

# Cocaine-induced preference conditioning: a machine vision perspective

Received: date / Accepted: date

**Abstract** Existing work on drug-induced synaptic changes has shown that the expression of perineuronal nets (PNNs) at the cerebellar cortex can be regulated by cocaine-related memory. However, these studies on animals have mostly relied on limited manually-driven procedures, and lack some more rigorous statistical approaches and more automated techniques. In this work, established methods from computer vision and machine learning are considered to build stronger evidence of those previous findings. To that end, an image descriptor is designed to characterize PNNs images; unsupervised learning (clustering) is used to automatically find distinctive patterns of PNNs; and supervised learning (classification) is adopted for predicting the experiment group of the mice from their PNN images.

Experts in neurobiology, who were not aware of the underlying computational procedures, were asked to describe the patterns emerging from the automatically found clusters, and their descriptions were found to align surprisingly well with the two types of PNN images revealed from previous studies, namely strong and weak PNNs. Furthermore, when the set of PNN images corresponding to every mice in the saline (control) group and the conditioned (experimental) group were characterized using a bag-of-words representation, and subject to supervised learning (saline vs conditioned mice), the high classification results suggest the ability of the proposed representation and procedures in recognizing these groups.

Therefore, despite the limited size of the dataset (1,032 PNN images of 6 saline and 6 conditioned mice), the results support existing evidence on the drug-related brain plasticity, while providing higher objectivity.

**Keywords** cerebellum · perineuronal nets · drug-related memory · computer vision · machine learning · unsupervised learning · supervised learning

## 1 Introduction

Addiction is characterized by an aberrant overconsolidation of drug-cue associative memory (Hyman et al., 2006). Even after protracted abstinence, drug-related cues have the ability to trigger craving and relapse (Shaham et al., 2003). Importantly, neuroimaging studies of cue-reactivity in drug addicts have consistently shown cerebellar activations when these memories are reactivated by presenting drug-related cues (Moulton et al., 2014, Moreno-Rius and Miquel, 2017). Thus, it seems that the cerebellum may represent drug-cue memories and thereby contribute to the persistent risks of relapse to addiction.

Interestingly, several external factors like drugs of abuse might control structural remodelling of brain circuitry by modulating the activity of regulatory molecules that restrict neuronal plasticity in order to stabilize circuits (Foscarin et al., 2011). These plasticity inhibitory mechanisms take place in a cartilage-like structure called Perineuronal net (PNN) consisting of molecules of extracellular matrix that enwraps the perikaryon of neurons (Brückner et al., 1993, Grimpe and Silver, 2002, Carulli et al., 2005, 2006, Frischknecht et al., 2009, Carulli et al., 2013). The PNN structure is quite resistant to intracellular degradative system (Toyama and Hetzer, 2013), leading to the suggestion thereby that PNNs might act as one of the stabilization mechanisms for long-lasting drug-induced synaptic modifications (Wright and Harding, 2009, Van den Oever et al., 2010, Xue et al., 2014, Slaker

et al., 2015, Vazquez-Sanroman et al., 2015a,b, Slaker et al., 2016, Sorg et al., 2016, Blacktop et al., 2017, Vazquez-Sanroman et al., 2017).

It is now clear that the reticular structure of PNNs could be modified by drug intake experiences resulting in conditioned memories (Xue et al., 2014, Slaker et al., 2015, 2016, Blacktop et al., 2017, Carbo-Gas et al., 2017). Recently, it has been described that cocaine-induced conditioned learning increased neuronal activity and expression of PNNs surrounding Golgi interneurons in the dorsal cerebellar cortex (Carbo-Gas et al., 2017). A fully condensed PNN around Golgi interneurons could “stamp in” the synaptic arrangement representing drug-cue association (Sorg et al., 2016). Golgi neurons are key components for the modulation of activity and plasticity in the cerebellum (D’Angelo and De Zeeuw, 2009, Carta et al., 2004); specifically, they are critical to the synchronization of granule cell clusters (Armano et al., 2000).

Although some statistical evidence of PNN remodeling has been found (Slaker et al., 2016), the fields of computer vision and machine learning have the potential for bringing more systematic and automatic forms of evidence. To that end, the goal of our study is to characterize PNN images of mice cerebella so that structural image information can be automatically inferred and used to recognize whether the PNNs of a mouse correspond to an animal that has acquired cocaine-induced preference memories for olfactory cues. This study is relevant since automatic recognition would back the hypothesis that PNN structure is stable under physiological conditions and could play a role in long-lasting drug-induced synaptic modifications.

Computer vision and machine learning (“machine vision” for short from now on) has been widely applied to the challenging fields of biological and medical imaging. As a brief overview of this vast field, some works closer to neuroimaging can be mentioned. Neuron reconstruction from images has been studied (Santamaría-Pang et al., 2015). Then, neuron reconstructions can be compared (Gillette et al., 2011), characterized (Wan et al., 2015), or segmented (Zhang et al., 2018). Shape and appearance features are explored for classifying dendritic spine into one of three possible classes (Ghani et al., 2017). Nevertheless, to the best of our knowledge, machine vision has not been previously studied in the context of neither PNN images in general nor in the particular problem addressed in this work.

To analyse our problem from the machine vision perspective, an image descriptor was designed for its potential ability to discriminate between types of PNNs. First, unsupervised learning is used to dis-

cover image patterns corresponding to different types of PNNs. Then, the resulting clusters are used to build a bag-of-words (BoW) model to characterize individually mice from their corresponding set of PNN images. Finally, the BoWs are used for supervisedly classifying mice into the established categories according to the experimental set-up for cocaine-induced preferences in mice.

## 2 Methodology

The methodology regarding mice conditioning experiments (Sects. 2.1–2.2) is first provided as background information to make the paper more self-contained. Then, the specific methodology corresponding to the computational aspects of the present work are detailed (Sects. 2.3–2.7).

### 2.1 Cocaine-induced conditioned preference protocol

Five-week-old Swiss male mice were purchased from Janvier (ST Berthevin Cedex, France) and housed in our animal facilities (Jaume I University, Spain). Experimental procedures started at the age of 7 weeks. Handling was performed daily for 5 minutes before the experiment began. Conditioning was developed in an opaque, oblong corridor that included two lateral black chambers ( $30 \times 15 \times 20$  cm) located on opposite sides. Two equally preferred olfactory stimuli (papaya and strawberry) were used. One of the odours acted as cocaine-paired stimulus (CS+) and was associated with cocaine hydrochloride (20 mg/kg) (Alcaliber S.A., Madrid, Spain), dissolved in a 0.9% saline solution and administered intraperitoneally (IP). The other one acted as saline-paired stimulus (CS-) and was associated with saline injections as the control vehicle. Olfactory stimuli were counterbalanced as CS+ and CS-. Four drops of papaya or strawberry scents were presented inside a steel ball at the end of the each conditioning box. During the training session, the animals remained confined in one of the lateral chambers, and access to the other side was blocked by a panel. Each pairing session lasted for 15 minutes.

A total of eight cocaine-cue paired sessions were conducted and the left and right locations in the corridor were also counterbalanced among the animals. Preference for the cocaine-related cue was evaluated 48 h after the last cocaine administration in a 30 min drug-free test in which the CS+ and CS- were presented simultaneously on both sides of the corridor in an opposite location to that in the training. Thus, for the first 10 min of the test session, the mice were allowed

to explore the new location of the odour cues, and consequently this period was not included in the analysis. All the test sessions were videotaped and scored by a blind observer. The preference score was calculated as  $100 \cdot \frac{t_+}{t_+ + t_-}$ , where  $t_+$  ( $t_-$ ) is the time (in seconds) spent in the CS+ (CS-) part of the corridor (Carbo-Gas et al., 2014a,b, 2017).

Additionally, we included two control groups: the saline group and a pseudo-conditioning group (the unpaired group); the latter was treated with the same number of cocaine injections than the former, but it was randomly associated with both olfactory stimuli. In a second step, we used an arbitrary cut-off point of 60% of time spent in the CS+ as a criterion to select the mice expressing preference for the cocaine-related cue (the conditioned group). Therefore, the experimental set-up in this study consisted of three groups of mice:

- The experimental or conditioned group (**C**) that was trained to associate an olfactory stimulus with the effects of repeated injections of cocaine.
- The saline group (**S**): animals undergoing the same conditioning sessions, but only received saline injections; and
- The unpaired group (**U**): animals receiving the same number of cocaine injections, but they were provided with random association between olfactory stimuli and cocaine.

## 2.2 PNN characterization by immunoanalysis

Drug-related modifications in PNNs have been mainly estimated by a densitometry analysis using a fluorescence labelling of the lectin wisteria floribunda agglutinin (WFA). WFA binds to the glycosaminoglycan chains of the proteoglycans (CSPGs) and labels the external structure of the PNN (Härtig et al., 1992). WFA immunolabeling was performed on free-floating sections. After several rinses with PBS 0.1 M triton X-100, cerebellar sections were exposed to a blocking buffer with donkey serum, and then they were incubated overnight at 4 °C in 1.5% donkey serum dissolved in PBS 0.1 M Triton X-100 with biotinylated Wisteria floribunda agglutinin (WFA) (1:200; Sigma Aldrich, Madrid, Spain).

To reveal WFA staining, cerebellar tissue was exposed for 2 hours to Cy3-streptavidin (1:200 Jackson ImmunoResearch Europe Ltd, Suffolk, UK). After fluorescence reaction occurred, the sections were mounted using Mowiol (Calbiochem, Merck Chemicals and Life Science, Madrid, Spain).

