Week 12 Document Submission

Jacob Dallaire

November 8, 2024

1. Paper

H. Chen *et al*., "Assessing impacts of data volume and data set balance in using deep learning approach to human activity recognition," *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Kansas City, MO, USA, 2017, pp. 1160-1165, doi: 10.1109/BIBM.2017.8217821.

SUMMARY

The introduction outlines the integration of various sensors in modern smartphones, which facilitates human activity recognition (HAR) by capturing diverse data types relevant to users' health and activity needs. The paper emphasizes the significance of training data volume and category balance on the accuracy of deep learning models, specifically Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), in recognizing human activities. It highlights the development of a modified SMOTE algorithm to address class imbalance in HAR datasets, demonstrating that balanced datasets can enhance recognition accuracy, particularly for minority activity categories, thereby improving overall model performance.

Not particularly relevant to vision classification but I am interested in genetics.

2. Scripts

Modification of the load_dataset_with_labels function with max_samples_per_class as a new parameter to correct for imbalance in the data set.

```
for filename in os.listdir(image_dir):
        if filename.endswith(('.jpg', '.jpeg', '.png')):
            image_path = os.path.join(image_dir, filename)
            label = extract_label_from_filename(filename)

        if class_counts[label] < max_samples_per_class or max_samples_per_class == 0:
                image_paths.append(image_path)
                labels.append(label)
                class_counts[label] += 1
```

Modification of the object detector model the include affine transformations of the dataset.

```
def rotate_bboxes(bboxes, angle):
    """Adjust bounding box coordinates based on the rotation angle."""
    rotated_bboxes = []
    for bbox in bboxes:
        xmin, ymin, xmax, ymax = bbox
        if angle == 90:
```

```
            # Rotate 90 degrees clockwise
            rotated_bbox = [ymin, 1 - xmax, ymax, 1 - xmin]
        elif angle == 180:
            # Rotate 180 degrees
            rotated_bbox = [1 - xmax, 1 - ymax, 1 - xmin, 1 - ymin]
        elif angle == 270:
            # Rotate 270 degrees clockwise
            rotated_bbox = [1 - ymax, xmin, 1 - ymin, xmax]
        else:
            rotated_bbox = bbox
        rotated_bboxes.append(rotated_bbox)
    return np.array(rotated_bboxes)

def rotate_image(image, angle):
    if angle == 90:
        return tf.image.rot90(image, k=1)
    elif angle == 180:
        return tf.image.rot90(image, k=2)
    elif angle == 270:
        return tf.image.rot90(image, k=3)
    else:
        return image
```

3. Documentation

I spent the week fairly ill and did not work on the project as much as intended.

This week was spent experimenting with different sizes of the training data set to see how reduced class imbalance impacts the training results. With a cap of 1000 samples the accuracy was 25%. Capped at 3000 samples the accuracy was 29.99% with top two accuracy of 53.25%. This is better than random prediction of 20% but leaves much room for improvement. At the end of my last epoch the model was still improving so more epochs and a larger dataset should still provide significant improvement.

The object detector model is outputting what is likely the mean bounding box of the entire image set. I did not work on this further this week.

4. Next Weeks Proposal

I will run experiments with larger datasets and a larger max epoch size. Implementation of an exit condition to epochs is necessary as well. To stop training when no further improvement is being made.

Weekly Report

Philip Woolley

2024-11-08

Time Log Reponse:

- Developed method for stitching adjacent frames together into one 3d volume
- What are you planning on working on next? - Continue segmenting training data. Retrain model with additional data and new metric. Change website structure to match request from Bree
- Is there anything blocking you? - Access to "Appearance" tab on wordpress needed to rearrange navigation menu

1 Abstract

Abstract

Automated identification and quantification of algae in microscopic images is a tool that allows high taxonomic resolution with reasonable technical efforts. However, in samples containing various non-algal objects, this is still not a satisfactorily solved problem. We show that autofluorescence information improves discrimination of algae from non-algal objects as well as phycoerythrin (PE) containing algae from others. We analyse the stability of the autofluorescence to estimate its constraints. Cold and dark storage of glutaraldehyde fixed samples maintains autofluorescence sufficiently for 3 weeks. Under repeated excitations, chlorophyll a (Chl a) or PE autofluorescence show an exponential decrease followed by an intermediate maximum. A peak also occurs in emission wavelength ranges without chlorophyll and PE fluorescence. The unspecific autofluorescence causing the peaks is at least partly identical with the blue–green fluorescence (BGF) in plant cells. BGF interferes with identification of algae, thus correction of pigment autofluorescence with such unspecific fluorescence allows a more reliable algal discrimination procedure. A classification scheme for discrimination of Chl a and PE-containing algae shows a high performance in a test with natural samples. Integration of fluorescence and bright-field image information provides a powerful tool for phytoplankton analysis in complex samples.

Summary This paper from the Journal of Plankton Research is an example of an ecology-focused paper detailing a machine learning based method. Similar to my recent conversation with Bree, I believe that this is the type of paper that my project is most suited to. I will be aiming to draft my manuscript similar to this paper, although I do not believe that the method itself is sufficiently novel to be published. This paper does a good job of describing the methods used, but the methods section takes up a comparatively smaller segment of the paper than I am seeking to write, because the methods are more immediately important than the results in my writing. The results of my work will be useful for future research in the lab, but will require expert analysis that I can not perform. Another thing this paper does well is compare their method to current state of the art in terms of accuracy and time needed to collect results.

Citation

Hense, Burkhard A., et al. "Use of fluorescence information for automated phytoplankton investigation by image analysis." Journal of Plankton Research 30.5 (2008): 587-606.

2 Scripts and Code Blocks

```python
def calc_overlap(set1, set2_intersect, set2_max):
    union = np.count_nonzero(set1) + np.count_nonzero(set2_max) - np.count_nonzero(set2_intersect)
    return np.count_nonzero(set2_intersect) / union
```

0.0s Python

```python
#choose which value on second frame corresponds best to area on first frame
def find_corresponding_tooth(source_pixels, target_frame, target_pixels):
    best_fit, max_overlap = 0, 0
    for val in np.unique(target_pixels):
        target_pixels = target_frame[target_frame == val]
        if((val !=0 and val !=1)):
            t = calc_overlap(source_pixels, target_pixels[target_pixels == val], target_frame[target_frame == val])
            if t > max_overlap:
                best_fit, max_overlap = val, t
    if max_overlap < 0.1:
        best_fit = 0
    return best_fit
```

0.0s Python

```python
tooth_label_val = 2


for idx, out in enumerate(outlist):
    frame = out['segmentation']
    f2 = copy.deepcopy(frame)
    segments_info = out['segments_info']
    for segment in segments_info:
        #modify tooth values to prepare for combining
        if segment['label_id'] == 2:
            frame[f2 == segment['id']] += 1000
        #change background and upper jaw values to 0
        elif segment['label_id'] == 0 or segment['label_id'] == 3:
            frame[f2 == segment['id']] = 0
        #change lower jaw values to 1
        else:
            frame[f2 == segment['id']] = 1
    #print(np.unique(frame))

#apply find_corresponding_tooth to each frame
```

```python
#apply find_corresponding_tooth to each frame
for idx, out in enumerate(outlist):
    frame = out['segmentation']
    if idx == 0:
        x = np.unique(frame)
        for value in x:
            if(value !=0 and value !=1):
                print(value)
                tval = value
                if value >= 1000:
                    tval = tooth_label_val
                frame[frame == value] = tval
                if value >= 1000:
                    tooth_label_val = tooth_label_val + 1
        print(np.unique(frame))
    if idx <= (len(outlist)-2):
        x = np.unique(frame)
        print(x)
        for value in x:
            if(value !=0 and value !=1):
                print(value)
                next_frame = outlist[idx + 1]['segmentation']
                corresponding_tooth_val = find_corresponding_tooth(frame[frame == value], next_frame, next_frame[frame ==
                tval = value
                if value >= 1000:
                    tval = tooth_label_val
                frame[frame == value] = tval
                if corresponding_tooth_val != 0 and corresponding_tooth_val != 1:
                    next_frame[next_frame == corresponding_tooth_val] = tval
                if value >= 1000:
                    tooth_label_val = tooth_label_val + 1
                outlist[idx + 1]['segmentation'] = next_frame
```

