**COMSATS University Islamabad,   
Abbottabad Campus**

**Project Proposal   
(SCOPE DOCUMENT)**

**for**

**Liver Tumor Segmentation in CT Scan images Using Light Weight Deep Learning Model**  
Version 1.0

***By***

**Faizan CIIT/FA21-BSE-011/ATD**

**Fawad Iqbal CIIT/FA21-BSE -012/ATD**

***Supervisor*Dr. Mubashir Ahmad**

***Bachelor of Science in Computer Science (2024-2025)***

|  |  |  |
| --- | --- | --- |
| **No.** | **Comment** | **Action** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**SCOPE DOCUMENT REVSION HISTORY**

**Supervisor Signature**

**Date:**

**Table of Contents**

Contents

[Abstract 6](#_Toc181677394)

[. 6](#_Toc181677395)

[Introduction 7](#_Toc181677396)

[Problem Statement 7](#_Toc181677397)

[Problem Solution for Proposed System 8](#_Toc181677398)

[Scope 8](#_Toc181677399)

[Modules 8](#_Toc181677400)

[1. User Interface (UI) Module 8](#_Toc181677401)

[2. Image Processing Module 8](#_Toc181677402)

[3. Segmentation Module 9](#_Toc181677403)

[4. Analysis & Metrics Module 9](#_Toc181677404)

[System Limitations/Constraints 9](#_Toc181677405)

[Software Process Methodology 10](#_Toc181677406)

[Tools and Technologies 10](#_Toc181677407)

[Project Stakeholders and Roles 10](#_Toc181677408)

[Team Members Individual Tasks/Work Division 11](#_Toc181677409)

[Data Gathering Approach 11](#_Toc181677410)

[Concepts 11](#_Toc181677411)

[Gantt chart 12](#_Toc181677412)

[Mockups 13](#_Toc181677413)

[Conclusion 15](#_Toc181677414)

[References 15](#_Toc181677415)

[Plagiarism Report 16](#_Toc181677416)

**Project Category: (**Select all the major domains of proposed project**)**

* **A-**Desktop Application/Information System **B-**Web Application/Web Application based Information System **C-** Problem Solving and Artificial Intelligence ** D-**Simulation and Modeling ** E-** Smartphone Application ** F-** Smartphone Game ** G-** Networks ** H-** Image Processing****Other (specify category) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This research aims to improve liver cancer diagnosis through automated segmentation of Liver and tumor in 3D CT images using the nnFormer architecture. By integrating convolutional layers with attention mechanisms and processing images as smaller patches, we expect to enhance the accuracy and reduce the computational complexity.

.

# Introduction

Liver cancer is the most common cause of death all around the world, past statistics shows that over 800,000 deaths have been caused by liver cancer [1, 2]. This is a serious threat to public health, which increases the demand for liver cancer diagnosis. This process of liver cancer diagnosis involves the accurate prediction of the tumor from the CT (Computed Tomography) images, which provides the facilitation of liver cancer treatment [4, 5].

However, manually diagnosing the liver tumor requires physician’s experience and is also a time-consuming process. It is also a challenge to figure out the sharpness of the edges of the liver with naked eyes. This emerges the demand of automating the liver tumor segmentation process [15].

The automation process of liver tumor segmentation brings along several challenges. 1) The shape of liver is affected by the presence of neighbouring organs, as the liver is a soft tissue. 2) Visualization is quite difficult because of the blurred boundaries with neighbouring organs, especially with organs such as spleen or stomach. 3) Diagnostics of lesion can affect shape and appearance of the liver [10]. To overcome the problems, we propose an automatic Liver Tumor Segmentation in CT image.

Our task is to segment the Liver from the CT scans and then segment the tumor out of the Liver segment. We propose liver tumor segmentation using nnFormer Architecture [12]. nnFormer Architecture uses a hybrid stem where convolution layer and self-attention are interlinked to exploit their ability. A light-weight convolutional embedding layer is put ahead of transformer blocks which extracts features whereas the transformer encodes the global context, with convolutional layer we encode pixel-level spatial information, more precisely compared to patch based positional encoding. This architecture also uses small kernel sizes for the convolutional layer to reduce computational complexity [12].

