1067-B

Machine learning spectroscopy for in-vitro recognition of tumoral and non-tumoral cells

Juan C. González, Paulo H. R. Amaral, Lidia M. Andrade, Maria I. N. da Silva

Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

Overview

- New in-vitro label-free spectroscopic approach for classification of tumoral and non-tumoral cell lines.
- Machine learning spectroscopy approach.
- Cultures of A431, FaDU, SCC-4, and HaCat cells.
- Measurements done by variable angle spectroscopic ellipsometry.
- Two-class classification → Accuracy: 77% - 99%.
- Multi-class classification → Accuracy = 79%.

Introduction

The development of new techniques for recognition of tumoral and non-tumoral tissues at the cellular level are of great importance for the diagnostic of cancer. In this work we present an important step in that direction with a new in-vitro label-free spectroscopic approach capable of differentiating samples of tumoral and non-tumoral cell lines by using machine learning spectroscopy techniques.

Methods

- Cultures of epidermoid carcinoma A431 cells, hypopharyngeal carcinoma FaDU cells, squamous carcinoma SCC-4 cells, and epidermal keratinocyte HaCat cells.
- Cells were grown on 22 mm x 22 mm Corning #1½ glass coverslips.
- Measurements: VASE, 240 nm to 1700 nm, AOI of 55°.
- ML models: SVM, XGBoost, MLP.
- ML approach: Two-class and Multi-class nested cross-validation.

Experimental Design

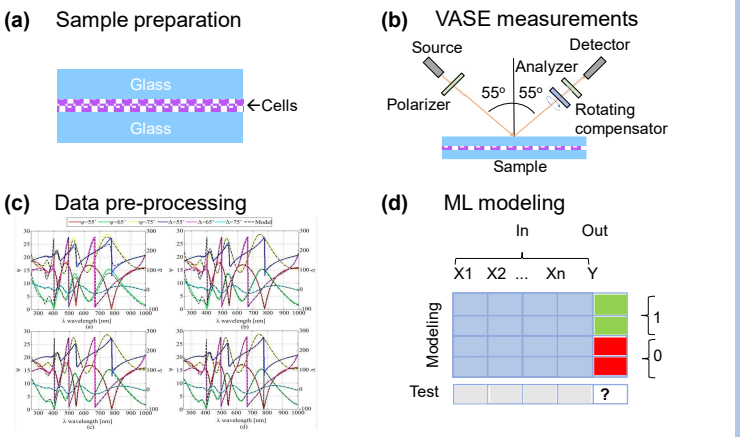

Fig. 1. Experimental design. (a) Sample preparation. (b) VASE measurements. (c) Pre-processing of data for ML. (d) ML modeling

Results

Fig. 2. Two-Class confusion matrix for optimized ML models SVM, XGBoost, and MLP.

Results

Table I. Performance metrics for all combinations of optimized two-class ML models and samples.

Model	Group	Cell	AUC	Acc (%)	Spec (%)	Sens (%)
SVM	Tumor vs. Non-Tumor	HaCat vs A431	0.88 ± 0.01	81.9 ± 0.3	79.0 ± 0.3	88.1 ± 0.7
		HaCat vs FaDu	1.00 ± 0.01	98.5 ± 0.2	99.0 ± 0.2	96.7 ± 0.2
		Hacat vs SCC4	0.99 ± 0.01	98.0 ± 0.1	98.4 ± 0.1	95.8 ± 0.3
	Non-tumor vs. Non-tumor	A431 vs SCC4	1.00 ± 0.01	98.7 ± 0.1	99.7 ± 0.1	96.0 ± 0.1
		A431 vs FaDu	0.98 ± 0.01	93.7 ± 0.2	95.7 ± 0.2	90.4 ± 0.5
		SCC4 vs FaDu	0.91 ± 0.01	87.5 ± 0.3	92.0 ± 0.3	80.8 ± 0.7
XGBoost	Tumor vs. Non-Tumor	HaCat vs A431	0.83 ± 0.01	76.9 ± 0.6	78.0 ± 0.6	74.0 ± 1.0
		HaCat vs FaDu	0.95 ± 0.01	92.3 ± 0.3	94.2 ± 0.3	85.6 ± 0.7
		Hacat vs SCC4	0.98 ± 0.01	96.2 ± 0.2	97.0 ± 0.2	92.3 ± 0.3
	Non-tumor vs. Non-tumor	A431 vs SCC4	1.00 ± 0.01	97.5 ± 0.2	98.5 ± 0.2	94.9 ± 0.5
		A431 vs FaDu	0.96 ± 0.01	92.8 ± 0.1	97.8 ± 0.1	84.3 ± 0.3
		SCC4 vs FaDu	0.83 ± 0.01	83.1 ± 0.3	94.4 ± 0.3	66.2 ± 0.6
MLP	Tumor vs. Non-Tumor	HaCat vs A431	0.85 ± 0.01	77.5 ± 0.4	74.9 ± 0.4	82.0 ± 1.0
		HaCat vs FaDu	1.00 ± 0.01	98.0 ± 0.2	98.3 ± 0.2	96.8 ± 0.2
		Hacat vs SCC4	0.99 ± 0.01	96.5 ± 0.4	97.0 ± 0.4	93.7 ± 0.6
	Non-tumor vs. Non-tumor	A431 vs SCC4	1.00 ± 0.01	98.7 ± 0.1	99.8 ± 0.1	95.8 ± 0.2
		A431 vs FaDu	1.00 ± 0.01	97.1 ± 0.2	98.9 ± 0.2	94.1 ± 0.4
		SCC4 vs FaDu	0.89 ± 0.01	80.9 ± 0.6	77.6 ± 0.6	86.0 ± 1.0

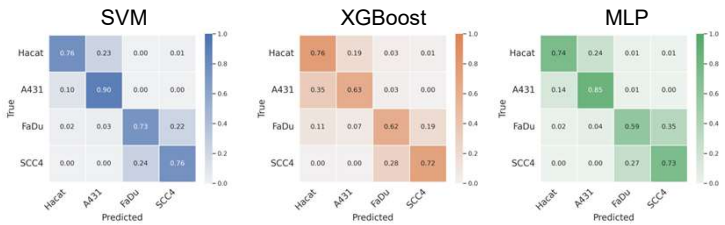

Fig. 2. Multi-class confusion matrix for optimized ML models SVM, XGBoost, and MLP.

Table I. Performance metrics for multi-class ML models.

Model	Accuracy (%)	Specificity (%)	Sensitility (%)
SVM	79 ± 1	93 ± 1	78 ± 2
XGBoost	71 ± 2	75 ± 2	68 ± 2
MLP	74 ± 1	91 ± 2	73 ± 2

Conclusions

- New in-vitro label-free spectroscopic approach for classification of cell lines: accurate, fast, and low-cost.
- VASE and SVM machine learning algorithm resulted in optimal experimental and data processing techniques for the diagnostic.
- Excellent performance for two-class.
- Good performance for multi-class.
- The method could be expanded for cellular diagnostic of cancer.

References

- ML CELL ID. 2022. Patent register number: 20210033. Register date: 07/02/2022. Universidade Federal de Minas Gerais. Amaral P. H., González J. C.; Andrade L. M.
- BR1020210252480: PROCESSO PARA CLASSIFICAÇÃO DE CÉLULAS QUANTO À LINHAGEM CELULAR E USO. INPI - Instituto Nacional da Propriedade Industrial. Brazil. Register date: 15/12/2021.

Acknowledments

MCTI Inteligência Artificial Inovação Aberta
CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
FAPEMIG
MAIN HR Minas Artificial Intelligence Network on Human Reproduction
CAPES