Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Geometric Feature Extraction In IVD Region Analysis: A Deep Learning And Region Growing Approach

Pradip Salve[1] , Harshal Patil[2] , Puja Saraf[3] , Sujata Satpute[4] , Vishwajeet Goswami[5] , Abhishek Shroff[6] , Sagar More[7]

[1]School Of Engineering, Ajeenkya Dy Patil University, Pune, India, Pradip.Salve@Adypu.Edu.In
[2]Assistant Professor, Rcpit, Shirpur, Ms, India. B
[3]Assistant Professor
Rcpit, Shirpur, Ms, India, Puja.Saraf@Rcpit.Ac.In
[4]Department Of Csit, Dr. B.A.M. University, Sambhajinagar, Ms, India, Sujata.Satpute@Gmail.Com
[5]School Of Engineering, Ajeenkya Dy Patil University, Pune, Ms, India, Vishwajeet.Goswami@Adypu.Edu.In
[6]School Of Engineering, Ajeenkya Dy Patil University, Pune, Ms, India, Abhishek.Shroff@Adypu.Edu.In
[7]Assistant Professor, Rcpit, Shirpur, Sagar.More1@Rcpit.Ac.In

This work develops an automated diagnostic system using medical image analysis methods such as feature extraction, segmentation, and localization. The system locates or identifies the Intervertebral Discs (IVDs) by using the Gabor features and the Mathematical Morphology approach on original spine MRI data. In order to detect the existence and location of IVD in pictures, the Gabor features are applied to MRIs of the spine. For IVD localization, we used the Stacked Hourglass deep learning model. For a given assignment, we investigate the efficacy of a feature combination strategy that makes use of geometric features and the Gabor features. Our investigation shows that combining these two different feature classes leads to a notable performance gain. The suggested approach addresses the intended problem with an astounding 88.06% accuracy, highlighting the effectiveness of feature combination approaches in computer vision applications. Moreover, the Dice coefficient was employed to assess the model's performance, highlighting the resilience and dependability of our methodology. These results highlight how feature combination might improve IVD segmentation and classification's robustness and accuracy.

Keywords: IVD Segmentation, IVD feature extraction, IVD classification, intervertebral disc analysis.

1. Introduction:

Chronic conditions such as low-back pain (LBP) are highly prevalent and pose a significant threat to human well-being. LBP, a widespread symptom with substantial societal and economic implications, affects a significant portion of the adult population in the United States, ranging from 25% to 50% [1]. The associated healthcare expenditures linked to spine-related pain, primarily attributable to LBP, are consistently increasing [2]. The main contributor to chronic LBP and disability is intervertebral disc (IVD) degeneration [3]. Magnetic resonance imaging (MRI) analysis plays a crucial role in evaluating the spinal cord, ligaments, lesions, and assessing degenerative disc disease [4]. Historically, manual disc localization and segmentation in spine MRI scans were conducted by radiologists, relying heavily on prior knowledge and experience. However, this approach was labor-intensive and lacked consistency among different observers [5]. Therefore, there is a critical need for automated methods to precisely localize and segment IVDs. The development of accurate techniques, aided by computer-aided diagnosis (CAD), holds the potential to facilitate the quantification of disc degeneration, disease diagnosis, and computer-assisted spinal surgeries [6–8]. However, challenges arise from variations in the size, shape, appearance, and intensity of different IVDs, along with the indistinct boundaries and resemblance in intensity to surrounding tissues, complicating the recognition process.

Geometric characteristics hold paramount importance in medical imaging, serving diverse objectives that enhance diagnosis, assist in treatment planning, and advance research. In the context of In Vitro Diagnostics (IVD) and region analysis, the extraction of geometric features from medical images has witnessed substantial advancement, contributing to a deeper understanding of anatomical structures, pathological conditions, and physiological processes. This dynamic field plays a pivotal role in uncovering intricate patterns, shapes, and spatial relationships within medical images. The exploration of IVD geometric feature extraction for region analysis stands as a progressive stride toward harnessing the power of computational analysis to enhance medical insights, paving the way for more precise and personalized healthcare interventions. This paper delves into the methodologies, challenges, and potential applications of IVD geometric feature extraction, emphasizing its transformative impact on medical imaging and diagnostics. In summary, IVD feature extraction and description are critical components of medical image analysis, enabling quantitative assessment, diagnosis, treatment planning, and research related to spinal conditions, thereby advancing medical knowledge and improving patient care in the field of spine and musculoskeletal disorders. The subsequent section provides an in-depth literature review on IVD Geometric Feature Extraction for Region Analysis, focusing on the involvement of deep learning and machine learning in this domain.

1.1. Literature Review:

2. Research Methodology:

This study employed a mixed-methods research approach of deep learning and machine learning to investigate the significance of automatic IVD segmentation and analysis. The research methodology encompassed quantitative methods to provide a comprehensive understanding of the IVD feature extraction. Fig. 1 depicts the overall methodology of IVD feature extraction and classification. It starts with data collection followed by pre-processing.

Prepocessed data is sailed through deep learning approach to localize IVD. Furthermore, localized IVD has been used for segmentation and feature extraction. We extract two type of features namely geometric features and GABOR features (textual features). The extracted features have been concatenates and used for classification.

