Toronto Metropolitan University | Data Science & Analytics Graduate Studies

2024-2025 M.Sc. in Data Science and Analytics
DEEP LEARNING VS MACHINE LEARNING: A PERFORMANCE COMPARISON FOR FUNGAL INFECTION CLASSIFICATION

Ahnaf Shahriyar Chowdhury　　　**Supervisor: Dr. Youcef Derbal**

Objective: The objective of this study is to compare the effectiveness of traditional machine learning models and deep learning models in the classification of fungal infections from microscopic images. Specifically, traditional algorithms such as Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) were tested using handcrafted feature extraction methods like Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG). These models were then compared against a deep learning model, ResNet50, which automatically extracts hierarchical features. In addition, this research also explores the use of Explainable AI (XAI) techniques such as Grad-CAM to provide visual insight into the decision-making process of deep learning models.

Background:

- Fungal infections are small, morphologically similar to other microorganisms, and difficult to diagnose with manual microscopy.
- Traditional methods are time-consuming and prone to observer variability.
- Deep learning (CNNs, ViTs) enables automated classification, reducing diagnostic time and improving accuracy.
- CNNs are particularly effective at extracting hierarchical features and capturing complex image relationships.

Methodology:

- **Dataset:** DeFungi (UCI repository) → 5 fungal classes (balanced to 500, 1500 and2500 images per class using augmentation and undersampling).

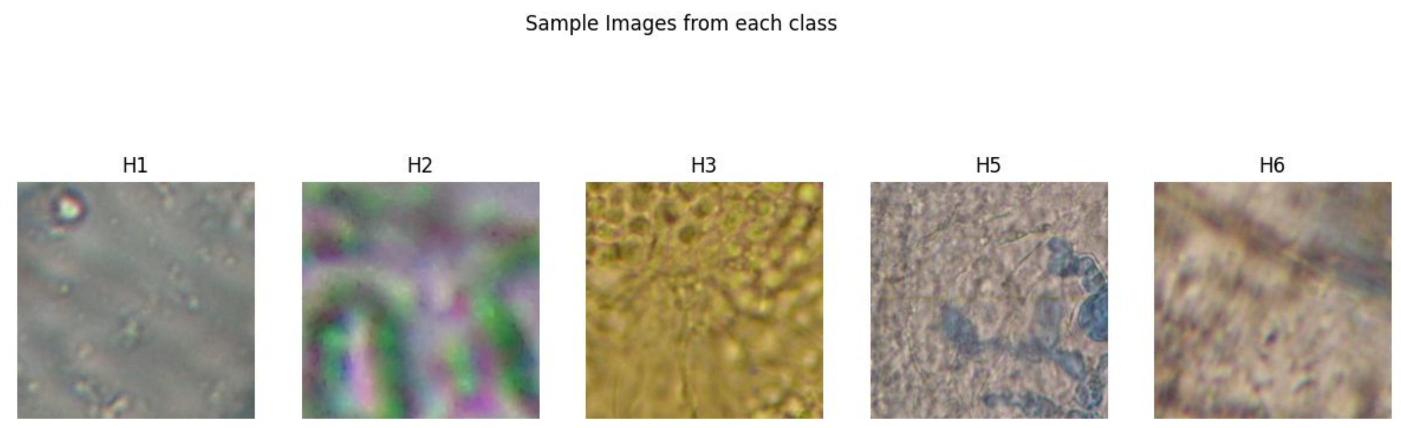

Figure 01 - Sample images from each class

- **Preprocessing:** Resized (224×224), converted to grayscale for ML models.
- **ML Pipeline:**
  - Feature Extraction: LBP, HOG, LBP+HOG
  - Classifiers: SVM, Random Forest, XGBoost
- **DL Pipeline:**
  - ResNet50 (pre-trained CNN) directly on RGB images.
- **Evaluation Metrics:** Accuracy, Precision, Recall, F1, AUC-ROC.
- **Interpretability:** Grad-CAM to visualize image regions influencing DL model predictions.

