Lung Cancer Classification via Convolutional Neural Network Model

Ava Ambuske, Cameron Emamian, Alyssa Fricker, Cece Hagen, Ellie Strande, Jonathan Vergonio

Chapman University Grand Challenges Initiative Project

CHAPMAN UNIVERSITY Grand Challenges Initiative

CHAPMAN UNIVERSITY Fowler School of Engineering

CHAPMAN UNIVERSITY Schmid College of Science and Technology

ABSTRACT

- **Lung cancer is a formidable global health challenge**, necessitating precise lesion classification to optimize treatment strategies for early diagnosis.
- Our **CNN model** demonstrated **remarkable accuracy** in discriminating between **benign and malignant lesions (97%)**, and **strong accuracy** in classifying **different lung cancer types (76%)**.
- This research provides a foundation for future investigations and innovations aimed at **diagnostic methodologies and therapeutic interventions** in the pursuit of improved patient outcomes.

INTRODUCTION

Lung Cancer Background

- Overall five-year survival rate of lung cancer is only **15.6%**
- Survival rate is **70% for stage 1 lung cancer**, indicating early diagnosis is critical
- Causes include air pollution, radiation, and fungal infections

Image Analysis in Cancer Detection

- Facilitates personalized treatment strategies
- Enhances prognostic accuracy

Challenges/Problems in Image Recognition

1) Varied tumor morphologies and texture
2) Image noise and variability
3) Limited interpretability for complex patterns

Progress in Image Analysis Technology

1) Advancements in high-resolution imaging
2) Sophisticated computer vision techniques
3) Improved computational capabilities

THE CHALLENGE

Our AI-driven solution aims to confront the challenge of **human error and inaccurate diagnostics in lung cancer**.

By leveraging advanced algorithms, our system **enhances diagnostic precision,** reducing the potential for human-related discrepancies and ultimately improving the reliability of lung cancer assessments.

Additionally, these AI tools must be **intuitive and catered toward medical professionals** with a simple, yet effective website application.

OUR APPROACH

Neural Network Architecture

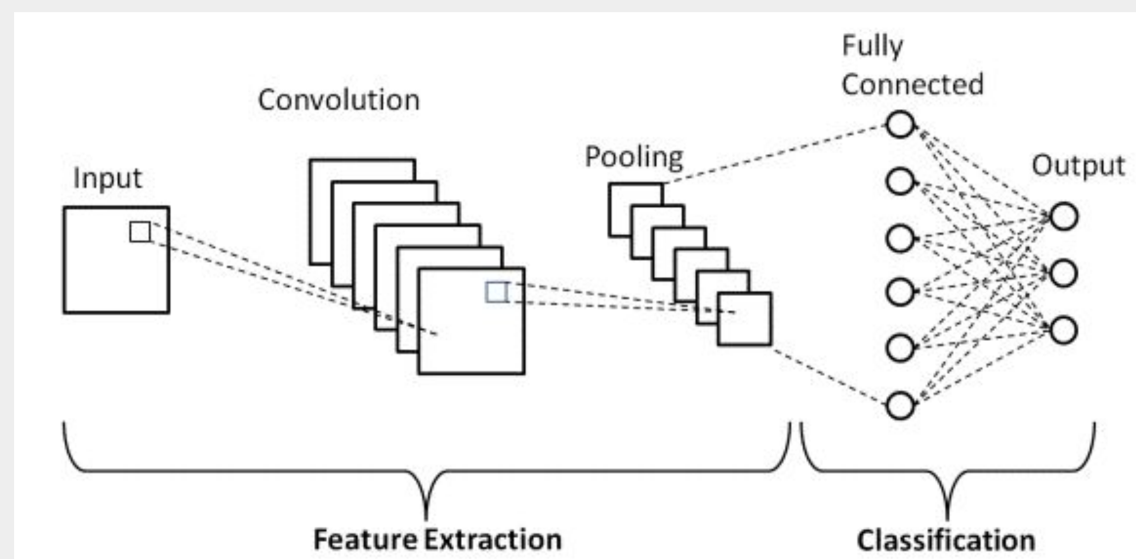


Figure 1. Convolutional Neural Network (CNN) - Deep Learning Architecture Framework

Convolutional Neural Network (CNN)

A widely recognized **computer vision technique** renowned for extracting **prognostic insights** from medical imaging scans, (i.e. CT scans).

