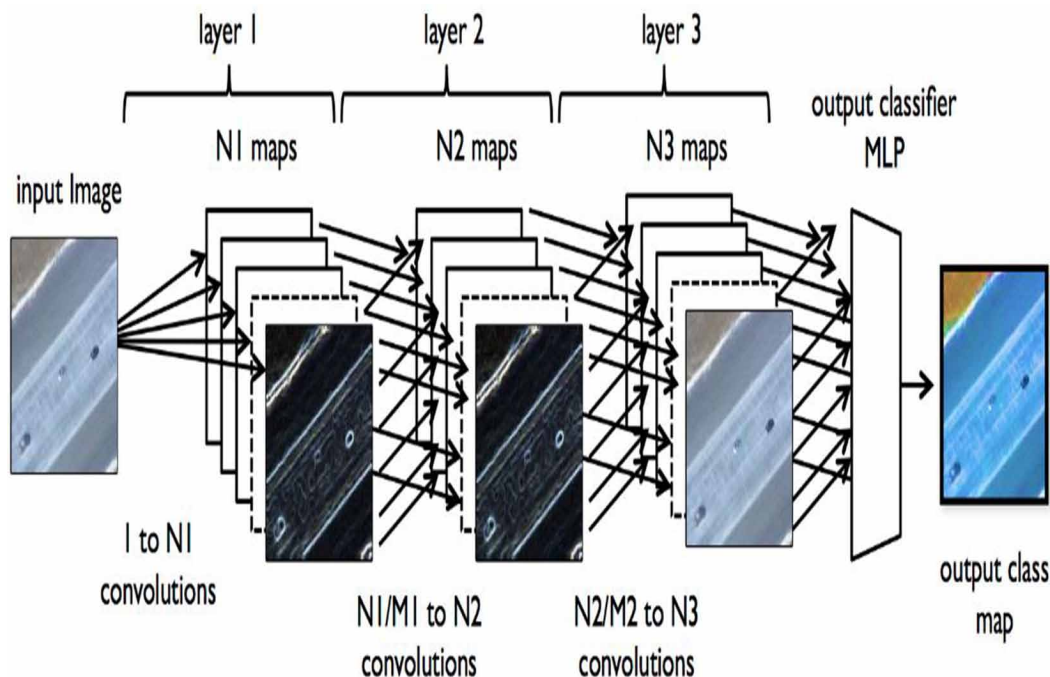
Figure 2. A phenomenon representing the image segmentation process where objects are in white pixel and background in black pixels


an area where there is an object detected. Considering this, we do object recognition with the help of recognition algorithms. Figure 2 represents an image segmentation process. All the objects present in the image should be traced out from the background. The objects that are present in the image are white and the background is of dark color after segmentation. At this stage, we can observe where the object resides in the image. These objects are then passed to the recognition stage.

Image Recognition and Labeling the Detected Objects

Once the segmentation process finished and the bounding box captures the key areas in the image a recognition algorithm tries to process the important information. A recognition algorithm is generally a neural net or a deep neural net (The author of Wang W. (2017), Ghosh, P. (2016), Zhu W. (2016) and Andrearczyk V. (2017) discussed in their research). Neural nets consist of different layers of neurons that perform different functions. For an image processing application, it is recommended to use convolutional neural network layers. A basic layout of a neural network or a deep neural net for image processing is shown in Figure 3.

As shown in Figure 1 there are three convolutional layers for the neural network in this case study. The more the number of the convolutional layer the computation load on the processing will also increase. The efficiency of understanding images increases with the number of convolutional layers. Due to the above two Parameters, it is always recommended to find a sweet spot to optimize the efficiency of the network. To build an optimized network with an optimized efficiency needs a lot of time frame. An alternative approach to achieve the same performance in a short period is the method of transfer learning. The following subsection enlightens us on how to do transfer learning. After a network is fully trained, the bounding box objects are passed through the neural network and the output will be a label. It can be any of the three categories as represented in the result section of

**Figure 3. The basic structure of a deep neural network**
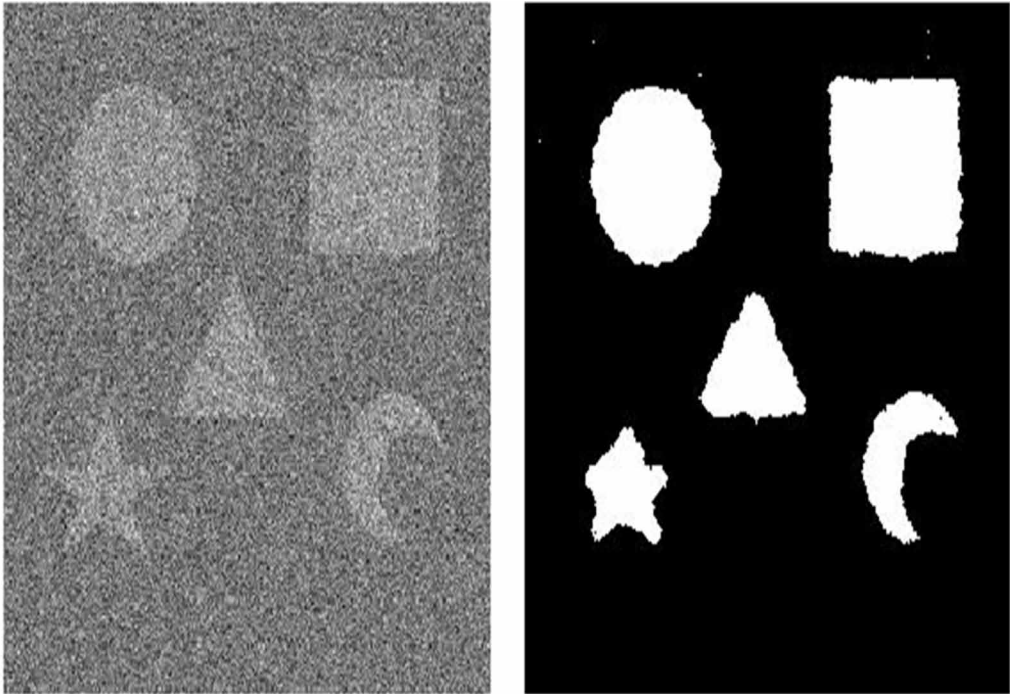
the manuscript. Each detected label is represented in the bounding box, which can be found in the result section of the manuscript.

