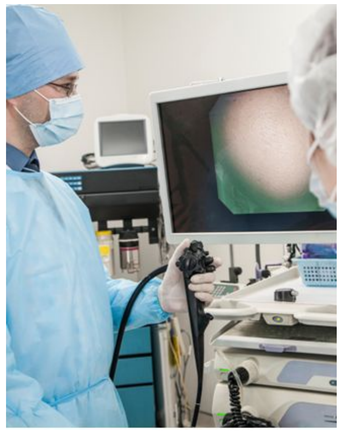
Assessing deep learning methods for the identification of kidney stones composition in endoscopic images

Andres Varelo-Silgado, Francisco Lopez-Tiro, Mauricio Mendez-Ruiz, Jonathan El Beze, Miguel Gonzalez, Vincent Estrade, Gilberto Ochoa Ruiz, Christian Daul

More info

1. Medical Context and Motivation



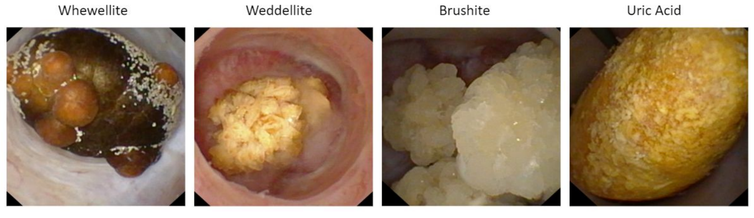
Urolithiasis disease is the formation of kidney stones

Morphological information provides very valuable information for diagnosis purposes

But this component it is lost in an endoscopic intervention known as **"dusting" (laser lithotripsy)**

Automated image analysis could alleviate these issues!!!

Not a trivial problem: ureteroscopy images are characterized by poor lighting and images from kidney stones present high **inter-class similarities and intraclass variations**



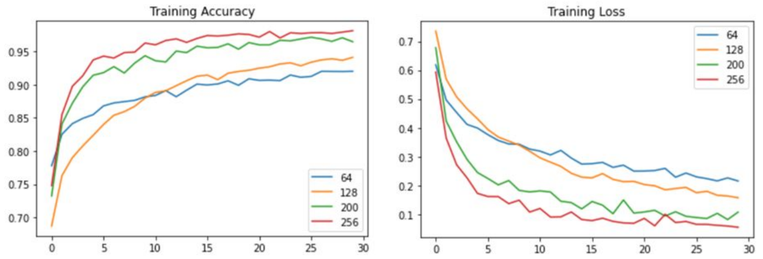
In this work we compare DL methods to best ML classification methods for automatically classifying **ureteroscopic in vivo kidney stones**

The study was carried by a group of machine learning and urology specialists from the institution listed below:

Most recent works in make use of **extracted (ex vivo)** kidneys stones, captured under highly controlled lighting conditions: **not very representative of clinical settings**

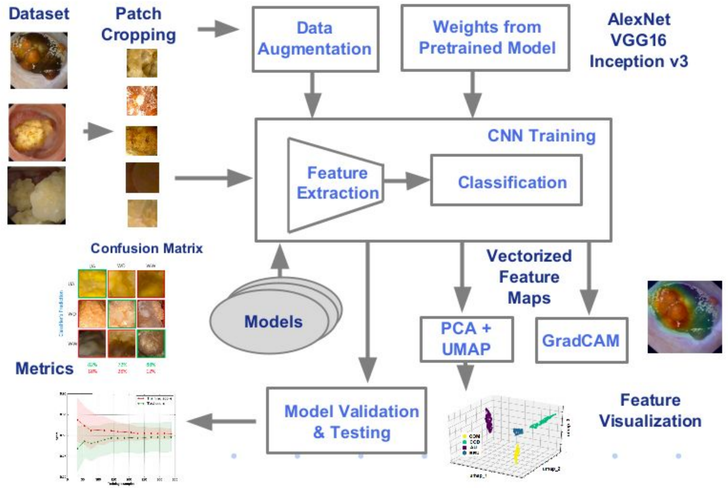
2. Materials and Methods

We build a **dataset** containing 90 surface and 87 cross section images from 4 classes, the dataset was extended scanning patches with severals dimensions and selecting the size of 256x256 pixels **6000 samples**.



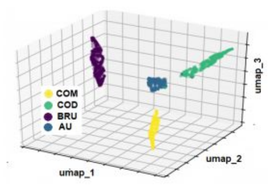
Accuracy changing patch sizes

To ML methods we performed **feature selection** studies choosing the **hue, saturation and lightness channel information** and **Locally Binary Patterns** (LBP)

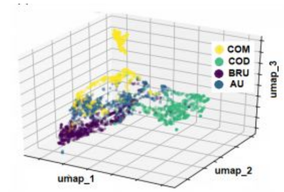


The figure above show **Deep Learning** architecture

Deep Learning Backbone helps to improve the classes separability!!



UMAP visualization using "deep features"



UMAP visualization using HSI+LBP Features

3. Results and Discussion

All models were trained with surfaces, sections and mixed patches.

The best results were obtained by **XGBoost** and **Inception v3**

Classifier	Surface		Section		Mixed	
	P	R	P	R	P	R
Random Forest	0.87	0.82	0.82	0.82	**0.91**	**0.91**
XGboost	0.93	0.93	0.89	0.89	**0.96**	**0.96**
AlexNet	0.93	0.95	0.83	0.82	0.92	0.92
VGG19	0.95	0.96	0.91	0.92	0.94	0.92
Inception	**0.98**	**0.97**	**0.94**	**0.96**	**0.97**	**0.98**

We demonstrate that is possible to classify **in vivo kidney stones** reliably!
Currently, we are trying to explain the DL models with **GradCAM**: