# Chapter 1: Introduction

## 1.1 Overview

Breast cancer is one of the most significant health concerns globally, affecting millions of women annually. Early and accurate diagnosis is critical for effective treatment and survival. Histopathological examination of tissue samples under a microscope remains the gold standard for breast cancer diagnosis. However, manual examination is a time-consuming process that relies heavily on the pathologist’s expertise, often resulting in inter-observer variability and diagnostic inconsistency (Litjens et al., 2017).

With the rise of artificial intelligence (AI) and machine learning (ML), particularly deep learning techniques, the automation of histopathological image analysis has emerged as a promising approach. Convolutional Neural Networks (CNNs) have demonstrated superior performance in feature extraction and pattern recognition, outperforming traditional machine learning methods in medical image classification tasks (Araujo et al., 2017). Datasets such as the Breast Cancer Histopathological Database (BreakHis) have become key benchmarks for developing and evaluating CNN models on breast cancer image classification (Spanhol et al., 2016).

## 1.2 Purpose of the Project

The primary aim of this project is to build an automated classification system using deep learning for the detection of benign and malignant breast cancer images from the BreakHis dataset. This project focuses on evaluating multiple CNN architectures, including VGG16 (Simonyan & Zisserman, 2014), ResNet50, and EfficientNet, to identify the most effective approach for accurate classification across various magnifications. Moreover, the project integrates explainability methods such as Grad-CAM and LIME to visualize model decision regions, enhancing interpretability and clinical trust.

## 1.3 Research Question

1. Can convolutional neural network (CNN) architectures, specifically VGG16, ResNet50, and EfficientNet, accurately classify benign vs malignant tumor tissue images using the BreakHis dataset?
2. Which visual features (texture, shape, patterns) are most influential in classification, and can explainability methods (e.g. Grad-CAM, LIME) highlight these regions?
3. How does performance vary across different image magnifications (200×, 400×), and can models generalize across magnification levels?

## 1.4 Project Aims

1. To design, train, and evaluate multiple CNN architectures (VGG16, ResNet50, EfficientNet) for binary classification of histopathological images.
2. To compare the models’ performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.
3. To investigate the impact of magnification levels on model performance.
4. To apply Grad-CAM and LIME explainability techniques to visualize and interpret model predictions.
5. To provide insights and recommendations for improving automated breast cancer diagnosis using deep learning.

# Chapter 2: Background

Automated analysis of histopathological images using deep learning has gained considerable attention in the medical and research communities. This section critically reviews significant peer-reviewed studies relevant to breast cancer image classification and explainable AI, highlighting their methodologies, datasets, results, and limitations in the context of the current project.

(Spanhol et al., 2016) present BreakHis, a curated dataset of 9,109 RGB images from 82 patients at 40×, 100×, 200× and 400× magnifications, labelled benign/malignant. The paper provides baseline experiments (mainly traditional texture descriptors and early deep models) and highlights challenges such as class imbalance and magnification effects. Its main strength is establishing a public, multi-magnification benchmark with patient-level provenance, which catalysed methodological progress. However, the baselines under-represent today’s deeper architectures and contemporary explainability. For this project, Spanhol et al. justify adopting patient-separated splits and evaluating both magnification-specific and magnification-independent performance, ensuring comparability and guarding against data leakage.

(Araujo et al., 2017) demonstrate one of the early CNN-based pipelines for breast histology classification using multi-scale image representations to capture both nuclear detail and global tissue architecture. Their model outperforms hand-crafted approaches on four-class and binary tasks, underscoring the value of deep feature learning and careful stain handling. Strengths include a principled multi-scale design and reproducible evaluation; limitations include modest dataset size and limited interpretability tooling. For this project, their results motivate scale-aware modelling and robust preprocessing/augmentation; we extend this by benchmarking modern backbones and adding systematic explainability.

(Hao et al., 2022) proposed a hybrid approach that combines deep semantic features extracted from a pre-trained DenseNet201 convolutional network with handcrafted Gray Level Co-occurrence Matrix (GLCM) texture descriptors to classify BreakHis breast histopathology images into benign and malignant categories. After feature extraction, an SVM classifier was applied for final prediction. Their experiments, performed on images at multiple magnifications (40×–400×), achieved high classification accuracies, above 95 % across several settings—and demonstrated that integrating texture information can complement deep feature representations. The authors concluded that hybrid feature fusion improves robustness compared to standalone CNNs. However, they also acknowledged that this multi-stage pipeline increases computational complexity and may reduce scalability for end-to-end deployment. This study provides a strong empirical baseline and motivates testing whether end-to-end CNNs (VGG16, ResNet-50, EfficientNet) can achieve comparable accuracy with greater simplicity and interpretability.

(Selvaraju et al., 2017) introduce Grad-CAM, producing class-discriminative heatmaps by projecting the gradient signal onto convolutional feature maps. Grad-CAM has become a de-facto standard in medical imaging to contextualise CNN predictions. Its strengths are intuitive visualisations and model-specific localisation; limitations include coarse resolution and sensitivity to layer choice. In this project, Grad-CAM supports auditing of benign/malignant decisions and complements metric-based comparisons by revealing the morphological evidence that drives predictions, crucial for clinical trust and for diagnosing failure modes.

# References

Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polonia, A., & Campilho, A. (2017). Classification of breast cancer histology images using Convolutional Neural Networks. *PLOS ONE*, *12*(6), e0177544. https://doi.org/10.1371/JOURNAL.PONE.0177544

Hao, Y., Zhang, L., Qiao, S., Bai, Y., Cheng, R., Xue, H., Hou, Y., Zhang, W., & Zhang, G. (2022). Breast cancer histopathological images classification based on deep semantic features and gray level co-occurrence matrix. *PLOS ONE*, *17*(5), e0267955. https://doi.org/10.1371/JOURNAL.PONE.0267955

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/J.MEDIA.2017.07.005

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-October*, 618–626. https://doi.org/10.1109/ICCV.2017.74

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/pdf/1409.1556

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, *63*(7), 1455–1462. https://doi.org/10.1109/TBME.2015.2496264