Week 13 Document Submission

Jacob Dallaire

November 15, 2024

1. Paper

Arnekvist, Isac, J. Frederico Carvalho, Danica Kragic, and Johannes A. Stork. "The effect of target normalization and momentum on dying relu." *arXiv preprint arXiv:2005.06195* (2020).

SUMMARY

The introduction discusses the challenges of training deep neural networks (NNs) using gradient-based optimization, particularly focusing on the phenomenon known as "dying ReLU," where the rectified linear unit (ReLU) activation function can become inactive and output constant zero values during training. The authors highlight the importance of target normalization and momentum in mitigating this issue, noting that improper scaling of target values can exacerbate the problem, especially in deeper architectures. Through theoretical analysis and empirical evidence, they aim to elucidate the underlying mechanisms that lead to dying ReLUs and propose a framework for understanding their behavior in relation to optimization techniques.

2. Scripts

No new scripts just experiment with different epochs and learning rate.

3. Documentation

I had my computer running a variety of different settings for epoch exit condition, training data under sampling rate, and learning rate. I believe what I was previously encountering was Relu death that caused my model to collapse to a single class prediction. I have had some success in getting the model to predict more variety in classes, but accuracy is around 35% but the error is spread out across classes. Currently a model is being trained with a smaller learning rate that hopefully see some improvement.

4. Next Weeks Proposal

More training to try and improve accuracy is necessary. Figuring out what I believe is a result of Relu death has given me a path forward to improve the model.

Weekly Report

Philip Woolley

2024-11-15

Time Log Reponse:

- Finalized registration method for aligning lizard heads to coordinate plane
- 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

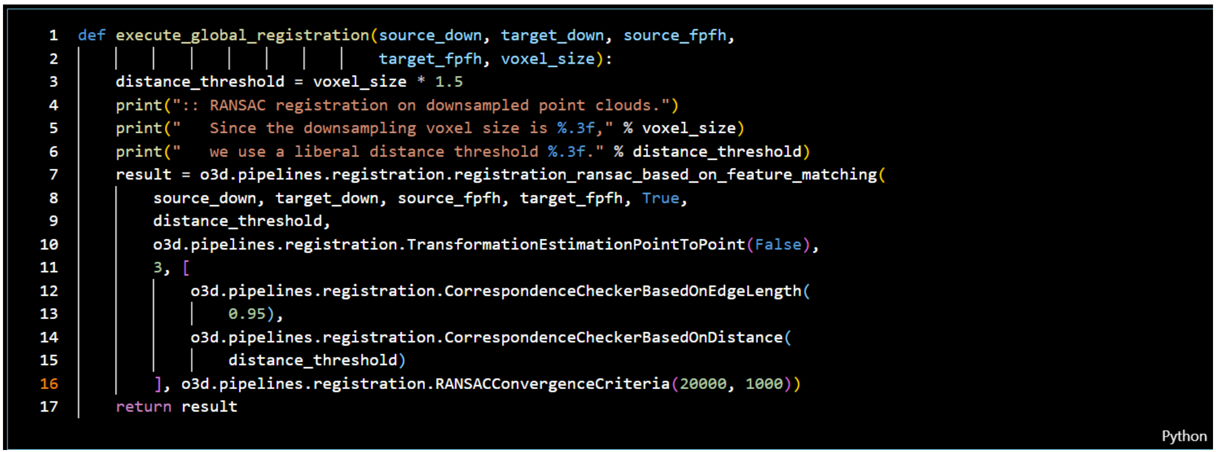
1 Abstract

Abstract Segmentation is crucial in medical imaging analysis to help extract regions of interest (ROI) from different imaging modalities. The aim of this study is to develop and train a 3D convolutional neural network (CNN) for skull segmentation in magnetic resonance imaging (MRI). 58 gold standard volumetric labels were created from computed tomography (CT) scans in standard tessellation language (STL) models. These STL models were converted into matrices and overlapped on the 58 corresponding MR images to create the MRI gold standards labels. The CNN was trained with these 58 MR images and a mean ± standard deviation (SD) Dice similarity coefficient (DSC) of 0.7300 ± 0.04 was achieved. A further investigation was carried out where the brain region was removed from the image with the help of a 3D CNN and manual corrections by using only MR images. This new dataset, without the brain, was presented to the previous CNN which reached a new mean ± SD DSC of 0.7826 ± 0.03. This paper aims to provide a framework for segmenting the skull using CNN and STL models, as the 3D CNN was able to segment the skull with a certain precision.

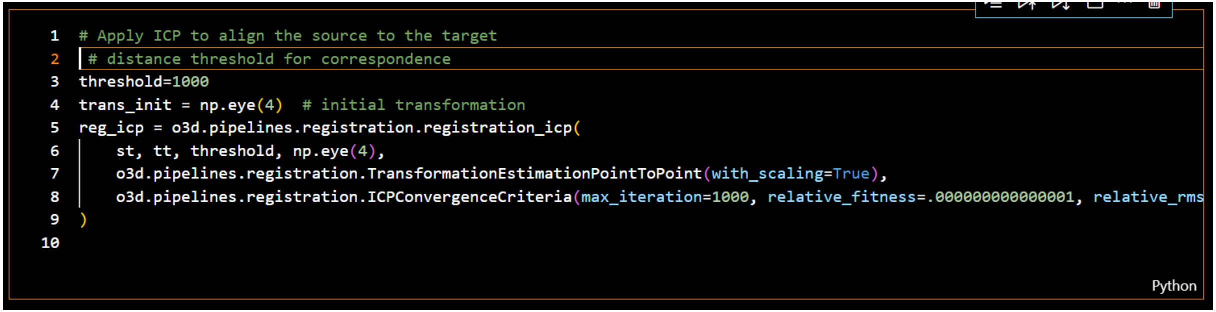
Summary This paper describes a deep learning approach for segmenting the brain from whole-head MRI images using labels generated from brain CT scans.

Citation Dalvit Carvalho da Silva R, Jenkyn TR, Carranza VA. Development of a Convolutional Neural Network Based Skull Segmentation in MRI Using Standard Tesselation Language Models. J Pers Med. 2021 Apr 16;11(4):310. doi: 10.3390/jpm11040310. PMID: 33923480; PMCID: PMC8074044.

2 Scripts and Code Blocks



```python
def execute_global_registration(source_down, target_down, source_fpfh,
                                target_fpfh, voxel_size):
    distance_threshold = voxel_size * 1.5
    print(":: RANSAC registration on downsampled point clouds.")
    print("   Since the downsampling voxel size is %.3f," % voxel_size)
    print("   we use a liberal distance threshold %.3f." % distance_threshold)
    result = o3d.pipelines.registration.registration_ransac_based_on_feature_matching(
        source_down, target_down, source_fpfh, target_fpfh, True,
        distance_threshold,
        o3d.pipelines.registration.TransformationEstimationPointToPoint(False),
        3, [
            o3d.pipelines.registration.CorrespondenceCheckerBasedOnEdgeLength(
                0.95),
            o3d.pipelines.registration.CorrespondenceCheckerBasedOnDistance(
                distance_threshold)
        ], o3d.pipelines.registration.RANSACConvergenceCriteria(20000, 1000))
    return result
```



```python
# Apply ICP to align the source to the target
 # distance threshold for correspondence
threshold=1000
trans_init = np.eye(4)  # initial transformation
reg_icp = o3d.pipelines.registration.registration_icp(
    st, tt, threshold, np.eye(4),
    o3d.pipelines.registration.TransformationEstimationPointToPoint(with_scaling=True),
    o3d.pipelines.registration.ICPConvergenceCriteria(max_iteration=1000, relative_fitness=.000000000000001, relative_rms
)

```

```python
bbox2 = st.get_minimal_oriented_bounding_box()
st.paint_uniform_color([1, 0, 0])
target_down.paint_uniform_color([0, 1, 0])
bbox2.extent = bbox2.extent * 1.15

```

by applying the inverse of the two transformations and cropping according to the bounding box, we can transform the initial model into an axis-aligned model of only the head.

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

Week13 report

Ruiqing Wang | CiChild CV team

- What progress did you make in the last week?

1. Trained network on Sleap
2. Run DEMO video export and validate results
3. Attend Cichild group meeting and discussed with Bree with current work on Sleap
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. Keep training network on SLEAP and refine dataset
2. Meet with Cichild CV team to discuss current progress

- Is anything blocking you from getting work done?

N/A

Paper abstract

Paper: The quest to develop automated systems for monitoring animal behavior
https://www.sciencedirect.com/science/article/pii/S0168159123001727

Abstract: Automated behavior analysis (ABA) is an emerging field focused on developing technologies to automatically detect and analyze animal behaviors across various species, potentially alleviating the labor-intensive task of manual video analysis by ethologists. However, the effectiveness of ABA is contingent upon the quality of training data, which is often limited due to a lack of formal ethograms and small, non-representative datasets that hinder generalizability. To enhance the robustness and applicability of ABA tools, interdisciplinary collaboration among ethologists, computer scientists, and data engineers is essential, alongside the establishment of standardized reference datasets and best practices for validation to ensure reliable behavioral assessments across diverse contexts.

