International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

An Overview of Inflammatory Spondylitis for Biomedical Imaging Using Deep Neural Networks

Kokkula Shiva Prasad
Research Scholar
Department of Electronics and Communication Engineering
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology(Deemed to be University)
No.42, Avadi-Vel Tech Road Vel Nagar, Avadi, Chennai, Tamil Nadu 600062
E mail: vtd1061@veltech.edu.in

Dr. Jana S
Professor
Department of Electronics and Communication Engineering
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology(Deemed to be University)
No.42, Avadi-Vel Tech Road Vel Nagar, Avadi, Chennai, Tamil Nadu 600062
E mail: drsjana@veltech.edu.in

ABSTRACT

Ankylosing Spondylitis (AS) is an axial spine inflammatory illness and also chronic that might present with a range of clinical symptoms and indicators. The illness is most frequently characterized by increasing spinal stiffness and persistent back discomfort. The affect of the sacroiliac joints, spine, peripheral joints, entheses and digits are the main cause of the illness. AS symptoms include reduced spinal mobility, aberrant posture, hip and dactylitis, enthesitis, peripheral arthritis, and buttock pain. With their exceptional picture classification ability, the diagnosis of AS illness has been transformed by deep learning techniques in artificial intelligence (AI). Despite the excellent results, these processes are still being widely used in clinical practice at a moderate rate. Due to security and health concerns, medical imaging applications utilizing deep learning must be viewed with caution. False instances, whether good or negative, have far-reaching effects on the well-being of patients and these are to be considered. These are extracted from the fact of the state-of-the-art of deep learning (DL) algorithms lack internal workings comprehension and have complicated interconnected structure, huge millions of parameters, and also a "black box" aspect compared to conventional machine learning (ML) algorithms. XAI (Explainable AI) approaches make it easier to comprehend model predictions, which promotes system reliability, speeds up the diagnosis of the AS disease, and complies with legal requirements.

Keywords: Ankylosing Spondylitis (AS), Predictive models, Diagnostic imaging, Black box, Features, Supervised learning, neural networks.

1. Introduction

Ankylosing spondylitis (AS) is an arthritic condition that causes discomfort in the lower back. The term "spondyloarthritis" (SpA) refers to any type of inflammatory arthritis that affects the spinal column.

The most prevalent of these conditions, AS, is classified by persistent inflammation of the ligaments and joints that envelop the spine. The sacroiliac joints, which attach the pelvis to the base of the spine, are frequently affected by AS (SI joints). The earliest symptoms include inflammation-related discomfort and lower back stiffness. The vertebrae may fuse if the disorder is allowed to go unchecked, decreasing mobility and increasing the risk of fractures [1-2].Complete joint injury from spinal vertebral fusion may call for joint replacement. AS affects men more frequently than it does women, and it usually first appears in young age (before the age of 45) [3]. It usually affects people under 45 and affects men more frequently than women [4] at a relatively young age.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

AS is a member of an axial spondyloarthritis (AxSpA) family of illnesses, which is more widely defined. Depending on whether radiographic imaging can identify changes to the sacroiliac joints and spine, AxSpA is further classified into non-radiographic and radiographic variants [5]. When these radiological changes manifest, AS, a radiographic AxSpA, can be diagnosed. Despite the perception, it is easy to diagnose radiographic AxSpA compared to non-radiographic AxSpA [6], and the disease typically progresses slowly before the occurrence of the symptoms on a conventional radiography. Because, at present there are no direct diagnostics that may reliably identify early stage signs of AS, medical imaging is employed to make the diagnosis [7]. A diagnosis for an AS patient often takes 7 to 10 years to come about. The patient is not receiving enough therapy during this time, and should the patient seek out additional care for their lingering symptoms, the healthcare system will be unnecessarily burdened. The different distinct imaging modalities have distinct capacities. to assess AS-related changes is given in Table 1.

Table 1: Ankylosing spondylitis imaging methods and the accuracy of the evaluation of acute and chronic alterations.

1. Inflammatory Lesions
Bone Marrow Oedema: Increase in bone marrow signal' on STIR images
2. Structural Lesions
Bone Erosion : Full-thickness loss of dark appearance of the cortical bone and change in normal bright appearance of adjacent bone marrow on T1-Weighted images''
Fat Infiltration: Focal increased signal in bone marrow on T1-weighted images''.
Bone Spur: Bright signal on T1-Weighted images extending from the vertebral endplate towards the adjacent vertebra (Spine)
Ankylosis : Bright signal on T1-Weighted images extending across the sacroiliac joints or extending from one vertebra being continuous with the adjacent vertebra(spine)

MRI is now the method that can diagnose spinal inflammation most accurately. Techniques for particle emission tomography not been adequately assessed for this usage. Patients with early AS are being categorized and identified more frequently using MRI. To evaluate spinal inflammation in individuals with AS as a measure of illness activities as well as a potential indication of therapy response, spinal MRI has been castoff. To ascertain the distribution among various therapy approaches, such as treatment with biologicals, MRI is still being assessed [8]. Figure 1 shows Sacroiliac joint (SJI) examinations with magnetic resonance imaging (SIJ) in people with Ankylosing spondylitis (AS).[14-31]

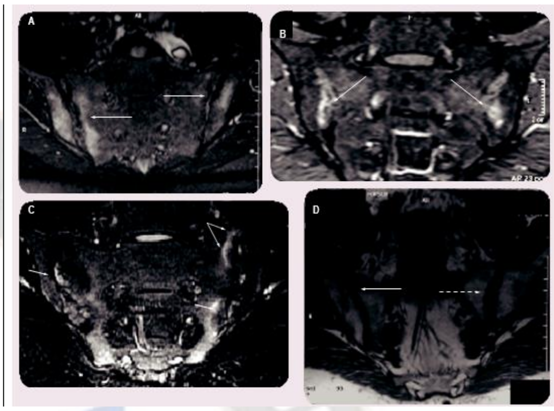

Figure 1. (A) Short tau inversion recovery magnetic resonance imaging (MRI). (B) T1w- MRI of the SIJ with contrast agent (C) T2w- MRI of the SIJ after fat suppression. (D) T1w- MRI of the SIJ without contrast agent.

The Assessment of Spondyloarthritis International Society (ASAS) states that an essential need for AS inclusion is sacroiliitis on imaging (MRI or radiography). Sacroiliitis has a specific plain radiography grading system that ranges from normal (IV) to the most severe (IV) and the grade is listed accordingly.

0: The joint width of SI is normal, and has sharp joint margins.

1: The case is suspicious

2: A few no. Of erosions and also Sclerosis

3: Erosions are severe, the joint space pseudo dilation and also have Ankylosis which may be partial.

3. Literature review

Computer Vision and Natural Language Processing (NLP) were utilized for image captioning to create textual descriptions for images, but deep learning significantly enhanced the encoder-decoder model's performance. The model that was recommended was split into two sections: a generating section for creating image captions, and a section for explaining caption terms for region weight matrices [9]. Visual Question Answering (VQA) approaches, which are applicable to the top CNN layer, aim to produce the responses

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

when given a pair of the picture and text-based queries. The features of the layer at the low-level are preferable for the semantic queries at low-level, according to a proposed Hierarchical Feature Network (HFnet). In order to provide an illustration of a class label and a classical prediction, EVCA (Visual Classification using Attributes) was suggested [28]. Using a CNN for picture analysis and an RNN for textual and justification generation, the system produced for specialists, the algorithm offered important visual characteristics; for non-experts, it extracted images that resembled the input image [9]. By utilizing a low-dimensional explanation space to learn an embedding DNN layer's high-dimensional activation vector, the Explainable Neural Network (XNN) was suggested to increase the explainability of DNN predictions. A SRAE (Sparse Reconstruction Auto encoder) was suggested to learn the embeddings in the explanation space. To predict, an XNN coupled to the DNN intermediate layer was connected to the DNN. It has been demonstrated that improving human-machine communication is possible when low-dimensional features are used to explain space visualization. For quantitative analysis, two datasets were used: Places365 and CUB-200-2011 [10]. The MAGIX (Model Agnostic Globally Interpretable Explanations) technique is utilized for the explainability of models in order to acquire classification issues. The suggested technique used a genetic algorithm to evolve the rules in order to determine the circumstances [30]. By arranging the rules and taking precision and cover into consideration, the unnecessary regulations were removed. A 500-tree random forest classifier and four datasets were used to assess the suggested approach [30]. A hybrid Deep type-2 fuzzy system (D2FLS) was created for XAI in order to overcome the challenge of high dimensional input for DNN. The technique coupled interpretable type-2 fuzzy systems with the idea of auto-encoders. Much like stacked auto encoders, the fuzzy logic system used if-then rules and was trained layer by layer to identify the primary impacted attributes [11].Using an objective function, the Explanations of the Black Box with the BETA (Transparent Approximations) method allowed for the learning of small decision sets that could be used to characterize the model's behavior in the feature space. Comparing the accuracy of the suggested method to Bayesian Decision Lists (BDL), Interpretable Decision Sets (IDS) and LIME, and it was shown to be larger.

By identifying the very important neurons for the prediction of every class and using the Graph analysis is to identify groups of similar classes, the strategy was tested using picture classification. Each neuron's activation values were shown as heatmaps on a graph that indicated the relationship between these values and the edges of the graph connecting the hidden layer and the input layer.

The complex information can be naturally captured by the graphs, especially in applications where multiple pieces of relationships and information interact. The counterfactual graphs were created to allow for an automated decision-making process that is interactive and involves a human. Knowledge graphs, a type of data structure, organize and describe items as nodes and their interactions as edges or links using an ontological schema. Knowledge graphs, a crucial tool for XAI, are studied in relation to rule-based machine learning, image recognition, recommender systems, prediction tasks, and natural language processing [13]. The X-NeSyL (eXplainable Neural-symbolic learning) method was introduced to facilitate DL and the acquisition of symbolic representations. The metric of explainability is used to evaluate the relation between the explanations provided by the machine and human experts. Knowledge-infused learning (K-iL) is a technique that was introduced for knowledge graphs [13] and these are combined with Deep Neural Network.