Fluorescent-labelled sections were examined under a Olympus FV1000 confocal microscope (Olympus Eu-

**Table 1** Number of images per mouse and group

Group	Mouse identifier						Total
	1	2	3	4	5	6	
<b>S</b>	65	35	55	67	63	51	336
<b>C</b>	88	71	81	63	70	19	392
<b>U</b>	61	75	63	30	75	–	304
				Total	.....		1,032

rope Holding GMBH, Hamburg, Germany). Confocal images were acquired in single planes with a 40× lenses, 2.5 zoom, a resolution of 1024 × 1024 and 100 Hz speed. Laser intensity, gain and offset were maintained constant in each analysis.

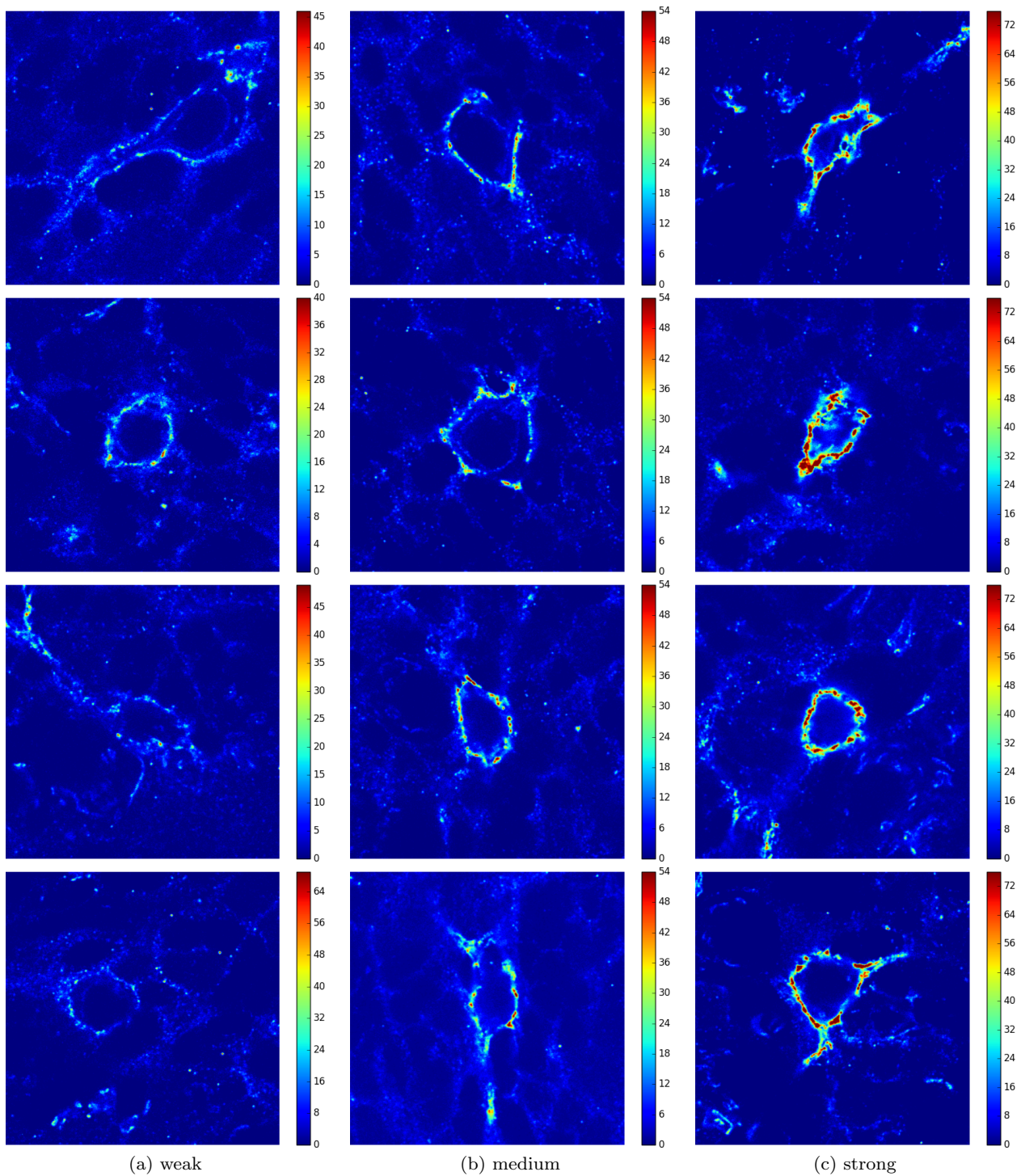
## 2.3 Image dataset

The image dataset was extracted from a total of 17 different mice divided into the three considered groups in the experiment: 6 saline, 6 conditioned and 5 unpaired. From the initial confocal microscope images of the WFA staining fluorescent-labelled sections, PNNs images were extracted in the following way. All PNNs in each image clearly identifiable were manually cropped and approximately centred in images of 400 × 400 pixel resolution (Fig. 1). Finally, a total of 1,032 PNNs images were extracted from all available confocal images of the 17 mice. The number of images per mouse and group are given in Table 1. We refer to a particular mouse in the saline and conditioned groups by  $S_i$  and  $C_i$ ,  $i \in \{1, \dots, 6\}$ , respectively.

Sample PNN images exhibiting weak, medium, and strong PNN patterns are given using the jet color map (Fig. 1). We will refer to weak ( $w$ ) and strong ( $s$ ) PNN as  $w$ -PNN and  $s$ -PNN. Since the original images are gray-level and low-contrast, the use of a color map is of key importance for human visual inspection. Even with the color map, these example images illustrate the difficulty of distinguishing  $w$ -PNN and  $s$ -PNN by a simple and general rule of thumb. However, one can appreciate that strong PNNs have denser regions with higher values (more red pixels in the jet color map). From a computational point of view, one of the challenges is to model this intuitive and generalist observation into a successful representation. This issue is discussed in the following section.

## 2.4 Image descriptor (HND)

To characterize the PNN images, the well-known Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) was used as a starting point. Like other keypoints detectors and descriptors, SIFT is invariant to rotation and scale,



**Fig. 1** Example images with clearly weak PNN (left), clearly strong PNN (right) and in-between PNNs (middle)

and robust to geometric deformations and illumination changes, which makes SIFT and its extensions and variants very useful in applications involving image matching (Liu et al., 2011), recognition (Scovanner et al., 2007) and other problems (Zhao and Ngo, 2013). Although other local keypoint detectors (Tuytelaars and Mikolajczyk, 2008) or descriptors, either general (Miko-

lajczyk and Schmid, 2005), or texture-based (Grigorescu et al., 2002) might be possible, in our case the general overall distribution of SIFT keypoints (Fig. 2) seemed to be potentially useful for discrimination purposes. A good image characterization was therefore required to capture this distribution. The proposed approach is based on these two key observations:

- In contrast to  $w$ -PNN images, in  $s$ -PNN images there are *many* keypoints *close* together, forming *some* dense clusters.
- In both kind of images, there are far more keypoint pairs that are far away than close together.

The first observation led to the concept of histogram of pair-wise keypoint distances, and the second one led to the concept of nearest neighbours. Thus, while histograms compactly summarize the distribution of pair-wise distances, the contamination of the representation with information derived from many big distances is prevented by limiting the distances to the nearest ones as well as by imposing a distance threshold. Additionally, as a way of separately modeling different distance magnitudes,  $K$  nearest neighbours of each keypoint are considered, and one specific histogram for each specific nearest neighbor is computed.

The descriptor, which we call *Histograms of Nearest Distances* (HND), is parameterized by the number of nearest neighbours considered,  $K$ , the number of bins per histogram,  $N$ , and a distance threshold,  $\theta_d$ . The following ones are the detailed steps for its computation, and Algorithm 1 (page 7) formalizes compactly this procedure.

1. The  $K$  nearest neighbours of each keypoint are first found, and their corresponding distances are kept. Since the closest nearest neighbours of a keypoint is itself, we exclude this zero distances from consideration.

Formally, let  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$  be set of the  $m$  detected keypoints,  $\mathbf{x}_i = (x_i, y_i)$ , in a given image, Let  $d(\mathbf{x}_i, \mathbf{x}_j)$  the Euclidean distance between any given pair of keypoints,  $1 \leq i, j \leq m$ , and  $d_{ik} = d_k(\mathbf{x}_i)$ ,  $1 \leq k \leq K$  the  $k$ -shortest distance in the set  $\{d(x_i, x_j) \leq \theta_d, 1 \leq j \leq m, i \neq j\}$ . Then, the neighbour distance matrix  $\mathbf{D}$  is computed,

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1K} \\ d_{21} & d_{22} & \cdots & d_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \cdots & d_{mK} \end{bmatrix} = [\mathbf{D}_1 \mathbf{D}_2 \cdots \mathbf{D}_K]. \quad (1)$$

Thus,  $\mathbf{D}$  has one row per interest point and one column  $\mathbf{D}_j$ ,  $1 \leq j \leq K$ , per nearest neighbour. Therefore, the distances corresponding to the  $j$ -th nearest neighbours are in column  $\mathbf{D}_j$ .

