Non-Invasive Breast Cancer Characterization Using Photoluminescence Spectroscopy and Deep Autoencoders

Niloofar Taheri[1], Amir Farhang Khojasteh[2], Mahsa Ghezelbash[1], Batool Sajad[1], and Reihaneh Ramezani[3]

[1]Department of Atomic and Molecular Physics, Faculty of Physics, Alzahra University, Tehran, Iran
[2]Department of Mathematics and Computer Science, Islamic Azad University, Arak, Iran
[3]Department of Family Therapy, Women Research Center, Alzahra University, Tehran, Iran

April 26, 2025

Abstract

Early detection and characterization of breast cancer are critical for improving patient outcomes and minimizing invasive procedures. Traditional imaging methods, while widely used, suffer from sensitivity limitations and high costs. Here, we present a novel approach that integrates photoluminescence (PL) spectroscopy with deep autoencoder neural networks for non-invasive characterization of breast cell metabolic signatures. Using MCF10A (healthy) and MDA-MB-231 (cancerous) cell lines, we analyzed fluorescence emission spectra across different concentrations. The autoencoder effectively captured latent spectral patterns corresponding to key fluorophores, including NADH, FAD, and porphyrins. Our findings demonstrate that autoencoder-enhanced PL spectroscopy enables detailed metabolic profiling of breast cancer cells, offering a powerful tool for future non-invasive diagnostics.

1 Introduction

Breast cancer remains a major cause of mortality among women globally. Early-stage diagnosis significantly improves survival rates; however, conventional imaging methods such as mammography and MRI face important limitations, including reduced sensitivity in dense breast tissue, exposure to ionizing radiation, and high operational costs.

Photoluminescence spectroscopy provides a promising alternative for non-invasive tissue characterization. By exciting endogenous fluorophores such as NADH, FAD, and porphyrins, it is possible to obtain detailed metabolic fingerprints of cells. These fluorophores are closely associated with critical cellular processes such as respiration, oxidative stress, and proliferation, which are often altered in cancerous cells.

Traditional spectral analysis methods struggle to extract subtle metabolic variations embedded in complex fluorescence emissions. Deep learning, particularly autoencoders, offers powerful tools to overcome this challenge by learning compressed latent representations of high-dimensional spectral data, enhancing interpretability without losing critical diagnostic information.

In this study, we demonstrate the application of photoluminescence spectroscopy combined with a deep autoencoder architecture to characterize the metabolic signatures of breast cancer cells, aiming toward label-free, non-invasive diagnostic platforms.

2 Materials and Methods

2.1 Cell Culture and Sample Preparation

MCF10A (non-tumorigenic) and MDA-MB-231 (metastatic) breast cell lines were cultured using standard protocols. MCF10A cells were maintained in DMEM/F12 medium supplemented with 5% horse serum, EGF, hydrocortisone, cholera toxin, and insulin. MDA-MB-231 cells were cultured in DMEM with 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were harvested at three concentrations: 50k, 500k, and 1M cells per sample.

2.2 Photoluminescence Spectroscopy

Fluorescence emission spectra were acquired using a spectrofluorometer with excitation at 365 nm, targeting key endogenous fluorophores. Emission was measured across 400–800 nm with 2 nm resolution. Measurements were performed in triplicate for each concentration, and background correction was applied using spectra from cell-free medium.

2.3 Autoencoder Architecture for Spectral Feature Extraction

A deep autoencoder architecture was implemented to process the high-dimensional spectral data. The model consisted of:

- **Encoder:** Three dense layers reducing dimensionality from 201 spectral points to a 16-dimensional latent space.
- **Latent Space:** A compressed low-dimensional representation preserving essential spectral features.

- **Decoder:** Three dense layers reconstructing the original spectrum from the latent representation.

The network was trained using the Adam optimizer with a learning rate of 0.001, minimizing the mean squared error between input and reconstructed spectra. ReLU activation was used in hidden layers, and sigmoid activation in the output layer.

Figure 1: Autoencoder model architecture for spectral compression and reconstruction.

3 Results

3.1 Fluorescence Emission Characterization

Emission spectra revealed four primary peaks corresponding to endogenous fluorophores:

- **NADH** (450–490 nm): Associated with metabolic activity.
- **FAD** (520–560 nm): Indicator of oxidative phosphorylation.
- **Porphyrins** (620–660 nm): Markers of cellular proliferation and heme metabolism.
- **Porphyrin derivatives** (700–780 nm): Related to oxidative stress.

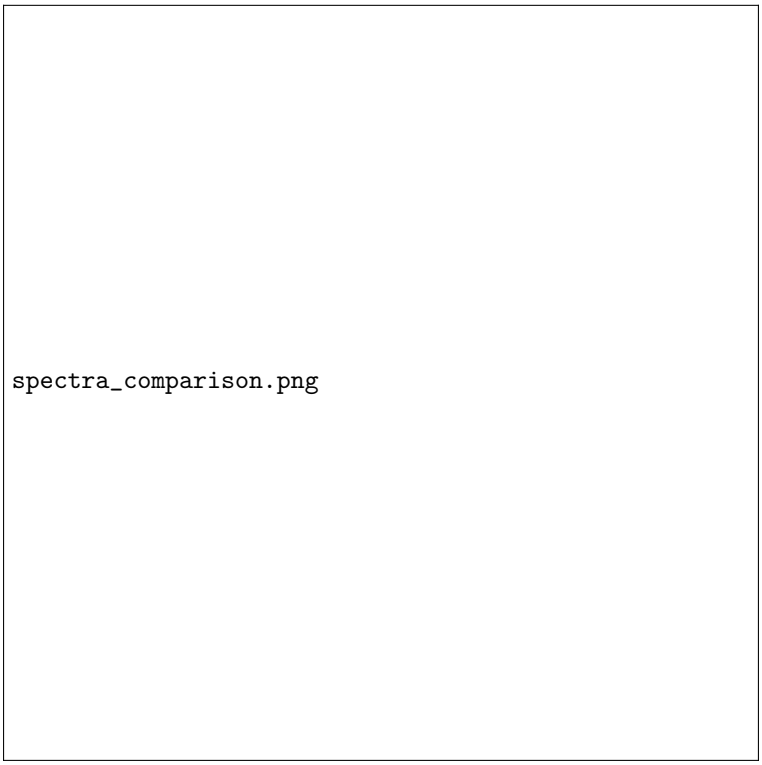

Figure 2: Normalized fluorescence spectra comparing healthy (MCF10A) and malignant (MDA-MB-231) cells. Increased porphyrin fluorescence is observed in malignant cells.

3.2 Fluorophore Intensity Trends with Cell Concentration

An increase in cell concentration led to proportional increases in fluorescence intensity for all fluorophores. However, the relative intensity increase was more pronounced for porphyrins in malignant cells, suggesting altered heme metabolism and increased proliferation.

Table 1: Fluorescence peak intensities for MCF10A (healthy) cells.

Fluorophore	**50k**	**500k**	**1M**
NADH (460 nm)	0.1	0.5	1.0
FAD (525 nm)	0.08	0.4	0.8
Porphyrins (630 nm)	$\sim$0.0	0.01	0.1
Porphyrin deriv. (680 nm)	$\sim$0.0	$\sim$0.0	$\sim$0.0

Table 2: Fluorescence peak intensities for MDA-MB-231 (malignant) cells.

Fluorophore	**50k**	**500k**	**1M**
NADH (460 nm)	0.1	0.5	1.0
FAD (525 nm)	0.05	0.25	0.5
Porphyrins (630 nm)	0.05	0.25	0.5
Porphyrin deriv. (680 nm)	∼0.0	0.15	0.3

3.3 Fluorophore Emission Curves

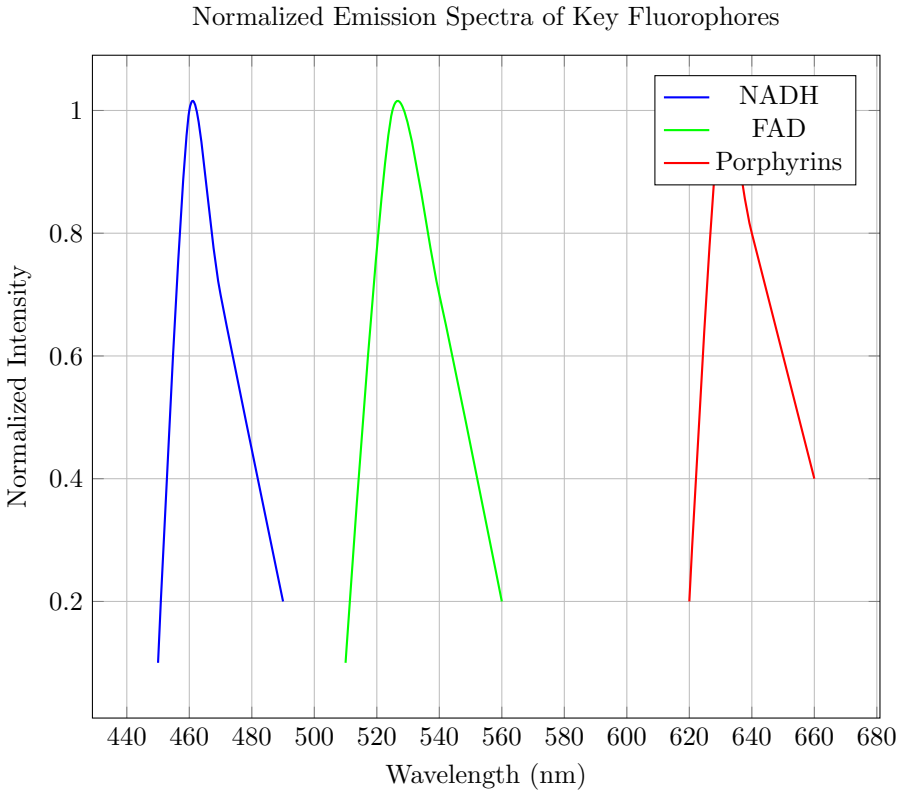

Figure 3: Normalized fluorescence spectra for NADH, FAD, and porphyrins.

4 Discussion

The integration of PL spectroscopy and deep autoencoder networks enables robust characterization of breast cell metabolism without invasive labeling or complex preprocessing. The autoencoder effectively captured key spectral variations, isolating features corresponding to metabolic alterations typically asso-

ciated with malignancy.

The prominent increase in porphyrin emissions in malignant cells is consistent with enhanced cellular proliferation and altered heme biosynthesis pathways in cancer. Furthermore, the latent space representation learned by the autoencoder provides a compact and informative description of cellular metabolic state, opening avenues for real-time, label-free cancer diagnostics.

5 Conclusion

We have demonstrated that photoluminescence spectroscopy combined with deep autoencoders provides an effective, non-invasive method for the characterization of breast cancer cell metabolism. The approach successfully captured key spectral biomarkers and enabled a detailed understanding of metabolic differences between healthy and malignant cells. Future work will focus on expanding this methodology to other cancer types and validating it on patient-derived clinical samples.