3 Documentation

The VisualizeModelResults.ipynb notebook is used for creating and viewing images of model output on validation data. Users provide a pretrained model and validation dataset, and this notebook inferences all of the images in the dataset and allows the user to review the output segmentations against the ground truth manual segmentations.

The DataProcess.ipynb notebook is used for converting slicer volume files (.nrrd and .seg.nrrd) into a HuggingFace dataset for use with the pretrained Mask2Former model. Volumes should be added to the "vols" folder, and segmentation volumes should be added to the "masks" folder.

https://www.morphosource.org/projects/0000C1059?locale=enpage=11sort=publication_status_s List of available MicroCT Datasets of anolis lizards that will be used for this project. When infrastructure for data storage is ready I will prepare documentation detailing the downloading and storage process.

https://slicermorph.github.io/ Documentation for SlicerMorph, an extension of the 3D slicer tool commonly used by Biologists. This is used for loading stacks of .tiff images as a volume in 3d slicer.

https://github.com/jmhuie/SlicerBiomech Documentation for the Dental Dynamics module, which is a 3D slicer extension for calculating tooth stress from jaw segmentations. the outputs from my segmentation pipeline will need to be compatible with this module for analysis.

4 Script Validation (Optional)

5 Results Visualization

6 Proof of Work

Please see code blocks section

7 Next Week's Proposal

- Continue segmenting training data for ML panoptic segmentation model
- Develop testing script for 3D image registration for converting coordinate systems
- Reformat blog page as requested by Bree

Week12 report

Ruiqing Wang | CiChild CV team

- What progress did you make in the last week?
  1. Creating data annotations on SLEAP
  2. Transferred all necessary videos, csv file and visualization from ICE to dropbox
  3. Attend Cichild group meeting and discussed with Bree with current work goal
  4. Review papers on SLEAP and pose estimation
  5. Help assembling paper report submissions and address submission situation.
- What are you planning on working on next?
  1. Train network on SLEAP
  2. Run DEMO video export and validate results
  3. Meet with Cichild CV team to discuss current progress
- Is anything blocking you from getting work done?
- My allocation run out of space, and I am trying to get extra spaces

N/A

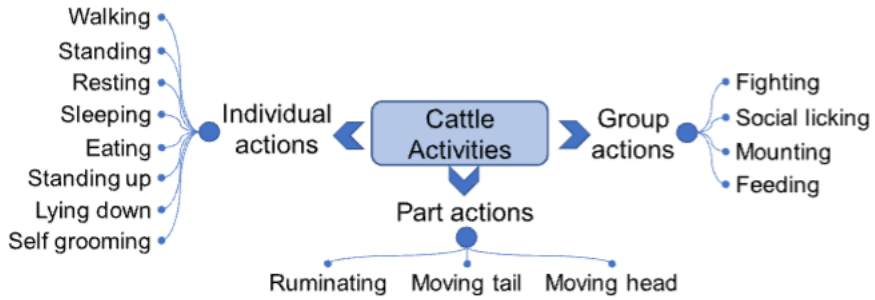
Paper abstract

Paper: Deep learning-based hierarchical cattle behavior recognition with spatiotemporal information
https://www.sciencedirect.com/science/article/pii/S0168169920307110

Abstract: Behavior serves as a critical metric for assessing animal welfare, traditionally monitored through video analysis or portable devices that track movements. However, these methods can inadvertently alter animal behavior and have practical limitations. This paper presents a deep learning-based framework for hierarchical cattle behavior recognition that leverages spatio-temporal information, allowing for the detection and localization of various cattle behaviors within video frames, supported by a comprehensive dataset collected under diverse conditions, demonstrating the system's efficacy in recognizing 15 distinct hierarchical activities.

Methodology:

The framework presented for automatic recognition of cattle behaviors integrates deep learning techniques to analyze video data, enabling real-time detection and localization of 15 hierarchical behaviors across individual, group, and part actions. By leveraging both spatial and temporal information, the system is designed to operate effectively under varying farm conditions, demonstrating its potential to enhance livestock management through automated monitoring.



The algorithm requires real-time performance, necessitating the use of deep learning frameworks that maintain high accuracy while processing data rapidly. YOLOv3, a state-of-the-art object detection model, excels in real-time scenarios and effectively identifies objects across various scales due to its pyramidal architecture and 53-layer network, which incorporates residual skip connections and up-sampling techniques. By initializing the model with pre-trained weights from the MSCOCO dataset and fine-tuning it on a specific dataset, the authors leverage YOLOv3 to detect regions of interest (ROIs) that correspond to specific cattle behaviors, utilizing keyframes extracted from video data for analysis.

Scripts and Code Blocks

This week I mainly read papers and user guide, install softwares and did data annotation. There do exist the issue related to computing at scratch folder:

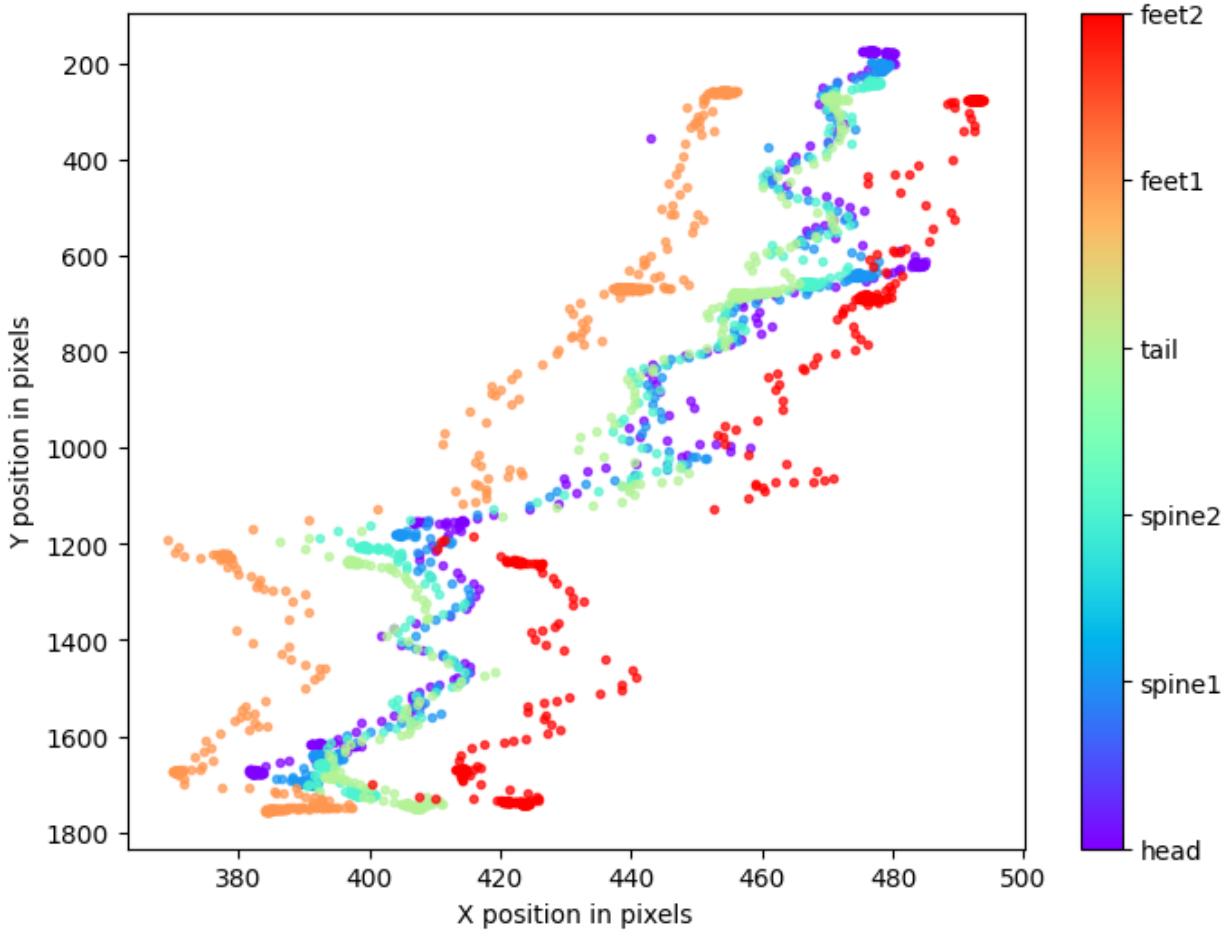
```
#!/bin/bash
#SBATCH -J DLC_analys_5_26                  # Job name
#SBATCH -N1 --ntasks-per-node=1                # Number tasks per node
#SBATCH --gres=gpu:H100:1      # reuqesting 1 GPU H100
#SBATCH --cpus-per-task=4                # request 4 cores
#SBATCH --mem=64G                 # Memory per gpu
#SBATCH -t 0-6:00:00                           # Duration of the job (Ex: 15 mins)
#SBATCH -o video_5_26_Report-%j.out                  # Combined output and error messages file
#SBATCH --mail-type=BEGIN,END,FAIL       # Mail preferences
#SBATCH --mail-user=rwang753@gatech.edu  # e-mail address for notifications
cd /home/hice1/rwang753/scratch/week9/trained_20videos_0603_RW_2024-10-10/20_videos_0603-RW-2024-10-10/videos/5_26_2024
nge to working directory created in

source /home/hice1/rwang753/anaconda3/bin/activate DEEPLABCUT
srun python analys_video.py
~
```

