Neural Computing and Applications (2022) 34:14427–14434
https://doi.org/10.1007/s00521-022-07310-5

ORIGINAL ARTICLE

A novel sagittal craniosynostosis classification system based on multi-view learning algorithm

Lei You[1] · Yang Deng[1,2] · Guangming Zhang[1] · Yanfei Wang[1] · Griffin Patrick Bins[3] · Christopher Michael Runyan[3] · Lisa David[3] · Xiaobo Zhou[1]

Received: 24 August 2021 / Accepted: 11 April 2022 / Published online: 6 May 2022

Abstract
Sagittal craniosynostosis (CSO) occurs when the sagittal suture of a growing child's skull is fused prematurely. Surgery, which involves re-moving the affected bones and increasing the volume of the cranium by repositioning the bone segments or using external forces to guide, is the primary treatment for CSO. For preparation of the surgery, physicians usually classify sagittal CSO subtypes by examining the reconstruction of skulls from CT images or the laser scans of the patients' 3D skull. We have proposed an automated sagittal CSO classification algorithm based on features extracted from projected 3D skull images. However, these features would become invalid if they were extracted from irregular skulls or skulls with completely closed sutures. In order to tackle this problem, we proposed a 3D skull multi-view images-based algorithm to classify the subtypes of sagittal CSO, like the physicians diagnosing these 3D reconstruction images of skulls or laser scans. Our innovations include: 1) the first to propose a multi-view learning algorithm to classify the subtypes of sagittal CSO cases, 2) 3D rendering and a patch-based erosion method are adopted for data augmentation to keep the original ratio and shape of the images, and 3) utilizing a transfer leaning training strategy to train the convolutional neural network (CNN) and a multi-view-based prediction strategy for classifying the cases.

Keywords Sagittal craniosynostosis · Convolutional neural network · Multiview learning · Transfer learning

1 Introduction

Craniosynostosis[1] (CSO) is an extremely serious birth defect that involves premature fusion of one or more sutures on an infant's skull, affecting 1 in 2000. Infants with CSO may suffer from intracranial hypertension and other serious neurological complications such headache, irritability, developmental delay and a lower IQ, due to the limited space for the infants' brain to grow. Thus, an early and accurate diagnosis of CSO is necessary for the infants, due to the cosmetic appearance and functional consequences. It is also crucial for surgery design, management, prevention of complications and early surgical correction. In addition, early surgical intervention allows a minimally invasive approach that decreases perioperative morbidity and mortality, whereas individuals undergoing surgery after the first year of life require a more extensive and higher risk intervention.

CSO can be classified into several subtypes based on the sutures involved. Sagittal CSO is dominate among these subtypes and accounts for 40% to 60% [2, 3]. There are several clinical options for diagnosing sagittal CSO [4–6]. Physical measurements [7] include evaluations on items such as the skull's circumference, cranial length, cranial breadth and so on. Kabbani et al. [8] defined sagittal synostosis by a ridged and fused sagittal suture, bitemporal narrowing, frontal bosing and occipital bossing. Massimi et al. [5] further classified sagittal CSO into three main variants, including anterior (evident frontal boosting),

✉ Xiaobo Zhou
Xiaobo.Zhou@uth.tmc.edu

1 School of Biomedical Informatics, The University of Texas Scientific Center at Houston, Houston, TX 77030, USA

2 School of Biomedical Enginmeering & Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, China

3 Department of Plastic and Reconstructive Surgery, Wake Forest Medical School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA

posterior (occipital bossing) and complex (combination). They [6] suggested that clinical findings of sagittal CSO are unequivocal and easily characterized, and computed tomography (CT) scans are not necessary for all patients. Instead, they suggested a 3D CT scan should be used to reconstruct the skull surface only cases in which a physician determines the scan to be of use for surgical planning and postoperative outcome evaluation. Lisa David et al. [4] provided a sagittal CSO classification system based on 3D CT slices, in which sagittal CSO was classified into four subtypes, including anterior, central, posterior and complex. The authors gave a brief description of each subtype based on the four views of the 3D reconstructed skull. Some 3D reconstruction samples of these four subtypes are shown in Fig. 1.

