IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⫘ Impact Factor 7.918 ⫘ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12320

A Survey on Deep Learning Techniques for Cervical Spine Fracture Detection

Madhuri B.H[1], Mamatha S[2], Manasa R[3], G. Punya[4]

Dept. of Computer Science and Engineering, Vidyavardhaka College of Engineering Mysore, Karnataka, India[1-4]

Abstract: Fracture is caused by a loss of bone mass (osteoporosis) that occurs as part of ageing. Lifting heavy- weighted things or falling from higher altitude will cause crack inback bone. Spinal cord is first seven bone of our neck region. Fracture in spinal cord may lead to loss of sensory functions or lead to death. These factures are suspected by computed tomography (CT) which is critical to patient management. This paper is a review on various methods required for detecting cervical spine fracture in a most efficient and effective manner using deep learning. It provides a brief description about the popular techniques available for detecting fracture and also gives comparison on various classification methods that can be used in the process.

Keywords: Medical imaging, Deep learning, Object detection, Classification, Cervical spine fracture, Convolutional Neural Network (CNN)

I. INTRODUCTION

The cervical spine is a flexible structure that protects the nervous system's innervation of the entire body while simultaneously preserving head and neck mobility. Seven stacked vertebrae, or bones, make up the cervical spine and are designated C1 through C7 [1].The cervical spine's bottom joins the upper back at roughly shoulder level and its top joins the skull. When viewed from the side, the cervical spine softlycurves toward the front of the body and then back, creating a lordotic curve. The term "broken neck" refers to a fracture of one of the cervical vertebrae. The majority of cervical fractures are caused by high-energy trauma, such as car accidents or falls.

Falls from a chair or other low-level objects can cause a neck fracture in older adults. A bone fracture is linked to 56% of cervical spinal cord injuries, and cervical spine fractures are a significant cause of mobility and mortality in trauma patients. According to the level affected, cervical spine fractures are often classified into one of three groups: C1, C2, or sub-axial spine (C3 to C7). The most serious kind of spinal cord injuries are often those to the cervical region. They could cause tetraplegia or quadriplegia,which would mean that all four extremities' muscles would lose strength[2].

Additionally, a neck X-ray can rule out uncommon and more severe causes of neck pain and stiffness include tumors, cancer, infections, or fractures. A cervical spine MRI can be used to detect issues including infections and tumors. Additionally, it can aid in the diagnosis of cervical spine herniated discs and spinal stenosis, a narrowing of the spinal canal.

Nowadays, computed tomography (CT) is almost completely used instead of radiography for the imaging diagnosis of adult spine fractures (x-rays). To stop neurologic degeneration and paralysis following trauma, it is crucial to locate any vertebral fractures as soon as possible[3]. The deep convolutional neural network (DCNN) used in was a 3D ResNet-101 [3] DCNN thatwas trained on 990 normal and 222 fracture patients. Area Under the Receiver Operating Characteristic (AUROC) and Area Under the Receiver Operating Characteristic (AUROC) performance of this approach.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁒ Impact Factor 7.918 ⁒ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12320

(a) c1 cervical spine fracture.

(b) c2 cervical spine fracture.

(c) c3 cervical spine fracture.

(d) c7 cervical spine fracture.

II. MEDICAL IMAGE PROCESSING

In recent years, medical domain produces a wide variety of image data related to patient's report, treatment, prescription, CT scans and MRI images which can be effectively used to automate tasks and produce quick and accurate results. The fundamental problem is that the report quality creates the impression of association because of the improper management of data. To extract and analyze these medical records elegantly and effectively, appropriate data processing techniques needs to be applied. There are various machine learning techniques that can be employed with particular classifiers to distribute data based on those qualities [5].

Digital image processing has become more significant in the field of health care, which is encouraged by the growing usage of direct digital imaging equipment. Digital sensors have now replaced computed tomography (CT) or magnetic resonance imaging (MRI), endoscopy and radiography digital methods. Digital images are composed of basic unit called pixels. They can be effectively processed and evaluated.

Processing a medical image can be mainly divided into five stages. The process of creating an image comprises each stage, from taking the picture through creating a digital image matrix. An image can be enhanced by using the right rotations, filters, and transformations [6].

