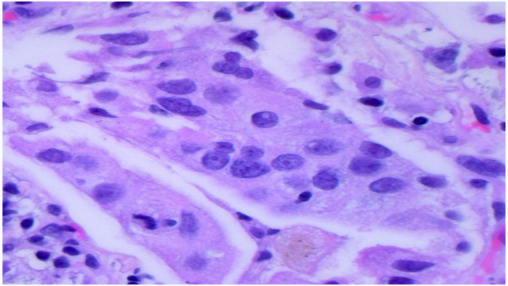
Visual Intelligence: Lung Cancer Histopathological Classification

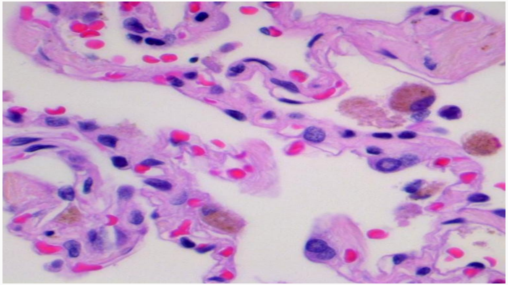
Lorenzo Mioso

March 2025

Introduction & Problem Statement

- **Dataset**: Lung cancer histopathological images (3 classes):
  - *Adenocarcinoma*
  - *Squamous cell carcinoma*
  - *Benign tissue*
- **Classification Task**: *Binary classification* (adenocarcinoma vs benign tissue)
- **Challenge**: Distinguishing *subtle tissue patterns* and *cellular structures*





Project Goals

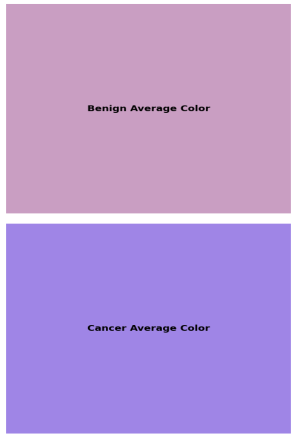
- Compare *traditional CNN* vs *ScatNet* approaches (in particular for feature extraction)
- Achieve *high accuracy* with *interpretable results*
- Apply *explainable AI techniques* to validate model decisions, in particular:
  - *Attribution analysis* with *guided backpropagation*
  - *Filter analysis* to understand learned patterns

Data Preprocessing & Setup

- *Dataset Organization*:
  - *K-fold cross-validation* with 10 folds
  - Class distributions already balanced
  - Image size kept at *768×768 pixels*
  - *Grayscale conversion* for analysis

Key Finding

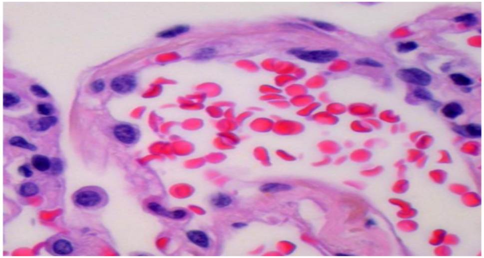
Using grayscale images for analysis because with only color images, the model achieved high accuracy by learning *color distributions*



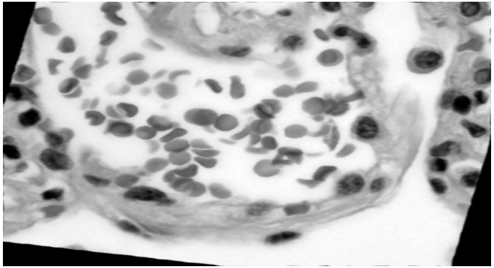
Preprocessing Pipeline

- Image *normalization* and *standardization* for each fold
- The dataset already includes augmentation, but I added some more:
  - *Random rotations*
  - *Random flips*
  - *Color jittering*
  - *Random cropping*
  - *Gaussian noise*

benign

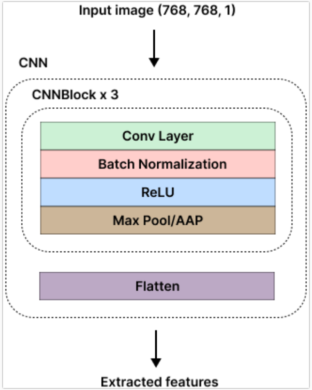


Augmented benign image



CNN Model Architecture

- *Efficient Feature Extraction Design*:
  - Similar to *ResNet architecture* without skip connections
  - Three progressive convolutional blocks (16→16→24 filters)
  - First convolutional layer with *11×11 kernel*
  - Strategic dimensionality reduction via *max pooling*
  - *Batch normalization* for improved training stability
  - *ReLU activations* for non-linear patterns

