# Performance Analysis of Feature-Based Automated Measurement of Mouse Social Behavioral

Mohammad Muntasir Rahman†*, Mst. Ashrafunnahar Hena‡, Bikash Chandra Singh§,
Mohammad Alamgir Hossain†

*†Department of Computer Science and Engineering,*
*Islamic University, Kushtia, Bangladesh*

*‡Department of Electrical & Electronic Engineering,*
*Islamic University, Kushtia, Bangladesh*

*§Department of Information and Communication Technology,*
*Islamic University, Kushtia, Bangladesh*

*Corresponding author: `muntasir@cse.iu.ac.bd`

***Abstract*—Automated social behavior analysis in the mammalian animal has become an increasingly popular and attractive alternative to traditional manual human annotation with the advancement of machine learning and video tracking system for automatic detection. In this work, we study a framework of how different features perform on the different classifiers to analyze automatic mice behavior. We conducted experiments on the Caltech Resident-Intruder Mouse (CRIM13) dataset, which provides two types of features: trajectory features and spatio-temporal features. With this feature, we train AdaBoost and Random Decision Forest (TreeBagger) classifiers to classify different mouse behaviors to show which features perform best on which classifier. The experimental result shows that the trajectory features are more informative and provide better accuracy than the widely used spatio-temporal features, and AdaBoost classifier shows better performance than the TreeBagger on these features.**

***Index Terms*—Social behaviors recognition, machine learning, trajectory features, spatio-temporal features, classification.**

## I. Introduction

Detecting and classifying the social behavior of experimental animals is an interesting issue in computer vision and neuroscience research. Social behavior analysis is very important for understanding the connection between neural activity and behavior. In order to understand this connection, many exciting methods have been developed over the years [1], [2], [3], [4], [5]. However, the traditional manual visual observation of animal activities takes a lot of time and manpower to analyze social behavior. With the rapid development of machine learning and video surveillance technology, automatic detection of unusual animal activities and behavior analysis have become popular to the researchers [6], [7], [8]. It is very difficult to conduct behavior analysis directly on humans. Therefore, research on animals provides a great opportunity for the development of automatic behavior analysis research.

In this work, we investigate the automatic mouse behavior analysis on the different features extracted from videos in the home care settings. Our main objective is to classify certain social behaviors of mice, such as *'sniff'*, *'attack'*, *'eat'* and *'walk'*. We use

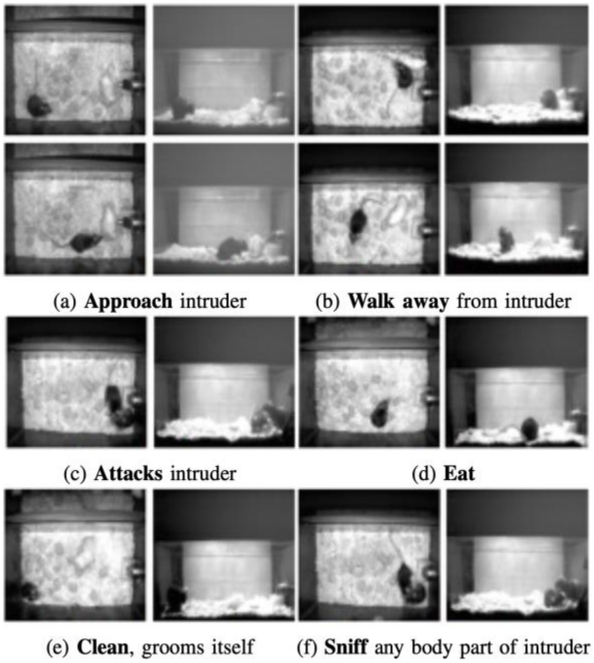


Fig. 1. Example video frames from CRIME13 [9] dataset.

the publicly available Caltech Resident-Intruder Mouse dataset (CRIM13) [9], which contains 237×2 videos (recorded with synchronized top and side view) of pairs of mice engaging in social behavior, divided into 13 different actions. Each video lasts about 10 minutes, a total of 88 hours of videos and 8 million frames. Each video is annotated frame by frame basis by some behavior experts.

The automatic behavior categorization typically requires a classification algorithm to characterize the visual information in the videos [10], [11], [9], [12]. In order to identify different behaviors, examples of labeled features are used to determine the parameters in the classification algorithm. For this task, the CRIM13 dataset provides two types of features, namely *trajectory features* and *spatio-temporal features*. In this work, supervised machine learning techniques AdaBoost [13] and Random Forest (TreeBagger) [14] classifiers are used to analyze and model the extracted feature data for training. We use trajectory features and spatio-temporal features separately, and we found that the weak trajectory features are superior to widely used spatio-temporal features on these classifiers, especially on the AdaBooste classifier.

The remainder of the paper is organized as follows. Section II presents related works. In section III, we formally describe the features of CRIM13 dataset. We conducted the experiment and provide results in section IV. Finally, section V concludes the paper.

## II. Related Work

Social behavior consists of some complex interactions that might be found in all mammalians, including humans. Various techniques have been applied to track social behavior in animals. An early example is [15], in which Khan *et al*. conducted experiments on ant tracking. Most of the techniques are applied to those animals, which have relatively few degrees of freedom in their movements [16], [17], [12], [18], [19], [20], [21], [22], [23], [24], [25]. Recently, the mouse model is being popular in automatic behavior analysis research, because mice are one of the social species that engaging in a high degree of social interactions [26], [27], [28], [29]. Branson *et al.* [30], [31] used a contour changing technique to mice by imaging the cage from a side view, while Pistori *et al.* [32] adopted a particle filtering approach to track the mice from the top view.

In the computer vision literature, a wide variety of behavioral analysis methods requires dual challenges to automatic classification: first, accurately extracting the correct representation from the data, and second, mapping the representation to the correct behavior for activity recognition [33], [34], [35], [36]. HOG/HOF, eSURF, and hierarchical spatio-temporal descriptors were used for feature extraction followed by a classifier in some works [33], [34] or for more complex behaviors analysis [35], [36]. In [37], Chaumont *et al*. proposed a physics-based method to track the position of two mice and monitor their interaction. On the other hand, Burgos *et al*. [9] adopted a machine learning based approach, where the behavior is learned automatically from the given examples. They also proposed a mice behavior analysis dataset called Caltech Resident-Intruder Mouse dataset (CRIM13), which provides a series of general-purpose features, such as trajectory features and spatiotemporal features. In this study, we used the CRIME13 datasets features for machine learning that can automatically detect and classify distinct social behaviors, especially those involving two mice in

TABLE I: Accuracy of each classifier on each behavior.

| Feature Type | Sniff | | Attack | | Eat | | Walk | |
|---|---|---|---|---|---|---|---|---|
| | **AdaBoost** | **R.Forest** | **AdaBoost** | **R.Forest** | **AdaBoost** | **R.Forest** | **AdaBoost** | **R.Forest** |
| WTF 75 | 58.94% | 56.98% | 80.87% | 56.23% | 49.97% | 56.58% | 75.47% | 56.72% |
| WTF 615 | 55.88% | 55.29% | 77.65% | 55.43% | 49.95% | 55.29% | 49.88% | 55.31% |
| WTF 75 + WTF 615 | 59.75% | 56.70% | 72.59% | 56.50% | 49.98% | 56.46% | 66.30% | 56.49% |
| STF Top | 50.69% | 54.72% | 82.33% | 54.72% | 50.13% | 54.72% | 50.91% | 54.72% |
| STF Side | 52.16% | 54.72% | 59.15% | 54.72% | 49.98% | 54.72% | 50.03% | 54.72% |
| STF Top+Side | 53.17% | 54.72% | 73.89% | 54.72% | 50.00% | 54.72% | 50.15% | 54.72% |

1WTF 75: Weak Trajectory Features computed using 75 Frames temporal window.
2WTF 615: Weak Trajectory Features computed using 615 Frames temporal window.
3STF Top: Spatio-temporal Features computed from the TOP videos.
4STF Side: Spatio-temporal Features computed from the SIDE videos.
5AdaBoost: Adaptive boosting classifier.
6R.Forest: Random decision forest classifier.

close and dynamic contacts in their home cage.

## III. Feature Selection

A common trend in automatic behavior analysis is to extract sparse and informative feature points. The use of such features makes the model easier to manage and enhance robustness. In the following sections, we describe the dataset and the features that we used in our experiment.

### *A. Dataset*

In this work, we use the CRIM13 [9] dataset, which consists of 237 videos, each video is about 10 minutes, recorded at 25fps, with a resolution of 640 × 480 pixels, 8-bit pixel depth and monochrome. Each scene uses two fixed synchronized cameras from the top and side views. The video always starts with a male "resident mouse", which is placed alone in the laboratory, and then at some point the second rat "intruder" is introduced into the cage. Therefore, the social interaction starts between the two mice, and finally, the intruder mouse is removed just before the video ends.