For both positive and negative automated evaluations, a counterfactual explanation of the prediction can be used to support a model conclusion. This explanation describes the smallest modification to the feature values that would affect the forecast. Conceptual and CoCoX (Counterfactual Explanations) CoCoX were introduced to illustrate CNN model judgments based on the features at semantic level, which are often referred to as fault lines in cognitive psychology. It was shown that explanations that are counterfactual and concept-based have value for both non-experts and experts. Using an image classification technique, the suggested model for "justified trust" and "explanation satisfaction" outperformed other approaches.

4. Deep Neural Networks

In comparison to other Artificial Intelligence (AI) techniques, Deep Neural Networks (DNN) deliver incredibly superior outcomes in picture categorization tasks, frequently outperforming human domain experts. These AI algorithms are used for image classification and segmentation in biomedical applications which has sparked considerable attention. Modern models for picture segmentation and classification in computer vision problems include convolutional neural networks (CNN) [14]. In according and order to separate the ROI, such as lesion, from the surrounding pixels in the medical image segmentation, each pixel in the image must be given a label so that comparable and similar objects in the image may be combined to make it into groups. The implant process of classification involves assigning a label to a medical image, such as benign or malignant in a binary classification, or for one of the many classes, such as pneumonia, normal, or AS for a multiclassification problem. A variety of imaging modalities, including traditional X-ray and ultrasound, are utilized to capture the images needed in biomedical and allied applications, which are then employed for illness diagnosis and management. For instance, Fundus photos are employed to identify eye problems [15]. The democratization of AI has made it possible for domain experts to implement machine learning algorithms without having a deeper performance grasp of the used techniques. This also help biomedical imaging because it can obtain more and better details about tumor size, texture, shape, and other factors than, say, human radiologists. The use of AI in radiology could revolutionize image-based diagnosis. It seems inevitable that disease diagnosis using these methods will assist in addressing the shortfall of, say, 2000 radiologists in the UK. For automated diagnosis to be used more widely in the regular healthcare sector, it is imperative to understand, suggest, and illustrate the predictions made by the AI model.

However, neural network techniques constitute the foundation of the models for image analysis that are currently gaining popularity. The Deep Neural Network imitates the biological connections between the neurons of the human brain, and it does so without having a complete grasp which involves the functioning of the brain and the interactions between neurons. Because it is not well known how a youngster recognizes an object, they can quickly adapt to recognize an animal in an image when given an object identification test. Neural networks are similar to a Black-Box that provides little affect into the decision-making process, requiring a large amount of time and training data before they might produce results that are to be satisfactory. Even the model creators find it difficult to comprehend how the model works because of its opaqueness. Simple deterministic models are frequently intrinsically comprehensible; for instance, it is comparatively simpler to comprehend a choice made by a simpler model. For situations that people can't fully understand, explainability is necessary in order to replicate or arrive at the steps from the problem to a solution. According to Fig. 2, explainable AI (XAI) strategies seek to increase confidence in model decisions by offering more details about them.[14] Explainability requires time and effort to incorporate into a model, which could impact how accurately it makes predictions. As a result, it should only be used in industries where health and safety are crucial, like the healthcare industry, as opposed to, say, an application that labels emails as spam.

The hyperparameters that control how well the trained model performs also influence the DNN model or algorithm design. "The explainable model is the one that offers reasons for the predictions for a particular task at the human level. An interpretable model is one from which certain inferences about the model's workings or predictions can be made. The ability of the model to understand has been improved via the development of numerous methodologies and tools. With the help of the online tool, one may learn about DNNs by viewing the activations at the layer-wise and also considering trained network characteristics. With image annotations and visual question-answering, the link between Machine Vision and NLP (Natural Language Processing) for AI explainability was examined. As a result of a partnership between the event's organizers, Google, and academic institutions, a XAI challenge seeks to improve the models of XAI for the financial services industry. In addition to imaging, contextual language models such as BERT(Bidirectional Encoder Representations from Transformers) suffer from similar interpretability gaps in present healthcare explainable systems. Knowledge graphs and other symbolic AI techniques can produce explanations that domain experts and end users can comprehend, but DNN cannot do so on its own.

Recent developments in the realm of autonomous vehicles and possibly in medicine have demonstrated that it is possible for a deep learning model to predict an incorrect result due to a lack of training data, poor data quality, or intentional intent to deliver incorrect data. The features that a model learns have given rise to several interesting questions. NVIDIA's PilotNet, a neural network-based system for autonomous vehicles, was used to justify the system's steering selections. In a comprehensive analysis of DNNs created for AS, several errors were found at various stages of data collecting, model creation, and explanation. From the perspectives of deep learning engineers and radiologists, the study highlighted the many errors caused by a lack of radiology knowledge and provides a checklist to ensure accurate analysis of lung pictures. The health domain is complicated because it straddles two worlds: clinical medicine at the bedside and medical science at the bench. DNN with the right visualization techniques can span this gap. The application of the techniques of deep learning in regulated fields is growing, and XAI is starting to be associated with this. Understanding how and why a choice was made, the reasons behind the model's failure, how to prevent failures, and ultimately how to improve the model are therefore becoming increasingly crucial. Given their significance, XAI techniques have been suggested for a

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

variety of uses, including electronic health records, manufacturing (3D images), industry (class imbalances), information retrieval, reinforcement learning, Generative Adversarial Networks (GAN), military training computer games, blockchain for trustworthy AI (saving decisions), text data explanation framework, pedestrian underpasses, and multimodal credibility analysis for online beauty Healthcare AI model clinical usage is increasingly need explainable AI.