The key thing is to specify folder access before source the environment

Documentation

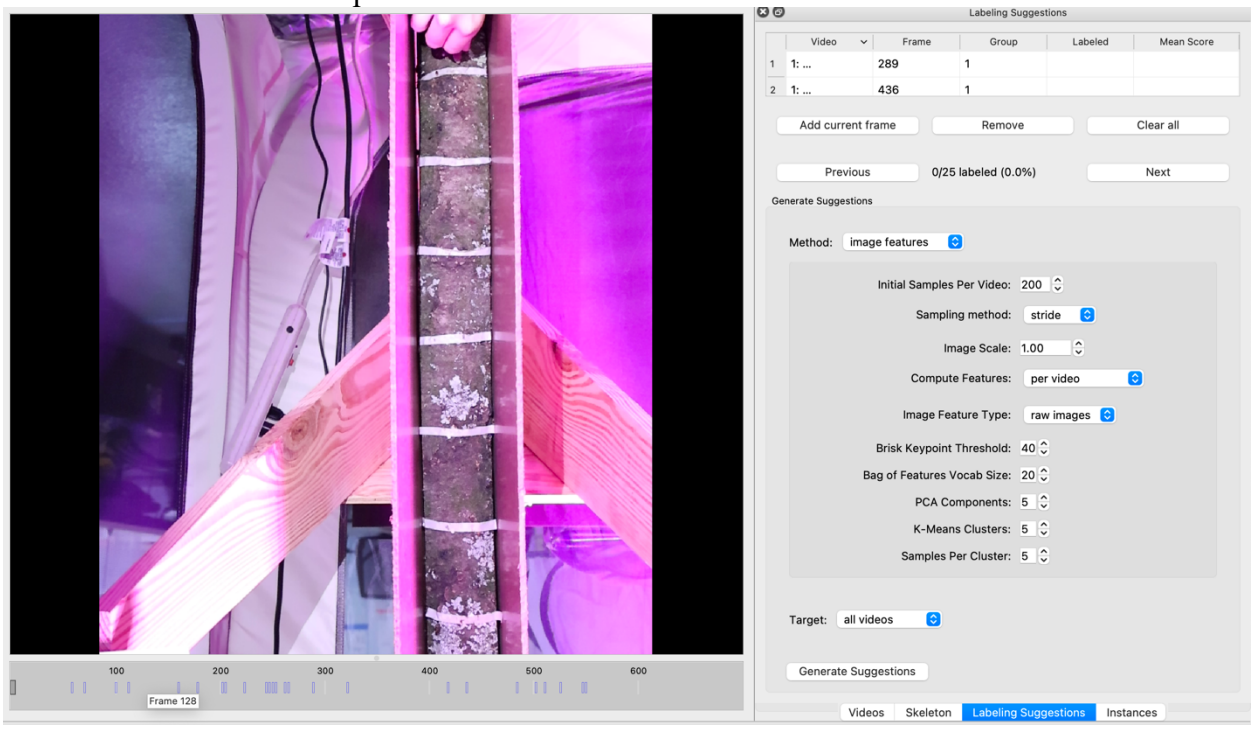
For dropbox, I have put my csv file, videos and visualizations for lizard datasets.



Here is the corresponding csv file:

scorer	DLC_Resnet50_20_videos_0603Oct10shuffle1_snapshot_200	DLC_Resnet50_20_videos_0603Oct10shuffle1_snapshot_200	DLC_Resnet50_20_videos_0603Oct10shuffle1_snapshot_200	DLC_Resnet50_20_videos_0603Oct10shuffle1_snapshot_200	DL
bodyparts	head	head	head	spine1	spir
coords	x	y	likelihood	x	y
0	450.8924	1523.7936	0.023387719	755.45294	
1	451.04587	1523.6476	0.049566478	755.44006	
2	450.9551	1523.7006	0.031934656	755.5543	
3	450.89334	1523.485	0.039829127	755.3232	
4	451.02097	1523.7048	0.03504911	755.3487	
5	451.03915	1523.528	0.04405748	755.3606	
6	451.0941	1523.6843	0.033583727	755.3973	
7	451.0625	1523.7017	0.03461857	755.5614	
8	450.8835	1523.5479	0.040243886	755.5875	
9	450.92932	1523.7587	0.029895231	755.5934	
10	411.7193	76.19147	0.018949805	755.5847	
11	450.56985	1523.7036	0.023614097	755.6527	
12	450.6947	1523.6785	0.026738338	755.63525	
13	450.95148	1523.6534	0.03680185	755.61646	
14	451.12033	1523.6615	0.047070604	755.5178	
15	451.08853	1523.786	0.030402774	755.3608	
16	451.15686	1523.7517	0.040336113	755.3448	
17	450.9848	1523.7882	0.029504487	755.2985	
18	451.04843	1523.5391	0.044539712	755.3143	
19	451.0614	1523.9434	0.025337484	755.5433	
20	659.7481	603.94025	0.022962674	755.6849	
21	451.15994	1523.892	0.030679714	755.36646	
22	411.6318	76.254326	0.021438936	755.67676	
23	451.20035	1523.9232	0.033234082	755.6303	
24	451.18857	1524.0265	0.02715397	755.48706	
25	659.78986	603.8074	0.021172663	755.51575	
26	451.17767	1523.8193	0.030324947	755.633	
27	451.17398	1524.0804	0.020310655	755.5873	
28	451.20468	1523.7506	0.038112305	755.39355	
29	451.14142	1524.1501	0.021569662	755.3984	
30	659.9974	603.8844	0.020293036	755.5621	
31	450.9933	1523.7104	0.033751406	755.56726	
32	660.0381	603.9226	0.025504304	755.47034	
33	450.92337	1523.6553	0.035111062	755.5906	
34	451.1423	1523.5846	0.040737808	755.4149	
35	659.98254	603.8879	0.023827244	755.3864	
36	660.08936	603.9308	0.026240086	755.45087	
37	451.24808	1523.7559	0.03944226	755.5792	
38	451.191	1524.0428	0.02618706	755.5659	
39	451.16562	1524.0314	0.024080213	755.48883	
40	411.63882	76.22808	0.018324904	755.4356	
41	451.1121	1523.9781	0.022682846	755.5973	

For data annotations at sleap: here is the details:

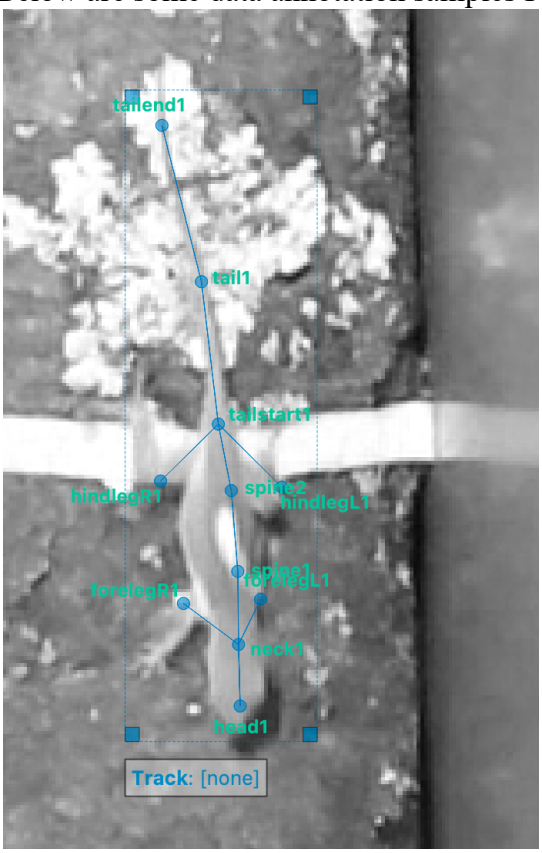


Results Visualization and Code Validation

The data visualization is very similar to data annotations last week. For skeleton design, Under “Skeleton" interface, we have lists nodes such as "head1," "neck1," "forelegL1," and others, indicating their labeled positions and symmetry attributes, which are key in defining the structure for motion analysis or behavioral studies. The annotations shows the structure of the labeled details.

Proof of Work and code validation

Below are some data annotation samples I did at the GUI:



Next Week’s Proposal

1. Train network on SLEAP
2. Run DEMO video export and validate results
3. Meet with Cichild CV team to discuss current progress

Week 12 Document Submission

Lizard X-RAY Landmark Group

Mercedes Quintana

What progress did you make in the last week?

- Continued to work on website
- Analyzed data from both trained models
- Commented code to make more readable

What are you planning on working on next?

- Add more images to auto processed images
- Create frontend for final product
- Continue to update the website

Is anything blocking you from getting work done?

- Nope

Abstracts:

URL: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13373

Camouflage detection: Optimization-based computer vision for Alligator sinensis with low detectability in complex wild environments

Alligator sinensis is an extremely rare species that possesses excellent camouflage, allowing it to fit perfectly into its natural environment. The use of camouflage makes detection difficult for both humans and automated systems, highlighting the importance of modern technologies for animal monitoring. To address this issue, we present YOLO v8-SIM, an innovative detection technique specifically developed to significantly enhance the identification precision. YOLO v8-SIM utilizes a sophisticated dual-layer attention mechanism, an optimized loss function called inner intersection-over-union (IoU), and a technique called slim-neck cross-layer hopping. The results of our study demonstrate that the model achieves an accuracy rate of 91 %, a recall rate of 89.9 %, and a mean average precision (mAP) of 92.3 % and an IoU threshold of 0.5. In addition, the model operates at a frame rate of 72.21 frames per second (FPS) and excels at accurately recognizing objects that are partially visible or smaller in size. To further improve our initiatives, we suggest creating an open-source collection of data that showcases A. sinensis in its native environment while using camouflage techniques. These developments collectively enhance the ability to detect disguised animals, thereby promoting the monitoring and protection of biodiversity, and supporting ecosystem sustainability.