Review Method:

Given the high density and homogeneity of these animals—often housed in large groups of similar breed, sex, and life stage—ABA faces significant challenges in accurately distinguishing individual behaviors due to frequent occlusion and overlap.

The critical importance of robust and diverse datasets for the effective training and validation of Automated Behavior Analysis (ABA) systems in livestock and poultry research. It highlights that many studies rely on small datasets, which may lead to solutions that are not generalizable across different contexts or animal populations due to insufficient variability in individual behaviors, environments, and data capture methods. Key considerations for dataset evaluation include the temporal and spatial independence of data, the representativeness of the sample in terms of age, breed, and social dynamics, and the contextual diversity of the environments in which the data were collected, all of which are essential for developing reliable and adaptable ABA tools.

The text highlights the limitations of current Applied Behavioral Analysis (ABA) solutions in ethology, particularly their focus on single behavioral issues rather than a comprehensive understanding of animal interactions and health indicators. Many existing ABA programs are designed for controlled laboratory settings, making them less adaptable for real-world applications, especially when dealing with complex social behaviors and environmental variability. Furthermore, the nascent state of the field and the limited market for commercial ABA products hinder the development of user-friendly, versatile tools that can effectively address the multifaceted nature of animal behavior and welfare.

Scripts and Code Blocks

This week I mainly trained the network and modify data annotations. There were not much code updated.

