Camouflaged Military Soldier Detection Using Dense Deconvolutional Networks

Mini Project-I report submitted in partial fulfillment of the Requirements for the Award of the Degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

by

JAMPANI L N K SURYA SRI CHARAN	218W1A0586
BOMMANA PHANI PRAHARSHA	228W5A0507
EDA GAYATRI	228W5A0508
THOTAKURA CHANAKYA	218W1A05C1

Under the Guidance of

Mr. SRIKAKULAPU SUBHAKAR. U, M.Tech.

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VELAGAPUDI RAMAKRISHNA

SIDDHARTHA ENGINEERING COLLEGE

Autonomous and Approved by AICTE, NAAC A+, NBA Accredited

Affiliated to Jawaharlal Nehru Technological University, Kakinada

Vijayawada 520 007

May 2024

VELAGAPUDI RAMAKRISHNA
SIDDHARTHA ENGINEERING COLLEGE

(Autonomous, Accredited with 'A+' grade by NAAC)

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Major Project Report entitled **"Camouflaged Military Soldier Detection Using Dense Deconvolutional Networks"** being submitted by

JAMPANI L N K SURYA SRI CHARAN	208W1A05I3
BOMMANA PHANI PRAHARSHA	228W5A0507
EDA GAYATHRI	228W5A0508
THOTAKURA CHANAKYA	218W1A05C1

in partial fulfillment for the award of the Degree of Bachelor of Technology in Computer Science and Engineering to the Jawaharlal Nehru Technological University, Kakinada is a record of bonafide work carried out during the period from 2023 - 2024.

Mr. Srikakulapu Subhakar, M.Tech
Assistant Professor & Guide

Dr.D.Rajeswara Rao, M.Tech, Ph.D
Professor & HOD,CSE

DECLARATION

We hereby declare that the Major Project entitled **"Camouflaged Military Soldier Detection Using Dense Deconvolutional Networks"** submitted for the B.Tech Degree is our original work and the dissertation has not formed the basis for the award of any degree, associateship, fellowship or any other similar titles.

Place: Vijayawada **JAMPANI L N K SURYA SRI CHARAN (218W1A0586)**

Date: **BOMMANA PHANI PRAHARSHA (228W5A0507)**

EDA GAYATRI (228W5A0508)

THOTAKURA CHANAKYA (218W1A05C1)

ACKNOWLEDGEMENT

We would like to thank **Dr. A. V. Ratna Prasad**, Principal of Velagapudi Ramakrishna Siddhartha Engineering College for the facilities provided during the course of Major Project.

We have been bestowed with the privilege of thanking **Dr. D. Rajeswara Rao**, Professor and Head of the Department for his moral and material support.

We would like to express our deep gratitude to our guide **Mr. Srikakulapu Subhakar**, Assistant Professor for his persisting encouragement, everlasting patience, and keen interest in discussion and for his numerous suggestions which we had at every phase of this project.

We owe our acknowledgments to an equally long list of people who helped us with Major Project work.

Place: Vijayawada **JAMPANI L N K SURYA SRI CHARAN (218W1A0586)**

Date: **BOMMANA PHANI PRAHARSHA (228W5A0507)**

EDA GAYATRI (228W5A0508)

THOTAKURA CHANAKYA (218W1A05C1)

Table of Contents

List of Tables

List of Figures

Abstract

Military camouflage is a set of techniques and patterns used by armed forces to hide personnel, equipment, and installations from observation by enemy forces. The primary objectives of military camouflage are to avoid detection, to deceive or confuse adversaries about the location or intentions of military assets, and to enhance survivability on the battlefield. This project aims to address the challenge of detecting camouflaged military personnel using deep learning techniques, through the utilization of Dense Deconvolution Networks (DDNs), an advanced variant of convolutional neural networks (CNNs). The project entails the collection of a diverse dataset comprising images containing camouflaged soldiers against various backgrounds. The DDN architecture is designed to effectively capture spatial hierarchies and feature representations within the input images, enabling the identification of camouflaged individuals with high precision through Binary segmentation. By using dense connections and deconvolutional layers, the DDN model can efficiently reconstruct spatial information lost during down-sampling operations, thus enhancing the detection outcomes.

Keywords: Dense Deconvolution Networks (DDN), Binary Segmentation, Upsampling, Feature Extraction, Object Detection.

Chapter 1

INTRODUCTION

Military camouflage offers soldiers a critical advantage by concealing them from enemy forces. However, current detection methods might not be entirely sufficient. This project tackles this challenge by proposing a deep learning approach for automatic detection of camouflaged personnel. The project leverages Dense Deconvolution Networks (DDNs), a powerful variant of Convolutional Neural Networks (CNNs), to analyze images. By collecting a diverse dataset of camouflaged soldiers in various environments, the DDN model will learn to capture key features and spatial relationships within the images. Notably, the DDN architecture utilizes dense connections and deconvolutional layers to recover information lost during processing, leading to more accurate detection. This project has the potential to significantly improve military operations and surveillance by achieving highprecision detection of camouflaged soldiers through deep learning techniques[1].

1.1 Basic Concepts

1.1.1 Data Preprocessing (Image Resizing)

The project likely involves preprocessing the dataset of soldier images before feeding them into the DDN model. This preprocessing might include resizing the images to a standard size. This ensures consistency for the model and allows it to focus on the key features of the soldiers rather than variations in background image dimensions.

1.1.2 DownSampling

This reduces the image resolution, ensuring all images have a consistent size for efficient processing by the DDN model. This allows the model to focus on identifying the camouflaged figures themselves, rather than being distracted by variations in background image sizes.

1.1.3 Upsampling

DDNs potentially mitigate a related issue. Down-sampling can sometimes lose spatial information within images. DDNs, through their use of deconvolutional layers, can help reconstruct this lost information, leading to more accurate detection of camouflaged soldiers.

1.1.4 Binary Segmentation

By performing this classification for every pixel, the DDN model can effectively create a ”mask” highlighting the camouflaged individuals within the image. This binary representation simplifies the detection process and allows for accurate identification of soldiers even when their camouflage blends in with the surroundings.

1.1.5 Feature map to Image conversion

The DDN architecture likely utilizes convolutional layers to extract features from the soldier images. These features are represented as internal data structures called ”feature maps.” While the project doesn’t mention directly converting these feature maps back to images, the DDN’s deconvolutional layers likely play a crucial role in utilizing this information.

1.1.6 IoU

IoU measures the overlap between the predicted ”mask” (areas classified as camouflaged soldier) and the ground truth (actual location of the soldier in the image). A higher IoU score indicates a better match between the model’s prediction and the true location of the camouflaged soldier.

1.2 Motivation

Military camouflage provides soldiers with a critical advantage by hiding from enemy forces. This translates to several benefits on the battlefield, including a reduced risk of detection (increasing survival rates and mission success), and the ability to deceive enemies about their location or intentions (providing tactical advantages). However, current detection methods for camouflage might not be as effective as needed. This project is driven by the need for more reliable and automated solutions to identify camouflaged personnel.

1.3 Problem Statement

Current military camouflage detection methods may not be sufficient, leaving soldiers vulnerable and limiting tactical advantages. This project proposes a deep learning solution using Dense Deconvolution Networks (DDNs) to automatically detect camouflaged personnel with high precision. By analyzing a diverse dataset of soldier images, the DDN model will learn to identify key features, even in complex backgrounds, potentially leading to significant advancements in military operations and surveillance

1.4 Objectives

1. To generate the feature map using convolutional layers from the Image input.
2. To Utilize Dense Deconvolutional network model that Upsample the feature map and extracts the data from the image.
3. To fit the layers densely to avoid loss of data.
4. To segment the Camouflaged Soldier in the image.
5. Train the model for detection and segmentation of the camouflaged soldier.

1.5 Scope

1. This model Detects camouflaged Soldiers in the Forests and Snow Fields of specific area..
2. Limited capability in segmenting very small or distant objects that might appear insignificant in the scene and struggle with multiple objects.

1.6 Advantages

1. This can significantly reduce the risk of soldiers being undetected, leading to increased battlefield safety.
2. This improves overall operational efficiency by freeing up valuable resources for other critical tasks.
3. This enhances the real-world applicability of the system in various operational theaters.

Chapter 2

LITERATURE REVIEW

2.1 Y. Zheng, X. Zhang, F. Wang, T. Cao, M. Sun and X. Wang, "Detection of People With Camouflage Pattern Via Dense Deconvolution Network," in IEEE Signal Processing Letters, vol. 26, no. 1, pp. 29-33, Jan. 2019

Yunfei Zheng and team presented The paper [1] which introduces a novel approach to address the challenging task of detecting camouflaged people within cluttered natural scenes. Addressing the lack of open evaluation and training data in this domain, the authors construct a specific dataset comprising camouflaged individuals in various natural settings. Their method leverages deep convolutional networks, specifically a dense deconvolution network (DDCN), to extract high-level semantic features essential for discerning camouflaged individuals. Through innovative training techniques and the fusion of semantic information, their proposed network effectively identifies camouflaged regions in images. Additionally, the authors employ superpixel segmentation and spatial smoothness constraints to further refine detection results. Experimental evaluations demonstrate the superiority of their method over classical approaches and typical CNN-based detection methods. This research not only presents a robust solution for camouflaged people detection but also contributes a valuable dataset and methodology to advance future research in this important field.

Advantages:

1. Effective detection of camouflaged individuals using deep convolutional networks.
2. Contribution of a dataset and methodology for advancing research in camouflaged. people detection.

Disadvantages:

1. Potential computational complexity due to deep networks.
2. Sensitivity to variations in natural scenes may impact performance.

2.2 D. Zhang, C. Wang and Q. Fu, ”Efficient Camouflaged Object Detection via Progressive Refinement Network,” in IEEE Signal Processing Letters, vol. 31, pp. 231-235, 2024

Zhang and team presented the paper introduces a Progressive Refinement Network (PRNet) [2] designed for the challenging task of camouflaged object detection (COD). The inherent difficulty lies in the similarity between camouflaged objects and their backgrounds, often compounded by ambiguous boundaries. To address these challenges, PRNet leverages human perception strategies through two key modules: a position-aware module (PAM) for rough object localization and an edge-guided fusion module (EGFM) for boundary refinement. These modules facilitate accurate and rapid identification of camouflaged objects within complex scenes. Extensive experiments conducted on four benchmark datasets demonstrate the superior performance of PRNet over 14 state-of-the-art algorithms. Notably, PRNet achieves real-time processing speeds while significantly outperforming existing methods in terms of accuracy, making it a promising solution for real-world applications requiring efficient and precise camouflaged object detection.

Advantages:

1. Enhanced accuracy via human perception-inspired modules (PAM and EGFM) for precise detection, even in complex scenes.

2. Real-time processing capability, vital for applications like surveillance or autonomous vehicles

Disadvantages:

1. Increased complexity due to multiple modules like PAM and EGFM, possibly making implementation and fine-tuning more challenging.

