M. Sc. Saúl Calderón
Instituto Tecnologico de Costa Rica /
De Montfort University

Artificial Intelligence for medical imaging applications: experiences in Costa Rica

Supervisors: Shengxiang Yang
Armaghan Moemeni
Simon Donelly-Conelly

Presentation summary

- What is Machine Learning (ML) / Deep Learning (DL)?
- Using DL for medical imaging in Costa Rica
- Human/organizational challenges of using DL for medical applications
- Technical challenges of DL for real world applications
- Questions and discussion

DE MONTFORT UNIVERSITY LEICESTER

PARMA Group

About me

- B. Sc. Computer Science, M. Sc. Electrical Engineering, PhD student at De Montfort University, UK
- Professor/Researcher at the Costa Rica Institute of Technology, and coordinator of the PAttern Recognition and MAchine Learning Group (PARMA Group,)
- A central-american country, with a centralized, universal and public health care system: Highest life expectancy in latinamerica with 80.98 years in average, higher than USA (BBC, CEPAL)

Some pics. Of Costa Rica

TEC main campus

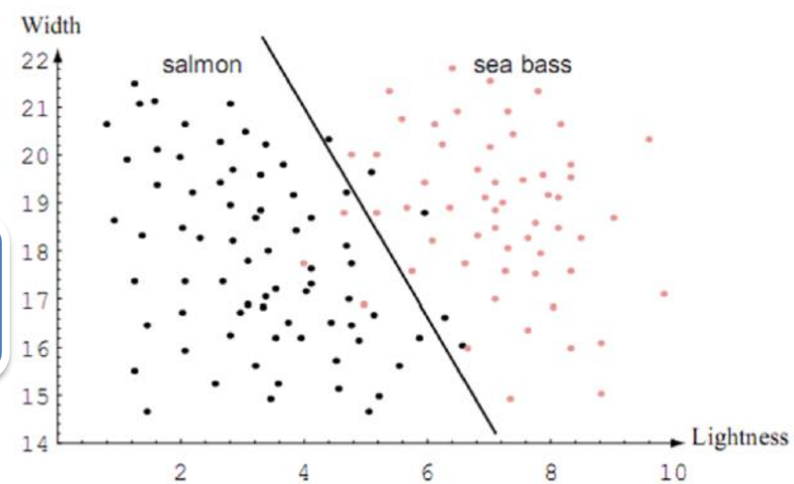
DE MONTFORT UNIVERSITY LEICESTER

PARMA Group

What is Machine Learning?

- Learn a mathematical model from data (measurements)
- **Feature space**: (semi) manually built from relevant data
- A Mathematical model defines a set of parameters $\vec{w} = [m, b]$, to be adjusted from the data (parametric models)
- Supervised methods use the ground truth data (labels y) in a dataset S_l to adjust the model parameters according to an **error measure or loss function** $\mathcal{L}(\vec{w}, S)$

