

Article

Deep Learning for Automatically Visual Evoked Potential Classification During Surgical Decompression of Sellar Region Tumors

Nidan Qiao^{1,*}, Mengju Song^{2,*}, Zhao Ye³, Wenqiang He³, Zengyi Ma³, Yongfei Wang³, Yuyan Zhang⁴, and Xuefei Shou³

¹ Shanghai Pituitary Tumor Center, Shanghai Neurosurgical Research Institute, Department of Neurosurgery, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China & Neuroendocrine Unit, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

² Department of Ophthalmology, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China & Putuo Oculopathy Dental Disease Prevention & Cure Clinic, Shanghai, China

³ Shanghai Pituitary Tumor Center, Shanghai Neurosurgical Research Institute, Department of Neurosurgery, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China

⁴ Department of Ophthalmology, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China

Correspondence: Yuyan Zhang, 12 Wulumuqi Zhong Road, Department of Ophthalmology, Huashan Hospital, Shanghai, China. e-mail: yuyan8688@163.com

Xuefei Shou, 12 Wulumuqi Zhong Road, Department of Neurosurgery, Huashan Hospital, Shanghai, China. e-mail: shouxf@hotmail.com

Received: 10 June 2019

Accepted: 17 September 2019

Published: 20 November 2019

Keywords: artificial intelligence; optic chiasm; intraoperative monitoring; neural network

Citation: Qiao N, Song M, Ye Z, He W, Ma Z, Wang Y, Zhang Y, Shou X. Deep learning for automatically visual evoked potential classification during surgical decompression of sellar region tumors. *Trans Vis Sci Tech.* 2019;8(6):21, <https://doi.org/10.1167/tvst.8.6.21>
Copyright 2019 The Authors

Purpose: Detection of the huge amount of data generated in real-time visual evoked potential (VEP) requires labor-intensive work and experienced electrophysiologists. This study aims to build an automatic VEP classification system by using a deep learning algorithm.

Methods: Patients with sellar region tumor and optic chiasm compression were enrolled. Flash VEP monitoring was applied during surgical decompression. Sequential VEP images were fed into three neural network algorithms to train VEP classification models.

Results: We included 76 patients. During surgical decompression, we observed 68 eyes with increased VEP amplitude, 47 eyes with a transient decrease, and 37 eyes without change. We generated 2,843 sequences (39,802 images) in total (887 sequences with increasing VEP, 276 sequences with decreasing VEP, and 1680 sequences without change). The model combining convolutional and recurrent neural network had the highest accuracy (87.4%; 95% confidence interval, 84.2%–90.1%). The sensitivity of predicting no change VEP, increasing VEP, and decreasing VEP was 92.6%, 78.9%, and 83.7%, respectively. The specificity of predicting no change VEP, increasing VEP, and decreasing VEP was 80.5%, 93.3%, and 100.0%, respectively. The class activation map visualization technique showed that the P2-N3-P3 complex was important in determining the output.

Conclusions: We identified three VEP responses (no change, increase, and decrease) during transsphenoidal surgical decompression of sellar region tumors. We developed a deep learning model to classify the sequential changes of intraoperative VEP.

Translational Relevance: Our model may have the potential to be applied in real-time monitoring during surgical resection of sellar region tumors.

Introduction

More than 20% of tumors in the central nervous system originate in the sellar region, which houses the pituitary gland.^{1,2} The most common among these

sellar tumors are pituitary adenomas, craniopharyngiomas, and meningiomas. Visual dysfunctions are usually one of the complaints in these patients and are indications for surgical decompression. The close relationship of tumors and the optic nerve or chiasm makes the latter vulnerable to any direct intraopera-

tive injury, blood supply reduction, and indirect insult due to heat conduction during coagulation. Although the visual outcomes are favorable in most patients, 10% of the patients are not able to fully recover after surgical decompression.³

Electrophysiological monitoring by the intraoperative real-time recording of visual evoked potential (VEP) can detect any possible VEP deviations during the surgical procedure and act as a surrogate for optic nerve injury. Several studies showed that this procedure had the potential to improve visual outcome.⁴⁻⁶ Nevertheless, detection of the massive amount of data generated in real-time requires labor-intensive work, which limits its wide usage in clinical settings. Moreover, the lack of experts who are experienced with electrophysiology and variations in different statuses of anesthesia make the real-time monitoring of VEP recordings difficult. The computer-aided analysis may have a solution to these issues.

