

# MODELLING A DEEP KERNEL-BASED LEARNING APPROACH FOR SPINAL CORD INJURY PREDICTION

P.R.S.S.VENKATAPATHIRAJU<sup>1</sup>, Dr.V.ASANAMBIGAI<sup>2</sup>, Dr. SURESH BABU MUDUNURI<sup>3</sup>

<sup>1</sup>Research Scholar, Department of CSE, Annamalai University, Annamalainagar, India

<sup>2</sup>Assistant Professor, Department of CSE, Annamalai University, Annamalainagar, India

<sup>3</sup>Associate Professor, Department of IT, SRKREC, Bhimavaram, India

E-mail: <sup>1</sup>prssvraju2@gmail.com, <sup>2</sup>tradingbaskeran@gmail.com, <sup>3</sup>sureshmudunuri@gmail.com

## ABSTRACT

Deep Learning (DL)-based spinal cord injury (SCI) prediction is an unswerving model in medical imaging. However, the prediction process is challenging during the segmentation and classification process. Conventionally, radiologists examine the spinal cord images to indicate the abnormalities in the spinal cord manually. The manual high dimensional spinal cord feature space interpretation makes it complex to identify the severity level. However, DL approaches help in faster and more accurate predictions. The model intends to predict the abnormal and normal spinal cord images automatically. Under certain conditions, the weight of the training images is based on similarity distribution which is likely to show better performance. It does not deal with unrepresentative data. Thus, the kernel-based learning process intends to reduce the difference between the testing and training data and explore the kernel learning value for image weighting. Here, a novel kernel-based weighting method reduces the maximal mean discrepancy (MMD) among the testing and training data, facilitating kernel and weighted image optimization. Experimental results demonstrate that the anticipated kernel-based image weighting model has a higher computational disorder prediction ability than other approaches. The numerical outcomes prove that the model has superior performance to the prevailing approaches regarding prediction accuracy, recall and precision.

**Keywords**— *Deep Learning, Spinal Cord, Weight Computation, Kernel Learning, Discrepancy Analysis*

## 1. INTRODUCTION

Every year about 2,50000 to 5,00000 patients worldwide are affected by the injury to the spinal cord, which is a crucial world health issue based on the World Health Organization (WHO) [1]. The fourth highest reason for death in the world is SCI having an estimated rate of mortality to rise by 2030 to the third most excellent place. Because of the trauma from the accidents and falls, spinal cord injury occurs. An important task is diagnostic imaging when a spinal cord injury is diagnosed. T1- and T2-weighted regions are commonly required to capture the spinal cord sequences [2]. The contrast and brightness of the tissue in the spinal cord need to be determined. Repetition time (RT) and Time to echo (TE) produce the T-1 weighted image. On the other hand, long TE and RT produce the T-2 image. The notification of T2 weight is not noticed when the injury occurs. Yet, hyper-intensity needs to be found on T2 images when the injury happens. Henceforth, it is considered for the efficient prediction process in the primary phase [2]. Due to vasogenic edema and cytotoxic, hyper-intensity is

caused can turn into a hemorrhage. The enhanced T2 weighted imaging approach [3] is the T2\*. The methods are ineffective in diagnosing the spinal cord injury.

Mainly, spinal cord injury is a neurological injury that turns to emotional damage and dangerous medications. During the functional change prediction, the symptoms are not helpful always. Magnetic resonance imaging schemes are established in spinal cord injury to examine the soft tissue. Compared with the T2 weighted imaging technique, the diffusion-weighted MRI [4] is good for efficiency in diagnosis. Many conventional hemorrhage/edema features in the spinal cord need to analyze by using the MRI approach. Based on extensive review, edema/hemorrhage interstitial fibrosis are concerned with considerable signal intensity modification [5].

In the computer vision domain, the image segmentation is concerned as a significant role. The isolation of the image plane into the non-overlapping regions is the primary objective of the technique [6]. Henceforth, thresholds are differed based on the

feature space. The multilevel thresholding approach is implemented in most real-time applications to reduce the error rate and enhance accuracy. The histogram of the provided image is split into various groups, and the fixed intensity value is assigned in the multilevel thresholding approach to every individual group. The guided filter technique for segmentation of the image features is proposed by L. Guo et al. for non-overlapping regions [7]. The parametric and non-parametric approaches are preferred to detect the optimal thresholds. The statistical parameters are evaluated via the criteria with optimization in the non-parametric methods like the measures of entropy and the class variance maximization [8]. This technique is attained by segmentation, the most popular [9]. The upgraded thresholding technique is implemented in spinal cord injury (SCI) images and detects brain lesions [10].



*Fig.1 Sample Spinal Cord Injured Image*

Moreover, deterioration is the demerits of the supervised learning approach; when performing with the differences between the test data and the trained data like the usage of various imaging protocols, scanners, and the differences in the groups of patients. Such differences are selected to adapt by the human observers [11]. Moreover, the supervised learning techniques find it difficult between the different images due to the results obtained in the differences in the feature space of the test samples and the training distributions. Specific differences are handled to design the transfer learning techniques between the test and the trained data with the sample distribution differences [12]. Based on the weighted training samples, many transfer learning techniques are described in the segmentation of medical images. Complete training

images and individual weighting samples (voxels) are used. The advantage of image weighting is required with no labeled training data compared to the sample weighting for handling the differences among scanners from the test scanner [13]. The weighting techniques are presented previously in the weight training samples that are no steps considered in representing the feature values of the test and trained samples to reduce the differences [14]. The training images are used as giving a positive weight or not using this as the weight of zero by the image weighting method. Combining the image weighting representation of the feature is proposed with the transfer phase. They are (i) the distributions of data with more similarity between the test data and the training data, and (ii) the various classes are separated.

The two effective kernel learning techniques are proposed in conjunction with image weighting. The multiple kernel learning (MKL) technique is discussed to lower the class distances and enlarge the class distance. Secondly, the MKL technique adds the MMD technique of the Sanchez et al. [15] framework. The outcomes suit to classify in the kernel space and give the test and training distributions the same. The suggested techniques are trained on the 10,000 training samples per image, which contradicts other techniques. Compared to using a standard Gaussian kernel, the kernel learning techniques are investigated in the proposed method, resulting in better segmentation. When using image weighting, learning kernels are checked to determine whether the performance is improved or not in the proposed system. Furthermore, the MMD measure is shown in the proposed method for determining the kernel and the weights of the image. The kernel and the image weights are jointly optimized to facilitate the optimization. Also, this work examines whether the learned kernels enhance the performance while using image weighting, and the measurement is used not only for kernels; however, it relies on image weights. Moreover, the weights are optimized to facilitate optimization.

The work is structured as follows: Section 2 provides a comprehensive analysis of various prevailing approaches. The methodology section is

elaborated in section 3, where the input image weighting is done based on testing and training data. The numerical outcomes attained after evaluating the anticipated model are shown in section 4, and the summary is provided in section 5.

## 2. RELATED WORKS

In the past 15 years, many techniques have been developed to segment the spinal canal/cord. Archie et al. [16] modelled an automatic spinal cord segmentation approach for computed tomography (CT). The anatomical structural mapping is used to predict the organs quickly in the neighborhood region. The approximate spinal canal position is determined. In the case of failure, segmentation is performed using the growing region and the active contours like snakes in 2D. Burnett et al. [17] explained the deformable modes efficiency in the spinal canal segmentation and CT volumes to identify the regions of interest (ROI). The known tissue intensity is used in the CT image to identify the nearby structures to provide canal segmentation or spinal cord initialization. The distance among the surface vertices and extracted points is done using the edge detection to initial position surface. The deformable model theory relaxes the surface to the bone edges. Nyul et al. [18] presented the semi-automatic spinal cord and segmentation approach in CT with the help of growing regions and the active contour approaches, which requires the user to identify one point in the center of the spinal cord.

On the other hand, the feasibility is developed by Rangayyan et al. [19] for detecting the spinal canal like the spinal cord on the axial CT images with the help of Hough transform [20]. Multiple morphological operations like closing, thresholding, and so on are used to follow the detection to initiate the region growing technique for extracting the exact shape of the spinal canal. Recently, SC segmentation techniques are more automatic, which require only a few interventions manually, and few are fully automated [21]. In recent days, the multi-atlas related segmentation technique has been used by Asman et al. [22] to give the fully automated SC segmentation and the internal structures. However, the automated SC segmentation technique is related to the convex continuous max-flow structure by

Pezold et al. [23]. These researchers merge the features of tubularity and cross-sectional similarity before enhancing the robustness of the artifacts and the noise. Alternative techniques are presented in these two techniques. Among these two, one is offered.

Moreover, no automated published techniques are segmented on the spinal canal and the SC. On the other hand, being powerful is essential compared to any field of view. Finally, there is no automatic generation of 3D models of the spinal canal and the SC where the vertebral level helps label the sub-regions to enable the intra-patient and the inter-patient precisely when compared to the thoracic spine and cervical anatomy. In recent days, the automated spinal cord segmentation technique has been developed for the images of MR [24]. Various fields of view like thoracic and cervical and multiple types of images like T1w, T2\*w, and T2w are validated by this technique.

The other considerable concern is the public availability technique. On the other hand, some techniques exist through free software [25] or paid software [26]. Others do not find that complicated to use the published segmentation technique in the previous form provided with the implementation of code that needs a few minor modifications. It is frequent to optimize the data type that needs to be validated by the published technique. The fully automatic structure is used in the proposed system for the spinal canal and the spinal cord segmentation by the thoracic and cervical vertebral levels. The two methods are adopted concerning the T1 and T2 weighted images and the MRI. This paper has essential contributions [27]. They are (a) the new enhancements are added concerning the adaptive contrary to the proposed method, and the sensitivity of the parameter is evaluated in the process of optimization, and (b) the proposed technique is extended to segment the spinal canal and the SC, (c) the framework is presented to combine the segmentation of SC along with the vertebral labeling [28] to provide the innovative generic coordinate system depends on the vertebral levels. However, open-source project is the priorities of making the free and publicly available method [29]. The unified

pipeline is suggested as fully automated. The measurements of SC are provided at the particular vertebral levels to enable the intra-subject and inter-subject objectives compared with the more significant throughput [30].

### 3. METHODOLOGY

This section provides a detailed analysis of the anticipated kernel-based weighting to diminish the variations encountered during the testing and training process. The model attains superior performance compared to other approaches. The maximal mean discrepancy is reduced and facilitates kernel and weight-based optimization. The model attains superior performance and gives promising solution.

#### 3.1. Pre-processing

In the proposed model, all the images are modified for the non-uniformity intensity. The brain mask's scaling of 4 to 96 percentile to the 0 to 1 interval is used to normalize all the image intensities. Image weighing is effectually adopted for performing segmentation for every training image with the provided weight. It relies on reducing the distance measure among the probability density function (PDF) of the testing and training images for segmentation purposes. However, the training set is gathered based on the training image and corresponding weights using class labels and sampling voxels based on weighted distribution. Subsequently, a kernel-based support vector machine (k-SVM) is trained where the testing images are segmented using a voxel-wise classification process. Kernel-based learning is implemented for every testing image, and classification is further performed with k-SVM with a learnable kernel.