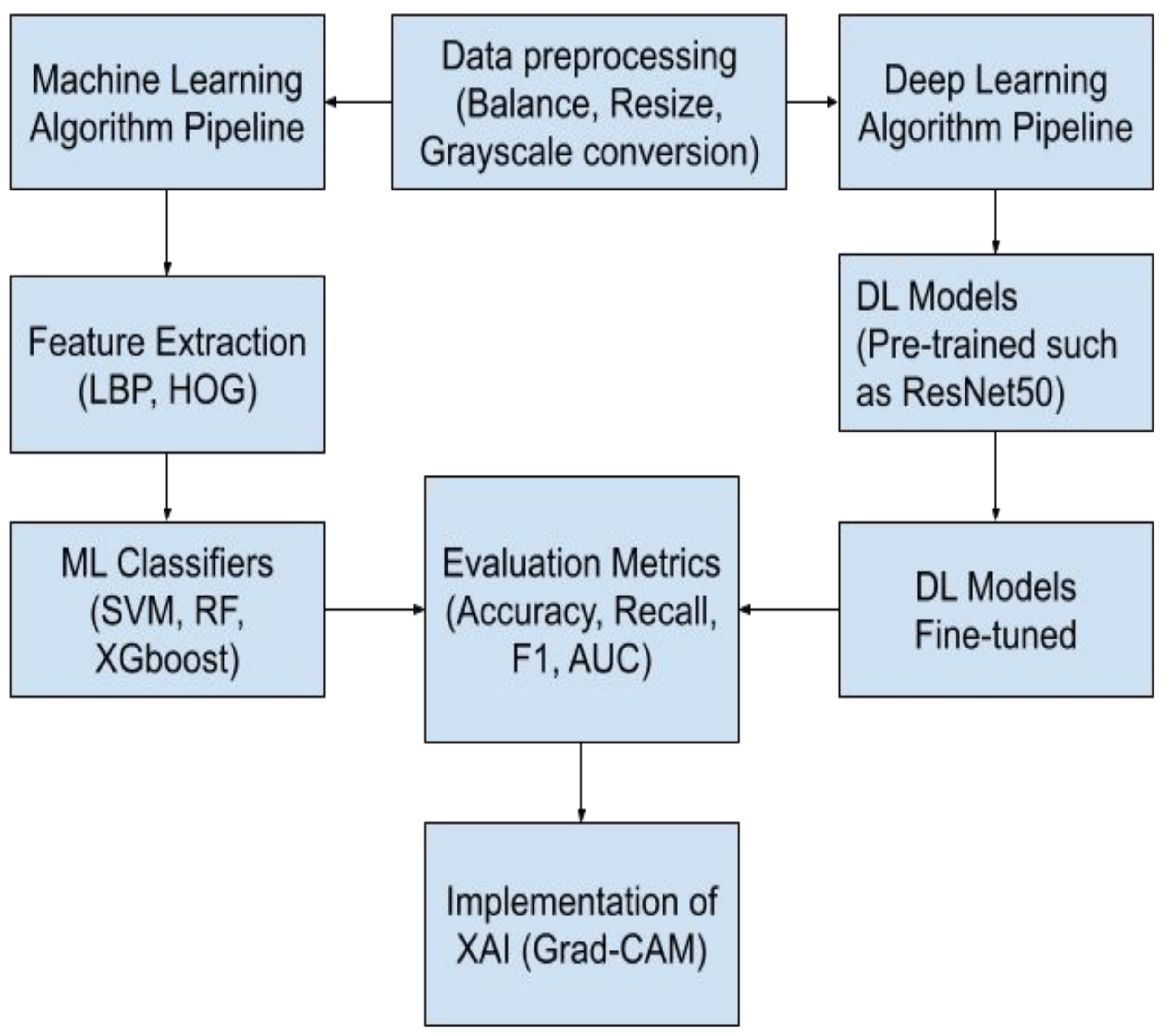

Figure 02 - Overall project workflow

Results:

- **ResNet50 (DL)** outperformed all ML models.
  - Accuracy: 0.68 to 0.73 as dataset increased (500 to 2500 images/class).
  - AUC improved from 0.917 to 0.941.
- **ML Models** showed modest improvement with larger datasets.
  - Best ML: XGBoost (0.58 accuracy at 2500 images/class).