A source of this assistance may lie in the rapidly progressing domain of artificial intelligence known as the neural network: a recent study has found that neural network algorithm performed well in classifying images from multifocal VEP data.⁷ Several studies also used an algorithm to predict steady-state VEP data.^{8,9} Whereas in the dynamic field, multiple studies applied artificial intelligence in real-time: a study used hospital-wide deployment of artificial intelligence to discern real-time clinician behaviors,¹⁰ and another study focused on early diagnosis of patients with heart failure.¹¹ No study was performed on the dynamic detection of abnormal VEP in real-time by using neural network algorithms. This study aims to determine whether deep learning algorithms can detect abnormal VEP. We hypothesize that VEP changing tendency can be identified by a deep learning algorithm during surgical decompression of sellar region tumors.

Methods

We enrolled 76 patients (36 female, 47.3%; mean age, 45.7 years old) with sellar region tumor, and the chiasm compression was indicated on the preoperative magnetic resonance imaging scan. All the patients underwent transsphenoidal tumor resection. The study was approved by Huashan Hospital Institutional Review Board and was in adherence to the Declaration of Helsinki. Informed consents were obtained from all the participants. Preoperational evaluations included thorough ophthalmic examinations as well as visual acuity and visual field

examination to exclude patients with fundus disease, glaucoma, or traumatic optic injury. Among 152 eyes included in this study, 60 eyes had decreased visual acuity, including 19 eyes that had visual acuity less than 0.1, and 87 eyes had an abnormal visual field (36 eyes had quadrantanopia and 51 eyes had hemianopia or worse).

Intraoperative VEP Recording

Flash VEP monitoring was applied using the NIM Eclipse system (Medtronic, Minneapolis, MN), with the International Society for Clinical Electrophysiology of Vision VEP standard.¹² The scalp for electrode placement (the midline active electrode was located at Oz, two lateral active electrodes were located at O₁ and O₂, and the reference electrode was located at Fz) was prepared by shaving and cleaning before recording channels were connected with gold cup electrodes. Conductive paste was applied to decrease the impedance of the electrodes. Light-proof goggles with a flashing light-emitting diode for visual stimulation were placed over the two eyes. Monocular flashing stimulation was made at a rate of 1 Hz alternatively.

The preoperational recording was obtained before the anesthesia. Propofol was used to induce anesthesia with a 3-mg/kg/h maintaining dosage. We did not use inhaling anesthesia. The baseline of postanesthesia was obtained 5 minutes after successful anesthesia. VEPs were recorded during the surgical process in three channels per eye.

Preanalysis Processing

We measured amplitude from the positive P2 peak at around 120 ms to the preceding N2 negative peak at around 90 ms. Latency was measured as the time from stimulus onset to the P2 amplitude. To minimize the noise during monitoring, we used the average response over 5 minutes after stable anesthesia to measure the amplitude and latency. Then, the amplitude and latency after anesthesia were used as the baseline, and change from the baseline was calculated in status after tumor resection.

To train the model for detecting VEP change, we identified no change, increase, and decrease (the criterion was defined as a >25% increase or >25% decrease in amplitude compared with the baseline) VEP sequences during surgical decompression. Preanalysis processing of the VEP sequences included extracting VEP images from all three channels in each eye and combining 14 VEP images over 5 minutes into

a sequence. Images with huge noise and artifacts were excluded. All the sequences were transformed into 350×90 -pixel size and vectorized using the value of each pixel.

