Week 15 Document Submission

Jacob Dallaire

November 29, 2024

1. Paper

A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, Świnoujście, Poland, 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

SUMMARY

The text discusses the prominence of Convolutional Neural Networks (CNNs) in image analysis and classification, highlighting their state-of-the-art performance alongside challenges such as overfitting, generalization issues, and vulnerability to adversarial attacks. It emphasizes the critical role of data augmentation in addressing the scarcity of high-quality training datasets, detailing traditional methods like affine transformations and advanced techniques such as Generative Adversarial Networks (GANs) and style transfer. The authors propose that style transfer can significantly enhance data augmentation by generating high-quality synthetic images that maintain the semantic content of original images while varying their stylistic attributes, thus improving the robustness and performance of deep learning models in various applications, particularly in medical imaging.

2. Scripts

Rewrote object detection training script to use a yolov8 architecture.

```
backbone = keras_cv.models.YOLOV8Backbone.from_preset(
    "yolo_v8_m_backbone_coco",
    load_weights=True
)

yolo = keras_cv.models.YOLOV8Detector(
    num_classes=1,
    bounding_box_format="xyxy",
    backbone=backbone,
    fpn_depth=3,
)
```

This involved reformatting input data to yolo format

```
def load_image_and_labels(image_path, annotation):
    # Load the image
    image = tf.io.read_file(image_path)
    image = tf.image.decode_jpeg(image, channels=3)

    # Store the original size for later scaling
    original_size = tf.shape(image)[:2]  # Height, Width

    # Resize the image to a standard size
    image_resized = tf.image.resize(image, [320, 320])

    # Get the corresponding annotations for this image
    bboxes= parse_annotations(annotation)
    #bboxes = np.concatenate([bboxes, c_score.reshape(-1, 1)], axis=-1)
    #bboxes = tf.ragged.constant(bboxes, dtype=tf.float32)
    classes = tf.convert_to_tensor(tf.zeros((tf.shape(bboxes)[0]),
dtype=tf.float32))

    # List to hold rotated images and boxes
    outputs = []
    angles = [90, 180]#, 270]
    outputs.append((image_resized, {"bbox": bboxes, "classes": classes}))
    for angle in angles:
        rotated_image = rotate_image(image_resized, angle)
        rotated_bboxes = rotate_bboxes(bboxes, angle)
        outputs.append((rotated_image,{"bbox": rotated_bboxes, "classes":
classes}))

    return outputs

def load_dataset(annotations_folder):
    images = []
    bboxes = []
    classes = []
    dataset = []
    bounding_boxes = {"boxes": [], "classes": []}
    # Iterate through JSON files in the specified folder
    for filename in os.listdir(annotations_folder):
        if filename.endswith('.json'):
            json_path = os.path.join(annotations_folder, filename)
```

```
            annotation = load_json_annotations(json_path)
            outputs =
load_image_and_labels(f'C:/Lizards/bbox/{annotation["imagePath"]}',annotation)
            for row in outputs:
                images.append(row[0])
                bboxes.append(row[1]["bbox"])
                classes.append(row[1]["classes"])
    #images_tensor = tf.convert_to_tensor(images)

    # Create dataset from image paths and annotations
    images = tf.stack(images)
    bounding_boxes["boxes"] = bboxes
    bounding_boxes["classes"] = classes
    # Return a tf.data.Dataset
    dataset = tf.data.Dataset.from_tensor_slices({
        "images": images,
        "bounding_boxes": bounding_boxes
    })

    dataset = dataset.cache().shuffle(1000).batch(4).prefetch(tf.data.AUTOTUNE)

    for batch in dataset.take(1):
        print("Images shape:", batch["images"].shape)
        print("Boxes shape:", batch["bounding_boxes"]["boxes"].shape)
        print("Classes shape:", batch["bounding_boxes"]["classes"].shape)
    return dataset
```

I have also been working on updating the testing script but have not completed this work.

3. Documentation

I have been working on changing the base models used to pretrain on. After updating the object detection model it unfortunately only is outputting -1 for all boxes and confidence scores. Training needs to be looked into further to find the cause.