2. For each  $\mathbf{D}_j$  its corresponding  $\mathbf{h}_j$  histogram is computed using  $N$  equally spaced values from 1 to  $\theta_d$ . In detail, let  $\mathcal{R} = \{r_0 = 1, r_1, r_2, \dots, r_N = \theta_d\}$  be the set of right-end point of intervals equally partitioning the distance range  $[1, \theta_d]$ , i.e.  $r_j - r_{j-1} \approx \frac{\theta_d - 1}{N}$ ,  $1 \leq j \leq N$ . Then, the  $b$ -th bin of histogram

$\mathbf{h}_j$ ,  $\mathbf{h}_j(b)$ , is computed by counting the number of distances in  $\mathbf{D}_j$  lying on the  $b$ -th interval in  $\mathcal{R}$ :

$$\mathbf{h}_j(b) = \sum_{i=1}^m \mathbf{1}_{[r_{b-1}, r_b)}(d_{ij}), \quad (2)$$

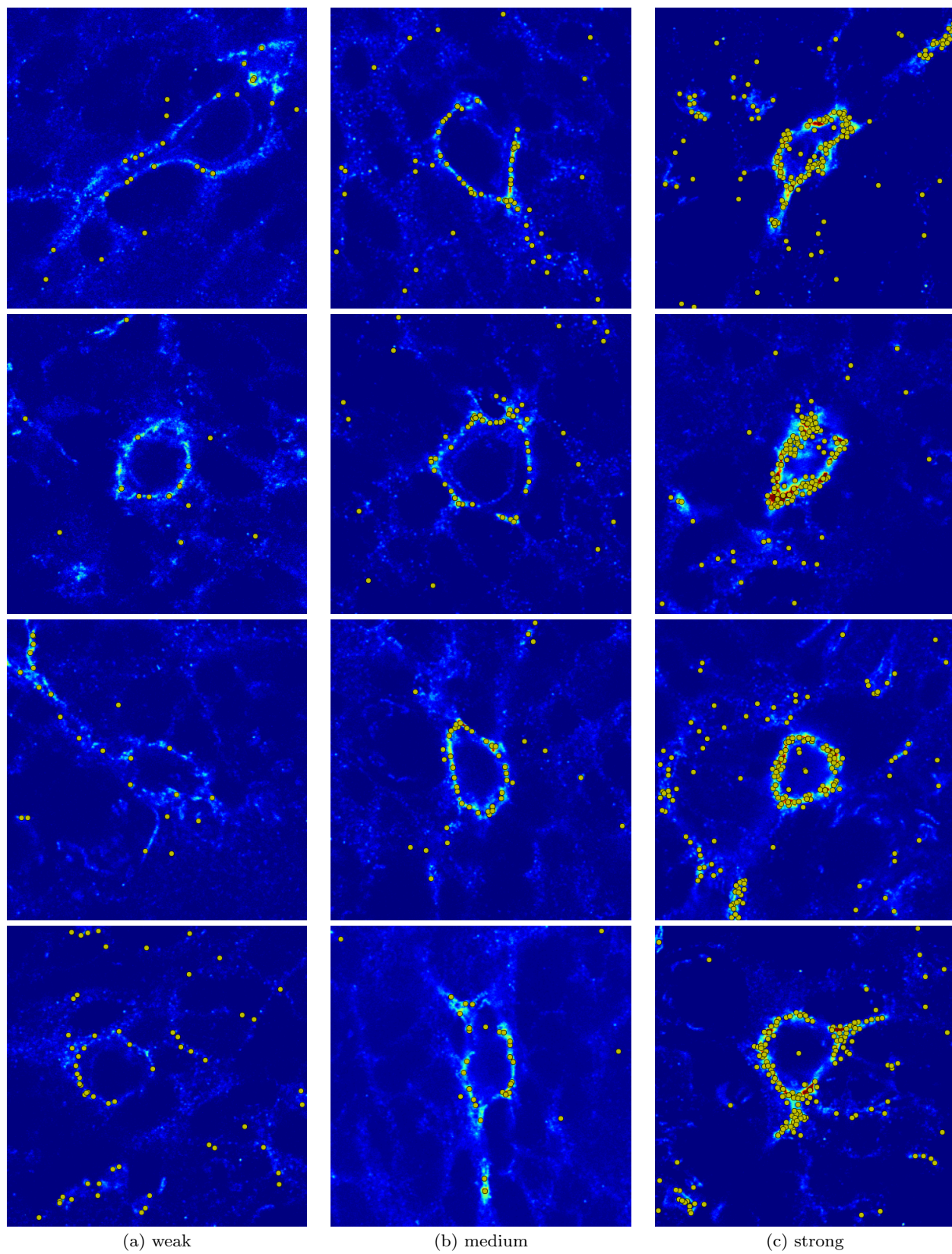
with  $\mathbf{1}_{\mathcal{A}}(x)$  being the indicator function which evaluates to 1 if  $x \in \mathcal{A}$  and to 0 if  $x \notin \mathcal{A}$ . To compute all bins of all histograms, Eq. (2) is evaluated for all  $j \in [1, K]$ ,  $b \in [1, N]$ . The resulting descriptor  $\mathbf{h}_{\text{HND}}$  has therefore dimensionality  $K \cdot N$ .

Notice that for many computer vision tasks, the SIFT *descriptor* is used in addition, or alternatively, to the SIFT detector. In our work, however, we only use the detected keypoints, in terms of a description of their geometric distribution. In other words, no photometric information is used in the HND descriptor. Additionally, it is worth noticing that HND is, by definition, translation and rotation invariant. Scale invariance is not sought because our images are all in the same scale; if they were not, some form of scale invariance would be required too.

To refer to a particular bin of the HND descriptor we use the notation  $k|n$ , where  $k \in \{1, \dots, K\}$  is the  $k$ -th nearest neighbour and  $n \in \{1, \dots, N\}$  is the  $n$ -th bin for the histogram for that particular nearest neighbour. From the normalized HND descriptor ( $K = 5$ ,  $N = 3$ ) averaged over mice per condition group (Fig. 3), it can be observed that the average number of keypoints in the conditioned group is higher than in the saline group for all the bins of the first nearest neighbour ( $B_1 = \{1|k, k \in \{1, 2, 3\}\}$ ) and the first two bins of the second nearest neighbour ( $B_2 = \{2|k, k \in \{1, 2\}\}$ ). For the rest of the bins, there are more points in PNNs of saline group. These bins ( $B_1 \cup B_2$ ) correspond to the shortest distances, and therefore, they may be capturing the denser regions occurring in  $s$ -PNNs which are expected to dominate in mice of the conditioned group. The number of points decreases with  $k|n$  (from left to right) in the conditioned group, whereas for the saline group the largest number of points happen at intermediate  $k|n$ . Again, this is indicative of the different distributions of keypoints in  $w$ -PNN and  $s$ -PNN images.

## 2.5 PNN image pattern discovery and mice classification

In an earlier stage of our work, we performed experiments with supervised learning using expert-provided labels of individual images as ground-truth class labels of weak and strong PNNs. Although there are some images which clearly correspond to weak PNNs and others



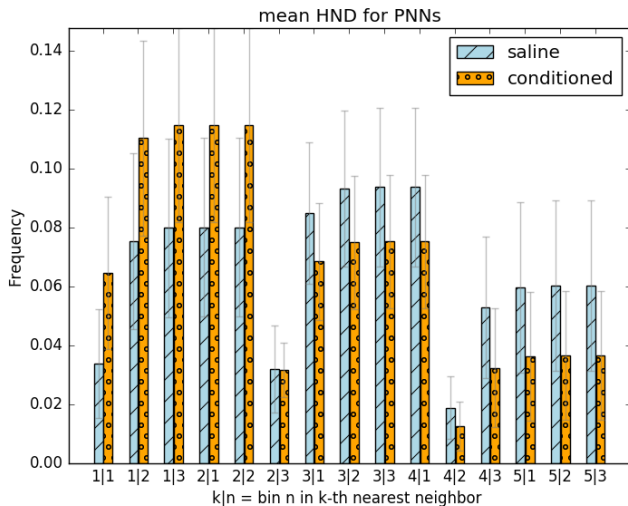
**Fig. 2** SIFT keypoints (yellow dots) detected on the images in Fig. 1

```

Input:  $\mathbf{X}$ , the set of interest points on an image
Output:  $\mathbf{h}_{\text{HND}}$ , the HND descriptor (i.e. the concatenation of distance histograms)
 $\mathbf{D} \leftarrow \text{computeDistancesNearestNeighbors}(\mathbf{X}; K)$  // Eq. (1)
 $\mathcal{R} \leftarrow \text{generateLinearlySpacedValues}([1, \theta_d], N)$ 
for each  $j \in \{1, 2, \dots, K\}$ 
  |  $\mathbf{h}_j \leftarrow \text{computeHistogram}(\mathbf{D}_j, \mathcal{R})$  // Eq. (2)
end
 $\mathbf{h}_{\text{HND}} \leftarrow [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]$ 
return  $\mathbf{h}_{\text{HND}}$ 

```

**Algorithm 1:** HND( $K, N; \theta_d$ ) computation



**Fig. 3** Mean (bars) and standard deviation (error bars) of histograms corresponding to HNDs

to strong PNNs, many more images are something in between which are hard to classify with an unequivocal label. Consequently, it was decided to use unsupervised learning at image level (Sect. 2.5.2), and supervised learning only at mouse level (Sect. 2.5.4). But before presenting the details for these two processes, some insights are given in graphical form (Sect. 2.5.1).

### 2.5.1 Visualizing the HND in a low-dimensional space

We used the t-Distributed Stochastic Neighbor Embedding (t-SNE) technique (van der Maaten and Hinton, 2008) to reveal the implicit structure in the 15-D HND space, and visualize it in the human-intuitive 2-D space. The result using a perplexity value of 20 (Fig. 4) shows that, as in most real-world problems, the two compared mice groups significantly overlap in this space. After all, PNN images within one group can be of many different types in a continuum ranging from clear w-PNN to clear s-PNNs. The output is not random, though, since the manifold underlying the data is captured and some regularities emerge. Indeed, it can reasonably be

argued that the overlapping pattern itself supports the hypothesis that saline mice do not develop so many s-PNN structures as conditioned mice do.

Interestingly, there is a noticeable difference between the sample distribution of both groups. In particular, there are areas where points of one of the groups dominate, such as the dense red dots (conditioned mice) in the upper right part of the scatter plot or the areas of denser blue squares (saline mice) in the bottom left part. Therefore, one can hypothesize that regions where images of conditioned mice dominate would mostly correspond to s-PNNs. As a matter of fact, by observing images corresponding to different areas in this 2D visualization space (Fig. 5), different image PNN structures in the dominant saline and conditioned areas can be noticed. The visualization of the set of images of particular mice (Fig. 6) provides further insight into the different distributions of the patterns of PNN images depending on the experimental group of the mouse.

The above discussion and series of plots not only provide a first evidence on the hypothesis of this work, but also point to the potential of the proposed approach, namely, an unsupervised image-level pattern discovery plus a supervised mice-level group classification.

### 2.5.2 Finding image patterns unsupervisedly

Following our approach, each PNN image is represented with its HND descriptor (Sect. 2.4). Our purpose now is to be able to automatically find similar patterns among the images in the dataset through their HND characterizations. In particular, the goal is to automatically indicate to which of the found patterns a new PNN image resembles the most. To that end, we use vector quantization; in essence, similar HND descriptors will be grouped together and will be assigned a representative prototype against which novel descriptors will be compared. Since several groups will be found, then a new descriptor will be assigned to the group whose prototype this descriptor is most similar to. It is important to note that this is performed unsupervisedly, which means that no prior information of the proce-

dence of the PNN images is used at all. In other words, at this stage, our algorithm ignores whether the PNN images come from mice in the saline, conditioned or unpaired groups. The detailed procedure is as follows.

Let  $\mathcal{H} = \{\mathbf{h}_{\text{HND}}^{(i)}\}_1^m$  be the set of HND descriptors for a given set of images. Let  $L$  be the desired number of groups (i.e. the patterns) we want to discover. We use the well known  $k$ -means clustering (Jain, 2010) to find these patterns (with  $k = L$ ). To find the  $L$  clusters  $\{\mathcal{C}_j\}_1^L$  and their corresponding prototypes  $\{\mathbf{p}_j\}_1^L$  from a given set of  $m$  data points  $\mathcal{X} = \{\mathbf{x}_i\}_1^m, \bigcup_{j=1}^L \mathcal{C}_j = \mathcal{X}, \mathcal{C}_i \cap \mathcal{C}_j = \emptyset, i \neq j$ ,  $k$ -means proceeds iteratively, usually starting with a random assignment of every data point  $\{\mathbf{x}_i\}$  to a single cluster  $c_i \in \{1, 2, \dots, L\}$ .

1. Given the current clusters  $\{\mathcal{C}_i\}_1^L$ , compute their prototypes  $\{\mathbf{p}_j\}$  as the mean of the points within each cluster:

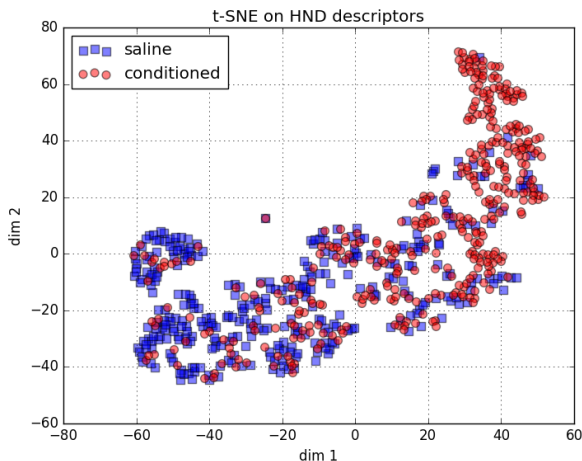
$$\mathbf{p}_j = \frac{1}{|\mathcal{C}_j|} \sum_{\mathbf{x} \in \mathcal{C}_j} \mathbf{x}, j \in \{1, 2, \dots, L\}. \quad (3)$$

with  $|\mathcal{A}|$  denoting the cardinality of the set  $\mathcal{A}$ .

2. Given the new prototypes  $\{\mathbf{p}_i\}$ , recompute the assignment of points  $\{\mathbf{x}_i\}$  to clusters  $\{c_i\}$  by finding their closest prototypes:

$$c_i = \arg \min_{j \in \{1, 2, \dots, L\}} d(\mathbf{x}_i, \mathbf{p}_j), i \in \{1, 2, \dots, m\}. \quad (4)$$

These two steps are repeated until convergence (the clusters become stable and no longer change) or until a maximum number of iterations are performed. In our



**Fig. 4** Output of t-SNE on the set of HND descriptors. Each point in this space corresponds to a single PNN image over the set of mice in the saline and conditioned groups. The manifold underlying the HND descriptor is somehow uncovered

case, the data points that are clustered are the set of HND descriptors, i.e.  $\mathcal{X} = \mathcal{H}$ . After that, to assign a new image descriptor  $\mathbf{q} \notin \mathcal{X}$  to a given cluster, the most similar existing prototype is found, exactly in the same way as data points are reassigned to a cluster in Eq. (4).

Examples of BoW averaged over mice per condition group (Fig. 7) illustrate that mice in different groups may differ in the number of PNNs assigned to different clusters. These differences suggest its potential discriminative effect.

### 2.5.3 Choosing representative images

Although all images in the identified clusters are used for the subsequent supervised learning stage (Sect. 2.5.4), we were also interested in showing a limited representative set of images for human experts to analyze. Simply choosing images randomly might result in a non-representative sample of each cluster. Therefore, a smarter sampling approach based on a probability density estimation (Scott, 1992) is applied over the feature space of the HND descriptors. To that end, given the  $m$  data points  $\{\mathbf{x}_i\}$ , each point in a  $n$ -dimensional space, the underlying unknown density function  $f$  is estimated non-parametrically through kernel density estimation (KDE) as:

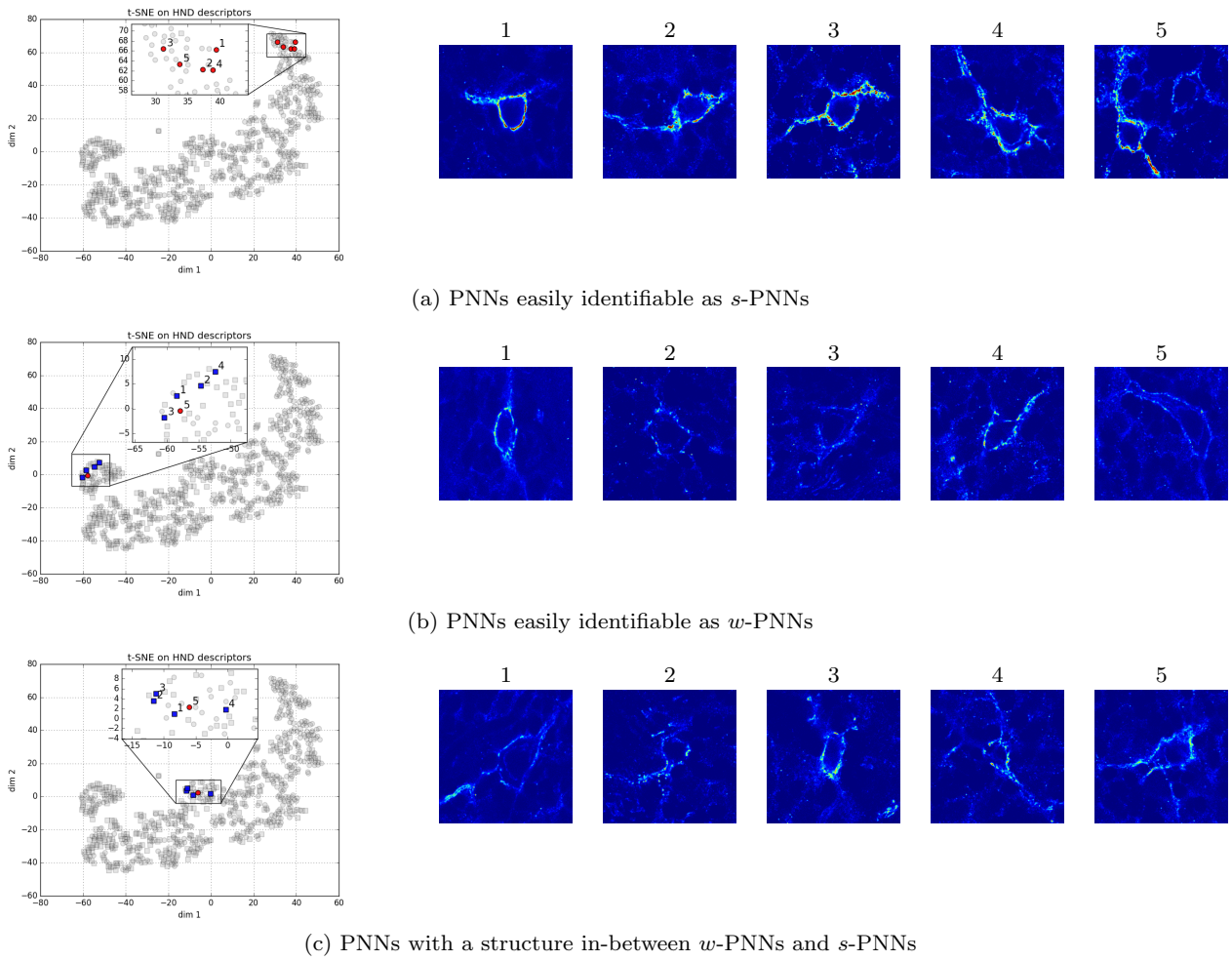
$$\hat{f}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^m K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i), \quad (5)$$

where  $K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-\frac{1}{2}} K(\mathbf{H}^{-\frac{1}{2}} \mathbf{x})$ , with  $K(\cdot)$  being a chosen kernel function, and  $\mathbf{H}$  being the  $n \times n$  bandwidth matrix. Intuitively, a kernel-based density estimation resembles the estimation of an histogram from data, but with benefits regarding smoothness and continuity. Since we used a Gaussian kernel, and automatic scale parameter selection with Scott's rule (Scott, 1992), then,

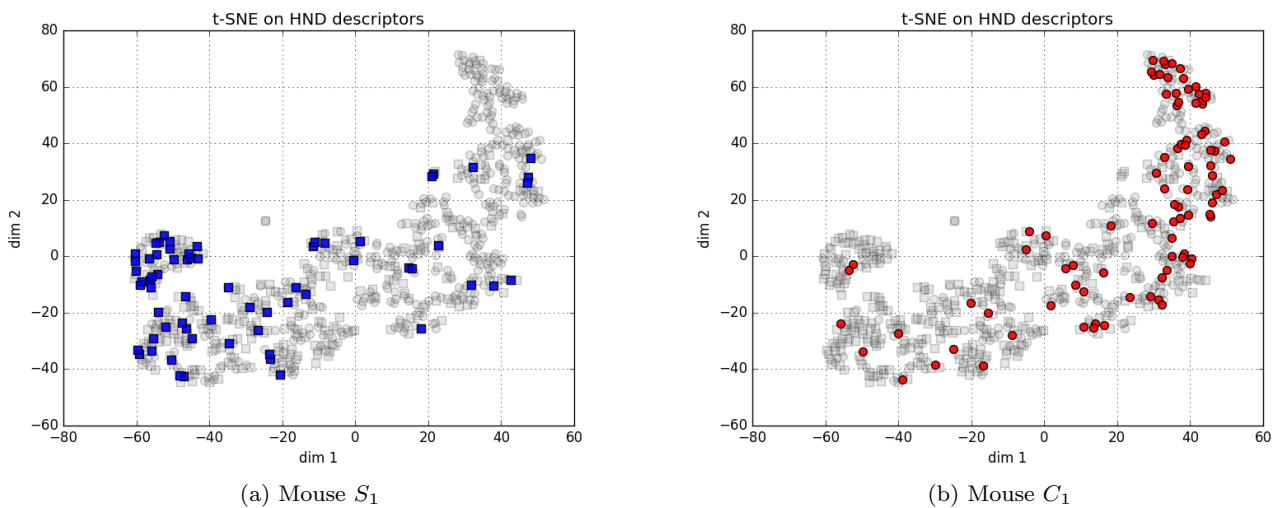
$$K_{\mathbf{H}}(\mathbf{x}) = (2\pi)^{n/2} |\mathbf{H}|^{-\frac{1}{2}} e^{-\frac{1}{2} \mathbf{x}^T \mathbf{H}^{-1} \mathbf{x}},$$

and  $\mathbf{H}_{ii} = \sigma_i^2 \cdot m^{-2/(n+4)}$ , with  $\sigma_i^2$  being the variance of the  $i$ -th variable. The way we use KDE in our setting consists of applying this procedure separately for the data points in each cluster. To avoid numerical issues due to singular matrices, zero-mean Gaussian noise of standard deviation  $10^{-3}$  was added to each variable of the data points prior to the KDE estimation. After  $\hat{f}$  is estimated, we evaluate it on the data points and normalize them to sum to 1, so we have  $p_i = \hat{f}(\mathbf{x}_i) / \sum_{i=1}^m \hat{f}(\mathbf{x}_i)$ . Then, for each cluster, we randomly select the desired number  $\ell$  of points according to the probability density given by the  $p_i$  values.





**Fig. 5** Examples of PNNs randomly selected from three very distinct areas of the output of t-SNE, which roughly correspond to easily identifiable kinds of PNNs

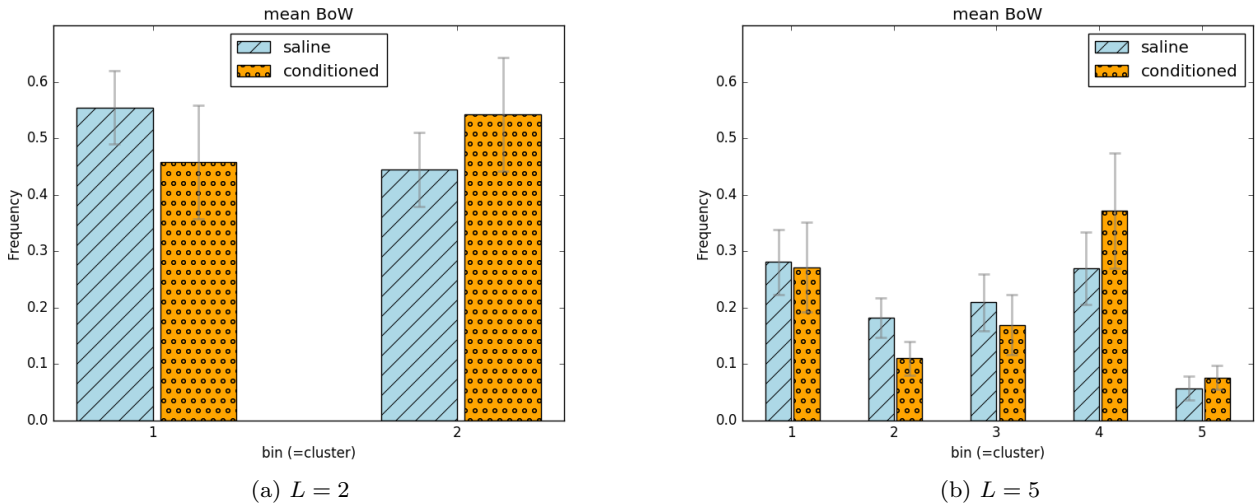


**Fig. 6** PNNs of one mouse in each group highlighted on t-SNE output

#### 2.5.4 Recognizing the mice group supervisedly

Supervised classification aims at learning a decision function that can predict the class  $y$  of a novel data in-

stance  $\mathbf{x}$ . To this end, a classifier learns the association



**Fig. 7** Mean (bars) and standard deviation (error bars) of histograms corresponding to BoWs of the two experimental groups. Two different vocabulary sizes  $L$  are given as examples

of instances  $\{\mathbf{x}_i\}$  to class labels  $\{y_i\}$  given as a training set. The adjective “supervised” comes from the fact that ground-truth class labels are used for this *training* stage.

In our problem, we consider two classes, namely the saline and the conditioned experimental groups of the mice. We want to be able to predict the class of an unseen mouse given the set of its PNN images. Since we can compute the HND descriptor for every PNN image, we might use the set of HND descriptors corresponding to the PNN images of a mouse to infer its class label. However, this approach is complicated for a number of reasons such as that different mice have different number of PNN images.

In cases like this one, some kind of pooling mechanism is usually required that “sums up” a set of features of an arbitrary size into a single, fixed-length representation. One widely known pooling approach consists of the bag-of-words (BoW). In essence, the BoW of a *visual entity* corresponds to an histogram-like count of how many of the features in a set for that entity are similar to each of a given set of “words”. This set of words is often referred to as vocabulary or dictionary. Although conceptually simple, this representation has shown to be useful in a number of problems in computer vision (Lazebnik et al., 2006, Niebles and Fei-Fei, 2007, Yu et al., 2013).

A key requirement to use the BoW is to define a vocabulary of words. In our case, the words correspond to each the clusters of HND descriptors found unsupervisedly (Sect. 2.5.2). Let  $\{\mathbf{x}_i\}_1^m$  be the set of  $m$  HND descriptors of the set of PNN images for a certain mouse. By using Eq. 4, each of these descriptors will

be assigned to a cluster (word)  $c_i$ . Then, we can count how many of the descriptors are assigned to each of the words, and this will be the BoW representation for that mouse. Since there are  $L$  clusters (words), the BoW representation will be a  $L$ -length histogram  $\mathbf{h}_{\text{BoW}}$ . Formally, each bin  $b$  of this histogram is therefore computed as

$$\mathbf{h}_{\text{BoW}}(b) = \sum_{i=1}^m \mathbf{1}_{\{b\}}(c_i), \quad b \in \{1, 2, \dots, L\}, \quad (6)$$

where  $c_i$  is the cluster which data point  $\mathbf{x}_i$  is assigned to, and we again use the indicator function  $\mathbf{1}_{\mathcal{A}}(\cdot)$  by using as the set  $\mathcal{A}$  the singleton whose element is the cluster (word) index for which we want to count the number of assignments. Remember that in our problem each data point  $\mathbf{x}_i$  corresponds to a HND descriptor.

By using this BoW representation, a set of mice  $\mathcal{M} = \{m_i\}_1^M$  can therefore be represented as the set of their corresponding BoW histograms,  $\{\mathbf{h}_{\text{BoW}}^{(i)}\}_1^M$  which, together with their true class labels  $\{y_i\}_1^M$ , are used for supervisedly learning a classifier. After the learning (training) stage, the class of a different mouse  $m \notin \mathcal{M}$  (i.e. whose images have not been used during training) can be predicted from the BoW histogram computed on the set of PNN images available for this mouse  $m$ . As we will see below, care was taken to use disjoint sets of training and testing sets of mice and PNN images, both when computing the vocabulary of words and when training the classifier in order to distinguish between conditioned and saline mice. Regarding the classifier, almost any general-purpose classifier can be used. In our case, the Support Vector Machines (SVM) model (Vapnik, 1998) is adopted since it is widely accepted and has generally good performance.

It may be worth noting that since the HND descriptor has very low dimensionality (15 bins), no dimensionality reduction is required for the unsupervised learning. This is also the case for the supervised learning, since the vocabulary size is also small (less than 100 words are used).

## 2.6 Settings and software

The VLFeat library (Vedaldi and Fulkerson, 2008) for SIFT detection, and Python packages `scikit-image` (van der Walt et al., 2014) for image processing and `scikit-learn` (Pedregosa et al., 2011) for machine learning, were used. For t-SNE, the Python implementation provided by its author was used (van der Maaten, 2018). For SIFT, the edge and peak thresholds were set to 10 and 2, respectively. The HND parameters were  $K = 5$ ,  $N = 3$  and  $\theta_d = 30$ , thus resulting in compact 15-dimensional feature vector. An SVM with radial-basis function (RBF) kernel was used. The regularization parameter  $C$  for the SVM, and the scale parameter  $\gamma$  of the RBF kernel were selected by validation from the sets  $C \in \{10^i : i \in \{-8, -7, \dots, 7, 8\}\}$  and  $\gamma \in \{10^j : j \in \{-4, -3, \dots, 3, 4\}\}$ , respectively.