Statistical Analysis

We applied three neural network models (a three-layer convolutional neural network, a pretrained convolutional neural network, and a combination of a convolutional and recurrent neural network) to detect different VEP responses. In the first model (Supplementary Table S1), convolutional neural network alone has the potential to distinguish among different responses. In this model, we used three convolutional layers and max-pooling layers. In the second model (Supplementary Table S2), a VGG19 architecture with preinitialization weights from the same network trained in our previous work⁷ was used. In the third model (Supplementary Table S3), the convolutional neural network first recognized the images from each sequence; then, the recurrent neural network distinguished the sequence. In this model, multiple convolutional layers were followed by a long-short memory layer. We used another simpler model (one long-short memory layer) with amplitude and latency from every single VEP image as inputs for comparison. Workflow for VEP analysis is provided in Figure 1.

During the training process, random dropout was used to prevent overfitting. Data were randomly split into training dataset (60%), validation dataset (20%), and test dataset (20%). We used bootstrap (a new dataset was created by sampling from the original data in every step of the bootstrap) to test the robustness of our built models in the test dataset. Confusion matrix of the whole cohort was calculated with 5-fold cross-validation. Mean and 95% confidence intervals of the accuracy were provided. To provide model interpretability, we used class activation map (CAM) visualization technique to demonstrate the model explanation.¹³ The CAM computes how important each location is concerning the classification. For instance, given an image fed into our “no change versus decreasing versus increasing” model, CAM visualization allows generation of an “importance” heatmap for any input image. All the analyses were performed on Python 3.6 with the Keras package (version 2.1.1).

All the post-processing data can be assessed by request. The code will be available on GitHub (<https://github.com/norikaisa/DeepiVEP>).

Results

Characteristics of VEPs of the Cohort

Preoperation baseline VEP amplitude was $4.0 \pm 2.4 \mu\text{v}$, and the latency was $99.3 \pm 27.2 \text{ ms}$. After anesthesia, VEP amplitude decreased to $1.5 \pm 1.4 \mu\text{v}$, and the latency increased to $104.7 \pm 29.3 \text{ ms}$ (Supplementary Figure S1). The change was significant in VEP amplitude (-107% ; interquartile range, -291% to -5%) but not in VEP latency (7% ; interquartile range, -25% to 24%).

During surgical decompression of the optic chiasm, we observed 68 eyes had increased VEP amplitude. We also identified 47 eyes with a transient decrease in VEP amplitude and 37 eyes without change during the surgical decompression (Table 1). These statuses were used as the ground truth for future training. Eyes with transient VEP amplitude decrease had better preoperative visual acuity. On the contrary, eyes without obvious change had worse visual acuity.

A sequence comprised 14 VEP images over 5 minutes, including the changing amplitude (Fig. 2). We created 1931 sequences during surgical decompression (887 sequences with increasing VEP, 276 sequences with decreasing VEP, and 768 sequences without change). Thus, on average, 4.2 sequences were generated per eye per channel. We further increased the sample size of VEP without change to 1680 sequences (adding 912 sequences from the baseline). The total sample size of our study was 2843 sequences (39,802 images).

Deep Learning Models

First, we built a simple convolutional neural network (Supplementary Table S1) to classify the three VEP status. The model got an accuracy of 82.7% (95% confidence interval [CI], 78.3%–86.5%). The sensitivity of predicting no change VEP, increasing VEP, and decreasing VEP was 84.6%, 79.3%, and 88.5%, respectively. The specificity of predicting no change VEP, increasing VEP, and decreasing VEP was 82.4%, 88.3%, and 99.3%, respectively.

In transfer learning using pretrained VGG16 structure (Supplementary Table S2), the accuracy was 84.1% (95% CI, 79.8%–87.8%). The sensitivity of predicting no change VEP, increasing VEP, and decreasing VEP was 90.4%, 80.5%, and 86.0%, respectively. The specificity of predicting no change VEP, increasing VEP, and decreasing VEP was 83.3%, 90.4%, and 99.3%, respectively.

Table 1. Gold Standard of the Classification in Visual Evoked Potential

Visual Measurements	No Change (<i>n</i> = 37)	Increasing (<i>n</i> = 68)	Decreasing (<i>n</i> = 47)	<i>P</i>
Visual acuity, no. (%)				
Normal	11 (29.7)	33 (48.5)	29 (61.7)	0.003
Abnormal	26 (70.3)	35 (51.5)	18 (38.3)	
Visual field, no. (%)				
Normal	15 (41.7)	29 (42.6)	21 (44.7)	0.484
Abnormal	22 (58.3)	39 (57.2)	26 (55.3)	
Preanesthesia VEP, mean ± SD				
Amplitude (μv)	4.4 ± 2.1	3.6 ± 2.1	4.2 ± 2.9	0.253
Latency (ms)	103.3 ± 28.4	97.2 ± 26.3	99.5 ± 27.9	0.586
VEP after stable anesthesia (baseline)				
Amplitude (μv), mean ± SD	2.2 ± 0.9	1.2 ± 0.7	1.5 ± 2.0	0.001
Amplitude change from preanesthesia, (CI), %	-53 (-70, -20)	-67 (-79, -45)	-71 (-107, -21)	0.037
Latency (ms), mean ± SD	112.8 ± 28.3	102.2 ± 28.7	102.0 ± 30.4	0.153
Latency change from preanesthesia, (CI), %	13 (-5, 31)	8 (-23, 40)	2 (-19, 32)	0.892
VEP during tumor decompression				
Amplitude (μv), mean ± SD	2.1 ± 0.9	3.0 ± 1.9	1.4 ± 1.2	<0.001
Amplitude change from baseline, (CI), %	0 (-10, 0)	146 (69, 270)	-75 (-270, -43)	<0.001
Latency (ms), mean ± SD	99.1 ± 29.3	95.6 ± 27.1	100.6 ± 25.7	0.604
Latency change from baseline, (CI), %	-5 (-17, 0)	-3 (-15, 5)	0 (-12, 17)	0.058

no., number; CI, 95% confidence interval.

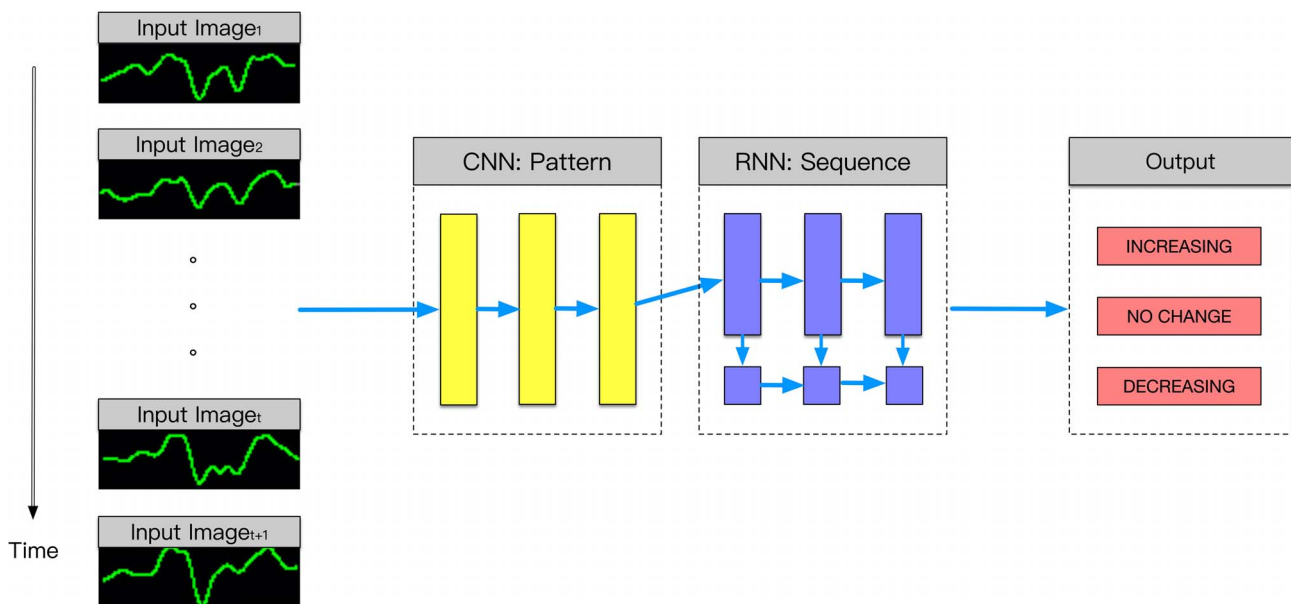


Figure 1. The proposed workflow of analyzing intraoperative visual evoked potential. Time sequential visual evoked potentials were inputted to a convolutional neural network, followed by a recurrent neural network to predict no change, increasing, or decreasing. CNN, convolutional neural network; RNN, recurrent neural network.

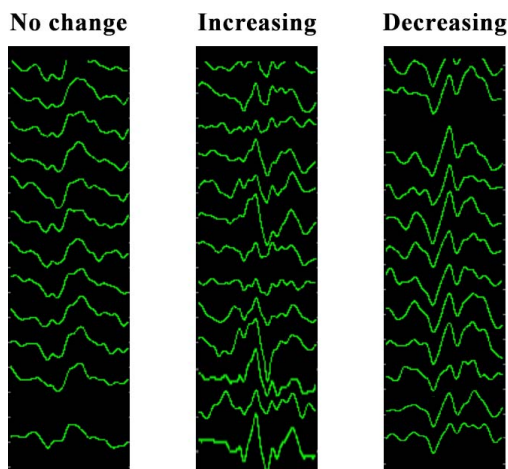


Figure 2. Examples of preprocessed visual evoked potential sequences.

In the model combining convolutional and recurrent neural network (Supplementary Table S3), the accuracy was 87.4% (95% CI, 84.2%–90.1%). The sensitivity of predicting no change VEP, increasing VEP, and decreasing VEP was 92.6%, 78.9%, and 83.7%, respectively. The specificity of predicting no change VEP, increasing VEP, and decreasing VEP was 80.5%, 93.3%, and 100.0%, respectively. The 5-fold cross-validation of this model demonstrated that 2488 sequences were correctly classified (Fig. 3).

The simpler model only using amplitude and latency from every single VEP images as prediction features yielded an accuracy of 83.1% (95% CI, 81.7%–84.5%).

Explanation of the Model

We calculated CAM in a typical image, and the visualization showed that VEP images in the bottom

	Actual label		
	no change	increasing	decreasing
Predicted no change	1558	186	40
Predicted increasing	122	700	6
Predicted decreasing	0	1	230

Figure 3. Confusion matrix of the whole cohort after cross-validation.

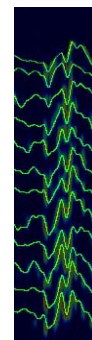


Figure 4. Class activation map visualization technique to demonstrate the model explanation. The visualization showed that visual evoked potential images in the bottom area (later time) and in the P2-N3-P3 complex were more important in determining the output.

area (later time) and the P2-N3-P3 complex were more important in determining the output (Fig. 4).

Discussion

In this paper, we built a workflow (extracting, preprocessing, and analysis) of the intraoperative monitoring VEP. Deep learning models were trained to detect the sequential change of intraoperative VEP. The model performance was comparable to human intelligence level in terms of differentiating VEPs without change, with increasing amplitude, or with decreasing amplitude. The results suggest our models can potentially assist or partially substitute human labor for VEP monitoring during surgical resection of sellar region tumors.

Intraoperative VEP monitoring was introduced to the field of neurosurgery in the 1970s¹⁴ when multiple publications demonstrated the usefulness of this technique to protect optic nerve from surgical injury.^{15,16} But a few publications indicated that this technique might be susceptible nonspecific influences, such as anesthesia,¹⁷ blood pressure, oxygen saturation, and bone procedures.^{18,19} We observed a huge discrepancy before the anesthesia and after anesthesia—the amplitude decreased by roughly 60% and the latency increased by roughly 5%, which corresponded to most of the previous studies.

Several publications have made criteria to predict postoperative visual functions; for example, Harding et al.¹⁶ argued that the absence of a previously normal VEP for more than 4 minutes during surgical manipulation within the orbit showed a correlation with postoperative impairment of vision. The decrease in latency was correlated with better visual prognosis

in several publications.¹³ Several groups used a 20% to 50% increase or decrease in amplitude as the criterion.^{20–22} Recent studies also argued the effectiveness of intraoperative VEP monitoring during surgeries that might have the risk of optic injury.^{4–6} On the contrary, several studies concluded that this procedure had no predictive value for postoperative prognosis.^{18,19} We cannot rule out publication bias where a negative result was less likely to be published.

In the previous paper, we built a convolutional neural network to differentiate normal VEPs from abnormal VEPs from signals obtained from multifocal VEP examination.⁷ Still images are more suitable for the convolutional neural network. In data with dynamic properties, a combination of the convolutional and recurrent neural network was more suitable. The recurrent neural network has been proven to be useful in analyzing data, such as clinical notes,^{23,24} anesthesia parameters,²⁵ and cardiographs.²⁶ Here, we combined a convolutional neural network and recurrent neural network with the assumption that the former can differentiate static images and the latter can recognize dynamic patterns. We chose the long-short memory layer because of its property of selectively remembering and forgetting patterns for long and short durations of time. The performance of the combining model outranged that of a simpler model using only amplitude and latency, the single convolutional neural network, and even the pretrained convolutional neural network, which suggested the usefulness of our model. We also argue that our model is simple enough in practice because the workflow (recording VEP sequences, extracting pixels, and feeding models) was automated.

Tumors for which craniotomy was warranted decades ago can now be resected using neuroendoscope with less risk of optic nerve injury. In our included cases, we always used the endoscopic transsphenoidal approach where direct injury to the optic apparatus was less likely to happen. Although we did not have enough cases for VEP decrease during monitoring, the differentiating power in these cases was not compromised. We will use generative adversarial networks to generate simulated decreasing VEP responses in future studies.

We have several limitations to our study. We did not have postoperative visual outcomes, but we argued the correlation of intraoperative monitoring and the outcome had been investigated by other studies,^{4–6,14–21} and our study did not focus on this topic. We have a relatively small sample size for a deep learning study. But, we generated more than

10,000 images from these samples, which were sufficient for deep learning training. The developed system should be investigated in the real-time setting to discern if the system can truly detect the increasing or decreasing amplitude, especially for those signals with noise artifacts. Future studies might include “noise” as one of the outcomes and let the net decide if these artifacts should be excluded or not. The generalizability of the model should be studied in a wider broad extend including other institutions.

Conclusion

In this paper, we developed a deep learning model to monitor the sequential change of intraoperative VEP. The model performance was comparable to human intelligence level in terms of differentiating VEP with increasing amplitude, decreasing amplitude, or with no change.

Acknowledgments

This study is supported by Shanghai Committee of Science and Technology, China (grant numbers 18441901400, 16ZR1404500, and 17JC1402100). Nidan Qiao is supported by the 2018 Milstein Medical Asian American Partnership Foundation translational medicine fellowship. We thank Brooke Swearingen for his language editing on the manuscript.

Disclosure: **N. Qiao**, None; **M. Song**, None; **Z. Ye**, None; **W. He**, None; **Z. Ma**, None; **Y. Wang**, None; **Y. Zhang**, None; **X. Shou**, None

*Nidan Qiao and Mengju Song contributed equally to the manuscript.

References

1. Bresson D, Herman P, Polivka M, Froelich S. Sellar lesions/pathology. *Otolaryngol Clin North Am.* 2016;49:63–93.
2. Freda PU, Post KD. Differential diagnosis of sellar masses. *Endocrinol Metab Clin North Am.* 1999;28:81–117, vi.
3. Qiao N, Ye Z, Shou X, Wang Y, Li S, Wang M, Zhao Y. Discrepancy between structural and functional visual recovery in patients after transsphenoidal pituitary adenoma resection. *Clin Neurol Neurosurg.* 2016;151:9–17.
4. Feng R, Schwartz J, Loewenstern J, et al. The predictive role of intra-operative visual evoked