Transfer Learning

As the name suggests it is a learning transfer from a trained network for a particular application to another type of neural network to perform some type of application. The tuned weights from the first neural network are used in the second neural network. In transfer learning, we do change the input size of the network and the softmax function in the output side of the network. The trained neural network used in the research mentioned in the manuscript uses Alex net. Alex net has three types of layers' convolutional layers, fully connected layers, and softmax function. In general, Alex-net is designed for color images but we have changed it to operate on grayscale images, as x-ray imaging is a greyscale image. There are two types of connections in the neural network. If all the neurons of one level are connected to all the neurons of the next level, then it is called a fully connected layer (The authors of Sarikaya D. (2017), Erickson B. (2017), Javanmardi M.(2018) and Suzuki K.(2017) discuss in their research). A layer may consist of different segments of the neuron. Softmax function decides the probability of the output categories that are present in the system. The output of the network will be the maximum possible category for a particular test case. Figure 4 shows the complete architecture of Alex-net.

The above-mentioned concepts are used in implementing the research described in the manuscript. A high level of optimization is done in achieving better performance in each stage of the process. Which in a way increases the overall performance of the system. The methodology part explores the method through which the object is being achieved. The result section reveals the outcome of the method followed. The conclusion section analyzes and summarizes the results. It also deals with the future scope and possibilities of the research to help the community in doing further improvement.

Figure 4. The left-hand side of this network represents the input; Conv: represents convolutional layers; FC: represents fully connected layers; and class number: represents the softmax function that decides the class to which the particular object belongs
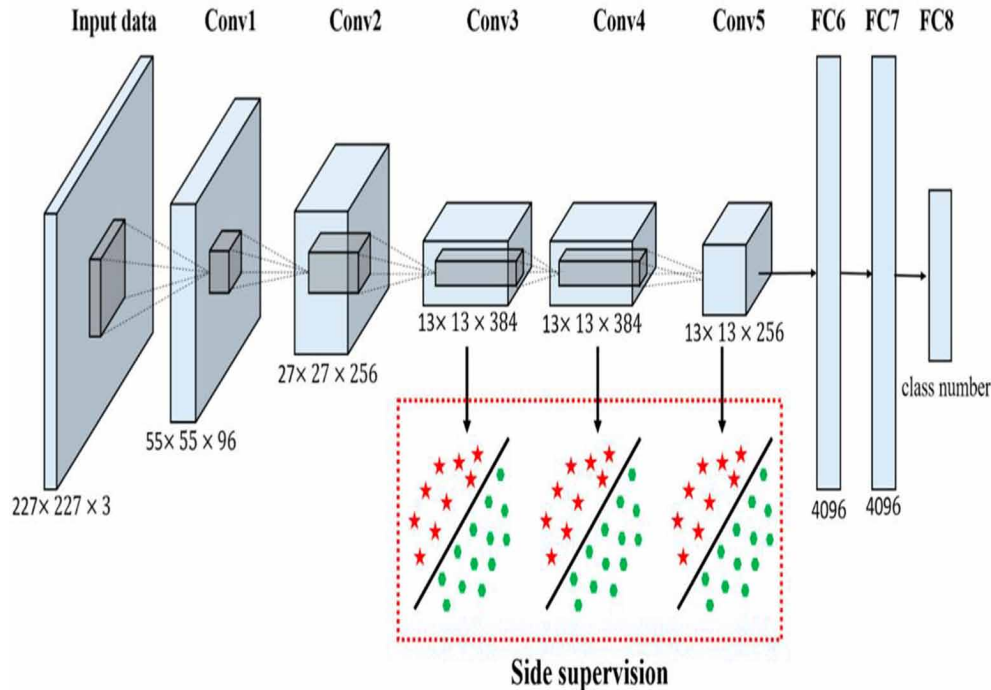

METHODOLOGY

The focus of this research is to detect and recognize abnormality in the human body such as metallic objects and tumors through X-Ray imaging. In the field of machine learning for image processing convolutional neural networks playing an important role for optimized performance. So a structured transfer-learned deep neural network is used in this case study. We have followed a process of supervised learning. We have labeled data for different categories of the image. These data are being collected through consistent google search over time. The images are being labelled with the corresponding objects that is present in the X-Ray image. For all the three categories we have collected around 500 X-Ray images to train, the deep neural network, which is 75 percent of the labeled data, is used for training and 25 percent of the labeled data is used for testing purposes. Before doing so we have done some pre-processing of the images to achieve better performance on the results. These preprocessing are being described in the background section of the manuscript (image thresolding and image segmentation). Figure 5 showcase the flow chart that describes the complete process in a sequential manner. The dataset used in the process is collected over the google search. The images are being resized to 227*227, as this is equal to the size of the input layer in alexnet.

Step 1: Start

At this stage, the algorithm will do the basic processing of the system that is memory check and all the related power supply checks. If this process fails, the complete system will not able to start. Followed by this process the X-Ray image acquisition system will be checked. Once the checks give positive response subsequently, the process will proceed to the second stage, which is Acquiring the X-Ray image.

