A Convolutional Neural Network Model for Detecting Sellar Floor Destruction of Pituitary Adenoma on Magnetic Resonance Imaging Scans

Tianshun Feng[1†], Yi Fang[2,3†], Zhijie Pei[2†], Ziqi Li[1], Hongjie Chen[2], Pengwei Hou[2], Liangfeng Wei[2], Renzhi Wang[3]* and Shousen Wang[1,2]*

[1] Department of Neurosurgery, Dongfang Affiliated Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, China, [2] Department of Neurosurgery, Fuzhou 900th Hospital, Fuzong Clinical Medical College of Fujian Medical University, Fuzhou, China, [3] Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Objective: Convolutional neural network (CNN) is designed for image classification and recognition with a multi-layer neural network. This study aimed to accurately assess sellar floor invasion (SFI) of pituitary adenoma (PA) using CNN.

Methods: A total of 1413 coronal and sagittal magnetic resonance images were collected from 695 patients with PAs. The enrolled images were divided into the invasive group ($n = 530$) and the non-invasive group ($n = 883$) according to the surgical observation of SFI. Before model training, 100 images were randomly selected for the external testing set. The remaining 1313 cases were randomly divided into the training and validation sets at a ratio of 80:20 for model training. Finally, the testing set was imported to evaluate the model performance.

Results: A CNN model with a 10-layer structure (6-layer convolution and 4-layer fully connected neural network) was constructed. After 1000 epoch of training, the model achieved high accuracy in identifying SFI (97.0 and 94.6% in the training and testing sets, respectively). The testing set presented excellent performance, with a model prediction accuracy of 96%, a sensitivity of 0.964, a specificity of 0.958, and an area under the receptor operator curve (AUC-ROC) value of 0.98. Four images in the testing set were misdiagnosed. Three images were misread with SFI (one with conchal type sphenoid sinus), and one image with a relatively intact sellar floor was not identified with SFI.

Conclusion: This study highlights the potential of the CNN model for the efficient assessment of PA invasion.

Keywords: pituitary adenoma, deep learning, magnetic resonance imaging, invasion, sellar floor

INTRODUCTION

Pituitary adenoma (PA) is a common intracranial neoplasm, with a frequency of invasiveness of 35–54% (Lee et al., 2015; Yang and Li, 2019; Principe et al., 2020). PAs with sellar invasion have a high rate of residual tumor and recurrence after surgery and pharmacologic tolerance (Dekkers et al., 2020; Trouillas et al., 2020). According to recent guidelines from the European

Society of Endocrinology, temozolomide was considered the first-line chemotherapy for aggressive PAs, radiologically invasive PAs with resistance to conventional treatment (Raverot et al., 2018). Therefore, the accurate radiological diagnosis of PA invasiveness is required to assist the clinician in making therapy strategies and prognosis assessment.

The current preoperative evaluation of sellar invasion is based on radiological characteristics (Bonneville et al., 2020). In previous studies, imaging grading systems of PA invasiveness, such as the Knosp and Hardy grading system, have been widely used to improve the rater reliability and evaluation efficiency (Yip and Aerts, 2016; Mooney et al., 2017b; Bonneville et al., 2020). However, the inter-observer agreement is weak for the Knosp and Hardy grading system (Mooney et al., 2017a). Dichotomizing the full scale was able to address the poor percent agreement of imaging grades. In a recent meta-analysis, PAs with Knosp grade 2, 3A, and 3B presented invasion rates of 30, 61.7, and 81.1%, respectively (Fang et al., 2021). Hence, the invasiveness and non-invasiveness cannot be completely dichotomized according to existing radiographic grades. Although magnetic resonance imaging (MRI) can reveal the sellar structures and tumor characteristics, distinguishing tumor invasion accurately using the naked eye is challenging.

As a branch of machine learning, deep learning has achieved significant improvement in multiple fields of image classification and computer vision, which has also prompted computer reading to become feasible in neuroscience. Machine identification and classification can facilitate faster, accurate, and stable preoperative assessment. Recent advancements in computational power have allowed artificial neural networks to achieve deep architectures due to the graphic processing units and the invention of gradient backward propagation (Wong et al., 2021). These deep neural networks present superior performance than other machine learning techniques and have been progressively applied in clinical practice for intracranial tumors (Akkus et al., 2017; Deepak and Ameer, 2019; Wong et al., 2021). As a deep neural network, convolutional neural networks (CNNs) utilize many learnable convolutional filters to facilitate imaging processing and recognition (Hosny et al., 2018).

This study aimed to construct a CNN model in combination with intraoperative evidence and to assist clinicians in identifying the sellar floor invasion (SFI) of PAs using a contrast-enhanced MRI.

MATERIALS AND METHODS

Patient Cohort

In this study, the keywords including "pituitary adenoma," "acromegaly," "Cushing's disease," and "hyperprolactinemia" were used to search for electronic medical records from 2015 to 2020. After screening, the clinical data of 695 PAs from 2 medical centers (Fuzhou 900th Hospital and Peking Union Medical College Hospital) were enrolled. Basic clinical data, imaging data, and surgical records were reviewed.