#### 3.2. Notation

$\phi$	Mapping with kernel space
$K$	Kernel tricks ( $K$ )
$m$	Training images

$m_i$	Randomly sampled images
$x_j^{m_i} \in \mathbb{R}^n$	Vectors value for every feature
$X^{tr}$	Matrixes of all training samples ( $m \cdot n^{tr} * n$ )
$X^{te}$	Matrixes of all testing samples ( $n^{te} * n$ )
$P_{m_i}(x)$ and $P^{te}(x)$	Probability of training and testing samples.
$w_{m_i}$	Weighted vector where the weight of training images are $w = [w_1, w_2, \dots, w_m]^T$
$\sum_{m_i=1}^m w_{m_i} =  w  = 1$	Normalized and non-negative values
$n^{te}$	Number of testing samples
$n^{tr}$	Number of training samples

Consider a training image  $m_i$  with an average weight  $w_i$  outcome in the total training PDF, i.e. the weighted sum of every individual's training PDF. It is shown in Equation. (1):

$$p^{tr}(x) = \sum_{m_i}^m w_{m_i} P_{m_i}(x) \tag{1}$$

Then, the optimal weights  $w^*$  are selected, reducing the distance criteria among the comprehensive training and testing PDF as in Equation. (2):

$$w^* = \arg \min_w distance_w(P^{te}, P^{tr}) \tag{2}$$

### 3.3. PDF-based weighted analysis

The PDF-based weighted analysis evaluates the PDF where  $P_{m_i}$  depends on the  $P^{te}$  distance function using the kernel density estimator (KDE). It utilizes a distance function over the PDFs being assessed. The PDFs are evaluated from the samples using KDE with the Gaussian, and the kernel parameter  $\sigma^S$  is depicted with weighted rule formation as in Equation. (3):

$$\sigma^S = \left(\frac{4}{n+2}\right)^{\frac{1}{n+4}} \cdot n^{tr \frac{-1}{n+4}} \sigma^{tr} \quad (3)$$

There are diverse techniques to evaluate the distance between the testing and training PDFs. Here,  $\sigma^{tr}$  specifies the SD of training samples, and the samples are averaged over certain features. The selected  $\sigma^S$  reduces the mean-based square error among the evaluated and actual PDF for multivariate Gaussian kernel. The Kullback divergence is expressed as in Equation. (4):

$$\begin{aligned} KL(P^{te} || P^{tr}) &= \int_D P^{te}(x) \log\left(\frac{P^{te}(x)}{P^{tr}(x)}\right) dx \quad (4) \\ &\approx \frac{1}{n^{te}} \sum_{j=1}^{n^{te}} \log P^{te}(x_j^{te}) \\ &\quad - \frac{1}{n^{te}} \sum_{j=1}^{n^{te}} \log\left(\sum_{m_i=1}^m w_{m_i} P_{m_i}(x_j^{te})\right) \quad (5) \end{aligned}$$

Here,  $D$  represents  $P^{tr}$  and  $P^{te}$  domain. Therefore, the optimal weight is provided by enhancing  $\sum_{j=1}^{n^{te}} \log(\sum_{m_i=1}^m w_{m_i} P_{m_i}(x_j^{te}))$  under certain constraints  $w \geq 0$  and  $|w| = 1$ . The technique is provided to evaluate the distance among PDF (Bhattacharyya distance) in Equation. (6):

$$\begin{aligned} BD(P^{te} || P^{tr}) &= - \int_D P^{te}(x) P^{tr}(x) dx \quad (6) \end{aligned}$$

$$\approx - \frac{1}{n^{tr}} \sum_{x_i \in D} \sqrt{P^{te}(x_i)} \sqrt{\sum_{m_i=1}^m w_{m_i} P_{m_i}(x_i)} \quad (7)$$

Where  $x_i$  represents a random  $D$  value, like KL, the criteria are reduced for  $w$  to the constraints  $w \geq 0$  and  $|w| = 1$ .

### 3.4. Maximal mean discrepancy (MMD) based on the weighted measure

A novel image-based weighted technique based on maximal mean discrepancy reduction. This approach helps depict the image weight immediately, and it does not need PDF estimation. It is like the MMD sample weight method; however, it evaluates the weight of samples (group) of an individual sample. The weights are determined by reducing the minimal MMD among the training  $X^{tr}$  and  $X^{te}$  samples. The MMD among the two different datasets,  $X$  and  $Y$  is depicted as the distance among the sample mean  $x \in X$  and  $y \in Y$  over kernel space  $\phi$ :

$$\begin{aligned} MMD(X, Y) &= \left\| \frac{1}{n^X} \sum_{i=1}^{n^X} \phi(x_i) - \frac{1}{n^Y} \sum_{j=1}^{n^Y} \phi(y_j) \right\|^2 \quad (8) \end{aligned}$$

Here,  $n^X$  and  $n^Y$  represent the total samples  $x$  and  $y$ , respectively. The weighted images are provided by reducing the maximal mean among the training set  $X^{tr}$ , which is equipped with a weight/image and test set  $X^{te}$ :

$$\begin{aligned} MMD(X^{tr}, X^{te}) &= \left\| \frac{1}{n^{tr}} \sum_{m_i=1}^m w_{m_i} \sum_{j=1}^{n^{tr}} \phi(x_j^{m_i}) - \frac{1}{n^{te}} \sum_{i=1}^{n^{te}} \phi(x_i^{te}) \right\|^2 \quad (9) \end{aligned}$$

$$\begin{aligned}
 MMD(X^{tr}, X^{te}) &= \frac{1}{n^{tr^2}} w^T K^{tr, tr} w \\
 &\quad - \frac{2}{n^{tr} n^{te}} w^T k^{tr, te} \\
 &\quad + \frac{1}{n^{te^2}} k^{te, te}
 \end{aligned} \tag{10}$$

Here,  $K^{tr, tr}$  represents  $m * m$  inner product matrix among training images,  $k^{tr, te}$  specifies the vector length  $m$  of the products among every training and testing image,  $k^{te, te}$  specifies the single value providing the testing sample product. The expressions for  $K^{tr, tr}$ ,  $k^{tr, te}$  and  $k^{te, te}$  are provided, and the supplementary files are expressed as in Equation. (11):

$$\begin{aligned}
 MMD(X^{tr}, X^{te}) & \\
 = \omega^T M \omega; \text{ where } \omega^T &= [w^T, 1]
 \end{aligned} \tag{11}$$

$$M = \begin{bmatrix} \frac{1}{n^{tr^2}} K^{tr, tr} & -\frac{1}{n^{tr} n^{te}} k^{tr, te} \\ -\frac{1}{n^{tr} n^{te}} k^{tr, te^T} & \frac{1}{n^{te^2}} k^{te, te} \end{bmatrix} \tag{12}$$

With the provided kernel  $K$ , the optimal weights are provided by reducing Eq. (10) for  $\omega$  under certain constraints  $\omega \geq 0$ ,  $|w| = 1$ .