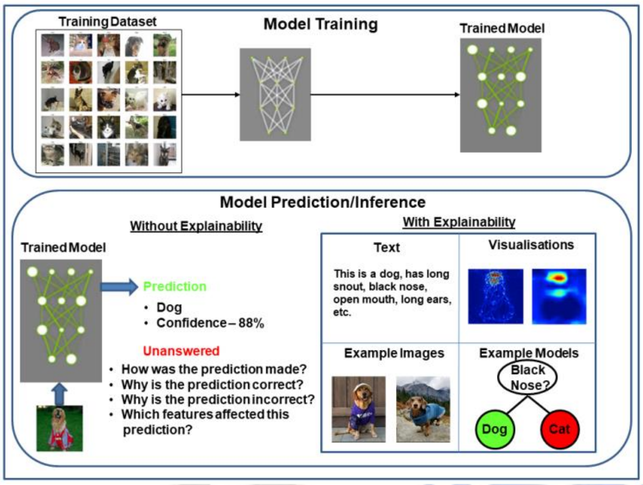

Figure. 2. XAI helping the stakeholder to understand model decision

The internal mechanisms of DNN, which possess a mysterious quality, are not fully comprehensible. Using a heat map, the hidden information about a characteristic's value can be seen visually. PESTLE analysis is a useful tool for assessing heat map quality, and it is commonly employed in the explanation of artificial intelligence systems. The following categories can be used to classify the various methods that are used to understand the categorization model at the pixel level, super-pixel level, or at the feature level:

- Perturbation: These techniques function by altering an image input and observing how the changes occur in model prediction. These may be accomplished through blurring, noise addition, or occlusion to the input. Visualizations: The goal of these strategies is to provide data visualizations that show how the model functions inside.

• Gradient/Back propagation: Using the backward pass, back propagation-based methods obtain the image characteristics which affect model prediction.

• Explanation by Example: By selecting instances that are comparable to the prediction, including novel features like text, or utilizing more data, these explanations improve the model's explainability.

The following categories can be used to group explainability approaches based on their nature, scope, and applicability:

• Range:

o Local: Procedures with a local scope offer an explanation for a prediction or specific outcome.

o Global: Methods that have a global scope aim to show how the model as a whole functions.

• Relevance:

o Model-Specific: This techniques are limited to use with particular models.

o Model Agnostic: These methods work with any model architecture, no matter what kind of model it is.

• Usage/Nature: Intrinsic: These models, such as decision trees, have explainability built in, making it simple to see the model's properties and decisions and to treat it like a white box.

o Post hoc: These techniques, which consider the model as a black box after it has been built, are used to explain to the model after it has been constructed. DNN is one example of a post hoc method.

3.1. Perturbation Based Approaches

An important map is produced by RISE (Randomized Input Sampling for the Explanation) of Black-box Models depending on the significance of each pixel to the forecast. The concept may additionally be employed to an image's caption. Additionally, the authors put forth a two model evaluation metric called the insertion and deletion, which refer to the addition and removal of pixels, respectively, along with the measurement of the resulting impact [13]. The model was assessed and shown to perform better than alternative explanatory approaches. Adaptive and Iterative Sampling with Spatial Attention (IASSA) uses a repetitive and adaptive sample module was developed to identify the map saliency of every pixel in the prediction model. RISE served as inspiration for the suggested approach. It was assessed utilizing the saliency maps' quality and compared to RISE and LIME [22].

High-level visual features can be encoded using capsule networks in the dimensions of the capsule vector. X-Caps, which is a human interpretable capsule network was developed to forecast using high-level image properties by encoding them within the capsule vectors. It has been shown that the suggested 2D capsule network technology achieves superior diagnostic accuracy and offers better visualizations when compared to 3D CNN. better visualizations when compared with the 3D CNN and also achieved better diagnostic accuracy.

Occlusion is the process of hiding a portion of a picture and measuring the resulting shift in classification. Two findings are used to suggest a unique hierarchical occlusion algorithm: first, the significant features are confined inside an image; and second, the aggregation of features at many scales might impact the prediction output. The regions with less significant traits were excluded once a hierarchy was introduced [16]. The image along with a classification model is necessary for the algorithm to function [16].

NICE (Natural Image Compression and Explanation) generates a mask for each pixel of the image, capturing its saliency based on how much it contributes for the prediction [22]. This mask is used for subsampling background pixels that are less significant to a low resolution, while maintaining the resolution of the critical pixels [17]. The sparse masks are more in accordance with human intuitions than back propagation-based methods [17].