2. Sensitivity to scene complexity, leading to varied performance based on scene intricacy and camouflage intensity

2.3 H. Li, C. -M. Feng, Y. Xu, T. Zhou, L. Yao and X. Chang, ”Zero-Shot Camouflaged Object Detection,” in IEEE Transactions on Image Processing, vol. 32, pp. 5126-5137, 2023

Haoran Li and team presented “Zero-Shot Camouflaged Object Detection” (ZSCOD) [3] with a dataset of Camouflaged Object Images from CAMO, CHAMELEON, COD10K, NC4K. The paper presents ZSCOD, a novel framework for Camouflaged Object Detection (COD) that addresses the challenge of detecting objects from unseen classes. Traditional methods struggle with unseen classes due to their reliance on dense annotations, making them impractical. ZSCOD introduces a two-part solution: a Dynamic Graph Searching Network (DGSNet) and a Camouflaged Visual Reasoning Generator (CVRG). DGSNet adaptively captures edge details and class information, while CVRG generates pseudo-features resembling real ones, transferring knowledge from seen to unseen classes. Through experiments on the COD10K dataset, ZSCOD achieves state-of-the-art performance in detecting both seen and unseen camouflaged objects, offering a promising approach to overcome the limitations of traditional methods.

Advantages:

1. Enables detection of unseen camouflaged objects, expanding the model’s applicability.
2. Eliminates reliance on large-scale annotated data, allowing for knowledge transfer from seen to unseen classes.

Disadvantage:

1. Requires sophisticated training techniques due to the complexity of zero-shot learning and camouflaged object detection.
2. Performance may vary depending on the quality and diversity of the semantic features used for knowledge transfer.

2.4 G. Chen, S. -J. Liu, Y. -J. Sun, G. -P. Ji, Y. -F. Wu and T. Zhou, ”Camouflaged Object Detection via Context-Aware Cross-Level Fusion,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 10, pp. 6981-6993, Oct. 2022

Geng Chen and team presented “Context-aware Cross-level Fusion Network for Camouflaged Object Detection” [4] with a dataset of Camouflaged people Images from CAMO, CHAMELEON, COD10K, NC4K Datasets. This paper introduces a novel Context-aware Cross-level Fusion Network (C2F-Net) to tackle the intricate challenge of camouflaged object detection (COD). Addressing the inherent difficulties associated with low boundary contrast and varied object appearances, the proposed model effectively integrates contextaware cross-level features using Attention-induced Cross-level Fusion Module (ACFM) and Dual-branch Global Context Module (DGCM). By leveraging rich global context information and innovative feature fusion techniques, C2F-Net outperforms state-of-the-art models in accurately identifying camouflaged objects, as demonstrated through extensive experiments on benchmark datasets. Moreover, the model exhibits promising potentials in downstream applications like polyp segmentation, highlighting its versatility and effectiveness in real-world scenarios.%.

Advantages:

1. Enhanced Detection Accuracy: C2F-Net integrates context-aware features for improved identification of camouflaged objects.

2. Versatile Applications: Beyond detection, C2F-Net shows promise in tasks like polyp segmentation, highlighting its adaptability.

Disadvantages:

1. Increased Complexity: Multiple modules and fusion techniques raise computational demands.

2. Implementation Challenges: Its intricate design may require expertise and effort for implementation and fine-tuning.

2.5 Y. Lv, J. Zhang, Y. Dai, A. Li, N. Barnes and D. -P. Fan, ”Toward Deeper Understanding of Camouflaged Object Detection,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 7, pp. 3462-3476, July 2023

CZhang and team presented “Towards deeper understanding of camouflaged object detection” [5] with a dataset of Camouflaged objects Images from CAM-LDR DATASET. This paper proposes a novel approach to camouflaged object detection (COD) by introducing two new tasks: camouflaged object localization (COL) and camouflaged object ranking (COR). It argues that the traditional binary segmentation setting fails to capture the intricacies of camouflage and proposes a triple-task learning framework to simultaneously localize, segment, and rank camouflaged objects. By employing an eye tracker to generate fixation data, the paper creates datasets for COL and COR tasks. The proposed framework aims to improve understanding of camouflage mechanisms and enhance the effectiveness of COD models by considering the conspicuousness and detectability of camouflaged objects. Experimental results demonstrate state-of-the-art performance and provide insights into the hierarchical levels of visual perception and animal evolution.

