Introduction

- Advanced technologies provide huge amount of video data
- Rich and precise information in videos
- Unlikely to process it manually due to limited resources
- Applied AI/ML tools to analyze videos
- Understand sea turtle’s behavior

Objectives

- Classify the behavior of sea turtle in the turtle footage
- Obtain frame-wise action labels for videos
- Behaviors include swimming, breathing, surface resting, and more

Method

Temporal Convolutional Network (TCN)

- Capture the relationship and dependencies between sequences of inputs
- Efficient and compact - less memory because of dilated convolutions over sequences of inputs

Results & Analysis

- Unable to distinguish between swimming and seafloor resting
- We conjecture that the model relies on the background information of frames for prediction instead of capturing motions across frames
- Majority of feeding behavior being predicted as either seafloor resting or swimming suggests their similarity
- Unbalanced dataset is one of the main reasons of non-optimum performance

Conclusion

The model has trouble distinguishing activities with similar backgrounds; swimming and resting both occur underwater. Our results suggest current SOTA models rely heavily on the spatial information and characteristics of frames to predict labels, instead of capturing the motions and correspondences across frames. Moreover, as our dataset is extremely unbalanced, the model tends to predict everything as the majority categories. Additionally, we show that using annotated dataset with higher temporal frequency does not improve the performance; one frame per second is sufficient for the purpose of video classification.

Future Work

- Better video classification frameworks such as Non-local network
- Synthetic dataset for fixing the unbalanced dataset
- Triplet loss for distinguishing similar activities

Reference


Acknowledgements

<table>
<thead>
<tr>
<th>ILLINOIS</th>
<th>NCSA</th>
<th>National Center for Supercomputing Applications</th>
</tr>
</thead>
</table>

AI System for Analyzing Footage from Animal-borne Cameras: a Sea Turtle Case Study
Brian Chen, Brian Allan, Nathan Robinson, Aiman S Soliman
Department of Computer Science, College of Engineering, University of Illinois at Urbana-Champaign