This study was approved by the review boards of Fuzhou 900th Hospital and Peking Union Medical College Hospital, and the requirement for informed consent was waived. The inclusion criteria were as follows: (1) patients with clear preoperative imaging data suitable for analysis, (2) transsphenoidal surgery patients with a detailed record of intraoperative invasion, and (3) patients with pathological diagnosis of PAs. Cases with other intracranial tumors, a previous history of surgery or trauma in the sellar region, and artifacts in the imaging were excluded.
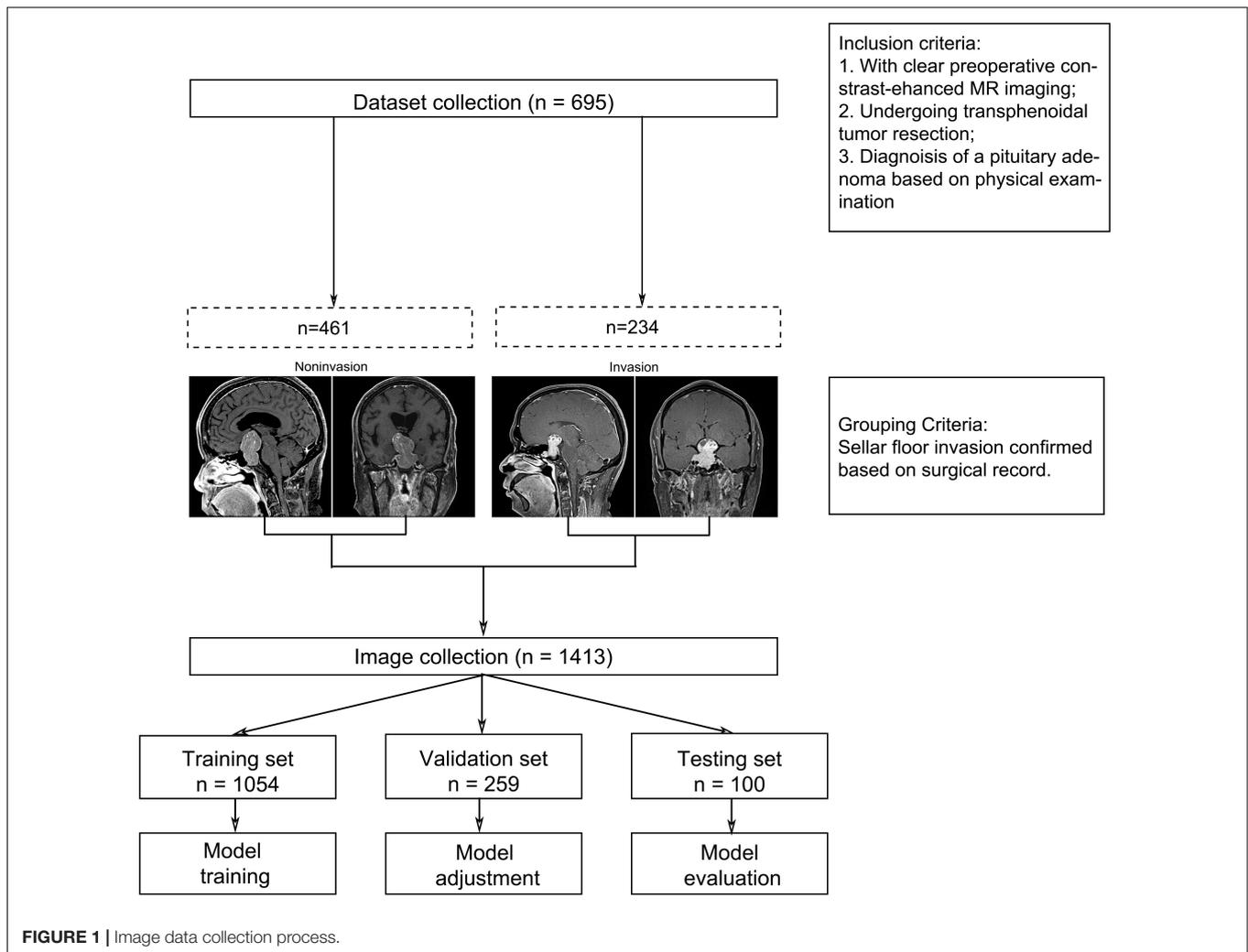
Image Acquisition

All imaging data were collected from contrast-enhanced MRI sequences, including coronal and sagittal scans (**Figure 1**). In this study, cases were divided into two groups according to surgical evidence. Images of PAs without SFI were collected in the slices with the largest tumor area in both coronal and sagittal scans. The SFI was located by experienced neurosurgeons on the basis of surgical evidence and CT scans, and the images were collected. In the sagittal and coronal MRI images of patients with SFI, except for patients with focal sellar floor destruction, other patients with multiple or diffuse sellar floor destruction were sampled from multilayer sections according to the location of the invasion. These images were then screened by neurosurgeons with more than 20 years of experience in the treatment PAs to remove some MRI images that were blurred or difficult to identify the SFI. After the acquisition, patient information was filtered and eliminated, and only the acquired image was retained.