Researchers [9, 10] try to design automated systems to diagnose CSO cases based on image processing and machine learning algorithms to reduce the labor cost of physicians. The pipeline usually consists of several steps, including image preprocessing, image feature extraction and classifier training. The most variant aspect comes from the image feature extraction set which can be divided into handcrafted features and deep learning-based features. For example, Mendoza et al. [10] extracted hand-crafted features from all the cranial bones and sutures by a statistical shape modeling to classify CSO cases. Qian et al. [9] extracted more than 100 features from projected skull images and sent them to a SVM classifier to classify four subtypes of sagittal CSO. In our previous work [11], we extended Qian's work into a deep learning based one by applying data augmentation and transfer learning strategies which extracts the high level representations from the projected images and increased the classification accuracy. These methods have achieved good performance in CSO analysis; however, they will lose effectiveness if one or more sutures completely closed (Fig. 2b) or the growth of the infant skull is abnormal, resulting in multiple broken areas (Fig. 2c) in the skull surface.

Since many physicians diagnose the CSO patients by rotating the 3D skull and looking at it from many views, researchers are trying to find more robustness features from 3D models to tackle the problems caused by 2D images. Authors [12] proved that 3D photography can provide good information for head shape qualification with a normative statistical head shape model. They utilized their head shape qualification model to classify normal skulls and patients craniosynostosis. Pouria Mashouri et al. [13] proposed a 3D photography-based neural network craniosynostosis triaging system which utilized 3D voxel points to train a 3D CNN and 2D projected top view images to training a 2D CNN. They showed an average classification accuracy of 80.3% on five objectives (sagittal, metopic, unicoronal, plagiocephaly and normal). Guido de Jong et al. [14] proposed a multi-fully connected layers network to learn the representations of the head shape raycasting from 3D photography. Riddhish Bhalodia et al. [15] proposed a metopic CSO severity qualifying system by extracting 2000 points from the 3D reconstruction of the skull and training a classifier with 8 of 2000 significant points selected by the principal component analysis.

From the above, we can see that analyzing 3D images with deep learning models is a better way of diagnosing the CSO related diseases. The major differences among those methods are the input of the networks which determines the choice of neural network. There are several branches of 3D learning architecture for 3D object detection and classification: point clouds-based models [16–18], volumetric occupancy grid-based models [19] and multi-view-based models [20–22]. 3D point clouds-based architectures take point clouds extracted from 3D objects (ModelNet40) as input and try to learn the points based discriminative presentations of 3D object. Due to the heavy cost of memory and computation restrictions, the resolutions of the input image should remain at a low level. The applications of this branch may be limited according to the specifications of input data and computation requirements of the network. On the contrast, multi-view-based ones could adopt 2D-based deep learning architectures to learn representations from the rendered views of objects at a much higher resolution. Since there are many popular available pre-trained models and no additional requirements for the input images, they have many prospects for the use in computer vision tasks, especially in medical imaging analysis.

In this paper, we proposed a multi-view-based algorithm to classify the subtypes of sagittal CSO to tackle the problems we met in our previous work. Our contributions are: 1) the first to propose a multi-view learning algorithm to classify the subtypes of sagittal CSO cases, 2) 3D rendering and a patch-based erosion method is adopted for data augmentation to keep the original ratio and shape of

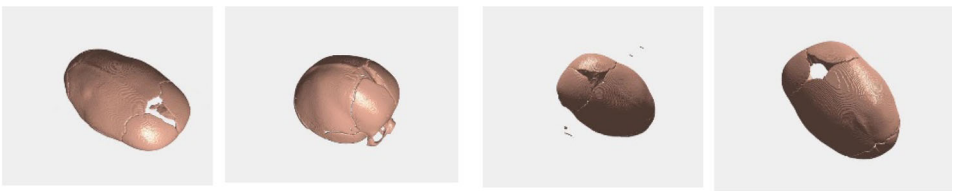

Fig. 1 The four subtypes of sagittal CSO

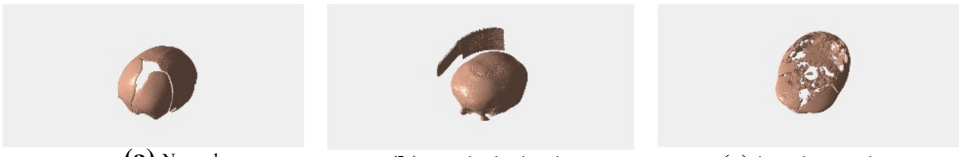

Fig. 2 Some confusing cases caused by the closed sutures or irregular growth of infant skull

(a) Normal **(b)** completely closed **(c)** irregular growth

the images and 3) utilizing a transfer leaning training strategy to train the convolutional neural network (CNN) and a multi-view-based prediction strategy for classifying the cases. As a result, our proposed algorithm is able to classify the sagittal CSO cases no matter whether the sutures completely closed or how abnormal the skulls are. Our following contents are organized as following: the detailed explanations of our method are shown in section II Method, section III presents the experiment and results, and the conclusion is summarized in section IV.