Third stage comprises of visualizing image by manipulating matrix to optimize it. Image analysis requires prior information about the nature and content of the images, which must be deeply abstracted into the algorithms. Image management provides many techniques to efficiently store, transmit and retrieve image data. Several compression techniques can be applied on images to reduce storage space [7].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 7.918 ⁒ Vol. 12, Issue 3, March 2023
DOI: 10.17148/IJARCCE.2023.12320

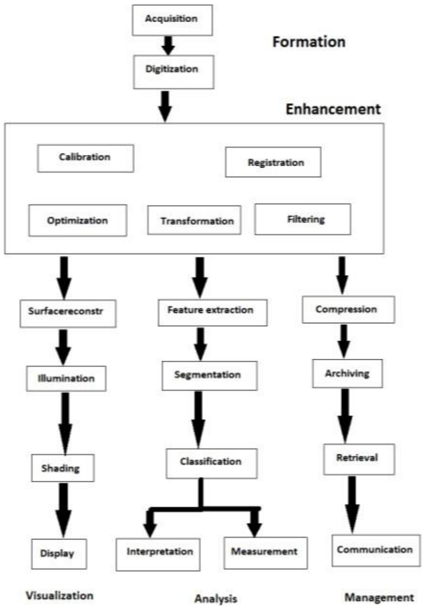

Fig. 1 Image Processing Stages

III. DEEP LEARNING TASKS IN MEDICAL IMAGING

Deep Learning is the advancement of machine learning which primarily employs neural networks for data learning and prediction. It comprises different algorithms ranging from simple to complex. Deep learning innovations have produced neural network algorithms that can now compete with humans in vision tasks like image segmentation and classification [8]. Transformation of such techniques into medical science has significantly improved analysis of an image. It is used to automate a variety of time-consuming radiology operations,including lesion detection, segmentation, classification, monitoring, and treatment response prediction, which are typically not possible without software [9].

Despite the major developments in deep learning techniques, there is less research on workflow to execute projects related to radiology which consists of variety of steps, including the selection of patient population, choice of index test and reference standard, model selection, and evaluation of performance [10].

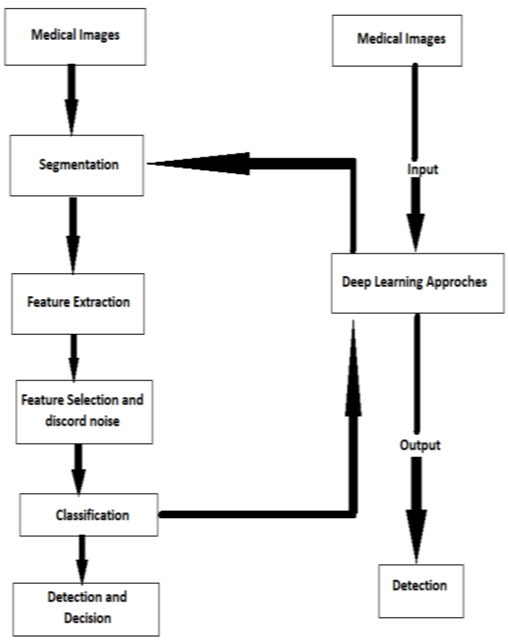

Fig. 2. Deep Learning algorithms workflow in medical image

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁒ Impact Factor 7.918 ⁒ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12320

A. Object Detection

Object detection in medical images involves the task of identifying the exact location of lesions and classifying either the presence of it or categorizing it into one among many classes. A bounding box is generated around the interested object region in the image. Faster R-CNN, Region-based Convolutional Nueral Networks(R-CNN), Region-based Fully Convolutional Network (R-FCN), Single Shot Detector (SSD), YOLO (You Only Look Once) are some of the popular algorithms for object detection [12].

1) Faster R-CNN: It is a two stage-detector. In an image there might be many objects that has to be detected. To determine to which grounding truth are we referring to can be computed in two steps. First step is to identify good candidate bounding box (bbox) and the second step is to classify and refine detected bounding boxes. Initially the image is passed into blocks of convolutional layers which outputs feature maps.These feature maps are passed into another branch to identify bounding boxes [13]. Once there is a good number of candidates, features are collected corresponding to bboxes and perform final regression and classification. It primarily comprises with the steps listed below.

- Discretise bbox space (x_c, y_c , h, w)
  where x_C, y_C are anchor points
  h, w is the height and width of the bbox
- N candidates per anchor
  Generally three different scales and ratios are chosen which gives nine candidates per anchor.
- A classifier is trained for each bbox that predicts the objectness score. The main goal is to find the presence of an object and not to classify it.
- Sort and keep the top scores
- Refine through regression

Fully connected layer is at the end which has four neurons which gives four co-ordinates of box[14] .

C) YOLO Detector: YOLO [15], which stands for You Only Look Once, is a real-time object detection system that can identify objects in an image or video with a high degree ofaccuracy. Unlike traditional object detection algorithms that use a two-step process of first detecting regions of interest and then classifying them, YOLO processes the entire image in one go, thus reducing the computation time.

