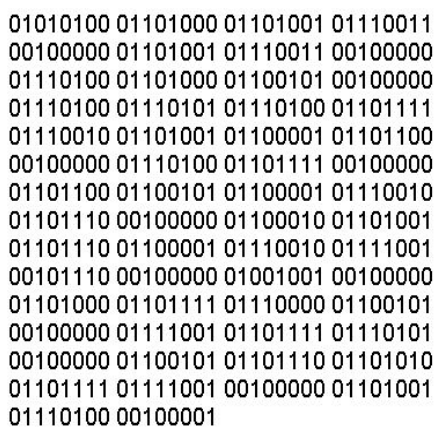
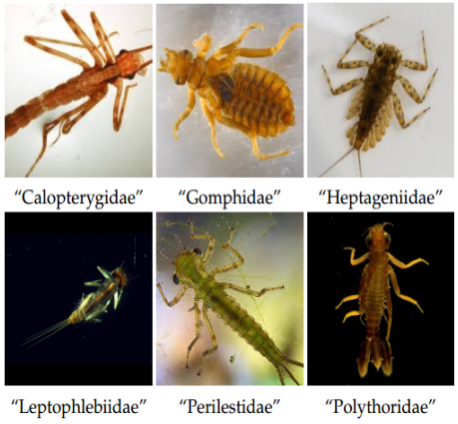




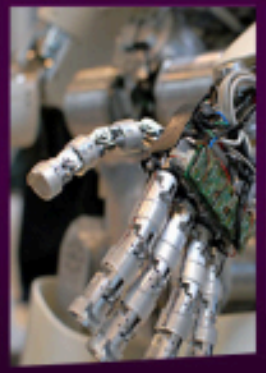
"Using Deep Convolutional Networks for the Automatic Recognition of MacroInvertebrate in Rivers and Affluents in Panama"

Javier Sánchez Galán, PhD
Universidad Tecnológica de Panamá
Grupo de Investigación en Biotecnología, Bioinformática y Biología de Sistemas – GIBBS
December 8th, 2018