## 2.7 Computational analysis

Asymptotic cost and actual running times on a i7-6700HQ CPU @ 2.60GHz are given in Table 3 for different parts of the proposed approach. Some explanations and observations are in order. The precise asymptotic cost for some algorithms can be implementation-dependent or can be particularly complex because it is much data- and parameter-dependent. Then, to be of practical use, a simpler and general expression is provided instead. This is particularly the case of SIFT, whose detailed computational breakdown can be found elsewhere (Vinukonda, 2011). The times for SIFT points detection (part  $A_1$ ) are an overestimate of the strictly required times since we used a command-line tool which computes also the SIFT descriptors (which are ignored for HND computation) and writes the resulting SIFT points to a file. After the SIFT points are found, the computation of the HND descriptor ( $A_2$ ) is very fast given the average number of points found in the PNN images in our dataset (Table 2).

Overall, it can be noted that the complete procedure is computationally quite affordable and reasonably scalable to more data and higher dimensions. For instance, using the mean times as an estimate (components  $A_1 + A_2 + B_2 + C_1$ , with  $L = 10$ ), classifying a mice would take about 265 ms. (i.e. a classifica-

**Table 2** Number of SIFT points detected on PNN images

Group	Average	Standar deviation
<b>S</b>	103.2	55.7
<b>C</b>	170.6	85.5
<b>S <math>\cup</math> C</b>	139.5	80.6

tion speed of 3.8 mice/s), even with unoptimized code. Certainly, with more mice and more or larger images per mice, some costs would increase, according to complexities given (Table 3), but this is hardly to become a practical issue in the contexts where these data are commonly used in neurobiological research, where hard time constraints are not usually imposed.

## 3 Experiments and results

The goal of our study is to assess how successful the proposed machine vision methods are in, first, qualitatively predicting functionally relevant PNN image patterns and, second, quantitatively inferring the experimental mice groups. Certainly, the ultimate interest of this assessment relies in gaining insight into how these results support the current evidence on the link between neuronal plasticity and drug addiction mechanisms.

### 3.1 Qualitative assessment of discovered PNN patterns

To gain some insight about the contents of the groups (clusters) of PNNs found by the unsupervised clustering procedure (Sect. 2.5.2), an small representative subset of images per group was selected (Sect. 2.5.3). To that end, the density-based sampling approach (Sect. 2.5.3), with  $\ell = 10$ , was applied to each of the  $L = 3$  clusters. The reason for this choice of the number of clusters was two-fold: on the one hand, it is a small number for human experts to analyse carefully without much cognitive and physical burden; on the other hand, from our discussions with them, some prior belief emerged that the number of relevant patterns might actually be small, about 3 or 4.

The set of the  $\ell = 10$  images selected per cluster were shown to the experts, who were asked to individually and carefully observe the images, and then describe in written form their patterns by characterizing the images within each group and how they differ with respect to those images in the other groups. A subset of these images is given in Fig. 8. To avoid judgement biases, we tried to inform the experts the least, and made an effort to use with them a jargon-free vocabulary regarding machine vision concepts. For instance, rather than talking about “clusters”, we used the more neutral term “pattern”. By reading experts’ annotations (Table 4),

**Table 3** Temporal complexity and running times (ms), average and standard deviation (in parentheses)

A: HND computation		
$A_1$ : SIFT detector		
Complexity	Values	Running time
$\theta(N^2)^*$	$N = 400$	263.4 (13.3) <sup>†**</sup>
* $N$ : width (height) of image <sup>†</sup> Average per image ** Includes SIFT description + I/O		
$A_2$ : $k$ -NN and histogram of distances (HD)		
Complexity	Values	Running time
$O(m^2Kd)^{*†}$	$m \approx 140^{**}, K = 5, d = 2^{\ddagger}$	1.5 (13.3) <sup>§</sup>
* Cost of $k$ -NN, subsumes that of HD which is $O(mK)$ <sup>†</sup> $m$ : No. SIFT points/image, $K$ : No. of NNs, $d$ : data dim. ** See Table 2 <sup>‡</sup> $(x, y)$ of each SIFT point <sup>§</sup> Average per image		
B: BoW computation		
$B_1$ : $k$ -means		
Complexity	Values	Running time <sup>‡</sup>
$\theta(nkd)^*$	$d = 15^{\dagger}, n = 728^{**}$	See below for varying $k$
	$k = 2$	31.5 (2.8)
	$k = 5$	60.9 (6.3)
	$k = 10$	84.9 (10.6)
	$k = 50$	176.2 (5.5)
* $n$ : No. of data points, $k$ : No. of clusters, $d$ : data dim. <sup>†</sup> That of HND ** All images of mice in saline and conditioned groups <sup>‡</sup> For 20 repetitions to account for variability		
$B_2$ : Computing BoW histogram of given set of images		
Complexity	Values	Running time <sup>‡</sup>
$O(mkd)^{*†}$	$m = 70^{**}$	See below for varying $k$
	$k = 2$	0.34 (0.027)
	$k = 5$	0.36 (0.011)
	$k = 10$	0.37 (0.008)
	$k = 50$	0.40 (0.007)
* $m$ : No. of images, $k$ and $d$ : same meaning as before <sup>†</sup> $m$ times the cost of finding the closest cluster, $O(kd)$ ** Similar set size as in the mice in the dataset (Table 1) <sup>‡</sup> For 20 repetitions		
C: Mice classification		
$C_1$ : SVM classification given the BoW histogram		
Complexity	Values	Running time (ms)
$O(Nd)^*$	$N = 11^{\dagger}, d = L = 10^{**}$	0.17 (0.07) <sup>‡</sup>
* $N$ : No. of training instances, $d$ : data dim. <sup>†</sup> Leaving one mouse out ** No. of clusters <sup>‡</sup> Averaged over 12 mice and 10 repetitions		

it can be observed that they all (1) found quite clearly distinctive features in each of the three groups of images; and (2) roughly agree among themselves in the features they independently found.

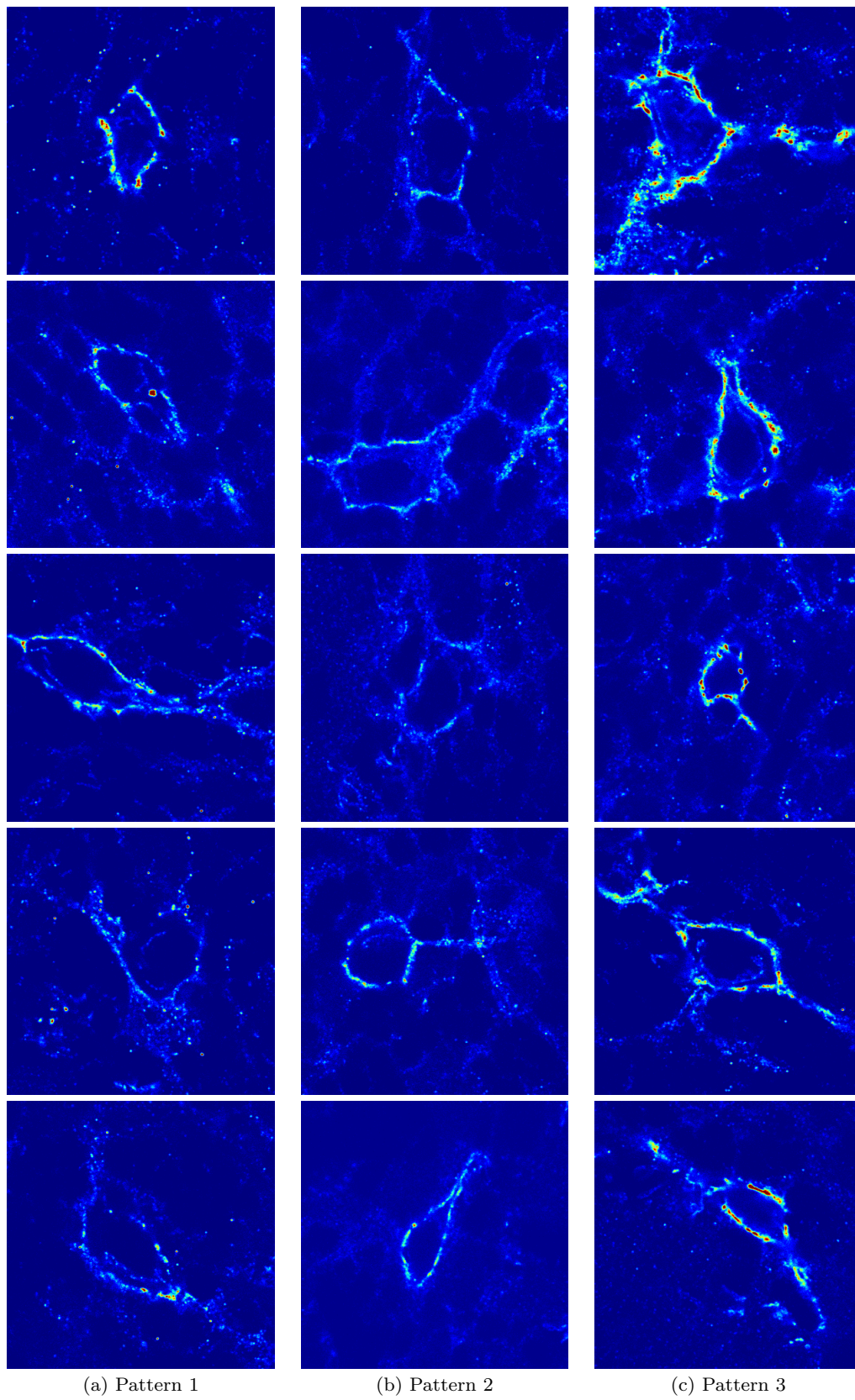
In an attempt to provide a unified general characterization of the three patterns, we identified the common traits that are approximately revealed from expert’s descriptions (Table 5). Roughly speaking, Pattern 1 can be assimilated to weak PNNs and Pattern 3 to strong PNNs, while Pattern 2 exhibits more intermediate, less clear-cut appearance, even though a distinctive round shape is still identified. It is worth noticing that Expert 3 associated the Pattern 3 directly with animals in the conditioned group. Interestingly, the 9 out of 10 images selected automatically for this pattern came from mice in the conditioned group, whereas in Patterns 1 and 2, only 3 out of 10 images came from conditioned mice. These observations suggest the proposed descriptor has the intended ability of capturing and characterizing neurons with different visual patterns corresponding to different PNNs structures.