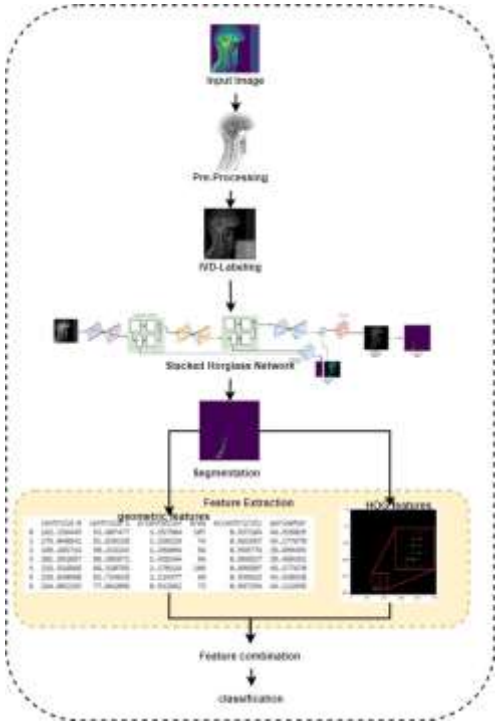


Figure 1. Proposed Methodology for IVD feature extraction and classification

2.1. Data:

We employed MRI spinal cord dataset [1]. The dataset consist of T1w MRI data from 235 subjects, the dataset includes inconsistent images since, images were captured from 40 different centers. The network was fed an average of each subject's six center slices as input images. For training and testing the dataset was divided into two parts: 1342 images and 658 images, respectively. Ground truth data was manually labeled. Labels drawn according to the sequence of cervical vertebrae from (C1 = top to C7= Bottom) using is shown in Figure 3.2(b). Nine patients' T1-weighted MR images were used to assess the performance of the suggested technique (using the Spine generic image dataset). Slice spacing ranges from 3.3 to 4.4 mm, with an in-plane resolution of 0.5*0.5 and 2 mm slice thickness. Slices of each image series are 320*320 pixels in size. There are 39 slices taken from each patient. Technical details are described in table 1 and 2.

Table 1. TechnicalDetails of the Dataset

Sl.No.	Parameters
1	FastSpinEcho(FSE)
2	Slicethickness=2mm
3	Slicespacing=0.5mm
4	Resolution=320x320

Table2. Training/Testing ratio

	Sample size
Training	1342
Testing	658
Total	2000

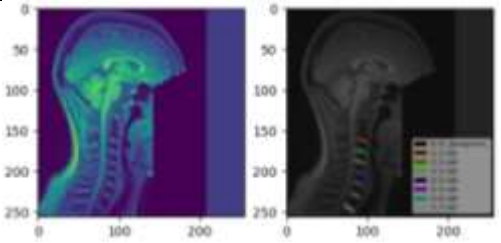


Figure. 2(a) Original Image of IVD Fig. **2(b)** Ground truth labels IVDs.

2.2. Pre-processing:

The Spinal Cord Toolbox (SCT) v4.0.1 was used to pre-process 3D volumes of the MRI data [2]. The images were re-sampled at 1 mm isotropic resolution and straightened using the spinal cord segmentation method to produce the spinal cord centerline [3]. The image was cropped to 256*256 pixels around the spinal region as part of the straightening procedure [4]. To reduce contrast variability in the image, a Contrast Limited Adaptive Histogram Equalization technique was used [5]. We increased the target size to deal with class imbalance by applying a 10-pixel Gaussian kernel to single-pixel labels.

Further we extract the average of 6 sagittal slices (centered in the middle slice) as a data sample for each subject. We normalize each image to be in range [0, 1] to reduce the effect of data variation. In order to prepare the ground truth data for the training process, first, we extract the intervertebral disc position (single pixel) from the ground truth data then we convolve the image with a Gaussian kernel to generate a smooth ground truth with increased target size (radius 10). We repeat this process for each IVD separately to produce V channel ground truth, where V is the number of intervertebral discs. Since the Spine Generic dataset consists of samples with variable number of IVDs (between C1-C7), we extract 6 IVDs for each subject. For any missing IVDs we consider unknown position and eliminate its effect on the training process by simply filtering out with the visibility flag on the loss function.

The proposed method starts with the pre-processing for the proposed model. The position of the intervertebral discs were extracted using the pose estimation method with attention mechanism of Hourglass model. Figure 6 Depicts the proposed method. Steps involved in the methodology are discussed in the subsequent section.

2.3. Contrast Limited Adaptive Histogram Equalization (CLAHE):

CLAHE, originally developed to enhance low-contrast medical images, offers a contrast-limited approach compared to ordinary AHE. It tackles the problem of noise amplification by implementing a clipping limit. Prior to computing the Cumulative Distribution Function, the CLAHE applies a predefined value to clip the histogram, thus limiting amplification. The technique divides the original image into non-overlapping contextual regions called sub-images, tiles, or blocks.

The CLAHE's performance is influenced by two primary parameters: Block Size (BS) and Clip Limit (CL). These parameters significantly impact image quality. Increasing CL results in a brighter image, especially when the input image has low intensity, as a larger CL makes the histogram flatter.

Moreover, increasing the BS expands the dynamic range and enhances image contrast. When selecting the two parameters based on maximum entropy curvature, the resulting image quality is subjectively perceived as good, as determined using image entropy.

Equation 1 in the entropy function yields a specific value that indicates the level of complexity present in the corresponding section of an image. This value represents the degree of intricacy or information content within that particular region of the image.

$$Entropy = -\frac{1}{4}\sum_{k=0}^{3}\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P(i,j;d,\theta_k) log_{10} P(i,j;d,\theta_k) \quad \text{Eq. 1}$$

The cosine weighting function, as depicted in equation 2, is utilized to measure the local contrast and luminance. In this equation, "p" represents the path radius, (xi, yi) denotes the location of theith pixel within the given patch, and (xc, yc) represents the center location of the given patch.

$$w_i = 0.5\left(cos\left(\frac{\pi}{p}\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}\right) + 1\right) \quad \text{Eq.2}$$

Equation 3 is employed to calculate the local luminance of the provided patch. Within this equation, the raised cosine weighting function is represented by "wi," the total number of pixels within the patch is denoted by "N," and the luminance of the ith pixel is indicated by "Li."

$$L = \frac{1}{\sum_{i=1}^{N} w_i}\sum_{i=1}^{N} w_i L_i \quad \text{Eq.3}$$

And contrast is calculated using equation 4.

$$C = \sqrt{\frac{1}{\sum_{i=1}^{N} w_i} \sum_{i=1}^{N} w_i \frac{(L_i - L)^2}{L^2}}$$ Eq. 4

In this research, the provided image of the cervical region ($Cervical_{img}$) (Figure 4) is processed using the CLAHE method with specific settings to obtain an enhanced image.