Universidad Tecnológica de Panamá





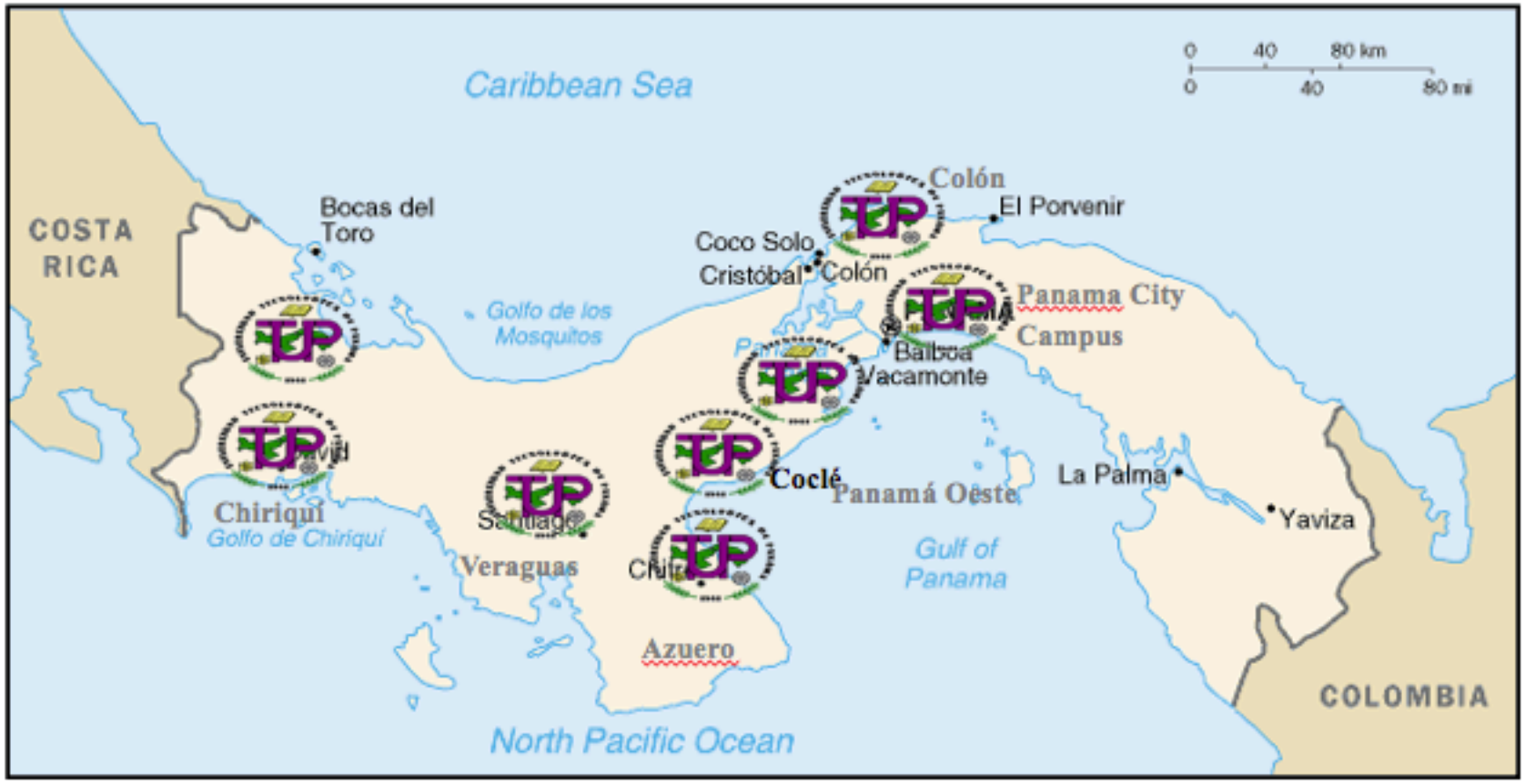


UNIVERSIDAD TECNOLÓGICA DE PANAMÁ

Panama in Brief

- **Oficial name:** Republic of Panama. **Demonym:** Panamanian
- **Location**: Central America, bordering both the Caribbean Sea and the North Pacific Ocean, between Colombia and Costa Rica.
- **Geographic coordinates**: 9N, 80W (Japan 35N, 139E).
- **Area**: total: 75,420 km^2 (Japan 377,972 km^2).
- **Administrative divisions**: 9 provinces and 5 autonomous indigenous regions.
- **Population**: 4,034,119 (est. 2016)
- **Climate**: tropical maritime; hot, humid, cloudy; prolonged rainy season (May to January), short dry season (January to May).
- **Economic activities**: tourism, port activities, canal fees, agricultural exports and mining.

Location of Main and Secondary Campuses



Academic Programs

- There are **131** careers at different levels, as follows: 2 Doctorate studies, 40 master degrees, 26 postgraduated courses, 1 Professor Career, 4 Specializations, 8 Diplomas, 14 Bachelor degrees in Engineering, 14 Bachelor Degrees, 8 Bachelor Degrees in Technology and 14 Technical careers.

Civil

Electrical

Mechanical

Industrial

Computer Systems

Science and Technology

Research Centers

Engineering Experimental Center

Information Technology and Communication Research and Innovation Center

Hydrotechnical and Hydraulics Research Center

Electrical, Mechanical and Industrial Research and Innovation Center.

Agro-industrial Production and Research Center

innovation and Technology Transfer Center

Project IDDS-2015-054: Participatory Biomonitoring of Water Quality with Rural Aqueduct Management Boards (JAAR)

Motivation

Project IDDS-2015-054 (SENACYT): **Participatory Biomonitoring of Water Quality with Rural Aqueduct Management Boards (JAAR): A Tool for the Sustainability of Water Resources in Panama**.
http://www.gorgas.gob.pa/aplicaciones/biomonitoreo/

Objectives

- In our country there are 5397 aquifer systems, in other words about 677,207 inhabitants (20% of the population)
- Verify the pollution of rivers and evaluate the quality of surface affluents.
- Use of macroinvertebrates as bioindicators.

Participatory and Educational Activities for the JAAR of the Chiriqui Region (Western Panama)



- Incorporate communities and especially collaborate with the AJAARCHI Committee.
- They will be given an app to collect the information.