$$y = mx + b$$



A simple Linear classifier: Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

DE MONTFORT UNIVERSITY LEICESTER

PARMA Group

PARMA Activities

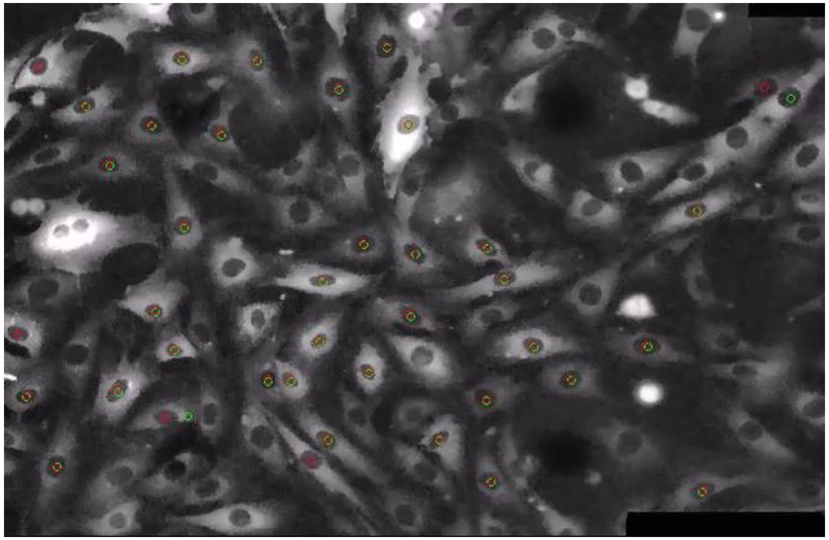
- 1st International Symposium on Machine Learning Applications 2017
- Escuela de Veranillo en HPC 2017, Coorganized with CNCA
- Costa Rica Big Data School 2018, Coorganized with CNCA
- 2nd International Symposium on Machine Learning Applications
- Escuela de Veranillo en Machine Learning (EVeMa 2018)
- 1st Symposium of AI applied to medical imaging 2019
  https://www.tec.ac.cr/eventos/simposio-inteligencia-artificial-aplicada-imagenes-medicas-0

1st Symposium of AI applied to medical imaging, Costa Rica

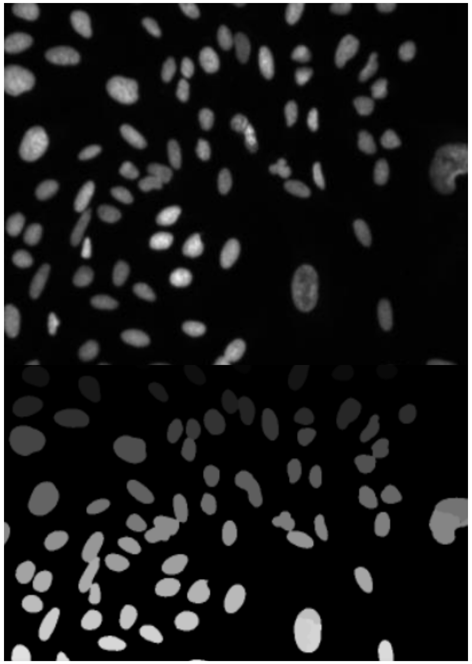
PARMA Group

Cell segmentation and tracking

- **Objective**: develop an in-silico model of regulation networks for proliferation, with DNA damage for malign cells, using per cell behavior data
- Accurately segment and track cells in fluorescence based microscopy for per cell growth estimation, to feed in-situ data of chemotherapy reaction to the model
- Dataset: Broad Bioimage Benchmark Collection, 84-well microplate containing U2OS (bone) cells stained with Hoechst 33342 nuclei markers, with 768 pixel-wise annotated images of 696 x 520 pixels



Fluorescence based microscopy, Microbiology lab, UCR



BBBC 06 dataset sample

PARMA Group

Cell segmentation and tracking

- Ground truth generation for cell segmentation is very time consuming and prone to variability
- **Active learning**: Ground truth generation tools with machine learning for speeding up the GT generation process

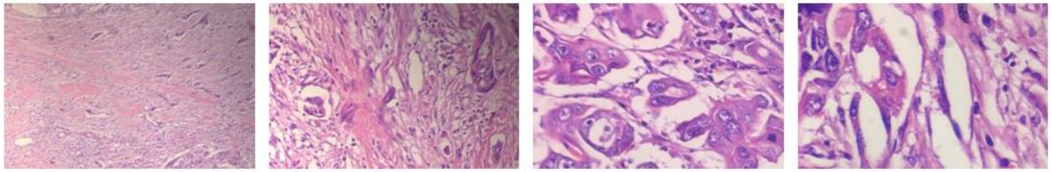
hulihealth

Ground truth authoring tool in development (PARMA-Huli Health): https://insight-gt.hulilabs.xyz/canvas/

