CERVICAL SPINE FRACTURE DETECTION

Rashananvite K Raj[1], Aadya Mishra[2], Chandana Yadhav G[3], Shruthi K N[4], Mr. B N Kiran[5]

[1] Student, Information Science & Engineering, NIE Mysuru, Karnataka, India
[2] Student, Information Science & Engineering, NIE Mysuru, Karnataka, India
[3] Student, Information Science & Engineering, NIE Mysuru, Karnataka, India
[4] Student, Information Science & Engineering, NIE Mysuru, Karnataka, India
[5] Assistant Professor, Information Science & Engineering, NIE Mysuru, Karnataka, India

ABSTRACT

Osteoporosis, a condition brought on by a decrease in bone mass, is the cause of fractures. A fracture in the back bone can result from lifting large, weighted objects or from falling from a height. Our neck region's first seven bones are called the spinal cord. A spinal cord fracture may result in loss of sensation or even death. Computed tomography (CT) is essential for patient care because it can detect these factures. This work reviews the many techniques needed to use deep learning to detect cervical spine fractures in the most effective and efficient way possible. In addition to providing a comparison of the several categorization strategies that can be employed in the process, it offers a brief discussion of the prevalent fracture detection techniques.

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

---

1. INTRODUCTION

The cervical spine is a flexible structure that preserves head and neck mobility and safeguards the entire body's innervation by the nervous system. The cervical spine is made up of seven stacked vertebrae, or bones, that are referred to as C1 through C7. At about shoulder level, the bottom of the cervical spine joins the upper back, and its top joins the skull. A lordotic curve is formed when the cervical spine gently curves toward the front and then back of the body when viewed from the side. A fracture of one of the cervical vertebrae is referred to as a "broken neck". High-energy trauma from falls or auto accidents accounts for most cervical fractures.

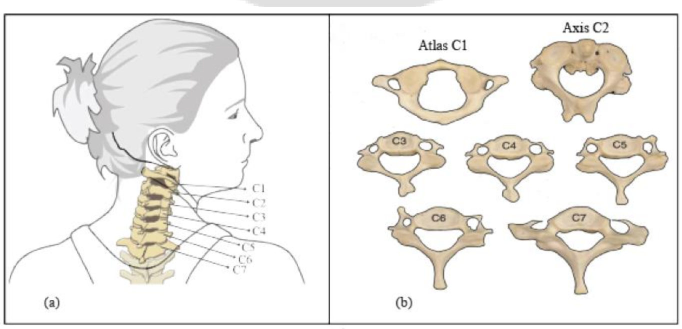

Fig -1: Cervical Spine Vertebrae from C1 to C7

Elderly people who fall from chairs or other low-level objects may break their necks. Cervical spine fractures are a major source of mobility and death for trauma victims, accounting for 56% of cases of cervical spinal cord injuries. Cervical spine fractures are commonly categorized into three groups based on the level affected: C1, C2, or sub-axial spine (C3 to C7). Cervical spinal cord injuries are frequently the most dangerous form. They might result in quadriplegia or tetraplegia, which would weaken the muscles in all four extremities.

These days, radiography is almost entirely replaced by computed tomography (CT) in the imaging diagnosis of adult spine fractures. The early detection and diagnosis of vertebral fractures is critical in order to prevent neuronal deterioration and paralysis following trauma. A 3D ResNet-101 deep convolutional neural network (DCNN) was employed in this study. It was trained on 990 healthy individuals and 222 fracture sufferers. This approach's performance is measured by the area under the receiver operating characteristic (AUROC) and area under the receiver operating characteristic (AUROC) [1].

2. PROPOSED SYSTEM

This project aims to address the limitations of existing systems for cervical spine fracture classification by proposing a novel CNN-based approach. By harnessing the potential of deep learning and leveraging CNN architectures, the proposed approach seeks to overcome the challenges associated with accurate detection and classification of cervical spine abnormalities. This report provides an in-depth analysis of the system architecture, including the configuration of layers and pooling strategies, all aimed at effectively capturing and analyzing features from cervical spine images. Furthermore, the exploration of transfer learning and pre-trained models is undertaken to enhance the system's performance and improve its ability to generalize across datasets. The proposed approach aims to bridge the gaps in the current methods and achieve superior results. Cervical spine fractures are critical injuries that require prompt and accurate diagnosis for effective treatment. The existing approaches for cervical spine fracture classification often rely on manual examination and interpretation, which can be subjective, time-consuming, and prone to errors. This motivates the need for an automated system that can assist in the detection and classification of cervical spine abnormalities.