Workflow of the Mobile System for Water Quality Registry & Macroinvertebrate Detection

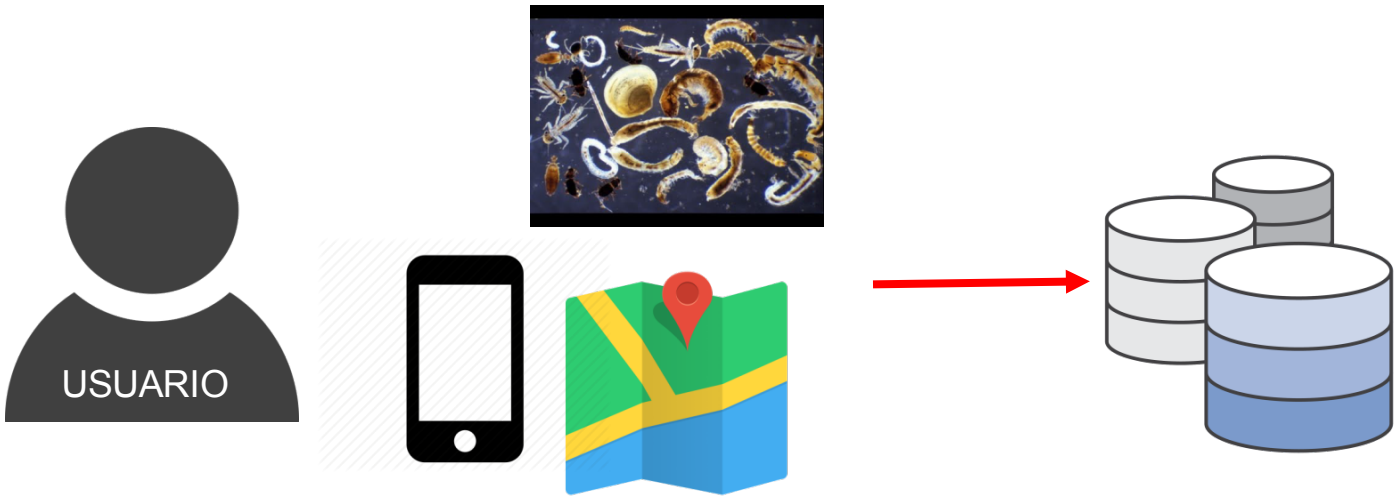
DB (TEMPORAL)

USER

EXPERT

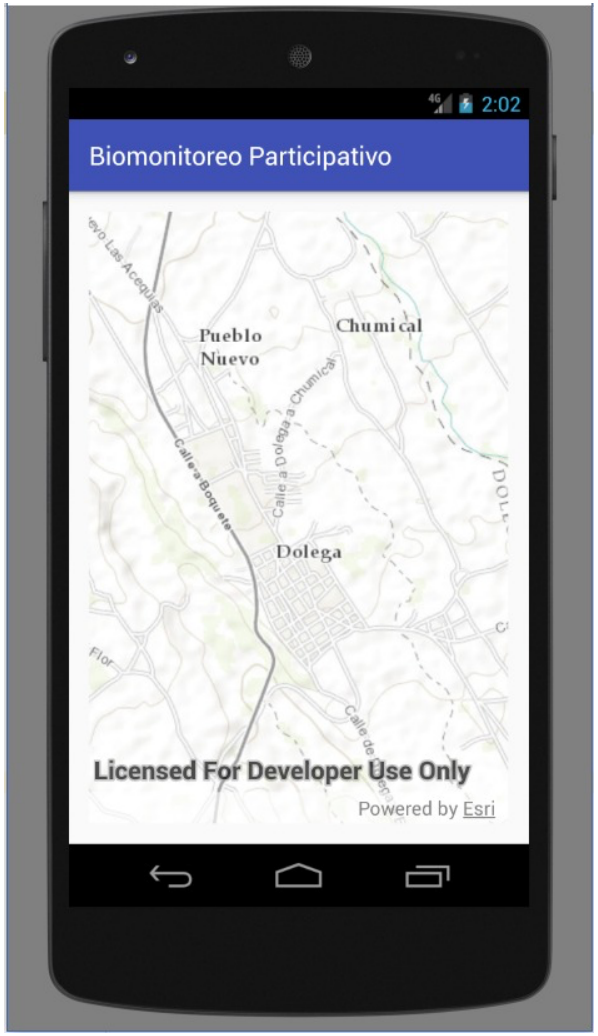
DB (FINAL)