4. Next Weeks Proposal

Over the weekend I will make sure everything is well documented for project handoff in the spring.

Weekly Report

Philip Woolley

2024-11-29

Time Log Reponse:

- Wrote Problem statement and Design sections of final project report.
- Retrain model with additional data and new metric. Continue final project documentation. Change website structure to match request from Bree
-

1 Abstract

Abstract Point cloud registration from laser scanning data is a technique to establish the mapping relationship between source and target point clouds, which has been widely used in automatic 3D reconstruction, pose estimation, localization, and navigation. While algorithms like Super4PCS and MSSF-4PCS can achieve registration without initial poses, they are relatively slow, less accurate, and require iterations. To address these issues, we propose a 3D point cloud registration algorithm based on interval segmentation and multi-dimensional feature. Firstly, the source and target point clouds are segmented internally and the point cloud curvature is designed to narrow down the search range for the registration between the segmented point clouds. Secondly, the corresponding four-point sets in the segmented areas of the source and target point clouds are determined using affine invariance constraints. Finally, a multi-dimensional feature vector based on curvature features and fast point feature histogram is established to determine the unique corresponding four-point set pairs, and the rigid body transformation matrix is solved accordingly. Our algorithm is tested on publicly available 3D point cloud data models Bunny, Dino, Dragon, and Horse from Stanford University. Results showed that our algorithm improved registration accuracy by 24.39% and registration efficiency by 46.21% compared to the MSSF-4PCS point cloud registration algorithm. Multiple sets of experimental results confirmed the effectiveness of our algorithm. The proposed 3D point cloud registration is proved to be fast with high accuracy, which can be utilized for automatic segmentation, reconstruction, and modelling from Laser Scanning Data.

Summary This paper describes 3D point cloud registration for models created using laser scanning. I chose to review this paper to see the data presentations they use, as I am looking for clear, concise ways to display the results of the registration portion of my 3D pipeline. I think the primary step-by-step model visualization the authors provide is very clear, and is similar to how I intend to present my process. The authors methods are meant for full model-to-model matching, and so are not particularly useful for my project. The authors present full algorithm pseudocode, which can be useful but would bog down my report. They also present a graph of RMSE vs hyperparameter values, which is an interesting idea I had not thought of.

Citation A. Xu, L. Rao, G. Fan and N. Chen, "Fast and High Accuracy 3D Point Cloud Registration for Automatic Reconstruction From Laser Scanning Data," in IEEE Access, vol. 11, pp. 42497-42509, 2023, doi: 10.1109/ACCESS.2023.3270502. keywords: Point cloud compression;Three-dimensional displays;Feature extraction;Real-time systems;Approximation algorithms;Matrix converters;Histograms;Point cloud registration;point cloud curvature;affine invariance;multi-dimensional features,

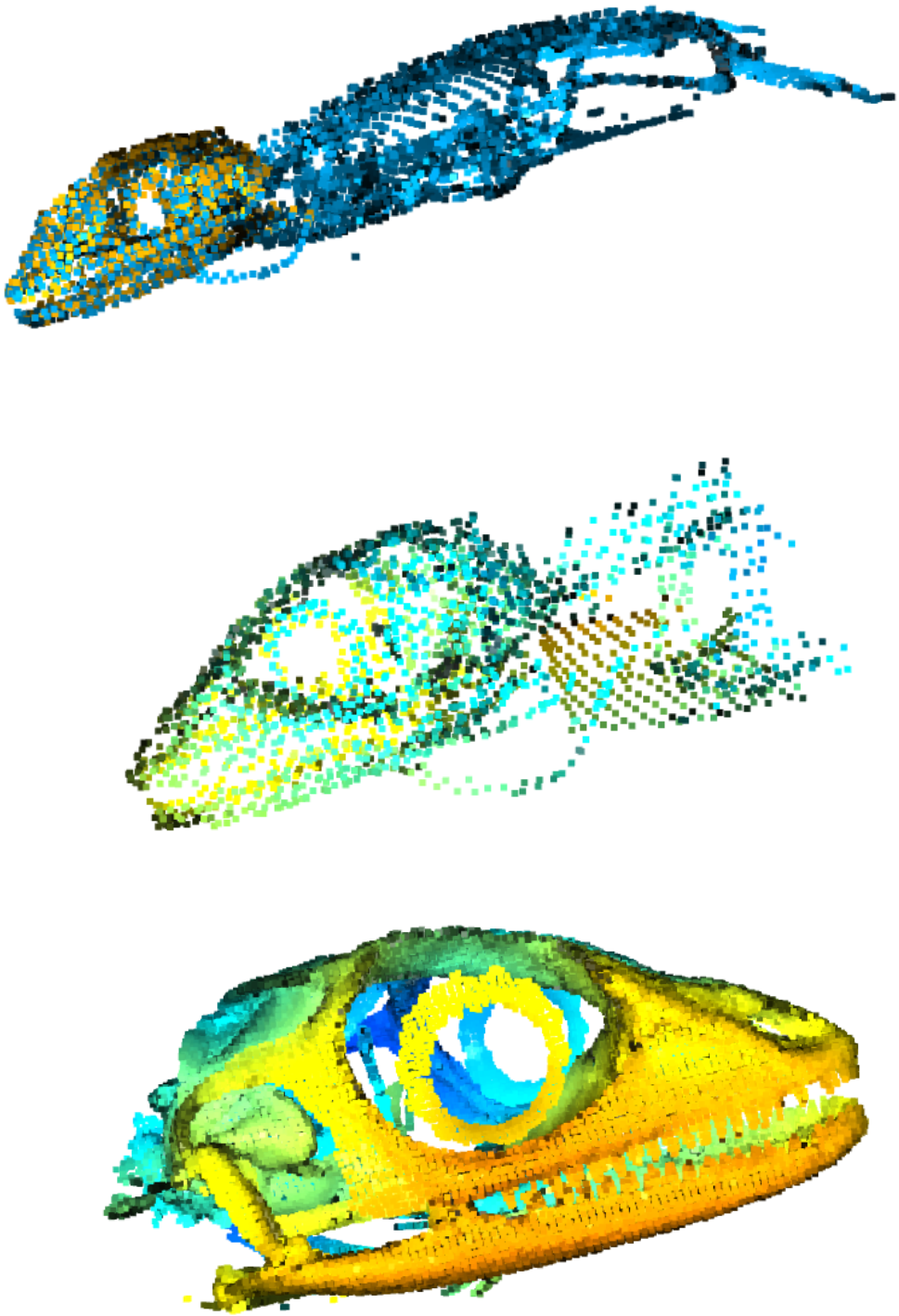
2 Scripts and Code Blocks

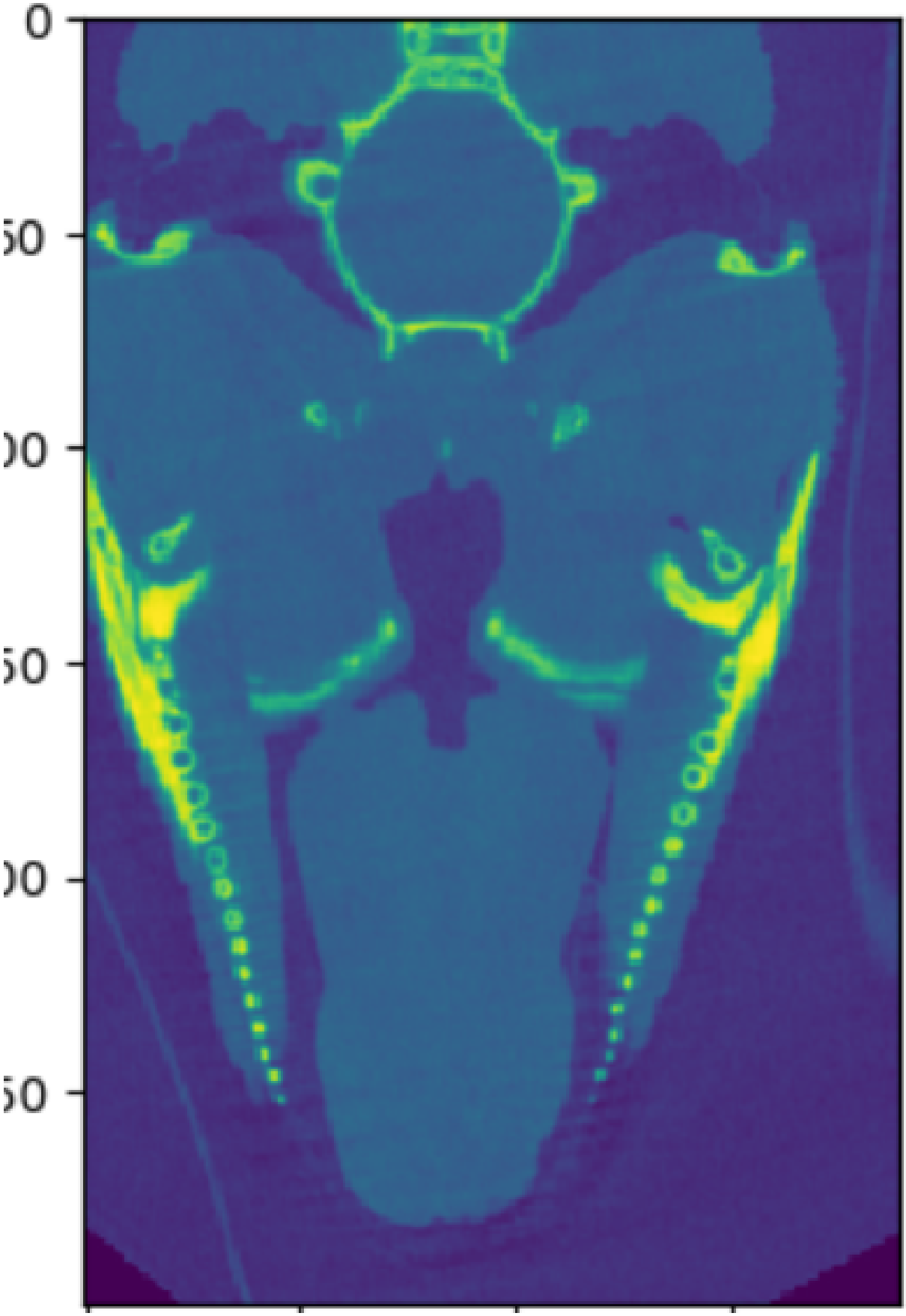
This week I worked only on documentation, so there are not code blocks.

3 Documentation

4 Script Validation (Optional)

5 Results Visualization




0
50
00
50
00
50

6 Proof of Work

Please see results visualization section

7 Next Week's Proposal

- Finish code cleanup
- Continue project documentation
- Reformat blog page as requested by Bree

Week14 report

Ruiqing Wang | CiChild CV team

- What progress did you make in the last week?
  1. Finish wrap up summary on DLC and SLEAP progress and protocol
  2. Rewrite up the liarzd report on DLC
  3. Attend Cichild group meeting with current work on Sleap
  4. Help assembling paper report submissions and address submission situation.
- What are you planning on working on next?
  1. Finish up writing on report
  2. Meet with Cichild CV team to discuss current progress
- Is anything blocking you from getting work done?

N/A

Paper abstract

Paper: Using machine learning and DeepLabCut in animal behavior
https://link.springer.com/article/10.1007/s10211-022-00397-y

Abstract: The study of animal behavior traditionally relies on direct observation, which is often time-consuming and resource-intensive. The commentary highlights the potential of automation through advanced software like DeepLabCut (DLC) for markerless tracking of animal movements, thereby alleviating the burdens associated with manual observation and the ethical concerns of physical tagging. By leveraging artificial intelligence and deep learning, DLC offers a cost-effective and efficient solution for pose estimation and tracking, enabling researchers to enhance data collection and analysis in behavioral studies while adhering to animal welfare standards.