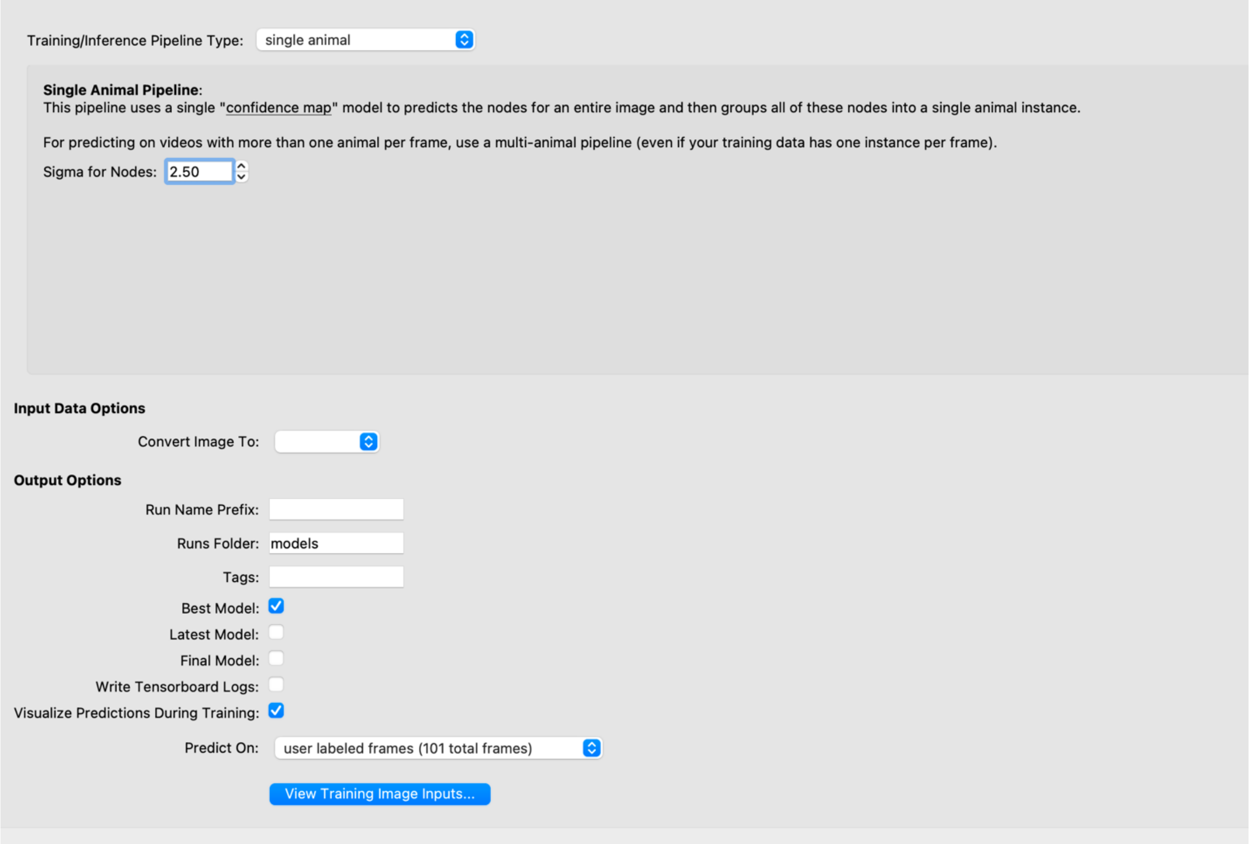
Documentation

For SLEAP network training, here is the link:
https://sleap.ai/notebooks/Training_and_inference_using_Google_Drive.html
Notes on training pipelines:

- for a **bottom-up** pipeline approach: multi_instance.json
- for a **top-down** pipeline, you'll have a different profile for each of the models: centroid.json and centered_instance.json,
- for a **single animal** pipeline: single_instance.json (this is the pipeline we assume here).

Here is the two algorithm notes:
In **top-down** mode for multiple animals, a network first finds each animal and then a separate network estimates the pose of each found animal:
In **bottom-up** mode, a network first finds all of the body parts in an image, and then another network groups them into animal instances using part affinity fields.

After that we need specify the hyperparameters.



Training/Inference Pipeline Type: single animal

Single Animal Pipeline:
This pipeline uses a single "confidence map" model to predicts the nodes for an entire image and then groups all of these nodes into a single animal instance.

For predicting on videos with more than one animal per frame, use a multi-animal pipeline (even if your training data has one instance per frame).

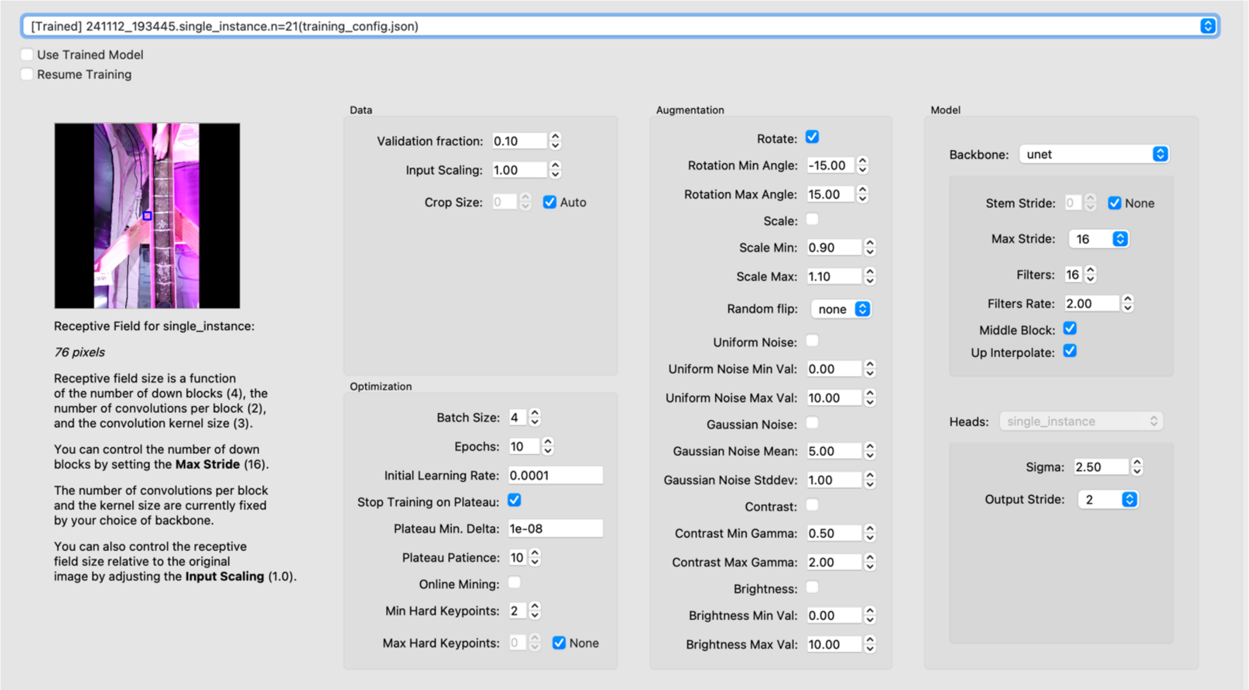
Sigma for Nodes: 2.50

Input Data Options

Convert Image To:

Output Options

Run Name Prefix:
Runs Folder: models
Tags:
Best Model: ☑
Latest Model: ☐
Final Model: ☐
Write Tensorboard Logs: ☐
Visualize Predictions During Training: ☑
Predict On: user labeled frames (101 total frames)



[Trained] 241112_193445.single_instance.n=21(training_config.json)

☐ Use Trained Model
☐ Resume Training

Receptive Field for single_instance:

76 pixels

Receptive field size is a function of the number of down blocks (4), the number of convolutions per block (2), and the convolution kernel size (3).

