

Re-Identification of Rats with Transfer Learning

Andres Molares-Ulloa, Maria del Rocio Ortega-Femia, Julian Dorado, Nieves Pedreira, and Alvaro Rodriguez

Universidade da Coruña, Department of Computer Science and Information Technology, Faculty of Computer Science, 15071, A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Department of Basic Psychology I, UNED, Ciudad Universitaria, Madrid, Spain.
Correspondence: andres.molares@udc.es

DOI: <https://doi.org/10.17979/spudc.000024.44>

Abstract: The study of animal behavior in laboratory experiments is key in ethology, ecotoxicology, neuroscience and other fields. Although modern studies use computer imaging techniques, current solutions cannot preserve the identity of multiple individuals in social experiments. Thanks to the use of Transfer Learning we seek to overcome this limitations while maintaining the effectiveness of Deep Learning and reducing its computational times. With this technique we achieved promising results in the re-identification of rats after an occlusion process.

1 Introduction

The study of animal behavior is an important part of laboratory experiments conducted in fields such as neuroscience, ecotoxicology, ethology, drug testing and discovery, and many others. These experiments are needed to assess the well-being and health of animals both in nature, where their behavior is impacted by pollution and climate change, and in human facilities, like farms and fish farms, where healthier animals result in lower risks of zoonosis and other pathogens entering the human diet. To conduct this experiments, modern studies rely on computer-imaging techniques, but this doesn't work when studying multiple animals in complex environments is required. Researches still lack techniques to monitor animals and obtain complex behaviors.

Animal behavior experiments are based on the study of behavioral phenotypes that can be defined as evolutionary adaptive traits that emerge as complex patterns of behavior (Cote et al., 2010).

This experiments require large series of tests that run for long time periods of time on large cohorts. In contemporary animal research, laboratories employ video recording and tracking software to simplify analysis, as outlined in (Rodriguez et al., 2018). A tracking application is a computer program designed to assist scientists by extracting motion and location data from video recordings. This is achieved by identifying organisms within each frame of the video and connecting them to the respective animals, enabling the monitoring of their movements. Researchers then utilize this data to gain insights into specific behaviors.

One of the primary limitations of current methods becomes apparent when dealing with multiple individuals. In such scenarios, maintaining the identities of the animals throughout the experiment is crucial. Frequently, animals obscure each other, leading tracking applications to lose sight of the target animals. Consequently, researchers must reassign identities to each animal after they cross paths or become obscured.

The intricacy of this issue is highlighted in (Pérez-Escudero et al., 2014), where they examined a situation involving multiple interacting animals. Even when correctly solving 99% of all crossings, only 11% of the animals were accurately identified after 2 minutes of tracking, due to error propagation. In summary, preserving identity is a complex and computationally intensive task, and only a handful of applications make significant contributions to this field, which remains unsolved.

In prior research (Pérez-Escudero et al., 2014; Rodriguez et al., 2017, 2018), probabilistic texture analysis was utilized to assess the similarity between animals. However, this approach is not suitable for tracking numerous targets over extended periods of time. However, it proves valuable in short experiments involving small groups of animals.

With the rise in popularity of Deep Learning techniques, some authors (Romero-Ferrero et al., 2019; Xu and Cheng, 2017) have employed a deep learning approach known as Convolutional Neural Networks (CNNs). CNNs currently rank among the most potent image classification methods available. In their 2019 work, Romero-Ferrero and colleagues achieved the most robust identity preservation algorithm using this approach. Nevertheless, the computational time required to analyze a standard experiment using this method can be as long as one hour per frame, rendering it impractical for the majority of experiments.

The main goal of this work is the development of a reliable and efficient identity preservation algorithm based on Deep Learning models of image classification, to track multiple identical organisms in social experiments, in scenarios where total occlusion occurs. To achieve this re-identification while maintaining the good results of CNNs but avoiding the long processing times usually required by this approach, we have developed a model based on Transfer Learning. This re-identification model is developed on the Xception architecture.

2 Materials and Methods

2.1 Dataset and its Construction

We have trained and validated our models with a dataset created in a learning trial with five adult laboratory rats of the species (*Rattus norvegicus domestica*) in conditioning cells (Operant conditioning chamber also known as a Skinner box) used for tasks such as teaching an animal to perform certain actions (like pressing a lever) in response to specific stimuli. We have had a dataset composed of color images from 5 different videos. Each video has been segmented to extract a total of 1500 images for each of the 5 individuals studied. These individuals are cataloged as AFH1, AFH2, AFH3, AFH4 and AFH5. In total we have a set of 1500 images of 5 individuals in 5 videos, totaling 37500 images. Each image has a size of $128 \times 128 \times 3$. Figure 1 shows examples of each individual.

In addition, we have applied a normalization process to the images to verify that the models do not base their classification on technical characteristics of the image. This process consists of a standardization (zero mean and standard deviation were passed to each color channel separately from the whole image), and then the output was normalized to the 0-255 range. We have applied this normalization in two different ways. The first one is applied on the total image before performing the individual segmentation and the second one is applied on the already segmented image window. This has resulted in a total of 3 datasets named: Standard (without normalization), Norm and Norm.Window.

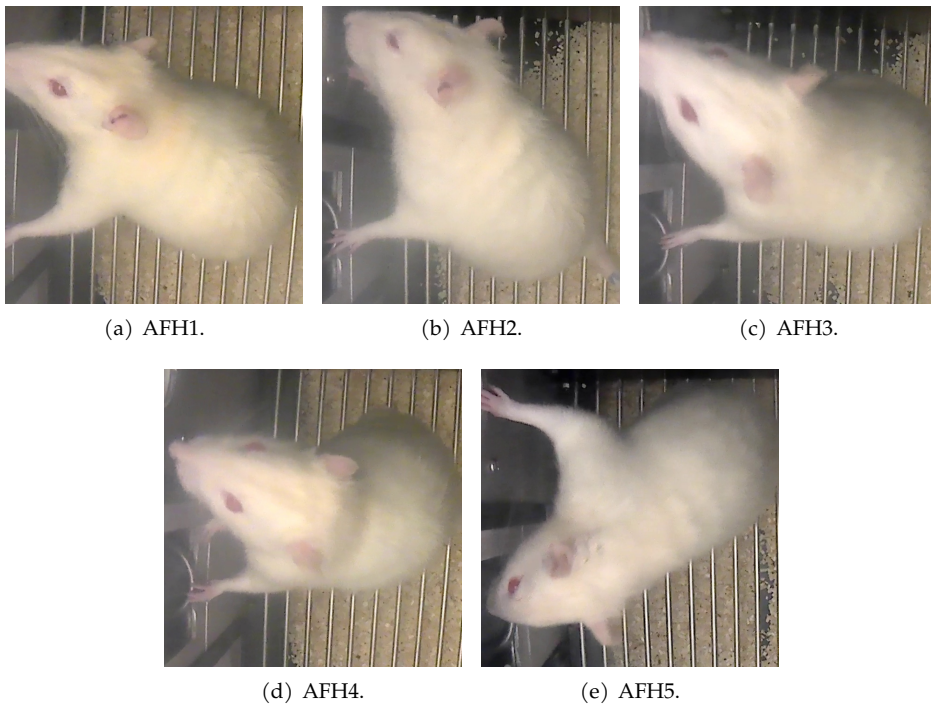


Figure 1: Images of unstandardized individuals.

2.2 Transfer learning and Xception

Transfer learning is a machine learning approach where an initially task-specific model serves as a foundational building block for constructing a new model customized for a different task. This strategy is widely embraced in the domain of deep learning, a subset of machine learning known for its frequent application of such techniques. As convolutional neural networks (CNNs) gained prominence in computer vision, various structured models leveraging CNNs were developed. One notable model is Inception, also referred to as Inception-v1 (Szegedy et al., 2015). The Inception architecture addresses challenges in image classification posed by objects of varying sizes by advocating the use of multiple filters of different sizes at the input stage. Furthermore, it recommends forwarding the output of this module to another Inception module.

The Xception network (Polat, 2021) used in this study can be called an interpretation of the Inception modules. The name Xception also comes from "extreme inception". In Xception architecture, differently from Inception architecture, a convolution operation almost same with depthwise separable convolution is used. This type of convolution contains a depthwise convolution and a pointwise convolution that follows it. In depthwise convolution each filter independently processes only one channel of the input image; and in pointwise convolution, 1x1 dimensional filter iterates every single point of the input. The flows and modules related to the Xception architecture are shown in the Figure 2.

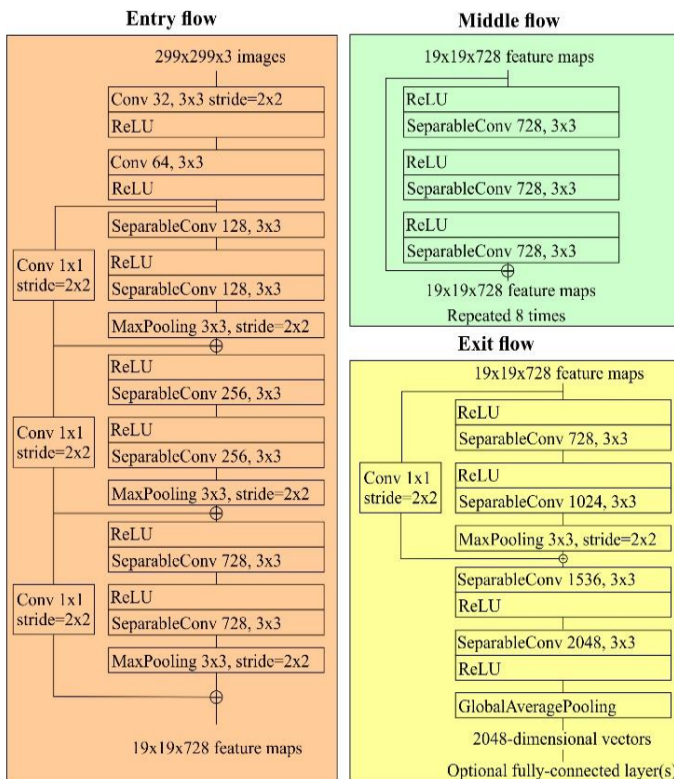


Figure 2: Xception architecture (Polat, 2021).

2.3 Experimentation Setup

To ensure the robustness of the models, we have applied the K-fold cross-validation strategy (Wong (2015)). In this study, the K-fold cross-validation strategy selected will be the 5-fold strategy, where k takes value 5. This is due to the use of 5 videos as a data set. When applying k-fold what we have done is to reserve the images from one video for testing while using the images from the other 4 videos for training. We perform this process until we obtain 5 metrics that robustly show the adequacy of the model. We have added a Data Augmentation layer to the model to increase the variability of the images during the training process. This layer randomly flips, rotates and zooms the training images to increase the generalisability of the system. To determine the values used in the tests, we have carried out an empirical experimentation. The specific configuration of this model is shown in the Table 1.

Table 1: Model configuration parameters

Module	Parameter	Value	
Xception	weights	imagenet	
	input shape	(128, 128)	
Data augmentation	RandomFlip	horizontal and vertical	
	RandomRotation	0.2	
	RandomZoom	0.1	
Output layer	Regular Fit	batch size	(32, 64 and 128)
		epochs	10
		optimizer	Adam(1e-3)
		loss	SparseCategoricalCrossentropy
	Fine tuning	batch size	(32, 64 and 128)
		epochs	5
		optimizer	Adam(1e-6)
		loss	SparseCategoricalCrossentropy

3 Results and Discussion

Table 2 shows the results of applying the model to the 3 datasets studied. We can see that there are no significant differences when applying the different normalization ways. Therefore, we decided to choose the Standard dataset as it has one of the highest means together with the N2 normalization but with a lower standard deviation. We can also observe that there are differences when applying the model depending on the video, with videos 2 and 4 being the most complicated.

Table 2: Model performance

Dataset	Accuracy						
	Video 1	Video 2	Video 3	Video 4	Video 5	Mean	σ
Standard	53.57%	37.51%	50.61%	30.45%	62.07%	46.84%	\pm 12.72%
Norm	54.07%	36.19%	52.93%	28.28%	62.85%	46.86%	\pm 14.18%
Norm.Window	53.51%	29.44%	52.52%	35.92%	53.25%	44.93%	\pm 11.42%

Figure 3 shows the confusion matrix showing the distribution of the results of the model applied to the unnormalised data set. In it we can see how there are differences in the classification of the different images, with those belonging to the AFH3 rat being the most complicated to distinguish and those belonging to the AFH5 rat the simplest. We can also observe how the model tends to confuse between individuals AFH1 and AFH2 and between AFH3 and AFH4 while AFH5 is the most recognizable individual. This may be due to the presence of characteristic features in some of the individuals. This difference in classification between individuals could also be due to behavioral elements reflected in the individual's habitual posture while others have similar routines.

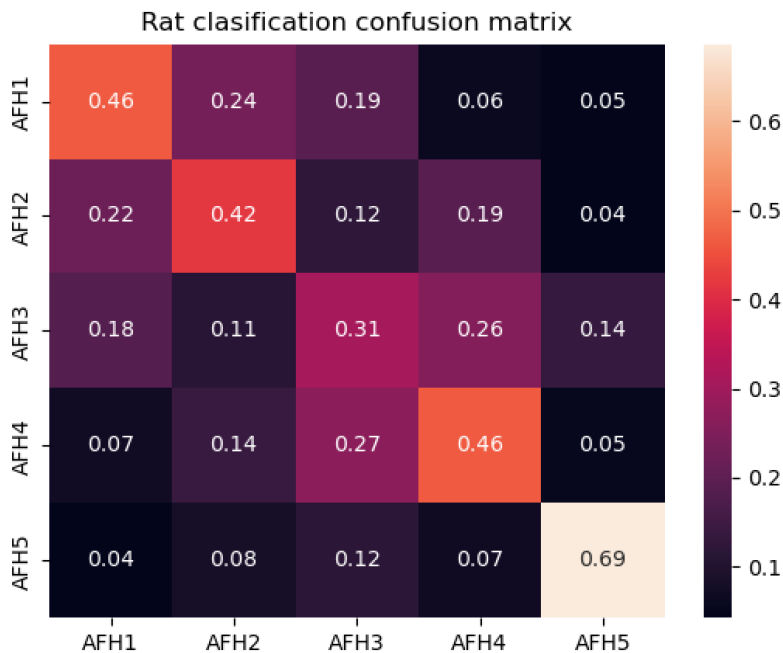


Figure 3: Confusion matrix of the model applied on the Standard dataset.

4 Conclusions

We have defined a performance framework when applying Deep Learning, and in particular Transfer Learning, to classify multiple individuals. We have achieved results of around 47% accuracy when re-identifying up to 5 individuals simultaneously. Being 20% accuracy the threshold from which we started (random classification) we can conclude that we have achieved promising results. We are therefore looking for further improvements in both the quality of the classification and the number of individuals that the system is able to re-identify simultaneously.

We have shown the importance of having a large data set that allows the development of a model with good generalization capacity and avoids biasing the system. Thus avoiding that the model bases its classification on artifacts present in the images and not on the individual's own characteristics. For this, it is important to have several sets of images obtained from different devices, as has been our case.

Acknowledgments

Grant PID2021-126289OA-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe. CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Credit authorship contribution statement

Andres Molares-Ulloa wrote the manuscript, trained and validated the models. **Maria del Rocio** performed and recorded the experiments. **Alvaro Rodriguez** processed the data, created the datasets, and coordinated and supervised the work. All authors revised and edited the manuscript.

Bibliography

- J. Cote, J. Clobert, T. Brodin, S. Fogarty, and A. Sih. Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1560):4065–4076, 2010.
- O. Polat. Detection of covid-19 from chest ct images using xception architecture: A deep transfer learning based approach. *Sakarya University Journal of Science*, 25(3):800 – 810, 2021.
- A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. de Polavieja. idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat Methods*, 2014.
- A. Rodriguez, H. Zhang, J. Klaminder, T. Brodin, and M. Andersson. Toxid: an efficient algorithm to solve occlusions when tracking multiple animals. *Scientific reports*, 7(1):14774, 2017.
- A. Rodriguez, H. Zhang, J. Klaminder, T. Brodin, P. L. Andersson, and M. Andersson. Toxtrac: A fast and robust software for tracking organisms. *Methods in Ecology and Evolution*, 9(3): 460–464, 2018.
- F. Romero-Ferrero, M. G. Bergomi, R. C. Hinz, F. J. Heras, and G. G. De Polavieja. Idtracker. ai: tracking all individuals in small or large collectives of unmarked animals. *Nature methods*, 16(2):179–182, 2019.
- C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
- T. T. Wong. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48, 2015.
- Z. Xu and X. E. Cheng. Zebrafish tracking using convolutional neural networks. *Scientific reports*, 7(1):42815, 2017.