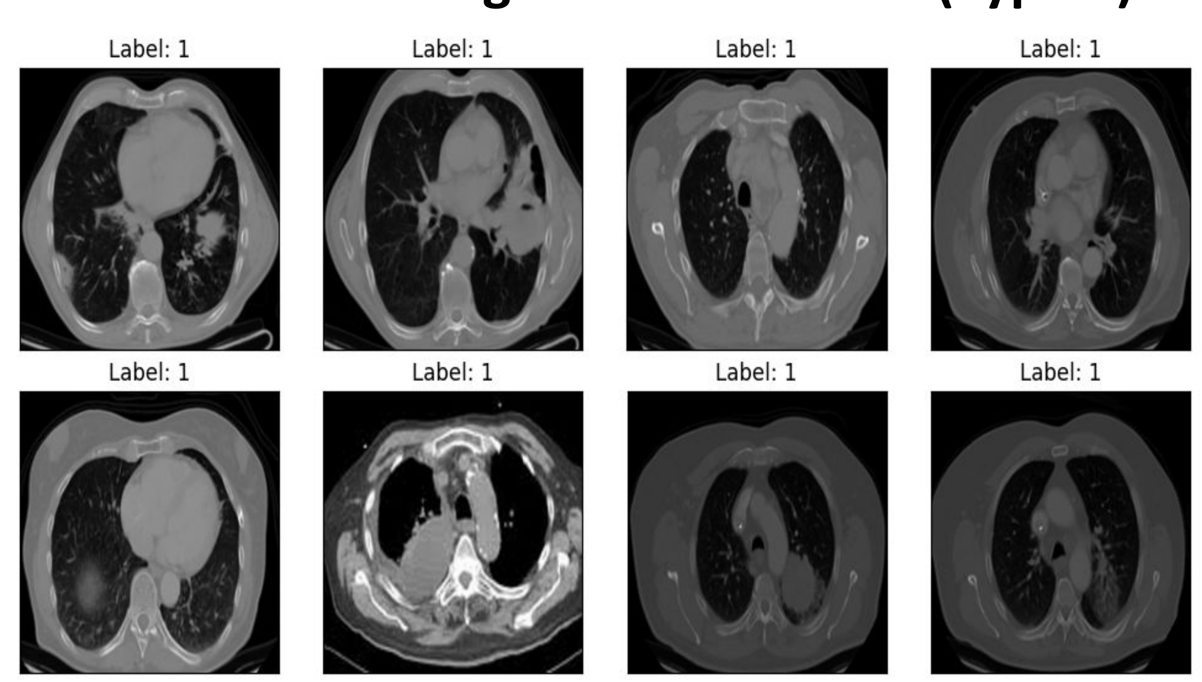
How do they work?

The **neural network layers** are strategically organized to initiate pattern recognition by detecting **simple elements** such as lines/ curves. As the network progresses, it captures **more intricate patterns, including nuanced representations of lung tissue**-a critical factor in effective cancer prognostics.

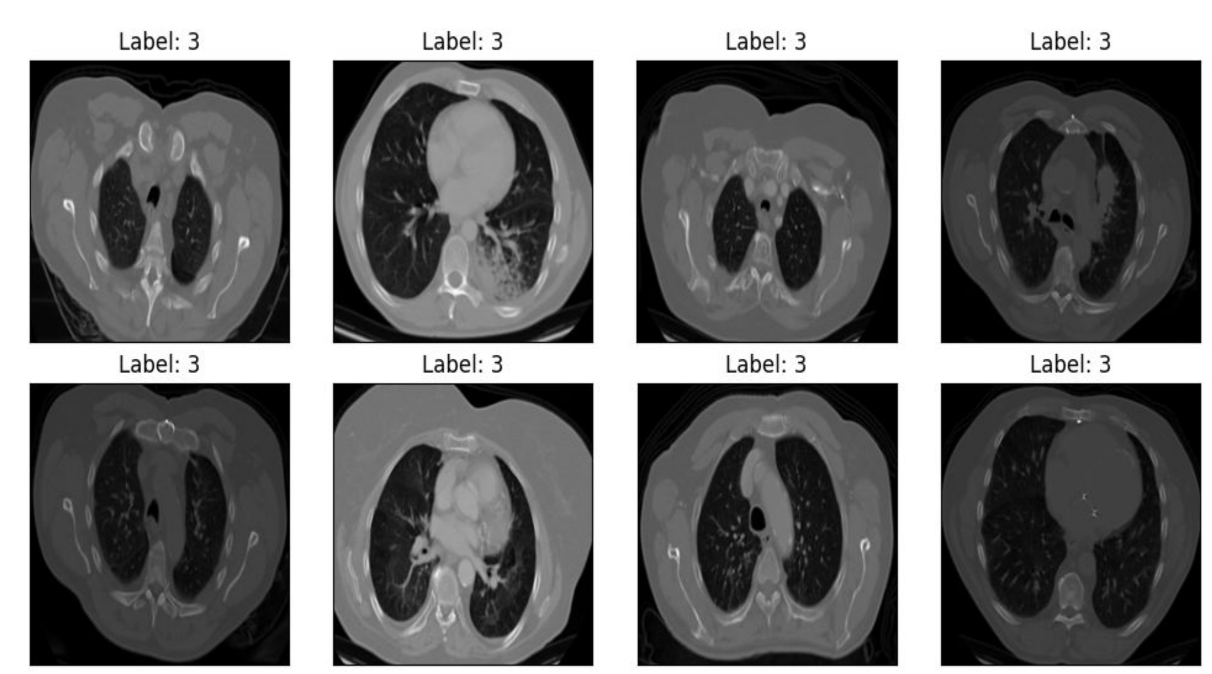
CT SCANS + DISCRIMINATION

Evaluating i) **Binary Classification Model**: Malignant vs. Benign ii) **Multiclass Classification Model:** Type 1 (Large Cell Carcinoma), Type 2 (Adenocarcinoma), and Type 3 (Squamous Cell Carcinoma) Lung Cancer

CT Scans of Large Cell Carcinoma (Type 1)



CT Scans of Squamous Cell Carcinoma (Type 3)