There are 12+1 mutually exclusive different behaviors are categorized in the dataset, of which there are 12 behaviors and one last category with no behavior named *other* are annotated carefully. For simplicity, we only use four behaviors in the dataset, namely *'sniff'*, *'attack'*, *'eat'* and *'walk'*. Fig. 1 shows some frames from CRIME13 dataset.

### *B. Features*

For the experiment, we used the spatio-temporal bag of words features and weak trajectory features provided by the CRIME13 [9] dataset. The features are described in the following:

**Spatio-temporal bags of words** are computed using existing methods outlined in [33], [38] by a sliding window centered at the current frame on each video. Spatio-temporal features are two types: one computed from the TOP videos; and another is computed from SIDE videos.

**Weak trajectory features** are computed from the set of positions $x_{m_i}(t)$, $y_{m_i}(t)$ of each mouse $m_i \in [1,2]$ for each top view video frame *t*. Then calculate the position and extract meaningful trajectory information, such as the distance between the mouse, the direction of movement, velocities and accelerations. After that, an algorithm is used to generate weak trajectory features in a similar way to what is done for object detection in [39]. Like the spatio-temporal features, there are two types of weak trajectory features: one is calculated using a 75-frame temporal window, and the other is calculated using a 615-frame temporal window.

## IV. Experiment and Results

Our main goal is to explore the use of supervised machine learning methods to automatically annotate social behaviors. Supervised learning is a method in which classifiers are trained using annotated datasets with the output of the desired classifier. The performance of the classifier is evaluated using a testing set of ground-truth videos that are not used in training. The training set and test set do not overlap and were obtained from separate videos. We use the same error metric defined in [9], where the error metric is calculated as the average of the diagonal of the confusion matrix, and the values of the confusion matrix are the average agreement per frame between annotations for each pair of behaviors. The average per-frame agreement, which is calculated across all frames, measures the similarity between annotations for that pair of behaviors. Finally, when taking the average of the diagonal, we favor classifiers that achieve a high similarity with the ground truth across all behaviors.

We used the spatio-temporal bag of words and weak trajectory features from the CRIM13 [9] dataset to train two supervised learning algorithms, adaptive boosting (AdaBoost) and Random Decision Forest (TreeBagger). Compared with random decision forest, AdaBoost provides the best performance in terms of prediction accuracy and training speed. We trained four social behavior classifiers (*'sniff'*, *'attack'*, *'eat'* and *'walk'*) using features from the weak trajectory and spatio-temporal features. These features contain ~100000 frames, and frames are manually annotated frame-by-frame basis. Finally, we compare which feature and which type of feature are more informative and provide better accuracy. For the AdaBoost classifier, we use a depth 2 tree for each weak classifier. For each behavior, a binary classifier is trained by boosting all training frames with labels that indicate the presence or absence of the behavior. Given $k = 1..K$ behavior types, each of the $k$ binary classifiers will output a confidence $h^k(i) \in R$ for that particular behavior being present in frame *i*. The only two parameters of the binary AdaBoost classifiers are the maximum number of weak classifiers (T) and the number of frames sampled at each training iteration (S). In the experiment, we chose T = 255 and S = 16. Compared with a larger number of frames, sampling a small number of frames can improve performance. As the number of frames increases, the weak classifiers will overfit, resulting in reduced performance. For the random decision forest (TreeBagger) classifier, we chose to generate 500 random decision trees.

Table I shows the experimental result of the two classifiers according to each behavior. From this table, we find that weak trajectory features outperform in both classifiers compared to the spatio-temporal features on the behavior *'sniff'*, *'attack'*, *'eat'* and *'walk'*, which indicate that the weak trajectory features are more informative than spatio-temporal features. Furthermore, adaptive boosting provide the best performance in prediction accuracy on the CRIME13 [9] dataset.

## V. Conclusion

In this work, we study video based animal behavior analysis on mouse. Here we use two types of features form the CRIM13 [9] dataset. We apply Adaboost and random decision forest classifier to each feature to classify four behavior *'sniff'*, *'attack'*, *'eat'*, and *'walk'*. From the experimental results, we found that the weak trajectory features outperform the spatio-temporal features and we also found that the Adaboost classifier performs better than a random decision tree on the features in the CRIME13 [9] dataset. However, due to limited resources, we experimented with only subset of frames of the CRIME13 dataset. Because CRIME13 is the largest and richest behavior dataset, containing over 8 million frames and 12+1 different behavior categories.

## References


[1] F. de Chaumont, E. Ey, N. Torquet, T. Lagache, S. Dallongeville, A. Imbert, T. Legou, A.-M. Le Sourd, P. Faure, T. Bourgeron et al., "Live mouse tracker: real-time behavioral analysis of groups of mice," *bioRxiv*, p. 345132, 2018.

[2] P. K. Thanos, C. Restif, J. R. ORourke, C. Y. Lam, and D. Metaxas, "Mouse social interaction test (most): a quantitative computer automated analysis of behavior," *Journal of Neural Transmission*, vol. 124, no. 1, pp. 3–11, 2017.

[3] S. Belongie, K. Branson, P. Doll´ar, and V. Rabaud, "Monitoring animal behavior in the smart vivarium," in *Measuring Behavior*. Wageningen The Netherlands, 2005, pp. 70–72.

[4] R. Ulrich, S. Dulaney, M. Arnett, and K. Mueller, "An experimental analysis of nonhuman and human aggression," in *Control of Aggression*. Routledge, 2017, pp. 79–111.

[5] G. Lavee, E. eyjolfsRivlin, and M. Rudzsky, "Understanding video events: a survey of methods for automatic interpretation of semantic occurrences in video," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 39, no. 5, pp. 489–504, 2009.

[6] Y. Nie, I. Ishii, K. Yamamoto, T. Takaki, K. Orito, and H. Matsuda, “High-speed video analysis of laboratory rats behaviors in forced swim test,” in *IEEE International Conference on Automation Science and Engineering*, 2008, pp. 206–211.
[7] H. Ishii, M. Ogura, S. Kurisu, A. Komura, A. Takanishi, N. Iida, and H. Kimura, “Development of autonomous experimental setup for behavior analysis of rats,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 4152–4157.
[8] X. Xue and T. C. Henderson, “Video-based animal behavior analysis from multiple cameras,” in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2006, pp. 335–340.
[9] X. P. Burgos-Artizzu, P. Doll´ar, D. Lin, D. J. Anderson, and P. Perona, “Social behavior recognition in continuous video,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 1322–1329.
[10] L. Giancardo, D. Sona, H. Huang, S. Sannino, F. Manag`o, D. Scheggia, F. Papaleo, and V. Murino, “Automatic visual tracking and social behaviour analysis with multiple mice,” *PloS one*, vol. 8, no. 9, p. e74557, 2013.
[11] W. Hong, A. Kennedy, X. P. Burgos-Artizzu, M. Zelikowsky, S. G. Navonne, P. Perona, and D. J. Anderson, “Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 38, pp. E5351–E5360, 2015.
[12] M. Kabra, A. A. Robie, M. Rivera-Alba, S. Branson, and K. Branson, “Jaaba: interactive machine learning for automatic annotation of animal behavior,” *nature methods*, vol. 10, no. 1, p. 64, 2013.
[13] R. E. Schapire, Y. Freund, P. Bartlett, W. S. Lee et al., “Boosting the margin: A new explanation for the effectiveness of voting methods,” *The annals of statistics*, vol. 26, no. 5, pp. 1651–1686, 1998.
[14] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
[15] Z. Khan, T. Balch, and F. Dellaert, “Mcmc-based particle filtering for tracking a variable number of interacting targets,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 11, pp. 1805–1819, 2005.
[16] H. Dankert, L. Wang, E. D. Hoopfer, D. J. Anderson, and P. Perona, “Automated monitoring and analysis of social behavior in drosophila,” *Nature methods*, vol. 6, no. 4, p. 297, 2009.
[17] E. Eyjolfsdottir, S. Branson, X. P. Burgos-Artizzu, E. D. Hoopfer, J. Schor, D. J. Anderson, and P. Perona, “Detecting social actions of fruit flies,” in *European Conference on Computer Vision ECCV*, 2014, pp. 772–787.
[18] K. Branson, A. A. Robie, J. Bender, P. Perona, and M. H. Dickinson, “High-throughput ethomics in large groups of drosophila,” *Nature methods*, vol. 6, no. 6, p. 451, 2009.
[19] H.-Y. Tsai and Y.-W. Huang, “Image tracking study on courtship behavior of drosophila,” *PloS one*, vol. 7, no. 4, p. e34784, 2012.
[20] A. Iyengar, J. Imoehl, A. Ueda, J. Nirschl, and C.-F. Wu, “Automated quantification of locomotion, social interaction, and mate preference in drosophila mutants,” *Journal of neurogenetics*, vol. 26, no. 3-4, pp. 306–316, 2012.
[21] A. Gomez-Marin, N. Partoune, G. J. Stephens, and M. Louis, “Automated tracking of animal posture and movement during exploration and sensory orientation behaviors,” *PloS one*, vol. 7, no. 8, p. e41642, 2012.
[22] K. J. Kohlhoff, T. R. Jahn, D. A. Lomas, C. M. Dobson, D. C. Crowther, and M. Vendruscolo, “The ifly tracking system for an automated locomotor and behavioural analysis of drosophila melanogaster,” *Integrative Biology*, vol. 3, no. 7, pp. 755–760, 2011.
[23] E. I. Fontaine, F. Zabala, M. H. Dickinson, and J. W. Burdick, “Wing and body motion during flight initiation in drosophila revealed by automated visual tracking,” *Journal of Experimental Biology*, vol. 212, no. 9, pp. 1307–1323, 2009.
[24] G. Card and M. Dickinson, “Performance trade-offs in the flight initiation of drosophila,” *Journal of Experimental Biology*, vol. 211, no. 3, pp. 341–353, 2008.
[25] F. W. Wolf, A. R. Rodan, L. T.-Y. Tsai, and U. Heberlein, “High-resolution analysis of ethanol-induced locomotor stimulation in drosophila,” *Journal of Neuroscience*, vol. 22, no. 24, pp. 11 035–11 044, 2002.
[26] E. Grant and J. Mackintosh, “A comparison of the social postures of some common laboratory rodents,” *Behaviour*, vol. 21, no. 3, pp. 246– 259, 1963.
[27] G. Gheusi, R.-M. Bluth´e, G. Goodall, and R. Dantzer, “Social and individual recognition in rodents: methodological aspects and neurobiological bases,” *Behavioural processes*, vol. 33, no. 1-2, pp. 59–87, 1994.
[28] A. Arac, P. Zhao, B. H. Dobkin, S. T. Carmichael, and P. Golshani, “DeepBehavior: A Deep Learning Toolbox for Automated Analysis of Animal and Human Behavior Imaging Data,” *Frontiers in systems neuroscience*, vol. 13, pp.20, 2019.
[29] Z. Zhang, Y. Yang, and Z. Wu, “Social Behavior Recognition in Mouse Video Using Agent Embedding and LSTM Modelling,” *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, 2019, pp. 530-541.
[30] K. Branson and S. Belongie, “Tracking multiple mouse contours (without too many samples),” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conferenc*e on, vol. 1. IEEE, 2005, pp. 1039–1046.
[31] K. Branson, “Tracking multiple mice through severe occlusions,” Ph.D. dissertation, UC San Diego, 2007.
[32] H. Pistori, V. V. V. A. Odakura, J. B. O. Monteiro, W. N. Gonc¸alves, A. R. Roel, J. de Andrade Silva, and B. B. Machado, “Mice and larvae tracking using a particle filter with an auto-adjustable observation model,” *Pattern Recognition Letters*, vol. 31, no. 4, pp. 337–346, 2010.
[33] P. Doll´ar, V. Rabaud, G. Cottrell, and S. Belongie, “Behavior recognition via sparse spatio-temporal features,” in *Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005. 2nd Joint IEEE International Workshop on*. IEEE, 2005, pp. 65–72.
[34] C. T. Hsu, P. Doll´ar, D. Chang, and A. D. Steele, “Daily timed sexual interaction induces moderate anticipatory activity in mice,” *PLoS One*, vol. 5, no. 11, p. e15429, 2010.
[35] H. Jhuang, E. Garrote, X. Yu, V. Khilnani, T. Poggio, A. D. Steele, and T. Serre, “Automated home-cage behavioural phenotyping of mice,” *Nature communications*, vol. 1, p. 68, 2010.
[36] E. Kyzar, S. Gaikwad, A. Roth, J. Green, M. Pham, A. Stewart, Y. Liang, V. Kobla, and A. V. Kalueff, “Towards high-throughput phenotyping of complex patterned behaviors in rodents: focus on mouse self-grooming and its sequencing,” *Behavioural brain research*, vol. 225, no. 2, pp. 426–431, 2011.

[37] F. De Chaumont, R. D.-S. Coura, P. Serreau, A. Cressant, J. Chabout, S. Granon, and J.-C. Olivo-Marin, "Computerized video analysis of social interactions in mice," *Nature methods*, vol. 9, no. 4, p. 410, 2012.

[38] H. Wang, M. M. Ullah, A. Klaser, I. Laptev, and C. Schmid, "Evaluation of local spatio-temporal features for action recognition," in *BMVC 2009- British Machine Vision Conference*. BMVA Press, 2009, pp. 124–1.

[39] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference* on, vol. 1. IEEE, 2001, pp. I–I.