3.2. Visualization

Individual Conditional Expectation (ICE) plots were proposed as a toolbox for the PDP (Partial Dependence Plots), a common method for visualizing the black box models [18]. In order to enhance the PDP, ICE plots exploit the relationship between the characteristics of individual observations along with the anticipated response [18]. Improvements in visualizations were seen in the trials using the suggested strategy [18]. The utility is offered as part of the ICEBOX R package [18]. A black box model, which is very challenging to directly interpret, can have its predictions approximated by using a simpler and more explicable model called the surrogate model. Ontology explains the subject of interest with appropriate logical language [19]. It was suggested that ontologies can be used in place of decision trees to increase explainability [19]. To extract the decision tree surrogate model for a DNN, the TREPAN Reloaded algorithm was introduced [29]. The assessment of the suggested approach revealed that incorporating semantic information can enhance how well explanations are understood by humans [19]. By extracting feature maps from the final CNN layer and generating a difference mask to provide the visual explanation and beauty of the prediction, the Similarity Difference and Uniqueness method (SIDU) was developed to estimate pixel saliency [20]. Grad-CAM and RISE served as the method's driving forces. On clinical and generic datasets, the proposed system was assessed [20].

The discriminative picture regions that determine the model prediction are found using a visual explanation technique called CNN Fixations, which makes use of the layers feature dependencies layers at the time of forward pass [21]. Without requiring more architectural modification, training or the gradient computation, the method has computed the image locations which are critical(CNN Fixations-similar to human eye fixations) [21]. The method finds the crucial pixel locations by dissecting the underlying forward pass procedure. Additionally, Long Short-Term Memory (LSTM) picture captioning model visual explanations were shown [21].

3.3. Backpropagation/Gradient based Approaches

The key input properties for model prediction are highlighted by the saliency map. These are accomplished by regular network testing to identify the components of the input that are affecting the output. The proposed model uses data randomization and parameter randomization tests were used to assess the saliency map-based picture classification methods. One study looked at two visualization methods; the later produced a class saliency map while the former developed an image using a maximizing class score [22]. For both convolution and deconvolution, saliency maps were utilized, along with a gradient-based visualization method [22].

One further study [23] expanded the capability of a CNN beyond simply using object localization to include locating the image areas that help in discrimination. Global average pooling was employed in the Class Activation Map (CAM) technique to create class activation maps [23]. For the advancement of the final output, the weighted sum of the spatial averaging among the feature maps available for each unit was employed [23]. The method had no appreciable impact on the model's classification accuracy 23]. Gradient-weighted Class Activation Mapping (Grad-CAM) was developed to detect and highlight the important portions of images using the final model layer gradients in order to produce visual explanations. The authors formally established that Grad-CAM (Generalized CAM) for several CNNs..Human studies were used to assess the approach for trust, class discrimination and faithfulness in elucidating the learned function. Grad-CAM's usefulness for image classification, captioning, and visual question answering was also shown [14]. By utilizing backpropagation techniques, Integrated Grad-CAM was suggested and ment to address difficulties of the gradient in CAM-based systems [25]. The proposed technique was evaluated and compared with Grad-CAM and Grad-CAM++ using Energy-based Point Game (EBPG) and Bounding box (Bbox) as object localization accuracy and feature visualization measures. When there is antagonistic noise present, in order to provide a visual explanation, the Eigen-CAM approach was devised. The results revealed an improvement for the proposed method when compared to Grad-CAM, Grad-CAM++, and CNN Fixations. Results comparing the suggested method to CNN Fixations, Grad-CAM, and Grad-CAM++ showed an improvement for the weakly supervised object localization [24]. It has been demonstrated that the proposed technique

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

can function in the presence of adversarial noise and independent of model correctness [24].

The contributions of the pixel represented as heat maps can be used in evaluating the accuracy of the categorization method. Layer-wise Relevance Propagation is a general approach that uses pixel-by-pixel breakdown. (LRP). In a range of computer vision applications, LRP first assigned relevance ratings or similarities to the various input dimension contributions with the model conclusion. It illustrates that the suggested method might work regardless of the accuracy of the model and even in the presence of adversarial noise [24].

To validate the model, the ILSVRC2012 dataset was utilized. In comparison to LRP, Contrastive LRP (CLRP), and Softmax Gradient LRP (SGLRP), the model was found to produce less noise, improve class discrimination, and hold entire target items [18].

By simulating CNN top-down attention using a novel back propagation approach named Excitation Back prop, task-specific attention maps were produced. Using the MS COCO, PASCAL VOC07, and Image Net datasets, the model's precision in the localization task was measured [19].

Deep Learning Important FeaTures (Deep LIFT) is a proposed method that operates by back propagating the network's neuron contributions to each input feature. Each neuron activation was compared to the reference activation, and the difference between the two values (20) was used to compute the contribution scores. Even in the event of zero gradients, the information can still be obtained by using differences [80]. The models trained with MNIST and simulated genomic data were used to assess the approach [20].

