Week 7 Document Submission

Jacob Dallaire

October 4, 2024

1. Paper

Z. -Q. Zhao, P. Zheng, S. -T. Xu and X. Wu, "Object Detection With Deep Learning: A Review," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

SUMMARY

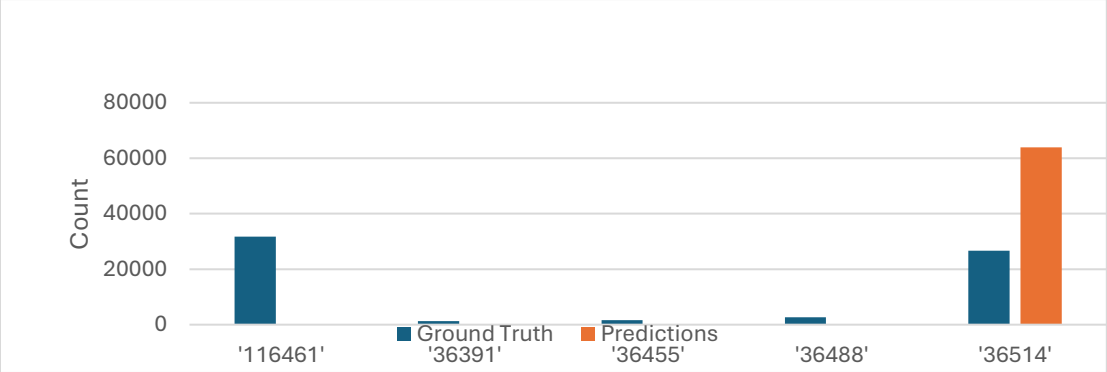
The text emphasizes the importance of object detection as a critical component of computer vision, which involves not only classifying images but also accurately locating and identifying objects within them. It outlines the evolution from traditional methods reliant on handcrafted features to advanced deep learning techniques, particularly convolutional neural networks (CNNs), which enable the automatic learning of complex, hierarchical features. The review also highlights various architectures and approaches, such as R-CNN and its derivatives, while discussing the challenges posed by variations in object scale, occlusion, and background complexity, ultimately suggesting future research directions to enhance detection performance across diverse applications.

2. Scripts

No finished scripts to share.

3. Documentation

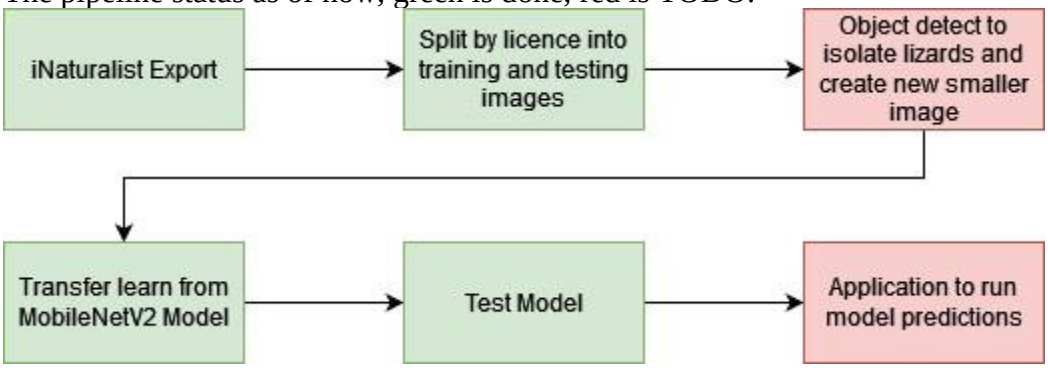
I did some analysis of the model I have already trained. I have also set up a new dataset just for testing how the model generalizes. While it has minimal loss and almost perfect f1score on the training data the new testing data only has a 42% accuracy. This accuracy is misleading however as it is a result of class imbalances. The model in fact predicts the same class, with 5 exceptions, for all input images as seen below.



Due to this poor performance, I am working on a script to do object detection to segment out the lizards and ignore as much of the background as possible however the pretrained yolo5 and mobilenetV2 models I have explored so far do not include a class for reptiles

or lizards. I am searching for a pretrained model one that does as the annotation of the training dataset needed to transfer learn with a new lizard class is too extensive to perform in our timeline.

The pipeline status as of now, green is done, red is TODO:



4. Next Weeks Proposal

I will continue to work on implementing the new object detection step to the pipeline.

Weekly Report

Philip Woolley

2024-09-27

Time Log Reponse:

- What Progress did you make in the last week? - Identified possible automatic alignment tools. Continued segmenting training data
- What are you planning on working on next? - Continue segmenting training data. Create a script for postprocessing model results
- Is there anything blocking you? - None at this time

1 Abstract

Abstract

We present VoxelMorph, a fast learning-based framework for deformable, pairwise medical image registration. Traditional registration methods optimize an objective function for each pair of images, which can be time-consuming for large datasets or rich deformation models. In contrast to this approach, and building on recent learning-based methods, we formulate registration as a function that maps an input image pair to a deformation field that aligns these images. We parameterize the function via a convolutional neural network (CNN), and optimize the parameters of the neural network on a set of images. Given a new pair of scans, VoxelMorph rapidly computes a deformation field by directly evaluating the function. In this work, we explore two different training strategies. In the first (unsupervised) setting, we train the model to maximize standard image matching objective functions that are based on the image intensities. In the second setting, we leverage auxiliary segmentations available in the training data. We demonstrate that the unsupervised model's accuracy is comparable to state-of-the-art methods, while operating orders of magnitude faster. We also show that VoxelMorph trained with auxiliary data improves registration accuracy at test time, and evaluate the effect of training set size on registration. Our method promises to speed up medical image analysis and processing pipelines, while facilitating novel directions in learning-based registration and its applications. Our code is freely available at this http URL.

