BiteFinder: APCOMP 215 Project Milestone 1

Team Members

Jack Hwang (jack_hwang@hms.harvard.edu)
Benet Fité Abril (benet_fite@hms.harvard.edu)
Zoe Zabetian (zoe_zabetian@hms.harvard.edu)
Irith Katiyar (irith_katiyar@hms.harvard.edu)

Problem Statement

Develop an (offline) bug bite classification app, where a user can take a photo of their skin on their phone and type their symptoms, and the app will predict what bug caused the bite and provide recommendations on which products to use or whether to see a doctor.

Background and Motivation

Global warming is expanding bug habitats into cooler regions (e.g., central Europe, northern Spain, New England), causing more unfamiliar bites (e.g., this summer's tick plagues, invasive wasps, new mosquitos). Seeing a dermatologist is inconvenient - appointments are scarce and delayed - so our target users are hikers and campers in remote areas without the Internet. We propose both an offline app and a web version.

Prior studies (Akshaykrishnan et al., Ilijoski et al., Khasanov et al.) have trained bug bite classifiers, and apps exist but receive poor reviews. None use multimodality (i.e. combining images with text). The only related work applies multimodal learning to skin disease (Yang). Our project fills this gap with a multimodal classifier that integrates photos and symptom descriptions, with an emphasis on accuracy and offline access beyond what ChatGPT provides.

Scope and Objectives

1. Preprocess a public image dataset of a defined set of bug bites.
2. Synthesize user narratives that describe common symptoms of each bug bite to create paired image-text dataset.
3. Develop and fine-tune a multimodal vision-language model to classify bug bites.
4. Create container infrastructure for development.
5. Develop frontend that takes image/text and shows results in user-friendly format.
6. Develop both an online web application and an offline mobile application using the cloud computing components of the course.
7. The lightweight LLM will both relay classification results conversationally and act as an agent to manage user interactions for symptom input and imaging.

Minimum Components for a Good Project

1. Large Data: Visually diverse set of over 1K skin images spanning many bug bites.
2. Scalability: Backend must support offline availability and multiple users.
3. Complex Models: Multimodal model consisting of image encoder, text encoder, and prediction head (e.g., ensemble classifier, vision-and-language transformer).
4. Computationally Expensive Inference: Make inference as lightweight as possible.

Data

Image data will come from the Bug Bite Images dataset on Kaggle (~1300 RGB images across 8 classes: ants, bed bugs, chiggers, fleas, mosquitos, spiders, ticks, and no bites). Images vary in lighting, skin tone, and anatomical location; we will apply augmentations during preprocessing. This dataset has been used in prior studies (Ilijoski et al., Khasanov et al.), supporting its credibility. Text data will be created by scraping symptom descriptions from reputable medical sources (e.g., Cleveland Clinic) and generating patient narratives via ChatGPT, yielding a paired image–text dataset. The main data challenges we will face are variations in skin tones and lighting, which we will address with targeted image augmentations. The data classes are fairly balanced, though we may randomly sample to ensure equally balanced classes for training.

Learning Emphasis

Our project is designed to emphasize MLOps concepts central to the course. By containerizing each component (data preprocessing, model training, and inference), we will ensure reproducibility and scalability across environments. Managing multimodal data pipelines will show our ability to handle heterogeneous data sources. We will also tackle the challenges of deploying resource-intensive models on lightweight devices.

Fun Factor

We are excited to work with multimodal learning and machine learning workflows, and some of us are avid hikers and could definitely see ourselves using this app!

Limitations and Risks

Making the model work offline will be technically challenging, while the web app format risks competing with ChatGPT (users may view ChatGPT as "good enough" for this task).