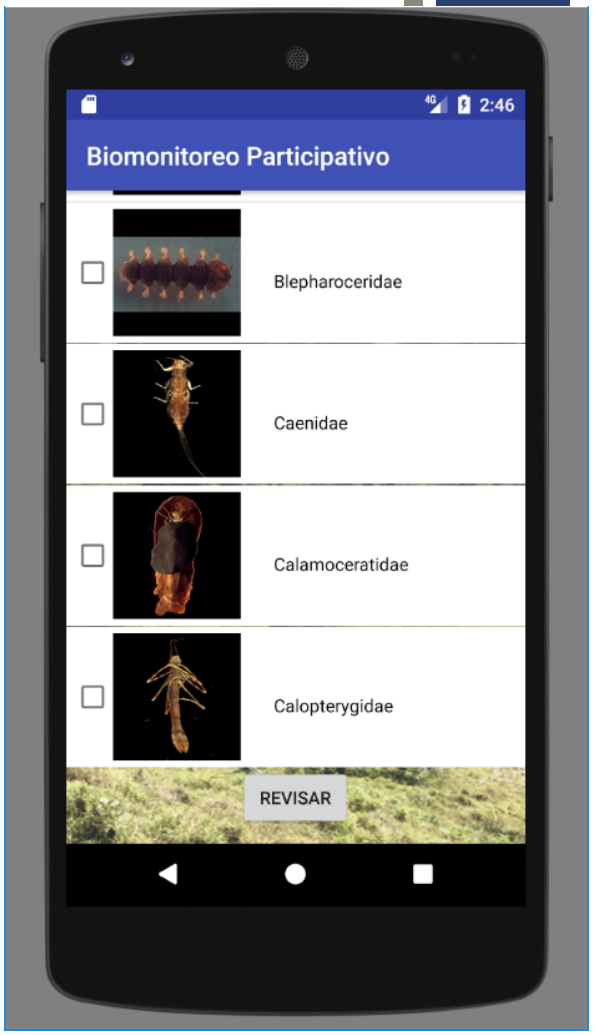
Mobile System for Water Quality Registry



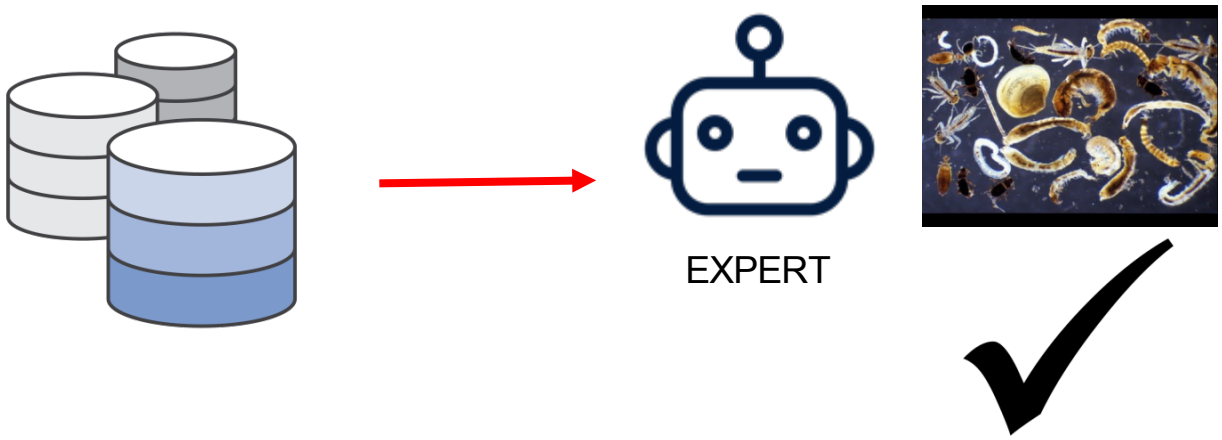
This project for the development of graphic interfaces and user experience is developing a thesis carried out by the student **Jesús Fuentes** from the School of Computer Systems Engineering and directed by **Dr. Elba Valderrama.**

Mobile System for Water Quality Registry





Workflow of the Automatic Macroinvertebrate Images Detection System

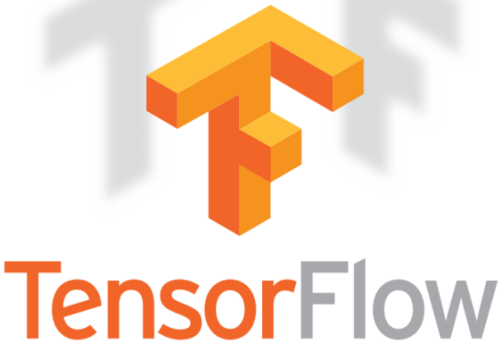


This automation project is part of the development of the thesis "*Use of Convolutional Neural Networks for the Automatic Recognition of Macroinvertebrate Images for Participatory Biomonitoring*" carried out by the student **Carlos Quintero** of the School of Computer Systems Engineering and Directed by **Dr. Javier Sánchez Galán.**