You can control the number of down blocks by setting the **Max Stride** (16).

The number of convolutions per block and the kernel size are currently fixed by your choice of backbone.

You can also control the receptive field size relative to the original image by adjusting the **Input Scaling** (1.0).

Data

Validation fraction: 0.10
Input Scaling: 1.00
Crop Size: 0 ☑ Auto

Optimization

Batch Size: 4
Epochs: 10
Initial Learning Rate: 0.0001
Stop Training on Plateau: ☑
Plateau Min. Delta: 1e-08
Plateau Patience: 10
Online Mining: ☐
Min Hard Keypoints: 2
Max Hard Keypoints: 0 ☑ None

Augmentation

Rotate: ☑
Rotation Min Angle: -15.00
Rotation Max Angle: 15.00
Scale: ☐
Scale Min: 0.90
Scale Max: 1.10
Random flip: none
Uniform Noise: ☐
Uniform Noise Min Val: 0.00
Uniform Noise Max Val: 10.00
Gaussian Noise: ☐
Gaussian Noise Mean: 5.00
Gaussian Noise Stddev: 1.00
Contrast: ☐
Contrast Min Gamma: 0.50
Contrast Max Gamma: 2.00
Brightness: ☐
Brightness Min Val: 0.00
Brightness Max Val: 10.00

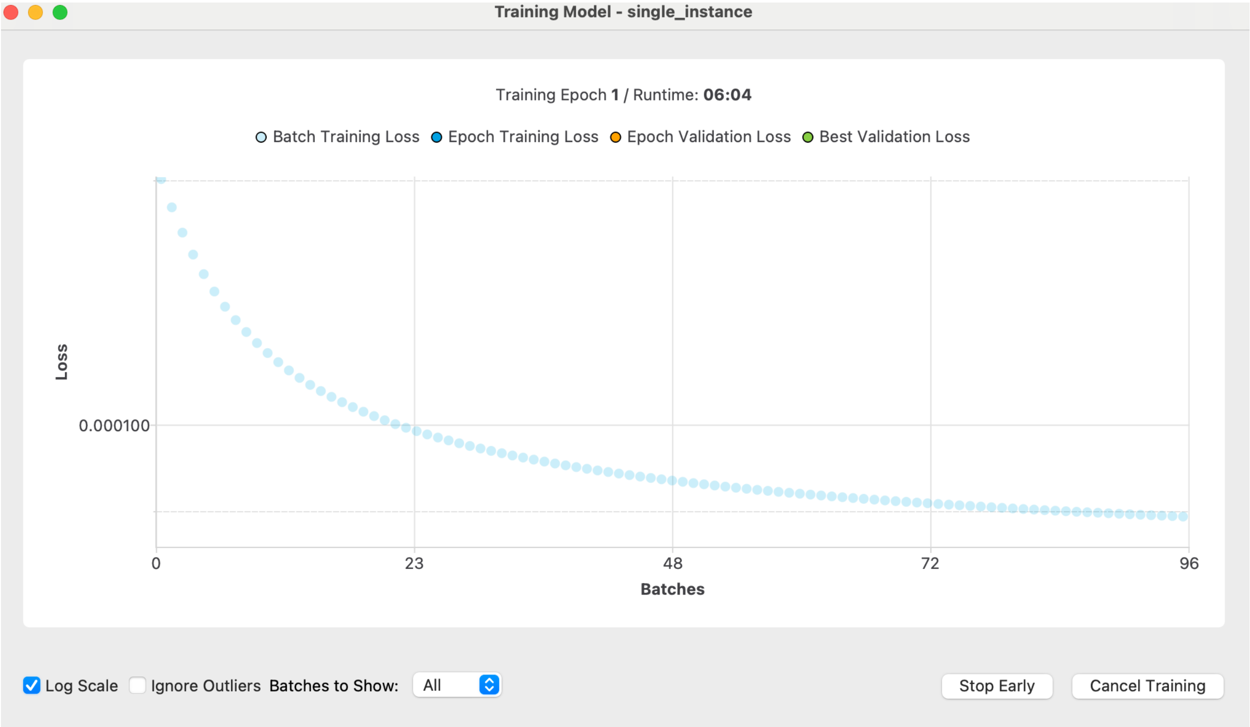
Model

Backbone: unet
Stem Stride: 0 ☑ None
Max Stride: 16
Filters: 16
Filters Rate: 2.00
Middle Block: ☑
Up Interpolate: ☑