Summary This paper proposes a Deep Learning method for image registration, which is the process of converting two images into a common coordinate system. In the context of my project, this is a possible solution to the problem of aligning the lizard head with the x,y,z planes of the slicer so that all of the teeth and the jaw a clearly imaged when sliced. This is necessary for performance of panoptic segmentation of the teeth and jaws. This paper focuses on a method for Deformable registration, which is similar to the task achieved by the ALPACA tool. My current plan is to use non-deformable registration, although there are other approaches which could make use of this. Regardless, this paper clearly presents the intuition behind their approach and the visualizations are very clear, both of which I can take inspiration for for explaining the registration step of my pipeline.

Citation

Balakrishnan, Guha, et al. "Voxelmorph: a learning framework for deformable medical image registration." IEEE transactions on medical imaging 38.8 (2019): 1788-1800.

2 Scripts and Code Blocks

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

This week, my work was continued segmentation as well as research for future implementation for image registration.

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

Week 7 Document Submission

Lizard X-RAY Landmark Group

Mercedes Quintana

What progress did you make in the last week?

- Continued to work on website (Everything actually shows now!)
- Investigated the images for ones that will best represent future Xrays
- Looked at the differences between landmarked results on manual and auto images

What are you planning on working on next?

- Find the best training data for the model
- Continue to update the website
- Find if the manual or auto images perform better

Is anything blocking you from getting work done?

- Nope

Abstracts:

URL: https://arxiv.org/pdf/2401.13191

TOWARDS MULTI-DOMAIN FACE LANDMARK DETECTION WITH SYNTHETIC DATA FROM DIFFUSION MODEL

Recently, deep learning-based facial landmark detection for in-the-wild faces has achieved significant improvement. However, there are still challenges in face landmark detection in other domains (e.g. cartoon, caricature, etc). This is due to the scarcity of extensively annotated training data. To tackle this concern, we design a two-stage training approach that effectively leverages limited datasets and the pre-trained diffusion model to obtain aligned pairs of landmarks and face in multiple domains. In the first stage, we train a landmark conditioned face generation model on a large dataset of real faces. In the second stage, we fine-tune the above model on a small dataset of image-landmark pairs with text prompts for controlling the domain. Our new designs enable our method to generate high-quality synthetic paired datasets from multiple domains while preserving the alignment between landmarks and facial features. Finally, we fine-tuned a pre-trained face landmark detection model on the synthetic dataset to achieve multi-domain face landmark detection. Our qualitative and quantitative results demonstrate that our method outperforms existing methods on multi-domain face landmark detection

Summary: The paper aims to expand the style of face that can be landmarked such as those from cartoons by creating synthetic data to handle different styles of faces, with landmarks intact, so that a model can learn and distinguish these faces.

Scripts and Code Blocks:

This week I did not create any scripts to upload to Github as I was manually checking the images to first map the automatically processed images to the manually processed images. Since the image names are different, this task required going through each image and putting together a document with the mapping. I met with Jon and we talked about better ways to accomplish this in the future.

Documentation:

In order to get the images:

1. Find image in automatically processed
2. Find corresponding image in the manually processed
3. Put together in txt file to map images to one another

Script Validation:

I have no validation steps now.

Results Visualization / Proof of Work:

This shows the by hand mapping to the automatically processed images to the manually processed images.

```
processed_May 16th Miami 1_05-16-2024 11_25_40_1-8.jpg,0081_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_30_18_1-12.jpg,0029_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_37_22_1-14.jpg,0102_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_44_30_1-18.jpg,0082_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_49_02_1-21.jpg,0160_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_54_56_1-23.jpg,0113_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 11_59_55_1-26.jpg,0173_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 12_10_01_1-28.jpg,0170_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_13_34_1-31.jpg,0157_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_19_42_1-35.jpg,0158_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_25_13_1-37.jpg,0084_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_31_55_1-39.jpg,0156_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_40_34_1-41.jpg,0028_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 13_49_30_1-43.jpg,0031_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 14_04_50_1-46.jpg,0060_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 14_09_20_1-50.jpg,0155_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 14_14_33_1-54.jpg,0073_dorsal.jpg
processed_May 16th Miami 1_05-16-2024 14_18_34_1-56.jpg,0047_dorsal.jpg
```

Next Week Proposal:

I plan to keep working on the website to keep it updated with the new meetings and work done. I plan to keep mapping images and then find how close each landmark is to it's counterpart with both statistics and phyical images of each landmark in the dataset. While

looking through landmarks, find the ones that will serve as good training data for the future model.