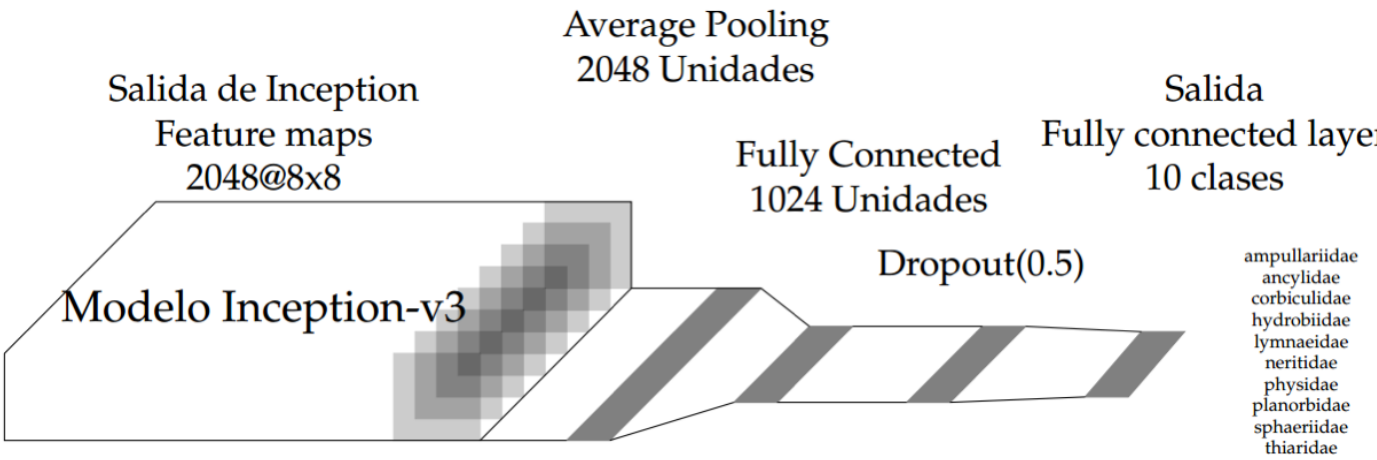
Macroinvertebrate Images Detection System

Objetives

- Macroinvertebrate recognition system using images and a deep neural network model
- Test different variants and compare their results
- Execution on Android mobile without the need of a server or internet.



Algorithm

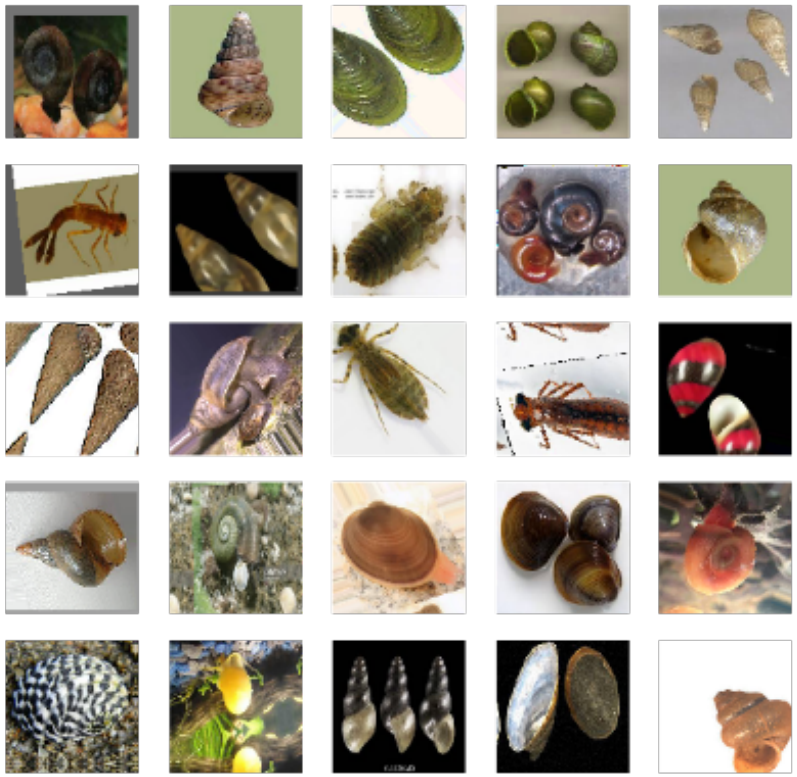


The Inception-v3 algorithm was used as the reference model, the Transfer Learning module (TL) was added, which implies that only the last layers of the network were trained and the classification of the family resulted in the output layer of macroinvertebrate.