Figure 5. The proposed workflow is mentioned in the above flow diagram. The detailed description of the process in a step-by-step manner is described in the following section.
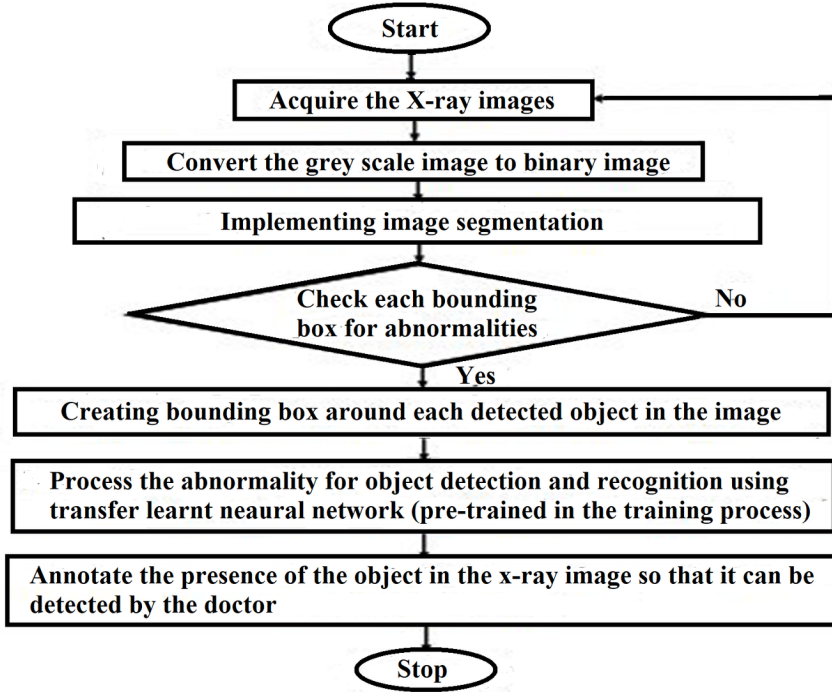
Step 2: Acquiring X-Ray Image

At this stage, the X-Ray is being acquired with the help of the X-Ray device and subsequently, it is subject to the process of image processing. The images are stored in the memory of the device and then it is subject to pre-processing. The pre-processing stage is needed in this process as it increases the efficiency of the detection of the objects and subjects the detected parts of the image to recognition. Pre-processing includes image thresholding and image segmentation, which is mentioned in the following stages.

Step 3: Converting the Grey Scale Image to a Binary Image

The image that is being collected in the X-Ray is a greyscale image. To distinguish between objects, it is recommended to the segment as mentioned in the segmentation sub-section in the background section of the manuscript. To achieve an optimized image segmentation, we need to convert the greyscale image into a binary image. For the conversion, we have used equation 1. Where "th" represents the threshold. Every pixel value, which is falling below the pixel, is represented as a dark pixel and every pixel value falling above the threshold value is represented as a bright pixel:

$$P = \begin{cases} 0, & P_x < Thresold(th) \\ 1, & P_x \geq Thresold(th) \end{cases} \tag{1}$$

For example, let p be the resultant image and p_x be the original greyscale image. The pixel of p will be a bright pixel for every value of the pixel of p_x which lies above the threshold value. The components of p are either zero or one but the components of p_x are within 0 to 255 as it is a greyscale

image. In this way, the grey image is being converted to a binary image that contains only dark and bright (or white pixels). After generating the binary image, we proceed for the process of image segmentation that is described in step 4.

Step 4: Image Segmentation

The image segmentation means dividing the image into different segments as described in previous sections. A defined region of the pixel that has the same bright pixel value is detected as an object. Whenever there is a sudden change in the pixel from dark to bright or vice-versa the corresponding region is detected as an object by enclosing it inside a bounding box. The coordinates of the bounding box are collected at this stage. The collected coordinates are used in subsequent processing stages in the proposed algorithm.

Step 5: Checking the Bounding Boxes for Abnormality

As mentioned in step 4 the bounding box coordinates are collected. These coordinates are used to capture the objects. Each of the bounding boxes is subjected to different abnormality classification morphologically. In this method, each object size is being checked in the bounding box region. If any abnormality is detected, the detection of abnormality represents the presence of abnormal conditions present in the patience's body. If it is so then the algorithm will proceed to step 6 If not, it will wait for the next frame.

Step 6: Creating Bounding Box Around Objects That is Being Detected

If any abnormality is being detected, for the defined bounding boxes collected in stage 4 then those bounding boxes are superimposed on the original X-Ray image. There is a predefined range of pixels for normal structures for human bones and organs according to the density of the tissues. If the pixel range in the bounding box is not falling in that range, then it is detected as an anomaly. So bounding boxes that contain the anomaly are used in step 7 of the process.

Step 7: Process the Abnormality of Object Detection and Recognition

At this stage, the bounding box containing the abnormal objects is analyzed with the help of the pre-trained neural network. The pre-trained neural network uses a convolutional neural network as mentioned in the background section of the manuscript. The proposed convolutional neural network contains a convolutional layer, four filters, and five fully connected layers and a softmax layer. The functionality of each type of layer is mentioned in the background section. The proposed architecture is being inspired by alexnet. In addition to alexnet architecture, proposed network has 4 additional filter layers. These additional filters are needed to deal with X-Ray imaging.

The training process for the neural network involves 1000 iterations for the training on different categories of X-Ray images. This X-Ray Image contains abnormal objects with their corresponding label. As the data set is huge, we have used image data store functionality of MATLAB to maintain track of the labeled images. For the training set, all the images are labeled with the correct labeling to avoid ambiguity. The result analysis section demonstrates the efficiency of the detection in the real scenario. It is found that the results are considerable. The two main objectives of the detection are to detect tumors and metallic content. In the case of a tumor cell, the pixel values will be less compared to a metallic content this is because X-rays will be reflected more from a metallic object as it is denser compared to a tumor cell. Due to this reason, the object which having metallic content will have a bright spot compared to the tumor in the X-Ray, which is taken into consideration for classification. It will help the doctors to trace out those areas, which are affected by these two things. After the detection and recognition are done, the result of recognition will be the object name, which is detected. It can be either a metallic part in the body or a tumor that can be at any stage. Once the recognition process is done the process shifts to the annotation phase. In the annotation phase, each

bounding box is being labeled and displayed on the screen superimposed on the original image. The annotation process is mentioned in detail in step 8.

Step 8: Annotation of the Bounding Box

At the end of the recognition process, individual bounding box, which contains the abnormality, is labeled. Here we keep a track of which object is being recognized in which bounding box. Correspondingly, it is stored in the memory. The labels along with the bounding box are annotated in the original X-Ray image. This represents the output of the proposed method. By viewing the output, doctors will be able to recognize which infected area they missing for their analysis. It will help them to avoid completely relying on their eyeball check.

This method enables experts to see the detected object and conclude on the operation procedure. MTALAB is used as an internal development environment (IDE) for the methodology implementation. The toolbox used in this process is image processing toll-box and neural network toolbox, deep neural network toolbox. The generated results are analyzed and compared with the pre-existing methods in the result section of the manuscript. Following the result section, the conclusion section evaluates the performance of the method proposed.

RESULTS

Following the proposed method experiment is conducted. Results are generated along with the performance evaluation. The data for the training are being obtained from google, which is suitable for the experimental criteria. Training is performed on 75 percent of the collected data and the remaining 25 percent of the data is used for testing purposes. For each detected object confidence factor (which is in percentage accuracy of detection) is plotted. There are two categories of the experiment is being conducted one for tumor detection another is for metal object detection. Whichever object is detected as an abnormal object and cannot be recognized with the help of the recognition algorithm is categorized into another object category. The output of the object detection is represented in Figure 6 to Figure 11 along with the confidence factor in percentage. The experiment was conducted automatically for the set of images using an image data store. The performance is plotted with the help of MATLAB plotting tools for visualization. Figure 12 represents the efficiency of detection for the three different categories. Figure 13 represents the time lag in microseconds, which is automatically being calculated for all the 100 tests. The performance is also evaluated in comparison to method 1 and method 2 in the literature review subsection. Figure 14 represents the performance evaluation for the same.

The above result obtained is discussed and analyzed in the conclusion section. The future scope of the work is analyzed in a short and precise manner in the conclusion section. The efficiency of the experiment is considerable. The time lag in the conducted experiment was negligible.

CONCLUSION

From the above experiment, it is analyzed and concluded that the experiment is giving a considerable amount of performance compared to the other two methods mentioned in the literature review. Which is seen from Figure 14 that accuracy of detection of method 1 and method 2 is around 87 and 89 percent respectively whereas it is 90 to 92 percent for the proposed method. There is no false positive detection in the experiment only the confidence factor for the object recognition is varied over the course of testing. This is why this concept can help many doctors to conduct operations more successfully. This can also help the doctors to do remote operations. This will also help the doctors from a remote place to discuss in a conference about the operation procedures and strategy, which will make it more successful. It is also found from Figure 13 that the time lag for each experiment is in the range of 2.5 to 8.5 nanoseconds, which make the time lag for the system as negligible compared to the studies proposed in the literature survey. This is achieved with an efficiency of around 90 to 92 percent while

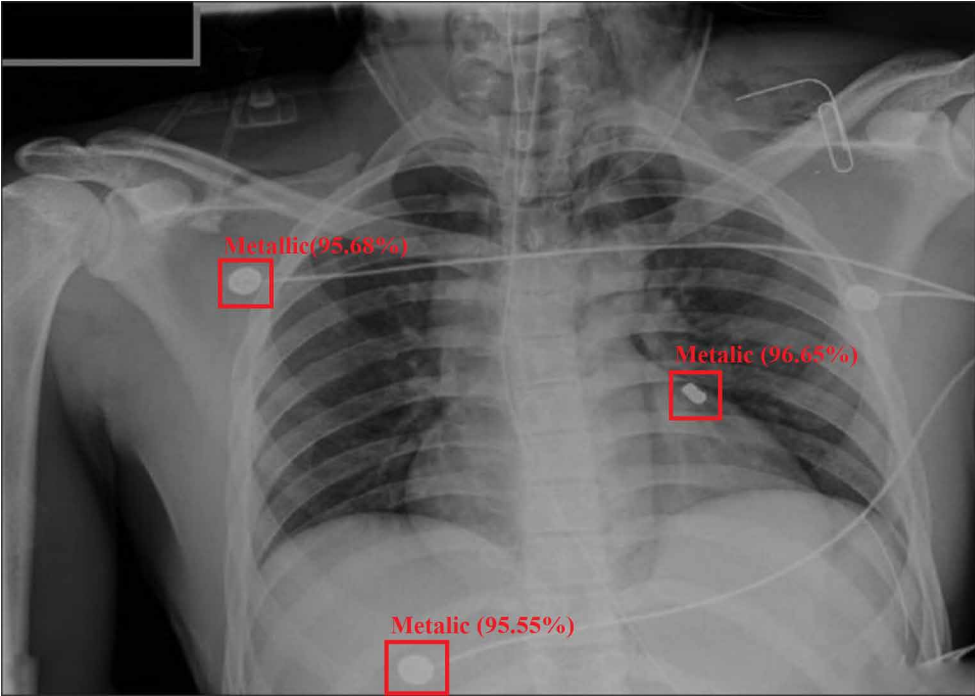Figure 6. Metallic particles detected in the human body near the lungs area and upper patient body

Figure 7. Bullets are detected in the spinal cord and near collar area of the patient's body
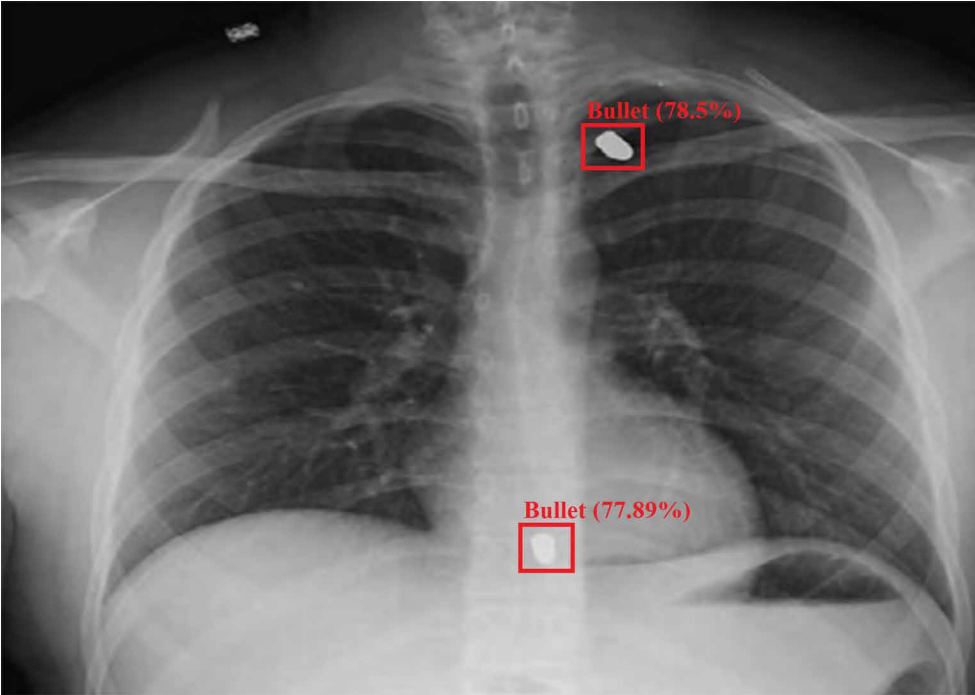
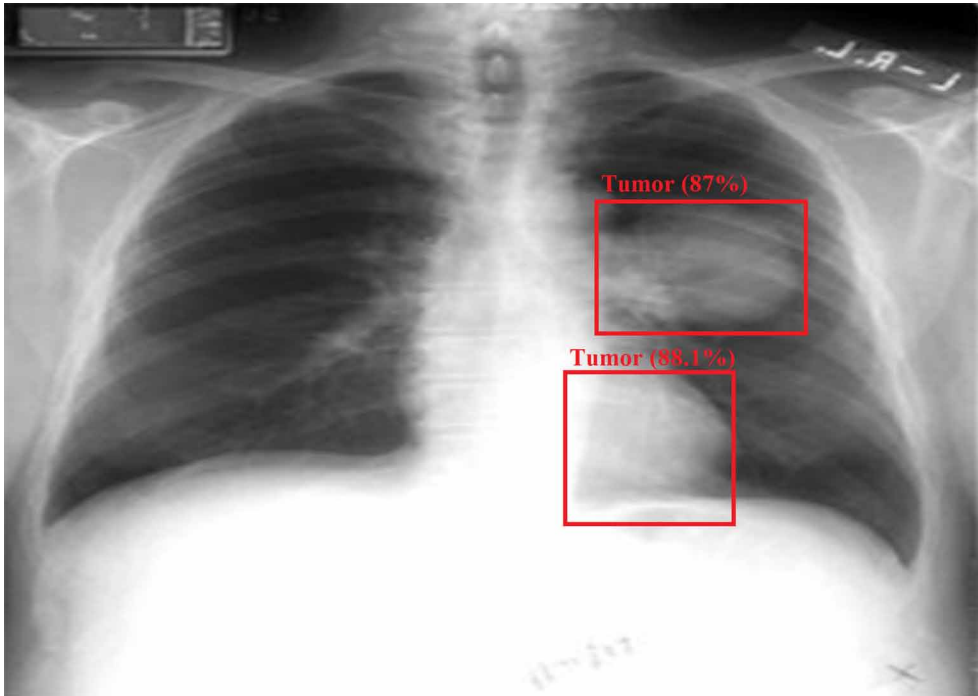**Figure 8. Tumor detected in different are of the patient's body**

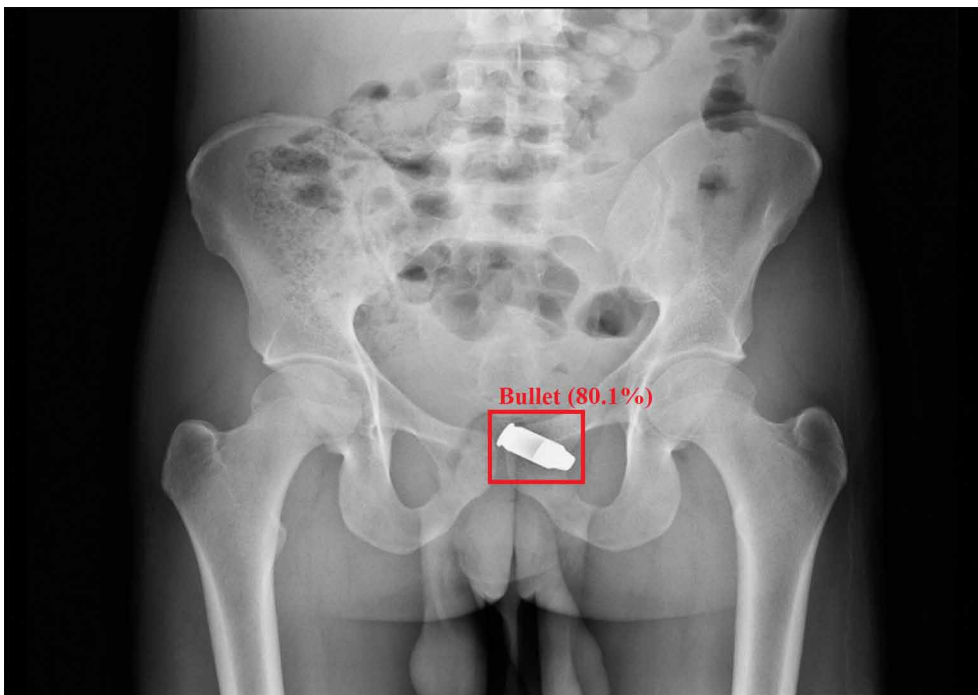**Figure 9. Bullet detected in the abdomen part of the patience body**

