

November 20, 2024

Classifier

1. Subject Line

Review and Discussion of Imbalanced Lizard Classification Model Progress

2. Two-Line Summary

This meeting focused on Jacob Dallaire's progress with a computer vision model for classifying lizards, emphasizing the challenges caused by dataset imbalances and species similarities. Suggestions were made to improve documentation, explore alternative models, and provide visual data examples for better context in future presentations.

3. Key Points and Repeated Topics

- Challenges with **overprediction** of the Equestrians (misclassification issues).
- Imbalanced dataset causing **low overall accuracy (35%)** and misclassification.
- Using **MobileNet V2** pretrained on ImageNet for transfer learning due to hardware limitations.
- Suggestions to enhance context in presentations, such as:
 - Including **visual examples of difficult classes**.
 - Showing **variance within and across species**.
- Focus on improving **documentation** for semester wrap-up and for continuity into next semester.
- Issues related to **low-quality images** in the dataset (blurry and indistinct images).
- Importance of exploring alternative models (e.g., BIOCLIP or other pretrained species classifiers).
- Current training set consists of **60,000 images across five species** with significant class imbalances.
- Feedback to improve model accuracy by addressing **variance and imbalance**, as well as incorporating expert feedback from collaborators.

4. Companies / Projects Mentioned

Company/Project Name	Description	Associated Contact/Company
MobileNet V2	Lightweight computer vision model for transfer learning.	Open-source, associated with Keras/Imagenet.
BIOCLIP	Pretrained model for species classification, possibly trained on iNaturalist data.	n/a
iNaturalist	Platform for biodiversity observations, used for labeling lizard images.	n/a

5. People Mentioned

Name	Description	Company/Association
Jacob Dallaire	Presenter, working on the lizard classification project.	n/a
Breanna Shi	Providing feedback on improving documentation and model performance.	n/a
Dr. Stroud	Leads the lab responsible for lizard species identification and labeling.	n/a

6. Facts/Statements Involving Numbers / Metrics

1. Overall model accuracy: **35%**.
2. Top-2 classification accuracy: **58%**.
3. Training dataset size: **60,000 images across five species**.
4. MobileNet V2 chosen due to **lighter weight and compatibility with hardware**.
5. Generalized species classifiers like BIOCLIP achieve **60-70% accuracy** for genus-level classification.

6. Testing size was reduced to allow for faster testing.
-

7. Action Items

Responsible Party: Jacob Dallaire

1. Investigate why **Equestrians are overpredicted**, regardless of ground truth.
2. Provide visual examples of challenging lizard classes for future presentations.
3. Include specific hypotheses about challenges with dataset and model performance.
4. Improve documentation for semester wrap-up, including:
 - Dataset challenges (class imbalance, variance within classes).
 - Model details (e.g., choice of MobileNet V2, performance metrics).
 - Examples of mislabeled or unclear data.
5. Explore alternative models such as BIOCLIP for comparison.
6. Retrieve and provide data on the least represented class sizes (within ~10 minutes).

Responsible Party: Breanna Shi

1. Provide feedback on Jacob's documentation when available.
2. Ensure Jacob's project is prepared for **continuation next semester** by recruiting new members and ensuring proper documentation exists.

Responsible Party: Dr. Stroud's Lab

1. Review low-quality images in the dataset and provide guidance on whether they should be excluded.
 2. Assist Jacob in identifying **key features** used in manual lizard species classification.
-

Final Notes

Does this cover the full transcript? If additional sections remain, please share them so I can incorporate all the relevant data.

Segmentation

1. Subject Line

Team Updates: Project Progress, Member Assignments, and Future Collaboration Planning

2. Two-Line Summary

This portion of the meeting addressed progress on various projects, including member allocation, documentation requirements, and next semester's plans. Discussions included the continuation of key initiatives, potential new collaborations, and skill set needs for group expansions.

3. Key Points and Repeated Topics

- **Team Updates and Assignments:**
 - Breanna Shi is evaluating whether certain projects (e.g., Philip Woolley's) need additional members for next semester.
 - Volunteering for projects can lead to varying levels of commitment and output, so member additions may be required to balance workloads.
- **Documentation and Semester Wrap-Up:**
 - All groups are expected to prepare **detailed documentation** for their projects, including source code cleanup and design documents.
 - Emphasis on preparing materials for **new members** or **continued work** next semester.
- **Exploration of New Collaborations:**
 - Shi noted potential collaboration with Dr. Stroud and another faculty member for new research next semester.
- **Specific Project Updates:**
 - **Philip Woolley's Project:**
 - Focused on **model performance improvements** and cleaning up the source code.

- Encouraged to explore the **slicer module element** and draft initial concepts for usability.
 - Documentation will guide decisions about skill sets needed for new members (e.g., UI/UX expertise).
 - Discussion on the success and separation of the **Ayesh and Mercedes project**, as it is a continuation from previous semesters.
- **Future Work Expectations:**
 - New researchers are allowed flexibility for exploration in their first semester to avoid excessive pressure.
 - Successful preliminary work will likely lead to expanded responsibilities and collaborations in the future.