Magnetic Resonance Imaging Protocol

All patients in the study were treated with 3.0T MRI machines (Siemens Medical Solutions, Erlangen, Germany, Fuzhou 900th Hospital; Discovery MR 750, GE Healthcare, Peking Union Medical College Hospital). Patients in the two centers for a contrast-enhanced MRI, using the same contrast agent of gadolinium-DTPA (Gd-DTPA) in a small dose. The contrast-enhanced MRI from Peking Union Medical College Hospital acquisition parameters included slice thickness of 3 mm, slice spacing of 0.39 mm, echo time of 9.2 ms, repetition time of 400 ms, and an image size of $512 \times 512 \times 8$ pixels. And, the parameters of contrast-enhanced MRI included scan field of view was 180×180 mm with matrix of $320-384 \times 240-252$, axial slice of 1.0 mm, gap of 1.0 mm, coronal and sagittal section of 1.0 mm with gap of 1.0 mm in Fuzhou 900th Hospital.

Image Preprocessing

One hundred images were randomly selected from the acquired images, which were not involved in model construction. The remaining 1314 image datasets were randomly grouped at a 80:20 ratio into the training and validation sets to develop a CNN model. The image preprocessing procedure is summarized as follows: (1) All images were converted to 256×256 square images using zero padding and image resizing as appropriate (Kim et al., 2019). (2) Images were converted into grayscale with a single channel. (3) Augmentation procedure using horizontal

Grouping Criteria:
Sellar floor invasion confirmed based on surgical record.

FIGURE 1 | Image data collection process.

flip and vertical flip and batch normalization were performed for data enhancement.

Classification With Deep Neural Network

The input images to the CNN model were 256 × 256 with single channel. Feature capture was performed using six-layer convolution layers, with 3 × 3 kernel size and zero padding without stride. Each layer included a convolution layer, an activation layer, a BatchNorm2d layer, and a maxpool layer. The maxpool layer used a 2 × 2 matrix. Batch normalization is a standard normal distribution with mean 1 and variance 1. It is used to solve the gradient hour problem and accelerate the convergence rate (parameter: eps = 1e−05, momentum = 0.1, affine = True, track_running_stats = True). Then, 4 × 4 × 256 feature maps were output through six-layer convolution layers, flattening, and connecting to a fully connected neural network of 256 neuronal nodes. The neural network consists of four layers. Finally, the binary classification results were output. In addition to the connecting and output layers, 2 layers of the hidden layer were also included, with 128 and 64 neural nodes, respectively. There were also two dropout layers with a probability of 0.7 and 0.5, respectively. The dropout layer reduces overfitting by randomly omitting partial feature detectors on each training case (Geoffrey et al., 2012). The full connection layer was activated using the ReLU function. A binary cross-entropy function was employed as the loss function, and Adam was used as the optimization function (learning rate = 0.0001). Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, can adjust different learning rates, is computationally efficient, and has little memory requirements (Kingma and Ba, 2014). The model structure is detailed in **Figure 2**. The trained model fixes and closes the dropout and batch normalization layers to fit validation set data through the eval function.

Evaluation

The external testing set was used to assess model generalization capabilities. The 100 testing images that did not participate in model development underwent simple transformation (converting to 256 × 256 square images, single-channel grayscale map, and normalization) into a model with adjusted weights for result testing and evaluation. The eval function that can fix

FIGURE 2 | The CNN structure.
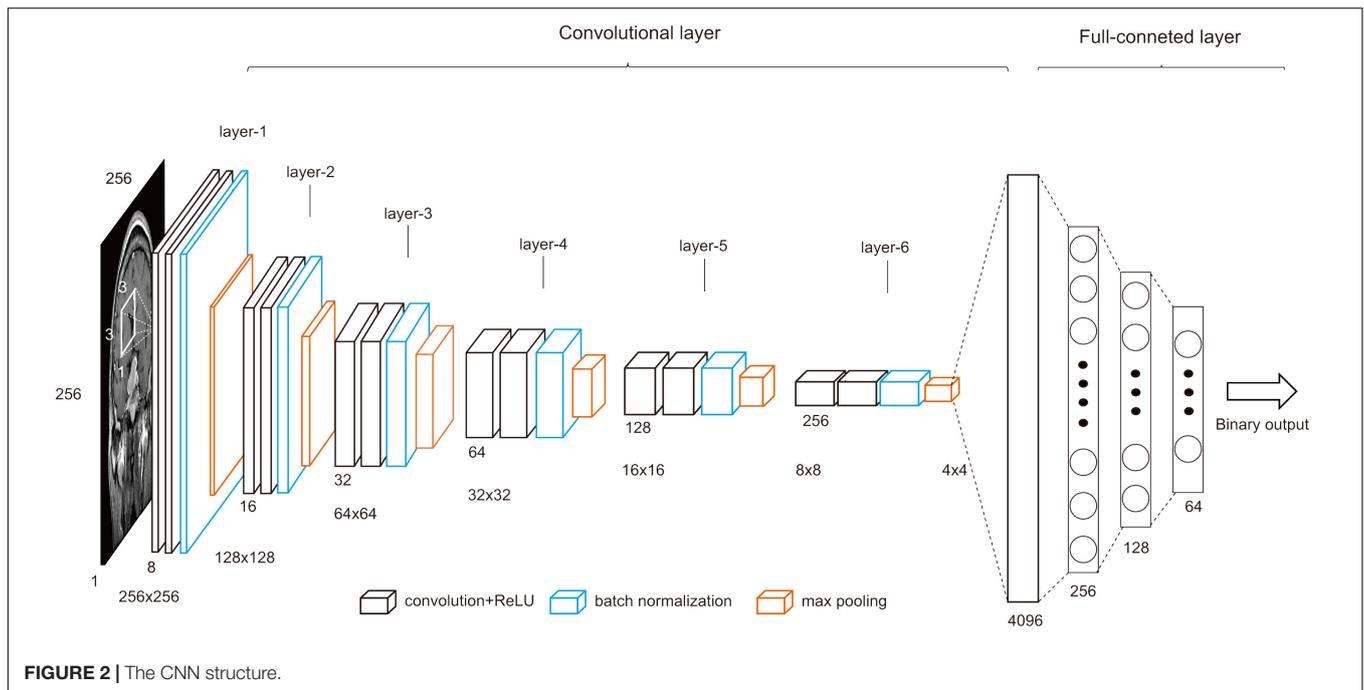
model batch normalization and dropout layer was also used in the testing set. The prediction results were combined with actual labels to establish a confusion matrix. The records included true positive, false positive, true negative, and false negative. The sensitivity, specificity, positive predictive value, negative predictive value, and area under the receptor operator curve

(AUC-ROC) were also calculated to assess the model's predictive and generalization ability.

In addition, the evaluation results of 100 images of all testing set were directly exported to reveal individual cases. The red mark represents the image with evaluation difference. The detailed code is shown in **Supplementary Material**.

Statistical Analysis and Software Availability

All imaging data processing and model methods were implemented using PyTorch (version 1.8.1)[1] and operated in Jupyter Notebook (version 6.4.0).[2] Several open modules, including Torch.nn, Torch.optim, and DataLoader, can be used to develop a CNN model. Open-source libraries such as Sklearn (version 2.1.0), NumPy (version 1.19.5), and Matplotlib (version 3.4.2) were also used for model performance evaluation and visualization.

The preprocessing of the image datasets depended on the Transforms module in Torchvision library (version 0.9.1). The interface for imaging data loading is the DataLoader module in Torch library (parameters of the training set and validation set: batch size = 4, shuffle = True). The model was developed using Torch.nn module. Optimization was performed using Torch.Adam. CrossEntropyLoss function was used as the loss function. Hiddenlayer library (version 0.3) was used to display the model training results dynamically. The confusion matrix was formed by Confusion_matrix function. AUC-ROC analysis was applied to assess the diagnostic value in identifying invasiveness, which was calculated using the Roc_curve function.

TABLE 1 | Summary of patient characteristics.

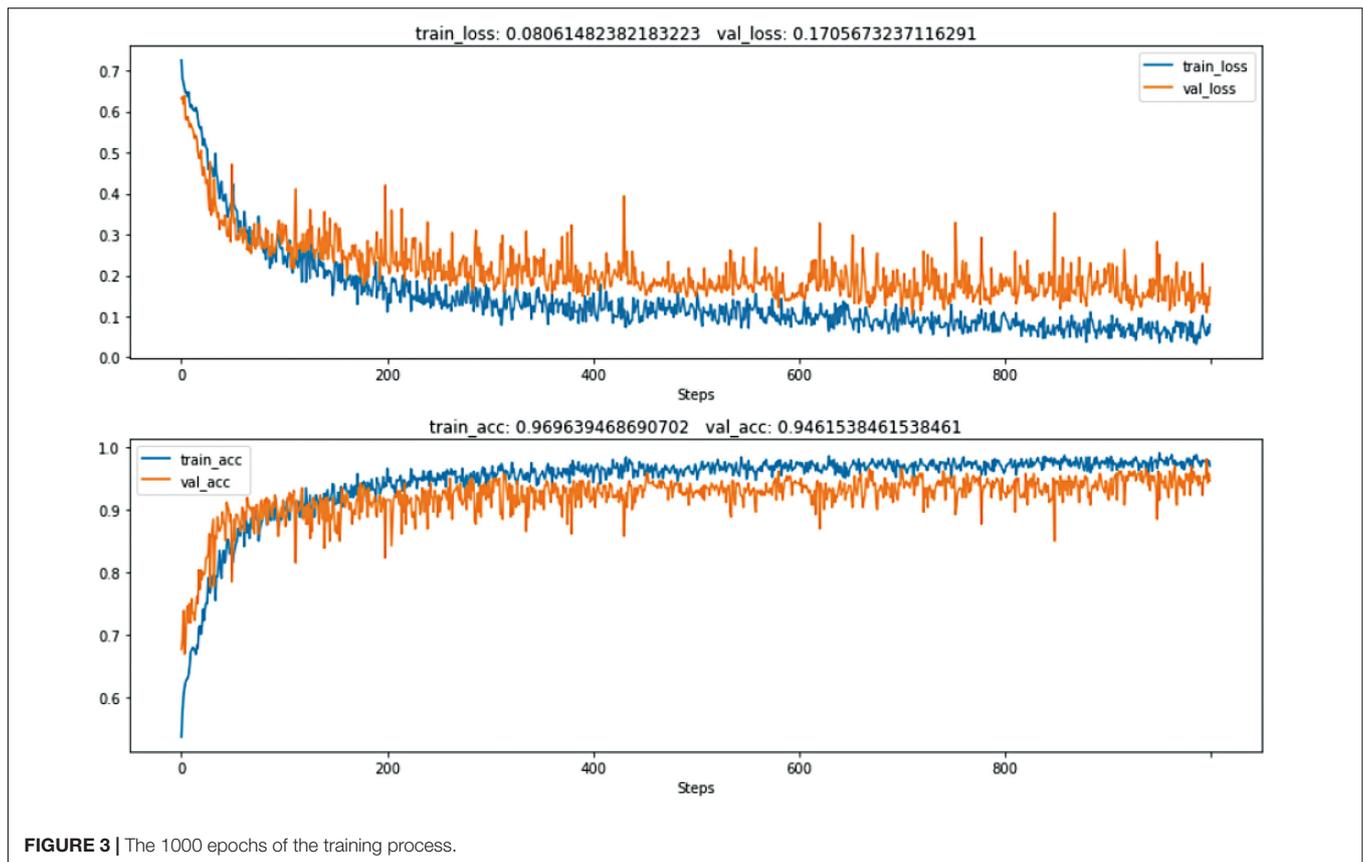| Characteristics N = 695	Value
Age	
Mean, years	48.6 ± 14.0
Range, years	12.0–83.0
Sex (%)	
Female	322 (46.33)
Male	373 (53.67)
Diameter	
Mean, mm	28.38 ± 11.00
With SFI, mm	38.91 ± 10.75
Without SFI, mm	27.20 ± 10.60
Sellar invasion (%)	
Yes	234 (33.7)
No	461 (66.3)
Model group (%)	
Training	1054 (80)
Validation	259 (20)
Testing	100
Accuracy (%)	
Training	97.0
Validation	94.6
Testing	96.0

FIGURE 3 | The 1000 epochs of the training process.

SPSS (version 25) was used for statistical analysis. Categorical variables were summarized as number (percentages) and analyzed by Pearson's Chi-squared test. Continuous variables were presented as mean ± continuous variables and analyzed by T-tests. The significance of differences was accepted at $p < 0.05$.

RESULTS

In the cohort, 234 (33.7%) cases were intraoperatively confirmed with SFI, and 461 (66.3%) cases had no invasion (**Table 1**). There were 373 males (53.67%) and 322 females (46.33%), and the incidence of SFI was significantly higher in males than in females ($p = 0.028$). The mean age was 48.6 ± 14.0 years (12–83 years). SFI was not significantly correlated with age ($p = 0.224$). The mean tumor diameter was 28.38 ± 11.00 mm. The tumor diameter in the group with SFI was significantly larger than that in the group without SFI (38.91 ± 10.75 versus 27.20 ± 10.60 mm, $p < 0.001$).

Finally, after sifting through all the images, a total of 1413 images were collected from 695 patients with a PA. Among them, 530 images were collected in cases with SFI, and 883 were collected in the group without SFI. Except for 100 images randomly selected for the external test set (28 in the invasion group and 72 in the non-invasive group; 32 coronal images and 68 sagittal images), the remaining 1313 images were randomly grouped at a ratio of 80:20. Consequently, 1054 images data were enrolled in the training set (400 in the invasion group and 654

in the non-invasive group; 518 coronal images and 536 sagittal images). A total of 259 images were enrolled in the validation set (102 in the invasion group and 157 in the non-invasive group; 135 coronal images and 124 sagittal images). No significant difference in the distribution of coronal and sagittal images has been found between the training set and the validation set ($p = 0.391$). Similarly, we also ascertained no significant difference in the number of invasive images between the training set and the validation set ($p = 0.671$).

With 1000 training epochs, the model presented convergence and achieved an accuracy over 90% (**Figure 3**). The diagnosis accuracy of SFI was 97.0% in the training set and 94.6% in the validation set. The confusion matrix is shown in **Figure 4**. The diagnostic accuracy of the testing set was 96%. The sensitivity was 0.964, the specificity was 0.958, the positive predictive value was 0.900, and the negative predictive value was 0.986, the positive likelihood ratio is 22.952. The model had an AUC-ROC value of 0.98. These results showed that the CNN model has excellent diagnostic efficacy to distinguish SFI.

In the four misdiagnosed images of the testing set, three were misread as having SFI and one as having no SFI. The diagnosis results of the testing-set images can be available are shown in **Supplementary Material**. In the images misread as SFI, one case was diagnosed with conchal type sphenoid sinus, and the other two cases with large PAs presented severe dilatation of the sellar floor. The frequency of conchal type sphenoid sinus was limited in this study, and the model might misunderstand the type of

FIGURE 4 | ROC curve of the training set, validation set, and testing set.
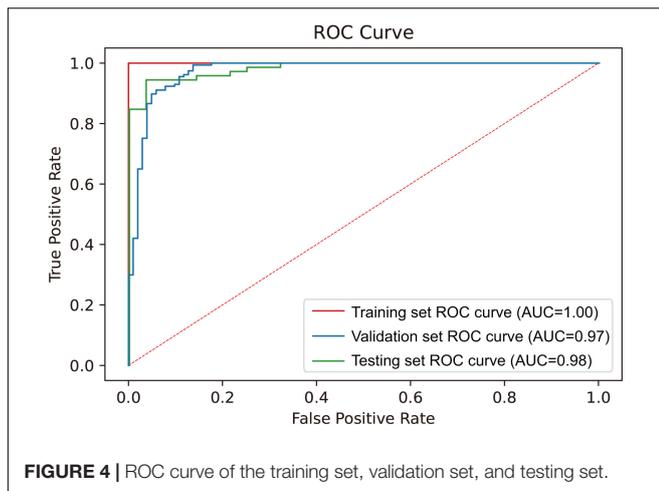
sellar floor. One misread image with SFI had a relatively intact sellar floor, which was not correctly identified by the model.

DISCUSSION

Imaging is essential for preoperative diagnosis of invasive and aggressive PA. Advanced imaging techniques can clearly reveal the sellar structure to assist clinicians in assessing the invasiveness of PAs (Cao et al., 2013; Bonneville et al., 2020). Furthermore, various imaging grades for PA invasiveness have been reported (Micko et al., 2019; Fang et al., 2021). In the sellar invasion scale of Hardy classification, the percent agreement among all raters improved from 16% (8/50 cases) for the full scale to 64% (32/50 cases) for the dichotomous scale (Mooney et al., 2017b). Although investigators try to develop novel sequences or imaging grades to accurately assess sellar structures, the invasiveness of PAs is difficult to accurately diagnose due to limitations of macroscopic identification, need for advanced imaging facilities, and time constraints (Yoneoka et al., 2008; Cao et al., 2013; Lang et al., 2018).

Machine learning can capture the information of each pixel as meticulously as possible through the algorithm to more accurately identify the details than the naked eye. Machine learning can significantly reduce the cost for centers that unconditionally update more advanced imaging facilities. There have been several tentative applications and investigations in the diagnosis and management of intracranial tumors, including the diagnosis of PAs and tumor characteristics (Deepak and Ameer, 2019; Fan et al., 2020; Wang et al., 2021). A total of 194 PAs with Knosp grade 2–3 was included in the latest radiomic study to identify invasion of the cavernous sinus (ICS) (Niu et al., 2019). This study extracted image data through manual delineation and segmentation, and 2553 image features were analyzed using SVM models to identify ICS. The AUC values of the training and test sets were 0.85 and 0.83, respectively, indicating that the model was reliable to distinguish cases with or without ICS in the Knosp grade. This literature confirms the feasibility of machine learning models in imaging identification

and classification of PAs. However, the most critical step in the application of radiomics is the delineation of tumor margins. The quality of delineation will directly affect the results of data analysis. In addition to manual bias, this method has a limited identification range and low efficiency, which is not conducive to clinical application and promotion, so it is more used to study clinical diseases (Yip and Aerts, 2016; Fan et al., 2020).

In the current study, a deep learning-based model was described to accurately classify the invasiveness of PA according to imaging data. With improved algorithms and computer modules, deep neural networks can train and generalize relatively small amounts of data with high sensitivity and specificity. There are few applications of deep learning in the pituitary tumor. Wei et al. (2020) used a deep learning convolution model to discriminate acromegaly ($n = 1139$), Cushing's disease ($n = 880$), and normal human facial images ($n = 12,667$). The accuracy of the external testing set ($n = 60$) was 91.7%, confirming the reliability of CNN for the diagnosis of PA. Li et al. (2021) recently reported a deep learning network that was constructed using 168 patients with PAs. The model can accurately assess functional and non-functional PAs according to imaging data. In our cohort, 1413 PA images were collected to develop a CNN to diagnose SFI. The accuracy of prediction was 97.0 and 94.6% in the training and validation sets, respectively. The model has high generalization ability. Through the performance evaluation of the 100-case testing set, the prediction accuracy of SFI was 96% with an AUC-ROC value of 0.98. The accuracy of the model for SFI diagnosis is much higher than that of Hardy classification. Therefore, the CNN model might become a valuable tool to identify and correctly diagnose the properties of PAs.

Recognizing the invasiveness of PA through deep learning provides not only an objective and stable basis for surgical strategies and prognostic evaluation, but also a more accurate invasive diagnosis for patients without surgical conditions, especially for aggressive PA requiring drug chemotherapy. The assessment of invasiveness is essential for diagnosing the presence of aggressive PAs that are significantly tolerated by traditional medical and surgical treatment (Lopes, 2017). In the latest guidelines for diagnosing and managing aggressive PAs, temozolomide was considered the first-line pharmacotherapy (Raverot et al., 2018; Luo et al., 2021). Consequently, a higher accuracy diagnosis of sellar invasion is required according to imaging data. The deep machine learning model in this study was confirmed to be stable and effective for diagnosing sellar invasion. In the future, machine models can be utilized to form image reports automatically. Furthermore, the location of invasion might be accurately simulated and marked by computer vision. Machine learning can compensate for the macroscopic shortcomings of low reading efficiency and recognition bias.

