## Diagnostic of mammary carcinoma in dogs: a machine learning spectroscopy approach

Maria I. da Silva<sup>1</sup>, Paulo H. R. Amaral<sup>1</sup>, Geovanni D. Cassali<sup>2</sup>, Ana P. Garcia<sup>2</sup>, Juan C. González<sup>1</sup>



<sup>1</sup>Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

<sup>2</sup>Laboratório de Patologia Comparada, Departamento de Patologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

#### Overview

- New diagnostic of mammary carcinoma in dogs.
- Machine learning spectroscopy approach.
- Mammary gland samples.
- Measurements done by variable angle spectroscopic ellipsometry.
- Two-class classification  $\rightarrow$  AUROC = 0.94.
- Label-free, low cost, results in 10 minutes.

### Introduction

The development of new techniques are of great importance for the diagnostic of cancer. In this work we present a new approach for diagnostic of mammary carcinoma from mammary gland samples using machine learning spectroscopy. Samples were measured by variable angle spectroscopy ellipsometry and the raw data modeled by machine learning algorithms.

#### Methods

Mammary gland samples.

(a)

(c)

- Samples mounted on 22 mm x 22 mm Corning #11/2 glass coverslips.
- Measurements: VASE, 240 nm to 1700 nm, six AOI.
- ML models: KNN, LR, SVM, XGBoost, MLP.
- ML approach: Two-class nested cross-validation.
- Approved by Ethics Committed on the Use of Animals at UFMG: 83/2021.

Experimental Design



0

1. Experimental design. (a) Sample preparation. (b) VASE Fia. measurements. (c) Pre-processing of data for ML. (d) ML modeling



Fig. 2. Rotating Compensator Variable Angle Ellipsometer used in the experiments.

Results (b) (a) 1.0 0.8 0.05 Benign 0.8 Sate 0.6 True 0.6 0.4 Malignant 0.24 Pos 0.4 0.2 LIP 0.2 mean AUROC = 0.940.0 0.0 0.2 0.4 0.6 0.8 1.0 Predicted False Positive Rate

Fig. 3. (a) Normalized confusion matrix and (b) ROC curve for the best ML model: Logistic Regression.



Fig. 4. Computer application user interface and scores for the optimized Logistic Regression model.

#### Conclusions

- A new technique and a computer application were developed for the accurate, fast, and low-cost diagnostic of mammary carcinoma in dogs using machine learning spectroscopy.
- Variable angle spectroscopy ellipsometry and logistic regression machine learning algorithm resulted in optimal experimental data processing techniques for the diagnostic.
- Excellent performance, with AUROC of 94%.
- · The method could be expanded for the diagnostic of breast cancer.

# Acknowledments