4. Companies / Projects Mentioned

Company/Project Name	Description	Associated Company/Contact
Dr. Stroud Collaboration	Research initiative with Dr. Stroud's lab, potentially expanding next semester.	Dr. Stroud's Lab
Ayesh and Mercedes Project	Ongoing project with notable success from previous semesters.	Ayesh and Mercedes

5. People Mentioned

Name	Description	Company/Association
Breanna Shi	Facilitator of the meeting, providing guidance on member assignments, project updates, and documentation.	n/a
Philip Woolley	Lead of a project focusing on model performance improvements and exploring a slicer module.	n/a
Ayesh	Contributor to the Ayesh and Mercedes project, noted for high commitment despite volunteering.	n/a
Mercedes	Co-contributor to the Ayesh and Mercedes project.	n/a

Name	Description	Company/Association
Dr. Stroud	Collaborator on research projects and lab leader contributing to species classification research.	Dr. Stroud's Lab

6. Facts/Statements Involving Numbers / Metrics

1. **60,000 images** in Jacob Dallaire's dataset were referenced in previous context.
 2. Philip Woolley's slicer module is in early exploration; full implementation is not required by semester's end.
-

7. Action Items

Responsible Party: Breanna Shi

1. Decide whether Philip Woolley's project requires additional members for next semester.
2. Determine appropriate skill sets (e.g., UI/UX expertise) for potential team expansion in Woolley's project.
3. Coordinate with Dr. Stroud and another faculty member regarding next semester's collaboration.

Responsible Party: Philip Woolley

1. Explore initial designs and feasibility for the **slicer module element**.
2. Focus on **model performance improvements** and clean up the source code.
3. Prepare a comprehensive **design document**, detailing decisions and technical aspects of the project.

Responsible Party: Ayesh and Mercedes

1. Continue their ongoing project, with recognition of their success despite volunteer status.
-

Final Notes

Does this cover the transcript sufficiently? If additional sections exist, please provide them for analysis and integration into the final summary.

X-Ray

1. Subject Line

Refining User Workflows, Landmarking Challenges, and Literature Review Guidance

2. Two-Line Summary

This segment of the meeting addressed refining project workflows, focusing on Ayush's image-processing and Mercedes' landmarking tasks. Discussions also emphasized literature reviews to understand the nuances of landmarking challenges and creating general descriptions for the project's website.

3. Key Points and Repeated Topics

Ayush's Project Workflow:

- Discussion about handling **ground truth** for current test images but ensuring new workflows exclude it.
- Proposed workflow:
 - User uploads an image.
 - The system processes the image and presents a **processed preview** to the user.
 - The system generates **landmarks**, which users can confirm or adjust if slightly off.
 - Export the data in a format suitable for **Dr. Stroud's lab**.
- Need to decide on synchronous vs. asynchronous processes for efficiency.

Mercedes' Landmarking Work:

- Errors in landmark placement (e.g., toes) are due to **distracting elements** in images that are not labeled, confusing the model.
- Importance of thorough labeling to reduce confusion in predictions.
- Suggested collaboration with **John** for toe labeling:
 - Mercedes could label and John could **review**, reducing workload.
- Mercedes also discussed overlay visualizations to preview landmark placements.

Literature Review and Research:

- Shi emphasized the need to conduct a **literature review** on landmarking challenges:
 - Understand differences between **pose estimation** tasks (e.g., videos) and **landmarking in static images** (e.g., X-rays).
 - High accuracy is more critical for static X-ray landmarking than for pose estimation.
 - Explore papers that highlight landmarking challenges and machine learning differences.
 - Find gaps in the literature and use the group's project as an answer to unresolved issues.
 - Slides summarizing key papers could be presented during meetings for feedback.

Website Content Guidance:

- Mercedes inquired about creating **project descriptions** for the website:
 - Suggested format: **one general paragraph** summarizing the group's work (not overly specific).
 - Should highlight **types of datasets** (e.g., lizard datasets) and methods (e.g., computer vision).
 - Include some specific examples but keep the description general enough to remain relevant as new projects are added.

4. Companies / Projects Mentioned

Company/Project Name	Description	Associated Contact/Company
Dr. Stroud's Lab	Lab collaborating on X-ray data and evolutionary research.	Dr. Stroud's Lab

5. People Mentioned

Name	Description	Company/Association
Breanna Shi	Meeting facilitator providing feedback on workflows, literature review, and website content.	n/a
Ayush Parikh	Working on image-processing workflows and automation of user adjustments.	n/a
Mercedes Quintana	Working on X-ray landmarking, labeling toes, and refining project descriptions.	n/a
Dr. Stroud	Oversees lab conducting X-ray and evolutionary lizard research.	Dr. Stroud's Lab
John	Suggested reviewer for toe-labeling tasks to assist Mercedes.	n/a

6. Facts/Statements Involving Numbers / Metrics

1. No specific numeric metrics were discussed in this section.

7. Action Items

Responsible Party: Ayush Parikh

1. Refine the **image-processing workflow**:
 - Include steps for uploading images, processing previews, generating landmarks, and user adjustments.
 - Decide on synchronous vs. asynchronous flow based on server response times.
 - Confirm data export format with Dr. Stroud's lab.

Responsible Party: Mercedes Quintana

1. Label toe landmarks and seek **John's review** to ensure accuracy.
2. Investigate errors in toe predictions caused by **distracting elements** in images.
3. Draft a **general project description** for the website:
 - Summarize the team's work in a paragraph with examples.

- Avoid overly specific details to keep the description adaptable for new projects.

Responsible Party: Breanna Shi

1. Guide team members to conduct a **literature review** on landmarking challenges:
 - Investigate differences between pose estimation and static X-ray landmarking.
 - Identify relevant papers and bring them to meetings for discussion.
2. Advise on crafting general, adaptable **website content** for project descriptions.

Responsible Party: John (if assigned)

1. Review toe labeling done by Mercedes to ensure quality and accuracy.

Final Notes

Does this capture the transcript adequately? If there are additional sections, feel free to share them for further integration.