Training Database

A database of images (subfolders) was created for each of the 14 families of macroinvertebrates in the study, examples of characteristic images of each of the families

Data Augmentation



Because there were few test images for some of the families studied, the Data augmentation (DA) technique was used to obtain new images from the initial images. To achieve this, different methods of rotation, cutting, translation and change in the intensity of colors are applied to each image, thus obtaining new images for the training of the model in the network.

Experimento #1

"Calopterygidae"

"Heptageniidae"

	Entrenamiento			Validación		
Modelo	top-1	κ de Cohen	error	top-1	κ de Cohen	error
TL	0.9167	0.8295	0.3265	0.8333	0.6606	0.4794
TL-DA	0.8906	0.7582	0.2798	0.8000	0.5982	0.3941
FT-DA	0.9922	0.9844	0.0150	0.9000	0.7982	0.3882

	Prueba		
Modelo	top-1	κ de Cohen	error
TL	0.7000	0.4000	0.5012
TL-DA	0.9000	0.8000	0.2914
FT-DA	0.9000	0.8000	0.1951

Experiment #2

"Calopterygidae" "Gomphidae" "Heptageniidae" "Heteragrionidae"

"Leptophlebiidae" "Perilestidae" "Polythoridae"

	Entrenamiento				Validación			
Modelo	top-1	top-3	κ de Cohen	error	top-1	top-3	κ de Cohen	error
TL	1.0000	1.0000	1.0000	0.0432	0.7191	0.9775	0.5979	0.7657
TL-DA	0.8750	1.0000	0.8246	0.3867	0.7416	0.9438	0.6350	0.7241
FT-DA	0.8828	1.0000	0.8326	0.3133	0.8090	0.9888	0.7354	0.5880

	Pruebas			
Modelo	top-1	top-3	κ de Cohen	error
TL	0.7742	1.0	0.6303	0.5671
TL-DA	0.8387	1.0	0.7359	0.4539
FT-DA	0.8710	1.0	0.7963	0.3352

Experiment #3

"Ampullaridae" "Ancylidae" "Corbiculidae" "Hydrobiidae" "Lymnaeidae"

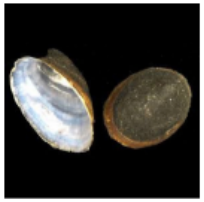
"Neritidae" "Physidae" "Planorbidae" "Sphaeriidae" "Thiaridae"

	Entrenamiento					Validación				
Modelo	top-1	top-3	top-5	κ de Cohen	error	top-1	top-3	top-5	κ de Cohen	error
TL	0.9860	1.0000	1.0000	0.9843	0.2041	0.6682	0.8925	0.9486	0.6278	0.9586
TL-DA	0.7188	0.9219	0.9766	0.6848	0.8281	0.6776	0.9019	0.9766	0.6306	0.9697
FT-DA	0.9062	0.9922	1.0000	0.8932	0.3480	0.7523	0.9579	0.9860	0.7160	0.7772

	Prueba				
Modelo	top-1	top-3	top-5	κ de Cohen	error
TL	0.6667	0.9306	0.9861	0.6189	1.0537
TL-DA	0.6528	0.9028	0.9722	0.6079	0.9759
FT-DA	0.7500	0.9305	0.9722	0.7094	0.8196

Experiment #4


"Ampullaridae"


"Ancylidae"


"Corbiculidae"


"Hydrobiidae"

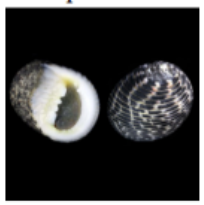

"Lymnaeidae"


"Calopterygidae"


"Gomphidae"


"Heptageniidae"


"Heteragrionidae"

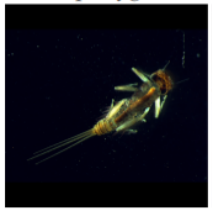

"Neritidae"


"Physidae"


"Planorbidae"


"Sphaeriidae"