Heads: single_instance
Sigma: 2.50
Output Stride: 2

This model configuration includes hyperparameters across five main categories: **Data**, **Augmentation**, **Optimization**, **Model**, and **Heads**. In Data, a 10% validation fraction is set with input scaling at 1.0 and automatic cropping. Augmentation applies rotation between -15° and 15°, scales images from 0.9 to 1.1, and adjusts contrast and brightness, but noise and random flip are off. Optimization uses a batch size of 4, a learning rate of 0.0001, and stops on a plateau after 10 patience steps. The Model section uses a UNet backbone with a max stride of 16 and 16 filters, while up-sampling is enabled. The Heads section sets a sigma of 2.5 and an output stride of 2, defining keypoint detection accuracy.

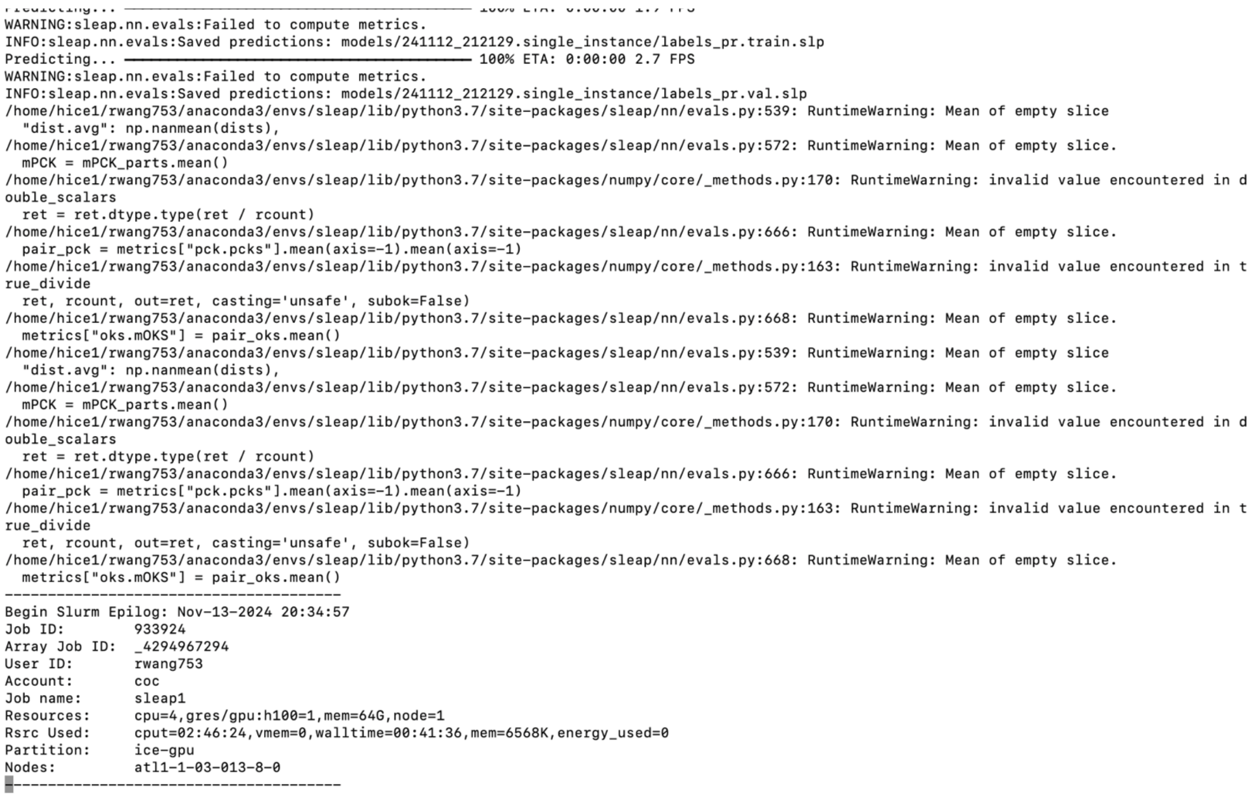
Results Visualization and Code Validation



Here is one example of running the training steps. There will be epochs loss shown in batch size.