2 Method

As shown in Fig. 3, our proposed multi-view-based sagittal CSO classification algorithm consists of 3D reconstruction of the skull, multi-view image generation, a patched-based erosion data augmentation, CNN model training and multi-view-based prediction. We will introduce the details in the following subsections.

2.1 Patient cohort and dataset

There are totally 180 CT scans collected from 2006 to 2019 from the Department of Plastic and Reconstructive Surgery of Wake Forest Medical School of Medicine. The patients age ranges from 1 to 12 months and the ratio of male and female is 76:24. More than 70% of patients are younger than 4 months old. Each case is assigned with the annotation from one subtype of sagittal CSO. We randomly pick up 160 cases for training the classification network and utilize the rest 20 cases for blind test. For the training process, the ratio for training, testing and validation is 8:1:1. All the CT scans will be transformed into the 3D views images as shown in Fig. 1 for the following processing.

From the first three parts of Fig. 3, we can see the data preprocessing process. The CT scans are segmented by a Hounsfield Unites (HU) thresholding method. Considering that the infant cranial tissue is still in ossification, we adopt a threshold of 300 for segmenting the soft tissues and the bone areas. The part above the Frankfort horizontal plane [23] of the skull is segmented and reconstructed as a 3D object which results in a computation effective 3D representation [15] of the skull.

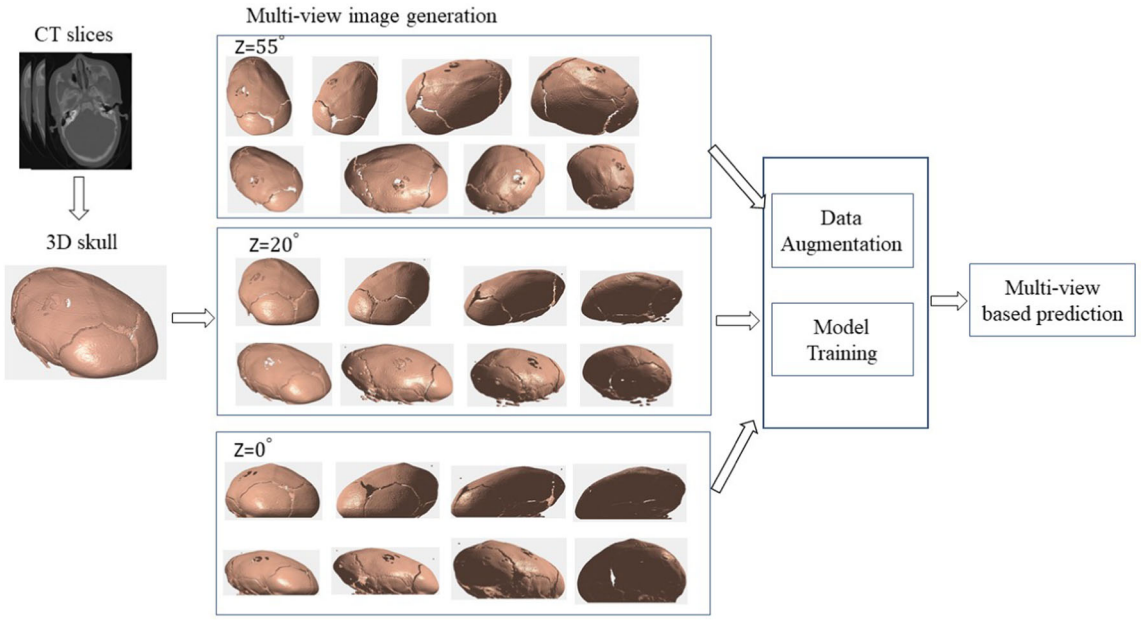

Fig. 3 Our proposed multi-view-based sagittal CSO classification algorithm

A camera is rendered to moving around the 3D reconstruction skull to get different view of images. We choose three angles along the Z direction which are 55 degree, 20 degree and 0 degree. Under each Z direction, we have 8 images for the skull by taking them at every 45 degree. The reason is that we want to use those data to train three CNNs and each CNN is supposed to learn a specific aspect of sagittal CSO. For example, the 0-degree CNN focuses more on the frontal bone and occipital bone which are mainly used for classifying the anterior and posterior subtypes. The other two CNNs could cover all of the subtypes since all the bones and sutures can be found in the images.

2.2 Date augmentation

After data preprocessing, we have 1280 images (232 anterior cases, 272 central cases, 448 posterior cases and 328 complex cases) of each Z direction for training and validating the CNN model. The image resolution may vary case by case which the input size of the network is fixed. Without proper transformation, a posterior case may become central by the resizing functions. To solve this, we pad each image with zero values to make them as a square one.

The distribution of these four subtypes is imbalanced which causes the network learn a biased distribution of the data. Additionally, traditional data augmentation methods such as rescale and cropping will change the image aspect ratio which may change the image into another subtype. To tackle this problem and increase the number of training images, we adopted a patch-based erosion method for data augmentation. As shown in Fig. 4, we randomly erase a patch from the original image by the size of one tenth of the image resolution. The back margins are the padded pixels to make the image as a square one. The erosion patch is applied according to the padded image size. We do two times patch-based erosion to anterior and central cases and one-time patch-based erosion to posterior and complex case. After data augmentation, we have 696 anterior cases, 816 central cases, 896 posterior cases and 656 complex cases.

2.3 Model design and training

Considering the available training images and the small inter-class differences among those four subtypes, we choose to utilize transfer learning to classify the four subtypes of sagittal CSO. Instead of training a network from scratch, a pre-trained Inception V3 model[24], which has been widely used in many medical image analysis tasks[25, 26], is chosen as the backbone of the network. We remove the original 1000 classification layer and added a regularization block to it. The regularization block consists

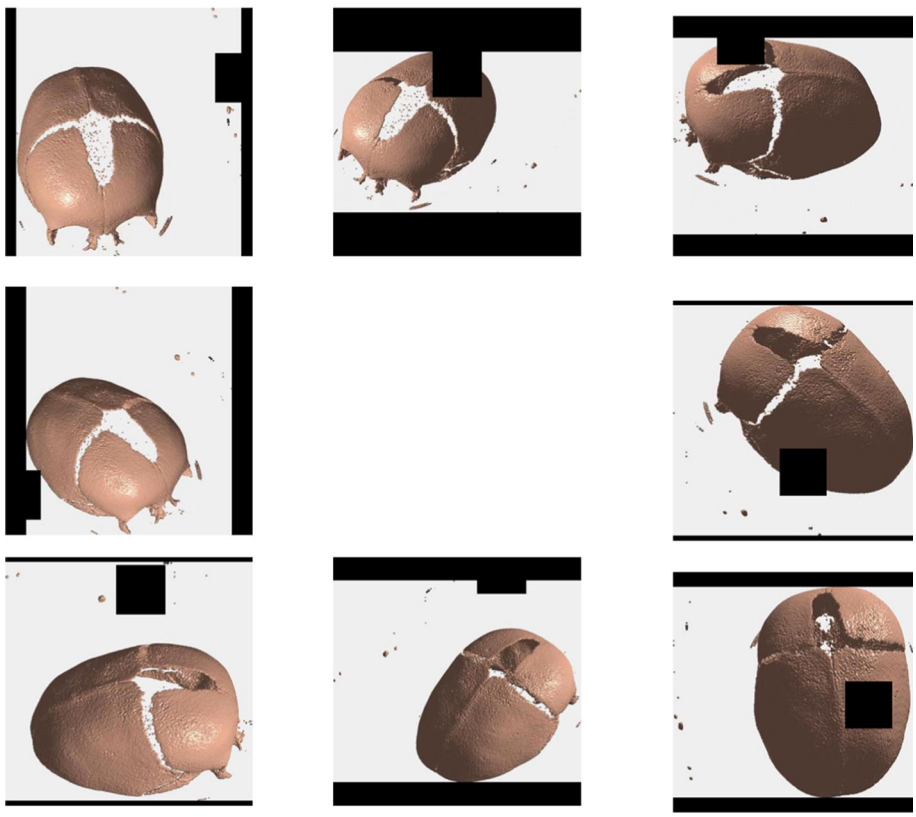

Fig. 4 Our patch-based erosion data augmentation

Table 1 The training performance of three CNN models

CNN model	Training accuracy (%)	Validation accuracy (%)	Test accuracy (%)
0-direction CNN	99.95	83.16	84.72
20-direction CNN	99.94	82.19	80.54
55-direction CNN (without patch-based augmentation)	99.87	60.31	58.73
55-direction CNN	99.56	90.64	90.71

Table 2 The prediction performance on 20 unseen data

CNN model	Accuracy (%)
0-direction CNN	80
20-direction CNN	75
55-direction CNN	80
Multi-view prediction	85

of a drop out layer, a dense connection layer and a following drop out layer. The drop-out rate of these layers is set to 0.3, and the number of nodes of the dense connection layer is 256. Instead of mapping the 2048-dimention vector from fully connected layer to the 4-class classification layer, the added regularization block can reduce the dimension of the features gradually and increase the robustness of the network.

In the training stage, we train three CNN models with different Z-view images which are 55-direction CNN, 20-direction CNN and 0-direction CNN. The algorithm developing platform is Tensorflow [27] with Keras [28] under Python 3. We utilize a Tesla V100 32G GPU to train these networks. The optimize function is stochastic gradient descent [29] and the initial learning rate is 0.0005. The learning rate will decrease as the validation accuracy increases. All the network is trained for 300 epochs with 60 iterations in each epoch.

In the testing stage, we propose a multi-view-based classification algorithm. For each input image of a specific Z-direction, the CNN network outputs a four-dimension vector which indicates the possibilities of the image being predicted to each class. A step function is utilized to set the possibilities to 1 if they are more than a pre-defined threshold. All the eight images' output vectors are added up to make the voting prediction of the current Z-direction. In order to be combined with the outputs of other Z-directions, the voting prediction vector is divided by eight. Then, the final prediction can be got by adding the voting prediction vectors from all the Z-directions and picking out the category with the maximum possibility. The multi-view prediction can be summarized in Eq. 1.

$$V = \alpha\left(\sum_{i=1}^{8} G_{55}\left(f_{55}\left(x_{55}^{i}\right)\right)\right)/8 + \beta\left(\sum_{j=1}^{8} G_{20}\left(f_{20}\left(x_{20}^{j}\right)\right)\right)/8 + \gamma\left(\sum_{k=1}^{8} G_{0}\left(f_{0}\left(x_{0}^{k}\right)\right)\right)/8 \quad (1)$$

where the x_{55}^{i} denotes the i-th input image of the 55-direction CNN (f_{55}) and G_{55} denotes the step function. α, β and γ are parameters according to the validation results of each Z-direction CNN in the training process.

3 Experiments and results

We divide the augmented data generated from the randomly selected 160 cases into training, validation and testing at the ratio of 3:1:1. The overall performance is shown in Table 1. After 300 epochs of iterations, these CNN models fit well on the training data. However, they perform differently on the validation data and test data when there are no patch-erosion-based augmentation data. For the last two rows, model trained on the augmented data shows a much better performance on the validation and test data than the one without the augmented data. It shows the efficiency of our proposed data augmentation strategy which succeeds in solving some of the data bias and data imbalance problem. The generalization ability of the models trained on augmented data will be tested on the unseen data.

We test our CNN models on the other 20 unseen cases to test the model's generalization performance. CT slices of these 20 cases are translate into different Z-direction multi-views images by the same image preprocessing steps in Fig. 3. After that, they are sent to the corresponding Z-direction CNN model for prediction. The final multi-view prediction result is from our multi-view prediction algorithm. From Table 2, we can see that all of the models have a good prediction performance on the unseen data. It shows that our data augmentation and regularization block improve the model's prediction performance on the unseen data. Our multi-view prediction algorithm further increases

