

902005 - VARS Annotation Assistance

Week 51

1. Danelle Cline
 - a. Team meeting to discuss 2020 milestones and 2021 plans. No other meetings planned for 2020 and the team expects to begin VARS integration and testing in early 2021.
 - b. Finishing up database design for GraphQL API. This will serve as a data store for the machine learning visual event tracks and other ML related data needed to manage the ML lifecycle. Onboarding planned with Duane this week. Expect to deploy sometime next week. Wish me luck with the podman blues.
 - c. Training a larger D4 EfficientDet model to use to populate a test database.
 - d. Successfully tested several FathomNet and Sagemaker API calls in a web app taking shape that I'm calling DeepSeaAI. This will be both a dashboard and tool for submitting and monitoring the ML pipelines.
2. Duane Edgington
 - a. Team meeting Dec 10
 - b. Status meeting Dec 15
 - c. Update session with Danelle planned for this week.

Week 47

1. Duane Edgington
 - a. Spinning up on Amplify and Vue on AWS.

Week 46

1. Danelle Cline
 - a. Drafting documentation on JSON output for the team to work through VARS ingest issues.
 - b. Finishing up local mode SageMaker tests which has sped-up the debugging process considerably.
 - c. Waiting for approval to use a bigger instance for model training from AWS.
 - d. Working on the integrated video inference and tracking code. So far so good.

Week 44

1. Danelle Cline
 - a. Instrumentation of efficientdet for spot training almost complete. Testing in SageMaker Studio. Many issues in the entire stack from the Docker/CUDA libraries, SageMaker and Wandb logging to resolve.
2. Duane Edgington
 - a. Continuing AWS SageMaker Studio work with Danelle on tfEfficientDet detector/classification code
 - b. Successfully ran multi-epoch spot training (70% cost saving over on-demand run)

Week 43

1. Duane Edgington
 - a. Working on tfefficientdet detector/classifier to run on AWS SageMaker Studio

Week 42

1. Duane Edgington
 - a. ~Daily work sessions with Danelle.
 - b. Running training jobs on AWS SageMaker Studio.

- c. Coming up to speed on Spot instance use (saves cost, 70% in my experiment) for ML training.

Week 41

1. Duane Edgington
 - a. Successful 1 epoch training trials of deepsea-efficientdet model (Danelle's docker image) with voc_crop_896x896 image crops (and 512x512 image crops) in SageMaker Studio. Collected training time and AWS cost statistics. Interfaced with WANDB and SageMaker artifact reporting and plotting. Learning how to control costs in SageMaker Studio...

Week 39

1. Danelle Cline
 - a. CUDA deinterlacing docker image
 - b. SageMaker training almost ready for sweeps
 - c. Integrating C++ inference engine into tracking code.
2. Duane Edgington
 - a. Working through SageMaker Studio evaluation / training

Week 38

1. Duane Edgington
 - a. Work on accessing M3 prores files. Thanks to IS, we now have a mount point to titan M3 on our dev-box.
 - b. Uploaded "calpoly" mezzanine and prores file to AWS S3
 - c. Exploring AWS SageMaker Studio

Week 37

1. Danelle Cline
 - a. Instrumented [wandb](#) logging in new object detection code in preparation for training in SageMaker. This will be a more efficient model (smaller/faster).
 - b. Added data registry for FathomNet and Cal Poly data collections
 - c. Created merged repo for taxa common to MARS BOEM sites with data filtering, merging, and correct VOC formatted outputs
 - d. Sleuthed out why some data was missing in our Cal Poly data set.
 - e. Metaflow/Cortex evaluation - looking like Cortex is a better fit (supports auto scaling, batch modes, easier to debug).
 - f. Beginnings of [OpenVINO](#) docker image for model optimization
2. Duane Edgington
 - a. Exploring Metaflow, Cortex tools with Danelle

Week 36

1. Danelle Cline
 - a. Setup detection test for [NetFlex Metaflow with TensorFlow2 Retinanet](#). Metaflow works on laptop and local GPU. Next steps are to test in the Cloud.
 - b. Creating data registries for all available data for upcoming process tests. This versions the data for reproducibility and simplifies cloud testing.
 - i. <https://bitbucket.org/mbari/registry-benthic2017>
 - ii. <https://bitbucket.org/mbari/registry-fathomnet>
 - iii. <https://bitbucket.org/mbari/registry-i2map>
 - iv. CalPoly is next.

Week 35

1. Danelle Cline
 - a. After evaluating ML stacks project plans are to process collection of benthic dives for proof-of-concept. Training libraries will include Cal Poly and FathomNet labeled images for top species that we have enough localizations for.
 - b. Instrumenting TensorFlow 2 Retinanet module for Metaflow and midwater multiple object track tests.
 - c. Experimentation with [Cortex Machine Learning Stack](#)
2. Duane Edgington
 - a. Gearing up for Metaflow experiment
 - b. Reviewing AWS materials
 - c. Team meeting Aug 20 included phase 1 response discussion and preparation.

Week 34

1. Duane Edgington
 - a. Team meeting Aug 20. Work on Phase 2 proposal.
 - b. Spin up [Metaflow](#) framework -- workflow candidate

Week 33

1. Duane Edgington
 - a. Tri-weekly intern meetings. Presentation Aug 12.

Week 32

1. Danelle Cline
 - a. Summer intern meetings and working on fixing an issue with object detection model deployment failing in AWS but working okay on prem.
2. Duane Edgington
 - a. Tri-weekly intern meetings.

Week 31

1. Danelle Cline
 - a. Fixed an issue with metrics failing to log correctly following object detection training.
 - b. Meetings with summer intern, Peyton who is working on evaluating performance metrics for our multiple object [video tracker](#) for the VAA project. We are using Amazon Web Services AWS and [CVAT](#) for this effort.
2. Duane Edgington
 - a. Intern teleconferences

Week 30

1. Duane Edgington
 - a. Project tasks

Week 29

1. Duane Edgington
 - a. Proposal review
 - b. On-going intern communications

Week 28

1. Danelle Cline
 - a. Code housekeeping to get all of the dockerized components in the workflow working with semantic versioning and automated docker builds and pushes to docker hub.
 - b. Project proposal input
 - c. Conference with USCS creative coding lab about AI/ML human-performance literature review and suggestions for improved VARS workflows using AI/ML
 - d. Created AWS s3 bucket for training data collections versioned with [DVC](#) and working on code for reproducible preprocessing.
2. Duane Edgington
 - a. Validated M3 registration of MBARI2017 and SRI-Rockfish video files.
 - b. Informative zoom meeting with UCSC Leila Takayama, Henry Zhou on Human-HL interaction/infoviz

Week 26

1. Danelle Cline
 - a. Summer intern discussions on track metrics using benthic/midwater test track data. Python metric tracking package is working on our test data and should address key ML performance questions. Work is progressing quite well with Peyton! Our focus will shift to creating a 2-hour test sequence shortly.
 - b. Setup CVAT and MLFlow development servers
 - c. Working on schema design and experimenting with [Cortex machine learning API](#) as a possible tool to simplify devops and compute scaling needs.
 - d. Discussion with UCSC Creative Coding lab scheduled for Tuesday June 30th.
2. Duane Edgington
 - a. Intern Peyton is making great progress!
 - b. Team meeting scheduled for June 25.

Week 25

1. Duane Edgington
 - a. Meeting with Intern Peyton Lee Monday June 15. He is making great progress.

Week 24

1. Danelle Cline
 - a. Discussion with UCSC Creative coding lab about possible collaboration on user interface and machine learning efforts
 - b. Student discussions
2. Duane Edgington
 - a. Project kickoff with summer intern Peyton Lee. We will be planning 3x/week virtual meetups while he gets going.
 - b. Project status meeting June 11.

Week 23

1. Danelle Cline
 - a. Finished week-long GraphQL/Prisma online training - all 110+ videos and exercises. Whew.
2. Duane Edgington
 - a. Explored Google Cloud Platform buckets and integration with CoLab environments. Works!
 - b. Intern Peyton Lee starts ~June 8

Week 22

1. Duane Edgington
 - a. Some preparation for intern Peyton Lee, including CoLab / Google Cloud exploration and setup info.

Week 21

1. Danelle Cline
 - a. Meeting with UCSC Creative coding lab; this group is interested in helping with human-machine learning visualization aspects of the project.
 - b. Meeting with summer intern, Peyton Lee, Univ Washington student, to discuss intern project.
2. Duane Edgington
 - a. Work continues on registering video files into M3. Needed to update the filenames to match creation_timestamp metadata
 - b. Weekly team meeting via teleconference. VideoLab presentations to UCSC folks.
 - c. Teleconference with 2020 intern Peyton Lee.
3. Brian Schlining
 - a. Participated in demos for UCSC of our video annotation ecosystem
 - b. [Worked on verification of localized videos](#) that will be integrated into M3.
 - c. [Verified loaders of Pascal VOC localizations into M3/VARS](#)

Week 20

1. Danelle Cline
 - a. Working on python notebooks for data localized data clean-up using the tools designed for assisting annotation. Two data sets ready for ingest so far. This set may serve well for summer intern work
 - b. Continued evaluation of amplify/vue/graphql web app
 - c. Student project discussion
 - d. Team meeting scrubbed this week
2. Duane Edgington
 - a. Work continues on registering video files into M3. Waiting for news that files have been moved to The Right Place on Titan...
 - b. Weekly team meeting via teleconference.
 - c. Teleconference scheduled with 2020 intern Peyton Lee.

