



Towards Automatic Rat's Gait Analysis Under Suboptimal Illumination Conditions

Ana F. Adonias^{1,2}(✉), Joana Ferreira-Gomes^{3,4,5}, Raquel Alonso^{3,4,5},
Fani Neto^{3,4,5}, and Jaime S. Cardoso^{1,2}

¹ Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
anafilipaadonias@gmail.com

² Instituto de Engenharia de Sistemas e Computadores,
Tecnologia e Ciência, Porto, Portugal

³ Instituto de Investigação e Inovação em Saúde,
Universidade do Porto, Porto, Portugal

⁴ IBMC - Instituto de Biologia Molecular e Celular,
Universidade do Porto, Porto, Portugal

⁵ Departamento de Biomedicina, Unidade de Biologia Experimental,
Faculdade de Medicina, Universidade do Porto, Porto, Portugal

Abstract. Rat's gait analysis plays an important role in the assessment of the impact of certain drugs on the treatment of osteoarthritis. Since movement-evoked pain is an early characteristic of this degenerative joint disease, the affected animal modifies its behavior to protect the injured joint from load while walking, altering its gait's parameters, which can be detected through a video analysis. Because commercially available video-based gait systems still present many limitations, researchers often choose to develop a customized system for the acquisition of the videos and analyze them manually, a laborious and time-consuming task prone to high user variability. Therefore, and bearing in mind the recent advances in machine learning and computer vision fields, as well as their presence in many tracking and recognition applications, this work is driven by the need to find a solution to automate the detection and quantification of the animal's gait changes making it an easier, faster, simpler and more robust task. Thus, a comparison between different methodologies to detect and segment the animal under degraded luminance conditions is presented in this paper as well as an algorithm to detect, segment and classify the animal's paws.

Keywords: Computer vision · Gait analysis · Rat · Osteoarthritis

1 Introduction

Osteoarthritis is a common degenerative disease that has a significant incidence worldwide, resulting in a considerable economic and social impact [1].

This disorder damages the joints, structures responsible for the movement of the body, causing pain, stiffness and functional impairment which as a tremendous impact on the patients' routines. Although there are some treatments that aim to relieve osteoarthritis' symptoms, these therapeutic strategies remain unsatisfactory once they fail to relieve pain or they trigger undesirable side-effects. Consequently, several researchers have attempted to develop more efficient drugs for the treatment of osteoarthritis by relying on the use of animal experimentation [2]. Among the animal species used in research, mice and rats are the most widely used models to study the pathophysiological mechanisms underlying osteoarthritis and to test the therapeutic efficacy of targeted drugs [3]. One of the experimental approaches consists on inducing osteoarthritis in the animal, and afterwards administering a drug and quantifying the changes in the behavioral signs of the animal over several time points of the treatment and/or disease progression. Since movement-evoked pain is a common characteristic of osteoarthritis, the affected animal usually modifies its behaviour after the induction of the model as an attempt to protect the injured joint from load while walking, and therefore decrease pain [4]. As consequence, the gait's parameters will vary over time and will differ between a normal (not ill) and an arthritic rodent. The analysis of these differences can be an useful tool to evaluate, through a continuous monitoring, the therapeutic impact of distinct approaches, namely drugs, on osteoarthritis treatment.

The gait changes can be assessed through the use of videos, which can be analyzed manually or automatically. The manual observation and quantification of the gait changes is a laborious and time-consuming intensive task that is prone to high user variability [5]. Currently, there are some commercially available video-based gait systems capable of detecting and identifying the animal's limbs and automatically compute spatial, temporal, kinematic and dynamic gait parameters. However, they are often expensive closed solutions and some even induce stress to the animal by forcing it to walk on a treadmill at a certain speed in a certain trajectory, which might affect the results [6]. Thus, researchers often select different gait pattern descriptors to report in the literature and some opt for the development of a customized system for the acquisition of the videos and their manual analysis. One approach to do so is to carefully examine videos and choose a set of frames with the animal walking (two paws on the platform) and with the animal standing still (four paws on the platform). After selecting the frames, a software like ImageJ¹ may be used to extract the metrics of interest: the footprints are manually selected and a threshold value is defined by the user, above which the number and intensity of pixels are quantified, allowing the comparison of the area and mean intensity applied by each paw [7].

In order to automate this task and make the system more robust, flexible and able to solve problems related not only with the time and cost but also with the reproducibility inherent to the assessment process conducted by the researchers, a demand for an automated system of gait analysis in laboratory animals like rats and mice arises naturally. Moreover, the availability of such

¹ imagej.nih.gov/ij/.

systems opens the possibility to rethink gait analysis itself, since the typical testing time scale can be easily extended, thus diversifying the gait parameters extracted and the context of evaluation. In this way, this paper contributes with the development of an automated, quantitative tool for rat gait assessment using recent computer vision techniques. A comparison between different methodologies to detect and segment the animal under degraded luminance conditions is made and an algorithm to detect, segment and classify the animal's paws is proposed. Furthermore, in order to evaluate the proposed methodologies, it was created a database with the rats' gait and its manual annotation which is made available to the community.

2 Related Work

Advances in the fields of computer vision and machine learning have been providing the detection, recognition and tracking of objects. Since mice and rats are the most used animal models in the studies of the mechanisms related with human diseases, some sophisticated algorithms have been proposed for mice and rats detection and tracking to ease and optimize time-consuming and laborious tasks usually done manually. The majority of these studies are, however, related with mouse behaviour analysis having a different goal and a different setup from the one addressed in this work [8–11]. On those studies, the camera records the animal from above in normal light conditions.

As previously mentioned, there are some commercially available video-based gait systems which combine video-tracking technology with image analysis methodologies to quantify and characterize mice and rats gait [12–16]. Besides all the commercial systems quoted, it is also worth to mention the work of Mendes *et al.* who developed the MouseWalker, an open source software which aims to evaluate mice gait [17]. This work has the same acquisition procedure as the one used in this paper but it has a much more controlled environment since there are not background movements and the trajectory of the animal is restricted to a walkway, being the mouse always within the frame and walking in the same direction. Another work found, authored by Leroy *et al.*, proposes an automated gait analysis for laboratory mice with the same setup but different lighting conditions [18]. In this study, a background subtraction method is applied to segment the mouse and a motion analysis is performed. Unlike the setting of our work, their hardware allows them to detect the footprints by resorting to a color filter, enhancing the pixels that correspond to the region of the paws with a strong red color component.

Nonetheless, as much as is possible to ascertain, there is no work that meets the same degraded luminosity conditions and works robustly in open field platforms.

3 Methodology

3.1 Data Acquisition

Animal Model. The data used for this study was acquired at Faculdade de Medicina da Universidade do Porto, Portugal. All experimental procedures were performed in compliance with all ethical norms required [19] and appropriate measures were taken to minimize pain or discomfort of the animals. All the conditions were assured in order to maintain reliable results without external interference. Under brief anaesthesia, the osteoarthritis model was induced on the right hind knee of adult male rats [7]. In order to have a control of the disease, a similar number of animals were injected with saline, under the same conditions.

CatWalk Test. This paper focuses on data acquired by a custom-made gait analysis system that allows the animal to walk freely on an open field platform located in a darkened room, with no stressors or rewards. The hardware configuration used is based on the CatWalk system, illustrated in Fig. 1 (adapted from [20]). A fluorescent light beam is sent through a glass platform being reflected internally in all the plane with the exception of the areas where the paws are placed, where light is refracted (A); images are recorded by a video camera placed under the glass platform in which animals are allowed to walk freely. The video camera is connected to a computer equipped with video acquisition software (Ulead Video Studio, Fremont, CA) (B); the signal intensity depends on the contact area of the paw with the surface and increases with the pressure applied by the paw (C); an example of the obtained images is represented in (D).