Binary Classification Results

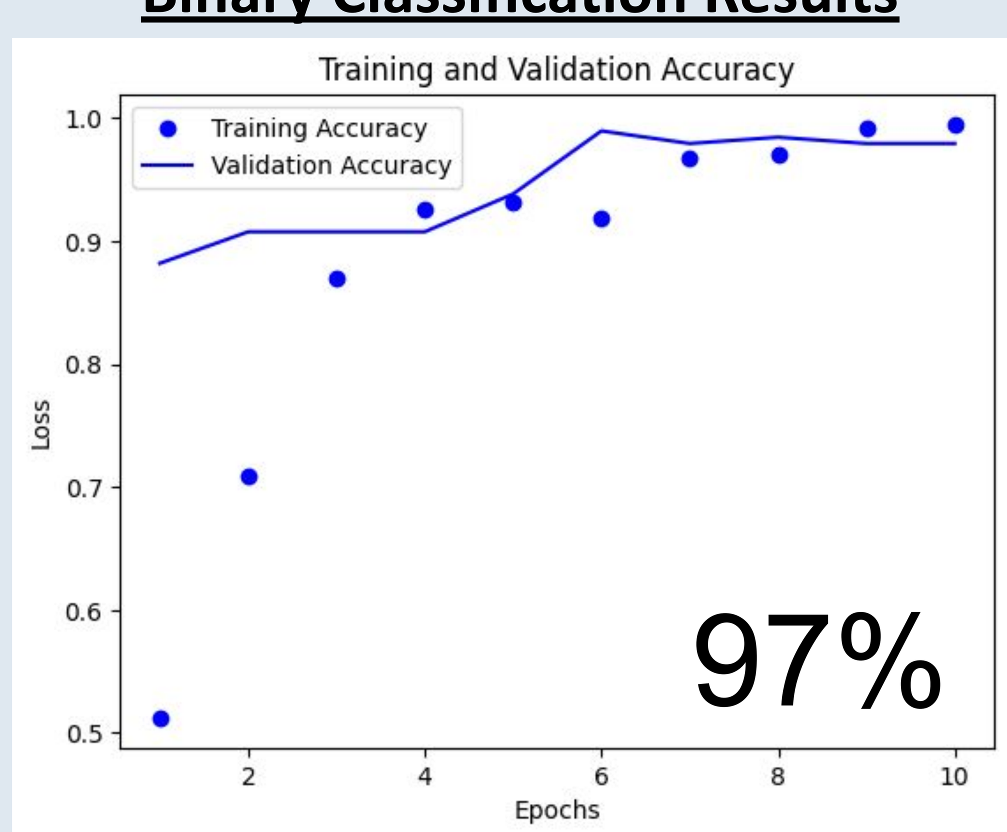


Figure 2. Graph of Malignant and Benign Cancer Accuracy and Model Discrimination

Multiclass Classification Results

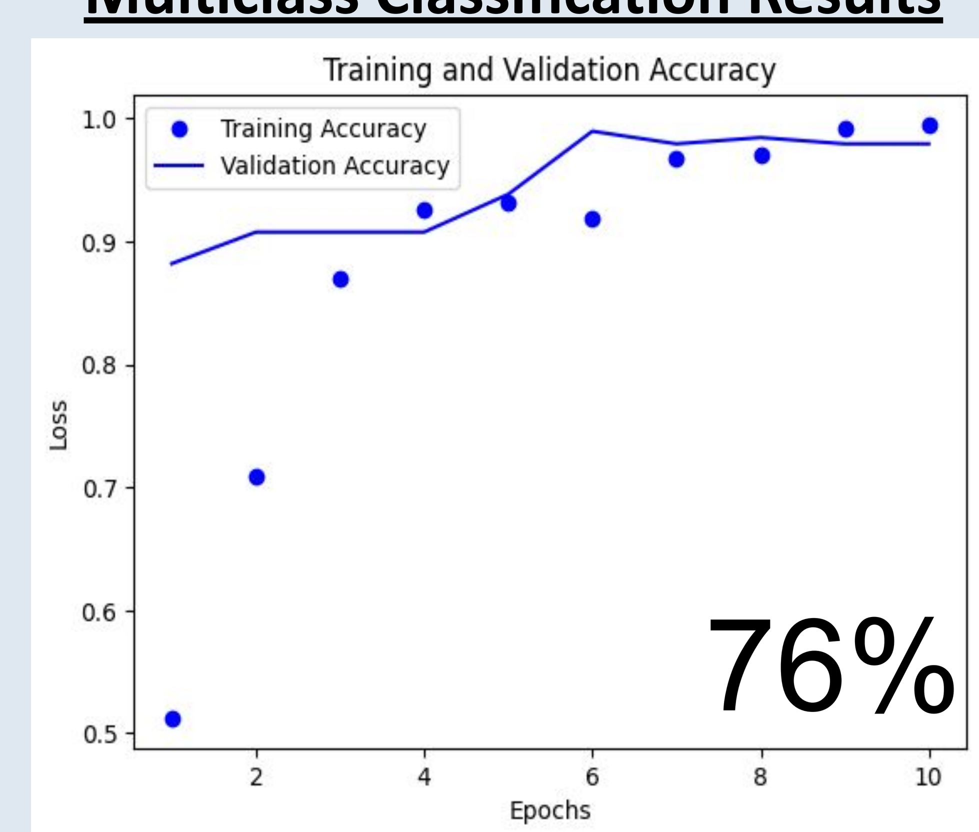
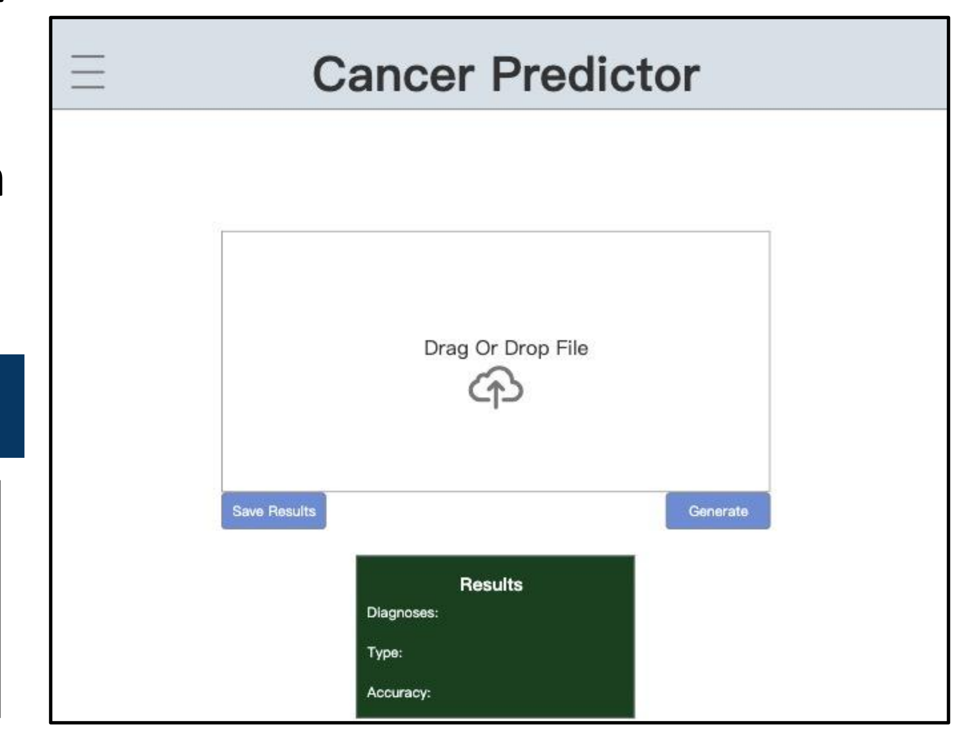