The network is designed to produce a set of bounding boxes around objects contained in an image, along with a class label and confidence score for each box. One of the key benefits of YOLO is its ability to detect objects in real-time, making it suitable for use in a variety of applications such as self-driving cars, surveillance systems, and even gaming. Additionally, YOLO is also highly scalable and can be trained on custom datasets, allowing it to detect objects specific to a particular use case. The most recent version, Tiny-YOLOV3, features a compact variant that can operate on embedded systems. One of the quickest methods for object detection, YOLO, has seen multiple improvements since it was first introduced, including YOLO-V1, YOLO-V2, and YOLO-V3 [16]. It performs well in real time and has a high degree of precision.

In conclusion, YOLO is a cutting-edge object detectionalgorithm that has revolutionized the field of computer vision. Its ability to detect objects in real-time with high accuracy and scalability make it a popular choice for a wide range of applications. So, whether you're working on the next big breakthrough in autonomous driving or just looking to add some cool features to your mobile app, YOLO is definitely worth a closer look [17].

The most recent version of YOLO with a reasonably modest model size for confined situations is called Tiny-YOLO-V3. On devices with little processing power, its detection accuracy isn't very good and the real-time performance is still unsatisfactory.

3) Single Shot Detector: It is high-accuracy object detecting technique is substantially faster than the abovementioned algorithms. SSD uses VGG-16 as base network for feature extraction. More convolutional layers are added for base VGG-16 network for detection. The final convolutional layers in the base network get gradually smaller. This aids in object detection at various scales. Each feature layer has a distinct convolutional detection model. Instead of using a single feature map of a single size to predict the bounding boxes and

of varying sizes representing various scales are used. SSD takes an input image with object-specific ground truth bounding boxes and predicts the bounding box with confidence score for all object categories [18].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 3, March 2023
DOI: 10.17148/IJARCCE.2023.12320

Table 1: Comparison between speed and accuracy of different object detection models on VOC2007.

Model	Speed in FPS	Mean Average Precision (mAP) in %
SSD500	22	76.9
Faster R-CNN	7	73.2
YOLO	45	63.4

IV. CLASSIFICATION IN MEDICAL IMAGING

A. 3D Convolutional Neural Network (CNN)

The 3D CNN has potential for helping radiologistsprioritize their work lists and find cervical spine fractures on CT scans. Prioritizing fracture-positive exams on the work list is greatly influenced by 3D CNN. 3D CNN's diagnostic utility will increase with additional sensitivity improvements. It is crucial to comprehend the 3D CNN's advantages and disadvantages before successfully implementing it in clinical practise. The current role of the 3D CNN in fracture identification is secondary to a comprehensive assessment by aradiologist in the evaluation of individual tests, and it should always be examined before the report is finalised. In regular CNN the kernels are of 2D shape. In 3D CNN the kernels are of 3D shape that slide across the space and time to generate spatio-temporal feature maps [19].

All the properties of 2D convolutions apply to 3D as well. 3D convolutions are non-casual. When we compute activation for unit at time t, a frame at time t+1 is required assuming that 3X3 convolution filter is used. Receptive field for 3D convolutions are symmetric which means the activations are computed by looking into past.

B. Bidirectional LSTM (B-LSTM)

Bidirectional Long Short-Term Memory (LSTM) isa type of deep learning model that is commonly used in natural language processing tasks. It takes into account boththe past and future context of a sequence of data, making it more effective than traditional LSTMs that only consider the past. The model uses two separate LSTMs, one that processes the input sequence from beginning to end and another that processes it from end to beginning. Output of both LSTMs are then combined to produce a final prediction. Bidirectional LSTMs have proven to be effectivein tasks such as sentiment analysis and named entity recognition. They are now a key component in many state- of the-art models for NLP tasks [20].

