

Classification of rat behavior with an image-processing method and a neural network

J. B. I. ROUSSEAU

Utrecht University, Utrecht, The Netherlands

P. B. A. VAN LOCHEM

Noldus Information Technology, Wageningen, The Netherlands

and

W. H. GISPEN and B. M. SPRUIJT

Utrecht University, Utrecht, The Netherlands

Video recordings of 11 rats were digitized at five frames per second, and parameters describing the shape and the position of the rat were calculated. The behavior displayed by the rats was observed by an experienced observer. Separate neural networks were trained and validated, using the data for each individual rat. The neural networks correctly classified an average of 76.53% of the frames in the validation set and 98.18% of the frames in the training set. A single neural network was trained with 6 rats and validated with 5 rats. The neural network correctly classified 63.74% of the frames in the validation set and 82.85% of the frames in the training set.

To quantify and qualify the effects of experimental manipulations, observers or researchers often measure the overt, postural behavior of a subject or the activity a subject performs. To reduce the complexity of the registration, a selection of behaviors is made from the total number of behaviors that a subject can perform. This list of behavioral categories is known as the ethogram. The measurements of these categories, often expressed in units of time (duration) or in frequency of occurrence, are extensively used in multiple research areas, such as pharmacology, agricultural science, and biomedical research. The postures, transition of postures, or series of postures that constitute a behavioral element have been described for several species. These publications are used as a standard within a field to form a consensus on the definition of a behavioral category. Previous publications have described the typical postures that constitute the behavioral categories of the rat (*Rattus norvegicus*; Timmermans, 1978) or the mouse (*Mus musculus*; Van Abeelen, 1963). In the case of rats, the postures of a single animal (Renner & Seltzer, 1991; Timmermans, 1978), the ontogenical development of these postures (Geisler, Westerga, & Gramsbergen, 1993), and the postures of two socially interacting animals (Grant & Mackintosh, 1963; Timmermans, 1978) have been described.

Human observers record the majority of the behavioral measurements described in the literature, either from direct observation or from videotape. The observations are stored on such media as check sheets, audiotapes, and computer disks. The need for automated behavior registration is prompted by the acknowledgment of the limited capacity of human observers for registering behavior accurately. Registration of behaviors with short duration, of behaviors that alternate rapidly, or of observations of long duration is beyond the registration capacity of most human observers. These limitations of human observation capacity lead to methodological problems, such as low intra- and interobserver reliability and consequent low test/retest reliability, and they limit the validity of the comparison of results obtained at different laboratories or over a large time span at one laboratory.

From the early days of behavioral research, methods were developed that allowed automated quantification of behavior, such as capacitance measuring devices (Stoff, Stauderman, & Wyatt, 1983; Tomkins & O'Donovan, 1981), infrared beams and photobeam detectors (Delbarre, Dumas, & Guionniere, 1970; Makanjuola, Hill, Maben, Dow, & Ashcroft, 1977), Doppler shift radar (Marsden & King, 1979), touch panels (Kao, Shaw, Young, & Jan, 1995), and weight gauges (Gapenne, Simon, & Lannou, 1990; Schlingmann, van de Weerd, Baumans, Remie, & van Zutphen, 1998). In recent years, the use of image-processing hardware and software for behavior registration has gained popularity. Such techniques are used to quantify such behaviors as movement in an open field (Bonatz, Steiner, & Huston, 1995; Cools, Ellenbroek, Gingras, Engbersen, & Heeren, 1997; Dai & Carey, 1994), orientation behavior in Morris mazes (Gallagher, Burwell, &

This work was supported by Eureka Grant EU88011. The authors thank all the participants in the Eureka project who participated over the past years for their valuable discussions on rat behavior and behavior registration. Correspondence concerning this article should be addressed to B. M. Spruijt, Rudolf Magnus Institute for Neurosciences, Department of Medical Pharmacology, Utrecht University, P.O. Box 80040, 3508 TA Utrecht, The Netherlands (e-mail: spruijt@med.uu.nl).

Burchinal, 1993; Wolfer & Lipp, 1992), and the social interactions between 2 rats (Arakawa & Ikeda, 1991; Sams-Dodd, 1995; Spruijt, Hol, & Rousseau, 1992). Such methods only register the spatial position (x -, y -coordinates) of the animal by calculating the spatial coordinates of the center of gravity of the group of pixels representing the animal. These coordinates are subsequently used to calculate parameters that quantify distance moved, the time spent in certain areas of the observation area, and the path shape.

Alternatively, other groups have developed image-processing systems that qualify the ongoing behavior, using the whole digital representation of the animal (Kernan, Mullenix, & Hopper, 1988; Spruijt & Gispen, 1983). These systems calculate a large number of parameters describing the postures of a rat and use these to classify behavior with statistical classification or rule-based techniques. A promising alternative classification technique, which has been researched extensively in the past decade, is the classification using artificial neural networks implemented on standard computers. Neural networks are particularly successful in labeling classes from noisy and fuzzy data and are generally able to correctly generalize from examples used during the training to examples never presented before. The behavior of animals contains fuzzy transitions from one behavioral class to another, and various behavioral classes can occur irrespective of the orientation of the animal relative to the observer. These considerations possibly make traditional classification methods, such as statistical classification or rule-based classification, less suitable for the classification of behavior.

Neural networks have been successfully used for classification problems in such diverse fields as ECG classification (Simon & Eswaran, 1997), recognition of vocalizations of fallow deer (Reby et al., 1997), and pistachio nut sorting (Ghazanfari, Kusalik, & Irudayaraj, 1997). Neural network classification of digital images has also been applied to behavioral observations of pigs (Shao, Xin,

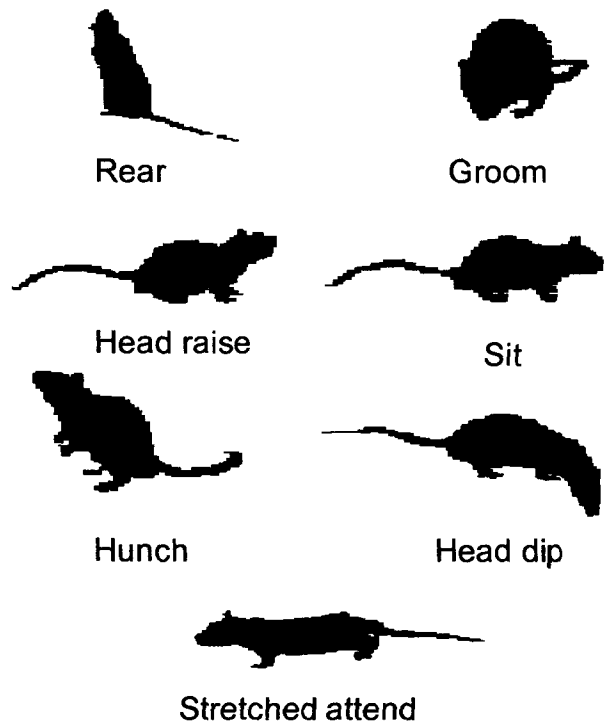


Figure 1. A number of typical examples of the postures corresponding to the behavioral classes used to train the neural network. Classes that consist of a number of consecutive postures such as walking and stationary rotation, are not displayed. The examples for sit and head raise are taken from two consecutive frames from the same videotape.

& Harmon, 1997). A neural network quantified whether the pigs were huddling or spread out, thus indicating the comfort of the animals in their pens. Neural network techniques have also been applied to behavior classification of rats, using data from a modified Stoelting electronic activity monitor (Gonzalez & Arnaldo, 1993).

An automated system using artificial neural network must address a number of issues to successfully classify the behavior of animals. The primary concern is to devise and select relevant features of the digital representation of the animal by which to discriminate the various behavioral classes. The second issue is to select a suitable neural network architecture and learning parameters. These selections include the number of layers, the number of neurons per layer, and the optimal learning function. The time-dependent nature of behavior and the modification over time must also be attended to in either the feature or the architecture. Finally, the generalization ability of neural network classification from one animal to an animal not used in the training has to be assessed.

The objective of this study was to address some of the issues mentioned above. The temporal nature of behavior is accounted for by presenting the parameters calculated from the digital representation of the animal over three consecutive time points. The applied architecture

Table 1
Description of the Behaviors (Ethogram)
as Scored by the Human Observer

Behavior	Description
Sit	The rat is immobile.
Walk	The rat displaces itself.
Head dip	The rat dips its head over the edge of the observation platform.
Head raise	The rat raises its head; the point of its nose is above the back.
Stretched attend	The body of the rat is stretched and immobile.
Groom	The rat licks and rubs its paws on its body.
Rear	The rat stands on its hind legs, and the back is straight.
Hunch	The rat stands on its hind legs while the back is bent.
Stationary rotation	The rat rotates the front part of the body without moving the hind legs.
Missing frame	The digitalization program misses a frame, owing to excessive random noise.

and learning rule of the neural network are a multilayer feed-forward network and standard backpropagation with a momentum term. The generalization is addressed by using two data sets. In one, the parameters belonging to a single rat are split at random in half; the first half is used to train the neural network, and the second half is used to validate the classification of the network. A separate network is trained for each animal. In the second data set, half of all the animals are assigned to the training set, and the other half to the validation set. A single network is trained, to assess the generalization capability over animals of the neural network.

MATERIALS AND METHOD

Eleven male rats (UWU/CPB-Wistar) bred at the GDL, Utrecht University (227 ± 28 gram) were used. The animals were placed on a nonreflecting black metal plate with a 50×50 cm area. The plate was placed on a box at a height of 33 cm from the floor. The plate was placed against a black background. A standard monochrome CCD camera (Sony AVC-D5CE) was placed on a tripod at a distance of 96 cm from the back of the plate, at a height of 70 cm. The camera was slightly inclined downward. Simultaneously, recordings were made from directly above the platform, but the data from this angle were not used. The recordings were performed under red light and lasted for 10 min. The observations were recorded with a standard VCR (Sanyo VHR-S700ES). The recordings were digitized with the EthoVision software at a sample rate of five frames per second over a period of 8 min. The resulting digital movies

could be replayed with a custom-developed computer program. All the software was developed in cooperation with Noldus Information Technology (Wageningen, The Netherlands).

An experienced human observer scored the digital recordings of 2,400 frames per rat according to the ethogram described in Table 1. The behaviors scored are head dip (the rat bends its neck over the edge of the observation platform), sit, walk, head raise (the rat raises its head so that the point of the nose is above the back), stretched attend (the rat stands on three or four legs and is slightly inclined forward; the head is bent forward), groom (the rat grooms itself; this category includes all kinds of grooming, such as flank, head, or anogenital grooming), rear (the animal stands supine on two legs, and the back is straight), hunch (the animal stands on two legs, and the back is slightly curved), stationary rotation (the rat rotates without actually moving), fall (the rat falls off the observation platform), and missing frame (a frame could not be digitized, owing to random noise). A number of typical examples of the postures of the rats for some behavioral categories are displayed in Figure 1.

By using a custom-designed software program, the parameters listed in Table 2 were calculated for each frame of the digital movies. Figure 2 shows an example of the location of a number of the parameters on a digital representation of a rat. The tail of the rat was eliminated from the image of the rat by means of image-processing techniques (erosion and dilation) for the calculation of the shape parameters (centralized moments, scalar invariant, and circularity). The relative positions of the nose and tail points within the rectangle bounding the rat were also calculated. Some frames could not be digitized properly, owing to random noise in the video recordings. These frames were coded with a missing value. The frames during which the rat was not visible on the recording, because it had fallen off the observation platform, were also scored

Table 2
Parameters Calculated From the Digital Image of the Rat

Parameter	Explanation	Parameter	Explanation
FRAME NUMBER	Sequential number of the frame in a movie	CIRCULARITY	Closeness (variance) of the digital shape to the circular shape
AREA	Area of the rat in pixels (equals CM00)	PRINCIPAL AXIS	Axis through the digital shape with the minimal dispersion
CENTERX	<i>x</i> ordinate of the center of gravity projected on the medial axis	ANGLE1	Angle of the line between the center of gravity and the nose point
CENTERY	<i>y</i> ordinate of the center of gravity projected on the medial axis	ANGLE2	Angle of the line between the center of gravity and the tail point
TAILBODYX	<i>x</i> ordinate of the base of the tail	ANGLE3	Angle of the lines between the nose point/center of gravity and the tail point/center of gravity
TAILBODYY	<i>y</i> ordinate of the base of the tail	LENGTH1	Length of the line between the center of gravity and the nose point
NOSEX	<i>x</i> ordinate of the nose	LENGTH2	Length of the line between the center of gravity and the tail point
NOSEY	<i>y</i> ordinate of the nose	LENGTH3	Length of the line between the tail point, the center of gravity, and the nose point
BOXX1	<i>x</i> ordinate of the upper left corner of the bounding rectangle	DM0	Displacement of the center of gravity over one frame
BOXY1	<i>y</i> ordinate of the upper left corner of the bounding rectangle	DM1	Displacement of the center of gravity over two frames
BOXX2	<i>x</i> ordinate of the lower right corner of the bounding rectangle	DM2	Displacement of the center of gravity over three frames
BOXY2	<i>y</i> ordinate of the lower right corner of the bounding rectangle	DM3	Displacement of the center of gravity over four frames
CM11 to CM40	Centralized moment 1,1 to centralized moment 4,0 (Masters, 1994)	DM4	Displacement of the center of gravity over five frames
SI1 to SI11	Centralized moments, scale and orientation independent (Kernan, Mullenix, & Hopper, 1988)		

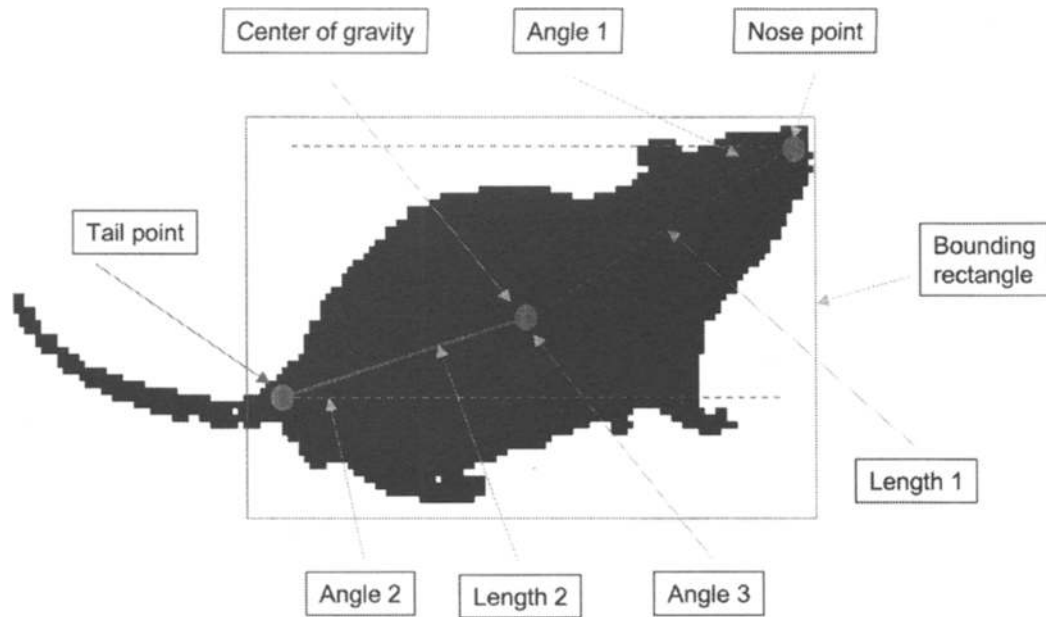


Figure 2. An example of some of the parameters calculated from the digital representation of a rat. The parameters displayed include the nose and tail points and the center of gravity. The various angles and the lengths between these points are calculated as well. The shape parameters (centralized moments) are not depicted in this figure.

with a missing value. All the parameters were normalized over all the observations with a z-transformation. Each behavior was coded in a separate data column in a 1-of-C format—that is, each behavioral class was coded with a 1 if that behavior occurred and with a 0 if the behavior did not occur.

The observations were split into two data sets. (1) The observation of each individual rat was split at random into a training and a validation set, with equal frequencies of each of the behavioral categories (Data Set 1). A separate neural network was trained for each of the observations. (2) Six of the observations were randomly assigned to a data set used to train a single neural network, and the five remaining observations were used to assess the performance of this neural network (Data Set 2). The comparison of the classifications for both data sets is indicative of the ability of a neural network classifier to generalize over observations.

Using the Stuttgart Neural Network Simulator (Institute of Parallel and Distributed High-Performance Systems, University of Stuttgart), a feed-forward neural network was constructed with three layers of 165 input neurons, 80 hidden neurons, and 10 output neurons. The input layer was fed with the parameters of three consecutive frames. Each of the neurons in the output layer corresponded to one of the behavioral classes (see Figure 3). The network was trained during 200 cycles, during which each of the frames from the training sets was presented to the network in random order. A file containing the state of the neural network was saved each 5 training cycles, and the classification on the training and validation sets was calculated. The learning function was backpropagation with a momentum term (Rumelhart, Hinton, & Williams, 1986). The values of the two parameters of the learning function were learning rate (α), .2, and momentum term, .01.

After training, the performance of the network was assessed, using the neural network with the optimal classification. For further analysis, in the case of Data Set 1, the neural nets obtained after the complete 200 training cycles were used. In the case of Data Set 2,

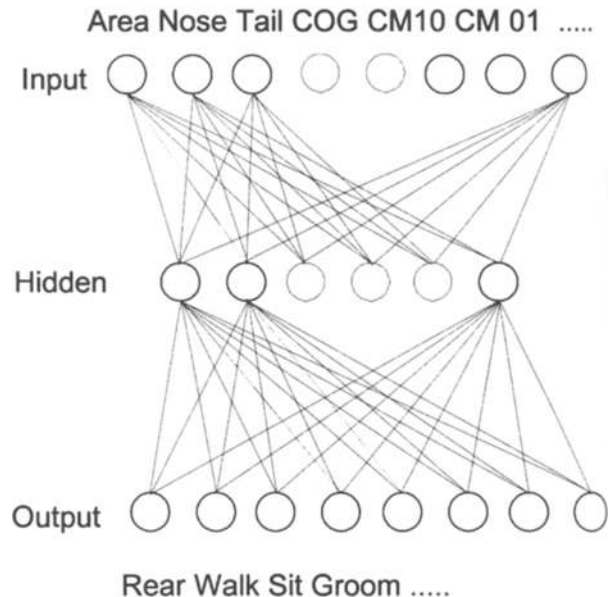


Figure 3. An example of a neural network architecture. The network is divided into three layers: the input, hidden, and output layers. Each neuron (unit) in the input layer receives as input the z-transformed value of one of the parameters. This value is fed forward to each unit in the hidden layer and is modified according to the weight of the connection. Similarly, the activations of the units in the hidden layer are fed to the final (output) layer. Each neuron in the output layer corresponds to one behavioral category. After training, the behavior classified by the neural network i is considered to be the output neuron with the highest activation.

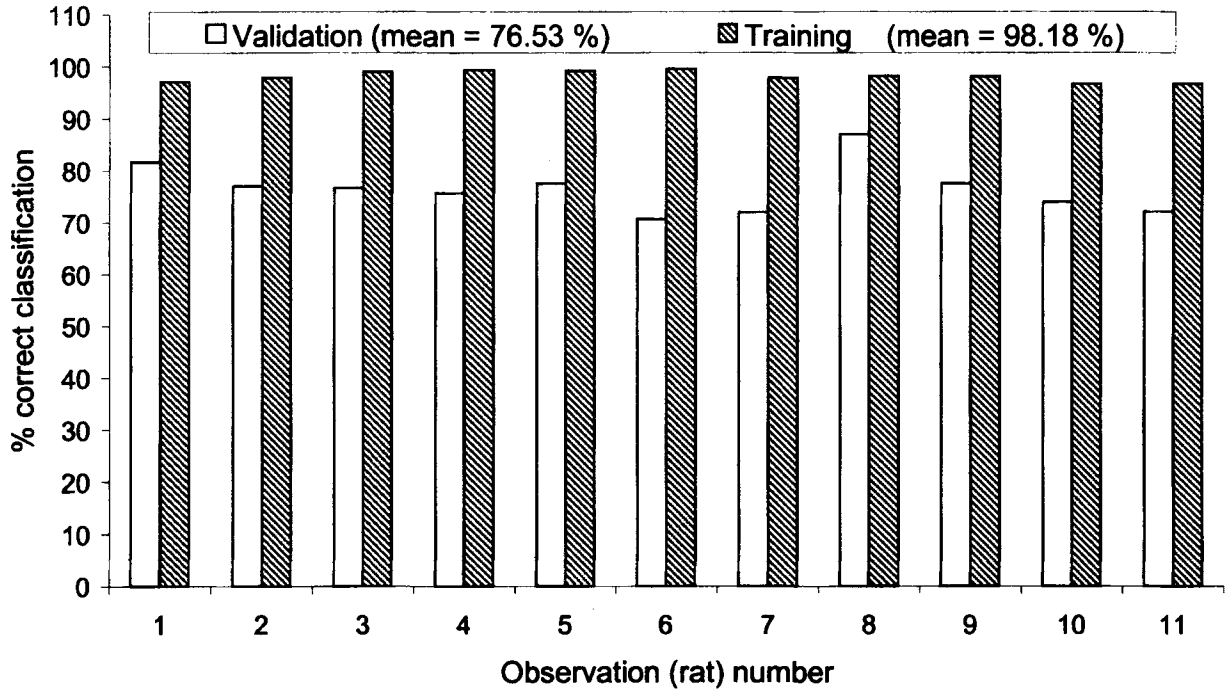


Figure 4. The classification of behavior (Data Set 1). A separate neural network was trained on each observation. This observation was split at random into a training set and a validation set. The average for correctly classified behaviors was 76.53% for the training sets and 98.18% for the validation sets.

the neural network with the maximal percentage of correctly classified behaviors was selected from the 40 available neural networks. Each of the data points in the validation set was presented to the network, and the activation values of each of the output neurons were written to a file. The behavioral category coded by the neuron with the highest activation level was considered to be the behavior classified by the neural network. This method is also known as *winner takes all*. Behaviors classified by the neural network with a duration of one frame were disregarded, and the previous behavior was considered to be still active.

The performance of the neural network classification can be assessed by using two approaches. The first approach compares the classification on the basis of single individual frames of 0.2-sec duration. The classification of the human observer and the neural network were compared in confusion matrices that included two mea-

asures of interobserver classifications appropriate for k-class situations: normalized mutual information (NMI; Forbes, 1995) and Cohen's kappa (κ ; Cohen, 1960). This approach was used for both Data Sets 1 and 2. In the second approach, the durations and frequencies of behaviors derived from the human and the neural networks were compared with a one-way analysis of variance (ANOVA), using SPSS for Windows Version 6.1.4.

RESULTS

Data Set 1

The percentage of correctly classified frames for Data Set 1 is displayed in Figure 4. The average percentage of correctly classified frames is 76.53% for the validation

Table 3
Confusion Matrix for a Single Rat of Data Set 1 (Rat 8) for the Validation Set

Neural Network Classification	Human Classification									
	Head Dip	Rear	Sit	Head Raise	Walk	Groom	Hunch	Stretched Attend	Missing Frame	Stationary Rotation
Head dip	25	0	10	0	0	0	0	0	0	0
Rear	0	0	2	0	1	3	0	0	1	0
Sit	6	0	561	4	15	3	0	0	0	4
Head raise	0	0	9	32	1	0	0	0	0	0
Walk	0	0	28	0	99	8	0	2	0	6
Groom	0	0	10	0	1	281	0	0	0	1
Hunch	0	0	1	0	1	1	0	0	0	0
Stretched attend	0	0	1	0	2	0	0	7	0	0
Missing frame	0	0	0	0	0	0	0	0	13	0
Stationary rotation	0	0	24	0	3	7	0	0	0	22
Total classified by human	31	0	646	36	123	303	0	9	14	33
Total percentage correctly classified	80.64	-	86.84	88.89	80.49	92.74	-	77.78	92.86	66.67

Table 4
Confusion Matrix for the Validation Set of Data Set 2 ($n = 5$)

Neural Network Classification	Human Classification									
	Head Dip	Rear	Sit	Head Raise	Walk	Groom	Hunch	Stretched Attend	Missing Frame	Stationary Rotation
Head dip	89	0	573	1	26	0	0	0	2	16
Rear	0	248	27	14	15	1	44	0	0	1
Sit	76	13	5,189	52	703	96	24	4	0	119
Head raise	1	24	366	217	67	0	27	4	0	12
Walk	4	1	659	5	1,413	22	7	0	0	70
Groom	0	0	216	0	5	151	0	0	0	7
Hunch	0	0	57	37	17	0	33	0	0	7
Stretched attend	0	0	119	3	18	0	0	0	0	0
Missing frame	0	0	1	0	0	0	0	0	100	0
Stationary rotation	13	0	576	2	153	22	2	0	0	195
Total classified by human	183	298	7,783	331	2,417	292	137	8	102	427
Total percentage correctly classified	48.63	83.22	66.67	65.56	58.46	51.71	24.09	0.00	98.04	45.67

set and 98.18% for the training set. An example of a confusion matrix is given in Table 3 (Rat 8, the neural network with the best result). The frequency of classification of a certain behavior by the human observer is tabulated in the horizontal direction, whereas the frequency of the neural network is tabulated in the vertical direction. Similar classifications of behaviors by the human observer and the neural network are to be found along the diagonal of the matrix. All nonzero values off this diagonal represent disagreements in the classification. The total percentage of correctly classified behaviors is given in the bottom row. For example, the total frequency of the groom behavior, as classified by the human observer, was 303. The neural network classified this behavior similarly in

281 cases and differently in 22 cases. Of these 22 cases, 7 were classified as stationary rotation by the neural network.

Data Set 2

The confusion matrix of the classification for Data Set 2, with different animals used for training and validating the neural network, is displayed in Table 4. The values in Tables 3 and 4 represent the classification of frames with a duration of 0.2 sec. The human classifications are plotted in the horizontal direction; the neural network classifications are plotted in the vertical direction. Nonzero values off the diagonal represent disagreements between the human classification and the neural network classification. The total percentage of correct classifica-

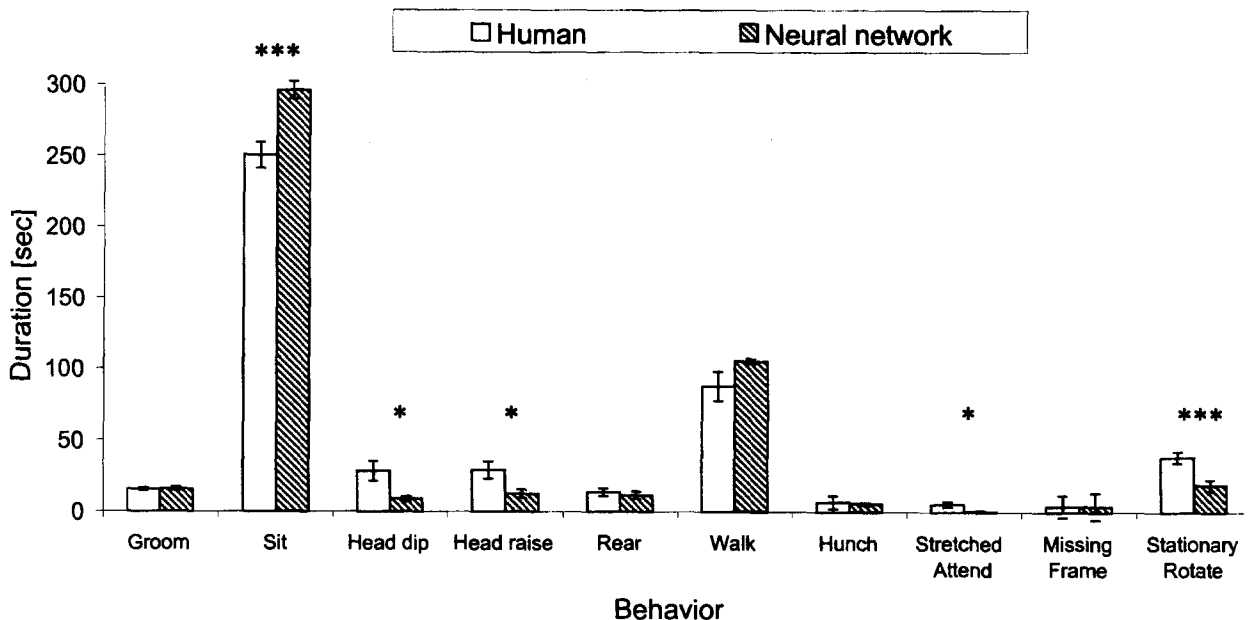


Figure 5. The total duration of behaviors for the validation set ($n = 5$), as classified by the human observer and the neural network (Data Set 2). The mean duration in seconds and the SEM are displayed. The neural network tends to overestimate the duration of the sit and walk behaviors, whereas it underestimates the durations of head dip, head raise, stretched attend, and stationary rotation ($*p < .05$, $***p < .001$). See also Table 3.

tions per behavioral category is displayed in the bottom row of the table. The overall percentage of correct classifications for the 0.2-sec frames for Data Set 2 was 82.85% ($\kappa = .729$, $NMI = .531$) for the training set and 63.74% ($\kappa = .411$, $NMI = .239$) for the validation set.

The results of the assessment of the neural network performance expressed in duration and frequencies are displayed in Figures 5 and 6. The total duration of each behavior category for the validation set ($n = 5$) is displayed in Figure 5. A one-way ANOVA yielded a significant difference for the duration of the following behaviors: sit [$F(8,1) = 19.20$, $p = .002$], head dip [$F(8,1) = 7.30$, $p = .027$], head raise [$F(8,1) = 6.17$, $p = .038$], stretched attend [$F(8,1) = 6.36$, $p = .036$], and stationary rotation [$F(8,1) = 17.29$, $p = .003$]. The other behaviors did not yield significant differences.

The frequency of each behavior is displayed in Figure 6. A one-way ANOVA revealed a significant difference in the frequency of behaviors for sit [$F(8,1) = 367.71$, $p < .001$], head raise [$F(8,1) = 6.44$, $p = .035$], and stationary rotation [$F(8,1) = 21.64$, $p = .001$].

DISCUSSION

The ability of neural networks to correctly classify behaviors is, in principle, demonstrated by the results obtained from Data Set 1, as displayed in Figure 4. The high percentage of correctly classified frames of the validation set (average, 76.53%) indicates that neural networks

are able to correctly classify behavior by using data from the same animal. The comparatively extremely high scores for the training sets (average, 98.18%) could indicate that a certain amount of *memorization* occurred. The data from Data Set 2 indicate that the neural network is able to generalize the classification of behaviors learned from the training set ($n = 6$) to the new observations in the validation set ($n = 5$). The smaller difference between the percentage of correctly classified frames in the validation set and that in the training set for Data Set 1 (63.74% vs. 82.85%) indicates that less memorization occurred than with the data from Data Set 1. When the classification of the neural network is compared with the human classification in terms of duration and frequency of behaviors, a mixed picture emerges. Some behaviors, such as rear and hunch, reveal a near-perfect match between the human and the neural network classifications. This indicates that the input parameters used enabled the network to correctly classify these behaviors. There were differences, however, between the human and the neural network classifications of the duration and frequency of other behaviors (e.g., sitting and stationary rotation). Several possible reasons for the misclassification are discussed below.

The misclassifications of the neural network could be the result either of a misclassification by the human observer of the behaviors which the neural network is trained or of actual misclassifications by the neural network. For instance, the behavior sit, which is the behav-

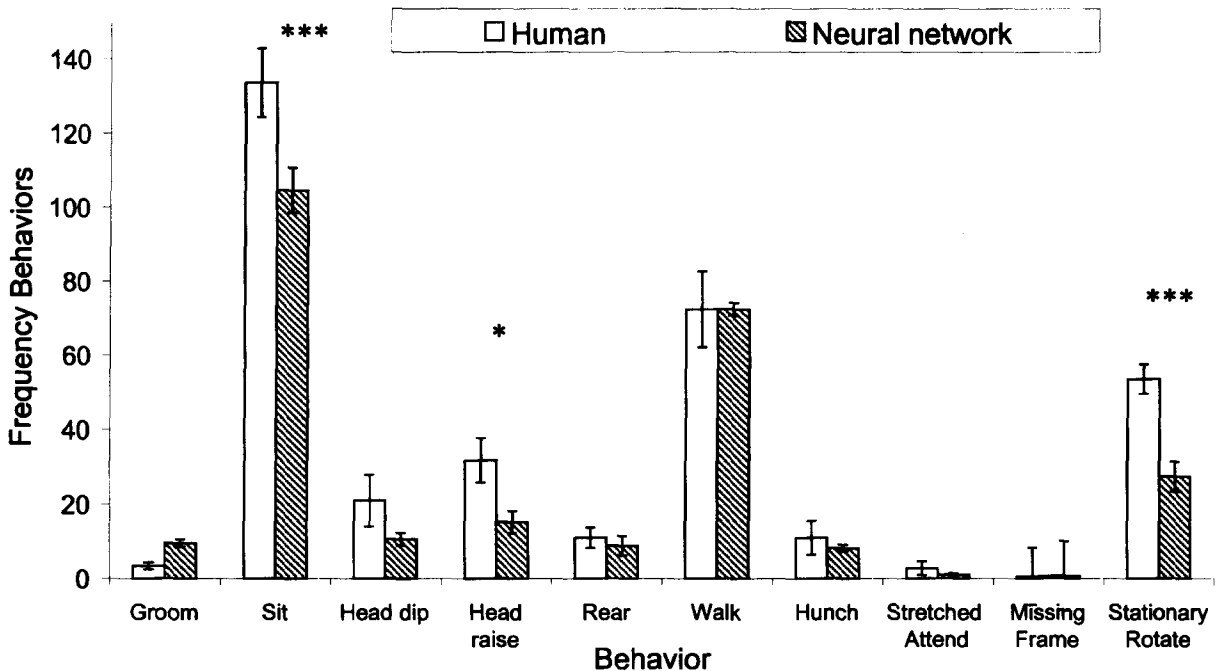


Figure 6. The frequency of behaviors for the validation set ($n = 5$), as classified by the human observer and the neural network (Data Set 2). The mean frequency and SEM are displayed. The neural network tends to underestimate the frequency of the sit, stationary rotation, and head raise behavior (* $p \leq .05$, *** $p \leq .001$).

ior with the highest frequency, is misclassified by the neural network as walk in 659 cases, as compared with the human observer. Conversely, the behavior walk, as classified by the human observer, is classified as sit in 703 cases by the neural network. A human observer has great difficulty in accurately determining the end time-point of the continuous movement of an animal; therefore, it is probable that this is the main source of the error in classification. If the misclassifications were resolved between the behaviors sit and walk, the total percentage of correctly classified behavior would rise to 75.11%. The most accurate method of quantifying the movements of animals is calculating the displacement of the center of gravity of the animal being tracked, using image-processing software and hardware that are similar to those used in this study.

Another possible strategy that could improve the classification is the elimination of irrelevant input parameters. The 55 parameters per frame that were used can be reduced, using techniques from statistical classification theory, such as hierarchical clustering, or techniques derived from artificial neural networks, which are generally known as feature selection techniques (Dash & Liu, 1997). This includes, for example, optimal brain surgery, selectively removing neurons from a network, and genetic algorithms. The generalization ability of neural networks is known to improve when the number of irrelevant input parameters (features) is reduced. Conversely, new features could be devised that could improve the classification of behaviors with large differences in classification between the human and the neural network. The neural network underestimates both the frequency and the total duration of the stationary rotation behavior. A new parameter could be devised that represents the displacements of the nose and tail points over a number of previous frames. This is analogous to the displacements of the center of gravity over five frames, currently used. A further improvement of the classification could be obtained if the number of observations in the training set were to be increased, as compared with the number of observations in the validation set.

To make possible the successful application of neural network classification of behavior in standard tests used in pharmacological research, several factors have to be taken into account. First, the sample of animals used to train a neural network must consist of a large and representative sample of all the possible behaviors that one wants to classify. Similarly, animals of different sizes must be used in the training set, although the possible confounding effects of size can be expected to be counteracted by the scale-invariant parameters used in this study. Further research is also required to assess the ability of neural networks to classify observations recorded from different camera positions relative to the observed animal and under different lighting conditions. Finally, experiments with drugs have to be conducted, to examine whether neural networks reveal drug effects similar to those revealed by human observations.

The fuzzy boundaries between behavioral classes—for instance, the transitions from walk to sit or rear to hunch—result in differences in classification of these behaviors by human observers. To ensure reliable and reproducible behavior classification, observers have to be trained extensively, and periodic evaluations of observers have to be conducted. These difficulties have led to a situation in which these safeguards are implemented only within laboratories, hampering the comparison of results obtained at different laboratories. Comparison among laboratories is not impossible, but a substantial effort has to be made to ensure similar behavior classifications. For example, the International Program on Chemical Safety collaborative study on neurobiological screening methods organized a 2-day training session to train a representative from each of the participating laboratories to classify the behaviors studied (Moser, Becking, MacPhail, & Kulig, 1997). The use of standard classification tools would eliminate such cumbersome protocols. Automated classification also allows continuous observations over long period of times, thus eliminating the effects of observer fatigue and allowing unexpected late-onset drug effects to be registered.

This study indicates that neural networks seem to offer a flexible classification framework that could be allowed to accommodate different behavioral categories. Further research is required to compare the interobserver reliabilities of automated classification techniques with the interobserver reliabilities of human observers before such techniques become widely accepted. However, the acceptance of such behavior classification techniques by the biomedical research community would eliminate some of the methodological deficiencies that behavioral observations are currently subject to.

REFERENCES

- ARAKAWA, O., & IKEDA, T. (1991). Apomorphine effect on single and paired rat open-field behavior. *Physiology & Behavior*, **50**, 189-194.
- BONATZ, A. E., STEINER, H., & HUSTON, J. P. (1995). Video image analysis of behavior by microcomputer: Categorization of turning and locomotion after 6-OHDA injection into the substantia nigra. *Journal of Neuroscience Methods*, **22**, 13-26.
- COHEN, J. (1960). A coefficient of agreement for nominal scales. *Educational & Psychological Measurement*, **20**, 37-46.
- COOLS, A. R., ELLENBROEK, B. A., GINGRAS, M. A., ENGBERSEN, A., & HEEREN, D. (1997). Differences in vulnerability and susceptibility to dexamphetamine in Nijmegen high and low responders to novelty: A dose-effect analysis of spatio-temporal programming of behaviour. *Psychopharmacology*, **132**, 181-187.
- DAI, H., & CAREY, R. J. (1994). A new method to quantify behavioral attention to a stimulus object in a modified open-field. *Journal of Neuroscience Methods*, **53**, 29-34.
- DASH, M., & LIU, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, **1**, 131-156.
- DELBARRE, B., DUMAS, G., & GUIONNIERE, M. (1970). An automatec open field method. *Psychopharmacologia*, **18**, 227-230.
- FORBES, A. D. (1995). Classification-algorithm evaluation: Five performance measures based on confusion matrices. *Journal of Clinical Monitoring*, **11**, 189-206.
- GALLAGHER, M., BURWELL, R., & BURCHINAL, M. R. (1993). Severity of spatial learning impairment in aging: Development of a learning

- index for performance in the Morris water maze. *Behavioral Neuroscience*, **107**, 618-626.
- GAPENNE, O., SIMON, P., & LANNOU, J. (1990). A simple method for recording the path of a rat in an open field. *Behavior Research Methods, Instruments, & Computers*, **22**, 443-448.
- GEISLER, H. C., WESTERGA, J., & GRAMSBERGEN, A. (1993). Development of posture in the rat. *Acta Neurobiologiae Experimentalis*, **53**, 517-523.
- GHAZANFARI, A., KUSALIK, A., & IRUDAYARAJ, J. (1997). Application of a multi-structure neural network (MSNN) to sorting pistachio nuts. *International Journal of Neural Systems*, **8**, 55-61.
- GONZALEZ, L. P., & ARNALDO, C. M. (1993). Classification of drug-induced behaviors using a multi-layer feed-forward neural network. *Computer Methods & Programs in Biomedicine*, **40**, 167-173.
- GRANT, E. C., & MACKINTOSH, J. H. (1963). A description of the social postures of some laboratory rodents. *Behaviour*, **21**, 246-259.
- KAO, S. D., SHAW, F. Z., YOUNG, M. S., & JAN, G. J. (1995). A new automated method for detection and recording of animal moving path. *Journal of Neuroscience Methods*, **63**, 205-209.
- KERNAN, W. J., JR., MULLENIX, P. J., & HOPPER, D. L. (1988). Pattern recognition of rat behavior. *Pharmacology, Biochemistry & Behavior*, **27**, 559-564.
- MAKANJUOLA, R. O., HILL, G., MABEN, I., DOW, R. C., & ASHCROFT, G. W. (1977). An automated method for studying exploratory and stereotyped behaviour in rats. *Psychopharmacology*, **52**, 271-277.
- MARSDEN, C. A., & KING, B. (1979). The use of Doppler shift radar to monitor physiological and drug induced activity patterns in the rat. *Pharmacology, Biochemistry & Behavior*, **10**, 631-635.
- MASTERS, T. (1994). *Signal and image processing with neural networks: A C++ sourcebook* (1st ed.). New York: Wiley.
- MOSER, V. C., BECKING, G. C., MACPHAIL, R. C., & KULIG, B. M. (1997). The IPCS collaborative study on neurobehavioral screening methods. *Fundamental Applied Toxicology*, **35**, 143-151.
- REBY, D., LEK, S., DIMOPOULOS, I., JOACHIM, J., LAUGA, J., & AULAGNIER, S. (1997). Artificial neural networks as a classification method in the behavioural sciences. *Behavioural Processes*, **40**, 35-43.
- RENNER, M. J., & SELTZER, C. P. (1991). Molar characteristics of exploratory and investigatory behavior in the rat (*Rattus norvegicus*). *Journal of Comparative Psychology*, **105**, 326-339.
- RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1986). Learning representations by back-propagating errors. *Nature*, **323**, 533-536.
- SAMS-DODD, F. (1995). Automation of the social interaction test by a video-tracking system: Behavioural effects of repeated phencyclidine treatment. *Journal of Neuroscience Methods*, **59**, 157-167.
- SCHLINGMANN, F., VAN DE WEERD, H. A., BAUMANS, V., REMIE, R., & VAN ZUTPHEN, L. F. M. (1998). A balance device for the analysis of behavioural patterns of the mouse. *Animal Welfare*, **7**, 177-188.
- SHAO, J., XIN, H., & HARMON, J. D. (1997). Neural network analysis of postural behavior of young swine to determine their thermal comfort state. *Transactions of the ASAE*, **40**, 755-760.
- SIMON, B. P., & ESWARAN, C. (1997). An ECG classifier designed using modified decision based neural networks. *Computers Biomedical Research*, **30**, 257-272.
- SPRUIJT, B. M., & GISPEN, W. H. (1983). Prolonged animal observation by use of digitized videodisplays. *Pharmacology, Biochemistry & Behavior*, **19**, 765-769.
- SPRUIJT, B. M., HOL, T., & ROUSSEAU, J. (1992). Approach, avoidance, and contact behavior of individually recognized animals automatically quantified with an imaging technique. *Physiology & Behavior*, **51**, 747-752.
- STOFF, D. M., STAUDERMAN, K., & WYATT, R. J. (1983). The time and space machine: Continuous measurement of drug-induced behavior patterns in the rat. *Psychopharmacology*, **80**, 319-324.
- TIMMERMANS, P. J. A. (1978). *Social behaviour in the rat*. Unpublished doctoral thesis, Nijmegen, The Netherlands.
- TOMKINS, P. T., & O'DONOVAN, D. J. (1981). A fully automated open-field apparatus incorporating rearing detection. *Physiology & Behavior*, **26**, 741-746.
- VAN ABELEN, J. H. F. (1963). Mouse mutants studied by means of ethological methods: I. Ethogram. *Genetica*, **34**, 79-94.
- WOLFER, D. P., & LIPP, H. P. (1992). A new computer program for detailed off-line analysis of swimming navigation in the Morris water maze. *Journal of Neuroscience Methods*, **41**, 65-74.

(Manuscript received March 10, 1999;
revision accepted for publication November 1, 1999.)