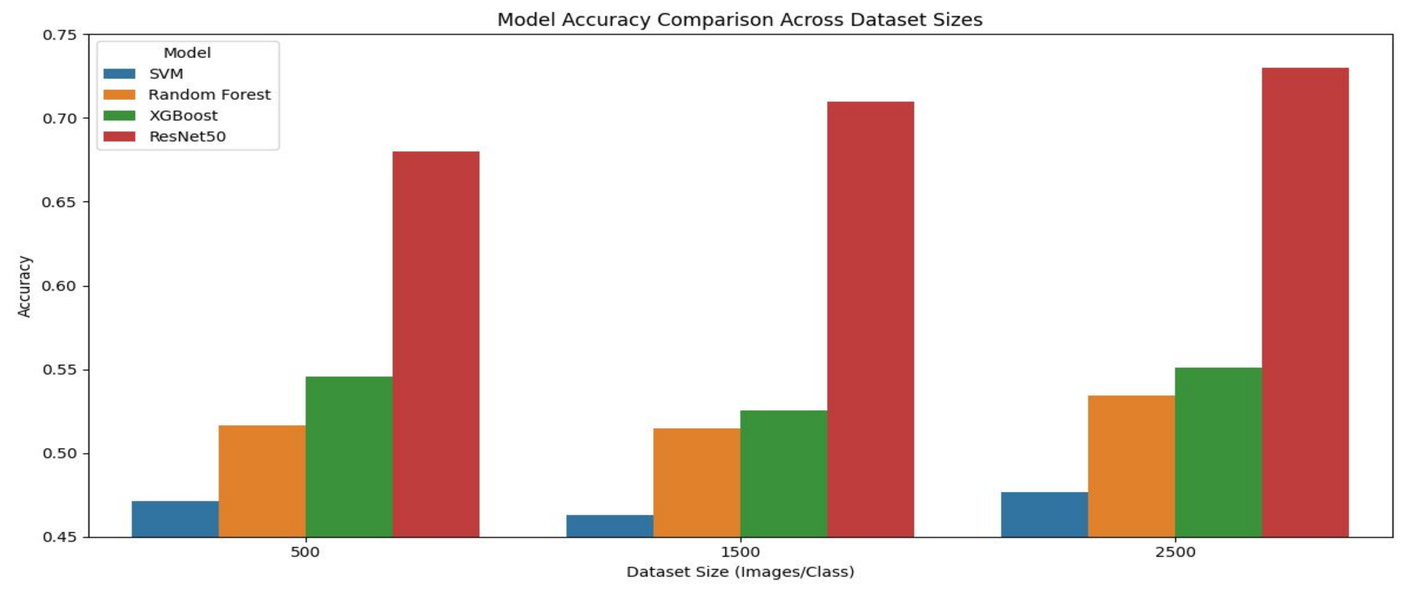

Figure 03 - Model accuracy comparison versus dataset size

- **Challenges**: Classes H1 and H2 had lower accuracy due to visual similarity.
- **Grad-CAM** confirmed DL models used biologically relevant features.

