Alignment, Segmentation and Pathfinding: A Novel Framework of Brain Midline Detection

Paper ID 991

Abstract. Midline related pathological image features are crucial for evaluating the severity of brain compression caused by stroke or traumatic brain injury (TBI). Automated midline detection not only improve the assessment and clinical decision making for patients with stroke symptoms or head trauma, but also reduce the time of diagnosis. Nevertheless, most of the previous methods model the midline by localizing the anatomical points, which are hard to detect or even missing in severe cases. This paper provides a novel framework of the automated midline detection for computer-aided decision support. The proposed framework firstly transforms an input CT image into the standard space by using an alignment network. Then the aligned image is processed by midline network integrated with the CoordConv Layer and Cascade AtrousConv Module. Finally, a pathfinding algorithm is proposed to obtain the optimal midline by modeling the global context. Experimental results show that our proposed framework can achieve superior performance on brain midline detection on two datasets. Moreover, we conduct abundant ablation studies to verify the effectiveness of each proposed component.

Keywords: Brain midline detection · Computer aided diagnosis · Segmentation · Dynamic programming

1 Introduction

Stroke and traumatic brain injury (TBI) are the major causes of death and disability in adults. For the patient with stroke symptoms or head trauma, Non-contrast head CT scan is commonly used as the initial imaging because of its wide availability and low acquisition time. In general, midline structures in the Non-contrast CT images are often associated with high intracranial pressure. Therefore, they provide abundant information for physicians to make an accurate diagnosis on the severity of stroke or TBI e.g. Immediate surgery may be indicated when there is a midline shift of over 5 mm [10].

Automated midline detection can quantify midline shift and speed up brain interpretation and decision making for the emergency physicians, illustrated in Fig. 1. Since the brain CT reading reliability of the emergency physicians is often questioned [4], the quantification results of the midline can greatly improve the assessment of stroke or TBI. Combined with other information (e.g. gender, age), the pathological image features related to midline can benefit the clinical diagnosis and prognosis treatment of stroke or TBI for neurologists, neurosurgeons, and neuroradiologists.

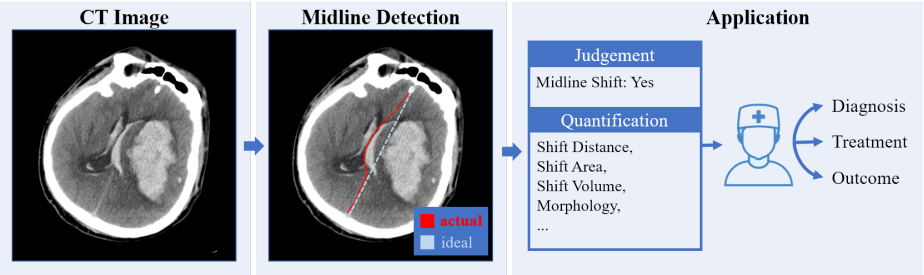


Fig. 1. Automated midline detection and its application.

Previous methods mainly focus on localizing the pre-defined points or parts based on anatomical information of the human brain. Liao *et al.* [5] proposed a deformed midline model according to the biomechanical properties of intracranial tissue. The model firstly localized the upper and lower straight parts of the tough meninges separating two hemispheres. Then a quadratic Bezier curve was proposed to model the deformed midline based on the symmetric structure of the brain. Chen *et al.* [2] estimated the position of the midline using shape matching among multiple regions. The feature points used for matching were localized based on the segmented ventricles from CT images referred to the ventricle templates on MRI. Similarly, Qi *et al.* [8] presented a variational level set to extract the ventricle contours. Then the position of midline was detected based on the identified right and left lateral ventricle contours. Liu *et al.* [7] proposed to delineate the midline by automatically localizing the anatomical points. They generated five candidate positions for each point and characterize the spatial relationships among the points to trace the deformed midline. However, on the severe cases, the predefined anatomical points or parts of the human brain may not be visible which limits the above methods. In addition, the detected midline referring to anatomical points is usually not smooth and decreases the quantification quality for accurate diagnosis.

Deep learning methods have shown the advantage on many biomedical applications, including midline analysis. By using convolutional neural network, Chilamkurthy *et al.* [3] proposed to predict the presence or absence of midline shift in each slice. The slice-level confidences are combined using a random forest to predict scan-level confidences. However, predicting the midline shift or not provides limited diagnostic information. Only segmenting the whole midline structures can provide the comprehensive quantification of pathological image features.