### 3.2 Quantifying the mice group recognition rate

To find whether these patterns are not only qualitatively meaningful and subjectively useful, classification of the experimental group (saline vs conditioned) *per mice* was performed using the BoW approach (Sect. 2.5.4). Given the stochastic nature of the  $k$ -means, the procedure (clustering, histogram coding, training and classification) was repeated ten times for four different cluster sizes ( $L \in \{2, 3, 4, 5\}$ ). A leave-one-out statistical validation was employed: each mice was used as test whereas the remaining 11 were used for training, and this procedure is repeated for each of the available mice. Despite the limited size of the dataset (12 mice), the results (Table 6) are interesting since they suggest that the designed image descriptor combined with unsupervised and supervised learning, is able to encode discriminative PNN properties and, eventually, predict with reasonable reliability the experimental condition (saline or conditioned) of the mice. The confusion matrices (Table 7) reveal that the classification rates per experimental condition are balanced.

### 3.3 Neurobiological relevance

In functional terms, the fact that an automated algorithm can discriminate PNNs expressed by saline animals from those of the conditioned group implies that the previously reported differences between these groups Carbo-Gas et al. (2017) are robust. Moreover,



**Fig. 8** A selection of patterns found when clustering with  $L = 3$  groups

**Table 4** Expert description of the patterns found by clustering.

Expert	Pattern		
	1	2	3
1	Except images 4 and 8, images have an identifiable morphological pattern. Cells are elongated. Expression degree in brightness terms is similar, intermediate.	Except image 5, all images are morphologically circular and show faint PNNs	Morphological inconsistency: there are both elongated and circular neurons. Recognizable high PNN expression
2	They seem to have less intensity. They are not very compact and have quite a few openings. They look more elongated than the others.	They present low intensity, with many gaps and look more round.	They are intense, are closer than the other patterns and even neurites are intense. Unlike the other patterns, they have branches.
3	Images represent the pattern of a non-conditioned animal. Less branching and condensation than Pattern 3. Less strong and fewer neurites than in Pattern 3. PNN is mostly observed in the soma part	Not clearly belonging to saline or conditioned group, but more to the latter. Branching PNNs, but of middle intensity and density. PNNs are bigger than in Pattern 1. Similar in size and shape to Pattern 3, but less intense.	They represent mostly the pattern of a conditioned animal. PNNs look like more condensed, more structured, denser. Besides being more intense, these neurons have larger branching, they have more neurites with PNN; an intense WFA labelling is not only observed in the soma. The thickness of PNNs seems to be higher. They are similar in size to those in Pattern 2, but not as much intense or dense. They are more intense, denser, and PNNs are observed in both soma and proximal dendrites. It seems they are bigger than those in Pattern 1.

**Table 5** Commonalities in experts' descriptions of the patterns

Pattern	Shape	Intensity	Density	Branching
1	Elongated	Middle	Low	Low
2	Circular	Low	Middle	Middle
3	Inconsistent	High	High	High

**Table 6** Performance of group classification (saline vs conditioned) at mouse-level with the BoW approach for different vocabular sizes  $L$  by leaving-one-out validation. Averages over the 12 mice of mean, standard deviation, and median of the accuracy for each mouse are reported. The corresponding confusion matrices for  $L \neq 2$  are given in Table 7

	Number of clusters ( $L$ )			
	2	3	4	5
mean	100	85	86.7	79.2
std. dev.	0	32.1	29.7	33.9
median	100	100	100	95.8

the present findings suggest that cocaine-induced memories are capable of up-regulating reliably one of the main plasticity mechanisms for synaptic stabilisation in the cerebellum. Notwithstanding these findings, the present analysis does not allow to rule out that the observed modifications in the PNN structure are memory unrelated and due to the neuropharmacological action of cocaine.

As an additional test, taking the BoW representation of images of both the saline and conditioned mice as training data, we classified the mice in the unpaired group. By tentatively assuming mice in the unpaired group actually belong to the saline group classification results (Table 8) are not as good as when classifying the mice in saline and conditioned groups. In fact, some better results are obtained if many more clusters than those tested with the conditioned and saline groups alone are used. This may be explained by the fact that the images in the unpaired group do not really distribute as those of the other two groups, and therefore might benefit from more fine-grained types of PNNs image patterns. Despite the decrease in recognition rate, the results still provide evidence that the unpaired mice could not establish a cue-drug association, and therefore they did not express strong fully condensed PNNs as the conditioned mice did. The variability (standard deviation of the estimated recognition rate) for these tests is in general quite large throughout all tested  $L$ . This further indicates that unpaired mice are not able to build PNNs according to the patterns observed and characterised for conditioned mice.

These results are interesting since, indeed, the unpaired group is different from the other two groups. Unpaired animals received the same cocaine treatment as the conditioned mice but, in this case, nothing could

**Table 7** Confusion matrices for different vocabulary sizes  $L$ , as complementary information to Table 6. Rows correspond to ground-truth groups and columns to predicted groups, for three number of clusters. Values are the number of mice and, in parentheses, the percentage computed group-wise (i.e. over 60). The overall accuracy is given by the sum of the correct classifications (on the main diagonal of each matrix) over the total (120)

	$L = 3$		$L = 4$		$L = 5$		Total
	<b>S</b>	<b>C</b>	<b>S</b>	<b>C</b>	<b>S</b>	<b>C</b>	
<b>S</b>	52 (86.7%)	8 (13.3%)	53 (88.3%)	7 (11.7%)	46 (76.7%)	14 (23.3%)	60
<b>C</b>	10 (16.7%)	50 (83.3%)	8 (15.0%)	51 (85%)	11 (18.3%)	49 (81.7%)	60
Accuracy (%)	85.0		86.7		79.2		

be learnt because the contingent cue-cocaine association was lacking. Hence, the results yielded by our algorithms indicate that in addition to cocaine-induced memory, PNN structure might be also regulated by a mere repeated experience with the drug independently of memory acquisition.

#### 4 Conclusions

Addiction results from long-lasting plasticity modifications that render the brain “inflexible” to drugs and drug-related stimuli. Thus, it seems crucial for the field to understand the mechanisms of stabilization that are linked to drug-induced synaptic changes. The present results using a computer vision and machine learning approach support prior findings involving PNN expression at the cerebellar cortex in cocaine-related conditioned memory. Also, these results open new avenues to describe specific patterns of brain plasticity structural modifications based on a machine vision perspective. The ability of the proposed image descriptor and the bag-of-words approach to characterize different patterns of PNN expression around Golgi cells is remarkable, particularly because these patterns roughly correspond to the treatment groups.

#### Information Sharing Statement

Data and code related to this work are available in GitLab: <https://gitlab.com/vtraver/PNNsMiceMachineVision> (RRID:SCR.016485). Details and explanations can be found in the [README.md](#) and the source files.

#### References

S. Armano, P. Rossi, V. Taglietti, and E. D’Angelo. Long-term potentiation of intrinsic excitability at the mossy fiber-granule cell synapse of rat cerebellum. *Journal of Neuroscience*, 20(14):5208–5216, Jul 2000.

J. M. Blacktop, R. P. Todd, and B. A. Sorg. Role of perineuronal nets in the anterior dorsal lateral hypothalamic area in the acquisition of cocaine-induced

conditioned place preference and self-administration. *Neuropharmacology*, 118:124–136, May 2017.

- G. Brückner, K. Brauer, W. Härtig, J. R. Wolff, M. J. Rickmann, A. Derouiche, B. Delpech, N. Girard, W. H. Oertel, and A. Reichenbach. Perineuronal nets provide a polyanionic, glia-associated form of microenvironment around certain neurons in many parts of the rat brain. *Glia*, 8(3):183–200, July 1993.
- M. Carbo-Gas, D. Vazquez-Sanroman, L. Aguirre-Manzo, G. A. Coria-Avila, J. Manzo, C. Sanchis-Segura, and M. Miquel. Involving the cerebellum in cocaine-induced memory: pattern of cFos expression in mice trained to acquire conditioned preference for cocaine. *Addiction biology*, 19(1):61–76, January 2014a.
- M. Carbo-Gas, D. Vazquez-Sanroman, I. Gil-Miravet, J. De las Heras-Chanes, G. A. Coria-Avila, J. Manzo, C. Sanchis-Segura, and M. Miquel. Cerebellar hallmarks of conditioned preference for cocaine. *Physiology & behavior*, 132:24–35, Jun 2014b.
- M. Carbo-Gas, J. Moreno-Rius, J. Guarque-Chabrera, D. Vazquez-Sanroman, I. Gil-Miravet, D. Carulli, F. Hoebeek, Chris D. Zeeuw, C. Sanchis-Segura, and M. Miquel. Cerebellar perineuronal nets in cocaine-induced pavlovian memory: Site matters. *Neuropharmacology*, 125:166–180, October 2017.
- M. Carta, M. Mameli, and C. F. Valenzuela. Alcohol enhances GABAergic transmission to cerebellar granule cells via an increase in Golgi cell excitability. *24(15):3746–3751*, April 2004.
- D. Carulli, T. Laabs, H. M. Geller, and J. W. Fawcett. Chondroitin sulfate proteoglycans in neural development and regeneration. *Current opinion in neurobiology*, 15(1):116–120, February 2005.
- D. Carulli, K. E. Rhodes, D. J. Brown, T. P. Bonnert, S. J. Pollack, K. Oliver, P. Strata, and J. W. Fawcett. Composition of perineuronal nets in the adult rat cerebellum and the cellular origin of their components. *The Journal of comparative neurology*, 494(4): 559–577, February 2006.
- D. Carulli, S. Foscarin, A. Faralli, E. Pajaj, and F. Rossi. Modulation of semaphorin3A in perineuronal nets during structural plasticity in the adult