### 3.5. Kernel-based learning model

This work attempts to predict kernel space that diminishes the variation among the datasets. While probing for a kernel matrix, the work intends to pose certain constraints to fulfil that the matrix is symmetric. It is a relatively more accessible way to predict a specific kernel matrix where the set of  $in^k$  pre-defined kernel  $K_k^b$  ( $k = 1, 2, \dots, n^k$ ). Kernel is depicted as the optimal linear representation of base kernels as in Equation. (13):

$$K = \sum_{k=1}^{n^k} v_k K_k^b \tag{13}$$

Here, kernel weight  $v = [v_1, v_2, \dots, v_{n^k}]^T \geq 0$  as  $K$  represents linear kernel combination, and it

is fulfilled as a kernel. Author et al. [30] anticipated a kernel learning-based approach that optimizes kernel weight  $v_k$  in Eq. (13) by maximizing the kernel alignment among the ideal and learned kernel ( $K^l$  and  $K$ ). The  $K^l$  is equal to 1 when the samples pose the similar class labels, and  $-1$  is identical to have various class labels. It is related to the situation where  $\Phi$  of all models with class labels are in the exact location and models with multiple labels are different. The kernel alignment centres the sample distribution over the kernel space, which specifies the sample expectation from kernel space to zero. By attaining the base kernels,  $K_k^{bc}$  ( $k = 1, 2, \dots, n^k$ ), weights  $v$  are selected to enhance the alignment among  $K^l$  and  $K$  and expressed in Equation. (14):

$$\begin{aligned}
 v^* &= \\
 \arg \min_K \theta \text{ distance}_K(P^{te}, P^{tr}) &+ \\
 f_K(X^{tr}, y^{tr})
 \end{aligned} \tag{14}$$

Here,  $distance_K$  specifies the PDF distance function of  $K$  kernel space,  $f_K$  specifies the classification error of  $K$  kernel space, and  $\theta$  establishes parameter trade-offs among various terms. With  $f_K$ , the structural function of SVM are computationally higher when different training samples are utilized. The kernel alignment is evaluated effectually for many training samples consisting of weighted subsets. Therefore, the kernel alignment values are used and multiplied with  $-1$  when the value is higher. From Eq. (9), the MMD among the testing and training distributions is evaluated when the training images are given weight ( $\omega = [\omega^T, 1]^T$ ) or when no weight is utilized ( $\omega = [\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}, 1]^T$ ). The kernel  $K$  is composed of a linear base kernel combination, and the matrix is expressed as in Equation. (15):

$$\begin{aligned}
 MMD(X^{te}, X^{tr}) &= \\
 \sum_{k=1}^{n^k} v_k \omega^T M_k \omega
 \end{aligned} \tag{15}$$

When the MMD weight and kernel learning are optimized with Equation. (16):

$$\begin{aligned}
 [w^*, v^*] = & \arg \min_{w,v} \theta \text{MMD}_{w,v}(P^{te}, P^{tr}) - \\
 & \text{kernel}_{alignment}(X^{tr}, y^{tr}) \quad (16) \\
 = & \arg \min_{w,v} \theta \sum_{k=1}^{n^k} v_k \begin{bmatrix} w \\ 1 \end{bmatrix}^T M_k \begin{bmatrix} w \\ 1 \end{bmatrix} \\
 & - \frac{(\sum_{k=1}^{n^k} v_k K_k^{bc}, K^l)}{\|\sum_{k=1}^{n^k} v_k K_k^{bc}\|_F} \quad (17)
 \end{aligned}$$

Equation. (17) is used for optimizing the kernel weight  $v$  when MMD kernel learning is done without MMD weighting (image-based). The numerical precision is evaluated based on MMD when kernel normalization intends to reduce the feature space, and the optimal  $\theta$  value is predicted effectually.

#### 4. NUMERICAL ANALYSIS

The results related to experiments are taken from the dataset on the SCI images. Therefore, the regions are segmented using the trained spinal cord vertebra dataset with the help of suggested thresholding-related segmentation models (<https://osf.io/nqjyw/>). The spinal cord patterns are presented in the dataset having the calculated data in the format of nii, which has all the images of both females and males. The simulation is done in Python environment. The proposed method is specified on the injury detection of real-time spinal cord vertebra is considered by the various hospital to test. The test of medical spinal cord images is described in Fig 2, which is regarded for automatically testing the segmented regions from hospital to data classification. The test spinal cord images are classified in the proposed system as having a less rate of error and a true high rate of positive.

##### 4.1. Image Weighting

The baseline technique is used, and it performs well than the weighting technique for all three applications. The weighting technique outperforms in various applications. Kernel-based weighting performance is better than other techniques for segmentation and is also analysed. The percentage of weighted voxels is used to

perform this. It generates the weighting technique for the overall image similarity to weigh considerably more than the similarity of the learning model (See Fig 1 to Fig 3). It is essential to consider that the weighted voxels are very high in intensity in the scan.

On the other hand, other voxels are not. Hence, the voxels of weighted learning are placed in a low  $P(x)$  position. Since the focus is on different parts of the distribution  $P(x)$ , i.e. generated using the log as the differences among the test and training distributions in the location where the intensity is the small contribution to the kernel learning and the weighting performance is better than maximal mean weighting. The differences are more minor among the three weighting techniques for the hippocampus and brain tissue segmentation. Here, the learning model performs better for the brain tissue segmentation for the hippocampus segmentation. Also, maximal mean is the secondary better in performance in the applications.

Image weighting performs better than kernel learning, especially brain tissue segmentation, during individual use. These two techniques are combined, giving them less extra enhancement in the performance than the segmentation of the hippocampus. Here, the performance is the same among the image weighting having and not having the kernel learning. Also, the images of weighted mean and the weighting kernels are joined to give no apparent difference in the performances when the comparison is made to the mean discrepancy image weighting kernel and every kernel learning. The segmentation of brain tissue and WML are shown in Fig 4a to Fig 4d, and both techniques were chosen the similar base kernels even though the weights are different slightly for the segmentation. The kernel weighting provides greater to lower parameter kernels, that is, concentration on the differences among all the samples for the experiments on the hippocampus. Moreover, the behavior happened in around one-third of the images which are segmented and are not appeared to force the performance considerably. It is due to the premature stopping in the optimization in the proposed model

that is the same as the kernel process, which was mentioned earlier. The distribution of image weights is studied during kernel learning and image weighting. These are very close to the computed for all three applications with no learning process.

All the techniques are incorporated kernel learning that has more time for computation. The time ranges with the number of training images utilized quadratically, and the kernel learning has around 45 images for the experiments of WML as 20-30 training images. Kernel learning requires approximately 2 hours for tissue segmentation as 30 training images. Fewer base kernels are used to speed up the technique significantly. Extraction of features in 10 and 30 seconds makes the kernel computation that relies on the number of extracted features.

The performance is improved considerably using the image weighting than the weighting on all the trained images equal for all the three applications such as hippocampus segmentation, voxelwise brain tissue, and white matter lesion (WML). The belief is convinced by image weighting, which relates to the earlier finding. The unbalanced classes and the Kullback-Leibler (KL) weighting perform better for the WML segmentation than other weighting techniques. Since more essential distribution have a small prior distribution  $P(x)$ , KL weighting performs better than the other weighting techniques. The best techniques are varied, and the differences among the three weighting techniques are considerably low for the segmentation of hippocampus and brain tissue.

The outcomes of the different methods are convinced to reveal that the combination of image weighting and kernel learning is beneficial for the segmentation of the scanner. The collections of training images as 20 to 134 are used in the proposed method having various characteristics. Hence the image weighting technique can choose the valuable images. The proposed system has experimented with the voxel-wise classification, yet the suggested approach is used for the patches or super voxels classification. It is used for the issues of image classifications like the computer-aided

diagnosis using the determination of image weights depending on the voxel distributions. A small number of classes only need to predict in the experiments as two or three. Yet the techniques are used for many class problems like the various brain subfields or structures separation. Since the weighting is unsupervised and the weights are based on the similarity between the complete images and weights based on the similarity among the entire images, the results are as the segmentation in the same image weights for the similar image into other classes. It is interesting to examine if this is possible to provide more importance for the class boundaries or particular class similarity for the cases where the classes are small as the instance in the subfields, structures of brain segmentation, or WMLs.

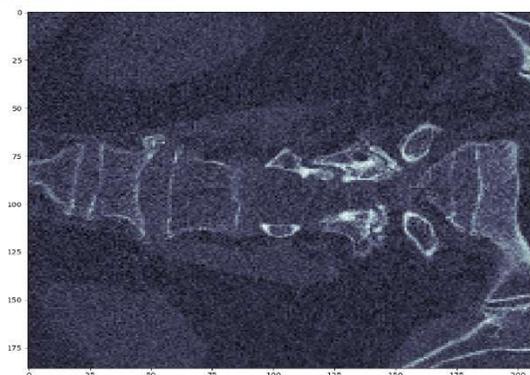
The proposed method is readily available for the practice. Consider an example since no labeled data is required from the distribution in multi-center research. Image weighting is beneficial when a few annotated images are available manually from the test scanner. The combination of image weighting and kernel learning is valid even though the setting is not tested in the proposed system. The labeled test data is utilized additionally for

ensuring the learned kernel using the kernel maximization. The added value is compared with the kernel learning and the image weighting in the proposed system (k-SVM) which has the space features of the Gaussian scale. Moreover, various features, classifiers, and pre and post-processing apply image weighting and kernel learning. Here, image weighting is used in the proposed framework, especially when the training samples are chosen that depend on the weights of the image. Moreover, the performance is not always good in the proposed system due to the normalization constraint of variance in the learned kernel spaces that provides the various components required to describe the data. The exciting direction is how to resolve the problem like having multiple constraints in kernel space or the varied way of extracting the features from the kernel space for further study. The transfer and combination of the feature representation by the other investigation techniques with image

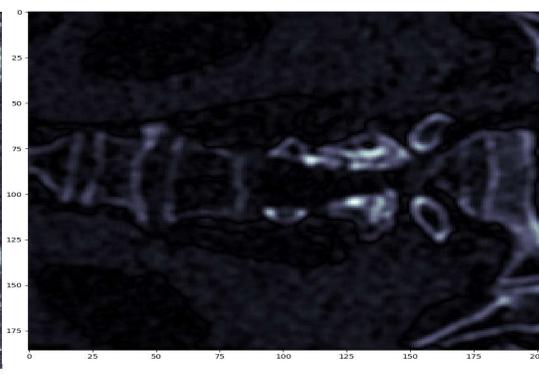
weighting is also an exciting factor. An example is the contrast synthesis technique used to perform the feature representation transformation. One needs to set an instance to think about transferring the representation for deep learning between the various datasets using the presented joint representation layer generation.

The image weighting combined with the feature representation is transferred via kernel learning, which would be a guaranteed technique for the supervised segmentation that is varied from the test dataset. Better outcomes have resulted in the suggested methods for the various tasks used for the wide range of segmentation tasks on medical images.