Summary: This paper demonstrates a method for tracking and monitoring rare, endangered crocodiles using an optimized version of YOLO v8 which is designed to separate background and the object of interest when they are similar colors.

Scripts and Code Blocks:

Last week I analyzed the data from the two machine learning models and prepared the visualization code to be used by others by adding comments and function descriptions. I found that the automatically processed images may be suitable for the final pipeline. I also prepared for the bi-weekly group meeting.

Documentation:

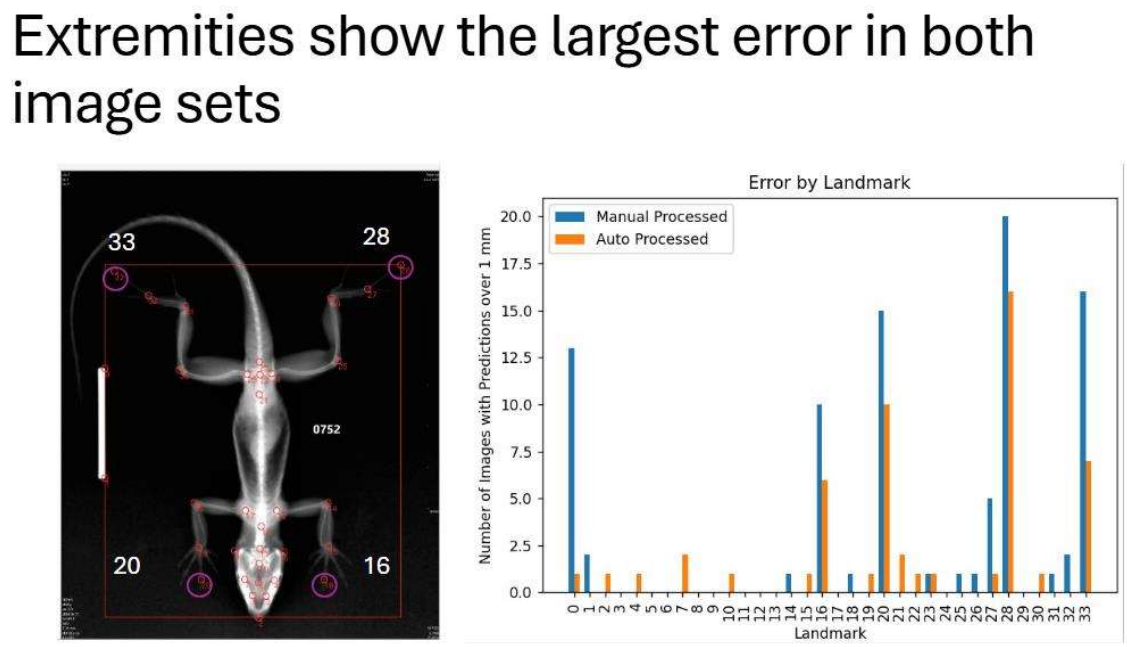
I have no new code to document as of now.

Script Validation:

I have no validation steps now.

Results Visualization / Proof of Work:

Here is a portion of the powerpoint shared at the meeting comparing the automatic and manually processed images. The models are both performing well but have issues with the toes.



Next Week Proposal:

Next week I am going to finish reading through the papers and journal suggestion Bree had set out for us in the meeting on Wednesday. I am going to continue to make sure the website is updated. I am going to add more images to the automatically processed model and see if that changes the accuracy for the better. I am also going to take a closer look at the current outliers in the model and see if they come from the same images.