**Table 8** Performance for different vocabular sizes  $L$  of group classification (saline vs conditioned) at mouse-level of unpaired mice taking saline as their ground-truth label

	Number of clusters ( $L$ )										
	2	3	4	5	6	8	10	15	20	50	100
mean	20	62	70	56	50	50	66	70	56	60	70
std. dev.	40	46.9	45.8	45.4	50	50	44.8	42.2	47.2	49	40.2
median	0	100	100	70	50	50	100	100	80	100	100

- cerebellum. *Molecular and Cellular Neuroscience*, 57: 10–22, November 2013.
- E. D’Angelo and C. I. De Zeeuw. Timing and plasticity in the cerebellum: focus on the granular layer. *Trends in neurosciences*, 32(1):30–40, Jan 2009.
- S. Foscari, D. Ponchione, E. Pajaj, K. Leto, M. Gawlak, G. M. Wilczynski, F. Rossi, and D. Carulli. Experience-dependent plasticity and modulation of growth regulatory molecules at central synapses. *PLOS ONE*, 6(1):1–14, 01 2011.
- R. Frischknecht, M. Heine, D. Perrais, C. I. Seidenbecher, D. Choquet, and E. D. Gundelfinger. Brain extracellular matrix affects AMPA receptor lateral mobility and short-term synaptic plasticity. *Nature neuroscience*, 12(7):897–904, July 2009.
- M. U. Ghani, F. Mesadi, S. D. Kamk, A. O Argunşah, A. F. Hobbiss, I. Israely, D. Ünay, T. Taşdizen, and M. Çetin. Dendritic spine classification using shape and appearance features based on two-photon microscopy. *Journal of Neuroscience Methods*, 279:13–21, March 2017.
- T. A. Gillette, K. M. Brown, and G. A. Ascoli. The DIADEM metric: Comparing multiple reconstructions of the same neuron. *Neuroinformatics*, 9(2): 233, April 2011.
- S. E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on Gabor filters. *IEEE Transactions on Image Processing*, 11(10): 1160–1167, October 2002. ISSN 1057-7149. doi: 10.1109/TIP.2002.804262.
- B. Grimpe and J. Silver. The extracellular matrix in axon regeneration. *Progress in brain research*, 137: 333–349, 2002.
- S. E. Hyman, R. C. Malenka, and E. J. Nestler. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annual Review Of Neuroscience*, 29:565–598, 2006.
- W. Härtig, K. Brauer, and G. Brückner. Wisteria floribunda agglutinin-labelled nets surround parvalbumin-containing neurons. *Neuroreport*, 3(10): 869–872, October 1992.
- A. K. Jain. Data clustering: 50 years beyond  $k$ -means. *Pattern recognition letters*, 31(8):651–666, 2010.
- S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2006.
- C. Liu, J. Yuen, and A. Torralba. SIFT flow: Dense correspondence across scenes and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):978–994, May 2011.
- D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10): 1615–1630, October 2005.
- J. Moreno-Rius and M. Miquel. The cerebellum in drug craving. *Drug and Alcohol Dependence*, 173:151–158, April 2017.
- E. A. Moulton, I. Elman, L. R. Becerra, R. Z. Goldstein, and D. Borsook. The cerebellum and addiction: insights gained from neuroimaging research. *Addict Biology*, 19(3):317–331, May 2014.
- J. C. Niebles and Li Fei-Fei. A hierarchical model of shape and appearance for human action classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2007.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- A. Santamaría-Pang, P. Hernandez-Herrera, M. Papadakis, P. Saggau, and I. A. Kakadiaris. Automatic morphological reconstruction of neurons from multiphoton and confocal microscopy images using 3D tubular models. *Neuroinformatics*, 13(3):297–320, July 2015. ISSN 1559-0089.
- D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, 1992.
- P. Scovanner, S. Ali, and M. Shah. A 3-dimensional SIFT descriptor and its application to action recognition. In *Proceedings of the 15th ACM International Conference on Multimedia*, pages 357–360, New York, NY, USA, 2007.



- Y. Shaham, U. Shalev, L. Lu, H. de Wit, and J. Stewart. The reinstatement model of drug relapse: history, methodology and major findings. *Psychopharmacology*, 168(1–2):3–20, July 2003.
- M. Slaker, L. Churchill, R. P. Todd, J. M. Blacktop, D. G. Zuloaga, J. Raber, R. A. Darling, T. E. Brown, and B. A. Sorg. Removal of perineuronal nets in the medial prefrontal cortex impairs the acquisition and reconsolidation of a cocaine-induced conditioned place preference memory. *The Journal of neuroscience*, 35(10):4190–4202, Mar 2015.
- M. L. Slaker, J. H. Harkness, and B. A. Sorg. A standardized and automated method of perineuronal net analysis using *Wisteria floribunda* agglutinin staining intensity. *IBRO Rep*, 1:54–60, Dec 2016.
- B. A. Sorg, S. Berretta, J. M. Blacktop, J. W. Fawcett, H. Kitagawa, J. C. Kwok, and M. Miquel. Casting a Wide Net: Role of Perineuronal Nets in Neural Plasticity. *The Journal of neuroscience*, 36(45):11459–11468, Nov 2016.
- B. H. Toyama and M. W. Hetzer. Protein homeostasis: live long, won’t prosper. *Nature Reviews Molecular Cell Biology*, 14(1):55–61, Jan 2013.
- T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: A survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3):177–280, July 2008.
- M. C. Van den Oever, B. R. Lubbers, N. A. Goriounova, K. W. Li, R. C. Van der Schors, M. Loos, D. Riga, J. Wiskerke, R. Binnekade, M. Stegeman, A. N. Schoffelmeer, H. D. Mansvelter, A. B. Smit, T. J. De Vries, and S. Spijker. Extracellular matrix plasticity and GABAergic inhibition of prefrontal cortex pyramidal cells facilitates relapse to heroin seeking. *Neuropsychopharmacology*, 35(10):2120–2133, Sep 2010.
- L. van der Maaten. t-SNE, 2018. <https://lvdmaaten.github.io/tsne>, Last acces: July 2018.
- L. J. P. van der Maaten and G. E. Hinton. Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, November 2008.
- S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014. ISSN 2167-8359. doi: 10.7717/peerj.453.
- V. N. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, Inc., New York, 1998.
- D. Vazquez-Sanroman, M. Carbo-Gas, K. Leto, M. Cerezo-Garcia, I. Gil-Miravet, C. Sanchis-Segura, D. Carulli, F. Rossi, and M. Miquel. Cocaine-induced plasticity in the cerebellum of sensitised mice. *Psychopharmacology (Berl.)*, 232(24):4455–4467, Dec 2015a.
- D. Vazquez-Sanroman, K. Leto, M. Cerezo-Garcia, M. Carbo-Gas, C. Sanchis-Segura, D. Carulli, F. Rossi, and M. Miquel. The cerebellum on cocaine: plasticity and metaplasticity. *Addiction biology*, 20(5):941–955, Sep 2015b.
- D. B. Vazquez-Sanroman, R. D. Monje, and M. T. Bardo. Nicotine self-administration remodels perineuronal nets in ventral tegmental area and orbitofrontal cortex in adult male rats. *Addiction biology*, 22(6):1743–1755, Nov 2017.
- A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/>, 2008.
- P. Vinukonda. A study of the scale-invariant feature transform on a parallel pipeline. Master’s thesis, Department of Electrical and Computer Engineering, Louisiana State University, 2011. [https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=3720&context=gradschool\\_theses](https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=3720&context=gradschool_theses), Last access: July 2018.
- Y. Wan, F. Long, L. Qu, H. Xiao, M. Hawrylycz, E. W. Myers, and H. Peng. BlastNeuron for automated comparison, retrieval and clustering of 3D neuron morphologies. *Neuroinformatics*, 13(4):487–499, October 2015.
- J. W. Wright and J. W. Harding. Contributions of matrix metalloproteinases to neural plasticity, habituation, associative learning and drug addiction. *Neural plasticity*, 2009:579382, 2009.
- Y. X. Xue, L. F. Xue, J. F. Liu, J. He, J. H. Deng, S. C. Sun, H. B. Han, Y. X. Luo, L. Z. Xu, P. Wu, and L. Lu. Depletion of perineuronal nets in the amygdala to enhance the erasure of drug memories. *The Journal of Neuroscience*, 34(19):6647–6658, May 2014.
- J. Yu, Z. Qin, T. Wan, and X. Zhang. Feature integration analysis of bag-of-features model for image retrieval. *Neurocomputing*, 120:355–364, 2013.
- D. Zhang, S. Liu, Y. Song, D. Feng, H. Peng, and Weidong Cai. Automated 3D soma segmentation with morphological surface evolution for neuron reconstruction. *Neuroinformatics*, January 2018.
- W. L. Zhao and C. W. Ngo. Flip-invariant SIFT for copy and object detection. *IEEE Transactions on Image Processing*, 22(3):980–991, March 2013.