Review Method:
DeepLabCut (DLC) operates in two modes—single-animal and multi-animal—each employing distinct workflows for pose estimation and tracking. In single-animal DLC, pose estimation yields x, y coordinates with timestamps, while multi-animal DLC separates pose estimation from tracking, necessitating precise pose identification to accurately associate body parts with individual identities across frames. This bifurcation allows for enhanced data collection efficiency and the potential for robust analysis of rare behaviors, although it introduces challenges such as the need for manual refinement in tracking and the risk of accumulating errors without human oversight, underscoring the trade-offs inherent in automated behavioral analysis.

Scripts and Code Blocks

This week I mainly working on writing up current report and so some finish-up analysis

Documentation

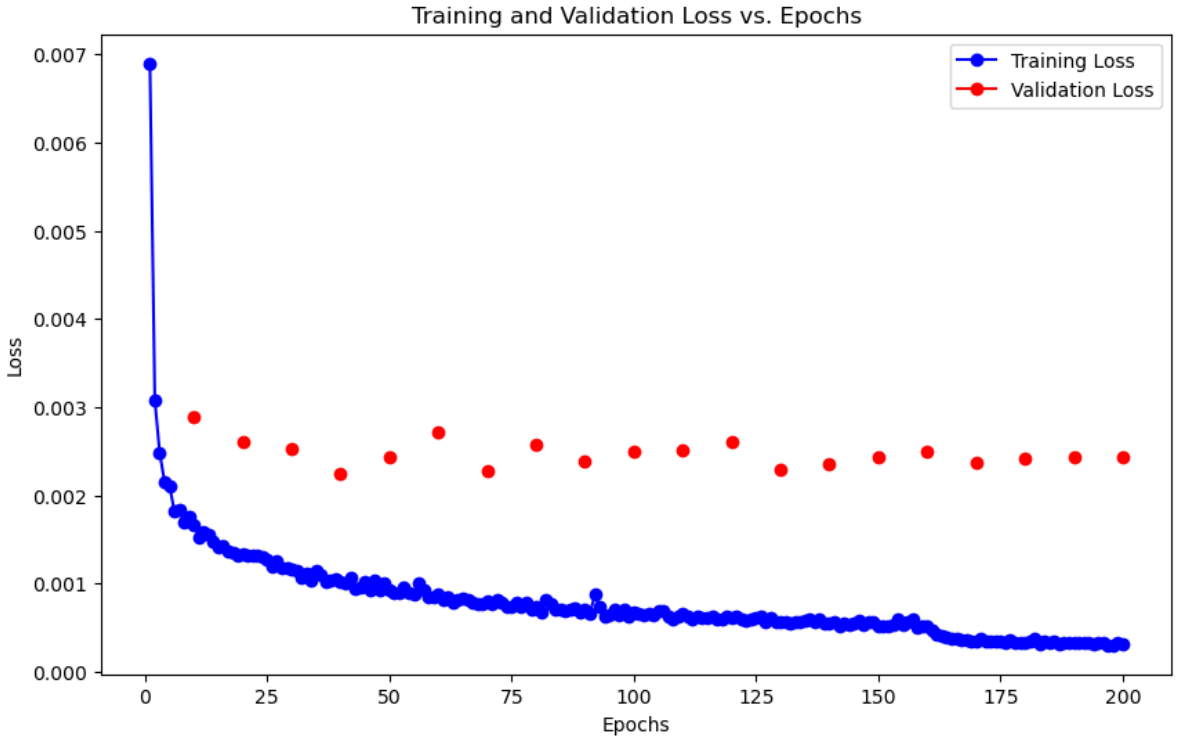
For user-guide of DLC and sleap, here is the link
https://gtvault-my.sharepoint.com/:w:/g/personal/rwang753_gatech_edu/EVD-vRI1ixlBrdBxLJMs-fEB2ZHJftgK1qz5CeJZoXsecQ?e=bZAUN2

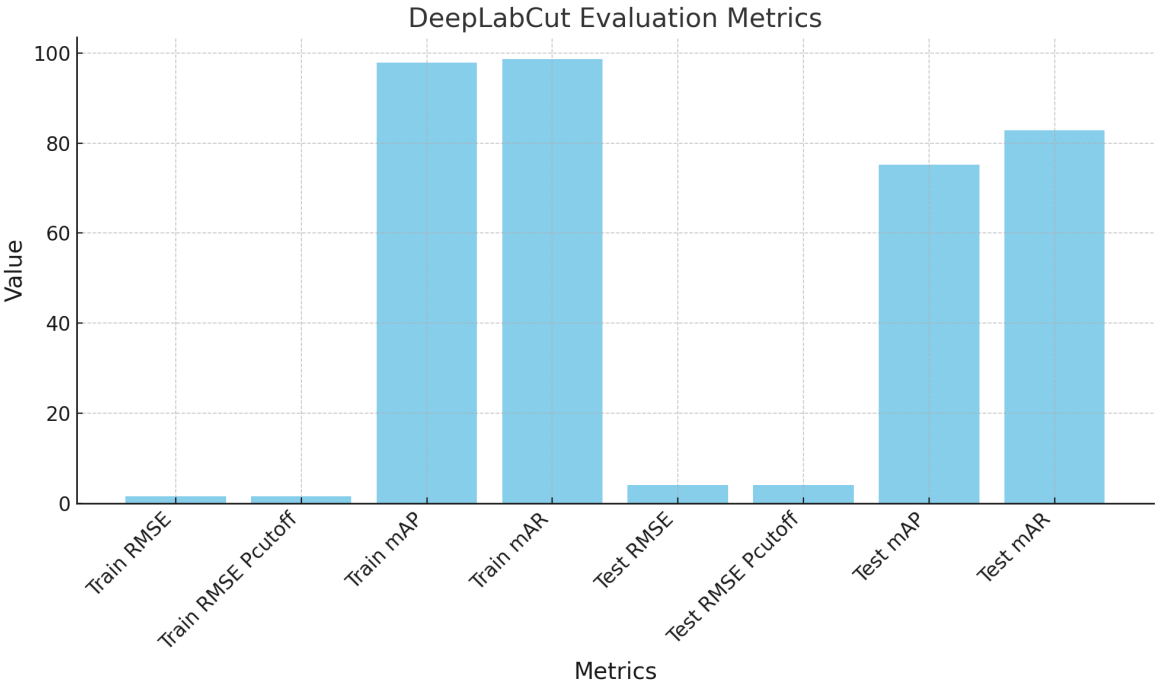
For specific questions on DLC and SLEAP method from Bree, here is the link:
https://gtvault-my.sharepoint.com/:w:/g/personal/bshi42_gatech_edu/EUfajG9yXahNoNQ7_aKjeukBm4NQ7uPagDRaasxQ-8TYYg?e=ExcYK1

For final lizard report, here is the link:
https://gtvault-my.sharepoint.com/:w:/g/personal/bshi42_gatech_edu/EYzZJ44yBrdOg1_6B6p4c-oBMG_zc3qVlx-UIuxUH3TWfw?e=5knKdp

Results Visualization and Code Validation

I rerun the model evaluation on DLC model, here is the result:





I added training loss to the top picture, and added training and test map result for current model evaluation. The evaluation results for the DeepLabCut model show strong performance during training and testing. The train RMSE and train RMSE Pcutoff values are both relatively low at 1.64, indicating good model accuracy during training. The train mAP (97.80) and train mAR (98.62) suggest that the model excels in both mean average precision and mean average recall, reflecting its high effectiveness in accurately predicting keypoints and capturing movement. The test RMSE and test RMSE Pcutoff values are slightly higher (4.11), indicating a slight increase in prediction error on the test set. Despite this, the test mAP (75.20) and test mAR (82.78) remain impressive, demonstrating that the model maintains strong generalization and performance on unseen data, though there is some drop-off compared to training.

Proof of Work and code validation

Please check above link.
The command for the model evaluation as :

```
deeplabcut.evaluate_network(config_path, Shuffles=[1],
plotting=True)
```

Next Week's Proposal

1. Determine the issue for SLEAP
2. Finish up writing
3. Meet with Cichild CV team to discuss current progress