Proof of Work and code validation

Currently I have system early stop which due to the small learning rate. I will try to fix this problem and get this work done soon.



```
Predicting... ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 100% ETA: 0:00:00 1.7 FPS
WARNING:sleap.nn.evals:Failed to compute metrics.
INFO:sleap.nn.evals:Saved predictions: models/241112_212129.single_instance/labels_pr.train.slp
Predicting... ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 100% ETA: 0:00:00 2.7 FPS
WARNING:sleap.nn.evals:Failed to compute metrics.
INFO:sleap.nn.evals:Saved predictions: models/241112_212129.single_instance/labels_pr.val.slp
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:539: RuntimeWarning: Mean of empty slice
  "dist.avg": np.nanmean(dists),
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:572: RuntimeWarning: Mean of empty slice.
  mPCK = mPCK_parts.mean()
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/numpy/core/_methods.py:170: RuntimeWarning: invalid value encountered in d
ouble_scalars
  ret = ret.dtype.type(ret / rcount)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:666: RuntimeWarning: Mean of empty slice.
  pair_pck = metrics["pck.pcks"].mean(axis=-1).mean(axis=-1)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/numpy/core/_methods.py:163: RuntimeWarning: invalid value encountered in t
rue_divide
  ret, rcount, out=ret, casting='unsafe', subok=False)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:668: RuntimeWarning: Mean of empty slice.
  metrics["oks.mOKS"] = pair_oks.mean()
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:539: RuntimeWarning: Mean of empty slice
  "dist.avg": np.nanmean(dists),
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:572: RuntimeWarning: Mean of empty slice.
  mPCK = mPCK_parts.mean()
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/numpy/core/_methods.py:170: RuntimeWarning: invalid value encountered in d
ouble_scalars
  ret = ret.dtype.type(ret / rcount)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:666: RuntimeWarning: Mean of empty slice.
  pair_pck = metrics["pck.pcks"].mean(axis=-1).mean(axis=-1)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/numpy/core/_methods.py:163: RuntimeWarning: invalid value encountered in t
rue_divide
  ret, rcount, out=ret, casting='unsafe', subok=False)
/home/hice1/rwang753/anaconda3/envs/sleap/lib/python3.7/site-packages/sleap/nn/evals.py:668: RuntimeWarning: Mean of empty slice.
  metrics["oks.mOKS"] = pair_oks.mean()
---------------------------------------
Begin Slurm Epilog: Nov-13-2024 20:34:57
Job ID:        933924
Array Job ID:  _4294967294
User ID:       rwang753
Account:       coc
Job name:      sleap1
Resources:     cpu=4,gres/gpu:h100=1,mem=64G,node=1
Rsrc Used:     cput=02:46:24,vmem=0,walltime=00:41:36,mem=6568K,energy_used=0
Partition:     ice-gpu
Nodes:         atl1-1-03-013-8-0
---------------------------------------
```

Early stopping and learning rate adjustments are crucial for optimizing model training. In this case, early stopping was triggered because the validation loss stopped improving, I will work on this part.

Next Week's Proposal

1. Keep training network on SLEAP
2. Refine the dataset and write documents
3. Meet with Cichild CV team to discuss current progress

Week 13 Document Submission

Lizard X-RAY Landmark Group

Mercedes Quintana

What progress did you make in the last week?

- Continued to work on website
- Looked through new training data
- Trained new data with new model

What are you planning on working on next?

- Analyze new model outputs
- Help Ayush create the website
- Continue to update the website

Is anything blocking you from getting work done?

- Nope

Abstracts:

URL: https://www.sciencedirect.com/science/article/pii/S1574954122003971

'I know the tiger by his paw': A non-invasive footprint identification technique for monitoring individual Amur tigers (Panthera tigris altaica) in snow