Figure 10. Bullet detected in the neck part of the patience body with the corresponding bounding box and confidence factor shown
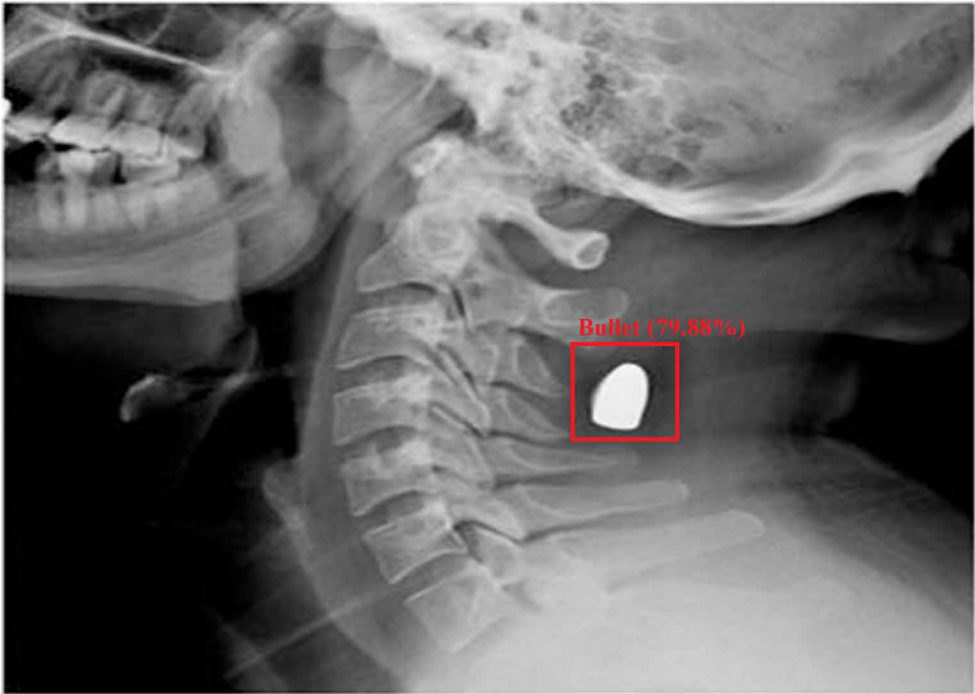

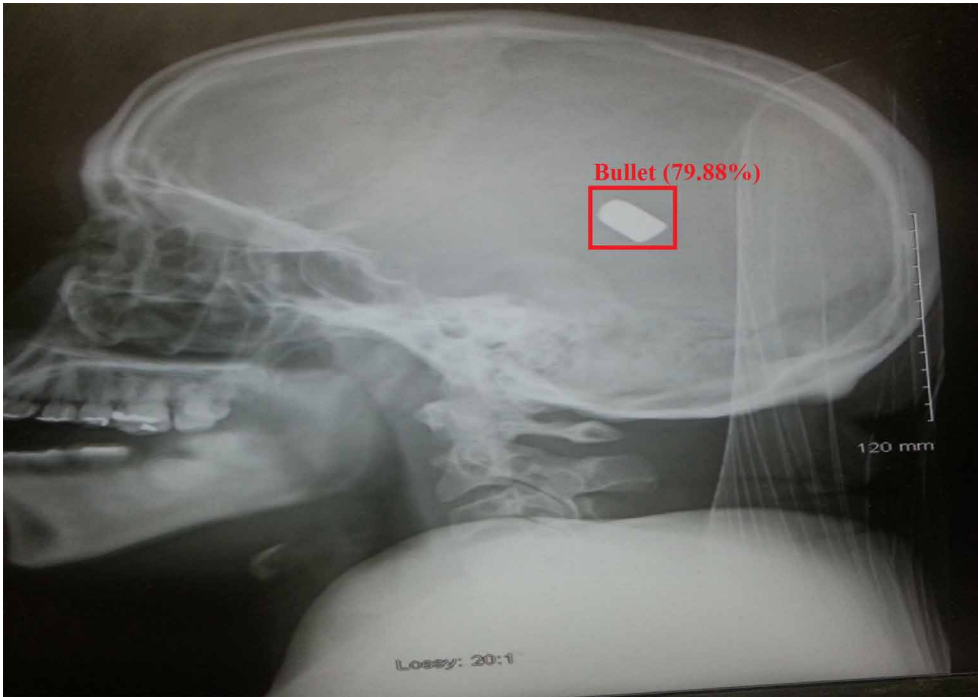**Figure 11. Bullet detected in the cranial part of the patience body**