Figure.3 Original Image

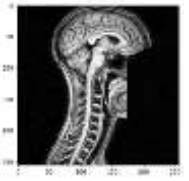


Figure.4 CLAHE Output Image

2.4. The mathematical morphology:

Mathematical morphology, initially formulated based on set theory by Martheron, was later expanded for image analysis by Serra et al [3]. This approach utilizes structuring elements, which represent specific structures and features, to evaluate the shape of an image and perform image processing. The objective is to extract information regarding the unique set structures present in the original image using the corresponding set of structuring elements. The resulting set is linked to certain characteristics of the structuring element given by Yu-Qian, et al. [4]. The fundamental operators in mathematical morphology include erosion and dilation.

2.4.1. Erosion:

Erosion, known as the shrinking operation in mathematical morphology, is a transformation that involves combining two sets with different dimensions of an image through vector subtraction. If X and Y represent sets in N-dimensional Euclidean space (E^N) with elements x = (x1, ..., xN) and y = (y1, ..., yN) respectively, the erosion of X by Y is defined as the set of all elements for which x + y belongs to X for every y⊆ Y. The erosion of X and Y can be denoted X Θ Y and it can be defined by equation 5.

$$X \Theta Y = \{x \in E^N | x + y \in X \; for \; every \; y \in Y\}$$ Eq. 5

In the proposed approach, the line filter is chosen as the structuring element to perform erosion on the Cervical IVD image in order to obtain the boundaries of the intervertebral discs (IVDs). The expected outcome of this erosion on Cervical IVD image is to obtain the pathological IVD boundaries in a sequential labeling order from C1 to C7. However, during the erosion process, the IVD boundaries are not obtained in the desired order. To address this issue, a horizontal-wise maximum response extraction technique is employed. This involves rotating the Cervical IVD image by 90 degrees counterclockwise and performing erosion using the line filter with specific length and degree settings, namely (15,0), (15,15), and (45,0). The purpose of these settings is to erode unwanted white spots and extract the IVD boundaries in the correct order. The resulting images are combined to obtain the horizontally-wise maximum response eroded image (HMaxR), as depicted in Figure 3(a).

2.4.2. Dilation:

Dilation, the counterpart of erosion in mathematical morphology, is a transformation that enlarges the input image. It is a morphological operation that combines two sets by using vector addition. If X and Y represent sets in EN with elements x = (x1, ..., xN) and y = (y1, ..., yN) respectively, the dilation of X by Y can be defined as the set of all possible sums of paired elements. It can be represented as X $\oplus$ Y and is mathematically defined as (Haralick, et al. 1987).

$$X \oplus Y = \{z \in E^N | z = x + y \; for \; some \; x \in X \; and \; y \in Y\} \qquad \text{Eq. 6}$$

In the proposed approach, dilation is employed to expand the HMaxR image by utilizing a disk as the structuring element with a radius of 2 mm. The resulting image, referred to as Dilationimg, is presented in Figure 3(b). Following that, unwanted white spots are eliminated from the dilated image by removing pixel values that fall below 100 or above 580. From the remaining connected components representing the valid upper and lower boundaries of the Cervical intervertebral discs (IVDs), the midpoint of the upper and lower boundaries is detected. This midpoint is then used to estimate the localization. Figure 3 illustrates the resulting images of the IVD localization process. Figure 3(a) clearly displays the combined erosion of the CervicalBimg, where unwanted edges and boundaries are eliminated, revealing the upper and lower boundaries of the vertebrae and IVDs. Figure 3(b) demonstrates the expanded boundaries of the vertebrae and IVDs obtained through dilation of the upper and lower boundaries. The final IVD boundary mask is applied to the Cervicalimg, as illustrated in Figure 4.3(c), and the IVD centers are identified using the midpoint Equation (7). Initially, the midpoint of the lower and upper boundaries is determined, and subsequently, the IVD centers are identified from the connected components' midpoints, resulting in the localized IVDs shown in Figure 3(d).

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right) \qquad \text{Eq. 7}$$

Where x and y represent the pixel positions of the corner of the connected components.

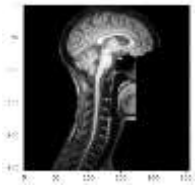


Figure 5 (a). Erosion result

Figure 5 (b). Dilation result

The output of the dilation is used to create masks, these mask send over the Stacked Hourglass Network Deep learning architecture. The mask image is depicted in figure 5.2 (b).

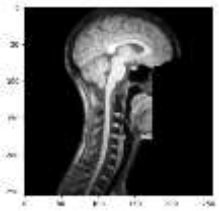


Figure 5.2 (a) Pre-processed image

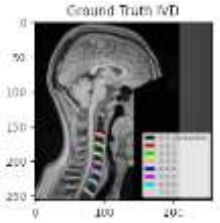


Figure 5.2(b) The IVD Mask

2.4.3. Localization using Stacked Hourglass Network:

The overall schematic of the proposed method can be observed in Figure 6. The initial stage of the method involves pre-processing the input data before feeding it into the model. The model itself employs a pose estimation technique with an attention mechanism to acquire knowledge about the position of intervertebral discs. In the following subsections, we will delve into these steps in detail.