By altering the various input patterns filters in Deconvolution Networks (deconvnet) are developed for the process of unsupervised learning which would capture the middle and high-level features which generalize the diverse classes [21]. The proposed de conv net determined which area of the image was responsible for activation by mapping the intermediate layers feature activations back to the input pixel space [21]. In according to map features to pixels, a convent does the job of a CNN in reverse [21]. Visualizations were used to demonstrate that the features had class selectivity and growing interlayer invariance rather than being random or incomprehensible patterns [21].contributions of dimensions to the model's choice. Using directional derivatives, the TCAV (Testing with Concept Activation Vectors) algorithm assessed the impact of a User-Defined Concept—with the existence of stripes—on the categorization of zebra images. The systems sought to offer explanations that are more approachable for humans [22]. Network structure was used by the Deep Taylor decomposition to back propagate explanations from the output layer to the input layer [23]. Every neuron was viewed as a function that could be broken down into its input variables. A sizable convolutional network and MNIST were employed to assess the technique [23]. The input pixels role for an unforeseen classification task was assisted by the generated heatmaps [23]..

By merging six existing approaches and assigning each character a value based on relevance, the SHAP (Shapley Additive explanations) framework was developed as a unifying framework for comprehending complicated model predictions [24]. For computing explanations, the traditional Shapley value estimation is based on the game theory [24]. In comparison to the other chosen ways, the proposed technique's evaluation was determined to be high in line with intuition and human understanding and to have a higher computational efficiency [24].Locality Guided Neural Network (LGNN), a suggested technique inspired by Self-Organizing Map (SOM), preserves the locality between adjacent neurons of each DNN layer [25]. Although it can be applied to various DNNs, a CNN was used to show it [25]. From one layer to another, groups of related filters can be used to highlight filters that are responsive to related semantic notions [25]. The information might be distributed among the neurons in the same layer thanks to the backpropagation process.

3.4. Libraries, Tools, and Frameworks

Through a uniform interface, the iNNvestigate package addresses the problem of systematically comparing several DNNs. Along with numerous others like SmoothGrad and GuidedBackprop, implementations for LRP, Pattern Net, and Pattern Attribution are also supplied. Backpropagation is used in all of the implemented algorithms, and perturbation analysis is used to offer a quantitative evaluation.

As a TensorBoard plugin, the explAIner Visual Analytics Framework for XAI was developed. Understanding the models, detecting any model defects using various XAI techniques, and fine-tuning the models were the three stages of the Interactive and Explainable Machine Learning framework.The system was subjectively assessed using the MNIST dataset by various categories of human users, and the results were highly encouraging.

A GUI that is easy to use provides CNN layers visualizations which was proposed as a part of the AI toolset known as Neuroscope for visualizing CNN classification and segmentation. By offering insights into the forecasts, the internal view of the model can be seen [18]. For classification and segmentation, various visualization techniques can be chosen, including Grad-CAM, activation maps, and saliency maps [18]. One can observe the data flow, model architecture, weights, layer processing, and other interpretable aspects of the entire model by utilizing an interactive web-based CNN visualization tool [19]. CNN and RNN visualization techniques are available via the interface [19]. The weights for the convolutional, pooling, and fully connected layers of the CNN model could be seen [19]. The effectiveness of the method was demonstrated by polls of both DNN specialists and non-experts that reviewed the interface [19]. The information used by AI algorithms for AS identification is explained via a publicly accessible user interface for radiologists [29].

The main cloud platforms offer their own assistance in creating and comprehending XAI models. Vertex XAI, a tool that may be used to comprehend the outputs of a classification model, is offered by the Google Cloud Platform (GCP) [24]. Both AutoML and specially trained image models are supported by the Vertex AI. To assist Machine Learning Developers and stakeholders in understanding how Machine Learning models produce predictions, Amazon Web Services (Amazon Cloud) offers the Amazon Sage Maker Clarify model [24]. To show the impact of a feature on the expected result, Clarify implements SHAP and generates PDP [24]. For both Black-Box and White-Box models, Microsoft offers the InterpretML toolset for XAI. The open-source toolkit Fairlearn is used to link the Auto ML graphical user interface with the Azure Machine Learning SDK. Users can validate their models with the help of domain experts, and models can be understood by the users based on primary influences [25].

Many radiologists believe that diagnosing brain-related problems might be challenging, particularly if the tumor has penetrated intricate tissue structures. The brain is so intricate that it is difficult to differentiate between the area affected by the tumor and the surrounding edematous zone. To improve image viewing for precise anomaly detection and infectious tissue segmentation, radiologists and oncologists need an advanced computerized diagnosing system [25–27]. Machine learning techniques have advanced in recent years to discover new characteristics because they best reflect the data. Deep neural networks (DNN) are used to address problems using data-driven and feature-driven approaches, and this is their fundamental idea. DNN applies the fully convolutional network (FCN) and the convolutional neural network (CNN) to broad. With DNN, the CNN (Convolutional Neural Network) and the FCN (Fully Convolutional Network) are applied to a wide range of appliances.