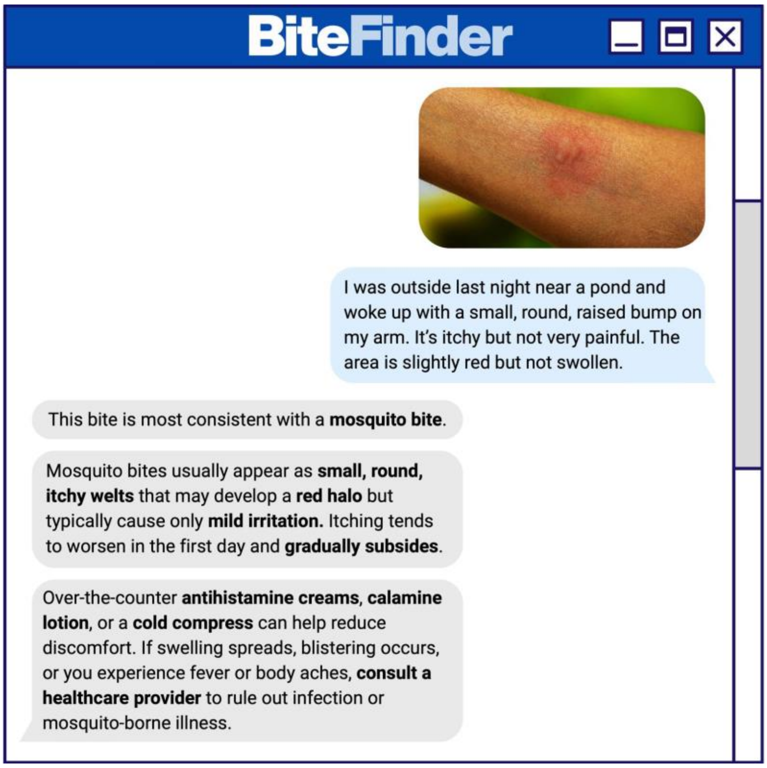
Milestones

09/25 – 10/05: finalize dataset (download, clean, augment images; scrape and synthesize symptom narratives).

: build and test preprocessing pipeline; establish container infrastructure; baseline unimodal image model. **10/21 – 11/05**: develop multimodal model (vision encoder + text encoder); experiment with prediction heads; initial evaluation. **11/06 – 11/15**: optimize model for inference efficiency (quantization, distillation, or pruning); begin lightweight LLM integration. **11/16 – 11/25**: build frontend prototype (web + mobile offline mode); integrate with backend containers. **11/26 – 12/10**: end-to-end testing; user experience refinements; prepare final demo and report.

Application Mock Design

The application will consist of a chatbot interface where the user can upload a photo of their bug bite and a text description and get responses back about what their bug bite is, reasoning as to why the bot made that assessment, and any recommendations. See the mockup below:



References

Akshaykrishnan, V., Sharanya, C., Abhinav, K., Aparna, C. K., & Bindu, P. V. (2023). A Machine Learning based Insect Bite Classification. 2023 3rd International Conference on Smart Data Intelligence (ICSMDI), 259–264. https://doi.org/10.1109/ICSMDI57622.2023.00111

Bug bites identification, Reddit. (n.d.). Retrieved September 21, 2025, from https://www.reddit.com/r/bugbites/

Bug_bite_images, Kaggle. (n.d.). Retrieved September 21, 2025, from https://www.kaggle.com/datasets/moonfallidk/bug-bite-images

Hareza, D., Langley, R., Haskell, M. G., & King, K. (2020). National Estimates of Noncanine Bite and Sting Injuries Treated in US Hospital Emergency Departments, 2011-2015. Southern medical journal, 113(5), 232–239. https://doi.org/10.14423/SMJ.0000000000001091

Ilijoski, B., Trojachanec Dineva, K., Tojtovska Ribarski, B., Petrov, P., Mladenovska, T., Trajanoska, M., Gjorshoska, I., & Lameski, P. (2023). Deep Learning Methods for Bug Bite Classification: An End-to-End System. Applied Sciences, 13(8), 5187. https://doi.org/10.3390/app13085187

Khasanov, D., Khujamatov, H., Shakhnoza, M., Abdullaev, M., Toshtemirov, T., Anarova, S., Lee, C., & Jeon, H.-S. (2025). DeepBiteNet: A Lightweight Ensemble Framework for Multiclass Bug Bite Classification Using Image-Based Deep Learning. Diagnostics, 15(15), 1841. https://doi.org/10.3390/diagnostics15151841

Yang, A. & Yang, E. (2024). A Multimodal Approach to The Detection and Classification of Skin Diseases. (n.d.). Retrieved September 21, 2025, from https://arxiv.org/html/2411.13855v1