In this paper, we mainly study the problem of automatically detecting the midline structures. Our contribution is threefold. Firstly, we propose a novel framework for midline detection based on image segmentation. By modeling the midline in a pixel-wise way, our method can accurately quantify the pathology features which are essential for clinical applications. Secondly, based on the U-shape network, we propose to inject a CoordConv Layer to leverage the spatial information of brain midline. In addition, a Cascade AtrousConv Module is uti-

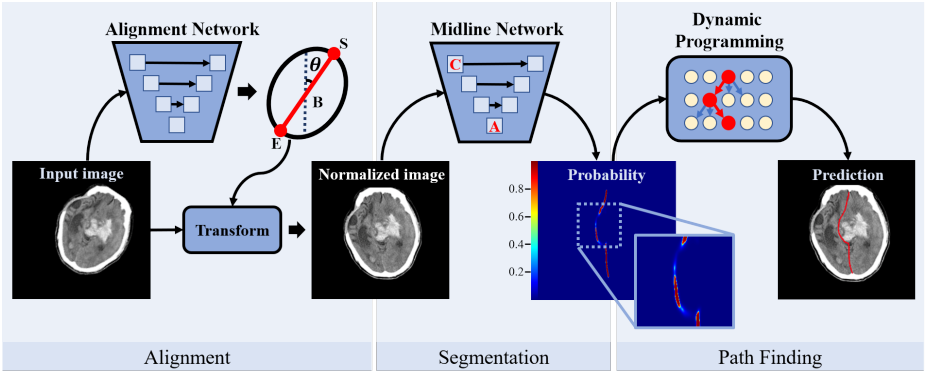


Fig. 2. The proposed framework consists of three parts: (a)Alignment: S, E, B are start, end, center point of the predicted midline respectively and θ denotes the offset angle from the vertical direction; (b)Segmentation: C in the Midline Network denotes the CoordConv Layer and A denotes the Cascade AtrousConv Module; (c)Pathfinding.

lized to enlarge the receptive field for the further refinement. Thirdly, a pathfinding algorithm is proposed to select the optimal midline structure based on the global context. In this way, the predicted midline tends to be both continuous and smooth. Experimental results show that our proposed framework achieves superior performance and promising generalization ability on automatic midline detection.

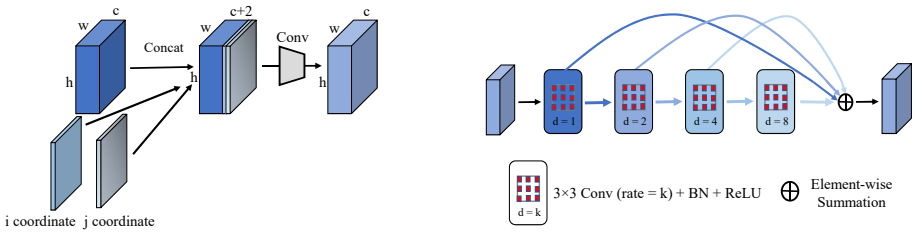
2 Method

Our proposed framework is illustrated in Fig. 2 and comprises of three components: alignment, midline segmentation and pathfinding. Next, we introduce the details of every key component of our framework.

2.1 Alignment Network

Due to the variations of patients' head positions during CT scanning, the relative position of brain in the CT image is usually not consistent. As the midline detection is sensitive to the location information, we attempt to map the image to the standard space. In this section, we propose an alignment network, based on UNet[9] to estimate the midline structure at a coarse scale, consequently the two endpoints (S and E in Fig. 2) of midline can be located. Therefore, we can calculate the offset angle from the vertical direction and the coordinate of brain center to obtain the transformation matrix as follows

$$T_1 = \begin{pmatrix} 1 & 0 & x_I \\ 0 & 1 & y_I \\ 0 & 0 & 1 \end{pmatrix}, R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, T_2 = \begin{pmatrix} 1 & 0 & -x_B \\ 0 & 1 & -y_B \\ 0 & 0 & 1 \end{pmatrix}, T = T_1 R T_2 \quad (1)$$



(a) CoordConv Layer (b) Cascade AtrousConv Module

Fig. 3. Detail structure of the proposed modules in Midline Network.

where (x_I, y_I) denotes the image center I, (x_B, y_B) denotes the brain center in original image B, θ denotes the offset angle from the vertical direction and T is the final transformation matrix.

