

Frank Cally Tabuco

cally.almaden@gmail.com | linkedin.com/in/phrungck/ | github.com/Phrungck

- Graduate research accepted at 33rd British Machine Vision Conference; awarded as "Most Outstanding Research in Cardiology" at Philippine Heart Association's 53rd Annual Convention.
- Expert in creating and processing large scale datasets for computer vision (CV) and natural language processing tasks (NLP).
- Experienced Python developer with knowledge in ReactJS, Streamlit, FastAPI.

Work History

(NLP) Founding Machine Learning Engineer (Nov. 2022 --)

(Remote) Emsisoft Ltd., New Zealand

Founding researcher, developer, and engineer of natural language processing models for cybersecurity analysis. Handled the end-to-end deployment of models from creation of NLP datasets to model implementation and development of embedding and transformer models, and, finally, model serving and monitoring.

- Technology stack: Python, PyTorch, C#, Hyper-V, PowerShell, ONNX, and Excel/CSV (~100 million rows), Azure.

(CV) Artificial Intelligence Researcher (Dec. 2021 -- Jan. 2023)

(Remote) University of the Philippines Diliman, Philippines

Lead researcher, developer, and engineer of computer vision models for predicting and analyzing future, localized weather events from Doppler radar images. Handled the end-to-end deployment of models from creation of computer vision datasets of Doppler radar data to model implementation and development of GANs and segmentation models, and, finally, model serving.

- Technology stack: Python, PyTorch, TensorFlow, Linux, Excel/CSV, OpenCV, PyArt.

Notable Projects

(CV) Two-View Left Ventricular Segmentation and Ejection Fraction Estimation in 2D Echocardiograms (Nov. 2022)

33rd British Machine Vision Conference, United Kingdom

Developed a novel deep learning framework for end-to-end estimation of ejection fraction using image segmentation and action recognition models.

- Awarded the "Most Outstanding Research in Cardiology" at Philippine Heart Association's 53rd Annual Convention.
- Technology Stack: Python, PyTorch, Excel/CSV, OpenCV.

Education

Master of Science in Computer Science (2020-2022), GPA: 4.00

University of the Philippines Diliman - Quezon City, Philippines

Skills

Advanced

- Computer Vision
 - Feature detection
 - Action recognition
 - Vision transformers
 - Image segmentation
 - Object detection
 - Generative models
- Natural Language Processing
 - Data preprocessing
 - word2vec
 - RoBERTa, GPT-2
- Machine Learning
 - Feature engineering
 - SVM, XGBoost
- Python, PyTorch
- OpenCV
- Microsoft Excel

Intermediate

- Reinforcement learning
 - Q-learning, SARSA
 - Cross-entropy method
 - Deep Q-network
 - A2C, PPO
- TensorFlow
- PowerShell, Hyper-V
- MySQL
- FastAPI
- ReactJS
- Streamlit
- Azure

Tools

Analysis Tools

- Numpy
- Pandas, Vaex, Polars
- PyArt, Librosa, MediaPipe
- Matplotlib, Jupyter Notebook

AI/ML Tools

- Albumentations, Torchvision
- Scipy, Scikit-learn, Skimage
- NLTK, Gensim, Hugging-face
- SMOTE, Einops
- ONNX

Education

Bachelor of Science in Management Engineering (2012-2016), GPA: 3.33

Ateneo de Manila University - Quezon City, Philippines

Graduate Research

URL-based Phishing Detection Using Ensembles and Feature Selection

- Used classical machine learning techniques for detection of phishing websites based o URL elements
- Models: XGBoost, AdaBoost, Random Forest

Detection of Myocardial Infarction Using Computer Vision and Deep Learning Techniques

- Applied advanced computer vision algorithms and a vanilla deep learning model to detect infarcted segments from 2D echocardiography videos
- Models: Harris Corner Detector, Histogram of Oriented Gradients, Unet++, ResNest50, LSTM

Predicting Rainforest Species Presence Using Acoustic Signal Classification

- Applied spectrogram transformations to classify multiple instances of bird and frog audios in a noisy environment
- Models: Fourier Transforms, Mel-spectrograms, ResNest50

Traffic Light Control with Multiple Vehicle Types Using Policy Gradient Methods and LSTM-Wrappers

- Utilized a traffic light control simulator to test the capacities of SOTA reinforcement learning algorithms on a customized multi-intersection network for traffic light control
- Models: Proximal Policy Optimization, A2C, LSTM

Modeling Bio-acoustic Signals Using Various Time Frequency Representations

- Compared space and time complexities of different frequency transformations on bio-acoustic dataset
- Models: Mel-spectrograms, Constant Q Transform, Fast Fourier Transform

Modeling and Verifying Lustre File System Data Operations Using CSP and PAT

- Tested properties of file system operations including non-termination, divergence-free, and deadlocks
- Models: CSP#, Process Analysis Toolkit