These techniques have generally been used for image processing in recent years, and they are particularly useful in systems for processing medical images. As a result, deep learning has made significant progress across a variety of fields with growing efficacy, particularly in the field of biomedical imaging systems. For specific tasks like segmentation, localization, and classification, CNN is a regularly used imaging system technology that uses regularization and normalization procedures to learn and extract target unique features. For CV and diagnostic imaging, the majority of researchers used ResNet, Inception-V3, VGG, and Google Net. Despite all of this work, detection, and categorization still need to be improved. To tackle these issues, in this work, we suggested an ideal deep-learning feature fusion for brain tumor classification. Our job is completed in multiple phases. However, the optimization of deep learning features and their subsequent integration into a single matrix are the primary topics of this paper.

In terms of DL-based techniques for medical image segmentation, U-Net and Fully Convolutional Network (FCN) are the most often used. Performance-wise, U-net has shown to be the most effective strategy. Although U-Nets can segment 2D images as accurately as a human, they cannot understand the relationship between neighboring slices in 3D images because they must be viewed as a collection of 2D slices for analyzing volumetric medical images. In order to generate finer localization, the following research also offers volumetric extensions of the U-Net. The inventor of this U-Net created the 3DU-Net [28], which substitutes the 3D convolutions for the 2D convolutions in the U-Net and provides a feasible solution to the volumetric segmentation problem. In a multi-deep supervision based on the 3D U-Net model, three stages in the synthesis process are called three separate levels. Moreover, Attention U-Net suggests emphasizing the more significant activations with soft attention modules. In particular, the authors assert that because activations in the synthesis pipeline are generated through up-sampling, they may be somewhat inaccurate. While several articles have offered a variety of unique designs for different segmentation needs, they usually depend on the 3D U-Net standard [30].

4. CONCLUSION

In this research, the artificial intelligence (XAI) with explainable approaches and also their use for the comprehension of AS disease detection tasks for biomedical imaging are surveyed. By giving the practitioner new information, the explainable AI which is combined with the deep neural network (DNN) models can aid in an early detection and identification of the disease. Opportunities to comprehend, enhance, and create superior AI algorithms also exist by using a model's decision-making process.

For biomedical imaging, there is a lack of algorithm standardization, despite several, unrelated efforts to make improvements. The development of algorithms can be coordinated by a single regulatory body to set benchmarks and guarantee continuous algorithm improvement. Before being used, the secretive algorithms ought to either be made public or at the very least go through a review procedure. A better quality data can be obtained with the help of a data-centric approach in enough quantity for model training, much like enhancing the deep learning model itself. The regulatory agencies can provide a benchmark biological data that can be used to evaluate and validate various explainability methodologies. Federated learning holds promise for biomedical applications as well because it allows for the contribution of data from several hospitals to the training and improvement of deep learning (DL) models without compromising with the privacy of the data.

The various fields of Biomedical Imaging such as ophthalmology and radiology, image-based diagnosis and also prognosis now require the expert judgment of practitioners who are already overworked. Utilizing the DL techniques with explainability and automated localization can help the existing scenario by hastening the decision-making method of medical professionals. We view a human-in-the-loop as a necessary condition for the clinical decision based on an XAI application. The goal of XAI approaches should be to assist clinical supervisors and other medical professionals in making the at most use of limited resources and enhancing service delivery.

REFERENCES

[1]. Bond, D., 2013. Ankylosing spondylitis: diagnosis and management. Nurs. Stand,Vol 28, (16-18), 52–59.

[2].Shaikh, S.A., 2007. Ankylosing spondylitis: recent breakthroughs in diagnosis and treatment. J. Can. Chiroract. Assoc. 51 (4), 249–260.

[3].Sieper, J., Braun, J., 2011. Ankylosing Spondylitis: In Clinical Practice. Springer.

[4].Sze, V., Chen, Y.-H., Yang, T.-J., Emer, J.S., 2017. Efficient processing of deep neural networks: a tutorial and survey. Proc. IEEE 105 (12), 2295–2329.

[5].Rudwaleit, M., van der Heijde,. The development of assessment of spondyloarthritis international society classification criteria for axial spondyloarthritis. Part II. Validation and final selection. Annu. Rheum. Dis. 68 (6), 777–783,2009.

[6].Navallas, M., Ares, J., Beltrán, B., Lisbona, M.P., Maymó, J., Solano, A., 2013.

[7].Sacroiliitis associated with axial spondyloarthropathy: new concepts and latest trends. RadioGraphics 33 (4), 933–956.

[8].Lacout, A., Rousselin, B., Pelage, J.-P., 2008. CT and MRI of spine and sacroiliac involvement in spondyloarthropathy. Am. J. Roentgenol. 191 (4), 1016–1023.

[9].Bollow M, Enzweiler C, Taupitz M et al.: Use of contrast enhanced magnetic resonance imaging to detect spinal inflammation in patients with spondyloarthritides. Clin. Exp. Rheumatol. 20(6 Suppl. 28),S167–S174 (2002).

[10].Hermann KG, Bollow M: Magnetic resonance imaging of the axial skeleton in rheumatoid disease. Best Pract. Res. Clin. Rheumatol. 18(6) 881–907 (2004).

[11].Puhakka KB, Melsen F, Jurik AG, Boel LW, Vesterby A, Egund N: MR imaging of the normal sacroiliac joint with correlation to histology. Skeletal Radiol. 33(1) 15–28 (2004).

[12].Rudwaleit M, van der Heijde D, Khan MA, Braun J, Sieper J: How to diagnose axial spondyloarthritis early. Ann. Rheum. Dis. 63(5) 535–543 (2004).

[13].Baraliakos X, Landewe R, Hermann KG et al.: Inflammation in ankylosing spondylitis: a systematic description of the extent and frequency of acute spinal changes using magnetic resonance imaging. Ann. Rheum. Dis. 64(5) 730–734 (2005).

[14].Braun J, Baraliakos X, Golder W et al.: Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. Arthritis and Rheumatism. 48(4) 1126–1136 (2003).

[15]. Marzo-Ortega H, Braun J, Maksymowych WP et al.: Interreader agreement in the assessment of magnetic resonance imaging of the sacroiliac joints in spondyloarthropathy-the 1st MISS study. Arthritis Rheum. 46, S428 (2002) (Abstract).

[16].Braun J, Sieper J, Bollow M: Imaging of sacroiliitis. Clin. Rheumatol. 19(1) 51–57 (2000).

[17].Landewé R, Hermann KG, van der Heijde D et al.: Scoring sacro-iliac joints by magnetic resonance imaging – a multiple- reader reliability experiment. J. Rheumatol. 32(10) 2050–2055 (2005).

[18].Braun J, Brandt J, Listing J et al.: Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet 359(9313) 1187–1193 (2002).

[19].Shenkman, Y., Qutteineh, B., Joskowicz, L., Szeskin, A., Yusef, A., Mayer, A., Eshed, I., 2019. Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings. Med. Image Anal. 57, 165–175.

[20].Van der Linden, S., Valkenburg, H., Cats, A., 1984. Evaluation of diagnostic criteria for ankylosing spondylitis. evaluation of diagnostic criteria for ankylosing spondylitis. a proposal for modification of the new york criteria. Arthr. Rheum. 27, 361–368.

[21] Z. Lin, J. Huang, Y. Chen, X. Zhang, W. Zhao, Y. Li, L. Lu, M. Zhan, X. Jiang, X. Liang, A high resolution representation network with multi-path scale for retinal vessel segmentation, Comput. Methods Programs Biomed. 208 (2021) 106206.

[22] S. Zheng, X. Cui, M. Vonder, R. N. Veldhuis, Z. Ye, R. Vliegenthart, M. Oudkerk, P. M. van Ooijen, Deep learning-based pulmonary nodule detection: Effect of slab thickness in maximum intensity projections at the nodule candidate detection stage, Comput. Methods Programs Biomed. 196 (2020) 105620.

[23]. X. Xu, W. Ding, M. D. Abra`moff, R. Cao, An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image, Comput. Methods Programs Biomed. 141 (2017) 3–9.

[24]. D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks, IEEE Trans. Biomed. Eng. 65 (12) (2018) 2720–2730. doi:10.1109/TBME.2018.2814538.

[25]. Arokia Jesu Prabhu and Jayachandran, A, "Mixture Model Segmentation System for Parasagittal Meningioma brain tumor Classification based on Hybrid Feature Vector‟ Journal of Medical System, vol 42, issues 12, 2018. .

[26]. Z. Yang, X. Peng, and Z. Yin, "Deeplab v3 plus-net for image semantic segmentation with channel compression," in Proceedings of IEEE 20th International Conference on Communication Technology (ICCT). IEEE, 2020, pp. 1320–1324.

[27] .Li, H., Liu, X., Boumaraf, S., Gong, X., Liao, D., and Ma, X. (2020). Deep distance map regression network with shape-aware loss for imbalanced medical image segmentation. In Proc. Int. Workshop on Machine Learning in Medi- cal Imaging, pages 231–240.

[28] Q. He, B. Zou, C. Zhu, X. Liu, H. Fu, L. Wang, Multi-Label Classification Scheme Based on Local Regression for Retinal Vessel Segmentation, Proc. - Int. Conf. Image Process. ICIP. (2018) 2765–2769. https://doi.org/10.1109/ICIP.2018.8451415.

[29] T.J. Jebaseeli, C.A. Deva Durai, J.D. Peter, Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM, Optik (Stuttg). 199 (2019) 163328. https://doi.org/10.1016/j.ijleo.2019.163328.

[30] D.A. Palanivel, S. Natarajan, S. Gopalakrishnan, Retinal vessel segmentation using multifractal characterization, Appl. Soft Comput. J. 94 (2020) 106439.

[31].Baraliakos X, Landewe R, Hermann KG et al.: Inflammation in ankylosing spondylitis: a systematic description of the extent and frequency of acute spinal changes using magnetic resonance imaging. Ann. Rheum. Dis. 64(5) 730–734 (2005).