2.2 Midline Network

CoordConv Layer. The midline is the junction of the left and right brains, which is highly correlated to position information. In this part, we adopt the coordination-guided convolutional layers (CoordConv Layer) [6] to model the spatial information. The CoordConv is a simple extension to the classic convolutional layer, which integrates position information by concatenating extra coordinate channels. As shown in Fig.3 (a), 2 extra channels are added respectively to represent x, y coordinates of the input. We add CoordConv layers in the first transition in the encoding path and the values of coordination channels are normalized to the range from -1 to 1.

Cascade AtrousConv Module. Atrous convolutions are widely used for semantic segmentation [1], which increase receptive field while keeping the feature map resolution unchanged. The receptive field of the encoder in the orignal UNet is 140×140, only covering part of the input CT image, which may neglect the importance of global context information on midline segmentation. Therefore, we propose a Cascade AtrousConv Module to explore a larger receptive field (620×620, covering the full image), as shown in Fig.3 (b).

Loss Function and Deep Supervision. Cross entropy is a classic loss function in semantic segmentation tasks. Dice coefficient loss can alleviate the problem of sample imbalance to a certain extent. So we combine the weighted cross-entropy and soft dice coefficient as the total loss function

$$\mathcal{L}_{total} = \mathcal{L}_{wce} + \mathcal{L}_{Dice} \tag{2}$$

To enhance the training of the encoder, we use a 1×1 convolution layer to predict a probability map for each scale of the feature in the decoder and deep supervision mechanism is further imposed on the semantic maps of different scale.

2.3 Midline Pathfinding Algorithm

In the standard space, the midline is a continuous line and composed of one point in each row. Given a segmented probability map, the objective is to minimize the following energy function

$$E(\boldsymbol{p}) = \sum_{i} \psi_i(p_i) + \sum_{i,j} \psi_{ij}(p_i, p_j) \tag{3}$$

where p_i represents the selected pixel in i-th row. We use a unary potential $\psi_i(p_i) = -logP(p_i)$, where $P(p_i)$ is the label assignment probability at pixel p_i as computed by Midline Network. The pairwise potential ψ_{ij} represents the smooth terms between selected points in the two rows. To simplify the problem, only smooth terms between adjacent rows are considered. And we limit the direction of connectivity to be one of *down*, *bottom-left* and *bottom-right*(see Fig. 2), so that $\psi_{ij}(p_i, p_j)$ is described as:

$$\psi_{ij}(p_i, p_j) = \begin{cases} 1, & \left|x_{p_i} - x_{p_j}\right| \leq 1, |i-j| = 1 \\ +\infty, & otherwise \end{cases} \tag{4}$$

where x_{p_i} denotes the horizontal coordinate of pixel p_i.

As we limit the connected directions between adjacent rows, the optimization problem is equal to find a path with the highest probability of summation from the start point selected in first row to the end point selected in last row, which can be solved by *dynamic programming*. Our proposed midline pathfinding algorithm has two advantages. Firstly, it takes the global information into account, so that the prediction of points is interdependent. Secondly, the constraints we add can guarantee the continuity and smoothness of the predicted midline.

3 Experiments

3.1 Dataset and Evaluation Metrics

Dataset. The proposed models are evaluated on an in-house dataset and a public dataset CQ500 [1]. For the in-house dataset, we collected 877 non-contrast head CT scans with 5-mm slice thickness from three hospitals. The dataset is randomly split into a train/val/test set of 708/87/82 stacks and the number of scans with midshift are 207/44/42 respectively. For the public dataset CQ500, we exclude the health CT scans and choose 235 CT scans (53 midline shift) with around 5-mm slice thickness. One senior radiologist marks the midline for all the data which are considered as the gold standard.

Evaluation Metrics. We use Hausdorff Distance as the evaluation metric which measures the distance between two sets of points. For one 3D CT series, we evaluate the Hausdorff Distance between the predicted midline and the ground truth on each slice. Then we use the average and the maximum Hausdorff Distance among all slices as our final patient-level evaluation metrics.