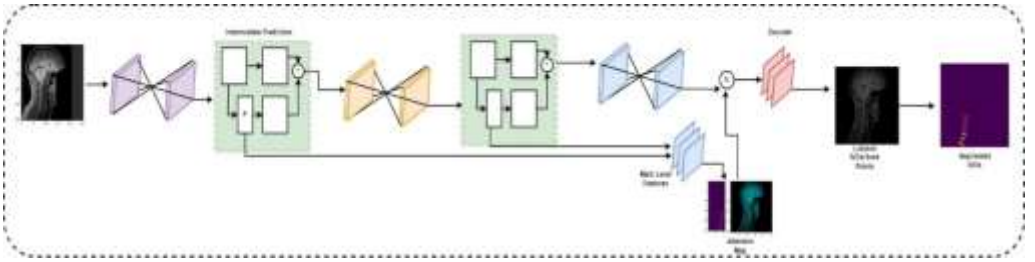


Figure6. The model utilized in this study is a stacked hourglass network that incorporates an attention mechanism. During training, the model calculates the loss function between each hourglass prediction and the corresponding ground truth mask, which allows for intermediate supervision. Additionally, the intermediate representation is passed through an attention layer to generate an attention map. This attention map serves as a guide for the decoder layer, enabling it to focus specifically on the intervertebral disc region.

To obtain data samples for each subject, we extract the average of six sagittal slices, with the middle slice being the center of the selection. To minimize the impact of data variation, we normalize each image to a range of [0, 1]. For the training data's ground truth preparation, we start by extracting the position of the intervertebral disc as a single pixel from the ground truth data. Next, we convolve the image with a Gaussian kernel to generate a smooth ground truth representation with an enlarged target size, using a radius of 10. We repeat this process for each intervertebral disc individually, resulting in a separate ground truth channel (referred to as the V channel) for each intervertebral disc. The number of intervertebral discs varies across samples in the Spine Generic dataset, ranging from 7 to 11. To account for this variability, we extract 11 intervertebral discs for each subject. If any intervertebral disc is missing, we consider its position as unknown and mitigate its impact on the training process by filtering it out using the visibility flag in the loss function given by equation 4.8.

$$X = \frac{1}{V * N} \sum_{i=1}^{V} \sum_{j=1}^{N} y_j - \hat{y}_j{}^2 \quad \text{Eq. 8}$$

Localization is a crucial initial step in achieving accurate segmentation of Cervical IVDs. In this study, the CLAHE technique and Deep Learning Stacked Hourglass Network are employed for localization in the given MRI images. The CLAHE technique effectively enhances and highlights the edges and boundaries of the cervical vertebrae. To further refine the image, an erosion-based mathematical morphological technique is applied, extracting the maximum response horizontally. This erodes the unwanted edges and boundaries, followed by a dilation process that expands the remaining edges and boundaries. The resulting image allows for the automatic localization of the midpoint of the cervical IVD.

2.4.4. Post-Processing:

While the proposed network achieves a high level of accuracy in learning IVDs, further post-processing of the predicted masks is necessary to reduce the false positive rate. To address this, we propose a skeleton-based approach for fine-tuning the predicted results. In this approach, we create a general skeleton model based on the training set. For each subject in the training set, we extract the location of the intervertebral discs and calculate the distance from each intervertebral disc (v^i) to the first intervertebral disc (v^1). This provides information about the relational structure among the intervertebral discs. To normalize this representation, we shift v^1 to the world coordinate (0, 0) and divide all the relational distances by the distance from v^1 to v^5. Each intervertebral disc $v^i$is calculated based on the average location of all subjects' IVDs. Using this generated skeleton, we define the skeleton model S as equation 9:

$$S = set\{v^i\}, i = 1,2, \ldots . V, v^i = v(x, y) \quad \text{Eq. 9}$$

The skeleton model S comprises V intervertebral discs with their respective 2D positions (x, y). During testing, for each predicted mask, a search tree is generated. Each node in this tree represents a possible combination of intervertebral disc locations in a specific order. Along each path in the tree (from leaf to node, forming a S'), the error function between the general skeleton S and the predicted skeleton S' is computed using equation 10. It's important to note that for each intervertebral disc (v^i), multiple candidates are available, denoted as (v^i, c) where c represents the c_{th} candidate for v^i. The solution to equation 10 is the S' that minimizes the error. To account for missing intervertebral discs, we introduce the flag δ in equation 10, indicating the availability of each candidate. Consequently, if the algorithm fails to detect a specific intervertebral disc, it will not impact the error function.

$$error(S, S') = \sum_{j=1}^{N} d(v^i, v^{j'}), d(v^i, v^{j'}) = \sqrt{\sum_{i=1}^{n} \delta (v_i^j - v'^j_i)^2} \quad \text{Eq. 10}$$

2.4.5. Feature Extraction:

Extracting geometric features from objects within an image is a common task in biomedical image processing, computer vision and image processing. Geometric features describe the shape, size, orientation, and position of objects in an image. Here are some commonly used geometric features and their mathematical formulas:

Area (A): The area of an object is a measure of its size. It is calculated by eq. 11.

A = Number of Pixels in the Object Eq. 11

Perimeter (P): The perimeter of an object is the total length of its boundary. Perimeter can be calculated using eq. 12.

Sum of the lengths of all boundary pixels $p = \sum I(b)$ Eq. 12

Eccentricity (E): Elongation measures how stretched or elongated an object is. Eccentricity can be calculated using eq. 13.

E = (Major Axis Length) / (Minor Axis Length) Eq. 13

Orientation (θ): Orientation represents the angle at which an object is oriented. It can be calculated using eq. 14

θ = 0.5 * arctan((2 * Covariance_xy) / (Covariance_xx - Covariance_yy)) Eq. 14

Centroid (X_c, Y_c): The centroid is the center point of an object. Coordinates of the centroid can be calculated using eq. 15&16.