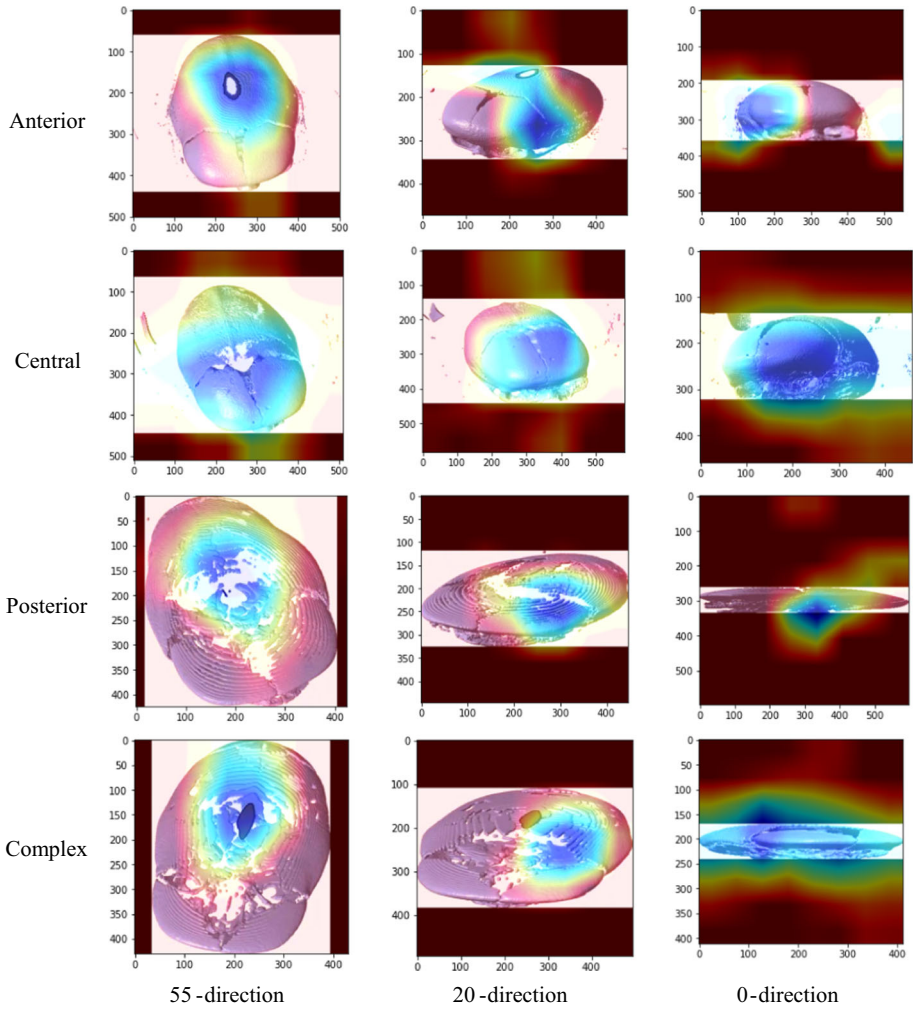

Fig. 5 The activation maps of four subtypes of sagittal CSO

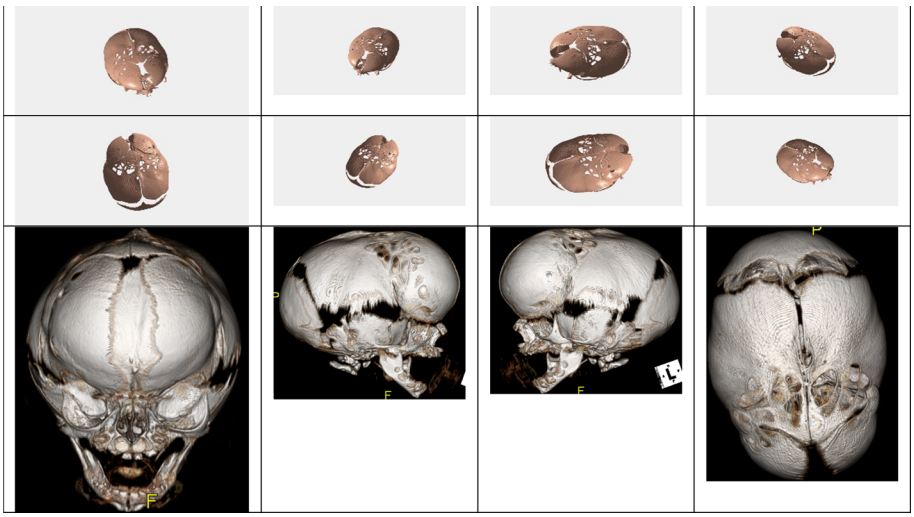

Fig. 6 A sample of difficult to classify

the prediction accuracy on the sagittal CSO classification task.