Figure 3. Graph of Type 1, Type 2, and Type 3 Lung Cancer Accuracy and Model Discrimination

METHODOLOGY

Data Collection Preprocessing: Histopathology Images from Kaggle → **Deep Learning Algorithm**: Implemented via TensorFlow and Keras Python libraries → **Model Training and Testing**: Evaluation of performance on unseen data

WEBSITE APPLICATION



To **enhance accessibility** for medical professionals, we developed a dedicated website. This platform allows uploads of CT Scans and **provides real-time analysis and classification results.**

Figure 4. Web Application Mock Demo utilizing HTML, CSS, Python, and Mockplus

CONCLUSIONS

- **Binary Classification:** Successfully developed a CNN model achieving a 97% accuracy in discriminating between benign and malignant lung cancer tumors.
- **Multiclass Classification:** Moderately-high accuracy (76%) in the classification of large cell carcinoma, adeno-carcinoma, and squamous cell carcinoma cancer types.

FUTURE PLANS

After meeting with a **medical professional in Computational and Data Sciences at CHOC**, our next steps are:

1) Focus on keeping neural network simple to perform best.
2) Investigate other performance metrics. (ROC Curve, etc.)
3) Implement a neural network to detect clinical biomarkers.

ACKNOWLEDGEMENTS + LITERATURE

This research project was supported by **Chapman University's Grand Challenge Initiatives (GCI) Program**. We would also like to acknowledge **Dr. Louis Ehwerhemuepha (CHOC), Dr. Javier Espeleta (GCI), and Dr. Indrajit Sen (GCI)** for providing professional advice and feedback on our project.

Liu, M., et al. (2023). The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis.