

Automated classification of dopaminergic neurons in the rodent brain

Azadeh Alavi, Brenton Cavanagh, Gervase Tuxworth, Adrian Meedeniya, Alan Mackay-Sim, and Michael Blumenstein

Abstract—Accurate morphological characterization of the multiple neuronal classes of the brain would facilitate the elucidation of brain function and the functional changes that underlie neurological disorders such as Parkinson's diseases or Schizophrenia. Manual morphological analysis is very time-consuming and suffers from a lack of accuracy because some cell characteristics are not readily quantified. This paper presents an investigation in automating the classification of dopaminergic neurons located in the brainstem of the rodent, a region critical to the regulation of motor behaviour and is implicated in multiple neurological disorders including Parkinson's disease. Using a Carl Zeiss Axioimager Z1 microscope with Apotome, salient information was obtained from images of dopaminergic neurons using a structural feature extraction technique. A data set of 100 images of neurons was generated and a set of 17 features was used to describe their morphology. In order to identify differences between neurons, 2-dimensional and 3-dimensional image representations were analyzed. This paper compares the performance of three popular classification methods in bioimage classification (Support Vector Machines (SVMs), Back Propagation Neural Networks (BPNNs) and Multinomial Logistic Regression (MLR)), and the results show a significant difference between machine classification (with 97% accuracy) and human expert based classification (72% accuracy).

I. INTRODUCTION

THE dopaminergic neurons of the midbrain play a role in cognition, reward pathways and movement. Their dysfunction underlies movement disorders such as Parkinson's disease ([1], [2], [3]). The midbrain dopaminergic neurons project to spatially discrete regions and are functionally

All experiments were carried out in the National Centre for Adult Stem Cell Research, Eskitis Institute for Cell and Molecular Therapies, and the Institute for Integrated and Intelligent Systems (IIIS), Griffith University.

Azadeh Alavi, School of Information and Communication Technology, Griffith University, Gold Coast, Queensland 4222, Australia (phone: +61 7 555 28186; Facsimile +61 7 555 28066; e-mail: a.alavi@griffith.edu.au).

Brenton Cavanagh, National Centre for Adult Stem Cell Research, Griffith University, Nathan, QLD 4111 (e-mail: b.cavanagh@griffith.edu.au).

Gervase Tuxworth, School of Information and Communication Technology, Griffith University, Gold Coast, Queensland 4222, Australia (e-mail: g.tuxworth@griffith.edu.au).

Dr. Adrian Meedeniya, National Centre for Adult Stem Cell Research, Griffith University, Nathan, QLD, 4111 (phone: +61 7 373 54417 e-mail: a.meedeniya@griffith.edu.au).

Prof. Alan Mackay-Sim, National Centre for Adult Stem Cell Research, Griffith University, Nathan, QLD, 4111 (phone: +61 7 373 57563; e-mail: a.mackay-sim@griffith.edu.au).

A. Prof Michael Blumenstein, School of Information and Communication Technology, Griffith University, Gold Coast, Queensland 4222, Australia (phone: +61 7 555 28271; Facsimile +61 7 555 28066; e-mail: m.blumenstein@griffith.edu).

divergent, yet the anatomical features of their cell soma are similar. These midbrain dopaminergic neurons are located within the A8, A9 and A10 nuclei in the rodent. Of these, neurons of the A9 nucleus form the substantia nigra pars compacta (SNC), pars reticulata (SNR), and pars lateralis (SNL) [4]. The A10 forms the midline nucleus, the ventral tegmental area (VTA). Another group of dopaminergic inhibitory interneurons are found within a spatially discrete region of the rostral brain, namely the olfactory bulbs (OB).

Whilst the primary neurotransmitter of these neurons is dopamine, neurons within a region may show subtle differences in neurochemistry, e.g. dopaminergic neurons of the VTA may or may not express the calcium binding protein calbindin.

Importantly, whilst the distribution, function and neurochemistry of these dopaminergic neurons were defined within the rodent brain, the cellular morphology of these closely adjacent yet functionally divergent neurons remain to be defined. This paucity of data remains despite early evidence that the morphology of neurons is closely correlated to their function [5]. Further, subtle morphological changes are often correlated to the functional status of neurons, including changes associated with disease.

Neurobiologists classify microscopic images manually, but in almost all cases, the accuracy is not high, due to some cell characteristics being difficult to recognize through manual analysis. In addition, some of the characteristics may not seem to be effective in classification, but play an essential role as a component of an effective feature set. Manual classification is also time-consuming; hence it is costly to use experts for classification purposes. Machine learning techniques are commonly used to resolve classification problems in variety of fields. Recent studies applying machine learning techniques for neurobiology problems show their significant advantage. For example applying a BPNN in a gene selection problem shows a relatively high level of accuracy [12]. A modified MLR performed at 92% accuracy classifying normal/tumour cells in the colon [16]; and a modified SVM classified gene expression data of cancer tissue (gene expression data contains a high level of noise) with a high level of accuracy [17]. Classification may be done very quickly and accurately using machine learning techniques, thus providing a significant advantage to neurobiological research.

Thus a method for unbiased, rapid, morphological analysis of neurons is needed. Employing feature extraction and machine learning techniques, this paper defines the morphology of three types of dopaminergic neurons of the basal ganglia and discriminates them from each other based on their shape.

II. NEURONAL MORPHOLOGY

All neurons are comprised of four major components: a cell body, axon, dendrites and synapses [6] (see Fig. 1). However, different types of neurons perform different functions and have subtle differences in their morphological properties. The number and form of the neuron's processes, in addition to its shape, are considered to be the best characteristics of neurons for classification purposes.

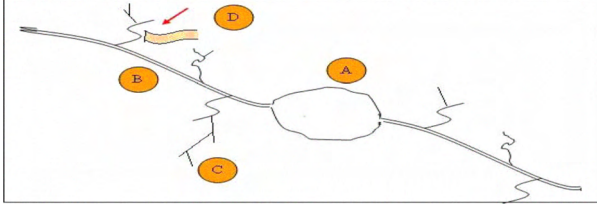


Fig. 1. Image of a neuron: A) cell body B) axon C) dendrite D) synapse

A. Dopaminergic neurons

Dopaminergic neurons were characterized by their expression of the enzyme Tyrosine Hydroxylase (TH) using immunochemical detection. Calbindin is a calcium-binding protein that is found in some populations of TH-positive cells but not others; hence cells were also characterized by their immunoreactivity to Calbindin. The sections were also stained with DAPI, which intercalates into DNA and so stains all cell nuclei. This provides important information for identifying brain regions and orienting the brain section (Fig. 2).

B. Areas of interest

We analysed dopaminergic neurons in three brain regions: SNC and VTA, in the midbrain, there are two areas that contain the largest dopaminergic populations; and the olfactory bulb (OB), located rostrally in the brain, which is also rich in dopaminergic neurons [8].

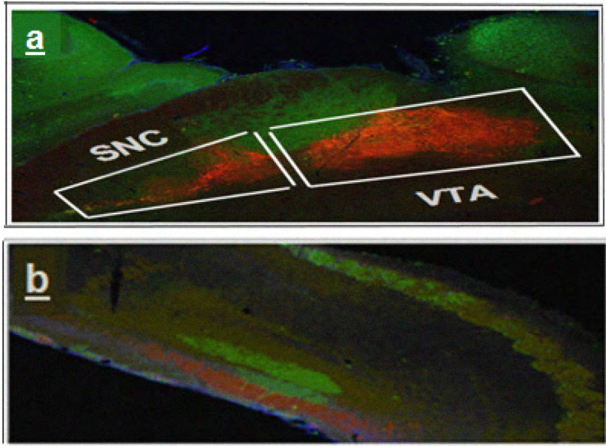


Fig. 2. Multiple labelling immunofluorescence images of a section of rat brain (a) an image from the midbrain containing the VTA and SNC showing tyrosine hydroxylase immunoreactive neurons (red) and calbindin immunoreactive neurons (green). (b) an image of the OB showing the distribution of the same markers.

In order to distinguish different types of dopaminergic neurons, sections were also stained to find cells expressing Calbindin.

III. AUTOMATIC CLASSIFICATION TECHNIQUE

Machine learning techniques were successfully used in a wide range of classification problems. This investigation will focus on three popular machine learning techniques (Support Vector Machines (SVMs), the Backpropagation Neural Network (BPNN) and Multinomial Logistic Regression (MLR)) comparing their performance on classifying dopaminergic neurons located at VTA, SNC and OB (from a custom dataset).

A. Support Vector Machines

SVMs are supervised machine learning techniques that have the advantage of being underpinned by a very well developed learning theory (statistical learning theory). SVMs map the input sample to high dimensional space, and seek a "separating hyper-plane" in this space.

Consider (x_i, y_i) as a training set, where $1 \leq i \leq N$, and where N is the number of training pairs. Each sample must conform to the C ("regularization parameter") that controls the trade off between the complexity of the hypothesis space used and the empirical error. SVMs are considered to be a robust classification method, and were used widely in image classification problems [9].

The space used by SVMs is a set of hyper-planes through the kernel k in the feature space.

$\{f : \|f\|_k^2 < \infty\}$ where K is the kernel that identifies a Reproducing Kernel Hilbert Space (RKHS). In the above formula, $\|f\|_k^2$ defines the RKHS norm of the function.

To minimize the trade off between the complexity of the hypothesis space and the empirical error, SVMs classification follows the formula:

$$\min_f \|f\|_k^2 + c \sum_{i=1}^l |1 - y_i f(x_i)|_+ \quad (1)$$

B. Backpropagation Neural Network

BPNN is considered a powerful classification method and is a popular Artificial Neural Network (ANN) classification technique. The BP algorithm is a non-parametric estimator [10]. It is a supervised algorithm and can be applied to multi-layer networks. A sigmoidal activation function is most commonly used in the calculation process, because of its ability to successfully handle both small and large signals with automatic gain control.

$$\text{Out} = f(\text{NET}) = 1 / (1 + e^{-\text{Net}}) \text{ when: } \text{NET} = \sum X_i * W_i \quad (2)$$

C. Multinomial Logistic regression (MLR)

MLR is a popular discriminative probabilistic classification model that performs particularly well in bio-image classification problems [11], [12]. Logistic Regression is considered to be one of the best probabilistic classifiers. It

measures both first best and log loss classification accuracy through a number of steps.

$$y_i = \begin{cases} 1, & \text{If } x \text{ belongs to class } i \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Where $i \in \{2, \dots, m\}$, and m represents the number of output classes.

If $m=2$ (binary problems) the technique is referred to as Logistic Regression, while when $m > 2$ the technique is known as MLR.

Using MLR, the probability that x belongs to class i is

$$P(y_i = 1|x; w) = \frac{\exp(w_i^T x)}{\sum_{j=1}^m \exp(w_j^T x)} \quad (4)$$

where $\sum_{i=1}^m p(y_i = 1|x; w) = 1$ (as a result of normalization).

In the above formula, P is the predicting variable, w_i denotes the weight vector, and superscript T is the vector transpose [13].

IV. DOPAMINERGIC NEURON CLASSIFICATION

To perform classification of dopaminergic neurons located in the VTA, SNC and OB, several steps were completed as described in Fig. 3.

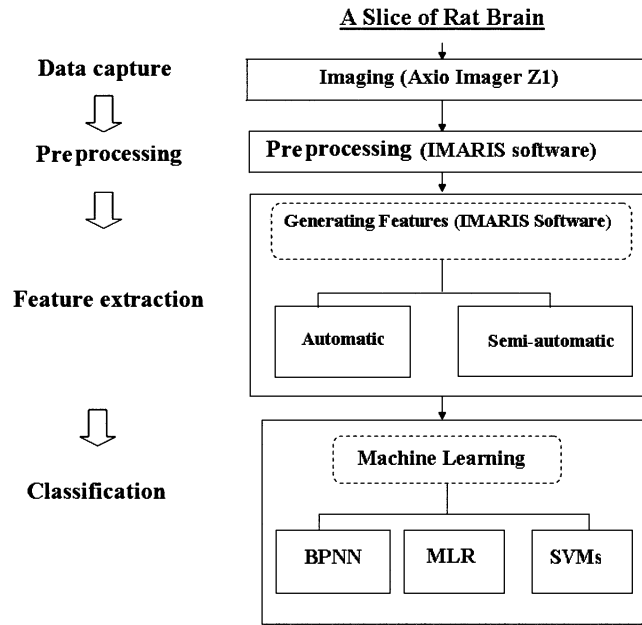


Fig. 3. Process followed to classify dopaminergic neurons from a dataset

A. Data capture

The data captured for this research was the result of analysing preprocessed images. Microscopic images were taken from slices of rat brain prepared for immunochemistry

and photographed using a Carl Zeiss Axio Imager Z1 microscope equipped with an ApoTome, an attachment for accurate optical sectioning, to provide z-stacks of images for 3D reconstruction of the cells of interest.

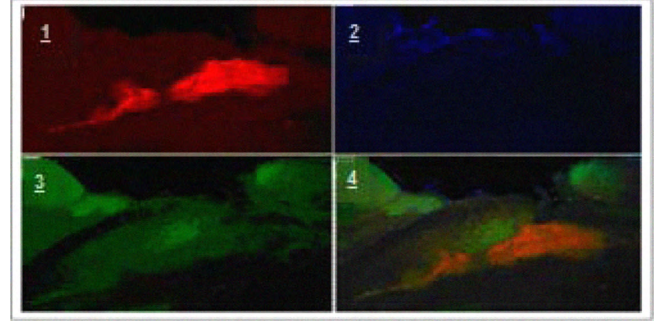


Fig. 4. Colocalisation of the key protein markers with (1) immunoreactivity to tyrosine hydroxylase (2) all cell nuclei (3) immunoreactivity to Calbindin and (4) the merged images (combining images 1, 2 and 3), as created by AxioVision software.

B. Processing and analysing images

To analyse the images, IMARIS software was used to undertake semi-automated preprocessing and segmentation. Before commencing segmentation, some preprocessing was required to eliminate background noise and remove any other detected objects in the image. This was done by applying a Gaussian filter and setting a suitable threshold. Finding the threshold is a semi-automated process, however IMARIS provided an automatic preview that showed how the image will appear if the selected threshold is applied.

Segmentation involved isolating the chosen neuron from other objects in the image that was being processed using IMARIS object detection. The IMARIS software was also used to identify the centre of the cell body and the neuronal filaments.

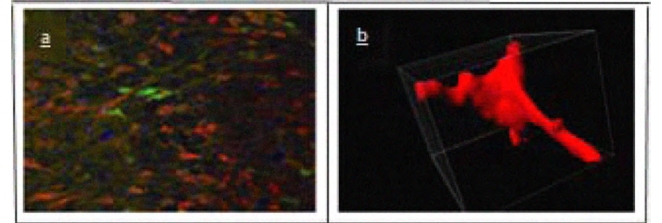


Fig. 5. (a) The unfiltered image of the VTA in a rat brain (b) a VTA neuron after performing preprocessing and segmentation

C. Feature extraction

We investigated the set of features that best describes the morphological differences between dopaminergic neurons located in the three different locations of the brain, for the purpose of classification.

Two sets of features were examined in this research:

- Automatically generated features (using IMARIS)
- Semi-automatically generated features (using IMARIS)

To morphologically analyse the segmented neurons, 16 features were used; 10 of these features are obtained as a result of IMARIS automatic image analysis, and 6 of them are obtained semi-automatically. These features were selected based on statistical information provided by IMARIS from previous work and empirical analysis.

1. Automatically generated features

When segmentation is complete and the desired object was separated, IMARIS automatically generates statistical information about the object (Table1).

TABLE I
AUTOMATICALLY GENERATED FEATURES

ID	FEATURE NAME	Description
1	Area	The sum of the area of all the surfaces
2	Ellipticity	The thickness of the calculated ellipsoid
3	Length	Total length of the neuron
4	No. Branch Points	The number of dendritic branching points
5	No. Edges	The number of object edges
6	No. End Segments	The number of dendrial endings
7	No. Segments	The number of paths between branch points
8	No. Vertices	The total number of vertices
9	Sphericity	A rating that describes how spherical an object is
10	Volume	The total volume

In the following sections each feature will be referred to by its 'ID'.

2. Semi-automatically generated features

Images were analysed semi-automatically using IMARIS software, demonstrating the shape of the objects (Table2).

TABLE 2
SEMI-AUTOMATICALLY GENERATED FEATURES

ID	FEATURE NAME	Description
11	R1	Fig .6
12	R2	Fig .6
13	Depth	Total thickness of neuron
14	Main angle1	Fig .7 and Fig .8
15	Main angle2	Fig .7 and Fig .8
16	Branch angle	Fig. 7 and Fig. 8

In the following sections each feature will be referred to by its 'ID'.

To be able to get R1 and R2 from the object, the widest 2D image was selected. If a segmented 3D image is made up of a stack of n 2D images, the widest image is the closest to the middle image ($\frac{n}{2}$). This 2D image will provide two features describing the width of the object (Fig. 6).

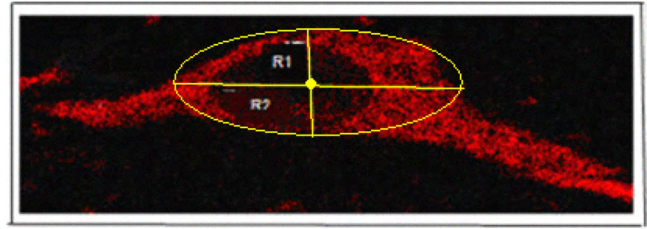


Fig. 6. The 2D image with the widest cell soma in the plane of the nucleus, which is a component (slice) of the final 3D image

Another geometric feature is the depth (thickness) of the object. In conjunction with R1 and R2, it provides an approximate description of the object's shape.

The image below (Fig. 7) clearly demonstrates the differences in the shape of these three dopaminergic neurons.

Although not all of the neurons from these three classes are this readily distinguishable, they follow the same basic morphology.

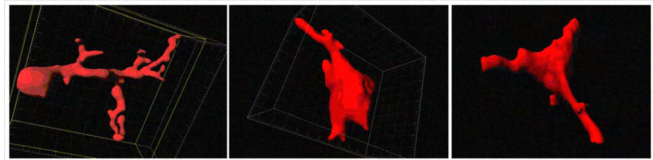


Fig. 7. Left to right: Olfactory Bulbs, SNC and VTA neurons

To describe this difference, the angle between the axon and two main dendrites are measured (Fig. 8), together with the angle between the two main dendrites (Fig. 8).

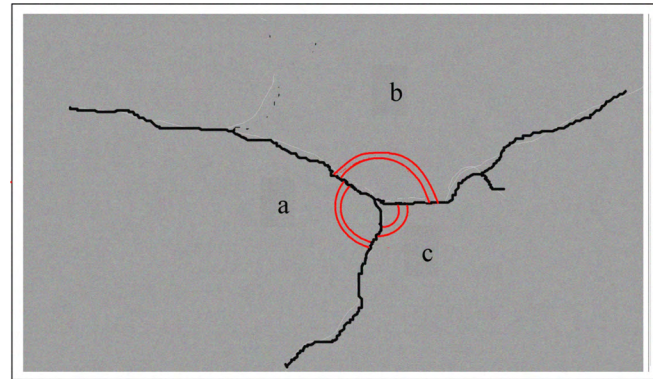


Fig. 8. The neural filament tracing of a VTA neuron. 'a' and 'b' are the angles between the axon and the main dendrites c. the angle of two main dendrites

3. Protein marker

The two types of Dopaminergic neurons within the VTA are readily distinguishable by identification of marker proteins using specific antibodies and epifluorescence microscopy. As noted above, we analysed two types of dopaminergic TH-immunoreactive neurons of the VTA, those with, and those without, Calbindin immunoreactivity (VTA.2).

D. Classification

Three popular classification techniques (BPNN, SVMs, MLR), were selected for comparison. For each of these techniques, all relevant parameters and settings were examined to ensure their maximum efficiency. In order to select the best feature vector for each classifier, attribute selection was performed based on examining different values for this parameter.

In this research, WEKA software was used to train the SVMs, because it implements John Platt's sequential minimal optimization algorithm for training support vector classifiers [15]. It has an advantage of replacing all missing values globally. It also transforms nominal attributes into binary ones, and normalizes all attributes by default [15].

WEKA software was chosen to train the MLR because it improves efficiency by applying a ridge estimator. In addition, original Logistic Regression does not deal with instance weights; however WEKA modifies the algorithm to enable it to handle the instance weights [15]. WEKA software was also selected to train the BPNN [15].

V. RESULTS AND DISCUSSION

The three popular classification techniques (BPNN, MLR, and SVMs) were compared, using a set of 100 images. 5-fold cross validation and 10-fold cross validation were used to randomly determine training and testing sets.

One aim of these experiments was to compare the automated classification techniques with each other, and with a human expert, in their ability to use morphological differences to discriminate between dopaminergic neurons in three regions of the brain affected by Parkinson's disease (VTA, SNC and OB).

A. Results

To analyse the morphological differences between these dopaminergic neurons (located in the VTA, SNC and OB), several training sets were used.

The first two sets of experiments focused on classifying dopaminergic neurons located in these three areas (Tables 3-5).

Then, as the OB neurons can be classified with 100% accuracy, the remaining experiments analysed morphological differences between VTA1, VTA2 and SNC (Tables 6-11).

In addition, to analyse the importance of semi-automatically generated features, the classifiers were trained twice; once only using automatically generated features (.1F), and secondly, integrating semi-automatically generated features (.2F).

All three classifiers were trained for each experiment comparing their performance, and all parameters were varied for each to ensure optimum settings.

1. SNC, VTA and OB

The first component of this experiment is to classify the dopaminergic neurons located in the VTA, SNC and OB, where both types of VTA neurons are marked as VTA, and the second component is when VTA.1 and VTA.2 are

treated separately. Tables 3, 4 and 5 summarise the results of this study using the BPNN, MLR and SVMs.

Each table details the following information: ID (experiment's unique identifier), Features (ID of the features used, detailed in section IV part C), TM (test mode (f-cv: fold-cross validation) AC (Accuracy), HN (Number of nodes in the hidden layer), LR (Learning Rate), M (Momentum), C (Complexity parameter) and \mathcal{E} (the epsilon for round-off error), with respect to the trained classifier.

TABLE 3
BPNN CLASSIFICATION RESULTS

ID	Features	LR	M	HN	TM	AC (%)
BPNN.1F.1	1,2,3,10	0.19	0.2	3	5 f-cv	86
BPNN.2F.1	1,2,3,8,10, 12,13,14	0.18	0.2	6	10 f-cv	88
BPNN.2F.2	1,2,3,8, 10, 12,13,14	0.17	0.15	6	10 f-cv	91

When training BPNN as a classifier, **BPNN.1F.1** indicates when VTA.1 and VTA.2 are both marked as VTA, and only automatically generated features are used. **BPNN.2F.1** indicates the same experiment but including semi-automatically generated features; **BPNN.2F.2** indicates when VTA.1 and VTA.2 are treated as separate, and both types of features are used.

TABLE 4
MLR CLASSIFICATION RESULTS

ID	Features	TM	AC (%)
MLR.1F.1	1,2,3,10	10 f-c v	85
MLR.2F.1	1,2,3,8,9,10, 12,13,14	10 f-c v	88
MLR.2F.2	1,2,8,10,13, 14,15	10 f-c v	90

When training MLR as a classifier, **MLR.1F.1** indicates when VTA.1 and VTA.2 are both marked as VTA, and only automatically generated features are used. **MLR.2F.1** indicates the same experiment but including semi-automatically generated features; **MLR.2F.2** indicates when VTA.1 and VTA.2 are treated as separate, and both types of features are used.

TABLE 5
SVM CLASSIFICATION RESULTS

ID	Features	C	\mathcal{E}	TM	AC (%)
SVM.1F.1	1,2,3,10	1.00E+ 03	1.00E-12	5 f-cv	83
SVM.2F.1	1,2,3,10, 11,12,13, 14	1.00E+ 03	1.00E-12	5 f-cv	86
SVM.2F.2	1,3,9,10, 13,14,15	1.00E+ 02	1.00E-12	10 f-cv	90

When training SVMs, **SVM.1F.1** indicates when VTA.1 and VTA.2 are both marked as VTA, and only automatically generated features are used. **SVM.2F.1** indicates the same experiment but including semi-automatically generated features; **SVM.2F.2** indicates when VTA.1 and VTA.2 are treated as separate, and both types of features are used.