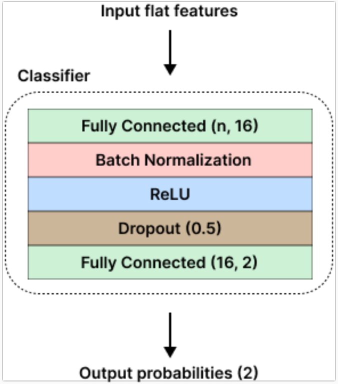


Key Findings

- Achieves *exceptional classification accuracy* despite minimal parameters
- First layer filters may not appear visually interpretable with small learning rates

Feature Classifier

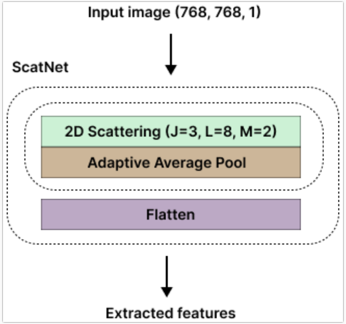
- Single hidden layer with 16 neurons
- *BatchNorm* for stability and faster convergence
- *ReLU activation* for non-linearity
- *Dropout (0.5)* for regularization
- Final layer with 2 output neurons (softmax activation)



ScatNet Model Architecture

- **Wavelet-based feature extraction**:
  - **J=3** scale parameter for wavelet decomposition
  - **L=8** orientations, **M=2** scattering order
  - **Translation**, **rotation**, and **scaling invariant**
  - Followed by global average pooling (4×4)
  - 3,472 features fed to the classifier

Key Finding
Requires **more complex classifier** layer for good performance

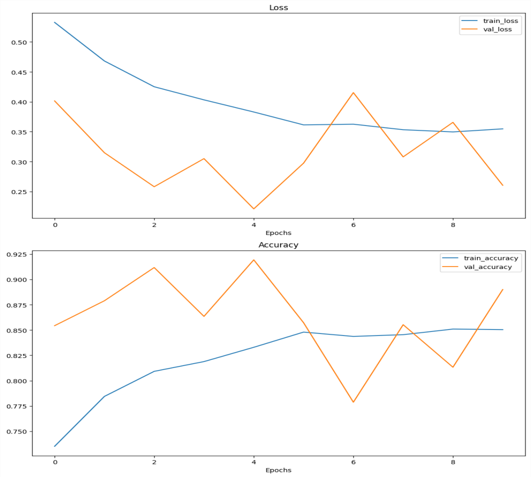


CNN Learning Curves Analysis

- **Rapid convergence** within 15 epochs
- **Consistent performance** across folds
- **Limited overfitting** due to effective regularization
- Final validation accuracy around **99%**

ScatNet Learning Curves Analysis

- **Slower convergence** but fewer epochs needed
- **More complex classifier** needed
- **Greater performance variation** across folds



Performance Results & Analysis

Metric	*CNN*	*ScatNet*
Mean Accuracy	*99.26%* $\pm$ *0.72%*	*92.99%* $\pm$ *1.59%*
Mean F1 Score	*99.27%* $\pm$ *0.72%*	*92.83%* $\pm$ *1.70%*
Accuracy Range	*97.70% - 99.90%*	*89.10% - 95.10%*
Training Speed	*Faster*	*Slower*

Key Performance Findings

- *CNN significantly outperforms* ScatNet (by *6.27%*)
- *K-fold validation* confirms *robust performance*
- *CNN* shows *less variance* between folds
- *CNN* achieves convergence in *fewer epochs*

CNN Performance Summary

Performance Range:

- Accuracy: **98.00% - 99.90%** across all folds
- F1 Score: **98.01% - 99.90%** across all folds
- Most consistent fold: **Fold 9** (99.90% accuracy)
- All folds achieved >**97.70% accuracy**

Overall Statistics:

- Mean Accuracy: **99.26%**
- Mean F1 Score: **99.27%**
- Standard Deviation: **0.72%**

ScatNet Performance Summary

Performance Range:

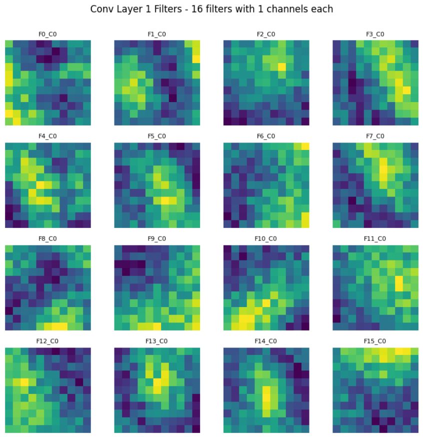
- Accuracy: **89.10%** - **95.10%** across all folds
- F1 Score: **88.68%** - **95.06%** across all folds
- Best performing fold: **Fold 6** (95.10% accuracy)
- Worst performing fold: **Fold 1** (89.10% accuracy)

Overall Statistics:

- Mean Accuracy: **92.99%**
- Mean F1 Score: **92.83%**
- Standard Deviation: **1.59%**

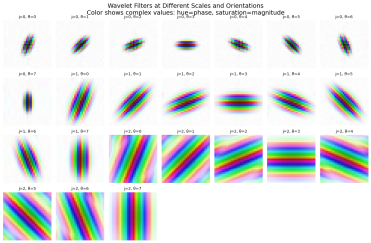
CNN Filter Analysis

- **Detected Filter Patterns**:
  - **Diagonal/Vertical Strips**: Detect edges, gaps, and transitions between textures
  - **Circular Points**: Identify spots, blobs, and localized features
  - **Two-Part Filters**: Capture gradients and contrast changes
- **Filter Implications**:
  - Learning localized and structured features
  - Capturing directional patterns and contrast
  - Detecting orientation-dependent structures

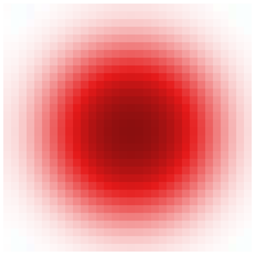


ScatNet Filter Analysis

- **Pre-defined wavelet transforms** (not learned)
- **Scale and rotation invariant** features
- **Lower discriminative power** despite theoretical advantages
- Fixed mathematical representation **limits adaptability**
- Data augmentation impact: **less significant**

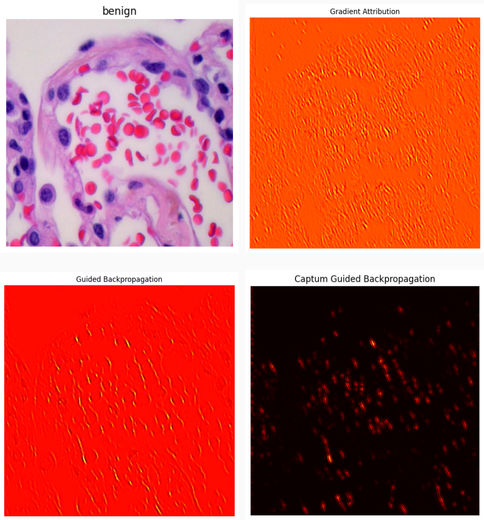


The corresponding low-pass filter, also known as scaling function.



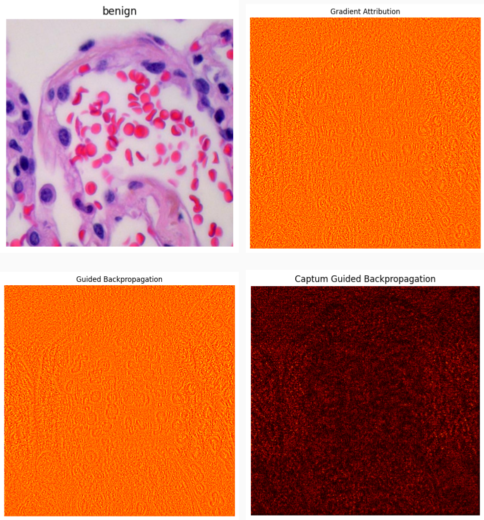
CNN Attribution Analysis

- Visualizes regions **most influential** for classification
- **Guided backpropagation** significantly reduces noise compared to regular backpropagation
- **Higher resolution** feature attribution with clearer patterns
- Captum implementation shows **similar patterns** despite different color scaling
- **Strong correlation** with pathological markers across all methods



ScatNet Attribution Analysis

- **Minimal differences** between guided and regular backpropagation
- Limited impact of guided backprop due to **ReLU only in classifier**
- **Non-recognizable patterns** in guided backprop visualization
- Attribution shows **diffuse, less interpretable regions**
- **Less coherent** with pathological indicators



Key Insights

- *Color features* are not crucial for lung cancer histopathology classification
- *Learned features* (CNN), in this scenario, are more effective than *pre-defined features* (ScatNet)
- *CNN* is harder to train but yields *better performance*

Thank you for your attention

Lorenzo Mioso