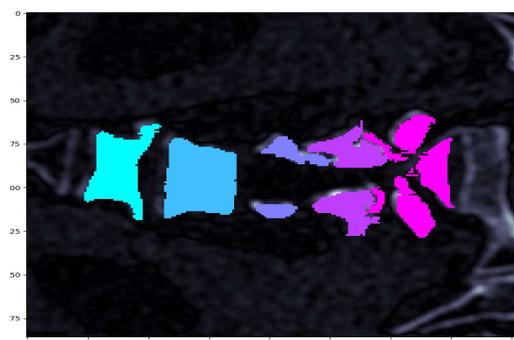
The segmentation for the 25 segmented regions is described in Fig 3. The suggested model effectively classifies the regions as the segmented region size maximizes with the help of the classification model, which is scalable. Therefore, every colour denotes one segmented region for the problem of classification of data.



*Fig 1 Input image*



*Fig 2 Filtered image*



*Fig 3 Over-segmented regions*

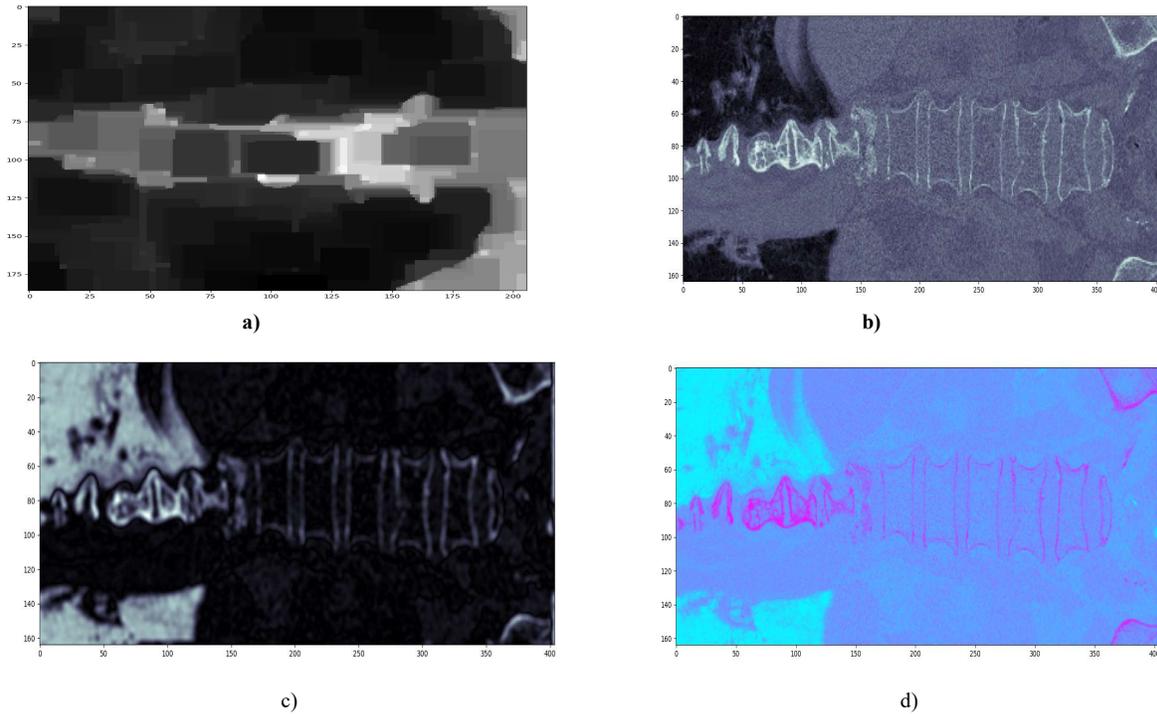


Fig 4 Segmented regions

## 4.2. Performance evaluation

These results are tested and simulated in the experiments to identify the quality segmentation and the accuracy in classifying spinal cord images. Various types of calculation metrics are utilized to improve the quality of the image concerning the spinal cord injury. Therefore, the two adjacent segmented regions are represented as  $A$  and  $B$  in the regions of spinal cord injury. The average mean and variance of  $A$  and  $B$  are presented as  $\theta_1$  and  $\theta_2$  (segmented region). Compared with the conventional methods, the three evaluation measures are required to assess the suggested classification and segmentation performance. At the same Time, the recall, accuracy, precision, and true positive rate are used to compute the classification accuracy.

### 4.2. i) Overall Accuracy

The overall accuracy is calculated by dividing the number of accurately sampled data points by the entire sample. The formula for the computation is expressed in Equation . (18):

$$P_{OA} = \frac{\sum_{i=1}^c x_{ii}}{\sum_{i=1}^c \sum_{j=1}^c x_{ij}} \quad (18)$$

Then, precision, accuracy, recall, and F-measure are among the performance measures evaluated for prediction. The performance indicators are based on the confusion matrix, with TP denoting a lot of positive instances that are positive, TN representing the number of negative instances that are negative, and FP indicating the series of adverse instances that are positive but are intended just to be positive, and FN indicating the number of positive instances that are negative but are designed to be positive. The model's capacity to handle a real example positively or negatively is solely dependent on the accuracy of the prediction process. The fraction of correctly predicted positive instances out of all optimistic predictions supplied by the predictor model is precision. Equation (19) expresses it as follows:

$$Precision = \frac{TP}{(TP + FP)} \quad (19)$$

The recall is represented as the percentage of projected positive events that are always positive. It's written like this in Equation. (20):

$$Recall = \frac{TP}{(TP + FN)} \quad (20)$$

The F1 score is a commonly used metric for classification issues. The recall and accuracy rates are calculated using the average harmonic technique, with the maximum value set to 1 and the minimum value set to 0. Eq. (21) is a mathematical expression for it:

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (21)$$

Specificity is defined as the ability to correctly forecast samples that are not deemed valid, i.e. true negative samples. Equation. (22) is a mathematical expression for it:

$$Specificity = \frac{TN}{FP + TN} * 100 \quad (22)$$

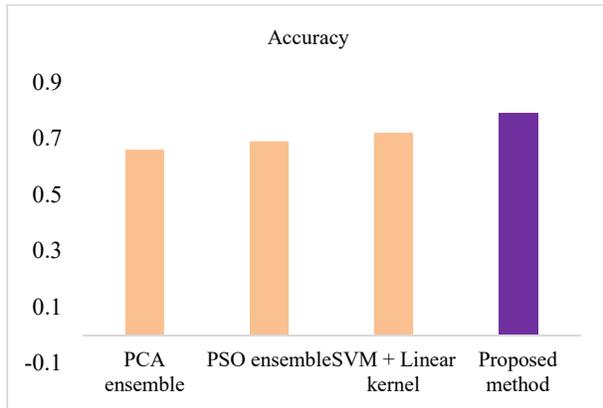


FIG 5 Accuracy Comparison

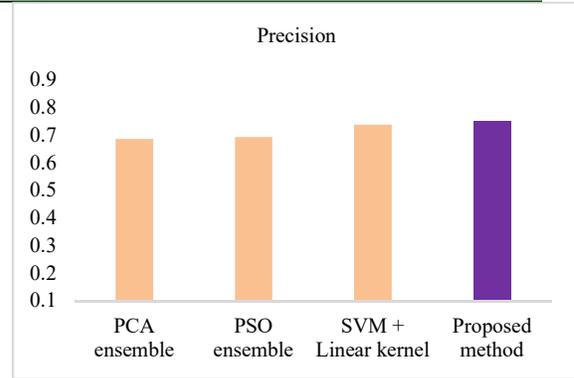


Fig 6 Precision Comparison

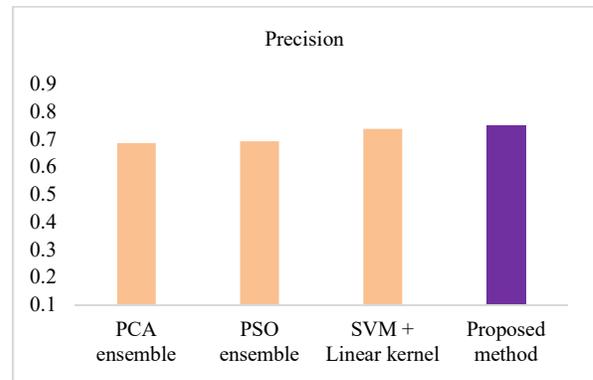


Fig 7 Recall comparison

Table 1 Performance Metrics Comparison

Methods	Accuracy	Precision	Recall
PCA ensemble	66%	68.64%	60.25%
PSO ensemble	69%	69.35%	62.33%
SVM + Linear kernel	72%	73.84%	64.65%
Proposed method	75%	75%	66.38%

Table 1 depicts the comparison of the anticipated model with other approaches like PCA ensemble, PSO ensemble, and SVM + Linear kernel model. The accuracy of the proposed model is 75% which is 9%, 6%, and 3% higher than the PCA ensemble, PSO ensemble, and SVM + Linear kernel model (See Fig 5). The precision of the anticipated model is 75% which is 6.36%, 5.65%, and 1.16% higher than the PCA ensemble, PSO ensemble, and SVM + Linear kernel model (See Fig 6). The recall of the anticipated model is 66.38% which is 6.13%, 4.05%, and 1.73% higher than other approaches. Based on these analyses, it

is proven that the anticipated model works well in the prediction process. The proposed model gives superior results in segmentation and prediction accuracy. However, the complexities encountered during the process of segmentation and classification need to be rectified to favor computational time. To be specific, the weighted learning-based model seems to be a promising solution for segmentation with various data from the provided dataset. Superior outcomes from various tasks specify that the anticipated model is applied over the wider range of segmentation tasks. In DL, the investigators need to think of the transfer representation among various datasets by providing layer representation.

Some existing learning approaches face some issues like interpretability and data deficiency. Also, the privacy factor leads to some complexities in medical data. Additionally, the high-quality data is unusual as the labelling and collection by clinicians are time-consuming. The learning approaches give brief explanation regarding the features that are chosen during training process that hinders the algorithm recognition. Thus, the weighted kernel analysis is encouraged to enhance the interpretability while performing the prediction process. Some existing research works concentrate on segmentation rather than prediction. However, the proposed model intends to give importance to prediction by enhancing the prediction efficiency and reliability

## 5. CONCLUSION

Many conventional methods are not dependent on the classification problem of the segmented features of the spinal cord. One of the fundamental problems in feature selection is predicting disease in higher dimensional spinal cord images in real-time applications. Because of the high computational memory and Time, the mentioned models failed to classify the new spinal cord images, which are added to the increased spinal cord segment size. A theoretical model segment-based classification is needed to overcome these problems and identify the injuries' severity. Also, the pattern of the disease needs to be predicated on the segmented features

and regions. The theoretical method is developed for minimizing the over-segmented features; when the segmented features in the SCI maximization and noisy areas are optimized in the images. A image threshold-based segmentation technique is used in the proposed system for lowering the over-segmented regions in the SC images. The segmented regions classify the images with the help of the classification technique. The proposed system is effective, proved in the experimental results rather than the previous techniques. The results are established in the experiments as the proposed model is better concerning the error rate, accuracy, and true positive when compared with the conventional methods. A theoretical multi-level segmentation-based classification technique is implemented in future work to enhance the accuracy and error rate on the gender-wise SC images. In addition, the classification accuracy needs to improve in the traditional images of SCI, and the regions of T1-weighted and T2-weighted noises need to optimize. The major research constraint is the lack of feature representation based on severity in spinal cord injury. The most influencing feature representation assists in enhancing the prediction rate.

In the future, this issue can be addressed with the inclusion of convolutional neural network (CNN) for feature representation (most influencing feature) as the model intends to enhance the prediction quality and works a better

## REFERENCES

- [1] Maier, C. Syben, T. Lasser and C. Riess, "A gentle introduction to deep learning in medical image processing", *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 86-101, 2019.
- [2] Giger, "Machine Learning in Medical Imaging", *Journal of the American College of Radiology*, vol. 15, no. 3, pp. 512-520, 2018.
- [3] Li, J. Xu, X. Chen, J. He and Y. Huang, "Phase Synchronization Between Motor Cortices During Gait Movement in Patients with Spinal Cord Injury," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 151-157, 2016.

