

Longitudinal MRI assessment: The identification of relevant features in the development of Posterior Fossa Syndrome in children

M. Spiteri^a, E. Lewis^a, D. Windridge^a and S. Avula^b

^aCentre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK, GU2 7XH;

^b Alder Hey Children's Hospital (NHS), Eaton Rd, Liverpool, UK, L12 2AP

1. ABSTRACT

Up to 25% of children who undergo brain tumour resection surgery in the posterior fossa develop posterior fossa syndrome (PFS). This syndrome is characterised by mutism and disturbance in speech. Our hypothesis is that there is a correlation between PFS and the occurrence of hypertrophic olivary degeneration (HOD) in lobes within the posterior fossa, known as the inferior olivary nuclei (ION). HOD is exhibited as an increase in size and intensity of the ION on an MR image.

Intra-operative MRI (IoMRI) is used during surgical procedures at the Alder Hey Children's Hospital, Liverpool, England, in the treatment of Posterior Fossa tumours and allows visualisation of the brain during surgery. The final MR scan on the IoMRI allows early assessment of the ION immediately after the surgical procedure.

The longitudinal MRI data of 28 patients was analysed in a collaborative study with Alder Hey Children's Hospital, in order to identify the most relevant imaging features that relate to the development of PFS, specifically related to HOD.

A semi-automated segmentation process was carried out to delineate the ION on each MRI. Feature selection techniques were used to identify the most relevant features amongst the MRI data, demographics and clinical data provided by the hospital. A support vector machine (SVM) was used to analyse the discriminative ability of the selected features. The results indicate the presence of HOD as the most efficient feature that correlates with the development of PFS, followed by the change in intensity and size of the ION and whether HOD occurred bilaterally or unilaterally.

Keywords: Posterior Fossa Syndrome, Cerebellar Mutism Syndrome, Inferior Olivary Nuclei, Hypertrophic Olivary Degeneration, Intra-operative MRI, brain, MRI, posterior fossa

2. INTRODUCTION

Up to 1 in 4 children who undergo brain tumour resection surgery in the posterior fossa develop a syndrome known as Posterior Fossa Syndrome (PFS).¹ This syndrome, also known as cerebellar mutism syndrome (CMS), describes a set of neurological symptoms, similar to stroke, which may develop from 24 to 107 hours after surgery.² Children suffering from PFS, characteristically suffer from disturbance in speech and mutism, but may also suffer from loss of muscle tone, incontinence, strabismus (cross-eyed), dysphagia, and personality changes such as anger, apathy, melancholy, crying and screaming.² The development of such a syndrome in children hinders their development and highly impacts their quality of life. It is therefore of great importance to identify biomarkers predictive of the onset of this syndrome in order to enhance clinician decision-making over treatment and prognosis, as well as to understand its underlying causes.

Further author information: (Send correspondence to M.S.)

M.S.: E-mail: m.spiteri@surrey.ac.uk, Telephone: +44 (0)1483 68 9856

E.L.: E-mail: e.lewis@surrey.ac.uk, Telephone: +44 (0)1483 68 6047

D.W.: E-mail: d.windridge@mdx.ac.uk Telephone: +44 (0)1483 68 6048

S.A.: E-mail: shivaram.avula@alderhey.nhs.co.uk Telephone: +44 (0)1512 52 5432

Our hypothesis is that there is a correlation between PFS and the occurrence of hypertrophic olivary degeneration (HOD) in lobes within the posterior fossa, known as the inferior olivary nuclei (ION). HOD is exhibited as an increase in size and intensity of the ION on an MR image.³ Intra-operative MRI (IoMRI) is currently being carried out during surgical procedures at the Alder Hey Children’s Hospital, Liverpool, England, in the treatment of such tumours. The final MR scan on the IoMRI allows early assessment of the ION immediately after the surgical procedure.

This study focuses on proving a link between PFS and HOD. Association of HOD and PFS will add to the existing evidence on the development of PFS and lead to a deeper understanding of the pathogenesis of the syndrome. A dataset of 28 patients was included in this study. The dataset was compiled from a number of patients treated for various types of posterior fossa tumours at Alder hey Children’s Hospital between 2007 and 2013. The patients were aged between 8 months and 18 years old, eleven of whom were diagnosed with PFS. Thirteen of these patients exhibited HOD, nine bilaterally (in both ION) and nine unilaterally (in either the left or right ION). Follow up MR imaging during one year post-surgery were reviewed and up to five MR images were acquired longitudinally for each patient across their treatment. The process undertaken in this study shall be explained in Section 2, with Section 3 exhibiting some preliminary results and analyses. Further results shall be released in the final paper.

3. METHODOLOGY

The study consisted of three sections: Feature Extraction, Feature Selection and Classification. These sections are explained in further detail below.

3.1 Image Processing

The aim of this investigation is to identify the most relevant imaging and clinical features that correlate with the development of PFS following tumour resection surgery in the posterior fossa. For this reason techniques were chosen to ensure that comprehensibility of imaging and clinical features is retained throughout the pipeline.

In order to extract information (features) about each ION it was required to segment these structures on the MR images. The modality of MRI used to acquire the images was spiral MRI, a modality which exhibits efficient k-space sampling as well as high temporal and spatial resolution. Such a modality limits the acquisition time, enabling it to be an ideal modality to acquire MR images of young restless patients. Due to the manner in which Spiral MR images are obtained a full volumetric representation is not obtained. For this reason segmentation was carried out in two-dimensional image slices.

The normal ION cannot be clearly delineated by the naked eye on MRI. This is due to very low contrast with the surrounding tissue as well as its relatively small cross-sectional area. For this reason, segmentation was carried out using a semi-automated seed-growing technique in two-dimensional space. The right and left ION were segmented separately. The process by which segmentation is carried out consists of three main steps: (1) the identification of a seed-point to initialise the region of interest (according to pixel intensity), (2) the application of a morphological closing operation to fill gaps. Step (2) is applied iteratively until a satisfactory region of interest is obtained.

The contrast between the ION and their surrounding tissue was calculated using the equation for Weber contrast in Equation 1 where I_{ION} refers to the mean grey level intensity of the ION and I_b refers to the mean grey level intensity of the surrounding tissue.

$$W = \frac{I_{ION} - I_b}{I_b} \tag{1}$$

For each MRI, the contrast of both the left and right ION was calculated separately. The segmentation of the left and right ION is exhibited in Figure 1.

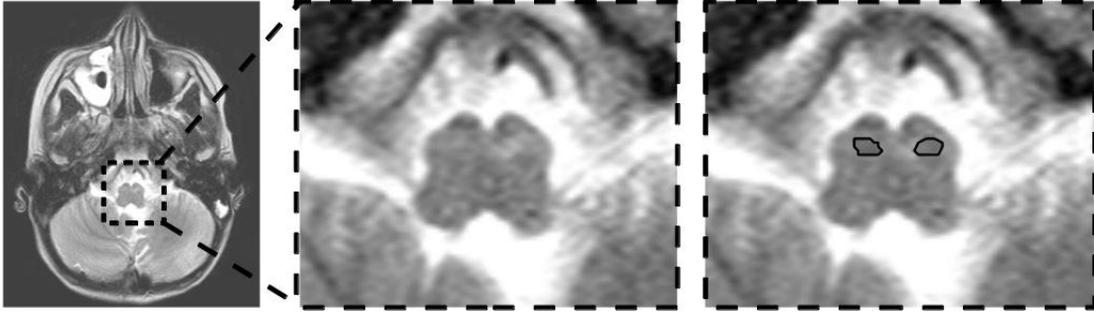


Figure 1: Segmentation of Inferior Olivary Nuclei (delineated in black)

3.2 Feature Extraction

Once the desired region was segmented it was possible to extract a set of features from each ION. Further clinical data was obtained from Alder Hey Children’s hospital. The features included are as follows: (1) ”Average slope of contrast in left nucleus”; (2) ”Variance of contrast in left nucleus”; (3) ”Average slope of area in left nucleus”; (4) ”Variance of area in left nucleus”; (5) ”Average slope of contrast in right nucleus”; (6) ”Variance of contrast in right nucleus”; (7) ”Average slope of area in right nucleus”; (8) ”Variance of area in right nucleus”; (9) ”Presence of HOD (clinical data)”; (10) ”Bilateral HOD”; (11) ”Unilateral HOD”; (12) ”Enlargement”; (13) ”Gender”; (14) ”Age at surgery”; (15) ”Random noise 1”; (15) ”Random noise 2”. Features 1 to 8 refer to data which was extracted from the patients’ MR images, whilst the features numbered 9 to 14 refer to clinical data provided by Alder Hey Children’s hospital. Features 15 and 16 refer to random noise which was added in order to assess the discriminative ability of the feature selection algorithms. Features 1 to 8 refer to data which was extracted from the patients’ MR images, whilst the features numbered 9 to 14 refer to clinical data provided by Alder Hey Children’s hospital. Features 15 and 16 refer to random noise which was added in order to assess the discriminative ability of the feature selection algorithms later on in the study.

The contrast for both the left and right ION was obtained for up to 6 MR Images per patient which were acquired at different time points throughout each patient’s treatment. The average slope of change in contrast between these time points was obtained for features 1 and 5. The variance in contrast, in features 2 and 6, was also calculated across each patient’s longitudinal set of MR Images.

Similarly, features 3, 4, 6 and 7 were calculated by obtaining the area of the ION from each MRI and calculating the average slope and variance across each longitudinal dataset for the left and right ION separately.

Features 9, 10 and 11 are clinical features which were determined by radiological assessment of each MR image. Feature 9 is a binary feature which indicates whether an expert radiologist identified HOD as present (1) or not (0) on MRI. Features 10 and 11 are also binary features with the former indicating whether HOD was present unilaterally (1) or not (0) and the latter indicating whether HOD was present bilaterally (1) or not (0). It is important to note that these features are not mutually exclusive and the lack of presence of HOD bilaterally may imply either unilateral HOD or no HOD.

3.3 Feature Selection

To avoid over fitting the training it is desirable to use only the most relevant features in classifying data into two groups: patients who have developed PFS and those who have not. This is known as Feature Selection and can be carried out using Filter or Wrapper Methods^{4,5}. Other dimensionality reduction techniques, such as Principal Component Analysis (PCA), results in loss of comprehensibility,⁶ rendering it inappropriate to be used in this scenario due to the need for medics and clinicians to interpret results.

In general, the problem of feature selection is NP-hard, and therefore intractable. Various techniques have therefore been applied, however, these are all prone to local minima. The techniques used to identify the salient

features out of the full feature set are: random subset feature selection (RSFS), sequential forward selection (SFS) and sequential floating forward selection (SFFS).

RSFS chooses a random subset of features from the entire feature set, the size of which is equal to the square root of the total number of features. A k-NN classification using three neighbours is carried out repeatedly on this chosen subset. Each feature is given a relevance score which is continuously updated according to its inclusion in the random subsets which perform well.⁷⁸ The relevance values of each feature are compared to random walk statistics and good features are chosen accordingly. The algorithm is carried out until the stopping criterion is reached, that is, if the size of the final feature set (consisting of the features with the highest relevance scores) has not changed by more than 0.5% in the previous 1000 iterations, or if the maximum number of iterations (300, 000) is reached. The RSFS algorithm was carried out 100 times, each time randomly dividing the dataset in two. Table 1 exhibits the most relevant features and their average relevance scores obtained by this method.

Unlike RSFS, SFS starts off with an empty data set. One feature is added at a time and a feature is kept or discarded depending on whether it exhibits the best classification performance when used together with the previously chosen features. SFS also makes use of k-NN classifier on the feature subset in order to obtain a classification score. Low-scoring features were discarded. In SFFS an attempt is made at finding the least useful feature in order to discard it from the final feature set. This process is repeated until the evaluation score becomes (and remains) better than the previous best score using a feature set of the same size.⁷⁹ Both the SFS and SFFS algorithms were carried out using 3 neighbours, 4 neighbours, 5 neighbours and 6 neighbours. This process was carried out 100 times and the average relevance scores were calculated.

A k-NN classifier was used in each algorithm as it is a generative technique and retains more information on the underlying distribution of data. A support vector machine (used for classification in Section 3.4) was not ideal for this task as it is a discriminative technique and hence more ideal for binary diagnostic classification. The relevance of each feature was scored using a UAR as in the case of the RSFS algorithm.⁷⁹

3.4 Classification

It was desired to classify patients into two groups: patients who developed PFS and patients who had not developed PFS. The binary classification was carried out in order to assess the discriminative ability of the efficient features chosen in the previous stage of the study. Three different feature subsets were used, the first subset included the entire feature set, the second subset included the most efficient features chosen by the RSFS algorithm, whilst the third subset included the most efficient features chosen by the SFS and SFFS algorithms. A simple support vector machine (SVM) was used to perform the classification task. 2.3 Due to the small size of the dataset it was not feasible to use half of the data as a training set and the other half as a test set. An efficient use of the small dataset, and an ideal way to maximise the use of the dataset, was to implement a leave-one-out cross-validation (LOOCV). In this validation technique only one observation was omitted from the entire set for training purposes; the same observation was then used as the test set. Classification was carried out in this way for each observation with an average score being calculated in order to represent the efficiency of the SVM classifier. A LOOCV for each feature subset was carried out, resulting in an average accuracy score for each subset. ROC graphs were then plotted in order to assess the difference in classification accuracy when using different feature subsets.

4. RESULTS

Table 1: The relevance scores calculated by the random subset feature selection algorithm

Feature	1	2	3	4	5	6	7	8
Average Relevance	5.5102	0.6258	1.5726	0.6442	0.9689	0.0900	0.0900	0.0933
Feature	9	10	11	12	13	14	15	16
Average Relevance	2.9436	1.1780	0.1389	0.1113	0.1113	0.3912	0.1554	0.0000

Table 2: The relevance scores calculated by the sequential forward selection algorithm and the sequential floating forward selection algorithm

Feature Key	Average Relevance Score over 100 iterations			
	k=3	k=4	k=5	k=6
1	86.1111	88.1944	92.3611	90.2778
11	0	90.2778	0	0

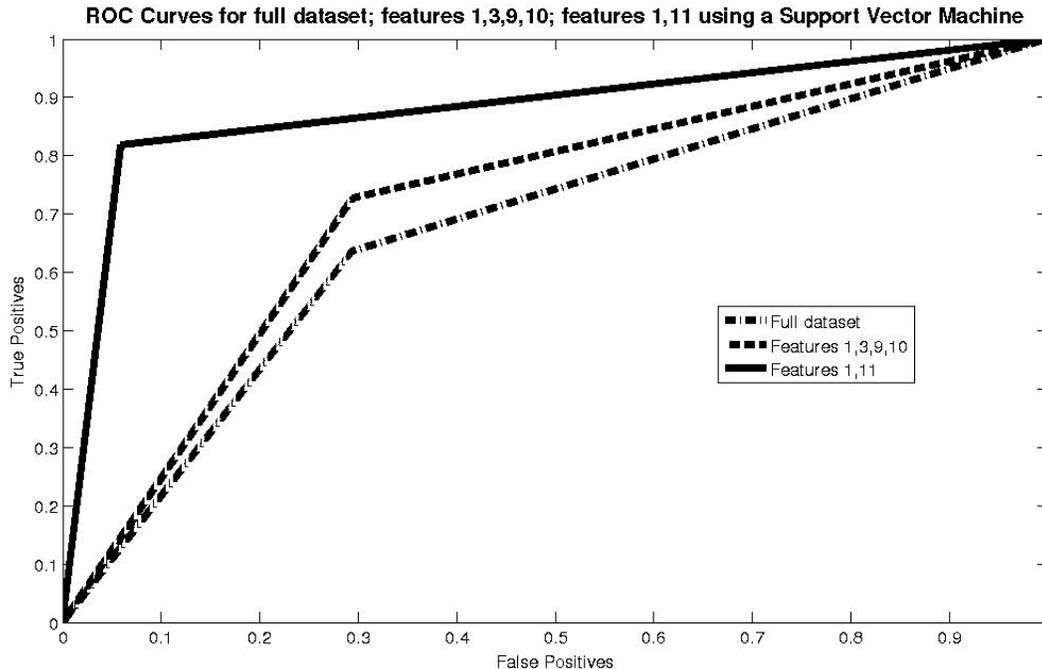


Figure 2: The receiver operating characteristic (ROC) curve for the SVM classifier used on: the full dataset, features 1,3,9,10, and features 1,11

The results yielded by the RSFS algorithm in Table 1 indicate features 1,9, 3 and 10 (in this order) as the most relevant features, scoring higher than other features considered in this study. These features correspond to the average slope of contrast in the left nucleus, the presence of HOD as identified by expert radiological assessment, the average slope of area in the left nucleus and the presence of bilateral HOD, respectively.

These results indicate that bilateral HOD is highly associated with development of PFS, furthermore the change in contrast and area in the left ION are of higher significance than the same changes occurring in the right ION. The highest scoring feature was the average slope of contrast in the left nucleus, implying that a high change in intensity of the left ION as seen on MRI is highly correlated to the presence of PFS. It is also evident that an increase in intensity in the left ION has a greater association with the development of PFS than an increase in size of the left ION. The diagnosis of HOD made by radiological assessment as a predictor of PFS is only around half as predictive as hyper intensity in the left ION as extracted from patient MRI. This is in keeping with the results of a recent study where damage to the right efferent cerebellar pathway, which communicates with the left ION, had a significant association with the development of PFS.¹⁰

The identical results yielded by SFS and SFFS, exhibited in Table 2, indicate feature 1 (Average slope of contrast in the left ION) and feature 11 (presence of unilateral HOD) as the most relevant features. Feature 1

was chosen as the most relevant throughout all four tests except the test using 4 neighbours, in which feature 11 was chosen as the most relevant feature. This further proves the relevance of an increase in intensity of the left ION in the development of PFS.

It should be noted that the search strategies used in this study are not optimal and are prone to local minima, with the exception of SFFS which makes an attempt at eliminating irrelevant features by carrying out a backward search in addition to the forward search. Notwithstanding this, the feature selection methods carried out in this study are ideal in a clinical scenario, more so than other methods, such as PCA, as the data retains interpretability after the feature selection techniques are applied.

The results yielded by the SVM classifier, shown in Figure 2, show an increase in classifier accuracy as the least efficient features were eliminated. The SVM classifier reached an accuracy of 89.29% when the only features included were those selected by the SFS and SFFS algorithms, that is, hyper-intensity in the left ION and the presence of unilateral HOD.

5. CONCLUSION

This study has identified intensity, or average slope of contrast, in the left ION as the most efficient feature that correlates with the development of PFS following tumour resection. Other efficient features found in this study include the presence of bilateral HOD, the presence of unilateral HOD and the average slope of area in the left ION. These features indicate that the presence of HOD, specifically in the left ION, highly correlates with the development of PFS following tumour resection surgery in the posterior fossa.

REFERENCES

1. Siffert, J., Poussaint, T., and Goumnerova, L., “Neurological dysfunction associated with postoperative cerebellar mutism,” *Journal of Neurooncology* **48**, 75–81 (2000).
2. Kirk, E., Howard, V, C., and Scott, C., “Description of posterior fossa syndrome in children after posterior fossa brain tumor surgery ,” *Journal of Pediatric Oncology Nursing* **12**(4), 181 – 187 (1995).
3. Patay, Z., Enterkin, J., Harreld, J., Yuan, Y., Ldobel, U., Rumboldt, Z., Khan, R., and Boop, F., “MR imaging evaluation of inferior olivary nuclei: Comparison of postoperative subjects with and without posterior fossa syndrome,” *American Journal of Neuroradiology* **35** (2013).
4. John, G., Kohavi, R., and Pfleger, K., “Irrelevant features and the subset selection problem,” *Proceedings of the International Conference on Machine Learning* **11**, 121–129 (1994).
5. Kojadinovic, I. and Wottka, T., “Comparison between a filter and a wrapper approach to variable subset selection in regression problems,” (2000).
6. Karamizadeh, S., Abdullah, S., Manaf, A., Zamani, M., and Hooman, A., “An overview of principal component analysis,” *Journal of Signal and Information Processing* **4**, 173–175 (2013).
7. Pohjalainen, J., Rasanen, O., and Kadioglu, S., “Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits,” *Computer Speech and Language* (2014).
8. Rsnen, O. and Pohjalainen, J., “Random subset feature selection in automatic recognition of developmental disorders, affective states, and level of conflict from speech,” *Proc. Interspeech’2013, Lyon, France* (2013).
9. Whitney, A. W., “A direct method of nonparametric measurement selection,” *IEEE Transactions on Computers* **C-20**, 1100–1103 (1971).
10. Law, N., Greenberg, M., Bouffet, E., Taylor, M., Laughlin, S., Strother, D., Fryer, C., McConnell, D., Hukin, J., Kaise, C., Wang, F., and Mabbott, D., “Clinical and neuroanatomical predictors of cerebellar mutism syndrome,” *Neuro-Oncology* **14**, 1294–1303 (2012).