# Problem Statement

Medical image segmentation plays an important role in accurately diagnosing and treating liver diseases. Traditional methods, like CNNs (Convolutional Neural Networks), mostly struggle in capturing detailed information within 3D medical images because they rely on local features [12]. Other methods like TransUNet (Transformer UNet), which is a hybrid CNN-Transformer architecture to gain both high-resolution spatial information from the extracted features using CNN, and the global context encoded by Transformers [13]. TransUNet relies on CNNs, which means that the benefits of transformers are not fully exploited, the model also uses one or two layers of transformers which are not enough to capture long-range dependencies. Researchers started using transformers as the main stem for medical image segmentation. SwinUNet improves upon TransUNet by using transformers as the main part of its architecture, Swin-UNet is the first transformer based UNet model for 2D medical image segmentation. It splits input images into patches, treating each patch like a “token”. These tokens are passed through an encoder to learn features, and then a decoder up-samples to restore the image resolution and predict the segmentation [14, 12]. However, they did not appropriately combine convolution and self-attention which can improve this task [12].

# Problem Solution for Proposed System

The nnFormer model overcomes these challenges by using a hybrid approach that interlinks convolution and transformer operations. This model effectively captures both local and global information, leading to improved segmentation results. it uses two types of attention mechanisms, Local Volume-based Multi-head Self-attention (LV-MSA) and Global volume-based Multi-head Self-attention (GV-MSA) to create feature layers that help represent liver tissues in much more detail [12].

This research aims to implement nnFormer for liver tumor segmentation in 3D medical images using Lits17 database, with the intention of comparing its performance against existing architectures like CNN based UNet, TransUNet, and SwinUNet. The goal is to demonstrate that nnFormer achieves better segmentation accuracy to contribute to better diagnostic decisions in Liver disease cases [12].

# Scope

This research focuses on developing an automated liver and tumor segmentation using the nnFormer architecture to improve diagnostic accuracy for liver cancer. The system will process 3D CT images converted to 2D slices from Lits17 database, segmenting the liver and any tumors within it. By using nnFormer architecture’s, hybrid approach—combining convolutional layers and transformer-based-self-attention mechanisms—this study aims to overcome current challenges in accurately detecting liver tumors, such as blurred boundaries and complex shapes. The performance of nnFormer will be compared with existing models, including CNN-based UNet, TransUNet, and SwinUNet, to ensure its effectiveness in liver tumor segmentation.

# Modules

## User Interface (UI) Module

* **Functionality:** Create user-friendly interface for users to upload CT scan image and view segmented results.
* **Components:**
  + **Image Upload Interface:** A drag and drop or file upload component to accept 2D CT scans.
  + **Result Display:** Visual display of original and segmented images.

## Image Processing Module

* **Functionality:** Preprocess uploaded images to prepare them for input into nnFormer model.
* **Components:**
  + **Thresholding:** Applying standard thresholds value for extracting Liver.
  + **Normalization:** Normalize the image data to a suitable format for model input.
  + **Resizing:** Resize to match the input size required by nnFormer architecture.
  + **Augmentation:** Implement techniques such as rotation, flipping, and scaling to enhance the dataset and robustness of the model.

## Segmentation Module

* **Functionality:** Implement the nnFormer architecture for performing segmentation on the preprocessed images.
* **Components**
  + **Encoder:** Convolution for down sampling to capture local features and Local self-attention.
  + **Bottleneck:** Global self-attention to capture global features across the image.
  + **Decoder:** Up-sampling with local self-attention and skip connections to restore spatial detail.
  + **Output Generation:** Segmentation mask generation for liver and tumor identification, based on feature maps from the decoder.