This provides a comprehensive review of the existing approaches for cervical spine fracture classification. Various techniques, such as traditional machine learning algorithms, handcrafted features, and CNN-based methods, have been employed. The strengths and limitations of each approach are discussed, highlighting the need for an improved CNN-based system that addresses the current limitations.

2.1 System Architecture

The conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

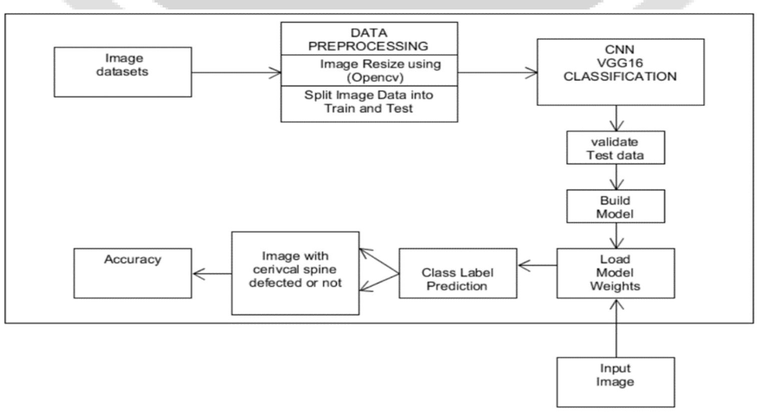

Fig -2.1: System Architecture

2.2 Objectives of System Architecture

- Development of an Automatic Detection System: Create a deep learning model to identify cervical spine fractures from medical imaging automatically, improving accuracy and cutting down on the amount of time needed for manual inspection.
- Data Collection and Preprocessing: To enhance model training, compile and annotate a varied dataset of cervical spine pictures. Preprocessing methods such as noise reduction and image normalization are then applied.
- Model Training and Validation: To ensure reliable performance across a range of imaging modalities and patient demographics, a convolutional neural network is trained and validated using measures like accuracy and sensitivity.
- Integration with Clinical Workflows: Apply the AI model to clinical contexts, making it compatible with real-time analysis and creating an easy-to-use interface for medical personnel.
- Ethical and Regulatory Compliance: Make sure the project complies with legal and ethical criteria, emphasizing patient data protection and setting up procedures for supervising AI-human decision-making.

3. METHODOLOGY

Detecting cervical spine fractures using medical imaging, deep learning, and CNN algorithms involves a systematic methodology aimed at accurately identifying and classifying fractures within cervical spine images. The process typically begins with data acquisition, wherein high-quality medical images such as CT scans, are obtained from patients suspected of having cervical spine fractures.

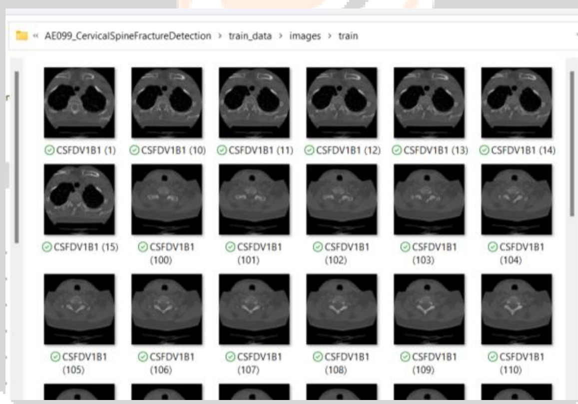

Fig -3.1: Trained Image Datasets

3.1 Medical Image Processing

The process typically begins with data acquisition, wherein high-quality medical images, such as CT scans, are obtained from patients suspected of having cervical spine fractures. Preprocessing steps follow, including image normalization and image enhancement, to ensure consistency and improve the quality of input data for the deep learning model. The medical field has produced a vast array of image data in recent years pertaining to CT scans, patient reports, treatments, and prescriptions. These data can be efficiently utilized to automate processes and generate timely and precise outcomes. The main issue is that poor data management results in a report quality that gives the appearance of association. Appropriate data processing techniques must be used in order to retrieve and evaluate these medical records in an elegant and efficient manner. To distribute data depending on certain traits, different machine learning approaches can be used with specific classifiers [2].