We use the class activation map (CAM) to demonstrate what the networks have learned in the sagittal CSO classification task and the differences of their focus. In Fig. 5,

we pick up four cases from these four categories and the CAMs of them. In each row, we present the CAM of 55-direction CNN, 20-direction CNN and 0-direction CNN for a same case, respectively. In each column, we present these four CAMs of different cases under the same Z-direction. The blue area indicates a higher confident of being predicted to a specific class by the network. As we expected, the 0-directrion CNN mainly focuses on the frontal bone and the brand along the coronal suture which are utilized to define the Anterior cases in David's literature [4]. The 20-direction CNN and 50-direciton CNN are able to cover most of the skulls and sutures of the skull at a specific view. By taking images around the 3D skull, all the texture and shape information can be learned by the network. Different Z-direction CNN can contribute to the prediction of this diagnosis task with small inter-class differences. From the figure, we also conclude that the 55-direction CNN succeeds in learning the high-level presentations around the sagittal suture areas. It matches the fact that all of the four subtypes are from sagittal CSO in which the symptom is caused by the pre-fused sagittal suture. The 20-direction CNN focus on the areas between the 55-direction and 0-direction which covers part of the vertex area and side area. As a result, the combination of them outperforms each single CNN in the unseen data prediction test. Because of the limited numbers of medical cases and the difficulty of annotating those case, we just adopt multi-view in the prediction stage instead of in the training state [20].

4 Conclusion

Our proposed multi-view-based sagittal CSO classification algorithm consists of three Z-direction CNNs. Combined with the patch-erosion data augmentation and the regularization block, it succeeds in preventing the network from overfitting issues in the classification task. By adopting multi-view strategy in the prediction stage, these complementary CNNs get a better prediction performance on the unseen data. The prediction accuracy on the unseen data outperforms the previous methods [9, 11] in the sagittal CSO classification task. Additionally, our proposed method can work on the cases with fewer limitations than the previous two. For example, these two methods will become invalid if the frontal suture completely closes since their methods are based on the projected images of the 3D skull. On the other hand, our multi-view based CNNs are able to deal with high-resolution images where 3D points-based networks cannot due to the high cost of computation resources. Other advantages, such as free choices of deep learning architectures, transfer learning from a well pre-trained model, easy to be applied to similar tasks, demonstrate the efficiency of our proposed method.

From the experiment results and the CAMs, even with limited number of data, our proposed multi-view learning algorithm is still able to learn different aspects of the target objects. This is critical in medical image processing tasks where the collection and annotation of them are expensive. Applications like this can make better use of existing medical resources and save the labor cost of physicians. We will extend it to a surgery suggestion system when the surgery information such as the spring force or the helmet used in these patients is available.

In Fig. 6, we demonstrate a complex case that can be misclassified by the physicians. In the first two rows, there are the eight images from 55-direction view. Images in the third row are from Philips software. This case has an obviously boosted front bone which can be classified as anterior case. An enlarged occipital bone can be seen at some specific view indicating a posterior case. As a result, it should be classified as complex case. Thus, there is a need to collect or generated more complex cases for the system. Considering the difficulties of generating 3D images or the CT slices of patients, generating multi-view 2D images has the advantages such as fewer cost of computation resources and being able to choose more complex deep learning models. We will further improve the performance of our method by utilizing the generative adversarial networks to generate more 2D images from our existing ones. In the future, we extend our current method into other tasks such as the prediction of severity of metopic CSO, the prediction of multi-craniosynostosis cases and other 3D medical image-related tasks.

Acknowledgements This work was supported in part by the National Institutes of Health (NIH) under Grant 1R01DE027027-04 and Grant 1U01 AR069395-04 (X.Z).

Declarations

Conflict of interest The authors clarify that there is no conflict of interest for report.