X_c = (Sum of x-coordinates of object pixels) / (Number of object pixels) Eq. 15

Y_c = (Sum of y-coordinates of object pixels) / (Number of object pixels) Eq. 16

2.4.6. GABOR feature extraction:

The GABOR descriptor operates by accumulating gradient directions across pixels within small spatial regions known as "cells." It then constructs a 1D histogram based on these accumulations, and the concatenation of these histograms forms the feature vector used for subsequent analysis. Let's assume there's an intensity (grayscale) function denoted as L that represents the image under examination. This image is divided into cells, each consisting of N×N pixels, as depicted in Figure 6(b). The orientation $\theta_{x,y}$of the gradient at each pixel is computed (as shown in Figure 6(b) and 6(d)) using the following rule:

GABOR represents a feature descriptor technique that centers on feature extraction. Within this image feature descriptor, only valuable information is retained, and irrelevant details are discarded. In the GABOR approach, the entire image is divided into small blocks, and a feature descriptor is constructed for each of these blocks [28]. Once the pertinent features from each image block are extracted, these blocks are assembled and subsequently normalized to derive contrast-normalized features. The process of Histogram of Oriented Gradients (GABOR) feature extraction encompasses multiple steps, as outlined in equation 8.1.

$$\theta_{x,y=tan^{-1}} \frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \quad \text{Eq. 8.1}$$

Furthermore, the orientations $\theta_i^j i = 1 \ldots . N^2$The gradient directions that fall within the same cell (let's call it "cell j") are quantized and combined into a histogram with M bins. Ultimately, all these individual histograms are arranged and joined together to form a single GABOR histogram. This consolidated GABOR histogram represents the final result of this algorithmic step, serving as the feature vector for subsequent processing.

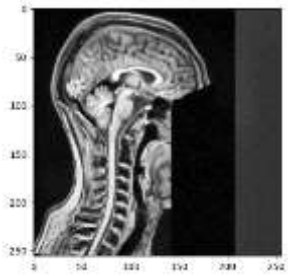


Figure 6(a).pre-processed image with CLAHE.

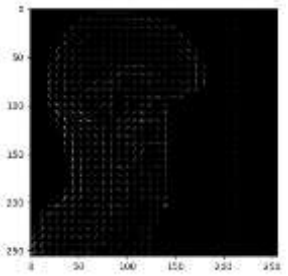


Figure 6(b). Extracted GABOR features

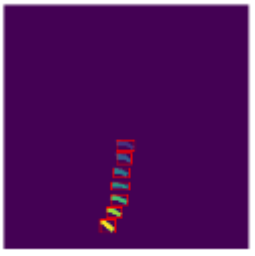


Figure 6(c). Segmented IVD

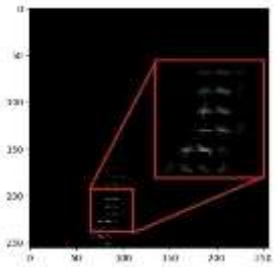


Fig 6(d). Extracted GABOR from IVD

3. Experimental setup

To assess the effectiveness of our proposed method, we utilize the Spine Generic Dataset [1]. This dataset comprises T1-weighted (T1w) contrasts for each subject, obtained from 42 different centers worldwide. The dataset exhibits significant variations in terms of image quality, scale, and imaging devices, presenting a challenging benchmark for intervertebral disc labeling.

The proposed model is trained for 150 epochs using Adam optimization, with a learning rate of 0.00025 and a batch size of 4. Through experimentation, we obtained the best results on the validation set using 2 stacks. The implementation and model training were conducted using ivadomed [12], and the method can be readily applied through the Spinal Cord Toolbox [2].

In order to compare our method with the existing literature, we adopt the same settings as described in [16].The figure 7 provided below illustrates the architecture of the model.

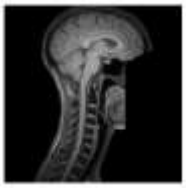


Fig. 7 (a) Original image and Corresponding ground truth mask

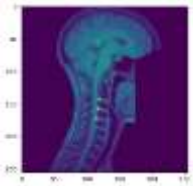


Fig. 7 (b) Attention map visualization.

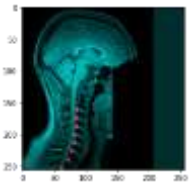


Fig. 7 (c) Prediction results in representative images.

3.1. IVD SEGMENTATION USING ACTIVE REGION GROWING TECHNIQUE:

The approach relies on identifying and extracting the most significant outlines in the image, which are distinguished by their considerable length relative to the entire image and by noticeable differences in color and texture between separated regions. Additionally, seeds are positioned in contour-free areas, or the central "core" of the regions. Each region is represented by a Gaussian distribution, where the mean and standard deviation initially derive from the seeds and depict the behavior of the homogeneous region. Consequently, the likelihood of a pixel (x, y) belonging to a region characterized by (μ, σ) can be determined.

$$P_{R(x,y)|(\mu,\sigma)=}\frac{1}{\sqrt{2\pi\sigma}}exp\left\{\frac{(I_{(x,y)}-\mu)^2}{2\sigma^2}\right\} \qquad \text{Eq. 17.}$$

Here, $I_{(x,y)}$represents the intensity of the pixel (x, y). The background is considered as a unified region with a uniform probability distribution denoted as P.

The objective of image segmentation is to divide the image into subregions with uniform intensity properties within their interiors and distinct boundaries from neighboring regions. To achieve an optimal segmentation, a global energy function is defined with two fundamental components. The boundary term quantifies the likelihood that boundary pixels represent actual edge pixels. This probability, denoted as$P_B(x, y)$, is directly proportional to the magnitude gradient of the pixel, indicating its probability of being on a real boundary. On the other hand, the region term assesses the homogeneity within the interior of the regions by measuring the probability of a pixel (x, y)belonging to a region modeled byμ, σ, denoted as$P_R(x, y)|\mu, \sigma$.

To establish these definitions, additional concepts are necessary. Let $\rho(R) = \{R_i: i\epsilon[0, N]\}$represent a partition of the image into $\{N + 1\}$non-overlapping regions, where R_0 corresponds to the background region. Furthermore, let $\partial_p(R) = \{\partial R_i: i\epsilon[1, N]\}$denote the boundaries of the partition $\rho(R)$. The energy function is then formulated as follows:

$$E\big(\rho(R)\big) = (1-\alpha)\sum_{i=1}^{N} -logP_B\big((x,y):\epsilon\partial R_i\big) + \alpha\sum_{i=1}^{N} -logP_R((x,y):(x,y)\epsilon R_i|(\mu_i,\sigma_i) \qquad \text{Eq. 18}$$

Here, represents a model parameter that determines the relative importance of the boundary probability and region homogeneity terms. The energy function is optimized using a region competition algorithm, which considers the neighboring pixels surrounding the current region boundaries $\partial\rho(R)$ to determine the next step. Specifically, a region incorporates a neighboring pixel if this updated classification reduces the overall segmentation energy.

The diagnosis process heavily relies on accurately identifying the boundaries of the intervertebral discs (IVDs) since disc bulges can cause shape deformations. If the boundaries are not properly segmented, it becomes difficult to detect abnormalities in the IVD. To address this, a new technique called active region growing is proposed to segment the IVDs from localized spine MR images. The underlying concept of this technique is to gather pixels with similar values to establish a boundary. In the Cervical IVD, there are two layers: the nucleus pulposes (inner layer) and the annulus fibrosis (outer layer). In MR images, the nucleus pulposes appears as a grey color, while the annulus fibrosis appears black. In grayscale images, pixel values range from 0 to 255, with 0 representing black and 255 representing white, and the values in between represent various shades of grey. In the proposed method, an initial pixel with a value less than 35 (closest to black) is selected from the boundaries formed by the CLAHE to mark the annulus fibrosis. Seed point for region growing is collected from Stacked Hourglass Network. The neighboring pixel values are then examined, and if they fall between 35 - 40, they are clustered with the initial pixel; otherwise, they are disregarded. This process is repeated until the entire image is examined, resulting in the identification of the annulus fibrosis, as shown in Figure 4.10 (b) where the boundaries of the annulus fibrosis are highlighted. Next, the inner layer, the nucleus pulposes, is filled using the hole filling method. In this method, pixels that cannot be reached within the marked boundaries are replaced with colored pixels. The resulting image is displayed in Figure 4.10 (c).

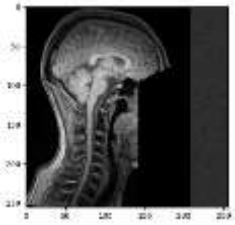


Figure8 (a). Original IVD image

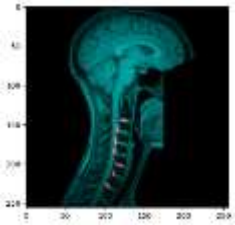


Figure8 (b). Region Growing Result

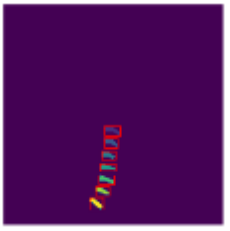


Figure8 (c). Segmented IVDs

Figure 9(a). original image **Figure 9(b).**Extracted IVD within bounding box.

	centroid-0	centroid-1	orientation	area	eccentricity	perimeter
0	162.336449	92.607477	1.357504	107	0.937184	46.935029
1	176.040541	91.635135	1.230129	74	0.961557	43.177670
2	189.285714	90.232143	1.296096	56	0.959776	35.899495
3	201.892857	88.303571	1.438244	56	0.955817	35.485281
4	215.518868	88.320755	1.170124	106	0.896507	45.177670
5	228.840580	82.724638	1.114377	69	0.939622	41.420310
6	244.082192	77.041096	0.912862	73	0.967294	44.112698

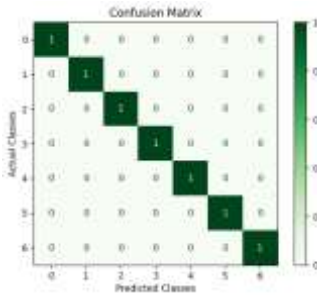
Fig. 10a Extracted features from sample 1.

3.2. Evaluation Metrics:

To showcase the effectiveness of our proposed method we employ various evaluation metrics. One of the metrics utilized is the Dice Overlap Coefficients, which measures the precision of the prediction by calculating the pixel overlaps percentage between each predicted intervertebral disc position and the corresponding ground truth along the superior-inferior axis.

4. Results and Discussion:

In this section we describe results obtained by using various performance measures. For reducing the space we have given some of the sample results.



	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	1
5	1.00	1.00	1.00	1
6	1.00	1.00	1.00	1
accuracy			1.00	7
macro avg	1.00	1.00	1.00	7
weighted avg	1.00	1.00	1.00	7

Fig. 11(a) Confusion matrix for Sample 1

Fig. 11(b) Precision, Recall and F1 score for sample 1

Fig11(a)depictstheeventsthatassignTruePositive(TP)7discscorrectly localized as per the ground truth , False Negative (FN) 0 discswrongly classified, True Negative (TN) 0 correctlyclassified,andFalsePositive(FP) 0wronglyclassifieddiscs. The performance metrics for classification is given in Section(3.3.1).

According to the figure 11(b) the precision measure evaluates the accuracy of positive

predictions made by the confusion matrix from Fig 11(a). A precision value of 1.00 indicates that 100% of the IVD instances predicted as positive by the Stacked Hourglass Model were indeed true positives, while the remaining (0)% were false positives. A recall value of 1.00 implies that 100% of the actual positive instances were successfully captured by the model, while the remaining (0)% were false negatives. Altogether, F1 score indicates a better trade-off between precision and recall that is 1.00 for the sample 1.

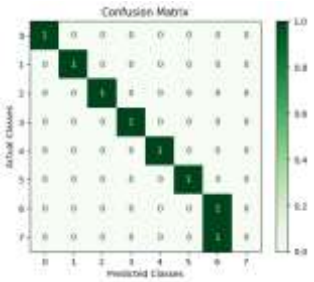


Fig. 11(c) Confusion matrix for Sample 2

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	1
5	1.00	1.00	1.00	1
6	0.50	1.00	0.67	1
7	0.00	0.00	0.00	1
accuracy			0.88	8
macro avg	0.81	0.88	0.83	8
weighted avg	0.81	0.88	0.83	8

Fig. 11(d) Precision, Recall and F1 score for sample 2

Fig11(c)depictstheeventsthatassignTruePositive(TP)6discscorrectly localized as per the ground truth , False Negative (FN) 1 discswrongly classified, True Negative (TN) 0 correctlyclassified,andFalsePositive(FP) 0wronglyclassifieddiscs. The performance metrics for classification is given in Section(3.3.1).

According to the figure 11(d) the precision measure evaluates the accuracy of positive predictions made by the confusion matrix from Fig 11(c). A precision value of 0.81 indicates that 88% of the IVD instances predicted as positive by the Stacked Hourglass Model were indeed true positives, while the remaining (19)% were false positives. A recall value of 0.88 implies that 88% of the actual positive instances were successfully captured by the model, while the remaining (12)% were false negatives. Altogether, F1 score indicates a better trade-off between precision and recall that is 0.88 for the sample 2.

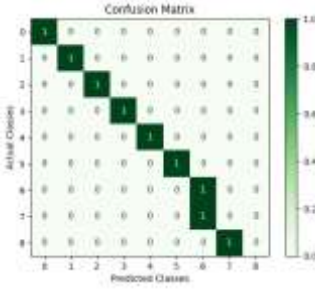


Fig. 11(e) Confusion matrix for Sample 3

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	1
5	1.00	1.00	1.00	1
6	0.50	1.00	0.67	1
7	0.00	0.00	0.00	1
8	0.00	0.00	0.00	1
accuracy			0.78	9
macro avg	0.72	0.78	0.74	9
weighted avg	0.72	0.78	0.74	9

Fig. 11(f) Precision, Recall and F1 score for sample 3

Fig11(e)depictstheeventsthatassignTruePositive(TP)6discscorrectly localized as per the ground truth , False Negative (FN) 2 discswrongly classified, True Negative (TN) 0 correctlyclassified,andFalsePositive(FP) 0wronglyclassifieddiscs. The performance

metrics for classification is given in Section(3.3.1).
According to the figure 11(f) the precision measure evaluates the accuracy of positive predictions made by the confusion matrix from Fig 11(e). A precision value of 0.72 indicates that 72% of the IVD instances predicted as positive by the Stacked Hourglass Model were indeed true positives, while the remaining (28)% were false positives. A recall value of 0.78 implies that 78% of the actual positive instances were successfully captured by the model, while the remaining (22)% were false negatives. Altogether, F1 score indicates a better trade-off between precision and recall that is 0.78 for the sample 3.

Table3. ClassificationMetrics

		Classification Metrics			
Sample. No.	**Total IVDs per Image**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
1.	7	100%	1	1	1
2.	8	88%	0.81	0.88	0.83
3.	8	88%	0.81	0.88	0.83
4.	9	78%	0.72	0.78	0.74
5	8	88%	0.81	0.88	0.83
6	8	100%	1	1	1
7	7	100%	1	1	1
8	9	67%	0.56	0.67	0.59
9	8	88%	0.81	0.88	0.83
10	7	100%	1	1	1
11	7	100%	1	1	1
12	8	88%	0.81	0.88	0.83
13	8	88%	0.81	0.88	0.83
14	9	78%	0.72	0.78	0.74

15	8	88%	0.81	0.88	0.83
16	8	100%	1	1	1
17	9	78%	0.72	0.78	0.74
18	7	100%	1	1	1
19	8	88%	0.81	0.88	0.83
20	7	100%	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮
658	7	100%	1	1	1

Table 4. Overall Accuracy

Total IVD tested	**Correctly localized IVDs**	**Wrongly localized IVDs**
3950	3500	450
Accuracy		88.60%

Overall accuracy, precision, recall and F1 score is shown in table x.

Accuracy	**Precision**	**Recall**	**F1 Score**
0.88	1.0	0.88	0.93

In this experiment, we analyze the accuracy of the localization and segmentation of cervical IVDs with the help of Dice coefficient metrics. The Dice coefficient accepts predicted pixels and ground truth pixels. This will help into calculate percent of pixels in image that are classified correctly.The predicted labels from the table 3, it was observed that out of 3950 IVDs from 658 MRI images 3500 IVDs are localized correctly. Similarly, 450 IVDs out of 3950 IVDs system has mis-classified. The overall accuracy of the proposed system is 88.60%.

The confusion matrix, precision, recall and F1 score curve are depicted in figure 11. The diagonaly conferred elements are correctly segmented IVD labels, while misclassified IVD are placed out of the diagonal. For the better visualization and simple understanding some the predicted IVD are given in the figure 11. The precision, recall and F1-score are computed from confusion matrix. The results were obtained from the input of the Dice Coefficient. The results obtained are tabulated in Table 3 and 4.

Conclusion

In this research study, we conducted an automatic localization of the Cervical Intervertebral Disc (IVD) using the feature combination of geometric features and texture feature descriptor GABOR. We have employed Stacked Hourglass Network a deep learning technique with the help of CLAHE for localization and segmentation of IVD. To simplify the diagnosis process for radiologists, we employed a novel active refined region growing technique for segmentation. The effectiveness of our proposed method was evaluated on an MRI dataset comprising 3950 IVDs from 39 subjects.The proposed method achieved accuracy in localizing the cervical IVDs, with 88.60% accuracy for both training and testing data, the significant advantage of this approach is its automated nature, eliminating the need for manual intervention while providing improved localization and segmentation of cervical IVDs. The segmented IVDs can now be categorized as normal, disc desiccation, and disc bulge.

References

[1] Ract I, Meadeb JM, Mercy G, Cueff F, Husson JL, Guillin R. A review of the value of MRI signs in low back pain. DiagnInterv Imaging. 2015;96:239–49.