[1] http://headctstudy.qure.ai/dataset

3.2 Implementation details

For image pre-processing, three adjacent slices are stacked as the input to model the inter-slice information. The image densities are normalized using brain window. The midline ground truth is expanded to a band with 5-pixel width, and the final prediction is transformed to the original space for evaluation. The model is trained by Adam with parameters $\beta_1 = 0.9$, $\beta_2 = 0.99$ for 100 epochs. The initial learning rate was 0.001. We employ the poly learning rate policy where the initial learning rate is multiplied by $(1 - \frac{iter}{total\_iter})^{power}$ with power = 0.9. Our implementation is based on Pytorch package.

3.3 Results

Table 1. Ablation Studies on the in-house dataset. - Alignment means the input images are not aligned. The row with + adds an extra component to the baseline.

Methods	Total		Shift		Not Shift	
	HD(avg)	HD(max)	HD(avg)	HD(max)	HD(avg)	HD(max)
- Alignment	10.593	29.639	13.662	38.406	9.962	27.834
Baseline	7.705	24.415	11.905	45.239	6.840	20.128
+ CoordConv	6.823	21.001	9.677	39.436	6.236	17.206
+ Cascades AtrousConv	6.567	18.352	8.226	24.425	6.226	17.102
+ DeepSupervisioin	7.081	18.012	9.325	30.751	6.619	15.390
Proposed	6.277	14.061	7.950	**23.272**	5.933	12.165
Baseline (w DP)	5.295	16.411	8.224	29.037	4.692	13.811
Proposed (w DP)	**4.852**	**12.961**	**6.474**	23.347	**4.518**	**10.823**

Table 2. Quantitative results on the CQ500 dataset.

Methods	Total		Shift		Not Shift	
	HD(avg)	HD(max)	HD(avg)	HD(max)	HD(avg)	HD(max)
- Alignment	7.313	25.009	9.141	36.807	6.780	21.573
Baseline	6.448	17.115	8.399	27.488	5.879	**14.094**
Proposed	6.170	17.293	7.307	25.923	5.839	14.780
Proposed (w DP)	**5.156**	**16.154**	**6.155**	**22.341**	**4.865**	14.353

Ablation Studies. In this section, we use UNet as the baseline method and conduct ablation experiments to investigate the proposed three components: alignment, segmentation and pathfinding. All ablation experiments are evaluated on our own test set. To validate the usefulness of the alignment network, we compare the baseline method with or without transforming the input CT images into the standard space and show the results in the first two rows of

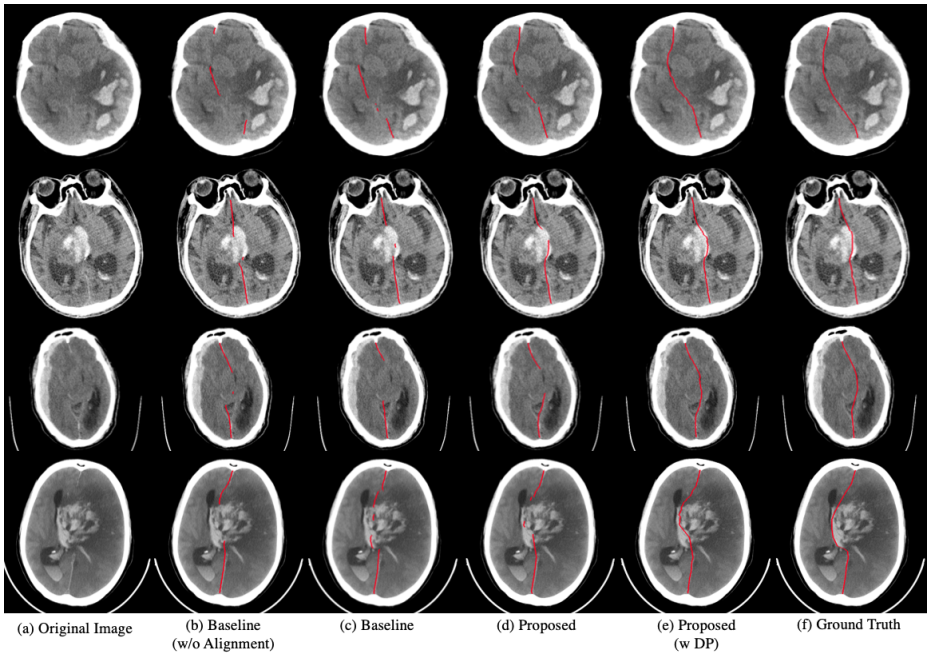


Fig. 4. Qualitative comparison between baseline(w/o alignment), baseline, proposed method and dynamic programming, showing segmentation results for our dataset(the first two rows) and CQ500(the last two rows).