PARMA Group

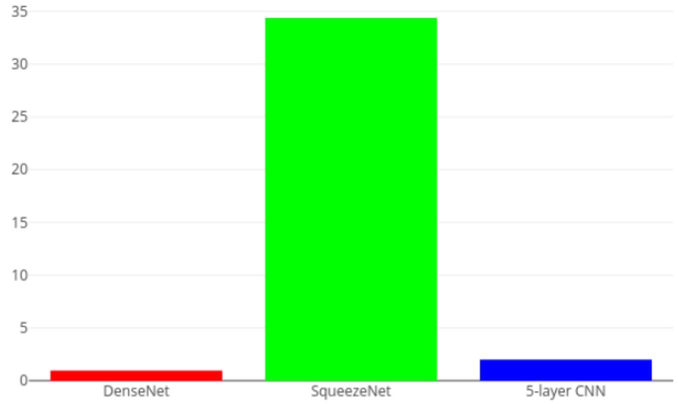
Breast Histopathologic images classification

- Classify breast histopathologic images in two major classes: benign and malign. These two classes are then subdivided into four subtypes each
- Data set: Breast Cancer Histopathological Database [24] known as BreakHis, composed of 7,909 microscopic images of breast tumor tissue collectedfrom 82 patients using different magnifying factors
- Squeezenet yielded the highest accuracy/parameter efficiency, allowing its usage in environments with limited computational resources (90% F1 Score)
- Web demo: http://martinvc96.pythonanywhere.com/demo/



40× 100× 200× 400×

Samples of the four magnifying factors presented in BreakHis dataset



PARMA Group

Organizational challenges for Using AI in medical imaging

- Define the right question (and data) to be answered by the model
- Challenges for the domain expert: Define proper error metrics for the model ¿How much is good enough? **¿Are the results statistically significant?**
- Build up **a labeling standard**, and follow it carefully, and gather colleagues in the field to participate during the GT generation… …many labeled observations needed

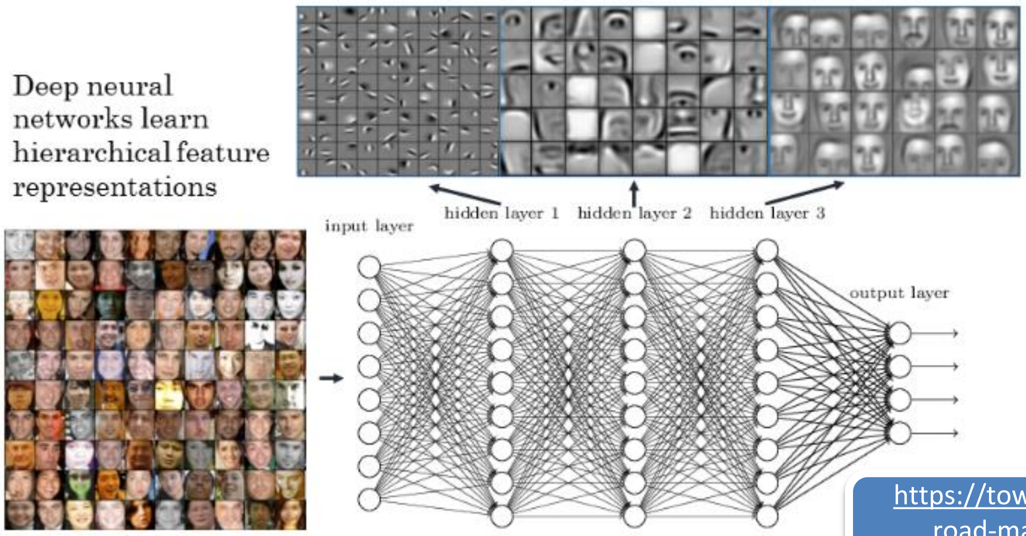
Our domain expert collaborators

- Dr. Mario Umaña, Radiologist, Marcial Fallas Clinic, CCSS
- Dr. Jose Luis Quiros, Pathologist, Max Peralta Hospital, CCSS
- Dr. Luis Fernando Chavarría, Costa Rican Ongologic Center
- M. Sc. Carolina Masis, Imaging Technologies, University of Costa Rica