Figure 12. The result showing the accuracy of the object detection that is conducted over the 100 different objects for all the categories of substance. Another remote object is that category of the remote object, which is not classified by the recognition algorithm, but it is detected as a remote object.
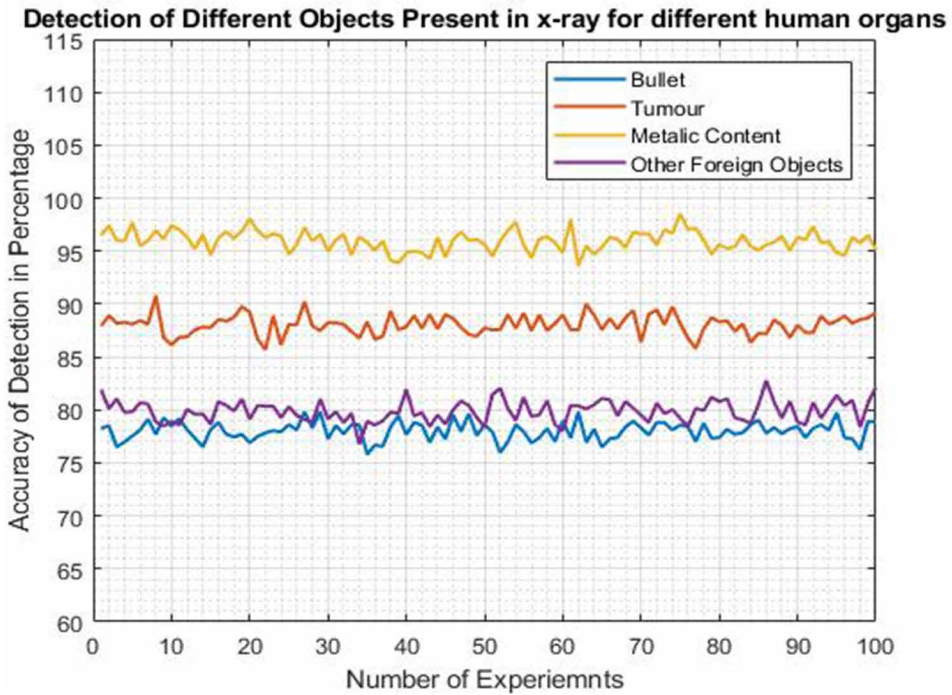

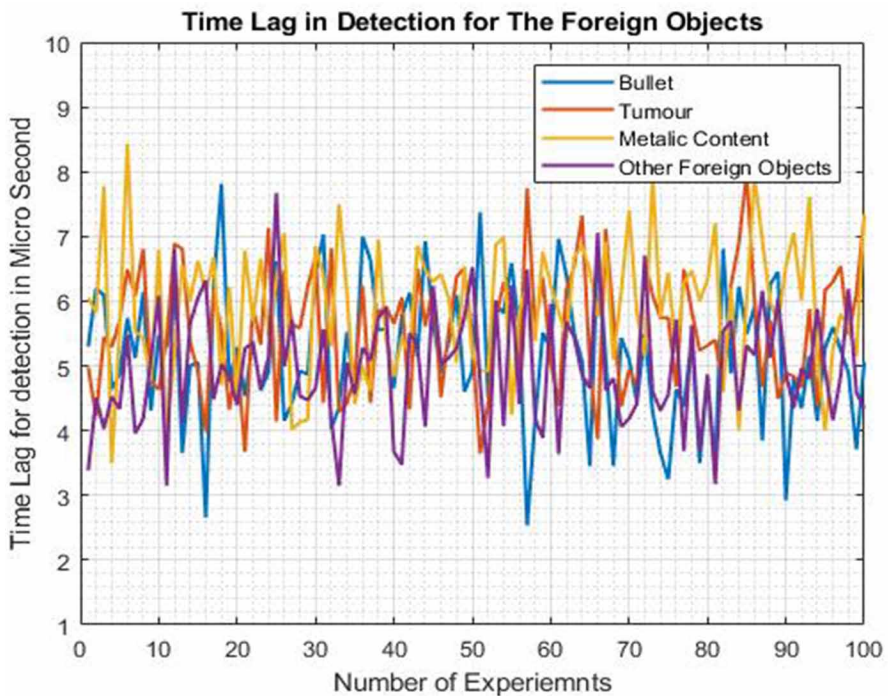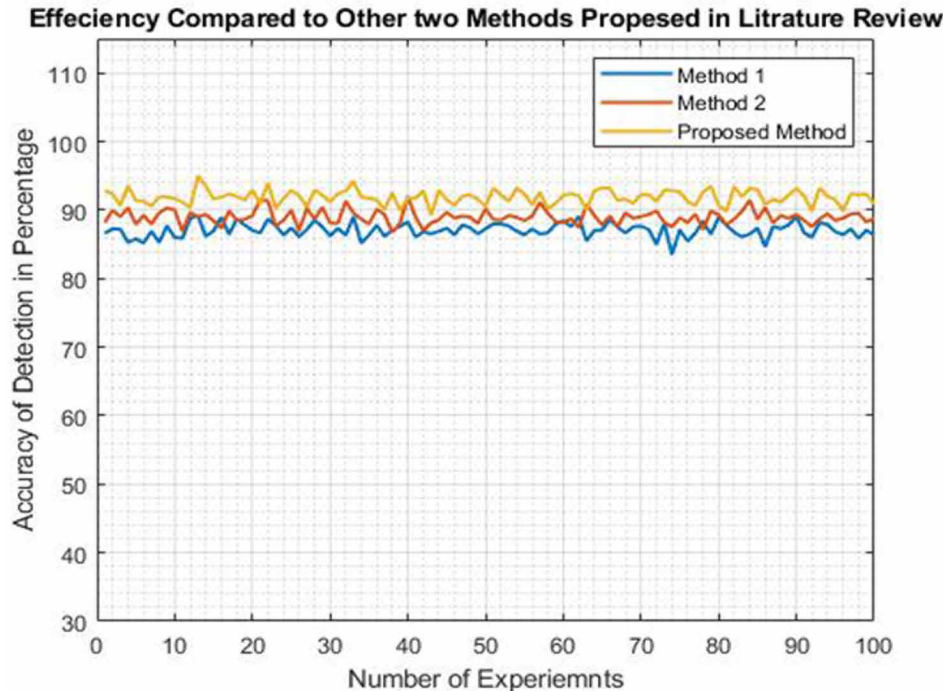**Figure 13. Time lag in microsecond for each conducted experiment**

