



# Visual Recognition of Great Ape Behaviours in the Wild

Faizaan Sakib and Dr. Tilo Burghardt

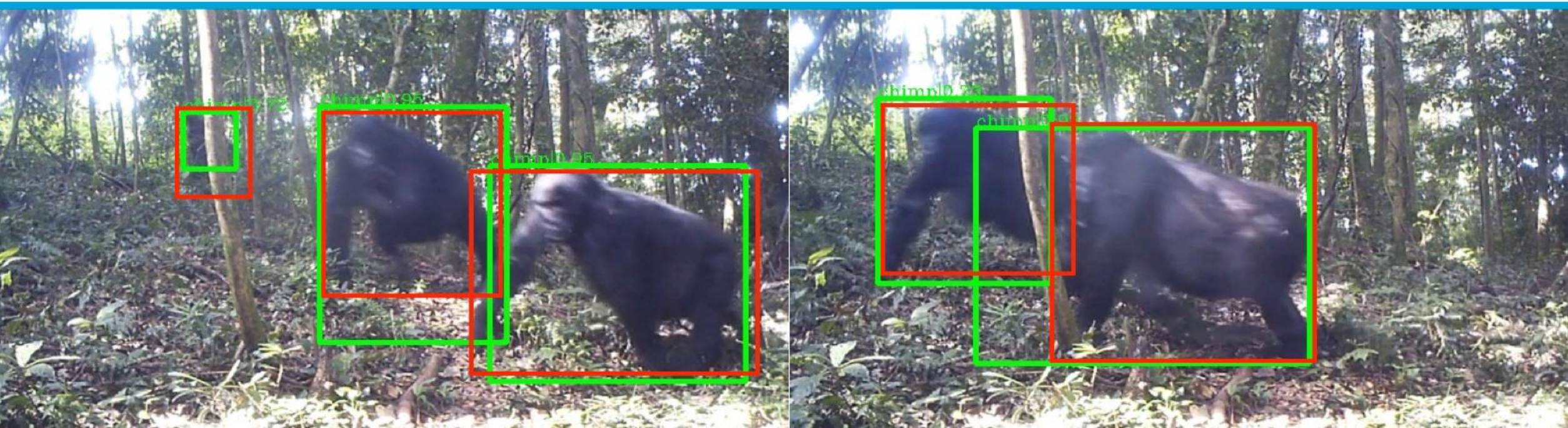
Visual Observation and Analysis of  
Vertebrate and Insect Behaviour Workshop