"Thiaridae"


"Leptophlebiidae"

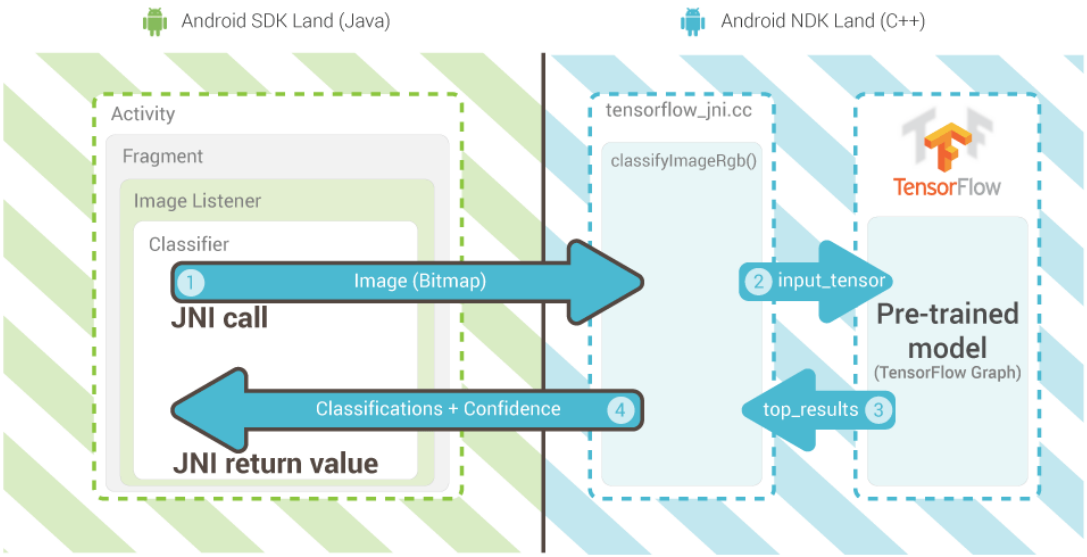

"Perilestidae"

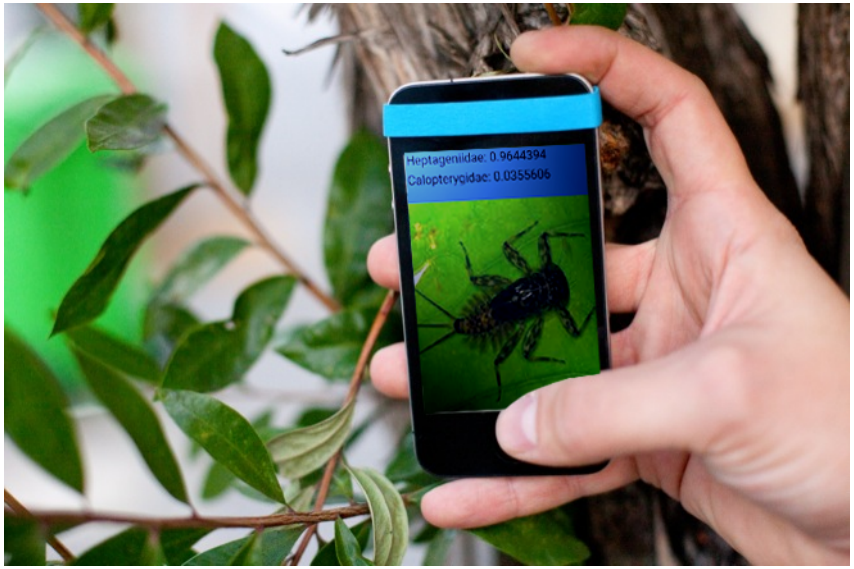

"Polythoridae"

	Entrenamiento					Validación				
Modelo	top-1	top-3	top-5	κ de Cohen	error	top-1	top-3	top-5	κ de Cohen	error
TL	0.9901	1.0000	1.0000	0.9892	0.0401	0.7007	0.8849	0.9309	0.6734	1.1625
TL-DA	0.6641	0.9062	0.9766	0.6299	0.9673	0.6809	0.9046	0.9605	0.6486	0.9643
FT-DA	0.8594	0.9844	1.0000	0.8436	0.5727	0.7270	0.9013	0.9737	0.6993	0.8500

	Prueba				
Modelo	top-1	top-3	top-5	κ de Cohen	error
TL	0.6863	0.8529	0.9706	0.6542	1.2423
TL-DA	0.7059	0.8922	0.9608	0.6760	0.9835
FT-DA	0.8235	0.9804	0.9902	0.8044	0.6856

Mobile Application in TensorFlow





Conclusions

The biggest discoveries obtained during the practice were the following:

1. It is possible to develop a macro invertebrate image recognition tool using a small volume database. This is possible when applying regularization techniques.
2. The best results in the recognition tool were obtained by using fine adjustment in the final layers of the convolutional model and a data augmentation module.
3. To obtain a good model, it is necessary to balance the volume of images between the different classes as much as possible or to weight the training in the classes with the lowest volume, this can be seen in the high value obtained in Cohen's κ coefficient.
4. Regularization techniques increase the capabilities of the model, but they are not sufficient to achieve maximum accuracy with a poor database.

Future Work

Future Work

USER

EXPERT

EXPERT

DB (FINAL)

Acknowledgements

Carlos Quintero

Universidad Tecnológica de Panamá

Javier Enrique Sanchez Galan F

2.19 · PhD in Experimental Medicine · Edit

Universidad Tecnológica de Panamá

Fernando Merchan

6 · PhD

Universidad Tecnológica de Panamá

Aydeé Cornejo

17.83

The Gorgas Memorial Institute for Health Studies (GMI)

Presentations

CONGRESO CIENTÍFICO
UNACHI '14 - 18/08/2017

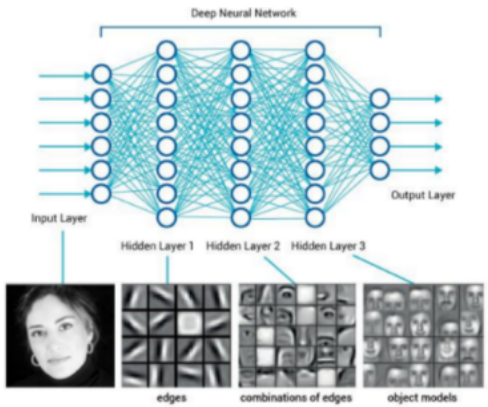
UNIVERSIDAD TECNOLOGICA DE PANAMA
1981

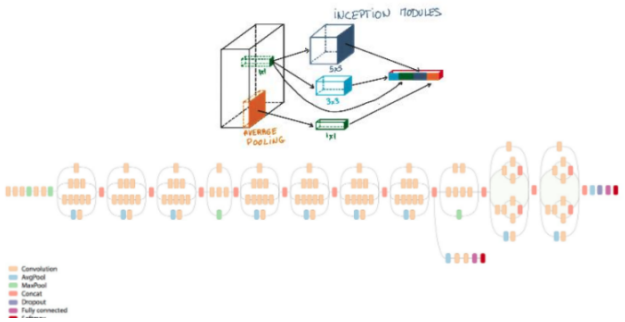
FACULTAD DE INGENIERÍA DE SISTEMAS COMPUTACIONALES - 1982 -
UNIVERSIDAD TECNOLÓGICA DE PANAMÁ

KnE Engineering

ESTEC Conference Proceedings
6th Engineering, Science and Technology Conference (2017)
Volume 2018

Knowledge E
Engaging minds



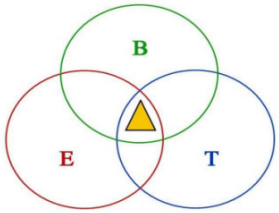


Conference Paper

Uso de Redes Neuronales Convolucionales para el Reconocimiento Automático de Imágenes de Macroinvertebrados para el Biomonitoreo Participativo

Carlos Quintero[1], Fernando Merchán[1], Aydeé Cornejo[2], and Javier Sánchez-Galán[1,3]

[1]Universidad Tecnológica de Panamá, Panamá
[2]Grupo de Investigación en Macroinvertebrados Dulceacuícolas de Panamá. Instituto Conmemorativo Gorgas de Estudios de la Salud (ICGES)
[3]Instituto de Investigaciones Científicas y Servicios de Alta Tecnología AIP (INDICASAT AIP)



"VITA SICUT SCIENTIAM ET PROPOSITUM"

Grupo de Investigación en Biotecnología, Bioinformática y Biología de Sistemas – GIBBS

Javier Sánchez Galán, PhD

javier.sanchezgalan@utp.ac.pa

@UTPPanama @CepiaUTP