Figure 14. The accuracy is compared with the other two proposed methods for all 100 experiments conducted in an automated fashion compared to the proposed method

conducting the experiment in an average for all the possible categories of classification this means the operation can be conducted efficiently and there will be no lag in the detection of the objects as the equipment need to operate in real-time. This field of research is an active area of research so the proposed research also has some future scope available in it. As discussed in the methodology section we have included three different categories of object and the remaining is categories as other categories. We can increase the number of categories that will give a diversity of the classification process. In case of tumor, there are different types and stages of tumor, which can also be classified, in further research. However, this can only be a justified classification category with the help of professional doctors in the related field, which is the current limitation of the proposed idea. As doctors need to seat together and discuss what are the possible scenarios they have encountered in their field of experience. By doing so young doctors can use the classification to comes to a conclusion. By this way, workload on the skilled doctors will be reduced and the treatment speed will increase without losing the effectiveness of the treatment. In addition to the above in absence of doctors, also an employee of a hospital can detect the objects present in the patient's body in case of emergency, which can be shared with an expert. Therefore, this technology can unravel a new era of medical treatment.

REFERENCES

Andrearczyk, V., & Whelan, P. F. (2017). Deep learning for biomedical texture image analysis. In *Proceedings of the Irish Machine Vision & Image Processing Conference*. Irish Pattern Recognition & Classification Society (IPRCS).

Cao, G., Xie, X., Yang, W., Liao, Q., Shi, G., & Wu, J. (2018, April). Feature-fused SSD: Fast detection for small objects. In *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)* (Vol. 10615, p. 106151E). International Society for Optics and Photonics.

Dou, Q., Chen, H., Yu, L., Qin, J., & Heng, P. A. (2016). Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, *64*(7), 1558–1567. doi:10.1109/TBME.2016.2613502 PMID:28113302

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, *37*(2), 505–515. doi:10.1148/rg.2017160130 PMID:28212054

Ghosh, P., Mitchell, M., Tanyi, J. A., & Hung, A. Y. (2016). Incorporating priors for medical image segmentation using a genetic algorithm. *Neurocomputing*, *195*, 181–194. doi:10.1016/j.neucom.2015.09.123

Javanmardi, M., & Tasdizen, T. (2018, April). Domain adaptation for biomedical image segmentation using adversarial training. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 554-558). IEEE. doi:10.1109/ISBI.2018.8363637

Khosravan, N., & Bagci, U. (2018, September). S4ND: single-shot single-scale lung nodule detection. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 794-802). Springer.

Koniar, D., Hargaš, L., Loncova, Z., Simonova, A., Duchoň, F., & Beňo, P. (2017). Visual system-based object tracking using image segmentation for biomedical applications. *Electrical Engineering*, *99*(4), 1349–1366. doi:10.1007/s00202-017-0609-0

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. doi:10.1016/j.media.2017.07.005 PMID:28778026

Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). *Deep learning and convolutional neural networks for medical image computing*. Advances in Computer Vision and Pattern Recognition. doi:10.1007/978-3-319-42999-1

Mohammed, M. A., Ghani, M. K. A., Arunkumar, N. A., Hamed, R. I., Abdullah, M. K., & Burhanuddin, M. A. (2018). A real time computer aided object detection of nasopharyngeal carcinoma using genetic algorithm and artificial neural network based on Haar feature fear. *Future Generation Computer Systems*, *89*, 539–547. doi:10.1016/j.future.2018.07.022

Nikolic, M., Tuba, E., & Tuba, M. (2016, November). Edge detection in medical ultrasound images using adjusted Canny edge detection algorithm. In *2016 24th Telecommunications Forum (TELFOR)* (pp. 1-4). IEEE. doi:10.1109/TELFOR.2016.7818878

Prall, M., Durante, M., Berger, T., Przybyla, B., Graeff, C., Lang, P. M., & Mariam, F. et al. (2016). High-energy proton imaging for biomedical applications. *Scientific Reports*, *6*(1), 27651. doi:10.1038/srep27651 PMID:27282667

Sarikaya, D., Corso, J. J., & Guru, K. A. (2017). Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. *IEEE Transactions on Medical Imaging*, *36*(7), 1542–1549. doi:10.1109/TMI.2017.2665671 PMID:28186883

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*(1), 221–248. doi:10.1146/annurev-bioeng-071516-044442 PMID:28301734

Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological Physics and Technology*, *10*(3), 257–273. doi:10.1007/s12194-017-0406-5 PMID:28689314

Wang, W., Stuijk, S., & de Haan, G. (2017). Living-skin classification via remote-PPG. *IEEE Transactions on Biomedical Engineering*, *64*(12), 2781–2792. doi:10.1109/TBME.2017.2676160 PMID:28278453

Zhu, W., Huang, W., Lin, Z., Yang, Y., Huang, S., & Zhou, J. (2016). Data and feature mixed ensemble based extreme learning machine for medical object detection and segmentation. *Multimedia Tools and Applications*, *75*(5), 2815–2837. doi:10.1007/s11042-015-2582-9

Garv Modwel has 18+ industry experience in various domains. Currently he is research scholar of Amity university from Department of Computer Science (ASET).

Anu Mehra is affiliated with Amity University, Noida, a private research university located in Noida, Uttar Pradesh, India. It was established in 2005 by an Act of the State Legislature of Uttar Pradesh. The University is a part of Amity education group.

Nitin Rakesh is affiliated with Sharda University, a private university in Greater Noida, Delhi NCR, India, established through an act of State Legislature of Uttar Pradesh. The university offers courses at the undergraduate and postgraduate levels. It is registered under Section 2 of UGC Act 1956 by University and Grants Commission.