After recording the animal's gait, the analysis of the CatWalk behavior consists on measuring the mean intensity and the contact area of each paw with the glass platform, to evaluate the disability induced by the model over time [7].

3.2 Dataset

The dataset is composed by 15 videos of 4 different animals in different time points of the experience, each one with different duration, averaging about 5 min each. The videos have a resolution of 640 horizontal \times 480 vertical pixels and a frequency of 25 frames per second. All these videos had already been analyzed by an expert researcher, who had already selected the frames of interest (approximately 9 frames per video) and the associated threshold used to compute the area and the mean intensity of the paws in each frame.

In order to evaluate the algorithm's performance, 130 frames, selected by experts, were annotated using LabelMe, an open source graphical image annotation tool. Six classes were used to annotate all the body parts of interest of the rat: the body, the tail and right-hind, left-hind, right-fore and left-fore paws. All these classes were annotated using polygons. The detailed masks of the paws (Fig. 2(C)) were obtained by applying the threshold defined by the experts. Rat and tail annotations were thereafter confirmed by an expert researcher.

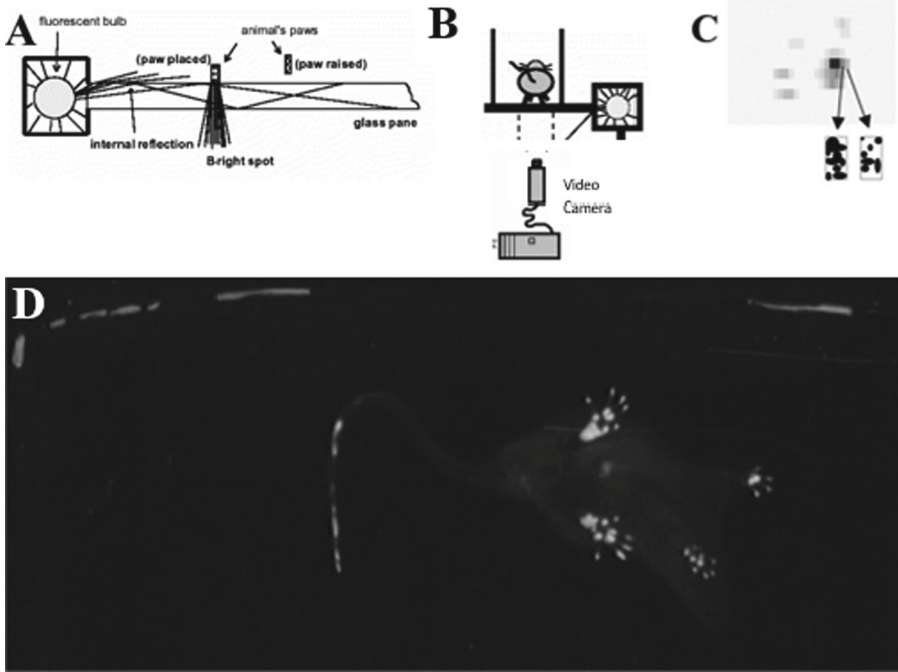


Fig. 1. Principle of the CatWalk setup.

The dataset is available for the community.²

Dataset Heterogeneity. Some artifacts and variations from video to video, and even within each video, hinder the automation process. Since the videos were acquired in the dark, degraded luminosity conditions is a feature of the dataset. Furthermore, although the data was acquired using a static camera in a slightly controlled environment, it is clear the presence of noise in the background, whether due to the interference of researchers with the animal or simply due to contamination of the platform by the animal. Since the platform area is larger than the area captured by the camera, periods when the animal leaves completely the field of view are also present in the dataset. Moreover, as the rat is a non-rigid object, its shape varies during the video sequence, making it more difficult to track. Some natural behaviors such as sitting, lying, scratching, getting up, among others, hinder the detection and segmentation of the paws since these behaviors promote the contact of the body with the platform, enhancing some non-paws derived pixels.

² To access the videos and the annotations, please contact the author Joana Ferreira-Gomes through the email jogomes@med.up.pt.

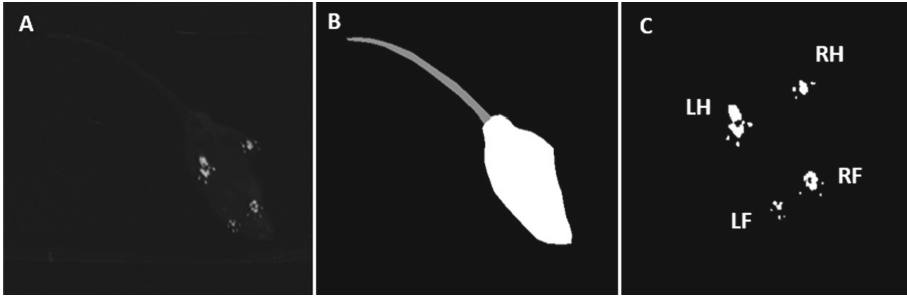


Fig. 2. **A.** input frame; **B.** *White*: ground truth rat’s body; *Gray*: ground truth tail; **C.** ground truth paws: *LH*: left-hind; *RH*: right-hind; *LF*: left-fore; *RF*: right-fore.

3.3 Methodologies

Based on the current state of the art, automated recognition of the animal’s gait pattern can be done by using different computer vision techniques. The generic video-based gait framework is presented in Fig. 3 and it may be briefly described as follows: rat detection and segmentation, paws detection, segmentation and classification and gait features extraction. We did not adopt the current trend of deep learning approaches since the dataset is small, not favouring deep learning solutions, and the strong field knowledge about the setting facilitates the feature-engineering approach.

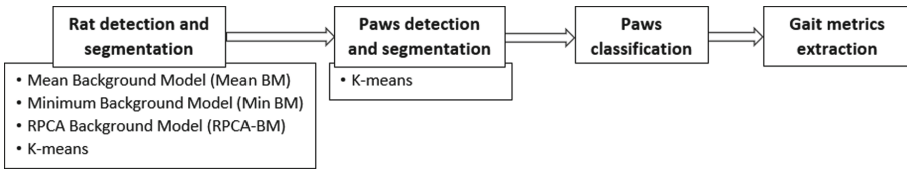


Fig. 3. The generic video-based gait framework.

Rat Detection and Segmentation. The goal of this first step is to robustly detect and segment the body of the animal. This operation is key in the adopted methodology since its success will impact the following steps. Several segmentation techniques were performed, evaluated and compared. Before the segmentation, a pre-processing step was applied to all frames in order to remove noise. This operation results in a smoother image and is performed by computing a simple convolution between each frame (f) and a 3-by-3 box linear kernel (k).

Since one of the dataset characteristics is the presence of background fluctuations, three different approaches were tested to model the background. The first consisted on averaging the whole video sequence (v), which was previously pre-processed (Eq. 1). In the second, because the animal is ideally the only bright

object in motion, the minimum value of each pixel (p_{ij}) in the entire video sequence was computed (Eq. 2).

$$BG(v) = \frac{1}{|v|} \sum_{f \in v} K \otimes f \quad (1)$$

$$BG_{ij}(v) = \min_{f \in v}(p_{ijf}) \quad (2)$$

Both of these offline approaches result in a static background image which is subtracted to each frame to obtain the foreground, the animal in motion (Eq. 3). A threshold operation is performed to obtain the mask of the rat and morphological operations are applied to remove some background noise.

$$FG(v, f) = |f - BG(v)| \quad (3)$$

The third tested approach to model the background was presented by Candès *et al.* [21] and it considers the data matrix (M) a sum of a low-rank component (L) with a sparse component (S). Each column in M corresponds to a frame of the video, properly vectorized. The intuition is that the observed data is the sum of the background information and the objects' data. Since the background is essentially static, it should lead to a low-rank matrix; since the objects tend to be small, they should correspond to a sparse matrix. The low-rank matrix can be then representative of the background, being able of comprising small fluctuations on it and on the other hand, the sparse matrix can formulate the objects in motion, in this case the rat, which can be seen as an outlier. In this method, the principal component pursuit is used to solve the robust principal component analysis (RPCA) following the Eq. 4, where $\|A\|_1$ denotes the vector l_1 norm of the matrix A and $\|A\|_*$ denotes the nuclear norm, the sum of the singular values of A. Morphological operations were applied to the sparse matrix to remove small blobs of the background by filtering it.

$$\min(\|L\|_* + \lambda \|S\|_1) \quad \text{subject to} \quad M = L + S \quad (4)$$

K-means was another of the methodologies used to segment both rat body and paws. This clustering algorithm aims the partitioning of the image into homogeneous regions containing pixels with similar characteristics. The number of clusters was set to three, where one represents the background, other the body of the rat and the third one the paws. This technique was applied to the foreground after a pre-processing algorithm that aims to improve the frames' quality, facilitating the segmentation process. This algorithm was divided in two main steps: a contrast enhancement, achieved by normalizing the frame histogram between zero and its maximum value, followed by a convolution between the enhanced image and a sharpening filter (S) (Eq. 5) which resulted in a well-defined image.

$$S = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & +1 & +1 & +1 & -1 \\ -1 & +1 & +9 & +1 & -1 \\ -1 & +1 & +1 & +1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix} \quad (5)$$

In all the mentioned methodologies the tail was removed by resorting to an opening morphological operation with a square structuring element of size 13×13 . The convex hull of the result of the opening operation was computed, obtaining thus a mask of the body of the rat. This was done to standardize the output of this step since the tail is not always detected and segmented together with the body of the animal.

Paws Detection, Segmentation and Classification. As mentioned before, K-means was the algorithm used to segment the paws. This methodology was performed to the foreground with and without the pre-processing algorithm mentioned above in order to test its efficacy. After having the paws segmented, their classification as left-hind (LH), right-hind (RH), left-fore (LF) and right-fore (RF) must occur. For this purpose, the animal's orientation must be first determined. As the body of the animal can be represented by an ellipse, this shape was used to fit the rat's body mask, the output of the rat segmentation step. Two approaches were performed to obtain the tail point, the extreme point of the ellipse nearest to the tail. The first consisted on a simple subtraction method between the mask obtained before the tail removal in the rat segmentation stage and after. By this way, when the tail is segmented simultaneously with the body there will exist a considerable difference in one of the ellipse's sides. However, as already referred, sometimes the tail is not visible on the frame, not being detected in the first place. In these cases, the difference between both images is not substantial and the tail point is computed by resorting to the second methodology. This second approach computes the minimum difference between the mean of the rat (region inside the ellipse) and its neighborhood in an extension of 250 pixels. A spatial distribution of the minimum points is obtained and each one of the points is associated to the nearest extreme point of the ellipse. The tail point is computed based on the median distance and on the standard deviation between the grouped points and its associated extreme. After determining the tail point, the ellipse is splitted by quadrants which are classified as LH, RH, LF and RF and each paw is assigned to the respective quadrant. This association is made based on euclidean distances.

4 Results

The algorithm was implemented in Python with the exception of the RPCA algorithm, which was implemented in MATLAB being available via the GitHub library of Sobral [22].

Rat Detection and Segmentation. In order to evaluate the different segmentation methods, the intersection level A_f between the ground truth bounding box of the animal's body GT and the tracked bounding box T , was computed to each frame (f) according to the expression $A_f = \text{Area}(GT_f \cap T_f) / \text{Area}(GT_f \cup T_f)$. A known measure to evaluate the performance on a video sequence is to count the number of successful frames whose A_f is larger than a defined threshold. According to the PASCAL criterion [23], this threshold value is 0.5, however, as the use of an unique threshold value may not be representative enough of this measure, the success plot showing the ratios of successful frames at thresholds ranging from 0 to 1 was computed for each of the methodologies (Fig. 4).

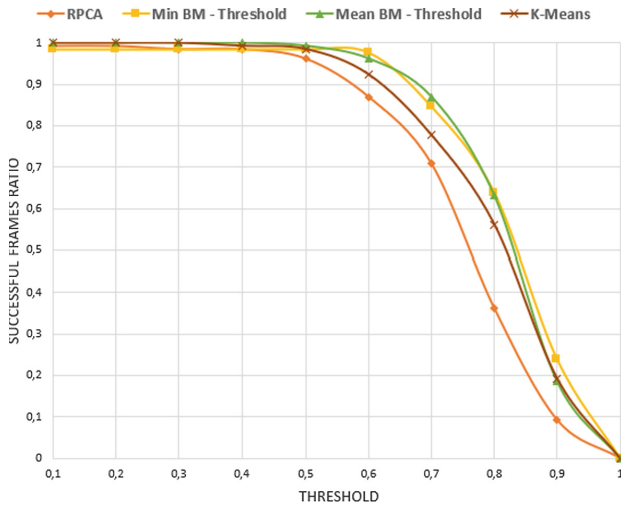


Fig. 4. Rat segmentation successful plot for the 130 frames manually labelled.

In general, all the methodologies are capable of detecting correctly the animal in the majority of the frames. The Mean BM methodology demonstrated to be the most robust, presenting the most similar results with the ground truth. The RPCA methodology showed to be the less vigorous due to the presence of the animal in the low-rank matrix. Since the algorithm was applied to the input video, without any previous processing and computed at intervals of 2000 frames due to computational costs, if the animal stays in the same place for a while, it will be assumed as belonging to the background, hindering its segmentation. The Min BM method showed also good results, being less able, however, of handling large background fluctuations and failing completely the detection of the rat in frames where the researcher interferes with the animal, for example, or when there is a contamination of the platform. Figure 5 shows two examples of the obtained bounding boxes for each methodology.

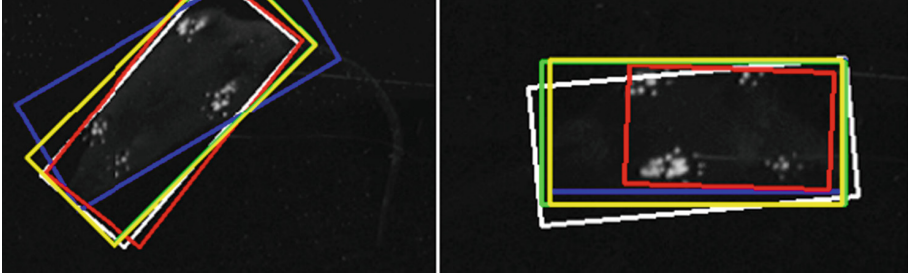


Fig. 5. Rat segmentation examples for the proposed methodologies: **White:** Ground truth; **Blue:** K-means; **Yellow:** Min BM; **Green:** Mean BM; **Red:** RPCA; (Color figure online)

Paws Segmentation. As the ultimate goal of this work is to quantify the rat's gait metrics, such as the area and the mean intensity of the paws, it is important to evaluate paws segmentation by using pixel-based metrics since the algorithm must be meticulous enough to return the referred gait metrics efficiently. While the true positives (TP) return the number of correctly detected paws pixels, the true negatives (TN) give the number of background pixels correctly detected. In contrast, the false positives (FP) and the false negatives (FN) are pixels that are falsely recognized as foreground and background, respectively. From these metrics, the true positive rate (TPR), false positive rate (FPR) and F-score were computed. The TPR reports how frequently the algorithm correctly detects the paws pixels, being given by $TPR = TP / (TP + FN)$. The FPR refers to the number of times the paws are falsely detected and it is given by $FPR = FP / (FP + TN)$. The F-score measure combines TPR and precision (P), given by $P = TP / (TP + FP)$, by their harmonic mean being a measure of the algorithm's accuracy.

The results of paws' segmentation are present in Table 1. Comparing both non pre-processed and pre-processed data, it can be concluded that this step introduces an improvement in the algorithm's performance. A visual example of this pre-processing step can be found in Fig. 6 and two examples of the K-means output applied to the pre-processed frames are present in Fig. 7.

Table 1. Results of paws' segmentation.

Method	TRP	FPR	F-score
K-means (with pre-processing)	$0,85 \pm 0,06$	$0,00 \pm 0,00$	$0,77 \pm 0,09$
K-means (without pre-processing)	$0,61 \pm 0,15$	$0,00 \pm 0,00$	$0,73 \pm 0,10$



Fig. 6. **Right:** input frame; **Left:** output from the pre-processing step.

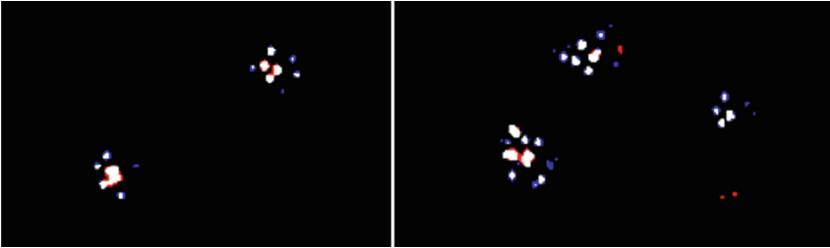


Fig. 7. Paws' segmentation examples: **White:** TP pixels; **Black:** TN pixels; **Blue:** FP pixels; **Red:** FN pixels (Color figure online)

Paws' Classification. The results obtained for the paws' classification are presented in Table 2. In the same way that the pre-processing improves paws' segmentation, it also improves their classification. A frame is only considered well classified when it detects exactly the same paws as the ones present in the GT image in the same region of the image. From the 32% of frames that the classification of frames fails, about one-third is because the algorithm fails in detecting the animal's orientation. This happens when the rat is occluded, when the rat segmentation is not meticulous, when the tail is not well removed or when the animal is near a bright object. The remaining frames' classification fails when the rat has its belly or other part of the body in contact with the platform, enhancing some non-paws derived pixels, or when the two fore paws are very close to each other and the algorithm assumes that it is only one footprint. When comparing the obtained results for each paw, the RH and LH paws' classification is correct in 82% and 85% of the frames, respectively, while RF and LF are successfully classified in 78% and 76% of the frames, respectively.

Table 2. Results of paws' classification.

Method	Frames well classified	Paws well classified
K-means (with pre processing)	68%	81%
K-means (without pre processing)	48%	75%

5 Conclusions

In this paper, an automated method for the assessment of the rat's gait is presented, with the aim of not only providing data that is less prone to user variability but also alleviating the work of scientists who resort to the analysis of the gait of animal models to evaluate the effectiveness of drugs used in the treatment of osteoarthritis. The contributions include an annotated dataset publicly available to the scientific community and a framework capable of detecting the animal, segmenting its body and detecting, segmenting and classifying its paws under degraded lighting conditions.

In between the proposed methodologies to segment the rat, modeling the background and subtract it from each frame gave the best results. Averaging the entire video sequence demonstrated to be the best approach to model the background. Quantitative and visual inspections of the results of paws' segmentation demonstrate a good performance of the algorithm. The proposed algorithm to classify the paws showed promising results.

Future work will focus on the development of an algorithm to select the frames of interest, and on the extraction of more complex gait metrics. We aim to keep extending our dataset and evaluating deep learning methodologies as an alternative to the adopted framework. We also plan the development of a web platform deploying the proposed framework as a free service available to researchers.

References

1. Arthritis Foundation: Arthritis by the numbers. Arthritis Foundation (2018)
2. Benson, R.A., McInnes, I.B., Garside, P., Brewer, J.M.: Model answers: rational application of murine models in arthritis research. *Eur. J. Immunol.* **48**(1), 32–38 (2018)
3. Hedrich, H., Bullock, G.: *The Laboratory Mouse: Handbook of Experimental Animals*. Elsevier Limited, London (2004)
4. Jacobs, B.Y., Kloefkorn, H.E., Allen, K.D.: Gait analysis methods for rodent models of osteoarthritis. *Current Pain Headache Rep.* **18**(10), 456 (2014)
5. Pathak, S.D., et al.: Quantitative image analysis: software systems in drug development trials. *Drug Disc. Today* **8**(10), 451–458 (2003)
6. Lakes, E.H., Allen, K.D.: Gait analysis methods for rodent models of arthritic disorders: reviews and recommendations. *Osteoarthritis Cartilage* **24**(11), 1837–1849 (2016)
7. Ferreira-Gomes, J., Adães, S., Castro-Lopes, J.M.: Assessment of movement-evoked pain in osteoarthritis by the knee-bend and CatWalk tests: a clinically relevant study. *J. Pain.* **9**(10), 945–954 (2008)
8. da Silva Monteiro, J.P., et al.: Automatic behavior recognition in laboratory animals using kinect (2012)
9. De Chaumont, F., et al.: Computerized video analysis of social interactions in mice. *Nat. Methods* **9**(4), 410 (2012)
10. De Chaumont, F., et al.: Live Mouse Tracker: real-time behavioral analysis of groups of mice. *BioRxiv* **16**(2), 345132 (2018)

11. Noldus, L.P.J.J., Spink, A.J., Tegelenbosch, Ruud A.J.: EthoVision: a versatile video tracking system for automation of behavioral experiments. *Behav. Res. Methods Instrum. Comput.* **33**(3), 398–414 (2001)
12. Digigait. <http://www.mousespecifics.com/digigait/>. Accessed 15 Apr 2019
13. CatWalk. <http://www.noldus.com/animal-behavior-research/products/catwalk>. Accessed 15 Apr 2019
14. TreadScan. <http://www.cleversysinc.com/CleverSysInc/wp-content/uploads/2013/02/Datasheet-TreadScan-2017.pdf>. Accessed 15 Apr 2019
15. RunwayScan. <http://www.cleversysinc.com/docs/software-RunwayScan.pdf>. Accessed 15 Apr 2019
16. GaitScan. <http://www.cleversysinc.com/CleverSysInc/wp-content/uploads/2013/02/Datasheet-GaitScan-2010.pdf>. Accessed 15 Apr 2019
17. Mendes, C.S., Bartos, I., Márka, Z., Akay, T., Márka, S., Mann, R.S.: Quantification of gait parameters in freely walking rodents. *BMC Biol.* **13**(1), 50 (2015)
18. Leroy, T., Silva, M., D'Hooge, R., Aerts, J.-M., Berckmans, D.: Automated gait analysis in the open-field test for laboratory mice. *Behav. Res. Methods* **41**(1), 148–153 (2009)
19. Zimmermann, M.: Ethical guidelines for investigations of experimental pain in conscious animals. *Pain* **16**(2), 109–110 (1983)
20. Vrinten, D.H., Hamers, F.F.T.: 'CatWalk' automated quantitative gait analysis as a novel method to assess mechanical allodynia in the rat; a comparison with von Frey testing. *Pain* **102**(1–2), 203–209 (2003)
21. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? *J. ACM (JACM)* **58**(3), 11 (2011)
22. Sobral, A., Bouwmans, T., Zahzah, E.: LRSLibrary: Low-Rank and Sparse Tools for Background Modeling and Subtraction in Videos. *Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing*. CRC Press, Boca Raton (2015)
23. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* **88**(2), 303–338 (2010)