- potentials (VEP) in visual improvement after endoscopic pituitary tumor resection in large and complex tumors: description and validation of a method. *World Neurosurg.* 2019;126:e136–e143.
5. Gutzwiller EM, Cabrilo I, Radovanovic I, Schaller K, Boëx C. Intraoperative monitoring with visual evoked potentials for brain surgeries. *J Neurosurg.* 2018;130:654–660.
 6. Toyama K, Wanibuchi M, Honma T, Komatsu K, Akiyama Y, Mikami T, Mikuni N. Effectiveness of intraoperative visual evoked potential in avoiding visual deterioration during endonasal transsphenoidal surgery for pituitary tumors. *Neurosurg Rev.* 2018;93:311–317.
 7. Qiao N. Using deep learning for the classification of images generated by multifocal visual evoked potential. *Front Neurol.* 2018;9:905–904.
 8. Palaniappan R, Mandic DP. Biometrics from brain electrical activity: a machine learning approach. *IEEE Trans Pattern Anal Mach Intell.* 2007;29:738–742.
 9. Ma T, Li H, Yang H, Lv X, Li P, Liu T, Yao D, Xu P. The extraction of motion-onset VEP BCI features based on deep learning and compressed sensing. *J Neurosci Methods.* 2017;275:80–92.
 10. Yeung S, Downing NL, Fei-Fei L, Milstein A. Bedside computer vision—moving artificial intelligence from driver assistance to patient safety. *N Engl J Med.* 2018;378:1271–1273.
 11. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Assoc.* 2017;24:361–370.
 12. Odom JV, Bach M, Brigell M, et al. ISCEV standard for clinical visual evoked potentials: (2016 update). *Doc Ophthalmol.* 2016;133:1–9.
 13. Ramakrishna V, Michael C, Abhishek D, Devi P, Dhruv B. Grad-CAM: visual explanations from deep networks via gradient-based localization. *2017 IEEE International Conference on Computer Vision (ICCV).* <https://doi.org/10.1109/ICCV.2017.74>
 14. Wright JE, Arden G, Jones BR. Continuous monitoring of the visually evoked response during intra-orbital surgery. *Trans Ophthalmol Soc UK.* 1973;93:311–314.
 15. Hussain SS, Laljee HC, Horrocks JM, Tec H, Grace AR. Monitoring of intra-operative visual evoked potentials during functional endoscopic sinus surgery (FESS) under general anaesthesia. *J Laryngol Otol.* 1996;110:31–36.
 16. Harding GF, Bland JD, Smith VH. Visual evoked potential monitoring of optic nerve function during surgery. *J Neurol Neurosurg Psychiatr.* 1990;53:890–895.
 17. Wiedemayer H, Fauser B, Armbruster W, Gasser T, Stolke D. Visual evoked potentials for intraoperative neurophysiologic monitoring using total intravenous anesthesia. *J Neurosurg Anesthesiol.* 2003;15:19–24.
 18. Chacko AG, Babu KS, Chandy MJ. Value of visual evoked potential monitoring during transsphenoidal pituitary surgery. *Br J Neurosurg.* 1996;10:275–278.
 19. Cedzich C, Schramm J, Fahlbusch R. Are flash-evoked visual potentials useful for intraoperative monitoring of visual pathway function? *Neurosurgery.* 1987;21:709–715.
 20. Chung S-B, Park C-W, Seo D-W, Kong D-S, Park S-K. Intraoperative visual evoked potential has no association with postoperative visual outcomes in transsphenoidal surgery. *Acta Neurochir (Wien).* 2012;154:1505–1510.
 21. Kamio Y, Sakai N, Sameshima T, et al. Usefulness of intraoperative monitoring of visual evoked potentials in transsphenoidal surgery. *Neurol Med Chir (Tokyo).* 2014;54:606–611.
 22. Luo Y, Regli L, Bozinov O, Sarnthein J. Clinical utility and limitations of intraoperative monitoring of visual evoked potentials. *PLoS One.* 2015; 10:e0120525.
 23. Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. *Sci Rep.* 2018;8:6085.
 24. Li R, Hu B, Liu F, et al. Detection of bleeding events in electronic health record notes using convolutional neural network models enhanced with recurrent neural network autoencoders: deep learning approach. *JMIR Med Inform.* 2019;7: e10788.
 25. Lundberg SC, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng.* 2018;2:749–760.
 26. Xiong Z, Nash MP, Cheng E, Fedorov VV, Stiles MK, Zhao J. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiol Meas.* 2018;39:094006.