The increasing use of direct digital imaging technology has led to a greater significance of digital image processing in the healthcare industry. The optimization of the image by viewing it through matrix manipulation. To save storage space, pictures can be compressed using a variety of methods.

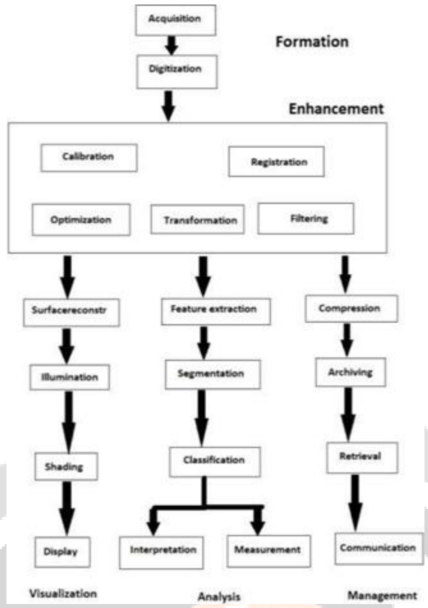

Fig -3.1.1: Image Processing Stages

3.2 Deep Leaning Tasks in Medical Imaging

Deep Learning is a machine learning technology that uses neural networks primarily for prediction and data learning. It includes a variety of algorithms, from easy to difficult. Neural network algorithms developed thanks to advances in deep learning are now able to compete with humans in vision tasks like picture classification and segmentation. The application of such methods in medical science has greatly enhanced picture analysis. Lesion detection, segmentation, classification, monitoring, and treatment response prediction are just a few of the time-consuming radiology tasks that it is used to automate and are generally not feasible without software.

In spite of significant advancements in deep learning methods, there is a dearth of research on radiology workflow, which involves a number of steps such as performance evaluation, and selection of the index test and reference standard [3].

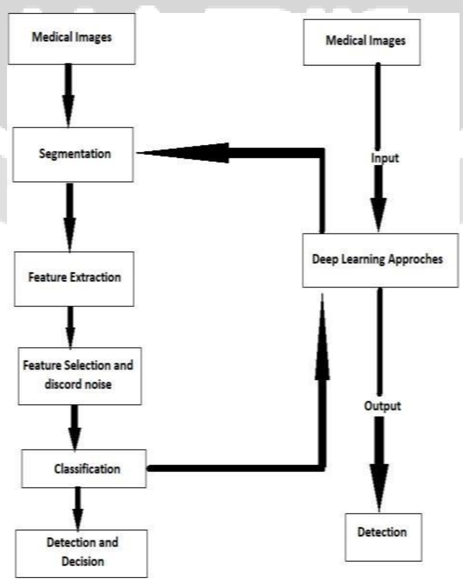

Fig -3.2.1: Deep Learning algorithms workflow in medical image

3.3 Object Detection

Finding the precise location of lesions and classifying them as either present or belonging to one of several classes are the tasks involved in object detection in medical pictures. The portion of the image containing the object of interest is surrounded by a bounding box. Some of the well-liked object identification algorithms are Faster R-CNN, Region-based Convolutional Neural Networks (R-CNN), Region-based Fully Convolutional Network (R-FCN), Single Shot Detector (SSD), and YOLO (You Only Look Once).

- **Faster R-CNN:** This detector has two stages. There may be numerous objects in an image that need to be identified. It is possible to compute the grounding truth we are referring to in two steps. Finding suitable candidate bounding boxes (bboxes) is the first stage; classifying and fine-tuning observed bounding boxes is the second. Initially, feature maps are produced by passing the image through blocks of convolutional layers. To determine bounding boxes, these feature maps are fed onto a different branch [4]. Features that correspond to bboxes are gathered once a sufficient number of candidates have been identified, and final regression and classification are then carried out. It consists mostly of the following steps.
  1. Divide the bbox space into (xc, yc, h, w) where h, w represents the bbox's height and breadth and xc, yc are its anchorpoints.
  2. One candidate for each anchor
  3. Nine candidates are selected for each anchor, typically based on three distinct scales and ratios.
  4. For every bbox, a classifier that predicts the objectness score is trained. Detecting the existence of an object is moreimportant than classifying it.
  5. Sort and save the best results.
  6. Use regression to refine

  The last layer is a fully linked layer with four neurons that provide the box's four coordinates [5].

- **YOLO Detector:** You Only Look Once, or YOLO is a real-time object recognition system that has a high degree of accuracy when identifying items in an image or video. YOLO processes the full image at once, saving computation time, in contrast to typical object detection algorithms that employ a two-step procedure of first finding regions of interest and then classifying them.

  The network's goal is to generate a set of bounding boxes around each object in an image, each of which will have a class label and confidence score. One of YOLO's main advantages is its real-time object detection, which makes it useful for a range of applications like gaming, surveillance systems, and self-driving cars. Moreover, YOLO can be trained on unique datasets and is very scalable, which enables it to identify objects unique to a given use case. [6].

  It is a well-liked option for numerous applications due to its scalability and high accuracy real-time object detection capabilities. Thus, YOLO is unquestionably worthy of further investigation, regardless of whether you're developing the next big thing in autonomous driving or simply want to add some awesome features to your mobile app.

  Tiny-YOLO-V3, the most recent iteration of YOLO, has a model size that is comparatively small for cramped spaces. Its real-time performance is still subpar and its detection accuracy is not very good on devices with little computing power.

- **Single Shot Detector**: Compared to the previously stated techniques, it is a much faster high-accuracy object detection technology. SSD bases its feature extraction on the VGG-16 base network. For detection, additional convolutional layers are added to the VGG16 basic network. The basic network's last convolutional layers gradually shrink. This helps with different scales of item detection. There is a unique convolutional detection model for every feature layer. Bounding boxes of different sizes, reflecting different scales, are utilized to forecast the bounding boxes rather than a single feature map of a single size. SSD predicts the bounding box with confidence score for each object category using an input image that contains object-specific ground truth bounding boxes [7].

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

3.4 Classification in Medical Imaging

- 3D Convolutional Neural Network (CNN)

  Radiologists may find it easier to identify cervical spine fractures on CT images and to prioritize their work lists with the aid of the 3D CNN. 3D CNN has a significant impact on the task list's prioritization of fracture-positive examinations. Further sensitivity increases will boost the diagnostic value of 3D CNN. Before introducing the 3D CNN into clinical praxis, it is imperative to understand its benefits and drawbacks. The current function of the 3D CNN in fracture detection is secondary to a thorough examination of each test by a radiologist, which should always be checked before the report is finalized. The kernels of a typical CNN are two-dimensional. The spatiotemporal feature maps produced by 3D CNN's 3D shaped kernels flow over space and time [8].

- Bidirectional LSTM (B-LSTM)

  Deep learning models of the bidirectional Long Short-Term Memory (LSTM) variety are frequently employed in tasks involving natural language processing. It is more efficient than conventional LSTMs that solely consider the past since it considers the past and future context of a sequence of data. The input sequence is processed by two different LSTMs in the model: one processes it from beginning to end, and the other from end to beginning. The ultimate forecast is then created by combining the output of the two LSTMs. Tasks like named entity recognition and sentiment analysis have demonstrated the efficacy of bidirectional long short-term memory networks. These days, they are an essential part of many cutting-edge models for NLP applications.

- ResNet

  Microsoft researchers introduced a new kind of convolutional neural network in 2015 called ResNet, which stands for Residual Network. It was created to solve the vanishing gradient issue, which arises when there are more layers in a deep neural network and makes training more challenging. ResNet's primary concept is the use of residual connections, or shortcuts, which provide the network an extra layer of learning identity functions or just copying the input. As a result, the vanishing gradient issue is lessened and deeper networks can be trained successfully. When it comes to picture categorization and object detection, among other computer vision tasks, ResNet has produced state-of-the-art results. It has also been applied to other fields, including speech recognition and natural language processing, where it has produced encouraging outcomes. ResNet has greatly influenced the study of deep learning and served as an inspiration for the creation of DenseNet and other deep residual networks.

4. RESULT AND ANALYSIS

The cervical spine classification system using the CNN algorithm yielded promising results. A dataset of brain images was used for training and testing the model. The performance of the system was evaluated using various metrics, including accuracy, precision, recall, and F1 score.