PARMA Group

Organizational challenges for Using AI in medical imaging

- Understand which data has been used for training and validating a given model. Is that data valid for the domain where I am using the model?
- Understand and gain insight about what is the model telling us, do we need to update our initial knowledge about the problem given the model behavior, How to use the learned features?
- Can we fuse also physiological, physical, socio-economical, molecular or genetical information along images to diagnose diseases?
- Computational Pathology, Computational Radiology….



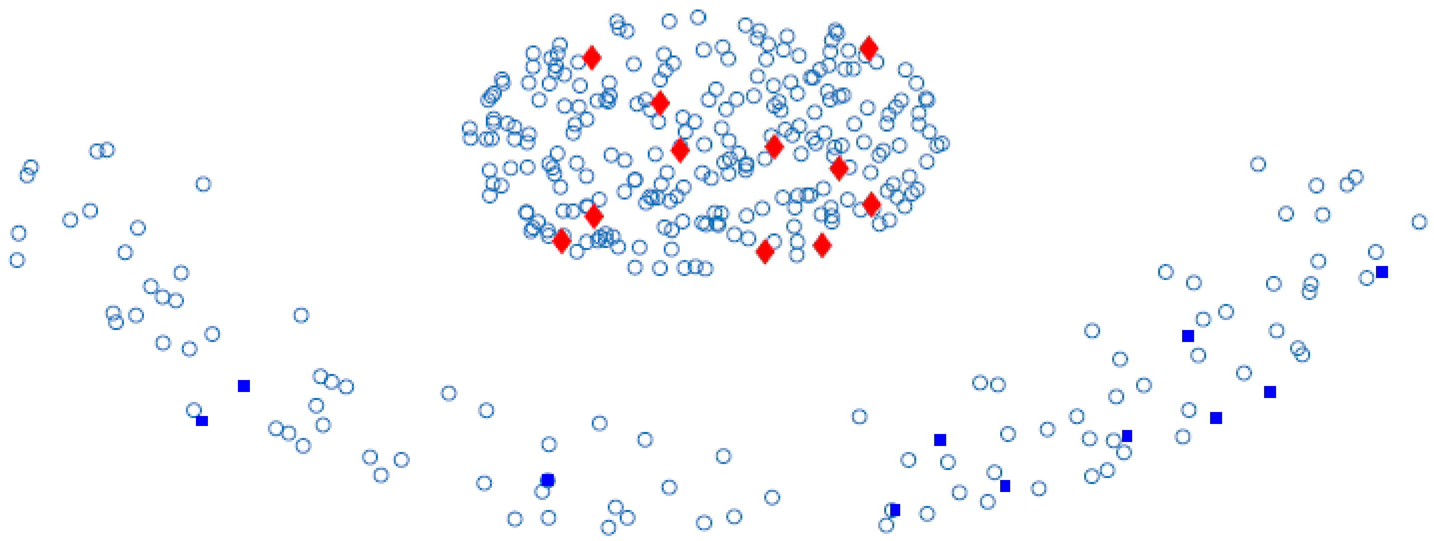
https://towardsdatascience.com/a-road-map-for-deep-learning-b9aee0b2919f

PARMA Group

Semi-supervised Learning (SSDL)

- Semi-supervised learning **uses the labeled dataset** $S_l = \{(\vec{x}_1, y_1), \dots, (\vec{x}_{n_l}, y_{n_l})\}$, and an unlabeled dataset $S_u = \{\vec{x}_1, \dots, \vec{x}_{n_u}\}$, to minimize a loss function