Week 19

1. Danelle Cline
 - a. Finished data transforms for 2 of 4 datasets to handoff to FathomNet team <https://bitbucket.org/mbari/vaa-localize-clean/>; working on using the dockerized tools developed for assisting annotation to clean the remaining datasets so they can easily be loaded into VARS and also be used for tracking experiments.
 - b. Team meeting and continued prototyping on Vue/GraphQL app.
 - c. Working with IS and KGomes on AWS web login. Joe Gomez made big strides in that last Friday spinning up a webapp in AWS redirected to our usual login. Thanks, Joe!
2. Duane Edgington
 - a. Work continues on registering video files into M3.
 - b. Weekly team meeting via teleconference.
3. Brian Schlining
 - a. Added documentation for registering video/image/annotation fro non-mbari videos at https://mbari-media-management.github.io/docs/mbari/registration_ex1.html
 - b. Working on Software Requirements Doc.

Week 18

1. Duane Edgington
 - a. Tasks on VAA-32 registering videos. Returning to docs and code examples.
2. Brian Schlining
 - a. Working on [data model](#) and [system design](#).

Week 17

1. Danelle Cline
 - a. Working on the beginnings of AWS web portal. Climbing learning curve of Vue, GraphQL and cloud deployment.
 - b. Dockerizing tracking code.
2. Duane Edgington
 - a. Update to project manager on time allocation etc for 2020
 - b. Nvidia webinar on Virtual GPU technology and offerings April 22
 - c. Slack call with Danelle; status update and plan for this week's work
 - d. Project manager has called a group meeting Thursday April 23.

Week 16

1. Brian Schlining
 - a. [New release of VARS Annotation which supports localization on video](#)
 - b. [Packaged Cthulhu for Mac and Linux](#)
 - c. Added [docs for localizing on video](#)

Week 15

1. Duane Edgington
 - a. Continuing work with Danelle on taking video file, image, and annotation datasets that have been generated over the last few years and preparing them to register with M3.
2. Brian Schlining
 - a. [First release of Cthulhu](#), a new video player that supports drawing bounding box annotations directly on videos. Video of it in action at <https://youtu.be/FKGuG8-UYC0>
 - b. Completed integration of Cthulhu with vars-annotation. [Prepping new release of vars-annotation](#).

Week 14

1. Danelle Cline
 - a. Finishing up ML tracker code and continuing work on VARS data loaders. Hoping to finish this up by next week
 - b. Working on high-level system requirements doc
2. Duane Edgington
 - a. Returning to Sprint with Danelle on taking video file, image, and annotation datasets that have been generated over the last few years and preparing them to register with M3.
3. Brian Schlining
 - a. Completed evaluation of the [cthulhu](#), a video player that supports annotation localization. Signed off as completed with external contractor.
 - b. Working on [integration of chullhu with vars-annotation](#).

Week 13

1. Danelle Cline
 - a. About half way through sprint to load localized data into VARS database [VAA-32](#).

- b. Example notebooks for benthic video detection, classification and tracking coming along.
 - c. Working on system requirements. [VAA-23](#).
- 2. Duane Edgington
 - a. Participating in Sprint with Danelle on taking video file, image, and annotation datasets that have been generated over the last few years and preparing them to register with M3.
 - b. Coordinate work over Slack, BitBucket, and M3 stack running on Digits-dev-box-fish (server running at MBARI Moss Landing). Jupyter notebooks, Python.
 - c. Brian, Danelle great to work with over slack (text and audio chats)!

Week 12

- 1. Duane Edgington
 - a. Danelle is leading us in Sprint to do VAA-32 Ingest localized images
 - b. Cloned repository into working directory on digits-dev-box-fish. Working over VPN through steps outlined by Brian S. and Danelle.

Week 11

- 1. Duane Edgington
 - a. Reviewed summer intern applicants. Prioritized list of applicants sent to George et al.
 - b. Successfully built an environment to do SageMaker model training locally while utilizing SageMaker infrastructure transparently. This allows easy decision of local vs cloud data, local vs cloud model training, and local vs cloud deployment without code changes or drama.
- 2. Brian Schlining
 - a. Testing Cthulhu, a new video player for localizing annotations

Week 10

- 1. Danelle Cline
 - a. Working on Sagemaker notebook to demo fast video processing workflow
 - b. Added unit test to [tfrecord](#)
 - c. Reviewing summer intern applications
- 2. Duane Edgington
 - a. Reviewing summer intern applicants.
- 3. Brian Schlining
 - a. Running tests on the video player cthulu, which can provide on-the-fly bounding box localizations on video. Integrates with VARS
 - b. Started integration of localization communication library into VARS in prep for cthulu delivery.