- [4] Geramipour, S. Makki and A. Erfanian, "Neural network-based forward prediction of bladder pressure using pudendal nerve electrical activity," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, pp. 4745-4748, 2015
- [5] Kachuee et al., "An Active Learning Based Prediction of Epidural Stimulation Outcome in Spinal Cord Injury Patients Using Dynamic Sample Weighting," 2017 IEEE International Conference on Healthcare Informatics (ICHI), Park City, UT, pp. 478-483, 2017.
- [6] O'Regan, "Putting machine learning into motion: applications in cardiovascular imaging", Clinical Radiology, 2019.
- [7] Qin and L. Chiang, "Advances and opportunities in machine learning for process data analytics", Computers & Chemical Engineering, vol. 126, pp. 465-473, 2019.
- [8] Amiri, N. Shoaib, and S. V. Hiremath, "A framework to enhance assistive technology-based mobility tracking in individuals with spinal cord injury," 2017 IEEE Global Conference on Signal and Information Processing, Montreal, QC, pp. 467-471, 2017.
- [9] Ahammad, V. Rajesh and M. Z. U. Rahman, "Fast and Accurate Feature Extraction-Based Segmentation Framework for Spinal Cord Injury Severity Classification," in IEEE Access, vol. 7, pp. 46092- 46103, 2019
- [10] About-i-Medina and S. Zafeiriou, "A unified framework for compositional fitting of active appearance models," International Journal of Computer Vision, vol. 121, no. 1, pp. 26–64, 2017.
- [11] Khatri KL, Tamil LS, "Early detection of peak demand days of chronic respiratory diseases emergency department visits using artificial neural networks", IEEE Journal of biomedical and health informatics, vol 22, no.1, pp. 285-90, 2017.
- [12] Wang, Xiang, David Sontag, and Fei Wang. "Unsupervised learning of disease progression models." In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 85-94. ACM, 2014.
- [13] Shi, P. Li, H. Yuan, J. Miao and L. Niu, "Fast kernel extreme learning machine for ordinal regression", Knowledge-Based Systems, vol. 177, pp. 44-54, 2019.
- [14] Xiao, T., Azeze, Y., Jayaraman, A., & Albert, M. V. (2018). Activity recognition for incomplete spinal cord injury subjects using hidden Markov models. IEEE Sensors Journal, vol.18, no.15, 6369- 6374, 2018.
- [15] Sánchez et al., "Machine learning on difference image analysis: A comparison of methods for transient detection", Astronomy and Computing, vol. 28, pp.100284, 2019.
- [16] Sok, P., Xiao, T., Azeze, Y., Jayaraman, A., & Albert, M. V. (2018). Activity recognition for incomplete spinal cord injury subjects using hidden Markov models. IEEE Sensors Journal, vol.18, no.15, 6369- 6374, 2018.
- [17] Zhang, C., Yu, J., Ishimatsu, T., & Shiraiishi, N. "Vision-based interface for people with serious spinal cord injury", In IEEE SENSORS, pp. 1-4, 2015.
- [18] Guo, L. Chen, C. L. P. Chen, and J. Zhou, "Integrating guided filter into fuzzy clustering for noisy image segmentation," Digit. Signal Process., vol. 83, pp. 235–248, Dec. 2018.
- [19] Kallel, S. Almouahed, B. Solaiman, and É. Bossé, "An iterative possibilistic knowledge diffusion approach for blind medical image segmentation," Pattern Recognit., vol. 78, pp. 182–197, Jun. 2018.
- [20] Tang, F. Zhou, Z. Gu, H. Zheng, Z. Yu, and B. Zheng, "Unsupervised pixel-wise classification for Chaetoceros image segmentation," Neurocomputing, vol. 318, pp. 261–270, Nov. 2018.
- [21] Min, J. Lu, W. Jia, Y. Zhao, and Y. Luo, "An effective local, regional model based on salient fitting for image segmentation," Neurocomputing, vol. 311, pp. 245–259, Oct. 2018.
- [22] Budde and N. P. Skinner, "Diffusion MRI in acute nervous system injury," J. Magn. Reson., vol. 292, pp. 137–148, Jul. 2018.
- [23] D'souza, A. Choudhary, M. Poonia, P. Kumar, and S. Khushu, "Diffusion tensor MR imaging in spinal cord injury," Injury, vol. 48, no. 4, pp. 880–884, 2017.
- [24] Goh, S. N. Basah, H. Yazid, M. J. A. Safar, and F. S. A. Saad, "Performance analysis of image thresholding: Otsu technique," Measurement, vol. 114, pp. 298–307, Jan. 2018.
- [25] He and S. Huang, "Modified firefly algorithm based multilevel thresholding for color image segmentation," Neurocomputing, vol. 240, pp. 152–174, May 2017.

- 
- [26] Khairuzzaman and S. Chaudhury, "Multilevel thresholding using grey wolf optimizer for image segmentation," *Expert Syst. Appl.*, vol. 86, pp. 64–76, Nov. 2017
- [27] Li, Z. Liu, and J. Zhang, "Unsupervised image co-segmentation via the guidance of simple images," *Neurocomputing*, vol. 275, pp. 1650–1661, Jan. 2018.
- [28] Min, J. Lu, W. Jia, Y. Zhao, and Y. Luo, "An effective local, regional model based on salient fitting for image segmentation," *Neurocomputing*, vol. 311, pp. 245–259, Oct. 2018.
- [29] Mlakar, B. Potočnik, and J. Brest, "A hybrid differential evolution for optimal multilevel image thresholding," *Expert Syst. Appl.*, vol. 65, pp. 221–232, Dec. 2016.
- [30] Nie, P. Zhang, J. Li, and D. Ding, "A novel generalized entropy and its application in image thresholding," *Signal Process.*, vol. 134, pp. 23–34, May 201.