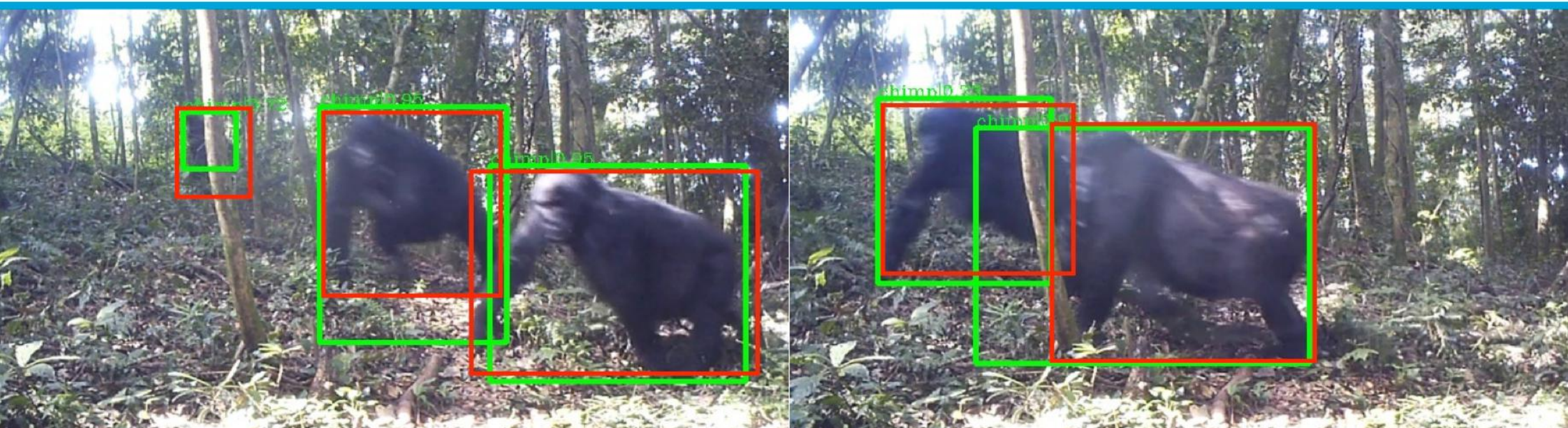
**Motivations**

# Great Apes

- A family of primates made up of eight different species
- Commonly seen as holding the closest genetic relationship to humans
- The majority of great apes are classified as “critically endangered”<sup>[1]</sup>







# Behaviour Recognition

**The**

**PanAfrican2020**

**Dataset**

# Video Data

- In the wild footage taken from the 'PanAfrican' archive [4]
- Subset of 500 videos (~180,000 frames) used
- Features gorillas and chimpanzees found throughout Central Africa
- Shot across multiple locations at varying times of day
- Bounding box annotations in place from previous work



# Annotations

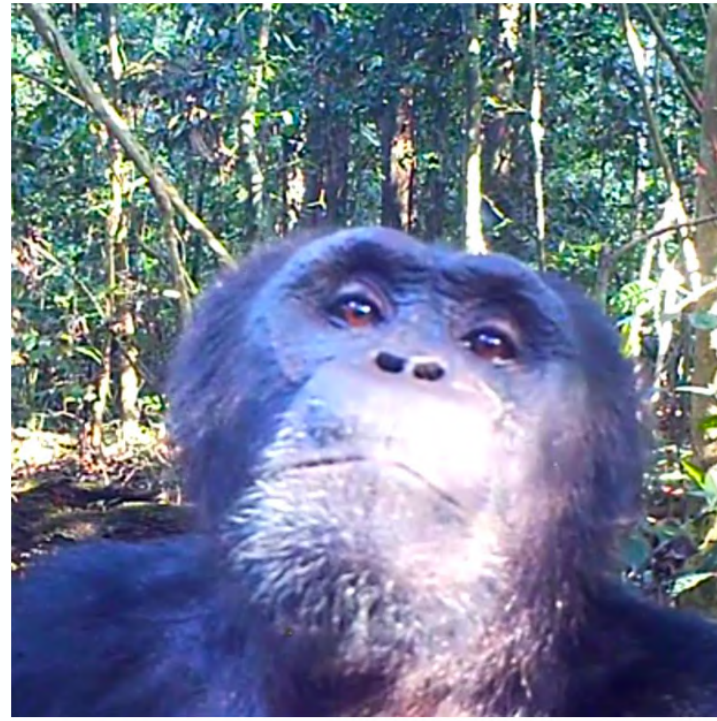
The existing dataset was extended with the following metadata:

- Behaviour
- Identification Number (ID)

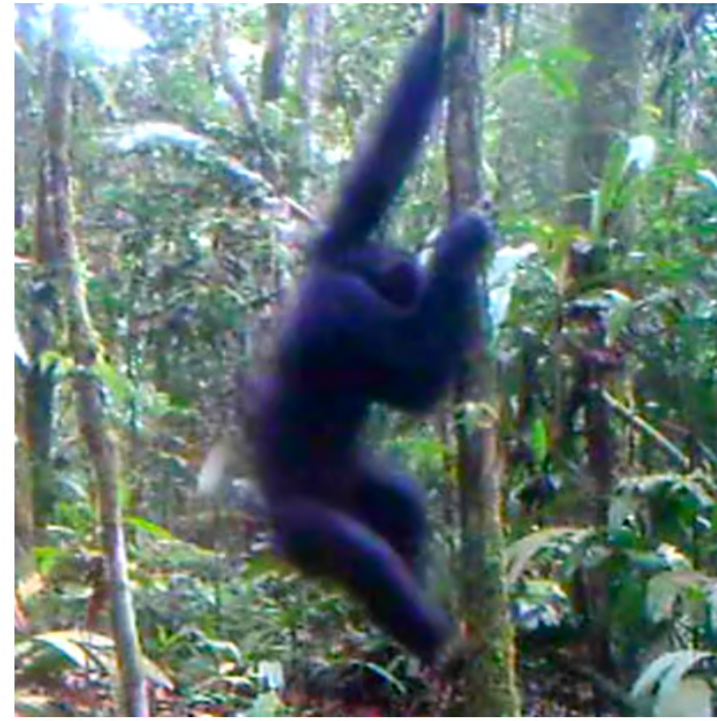
```
1 <annotation>
2   <videoname>1DtGa13tFZ</videoname>
3   <frameid>81</frameid>
4   <size>
5     <height>404</height>
6     <width>720</width>
7     <depth>3</depth>
8   </size>
9   <object>
10    <category>Great Ape</category>
11    <name>chimpanzee</name>
12    <id>0</id>
13    <activity>walking</activity>
14    <bndbox>
15      <xmin>0</xmin>
16      <ymin>276</ymin>
17      <xmax>82</xmax>
18      <ymax>396</ymax>
19    </bndbox>
20  </object>
21 </annotation>
```



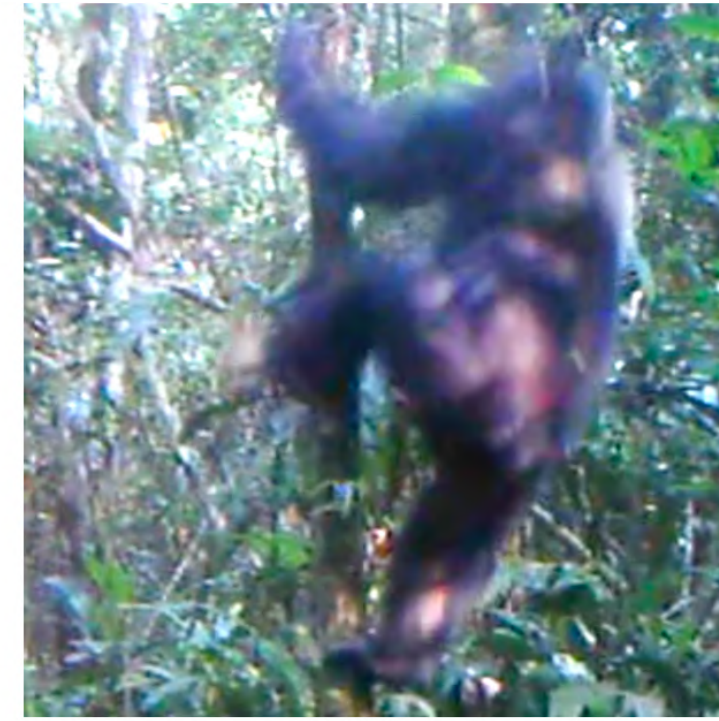
**Behaviours**



camera\_interaction



climbing\_down



climbing\_up



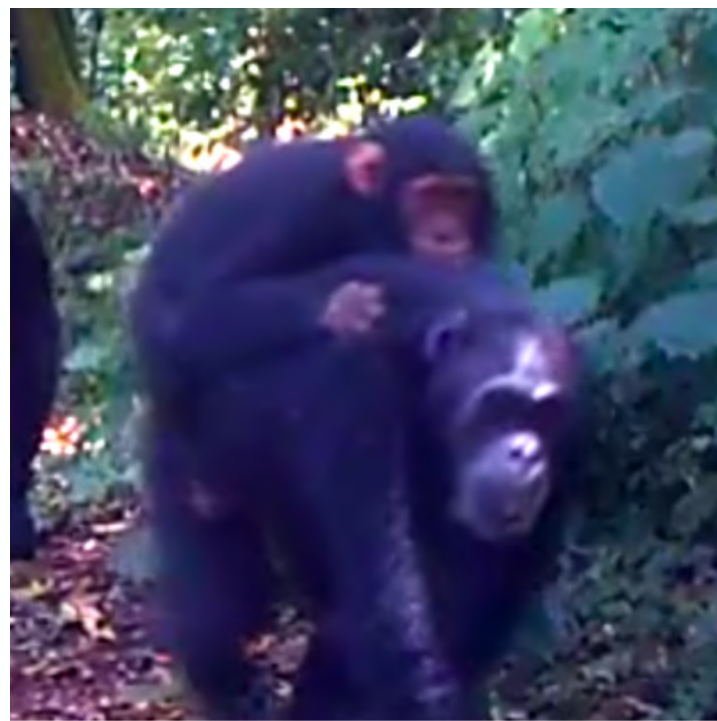
hanging



running



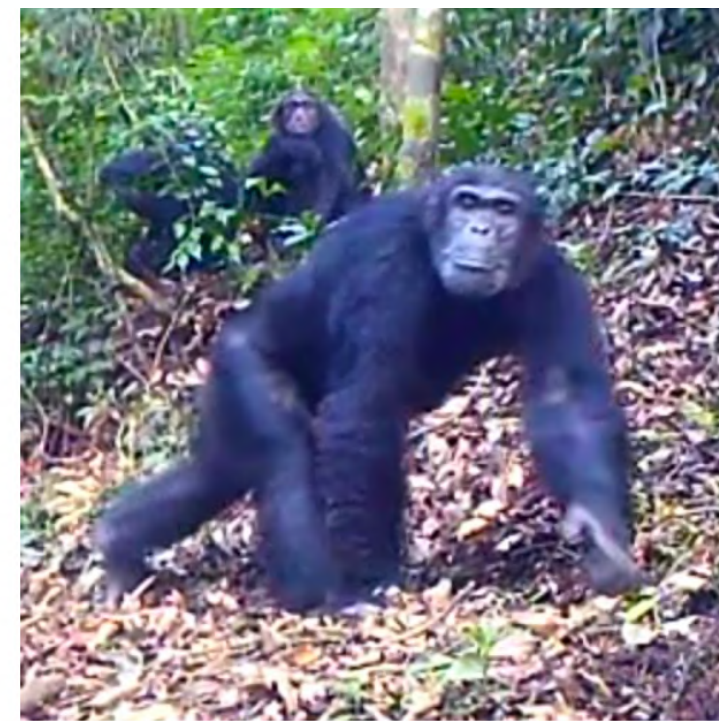
sitting



sitting\_on\_back



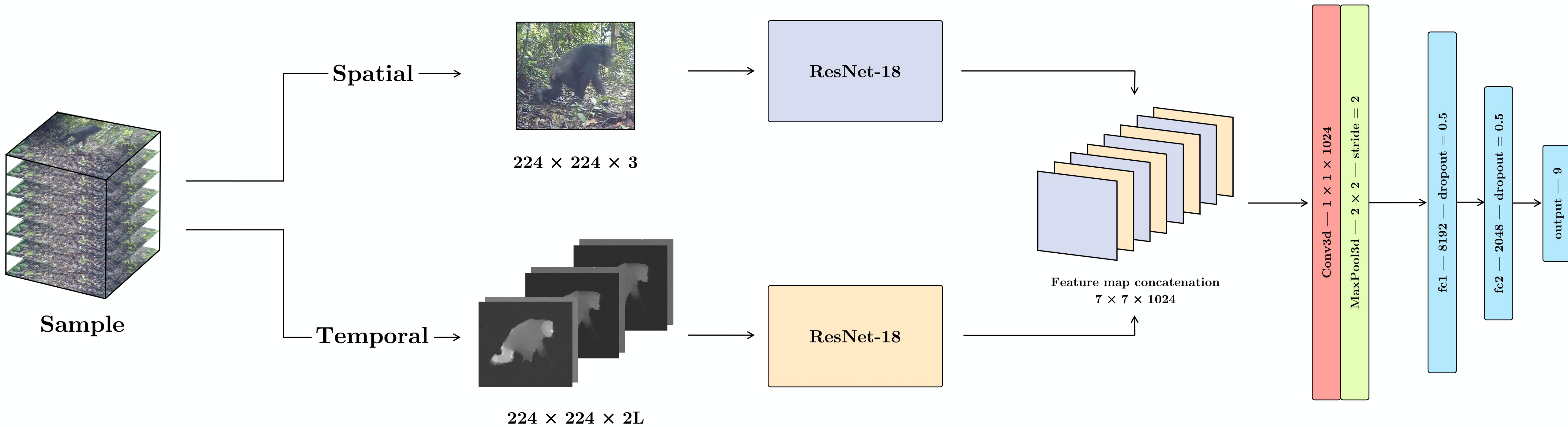
standing



walking

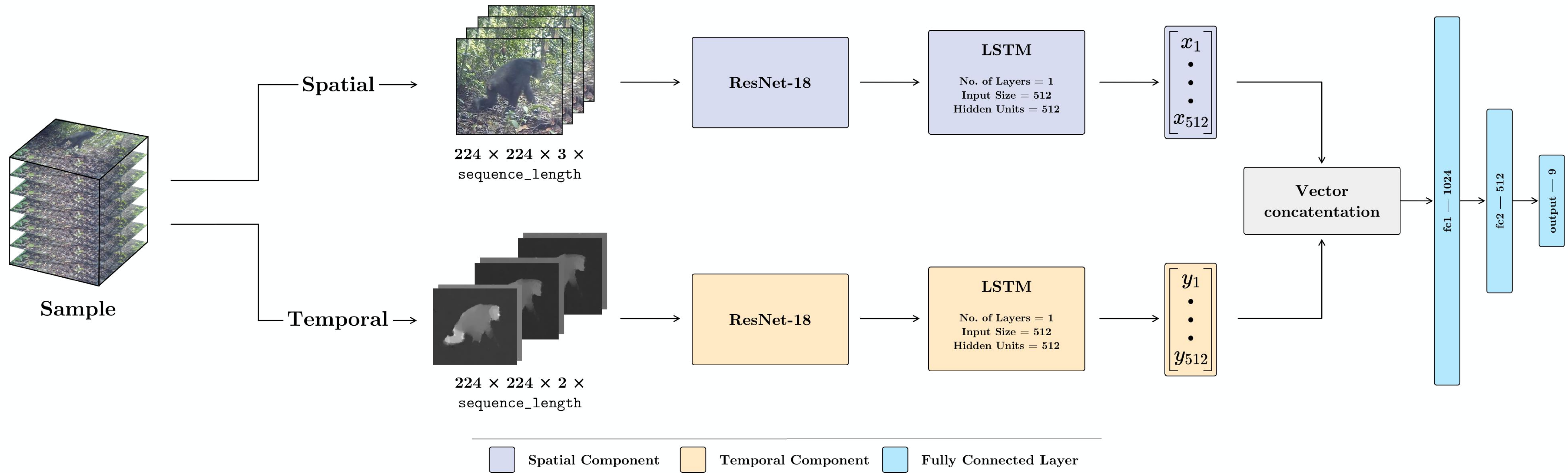
# The Model

# Baseline Architecture

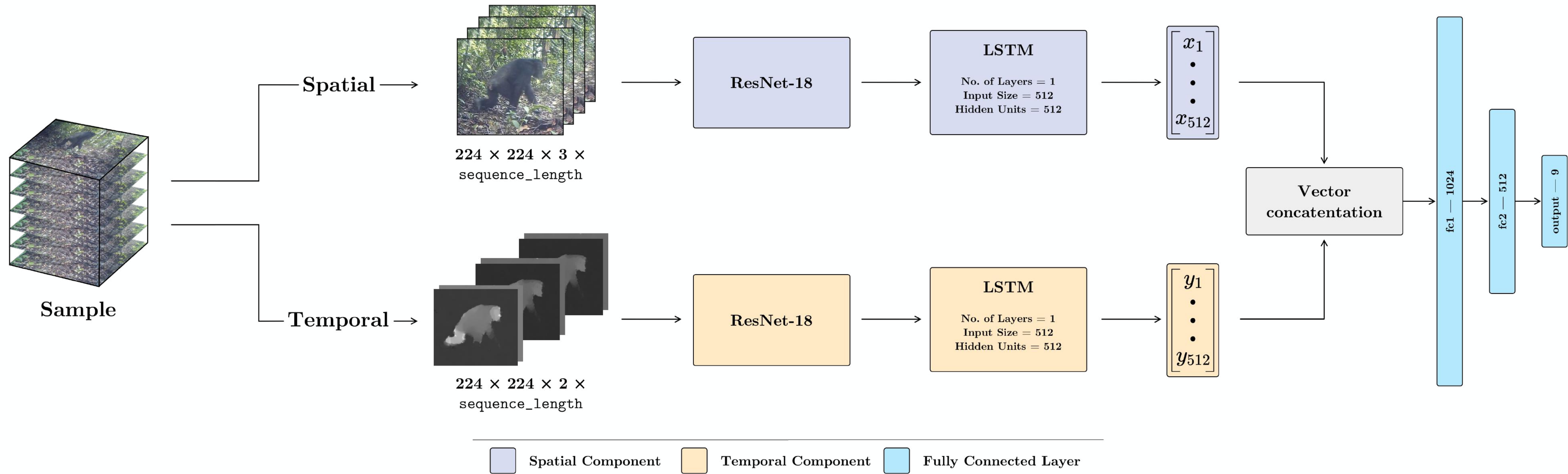


 Spatial Component    Temporal Component    Fully Connected Layer    Max-Pool Layer    3D Convolution Layer

# Final Architecture

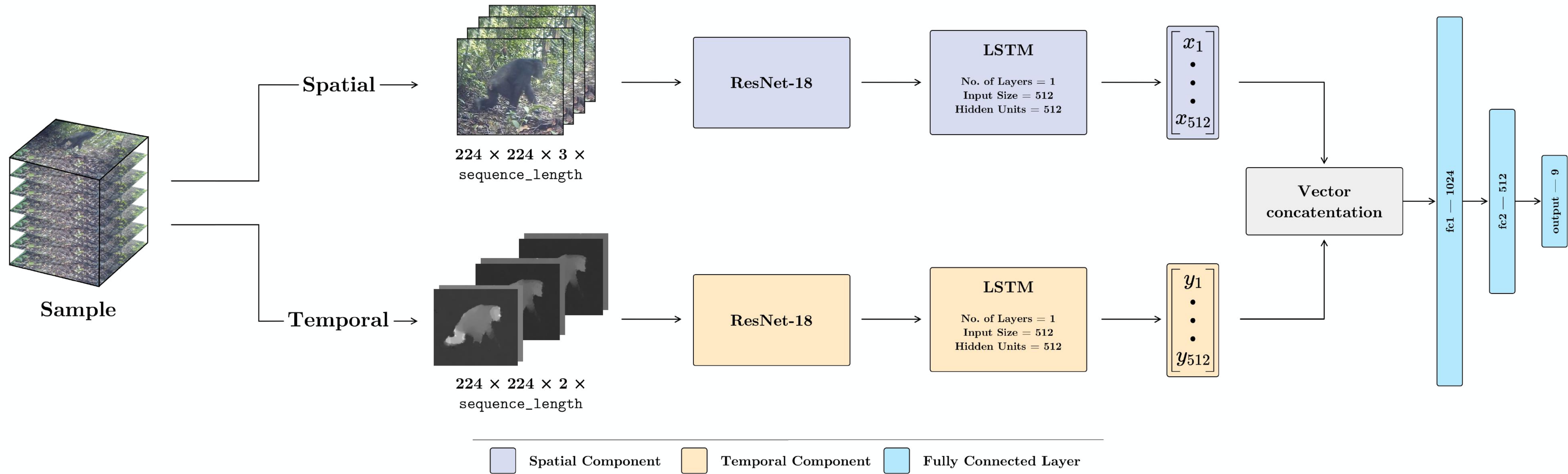


# Final Architecture



+ **Balanced Sampling**

# Final Architecture



+ Balanced Sampling

+ Focal Loss

# Sampling the Apes

To fulfil the multi-subject, multi-frame requirements of the model, every ape is considered for sampling in any given video.





# Sampling the Apes

Three parameters are used to determine the method of sampling:

1. Sequence Length
2. Sampling Rate
3. Behaviour Duration Threshold



**Results**



# Test Set Results

Model	Top1 Accuracy	Top3 Accuracy
Baseline	61.39%	86.34%
Final	<b>74.82%</b>	<b>95.49%</b>

# Cross Validated Results

- 4-fold Cross validation
- 375 training videos to 125 test videos per fold

Top1 Accuracy	Top3 Accuracy
<b>73.52%</b>	<b>94.07%</b>

# Qualitative Analysis

# Transitional Behaviours



# Transitional Behaviours (gone wrong)



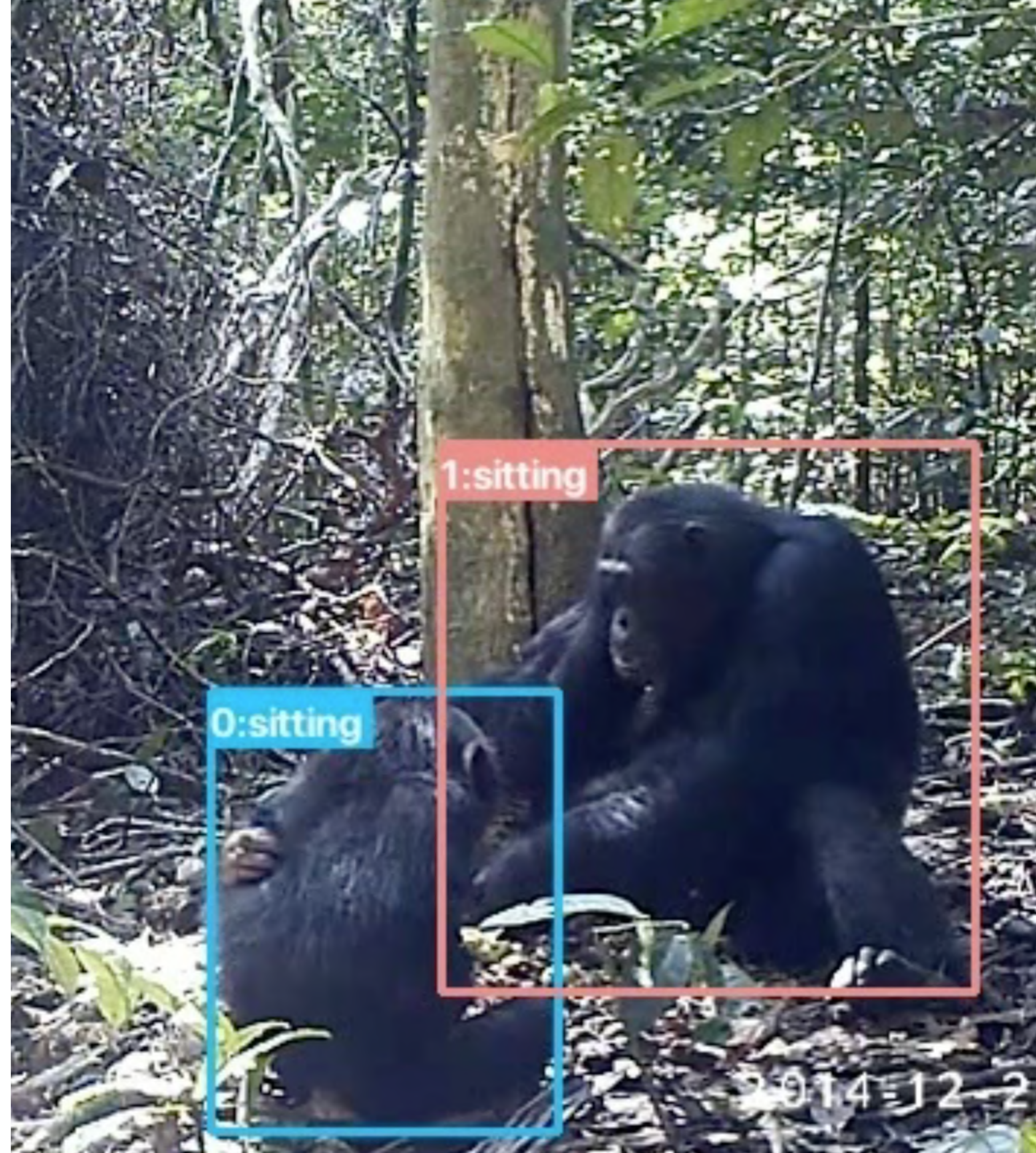


# Other Challenges



**Conclusion**

- We have introduced a great ape dataset annotated with bounding boxes and behaviours of 500 videos
- We evaluated a basic two-stream deep learning model using the challenging footage in the dataset, performing at a Top1 accuracy of **73.52%**



**Links to  
Contributions**

# Dataset Annotations

<http://bit.ly/PanAfrican2020>

# Dataset Videos

Please contact the Max Planck Institute for Evolutionary Anthropology.

# Code

<http://bit.ly/GreatApeBehaviourDetectorCode>

# GitHub

<https://github.com/fznsakib/great-ape-behaviour-detector>

# References

- [1] International Union for Conservation of Nature (IUCN). Red List of Threatened Species, 2020. <https://www.iucnredlist.org>.
- [2] X. Yang, M. Mirmehdi, and T. Burghardt, "Great ape detection in challenging jungle camera trap footage via attention-based spatial and temporal feature blending," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2019, pp. 255–262.
- [3] D. Schofield, A. Nagrani, A. Zisserman, M. Hayashi, T. Matsuzawa, D. Biro, and S. Carvalho, "Chimpanzee face recognition from videos in the wild using deep learning," *Science advances*, vol. 5, no. 9, 2019.
- [4] Max Planck Institute for Evolutionary Anthropology, "Pan African Programme: The Cultured Chimpanzee," <http://panafrican.eva.mpg.de/index.php>.

# Acknowledgements



We would like to thank the entire team of the Pan African Programme: 'The Cultured Chimpanzee' and its collaborators for allowing the use of their data for this paper. Please contact the copyright holder Pan African Programme at <http://panafrican.eva.mpg.de> to obtain the videos used from the dataset. Particularly, we thank: H Kuehl, C Boesch, M Arandjelovic, and P Dieguez. We would also like to thank: K Zuberbuehler, K Corogenes, E Normand, V Vergnes, A Meier, J Lapuente, D Dowd, S Jones, V Leinert, E Wessling, H Eshuis, K Langergraber, S Angedakin, S Marrocoli, K Dierks, T C Hicks, J Hart, K Lee, and M Murai.

Thanks also to the team at <https://www.chimpandsee.org>. The work that allowed for the collection of the dataset was funded by the Max Planck Society, Max Planck Society Innovation Fund, and Heinz L. Krekeler.

In this respect we would also like to thank: Foundation Ministre de la Recherche Scientifique, and Ministre des Eaux et Forêts in Cote d'Ivoire; Institut Congolais pour la Conservation de la Nature and Ministre de la Recherche Scientifique in DR Congo; Forestry Development Authority in Liberia; Direction des Eaux, Forêts Chasses et de la Conservation des Sols, Senegal; and Uganda National Council for Science and Technology, Uganda Wildlife Authority, National Forestry Authority in Uganda.

**Thank you!**