Week 9

- 1. Danelle Cline
 - a. Testing jupyter notebook with example benthic video processing pipelines using TensorFlow + MLFlow. A similar notebook will be developed for midwater. Ostensibly, the notebooks we are developing can also be adapted for Station M and drone still image processing. SageMaker has some great examples for those of you interested in this space. Please come talk to me or Duane Edgington for more details.
 - b. Spun up MLFlow in AWS for development testing
 - c. Added nose unit tests to tfrecord utility
 - d. Adding semantic versioning and bitbucket pipelines for automated releases
 - e. Working on system requirements
 - f. Sent out ideas to the team on user interface design and discussions with Carlos about design ideas
- 2. Duane Edgington
 - a. Continued work on implementing MLFlow workflow on SageMaker with Conda3 kernel jupyter notebook

- b. Exploring SageMaker Studio and SageMaker Debugger (new AWS tools in Preview)
- 3. Brian Schlining
 - a. Created and implemented [a communication protocol for exchanging localization/bounding box information](https://mbari1.atlassian.net/browse/VAA-17) between video player and external apps. <https://mbari1.atlassian.net/browse/VAA-17>
 - b. Documented a workflow and API calls for loading VOC localizations of downloaded VARS framegrabs into M3/VARS. <https://mbari1.atlassian.net/browse/VAA-30>

Week 8

- 1. Duane Edgington
 - a. Working on implementing MLFlow workflow on SageMaker with Conda3 kernel jupyter notebook
 - b. Tasked with loading FathomNet frames into VARS
- 2. Brian Schlining
 - a. Participated in a group meeting to discuss “What are we going to build?”
 - b. Create library/protocol for communicating information about bounding box localizations between two remote applications (e.g. cthulu and vars-annotation). See <https://mbari1.atlassian.net/browse/VAA-17> and [vcr4j 4.3.0.jre11](https://mbari1.atlassian.net/browse/VAA-17).

Week 7

- 1. Duane Edgington
 - a. Continuing work on integration of MLflow into AWS. Cloud9 and shared environments are cool!
 - b. Next step, configure a workflow in SageMaker...
- 2. Brian Schlining
 - a. Working with external contractor on development of a video player that supports annotation localization (e.g. drawing bounding boxes.). Reviewed latest milestone delivery. Wrote additional acceptance tests. Working on UDP protocol for inter-app communication.

Week 6

- 1. Danelle Cline
 - a. Team is working towards benchmarking cost/performance of cloud compute and preparations for testing with new ML-aware video player
 - b. Meeting scheduled for February 12@2pm with video lab to brainstorm about the future use of machine learning for enhancing and augmenting video and image analysis at MBARI. Contact Lonny Lundsten if interested in joining.
 - c. Fixed bug in Tensorflow batch inference
 - d. Successful deployment of MLFlow in dockerized container in AWS in Cloud9
- 2. Duane Edgington
 - a. Demonstrated serverless deployment of ML model for inference in AWS environment via [Chalis \(Lambda\)](#)
 - b. Weekly VAA working session Monday Feb 3
 - i. Continuing work on integration of MLflow into AWS
 - ii. Working on draft schematic diagram of VAA components and dataflow
- 3. Brian Schlining
 - a. Continue to work with external contractor on a video player that supports annotation localization (i.e. draw and display bounding boxes around organisms in real-time)
 - b. Assembled basic automation tests for the video player. See [cthulu-phase1](#)
 - c. Published new release of [VCR4J](#) which patches issues that appeared while creating automation tests.

Week 5

- 1. Danelle Cline

- a. Lead project “kick-off”. Great progress already on both Sagemaker and ML aware video player development!!
 - b. Feb 12 meeting planned to brainstorm with video lab about workflow.
 - c. Working on bug in tensorflow detection model inference served in MLFlow
 - d. Cleaned up wandb model reports organizing by project number instead of name
2. Duane Edgington
- a. Project “kick-off” meeting held Tues Jan 28.
 - i. Duane presented status of SageMaker exploration
 - b. Weekly VAA working session Monday Jan 27
 - i. Demonstrated integration of MLflow with Sagemaker
 - 1. SageMaker Docker image model creation in MLflow
 - 2. Model upload from MLflow to AWS Elastic Container Registry (ECR)
 - 3. Create SageMaker Model Inference Endpoint
3. Brian Schlining
- a. Our external contractor, [Caprica Software](#), has delivered a first iteration of a video player that interacts with [VARS](#) (as an alternative to [Sharktopoda](#)). This version will eventually support interactive feature localization that is required to support machine learning (e.g. draw and display bounding boxes on video.) Brian is testing the software and writing acceptance tests based on the [phase 1 requirements](#).
 - b. Attended VAA kickoff meeting.