The results demonstrate that using semi-automatically generated features increases the average accuracy by several percent. It also shows that separating VTA.1 and VTA.2 increases the average accuracy. BPNN provides the highest accuracy for these experiments.

2. VTA1 vs. VTA2

To find out how morphologically different VTA.1 and VTA.2 are, BPNN, MLR and SVMs were applied to classify these based on their morphological components. The tables below summarise the classification results following training of BPNN, MLR and SVMs.

TABLE 6
BPNN CLASSIFICATION RESULTS

ID	Features	LR	M	HN	TM	AC (%)
BPNN.1F.3	1,2,3,8,10	0.3	0.2	3	5 f-cv	72
BPNN.2F.3	8,10,12,13	0.17	0.2	3	5 f-cv	74

When using BPNN as a classifier, **BPNN.1F.3** indicates when only automatically generated features are used whereas **BPNN.2F.3** indicates when semi-automatically generated features are also included.

TABLE 7
MLR CLASSIFICATION RESULTS

ID	Features	TM	AC (%)
MLR.1F.3	1,2,3,8,10	5 f-c v	74
MLR.2F.3	1,8,10,13,14	5 f-c v	74

When using MLR as a classifier, **MLR.1F.3** indicates when only automatically generated features are used whereas **MLR.2F.3** indicates when semi-automatically generated features are also included.

TABLE 8
SVM CLASSIFICATION RESULTS

ID	Features	C	\mathcal{E}	TM	AC (%)
SVM.1F.3	1,2,3,8,10	1.00E+03	1.00E-12	5 f-cv	72
SVM.2F.3	1,8,10,12,13	1.00E+01	1.00E-12	5 f-cv	74

When using SVMs, **SVM.1F.3** indicates when only automatically generated features are used whereas **SVM.2F.3** indicates when semi-automatically generated features are also included.

The results show that VTA.1 and VTA.2 cannot be accurately distinguished from each other because of morphological overlaps. For these experiments, adding semi-automatically generated features improved accuracy but did not achieve optimal results. All classification techniques performed at a similar accuracy, obtaining an average of 74%.

3. VTA1 vs. SNC and VTA2 vs. SNC

The confusion matrix of experiment 2 shows that by separating VTA.1 and VTA.2, classification of VTA.1, SNC and OB reaches the average accuracy of 92%. While the confusion matrix of experiment 1 shows that the OB could be classified at 100% accuracy; therefore OB neurons were

removed from the dataset for the next experiment aiming at analysing the morphological difference between SNC and VTA dopaminergic neurons.

A manual analysis of the images shows that VTA.1's neurons have a morphological overlap with VTA.2 and SNC neurons. To study the morphological differences between VTA.1 and SNC, and VTA.2 and SNC, new experiments were formulated. The tables below summarise the results.

TABLE 9
BPNN CLASSIFICATION RESULTS

ID	Features	LR	M	HN	TM	AC (%)
BPNN.1F.VTA2	1,2,3,10	0.3	0.2	3	5 f-c v	84
BPNN.2F.VTA2	1,2,3,10,11,12,13,14	0.1	0.2	5	5 f-c v	95
BPNN.1F.VTA1	1,2,3,10	0.3	0.3	3	5 f-c v	85
BPNN.2F.VTA1	1,3,8,9,10,13,14	0.1	0.2	5	5 f-c v	87

When using BPNN as a classifier, **BPNN.1F** indicates when only automatically generated features are used; **BPNN.2F** indicates when semi-automatically generated features are also included.

TABLE 10
MLR CLASSIFICATION RESULTS

ID	Features	TM	AC (%)
MLR.1F.VTA2	1,2,3,10	5 f-c v	84
MLR.2F.VTA2	1,10,13,14	5 f-c v	89
MLR.1F.VTA1	1,2,3,10	5 f-c v	86
MLR.2F.VTA1	1,3,10,12,13	5 f-c v	86

When using MLR as a classifier, **MLR.1F** indicates when only automatically generated features are used; **MLR.2F** indicates when semi-automatically generated features are also included.

TABLE 11
SVM CLASSIFICATION RESULTS

ID	Features	C	\mathcal{E}	γ	TM	AC (%)
SVM.1F.VTA2	1,10	1.00 E+03	1.00 E-12	1.00 E-02	5 f-c v	85
SVM.2F.VTA2	1,2,3,10,11,12,13,14	1.00 E+03	1.00 E-12	1.00 E-02	5 f-c v	93
SVM.1F.VTA1	1,2,3,10	1.00 E+03	1.00 E-12	1.00 E-02	5 f-c v	83
SVM.2F.VTA1	1,3,10,12,13	1.00 E+02	1.00 E-12	1.00 E-02	5 f-c v	85

When using SVMs as a classifier, **SVM.1F** indicates when only automatically generated features are used; **SVM.2F** indicates when semi-automatically generated features are also included.

The results demonstrate that VTA.2 and SNC could be distinguished with a 95% accuracy, while VTA.1 and SNC can only be distinguished with a 87% accuracy; it confirms the results of manual analyses showing the morphological overlap of VTA.1's and SNC's dopaminergic neurons.

The results also confirm the positive impact of including semi-automatically generated features.

B. Discussion

1. Comparison of classification techniques

Fig. 9 illustrates that classification of dopaminergic neurons in the VTA, SNC and OB was improved when the VTA.1 and VTA.2 categories were separated. The figure shows that before separating VTA.1 from VTA.2, the neurons could be identified with an average accuracy of 88%, whereas after separating them, the accuracy increased to 91%. Fig. 9 also shows that the accuracy was similar between different machine learning techniques.

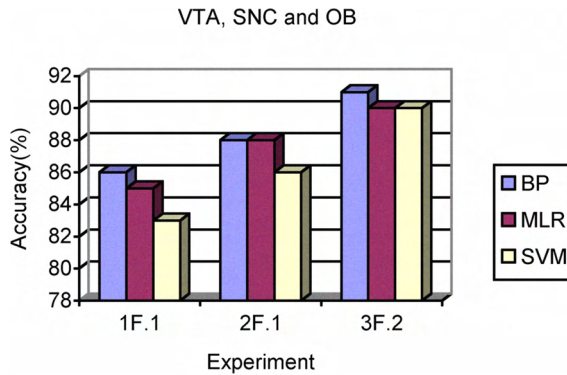


Fig. 9. The bar chart Compares the results of applying each type of machine learning technique. 1F.1 indicates when VTA.1 and VTA.2 are both marked as VTA, and only automatically generated features are used. 2F.1 represents the same experiment but includes semi-automatically generated features. 3F.2 refers to when VTA.1 and VTA.2 are separated, and both types of features are used.

Fig. 9 also shows no considerable variation in accuracy when different machine learning techniques are employed.

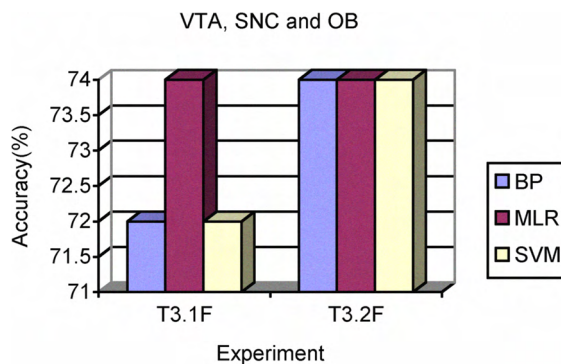


Fig. 10. The bar chart compares the results upon applying each machine learning technique to distinguish VTA.1 and VTA.2; T3.1F indicates when only automatically-generated features are used; T3.2F indicates when semi-automatically generate features are also included.

Fig. 10 shows that VTA.1 and VTA.2 were morphologically similar and difficult to separate. Comparing T3.1F (when only automatically generated features were used) and T3.2F (when semi-automatically generated features are also included) confirms that including semi-automatically generated features, had a positive effect on classification accuracy.

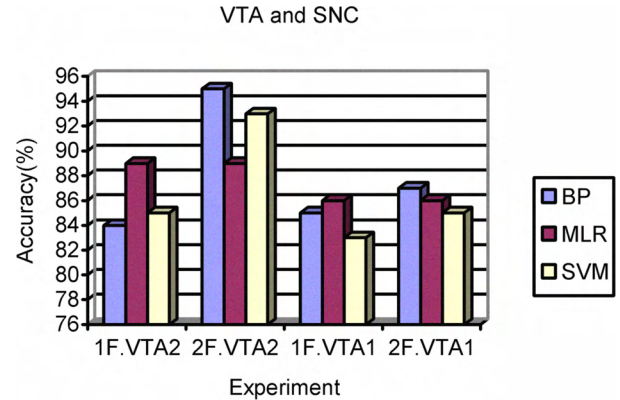


Fig. 11. The bar chart compares the results upon separating VTA.1 and VTA2, applying different machine learning techniques. 1F indicates when only automatically generated features are used. 2F represents the same experiment but also includes semi-automatically generated features.

Fig. 11 shows that VTA.2 and SNC dopaminergic neurons were distinguishable with an accuracy of 95%; while VTA.1 and SNC were only identified with an accuracy of 87%.

Fig. 11 demonstrates that consistency in accuracy was observed when employing different classification techniques.

2. Comparing the results from machine learning techniques with a cell expert

The performance of the machine learning techniques was compared with the classification accuracy by a human expert using a t-test. Among the machine learning techniques, BPNN was chosen as the most accurate when the combination of automatically and semi-automatically generated features was used (Figs. 9 and 11). Using these parameters, BPNN classified the cells with a 91% accuracy compared to 73% accuracy for the human expert. This difference was statistically significant ($p < 0.025$, t-test). Therefore, on average, BPNN outperformed a human expert with a 97.5% confidence interval.

3. Comparing the results with that of previous work

Previously, 72% was the best accuracy achieved for classifying dopaminergic neurons located at VTA and SNC (using a data set of 18 images) [14]. That method used only automatically generated features (using the IMARIS software). The present study shows that by adding semi-automatically generated features the accuracy improved to 88%, and by separating VTA.1 from VTA.2, the performance increased to 91%. The reason for these improvements in accuracy can be summarized as follows:

- Semi-automatically generated features are effective in providing better performance
- VTA.1 overlaps morphologically with VTA2 and SNC; hence identifying VTA.1 and VTA.2 separately, improves the classification performance

VI. CONCLUSION

The primary objective of this research was to develop a method for automating the classification of dopaminergic neurons in several regions of the brain using morphological differences (employing 16 features).

The results obtained for dopaminergic neuron classification demonstrates a significant difference between the machine classification result (91% accuracy with BPNN) and the human expert result (73% accuracy). It also illustrates the efficacy of adding semi-automatically generated features with respect to classification performance. The data from the two types of dopaminergic neurons located at VTA demonstrated the impact of cellular homogeneity on the classification rate (when classifying dopaminergic neurons located at VTA versus the SNC and OB). A significant morphological difference between the two types of VTA dopaminergic neurons was not seen. The results indicate that VTA.1 (whose cells only express TH) overlaps morphologically with SNC and VTA.2 (whose cells also express Calbindin). Hence by separating VTA.1 from VTA.2 the classification accuracy was improved.

The accuracy of the data and the sensitivity of the machine learning techniques in resolving subtle cellular morphometry suggest its potential for multiple applications in future studies. This would include differentiation between specific developmental stages of neurons and their responses to the growth environment, and defining cellular changes within normal and diseased brains.

REFERENCES

- [1] K. Hassler, "Zur pathologischen anatomie des senilen und des parkinsonistischen tremor". J. Psychol. Neurol, pp. 13-55, vol. 4, 1939.
- [2] A. Carlsson, "Antipsychotic drugs, neurotransmitters and schizophrenia", Am. J. Psychiatry, vol. 135, pp: 164-173, 1978.
- [3] S. H. Snyder, "Catecholamines in the brain as mediators of amphetamine psychosis", Arch. Gen. Psychiat, vol. 27, pp.169-179, 1972.
- [4] A. Dahlesstrom, K. Flux, "Evidence for the existence of monoamine-containing neurons in the central nervous system". I. Demonstration of monoamines in the cell bodies of brain stem neurons. Acta physiol. scand. vol. 62, Suppl. 232, pp.1-55, 1964 a.
- [5] E.L. Nelson, C.-L. Liang, C.M. Sinton, D.C. German, "Midbrain dopaminergic neurons in the mouse: Computer-assisted mapping", the journal of comparative Neurology, vol. 369, issue. 3, pp. 232: 1-55, 1998.
- [6] J. Nolte, *the Human Brain an Introduction to Its Function Anatomy*. London: Mosby, 1988, pp. 120-125.
- [7] K. Chung and T. Dawson, "Parkin and Hsp70 Sacked by BAG5", Neuron, vol. 44, No. 6, pp. 899-904, 2004.
- [8] I. Ina Brünig, M. Sommer, H. Hatt and J. Bormann, "Dopamine receptor subtypes modulate olfactory bulb γ -aminobutyric acid type A receptors", The National Academy of Sciences, vol 95. No.5, pp. 2456-2460, 1999.

- [9] S. Haykin, *Neural Network: a comprehensive foundation*. New Jersey: Prentic Hall, 1999, pp.10-20.
- [10] P. D. Wasserman, *Neural Computing Theory and Practice*. New York: Chapman & Hall, 1989.
- [11] M. Boland, R. Murphy, "A neural network classifier capable of recognizing a patern of all majors subcellular structures in fluorescence microscope images of Hela cell", Bioinformatics, vol.17, No. 13, pp.1213-1223, 2001.
- [12] J. Zhaung, S. Liu, Yue Wang, "Gene association study with SVMs, MLP and cross-validation for the diagnosis of disease.", Progress in Natural Science, vol. 18, Issue.6, 2008.
- [13] Le Cessie and J. C. Van Houwelingen, "Ridge Estimators in Logistic Regression" Applied Statistics, Vol. 41, No. 1, pp. 191-201, 1992.
- [14] G. Tuxworth, "Automatic Neural Cell Classification", Houners thesis, Griffith University, 2007.
- [15] I. H. Witten, E. Frank, *Data Mining: Practical machine learning tools and techniques*, San Francisco: Morgan Kaufmann, 2005.
- [16] Guo-Liang, Tiana, M. Tangb, H. Fanga and M. Tana "Efficient methods for estimating constrained parameters with applications to regularized (lasso) logistic regression", Computational Statistics & Data Analysis, vol. 52, Issue 7, 2008.
- [17] Z. Chen, J. Li and Liwei Wei "A multiple Kernel support vector machine scheme for feature extraction from gene expression data of cancer tissue", Artificial Intelligence in Medicine, vol. 41, Issue. 2, pp.161-175, 2007.