Apex predator populations are in decline around the world. Many exist at low density and are elusive, making the acquisition of reliable data on their numbers and distribution a considerable challenge. The Amur tiger (Panthera tigris altaica) is the largest of the five extant sub-species of tiger. The single most significant, contiguous population, an estimated 550 animals, exists in the Russian Far East, with smaller populations on the far eastern SinoRussian border. For the last few decades, active efforts on the part of Chinese authorities have encouraged the recolonization of these populations back to their former ranges in Northeast China. Reliable data on Amur tiger numbers and distribution are required to assess population recovery at the landscape scale. Footprints, ubiquitous in the snow over range areas, could inform on these baseline data. This paper describes a statistically robust, cost-effective and non-invasive footprint identification technique (FIT) to identify individual tigers from footprints in snow. It is based on a rigorous data collection and data-processing protocol, combined with a crossvalidated discriminant analysis method. A Ward's clustering technique provides a visual output of individual classification. The analytical tools are packaged in a user-friendly analytical interface. Between December 2011 and December 2012, we collected a series of 605 footprint images from 44 captive individual Amur tigers for a reference database from which to derive a classification algorithm. The 23 females and 21 males ranged in age from 3 to 13 years (female mean age 7.95 +/− 0.18; male mean age 8.08 +/− 0.19). 128 measurements (areas, lengths and angles) were taken from each print and analyzed with the FIT add-in to JMP software. The derived classification algorithm was then applied to 21 footprint trails collected from an unknown number of freeranging Amur tigers during 2012 and 2015/2016. The algorithm predicted 7 Amur tigers at the site surveyed in 2012, and 4 tigers surveyed at two sites in 2015/16. We demonstrate that the footprint identification technique translates traditional tracking methodologies into a statistically robust and objective analytical tool that can be deployed by both scientists and local communities to monitor the recovery of big cat populations.

Summary: This paper demonstrates a method for tracking and monitoring rare, endangered tigers through their footprints in the snow. This model uses an interface to make the technology use easier and was validated on populations of captive and free-ranging tigers.

Scripts and Code Blocks:

Last week, now that using the auto processed images was decided, I went through and found the best training data from the rest of the data that we have not used yet. This involved grading the images, which are now available on the github under 2411_grade_rest_images.txt. Now that the images are added in, we have the lowest error yet:

Training error (average pixel deviation): 0.7939865005902391

Testing error (average pixel deviation): 6.610369045318518

Documentation:

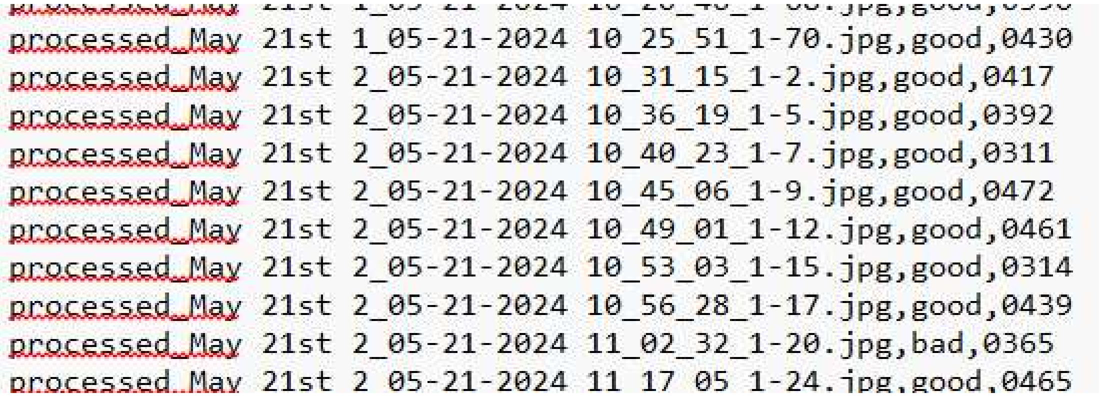
I have no new code to document as of now.

Script Validation:

I have no validation steps now.

Results Visualization / Proof of Work:

Here is some of the file mapping and image grading I did before training the new model:



```
processed_May 21st 1_05-21-2024 10_25_51_1-70.jpg,good,0430
processed_May 21st 2_05-21-2024 10_31_15_1-2.jpg,good,0417
processed_May 21st 2_05-21-2024 10_36_19_1-5.jpg,good,0392
processed_May 21st 2_05-21-2024 10_40_23_1-7.jpg,good,0311
processed_May 21st 2_05-21-2024 10_45_06_1-9.jpg,good,0472
processed_May 21st 2_05-21-2024 10_49_01_1-12.jpg,good,0461
processed_May 21st 2_05-21-2024 10_53_03_1-15.jpg,good,0314
processed_May 21st 2_05-21-2024 10_56_28_1-17.jpg,good,0439
processed_May 21st 2_05-21-2024 11_02_32_1-20.jpg,bad,0365
processed May 21st 2 05-21-2024 11 17 05 1-24.jpg,good,0465
```

Next Week Proposal:

Next week I am going to continue to make sure the website is updated. I am going to analyze the new data from the model and prepare to share it on Wednesday. I am going to continue to prepare the new website to share on at the meeting on Wednesday.