

An annotated image dataset for training mosquito species recognition system on human skin

ABSTRACT

This paper introduces a new mosquito images dataset that is suitable for training and evaluating a recognition system on mosquitoes in normal or smashed conditions. The images dataset served mainly for the development a machine learning model that can recognize the mosquito in the public community, which commonly found in the smashed/damaged form by human. Especially the images of mosquito in hashed condition, which to the best of our knowledge, a dataset that fulfilled such condition is not available. There are three mosquito species in the dataset, which are *Aedes aegypti*, *Aedes albopictus* and *Culex quinquefasciatus*, and the images were annotated until species level due to the specimen was purely bred in a WHO accredited breeding laboratory. The dataset consists of seven root files, six root files that composed of six classes (each species with either normal landing, or random damaged conditions) with a total of 1500 images, and one pre-processed file which consists of a train, test and prediction set, respectively for model construction.