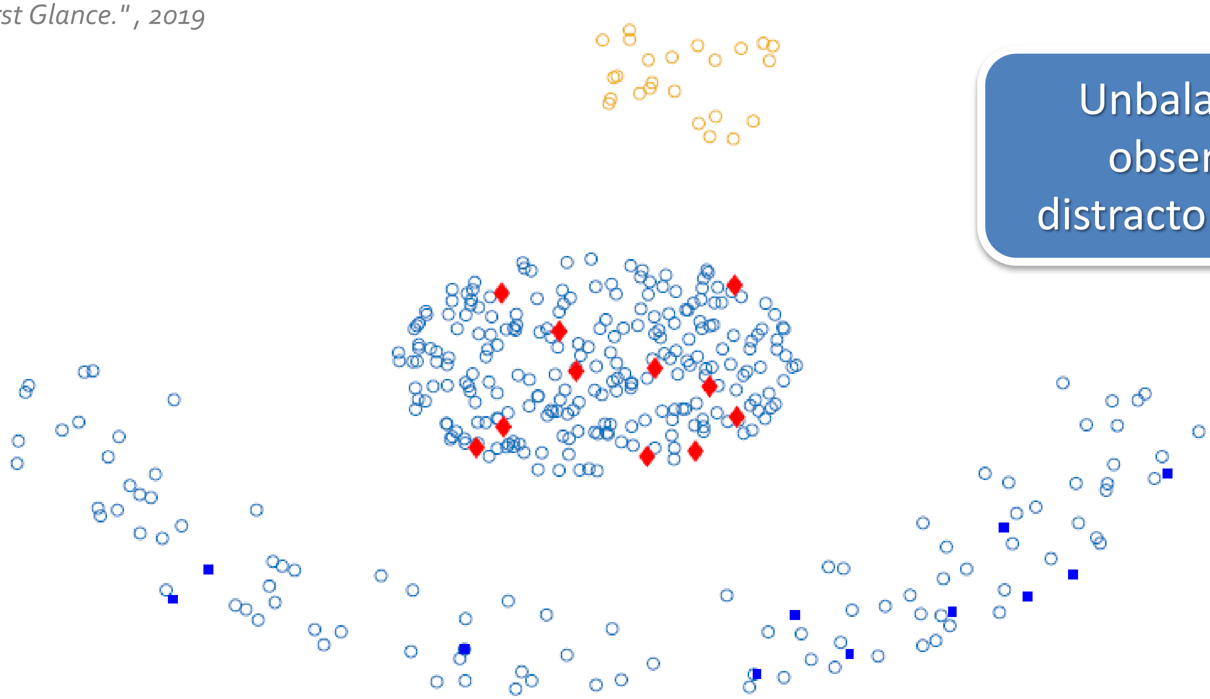
$$t(S_l \cup S_u) = \mathrm{argmin}_{\vec{w}} \mathcal{L}(S_l \cup S_u, \vec{w})$$



Circles correspond to unlabeled observations, diamond and squares to labeled observations

SSDL state of the art: Pending gaps

- Unbalanced data with outliers and observations from distractor classes in the unlabeled dataset impact and how to become robust to it has not been fully explored yet
- “We studied all suffered when the unlabeled data came from different classes than the labeled data — a realistic scenario that to our knowledge is drastically understudied” (Goodfellow et. Al, 2018)
- Furthermore, a proper metric to assess if the dataset is big enough is pending, Mendez, M., *S. Calderon, and P. N. Tyrrell. "Using Cluster Analysis to Assess the Impact of Dataset Heterogeneity on Deep Convolutional Network Accuracy: A First Glance."*, 2019



Unbalanced data and observations from distractor classes (yellow)

DE MONTFORT UNIVERSITY LEICESTER

PARMA Group

sacalderon@itcr.ac.cr

Facebook: https://www.facebook.com/ParmaResearchGroup