References

1. Church MW, Parent-Jenkins L, Rozzelle AA, Eldis FE, Kazzi SNJ (2007) Auditory brainstem response abnormalities and hearing loss in children with craniosynostosis. Pediatrics 119(6):e1351–e1360
2. Kimonis V, Gold J-A, Hoffman TL, Panchal J, Boyadjiev SA (2007) Genetics of craniosynostosis. Semin pediatr neurol 14:150–161
3. Ruiz-Correa S et al (2006) New scaphocephaly severity indices of sagittal craniosynostosis: a comparative study with cranial index quantifications. Cleft Palate Craniofac J 43(2):211–221

4. David L, Glazier S, Pyle J, Thompson J, Argenta L (2009) Classification system for sagittal craniosynostosis. J Craniofacial Surg 20(2):279–282
5. Massimi L, Caldarelli M, Tamburrini G, Paternoster G, Di Rocco C (2012) Isolated sagittal craniosynostosis: definition, classification, and surgical indications. Childs Nerv Syst 28(9):1311–1317
6. Jane JA, Lin KY (2000) Sagittal synostosis. Neurosurg Focus 9(3):1–6
7. Weathers WM et al (2014) A novel quantitative method for evaluating surgical outcomes in craniosynostosis: pilot analysis for metopic synostosis. Craniomaxillofacial Trauma Reconstr 7(01):001–008
8. Kabbani H, Raghuveer TS (2004) Craniosynostosis. American Fam Phys 69(12):5545–5558. https://doi.org/10.1118/1.4928708
9. Qian X et al (2015) Objective classification system for sagittal craniosynostosis based on suture segmentation. Med Phys 42(9):5545–5558
10. Mendoza CS, Safdar N, Okada K, Myers E, Rogers GF, Linguraru MG (2014) Personalized assessment of craniosynostosis via statistical shape modeling. Med Image Anal 18(4):635–646
11. You L, Zhang G, Zhao W, Greives M, David L, Zhou X (2020) Automated sagittal craniosynostosis classification from CT images using transfer learning. Clin Surg 5:1–10
12. Porras AR et al (2016) Quantification of head shape from three-dimensional photography for pre-and post-surgical evaluation of craniosynostosis. Plast Reconstr Surg 144(6):1051e
13. Mashouri P et al (2020) 3D photography based neural network craniosynostosis triaging system. In: Machine Learning for Health, pp 226–237: PMLR
14. de Jong G et al (2020) Combining deep learning with 3D stereophotogrammetry for craniosynostosis diagnosis. Sci Rep 10(1):1–6
15. Bhalodia R, Dvoracek LA, Ayyash AM, Kavan L, Whitaker R, Goldstein JA (2020) Quantifying the severity of metopic craniosynostosis: a pilot study application of machine learning in craniofacial surgery. J Craniofacial Surg 31(3):697
16. Zhou Y, Tuzel O (2018) Voxelnet: End-to-end learning for point cloud based 3d object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 4490–4499
17. Qi CR, Su H, Mo K, Guibas LJ (2017) Pointnet: deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 652–660
18. Qi CR, Yi L, Su H, Guibas LJ (2017) Pointnet: deep hierarchical feature learning on point sets in a metric space. Adv Neural Inf Process Syst 5099–5108
19. Maturana D, Scherer S (2015) Voxnet: a 3d convolutional neural network for real-time object recognition. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 922–928: IEEE
20. Su H, Maji S, Kalogerakis E, Learned-Miller E (2015)Multi-view convolutional neural networks for 3d shape recognition. In: Proceedings of the IEEE international conference on computer vision, pp 945–953
21. Ma C, Guo Y, Yang J, An W (2018) Learning multi-view representation with LSTM for 3-D shape recognition and retrieval. IEEE Trans Multimedia 21(5):1169–1182
22. Dai A, Ruizhongtai Qi C, Nießner M (2017) Shape completion using 3d-encoder-predictor cnns and shape synthesis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 5868–5877
23. Bayome M, Park JH, Kook YA (2013) New three-dimensional cephalometric analyses among adults with a skeletal class I pattern and normal occlusion. Korean J Orthod 43(2):62–73
24. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
25. Motlagh NH et al (2018) Breast cancer histopathological image classification: a deep learning approach. BioRxiv. https://doi.org/10.1101/242818
26. Ting DS, Liu Y, Burlina P, Xu X, Bressler NM, Wong TY (2018) "AI for medical imaging goes deep. Nat Med 24(5):539–540
27. Abadi M (ed) et al (2015) Tensorflow: large-scale machine learning on heterogeneous systems
28. Ketkar N (2017) Introduction to keras. In: Deep learning with Python. Springer, pp 97–111
29. Johnson R, Zhang T (2013) Accelerating stochastic gradient descent using predictive variance reduction. Adv Neural Inf Process Syst 26:315–323