C. ResNet

ResNet, short for Residual Network, is a type of convolutional neural network that was introduced in 2015 by Microsoft researchers. It was designed to address the vanishing gradient problem in deep neural networks, where training becomes increasingly difficult as the number of layers increases. The main idea behind ResNet is to use residual connections, or shortcuts, that allow the network to learn identity functions, or simply copy the input, as an additional layer. This helps mitigate the vanishing gradient problem and allows for deeper networks to be trained effectively. ResNet has achieved state-of-the-art results in a number of computer vision tasks, including image classification and object detection. It has also been extendedto other domains, such as natural language processing and speech recognition, where it has also shown promising results. ResNet has had a significant impact on the field of deep learning, inspiring the development of other deep residual networks, such as ResNet and DenseNet [21].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 7.918 ⁒ Vol. 12, Issue 3, March 2023
DOI: 10.17148/IJARCCE.2023.12320

V. CONCLUSION

This paper covers up all the details regarding what acervical spine fracture is and the ways to tackle it. This includes a description of medical image processing stepswhich helps in providing a wide variety of data related to a patient's report and object detection methods like Faster-RCNN, YOLO and single shot detector. It includes the detailing of the classification techniques such as CNN,Resnet and BILSTM. Neocortex neural network algorithmscan now compete with humans in vision tasks like imagesegmentation and classification because to advancements indeep learning. It is exceedingly difficult to automaticallydetect fractures in cervical spine CT scans. The machinelearning model in this paper, based on the ResNet-50+BLSTM layer, exhibits the capabilities of deep neuralnetworks to take on this issue. We encourage the researchcommunity to tackle this issue head-on and are in theprocess of making our sizable labeled dataset available for analysis.

REFERENCES

[1] Hojjat Salehinejad, Edward Ho Hui-Ming Lin, “Deep sequential learning for cervical spine fracture detection on computed tomography imaging”, in Proceedings of the IEEE 18th international Symposium on Biomedical Imaging, April, 2021, France.
[2] P.M Phal, L.P Riccelli, P.Wang, G.M. Nesbit, “Fracture detection in the cervical spine with the multidetector CT: 1-mm versus 3-mm axial images”, in American Journal ofNeuroradiology, September, 2008.
[3] J.E. Small, P. Osler, A.B. Paul and M. Kunst, “CT cervical spine fracture detection using a convolutional neural network, AJNR Am J Neuroradiol 42:1341-47, July, 2021.
[4] Irsoy, Cardie, O. Irsoy, C. Cardie, “Opinion mining with deep recurrent neural networks”, Proceedings of the 2014 conference on empirical methods in image processing (EMNLP) (2014), pp. 720-728
[5] LeCun, Y., Bengio, Y., Hinton, G., “Deep learning. Nature”, 2015, 521, 436–444.
[6] DorisPE et al, “The next logical step in the emergency radiographic evaluation of the cervical spine: The five view trauma series”, 2018.
[7] Aasia Rehman, Muheet Ahmed Butt, Majid Zaman "A survey of medical image analysis using deep learning approachs", IEEE.
[8] Mingyu Kim, Jihye Yun, Yongwon Cho, Keewon Shin, Ryoungwoo Jang, Hyun-jin Bae, Namkug Kim " Deep learning in medical imaging", Neurospine 2019, 16(4):657-668.
[9] Geert Litjens, Babak Ehteshami bejnordi, Arnaud ArindraAdiyoso Setio, Thijs kool, "A survey on deep learning in medical image analysis", arXiv:1702.05747v1, Feb 2017.
[10] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT Press.
[11] Aasia Rehman, Muheet Ahmed Butt, Majid Zaman, “A Survey of Medical Image Analysis Using Deep Learning Approaches”, published in 5th International conference on computing methodologies and communication (ICCMC), 2021.
[12] Abeer Aljuaid1, Mohd Anwar1, “Survey of Supervised Learning for Medical Image Processing”, Springer Nature Journal, 2022.
[13] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, “A Survey on Deep Learning in Medical Image Analysis”, arXiv:1702.05747, 2017.
[14] Shaoqing Ren, Kaiming He, Ross Girshick, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, 2022.
[15] C. Szegedy, A. Toshev, and D. Erhan. Deep neural networksfor object detection. In NIPS, 2013.
[16] Joseph Redmon; Santosh Divvala; Ross Girshick, “YouOnly Look Once: Unified, Real-Time Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
[17] Chengji Liu, Yufan Tao, Jiawei Liang; Kai Li, “Object Detection Based on YOLO Network”, IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), Dec 2018.
[18] Wei Liu, Dragomir Anguelov, Dumitru Erhan, “SSD:SingleShot MultiBox Detector”, arXiv:1512.02325**.**
[19] Du Tran, Lubomir Bourdev, Rob Fergus, “LearningSpatiotemporal Features with 3D Convolutional Networks”, arXiv:1412.0767.
[20] Adullah Aziz Sharfuddin, Md. Nafis Tihami, Md. Saiful Islam, “A deep recurrent neural network with BiLSTM model for classification”, IEEE, 2018.
[21] Sasha Targ, Diogo Almeida, Kevin Lyman, “Resnet in Resnet: Generalizing residual architectures”, ICLR 2016.