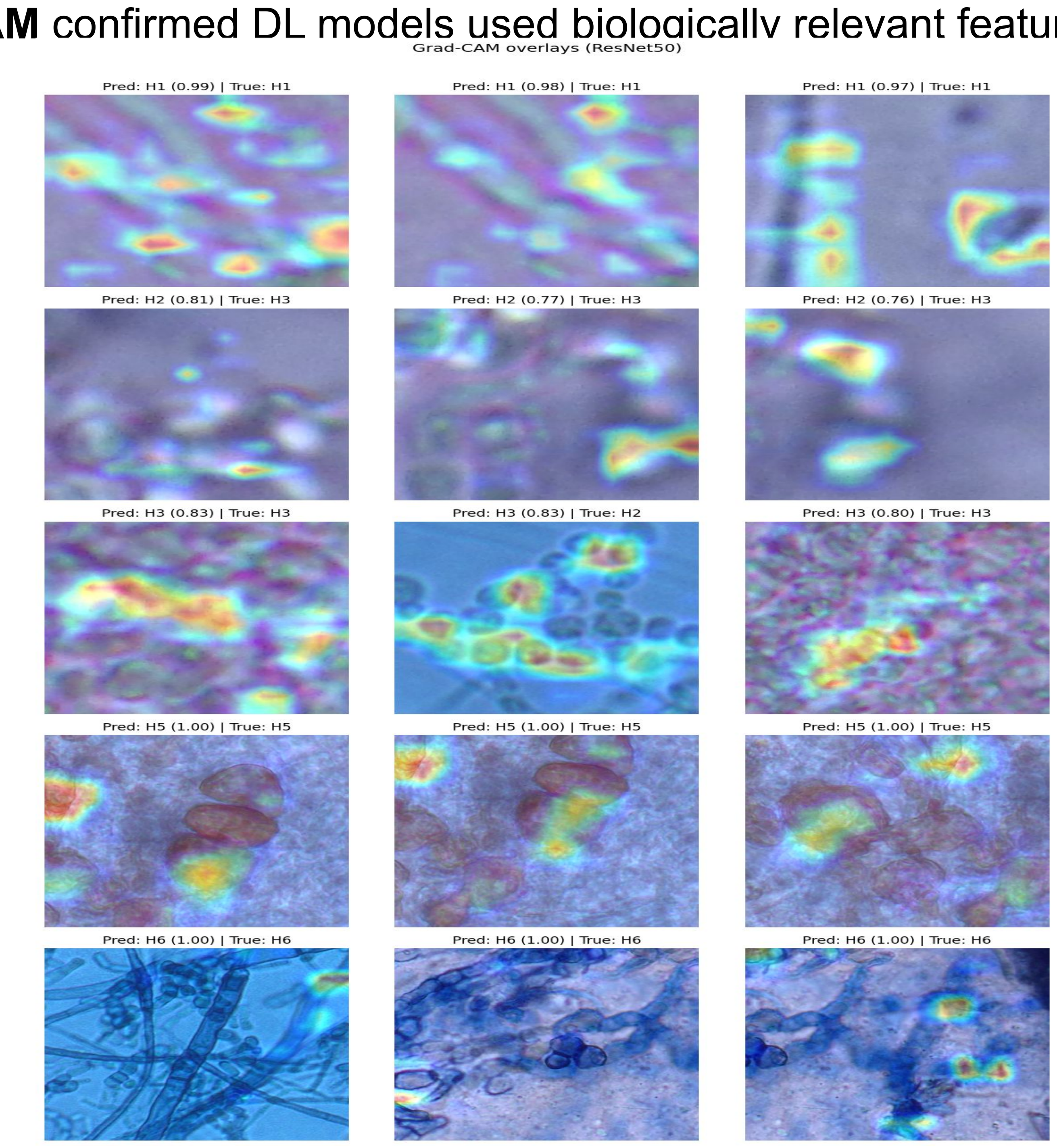

Figure 04 - Grad-CAM implementation on ResNet50

Conclusions

- ResNet50 outperforms ML models, especially with larger datasets.
- Dataset size significantly boosts DL performance; ML shows limited gains.
- Grad-CAM demonstrates interpretability, showing biologically relevant focus regions.
- Challenges remain for visually similar classes (H1 & H2).
- Future work: higher-quality datasets, advanced augmentation, and newer DL architectures (ResNet101, EfficientNet, ViTs, ensembles).
- Research supports development of reliable AI-driven diagnostic tools for clinical use.