In addition, the model does not have any prior medical knowledge except for two groups of images and labels. It spontaneously discovers appropriate interpretable features to assess sellar invasion. This suggests that deep learning methods can extract human-understandable domain knowledge from supervised data and have the ability to predict on the basis of the extracted knowledge.

Strengths and Limitations

This study used the convolutional depth neural network model to identify SFI and achieve high diagnostic efficacy. After training, the model was gradually stable, and the generalization ability was excellent. This model is expected to be applied in clinical practice to assist clinicians in screening and distinguishing sellar invasion and improve reading efficiency. In addition, manual segmentation was not used in this study, which reduced the bias caused by manual factors during sampling. At the same time, this study has some limitations. Given that only contrast-enhanced sequences were collected in this study, no healthy population was used as a control group to promote the model to recognize normal sellar structures. In addition, although the model constructed in this study had good evaluation performance, the number of input images was not enough to support the model to learn more details of the sellar region. For example, the rare cases of conchal type sphenoid sinus were extremely limited in this study, the recognition ability of the model for this type was weak, and misinterpretation occurred in the testing set. Expanding the dataset and adding sequence types and healthy human saddle area images can further improve the performance and generalization ability of the model.

CONCLUSION

The convolutional deep learning neural network can objectively and stably identify SFI. The CNN model has the potential to assist clinicians in accurately evaluating PA invasiveness to improve medical strategies.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Institutional Review Board of Fuzhou 900th Hospital of Fujian Medical University and Peking Union Medical College Hospital. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ZL participated in the revision of the manuscript. SW and RW approved the final version to be published and contributed to reviewing the manuscript. LW carried out the statistical analyses. TF, YF, and ZP provided the medical writing and editorial support. All authors made a substantial contribution to the research design, acquisition, analysis, or interpretation of data; revised the manuscript critically; and approved the final version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.900519/full#supplementary-material

REFERENCES

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., and Erickson, B. J. (2017). Deep learning for brain MRI segmentation: state of the art and future directions. *J. Digit. Imaging* 30, 449–459. doi: 10.1007/s10278-017-9983-4

Bonneville, J. F., Potorac, J., and Beckers, A. (2020). Neuroimaging of aggressive pituitary tumors. *Rev. Endocr. Metab. Disord.* 21, 235–242. doi: 10.1007/s11154-020-09557-6

Cao, L., Chen, H., Hong, J., Ma, M., Zhong, Q., and Wang, S. (2013). Magnetic resonance imaging appearance of the medial wall of the cavernous sinus for the assessment of cavernous sinus invasion by pituitary adenomas. *J. Neuroradiol.* 40, 245–251. doi: 10.1016/j.neurad.2013.06.003

Deepak, S., and Ameer, P. M. (2019). Brain tumor classification using deep CNN features *via* transfer learning. *Comput. Biol. Med.* 111, 103345–103352. doi: 10.1016/j.compbiomed.2019.103345

Dekkers, O. M., Karavitaki, N., and Pereira, A. M. (2020). The epidemiology of aggressive pituitary tumors (and its challenges). *Rev. Endocr. Metab. Disord.* 21, 209–212. doi: 10.1007/s11154-020-09556-7

Fan, Y., Chai, Y., Li, K., Fang, H., Mou, A., Feng, S., et al. (2020). Non-invasive and real-time proliferative activity estimation based on a quantitative radiomics approach for patients with acromegaly: a multicenter study. *J. Endocrinol. Invest.* 43, 755–765. doi: 10.1007/s40618-019-01159-7

Fang, Y., Pei, Z., Chen, H., Wang, R., Feng, M., Wei, L., et al. (2021). Diagnostic value of Knosp grade and modified Knosp grade for cavernous sinus invasion in pituitary adenomas: a systematic review and meta-analysis. *Pituitary* 24, 457–464. doi: 10.1007/s11102-020-01122-3

Geoffrey, E. H., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012). Improving neural networks by preventing. *Comput. Sci.* 3, 212–223. doi: 10.9774/GLEAF.978-1-909493-38-4_2

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., and Aerts, H. (2018). Artificial intelligence in radiology. *Nat. Rev. Cancer* 18, 500–510. doi: 10.1038/s41568-018-0016-5

Kim, T., Heo, J., Jang, D. K., Sunwoo, L., Kim, J., Lee, K. J., et al. (2019). Machine learning for detecting moyamoya disease in plain skull radiography using a convolutional neural network. *EBioMedicine* 40, 636–642. doi: 10.1016/j.ebiom.2018.12.043

Kingma, D., and Ba, J. (2014). Adam a method for stochastic optimization. *Comput. Sci.* 13, 1–13. doi: 10.48550/arXiv.1412.6980

Lang, M., Silva, D., Dai, L., Kshettry, V. R., Woodard, T. D., Sindwani, R., et al. (2018). Superiority of constructive interference in steady-state MRI sequencing over T1-weighted MRI sequencing for evaluating cavernous sinus invasion by pituitary macroadenomas. *J. Neurosurg.* Online ahead of print. doi: 10.3171/2017.9.Jns171699

Lee, M., Lupp, A., Mendoza, N., Martin, N., Beschorner, R., Honegger, J., et al. (2015). SSTR3 is a putative target for the medical treatment of gonadotroph adenomas of the pituitary. *Endocr. Relat. Cancer* 22, 111–119. doi: 10.1530/erc-14-0472

Li, H., Zhao, Q., Zhang, Y., Sai, K., Xu, L., Mou, Y., et al. (2021). Image-driven classification of functioning and nonfunctioning pituitary adenoma by deep convolutional neural networks. *Comput. Struct. Biotechnol. J.* 19, 3077–3086. doi: 10.1016/j.csbj.2021.05.023

Lopes, M. B. S. (2017). The 2017 world health organization classification of tumors of the pituitary gland: a summary. *Acta Neuropathol.* 134, 521–535. doi: 10.1007/s00401-017-1769-8

Luo, M., Tan, Y., Chen, W., Hu, B., Wang, Z., Zhu, D., et al. (2021). Clinical efficacy of temozolomide and its predictors in aggressive pituitary tumors and pituitary carcinomas: a systematic review and meta-analysis. *Front. Neurol.* 12:700007. doi: 10.3389/fneur.2021.700007

Micko, A., Oberndorfer, J., Weninger, W. J., Vila, G., Höftberger, R., Wolfsberger, S., et al. (2019). Challenging knosp high-grade pituitary adenomas. *J. Neurosurg.* 132, 1739–1746. doi: 10.3171/2019.3.Jns19367

Mooney, M. A., Hardesty, D. A., Sheehy, J. P., Bird, C. R., Chapple, K., White, W. L., et al. (2017b). Rater reliability of the hardy classification for pituitary adenomas in the magnetic resonance imaging era. *J. Neurol. Surg. B Skull Base* 78, 413–418. doi: 10.1055/s-0037-1603649

Mooney, M. A., Hardesty, D. A., Sheehy, J. P., Bird, R., Chapple, K., White, W. L., et al. (2017a). Interrater and intrarater reliability of the Knosp scale for pituitary adenoma grading. *J. Neurosurg.* 126, 1714–1719. doi: 10.3171/2016.3.Jns153044

Niu, J., Zhang, S., Ma, S., Diao, J., Zhou, W., Tian, J., et al. (2019). Preoperative prediction of cavernous sinus invasion by pituitary adenomas using a radiomics method based on magnetic resonance images. *Eur. Radiol.* 29, 1625–1634. doi: 10.1007/s00330-018-5725-3

Principe, M., Chanal, M., Ilie, M. D., Ziverec, A., Vasiljevic, A., Jouanneau, E., et al. (2020). Immune landscape of pituitary tumors reveals association between macrophages and gonadotroph tumor invasion. *J. Clin. Endocrinol. Metab.* 105, 520–531. doi: 10.1210/clinem/dgaa520

Raverot, G., Burman, P., McCormack, A., Heaney, A., Petersenn, S., Popovic, V., et al. (2018). European society of endocrinology clinical practice guidelines for the management of aggressive pituitary tumours and carcinomas. *Eur. J. Endocrinol.* 178, 1–24. doi: 10.1530/eje-17-0796

Trouillas, J., Jaffrain-Rea, M. L., Vasiljevic, A., Raverot, G., Roncaroli, F., and Villa, C. (2020). How to classify the pituitary neuroendocrine tumors (PitNET)s in 2020. *Cancers (Basel)* 12, 514–531. doi: 10.3390/cancers12020514

Wang, H., Zhang, W., Li, S., Fan, Y., Feng, M., and Wang, R. (2021). Development and evaluation of deep learning-based automated segmentation of pituitary adenoma in clinical task. *J. Clin. Endocrinol. Metab.* Online ahead of print. doi: 10.1210/clinem/dgab371

Wei, R., Jiang, C., Gao, J., Xu, P., Zhang, D., Sun, Z., et al. (2020). Deep-Learning approach to automatic identification of facial anomalies in endocrine disorders. *Neuroendocrinology* 110, 328–337. doi: 10.1159/000502211

Wong, L. M., King, A. D., Ai, Q. Y. H., Lam, W. K. J., Poon, D. M. C., Ma, B. B. Y., et al. (2021). Convolutional neural network for discriminating nasopharyngeal carcinoma and benign hyperplasia on MRI. *Eur. Radiol.* 31, 3856–3863. doi: 10.1007/s00330-020-07451-y

Yang, Q., and Li, X. (2019). Molecular network basis of invasive pituitary adenoma: a review. *Front. Endocrinol. (Lausanne)* 10:7–15. doi: 10.3389/fendo.2019.00007

Yip, S. S., and Aerts, H. J. (2016). Applications and limitations of radiomics. *Phys. Med. Biol.* 61, 150–166. doi: 10.1088/0031-9155/61/13/R150

Yoneoka, Y., Watanabe, N., Matsuzawa, H., Tsumanuma, I., Ueki, S., Nakada, T., et al. (2008). Preoperative depiction of cavernous sinus invasion by pituitary macroadenoma using three-dimensional anisotropy contrast periodically rotated overlapping parallel lines with enhanced reconstruction imaging on a 3-tesla system. *J. Neurosurg.* 108, 37–41. doi: 10.3171/jns/2008/108/01/0037