## Analysis & Metrics Module

* **Functionality**: Provides insights on the segmented images, including tumor size and other key statistics for medical assessment.
* **Components**:
  + **Tumor Size Calculation**: Estimates the tumor volume from segmented images.
  + **Statistical Analysis**: Generates metrics like segmentation confidence, tumor area percentage, and location within the liver.
  + **Visualization**: Graphs and charts to help users understand the segmentation results.

# System Limitations/Constraints

1. **Limited Computational Resources for 3D Processing:** While 3D CT images provide valuable depth and detail for liver tumor segmentation, processing them requires high computational power and memory. Due to limited resources, the system converts 3D CT images into 2D slices, which may lead to a loss of depth information, impacting segmentation accuracy.

# Software Process Methodology

We’re using the **Object-Oriented Methodology (OOM)** for developing our project. This approach organizes our system around "objects" that represent key parts like images, segmentation results, and analysis metrics. OOM helps us keep our code modular and easy to manage, as each part of the system can be built separately. This methodology is a good choice because it makes our code reusable and easier to expand, which suits the structure and needs of our project well.

# Tools and Technologies

|  |  |  |  |
| --- | --- | --- | --- |
| **Tools**  **And**  **Technologies** | **Tools & Technologies** | **Version** | **Rationale** |
| Python | 3.12.4 | Main programming language for the project |
| PyTorch | 2.5.1 | Deep learning framework to implement nnFormer |
| RStudio | 2024.09.0 +375 | Data analysis and visualization |
| Matplotlib (Python) | Latest | Visualization libraries for displaying results. |
| NumPy, Pandas | Latest | Data manipulation and preprocessing |
| Anaconda | 24.5.0 | Python package manager |
| PyCharm | Community 2023.3 | IDE for code development and debugging |
| MS Power Point | 2015 | Presentation |
| Flask | 2.1.3 | Creating web application |
| Lits17 Dataset | 2017 | Medical dataset for liver tumor segmentation |
| Streamlit | Latest | Visualizing the output images |

Table 1: Tools and Technologies for the proposed system

# Project Stakeholders and Roles

|  |  |
| --- | --- |
| **Project Sponsor** | COMSATS University Islamabad, Abbottabad Campus. |
| **Students** | Fawad Iqbal – Responsible for data preprocessing, implementation of the architecture, and performance evaluation.  Faizan – Responsible for developing the nnFormer model and conducting experiments for liver tumor segmentation. |
| **Project Supervisor** | Dr. Mubashir Ahmad – Guides the project, providing mentorship and ensuring academics are met. Reviews the progress and gives feedback on development. |
| **Final Year Project Committee** | 1. Dr. Mubashir Ahmad 2. Mukhtiar Zamin 3. Ehzaz Mustafa 4. Bushra Mushtaq |

Table 2: Project Stakeholders for Proposed Project

# Team Members Individual Tasks/Work Division

**Table 4Team Member Work Division for Proposed Project**

|  |  |  |
| --- | --- | --- |
| **Student Name** | **Student Registration Number** | **Responsibility/ Modules** |
| Faizan | FA21-BSE-011 | 1. User Interface (UI) Module 2. Image Processing Module |
| Fawad Iqbal | FA21-BSE-012 | 1. Segmentation Module 2. Analysis & Metrics Module |

# Data Gathering Approach

We are using the LiTS17 dataset, which contains liver CT scan images for identifying and marking liver tumor. This data helps us train our model to accurately detect and segment liver tumor.

# Concepts

**1. Deep Learning for Medical Imaging**  
We'll learn how to apply deep learning, especially for processing medical images. This means understanding how neural networks like nnFormer can help detect liver and tumor regions in CT scans.

**2. Convolutional Neural Networks (CNNs)**  
CNNs are key for working with images in our project. We’ll learn about how CNNs recognize patterns in images, which is useful for accurately identifying areas in the CT scans.

**3. Attention Mechanisms**  
We’ll study attention mechanisms to make our model more precise. These allow the model to focus on specific parts of the image, which helps it better spot tumors within the liver tissue.

**4. Image Preprocessing**  
We’ll learn image preprocessing techniques like resizing, normalizing, and adding variations to make our model more reliable. This step prepares the CT images so that our model can work with them correctly.

**5. Data Visualization**  
Finally, we’ll learn to display our results in a clear way using tools like Matplotlib and Streamlit. This helps make the output easy to understand, so we can effectively show where the tumor is and how much of the liver it affects.

# Gantt chart

A grid of lines with small black and white dots

Description automatically generated with medium confidenceA close-up of a computer screen

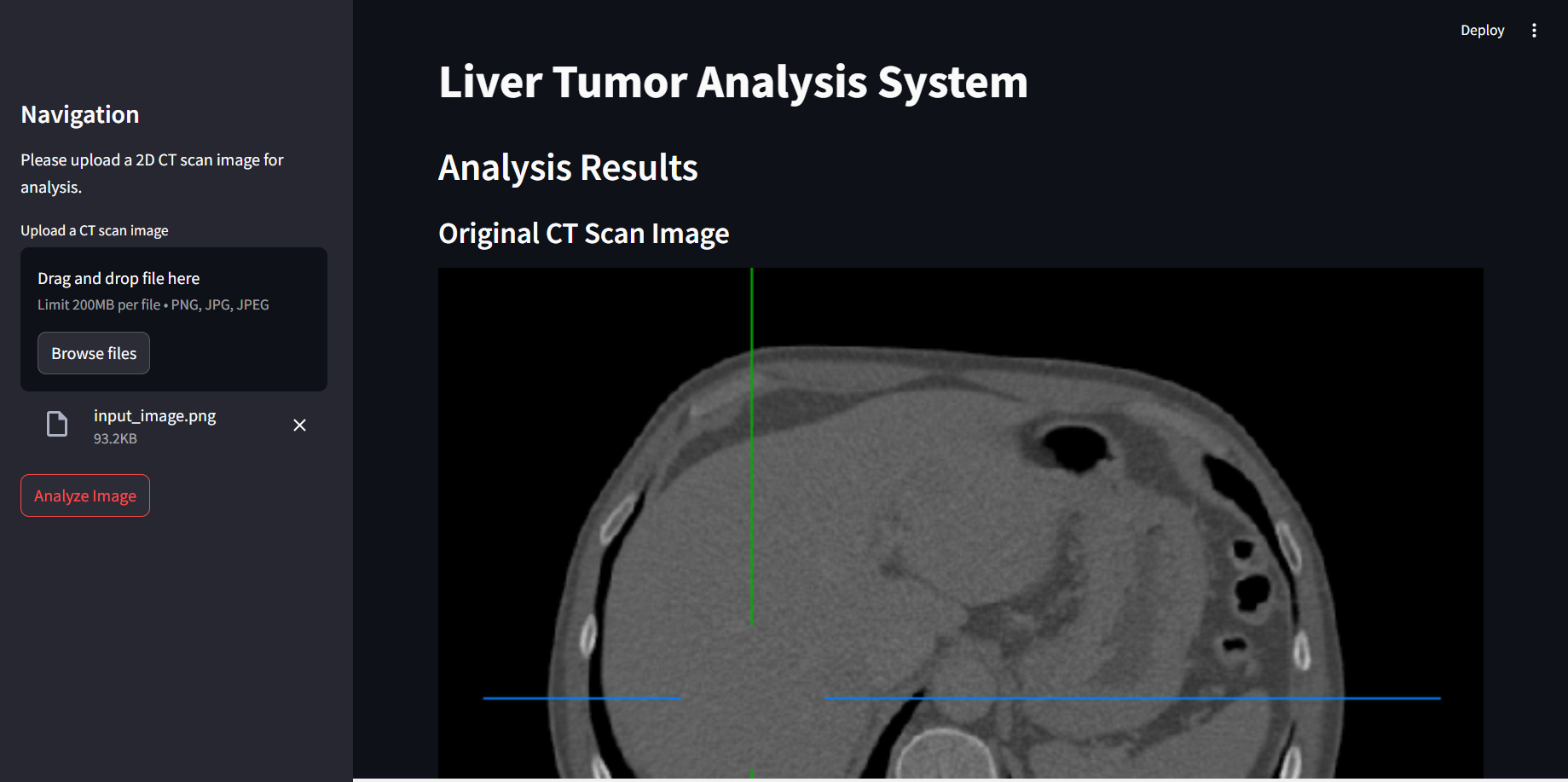
Description automatically generated

A grid of white squares

Description automatically generated

Figure 1 Gantt chart

# Mockups



A close-up of a ct scan

Description automatically generated

A screenshot of a computer screen

Description automatically generated

**A screenshot of a graph

Description automatically generated**

# Conclusion

In conclusion, our research is based on advanced deep learning techniques, specifically the nnFormer architecture, to enhance the accuracy and efficiency of liver tumor segmentation in 2D CT scans.

Our system aims to assist medical professionals to diagnose liver tumors with better precision and less time, providing an accessible and user-friendly interface.

# References

[1] Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, et al. Global Burden of 5 major types of gastrointestinal cancer. Gastroenterology.

[2] Chon YE, Park SY, Hong HP, Son D, Lee J, Yoon E, et al. Hepatocellular carcinoma incidence is decreasing in Korea but increasing in the very elderly. Clin Mol Hepatol.

[3] C. Lago-Hernandez, N.H. Nguyen, R. Khera, et al., Cost-related nonadherence to medications among US adults with chronic liver diseases, Mayo Clinic Proceedings. Elsevier 96 (10) (2021) 2639–2650.

[4] L. Liu, F.X. Wu, Y.P. Wang, et al., Multi-receptive-Field CNN for semantic segmentation of medical images, IEEE J. Biomed. Health Inform. 24 (11) (2020) 3215–3225.

[5] L. Yu, X. Yang, H. Chen, et al. Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI’17). AAAI Press, 2017.

[6] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of 2005 IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

[7] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015.

[8] O. Ronneberger, P.Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In Proc. MICCAI, volume 9351 of LNCS, pages 234–241, 2015.

[9] EG-UNETR: An edge-guided liver tumor segmentation network based on cross-level interactive transformer Dongxu Cheng \* , Zifang Zhou , Jingwen Zhang School of Mathematics and Information Science, Zhongyuan University of Technology, Zhengzhou, Henan 450007, China.

[10] Liver segmentation from computed tomography images using cascade deep learning

[11] UNETR: Transformers for 3D Medical Image Segmentation Ali Hatamizadeh NVIDIA Yucheng Tang Vanderbilt University Vishwesh Nath NVIDIA Dong Yang NVIDIA Andriy Myronenko NVIDIA Bennett Landman Vanderbilt University Holger R. Roth NVIDIA Daguang Xu NVIDIA

[12] nnFormer: Volumetric Medical Image Segmentation via a 3D Transformer HongYu Zhou, Student Member, IEEE, Jiansen Guo, Yinghao Zhang, Xiaoguang Han, Lequan Yu, Liansheng Wang, Member, IEEE, and Yizhou Yu, Fellow, IEEE. 7

[13] TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation Jieneng Chen 1 , Yongyi Lu 1 , Qihang Yu 1 , Xiangde Luo 2 , Ehsan Adeli 3 , Yan Wang 4 , Le Lu 5 , Alan L. Yuille 1 , and Yuyin Zhou 3 1 Johns Hopkins University 2University of Electronic Science and Technology of China 3Stanford University 4 East China Normal University 5PAII Inc.

[14] Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation Hu Cao 1 † , Yueyue Wang 2 † , Joy Chen 1 , Dongsheng Jiang 3 ∗ , Xiaopeng Zhang 3 ∗ , Qi Tian 3 ∗ , and Manning Wang 2 1 Technische Universit¨at M¨unchen, M¨unchen, Germany 2 Fudan University, Shanghai, China 3 Huawei Technologies, Shanghai, China

# Plagiarism Report

Attach the Plagiarism report of your project scope document from library staff of turnitin tool (http://turnitin.com