[2] Dagenais S, Galloway EK, Roffey DM. A systematic review of diagnostic imaging use for low back pain in the UnitedStates. Spine J. 2014;14:1036–48.

[3] Lzzo R, Popolizio T, D'Aprile P, Muto M. Spinal pain. Eur J Radiol. 2015;84(5):746–56.

[4] Siemund R, Thurnher M, Sundgren PC. How to image patients with spine pain. Eur J Radiol. 2015;84:757–64.

[5] Ghosh S, Malgireddy MR, Chaudhary V, Dhillon G. A new approach to automatic disc localization in clinical lumbarMRI: combining machine learning with heuristics. 9th IEEE International Symposium on Biomedical Imaging (ISBI).2012;114–117.

[6] Alomari RS, Corso JJ, Chaudhary V. Labeling of lumbar discs using both pixel-and object-level features with a two-level probabilistic model. IEEE Trans Med Imaging. 2011;30:1–10.

[7] Michopoulou SK, Costaridou L, Panagiotopoulos E, Speller R, Panayiotakis G, Todd-Pokropek A. Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine. IEEE Trans Biomed Eng. 2009;56:2225–31.

[8] Ghosh S, Alomari RS, Chaudhary V, Dhillon G. Composite features for automatic diagnosis of intervertebral disc herniation from lumbar MRI. Paper presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2011; IEEE.

[9] Chwialkowski M, Shile P, Peshock R, Pfeifer D, Parkey RW. Automated Detection and Evaluation of Lumbar Discs in MR Images. Paper presented at: Images of the Twenty-First Century.Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 1989; IEEE.

[10] Hao S, Jiang J, Guo Y, Li H. Active learning based intervertebral disk classification combining shape and texture similarities. Neurocomputing. 2013;101:252-257. https://doi.org/10.1016/j. neucom.2012.08.008.

[11] Beulah A, Sharmila TS. Classification of Intervertebral Disc on Lumbar MR Images using SVM. Paper presented at: 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT); 2016; IEEE.

[12] Tsai M-D, Jou S-B. Hsieh M-S. a new method for lumbar herniated inter-vertebral disc diagnosis based on image analysis of transverse sections. Comput Med Imaging Graph. 2002;26(6):369-380. https://doi.org/10.1016/S0895-6111(02)00033-2.

[13] Michopoulou S, Boniatis I, Costaridou L, Cavouras D, Panagiotopoulos E, Panayiotakis G. Computer assisted characterization of cervical intervertebral disc degeneration in MRI.J Instrum. 2009;4(05):P05022. https://doi.org/10.1088/1748-0221/ 4/05/P05022.

[14] Beulah A, Sharmila TS, Pramod V. Disc bulge diagnostic modelin axial lumbar MR images using intervertebral disc descriptor(IdD). Multimed Tools Appl. 2018;77(20):27215-27230. https://doi.org/10.1007/s11042-018-5914-8.
[15] Unal Y, Polat K, Kocer HE, Hariharan M. Detection of abnormalities in lumbar discs from clinical lumbar MRI with hybrid models.Appl Soft Comput. 2015;33:65-76. https://doi.org/10.1016/j.asoc.2015.04.031.
[16] Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D.Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. Neuroimage. 2013;65:167-175.https://doi.org/10.1016/j.neuroimage.2012.09.065.
[17] Lee SLA, Kouzani AZ. Hu EJ. Random forest based lung nodule classification aided by clustering. Comput Med Imaging Graph. 2010;34(7):535-542. https://doi.org/10.1016/j.compmedimag.2010.03.006.
[18] Azar AT, Elshazly HI, Hassanien AE, Elkorany AM. A random forestclassifier for lymph diseases. Comput Methods Programs Biomed.2014;113(2):465-473. https://doi.org/10.1016/j.cmpb.2013.11.004.
[19] Aboul-Yazeed RS, Mohamed AS, El-Bialy A. Edge-Based IVDSegmentation System. Middle East Conference on BiomedicalEngineering (MECBME); 2014; IEEE.
[20] Peng ZG, Zhong J, Wee W, Lee JH. Automated vertebra detection and segmentation from the whole spine MRimages. 2005 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2005;2527–2530.
[21] Castro-Mateos I, Pozo JM, Lazary A, Frangi AF. 2D segmentation of intervertebral discs and its degree of degeneration from T2-weighted magnetic resonance images. Medical imaging 2014. Comput Aided Diagn. 2014;9035:17.
[22] Haq R, Aras R, Besachio DA, Borgie RC, Audette MA. 3D lumbar spine intervertebral disc segmentation and compression simulation from MRI using shape-aware models. Int J Comput Assist Radiol Surg. 2015;10:45–54.
[23] Law MWK, Tay K, Leung A, Garvin GJ, Li S. Intervertebral disc segmentation in MR images using anisotropic oriented flux. Med Image Anal. 2013;17:43–61.
[24] Chevrefils C, Cheriet F, Grimard G, Aubin C. Watershed segmentation of intervertebral disk and spinal canal from MRI images. Image Anal Recognit. 2007;4633:1017–27.
[25] Neubert A, Fripp J, Engstrom C, Schwarz R, Lauer L, Salvado O, et al. Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys Med Biol. 2012;57:8357–76.
[26] Neubert A, Fripp J, Engstrom C, Walker D, Weber MA, Schwarz R, et al. Three-dimensional morphological and signal intensity features for detection of intervertebral disc degeneration from magnetic resonance images. J Am Med Inform Assoc. 2013;20:1082–90.
[27]Oktay AB, Akgul YS. Simultaneous localization of lumbar vertebrae and intervertebral discs with SVM-based MRF.IEEE Trans Biomed Eng. 2013;60:2375–83.
[28]Carcagnì, Pierluigi, Marco Del Coco, Marco Leo, and Cosimo Distante. "Facial expression recognition and histograms of oriented gradients: a comprehensive study." SpringerPlus 4, no. 1 (2015): 1-25.