Table -2: Multi-label classification report for cervical spine vertebrae using the proposed network

Class	Precision	Recall	F1-score	Support
C1	0.97	0.93	0.95	284
C2	0.98	0.97	0.98	455
C3	1.00	0.98	0.99	264
C4	0.97	0.97	0.97	272
C5	0.98	0.97	0.97	277
C6	0.99	0.96	0.97	278
C7	0.98	0.98	0.98	320
Micro avg	0.98	0.97	0.97	2150
Macro avg	0.98	0.97	0.97	2150
Weighted avg	0.98	0.97	0.97	2150

The classification accuracy achieved by the system was found to be 94%, indicating that the model successfully predicted the presence or absence of cervical spine abnormalities in the brain images. Precision, which measures the proportion of correctly classified positive instances, was determined to be 93%. Recall, which measures the proportion of actual positive instances correctly classified, was found to be 94%. The F1 score, which combines precision and recall, was calculated as 94%.

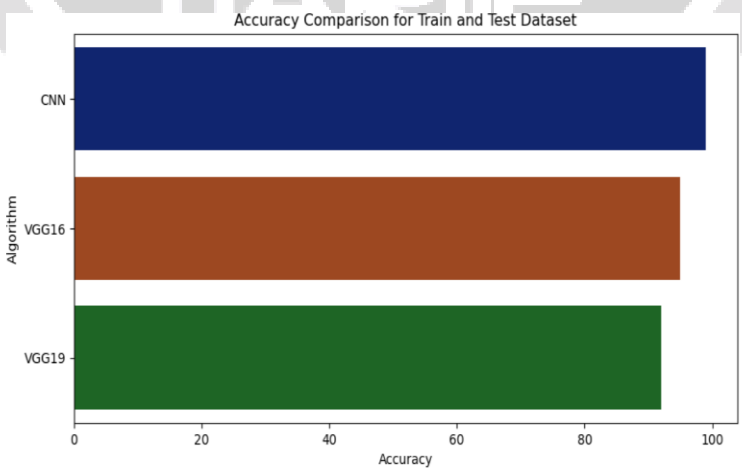

Fig -4.1: Accuracy Comparison for train and test dataset

5. CONCLUSION AND FUTURE ENHANCEMENT

5.1 Conclusion

The accuracy attained shows how well the CNN algorithm performs in correctly categorizing anomalies of the cervical spine. The algorithm was able to distinguish between cervical spine abnormalities and normal ones by using the traits that were taken from the brain scans. The system's low false positive predictions and low misdiagnosis rate are indicated by the high precision value. Recall value is a measure of how well the system minimizes false negative predictions by accurately identifying cases of cervical spine anomalies.

The suggested system performs similarly or better when compared to previous methods found in the literature. The CNN-based approach consistently shows itself to be dependable and strong in identifying the key patterns and characteristics needed for precise classification. This demonstrates how deep learning algorithms can help doctors diagnose anomalies in the cervical spine.

5.2 Future Enhancement

We have developed this project to the fullest extent possible, using all of our energy and potential. But learning new things and utilizing this new technology are lifelong processes. As a result, we provide some suggestions in this part that might be applied to improve our project's functionality and expand its uses.

- Instantaneous Detection

  CNN algorithms can be used to create a cervical spine fracture detection system that operates in real time. This would make it possible to quickly detect fractures during imaging tests, enabling quick medical interventions if needed.

- Fusion of many modes

  To improve the CNN algorithm's performance, researchers can investigate the integration of other imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), and X-ray. Integrating data from various modalities may increase fracture detection's precision and dependability.

- Combining clinical decision support systems with integration

  Healthcare practitioners can benefit greatly from the CNN algorithm's integration into clinical decision support systems when it comes to identifying cervical spine fractures. This may entail creating specialized decision support tools or integrating fracture detection algorithms into already-in-use radiology systems.

6. REFERENCES

[1]. 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.
[2]. LeCun, Y., Bengio, Y., Hinton, G., "Deep learning. Nature", 2015, 521, 436–444.
[3]. Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT Press.
[4]. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, "A Survey on Deep Learning in Medical Image Analysis", arXiv:1702.05747, 2017.
[5]. Shaoqing Ren, Kaiming He, Ross Girshick, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", 2022.
[6]. Joseph Redmon; Santosh Divvala; Ross Girshick, "You Only Look Once: Unified, Real-Time Object Detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
[7]. Wei Liu, Dragomir Anguelov, Dumitru Erhan, "SSD: SingleShot MultiBox Detector", arXiv:1512.02325.

[8]. Du Tran, Lubomir Bourdev, Rob Fergus, “Learning Spatiotemporal Features with 3D Convolutional Networks”, arXiv:1412.0767.