Table 3.3. Without the alignment component, the performance drops greatly since the midline detection is highly correlated with the spatial information. For a comparison with the proposed three modules(resive) of the midline network, we combine each module with the baseline method independently. As shown in Table 3.3, each of the proposed module can improve the performance by a large gap especially in the cases with severe midline shift. Moreover, by the combining all the modules together, our proposed method achieves the best performance according to both avg distance and max distance on all types of cases. Finally, we process the output probability map of the proposed method with the pathfinding algorithm. The performance shown at the bottom of Table 3.3 demonstrates the advantage of using dynamic programming algorithm.

Several examples are shown in Fig. 4. For the baseline (w/o alignment) in row one, the predicted midline are spited into two parts and the orientation of the two parts are highly inconsistent. By using the alignment module, the predict improve greatly by correcting the orientation of the bottom part. Comparing with baseline method, the proposed method performs well especially on the severe part of midline. Furthermore, by processing the probability map with the pathfinding algorithm, the prediction is guaranteed to be continuous which highly improves the qualitative quality of the midline detection.

Generalization Ability Evaluation. We use the models trained on the in-house dataset to predict the midline of the scans of CQ500. As show in Table 3.3, each of the proposed component can improve the performance to a certain extent which is consistent with the results on the in-house dataset. We also visualize some cases in Fig. 4 to evaluate the comparing methods qualitatively.

4 Conclusion

We proposed a novel framework for midline detection on Non-contrast head CT scans. The proposed framework contains three components: alignment, segmentation and pathfinding and each of them resolves the problem from different perspectives. Substantial experiments are conducted to demonstrate the effectiveness of each proposed module. In addition, we applied the proposed method on a public dataset to show its generalization ability.

References

1. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv:1412.7062 (2014)
2. Chen, W., Najarian, K., Ward, K.: Actual midline estimation from brain ct scan using multiple regions shape matching. In: 2010 20th International Conference on Pattern Recognition. pp. 2552–2555. IEEE (2010)
3. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Campeau, N.G., Venugopal, V.K., Mahajan, V., Rao, P., Warier, P.: Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. The Lancet **392**(10162), 2388–2396 (2018)
4. Dolatabadi, A.A., Baratloo, A., Rouhipour, A., Abdalvand, A., Hatamabadi, H., Forouzanfar, M., Shojaee, M., Hashemi, B.: Interpretation of computed tomography of the head: emergency physicians versus radiologists. Trauma monthly **18**(2), 86 (2013)
5. Liao, C.C., Chiang, I.J., Xiao, F., Wong, J.M.: Tracing the deformed midline on brain ct. Biomedical Engineering: Applications, Basis and Communications **18**(06), 305–311 (2006)
6. Liu, R., Lehman, J., Molino, P., Such, F.P., Frank, E., Sergeev, A., Yosinski, J.: An intriguing failing of convolutional neural networks and the coordconv solution. In: Advances in Neural Information Processing Systems. pp. 9628–9639 (2018)
7. Liu, R., Li, S., Su, B., Tan, C.L., Leong, T.Y., Pang, B.C., Lim, C.T., Lee, C.K.: Automatic detection and quantification of brain midline shift using anatomical marker model. Computerized Medical Imaging and Graphics **38**(1), 1–14 (2014)
8. Qi, X.: Automated midline shift detection on brain CT images for computer-aided clinical decision support. Virginia Commonwealth University (2013)
9. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
10. Tu, P.H., Liu, Z.H., Chuang, C.C., Yang, T.C., Wu, C.T., Lee, S.T.: Postoperative midline shift as secondary screening for the long-term outcomes of surgical decompression of malignant middle cerebral artery infarcts. Journal of Clinical Neuroscience **19**(5), 661–664 (2012)