Advantages:

1. Comprehensive Understanding: Triple-task learning allows localization, segmentation, and ranking, deepening understanding.

2. Improved Performance: Including COL and COR tasks enhances detection accuracy, achieving state-of-the-art results.

Disadvantages:

1. Data Intensiveness: Creating datasets with eye tracking is resource-heavy, limiting scalability.

2. Increased Complexity: Additional tasks and framework complexity may hinder efficiency and require careful management

2.6 Y. Yang and Q. Zhang, ”Finding Camouflaged Objects Along the Camouflage Mechanisms,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 34, no. 4, pp. 2346-2360, April 2024

Zhang and team presented “Finding Camouflaged Objects along the Camouflage Mechanisms” [6] with a dataset of Camouflaged objects Images from CAMO, COD10K. This paper introduces a novel approach to camouflaged object detection (COD) by reexamining camouflage mechanisms and proposing a de-camouflaging perspective. By incorporating salient object detection (SOD) into the COD model, they aim to alleviate camouflage effects and improve detection accuracy. Specifically, they propose a multi-task learning framework that models task-conflicting and task-consistent attributes between SOD and COD. The framework includes components such as a Region Distraction Module (RDM) and a Gate Classification (GC) strategy to suppress interference from salient objects, and an adversarial learning (AL) scheme and a Boundary Injection Module (BIM) to enhance boundary differences for accurate detection. Extensive experiments demonstrate the effectiveness of their proposed model over existing ones.

Advantages:

1. Enhanced accuracy through multi-task learning.

2. Comprehensive approach integrating various mechanisms.

Disadvantages:

1. Increased complexity due to multiple components.

2. Dependency on accurate performance of SOD.

2.7 W. Zhai, Y. Cao, H. Xie and Z. -J. Zha, "Deep Texton-Coherence Network for Camouflaged Object Detection," in IEEE Transactions on Multimedia, vol. 25, pp. 5155-5165, 2023

Zhai and team presented "Edge-Aware Mirror Network for Camouflaged Object Detection" [7] with a dataset of Camouflaged objects Images from CAMO, COD10K, NC4K. The proposed Edge-aware Mirror Network (EAMNet) addresses the challenge of camouflaged object detection (COD) by introducing a novel cross refinement process between edge detection and segmentation. Unlike existing methods that suffer from erroneous foregroundbackground predictions, EAMNet integrates edge and segmentation branches, guiding each other through specialized modules to enhance accuracy. Specifically, it employs edgeinduced integrity aggregation and segmentation-induced edge aggregation modules to refine features at multiple scales. Additionally, a guided-residual channel attention module helps extract structural details from low-level features. Experimental results demonstrate EAMNet's superiority over state-of-the-art methods on three COD datasets, showcasing its potential for advancing COD research and applications

Advantages:

1. EAMNet enhances accuracy through cross refinement between edge detection and segmentation.

2. The guided-residual channel attention module improves fine-grained detail extraction from low-level features.

Disadvantages:

1. EAMNet's complex architecture demands substantial computational resources and training time.

2. Performance of EAMNet might be sensitive to dataset variations, requiring meticulous tuning for optimal results.

Chapter 3

SOFTWARE REQUIREMENTS ANALYSIS

Requirements analysis, also known as requirements engineering, involves evaluating consumer demands for a new or modified product. It's a collaborative endeavor that requires a blend of engineering expertise and effective communication skills. These criteria, often termed as functional specifications in software engineering, must be specific, measurable, and comprehensive. A software requirements specification offers a detailed overview of the intended functionality and environment for a program under development. This document delineates what the program aims to accomplish and how it should function. By defining these requirements, the specification streamlines development efforts, allowing programmers to allocate their resources efficiently and reduce production costs. Moreover, the software requirements specification is vital in real-world scenarios as it dictates how the application interacts with hardware, other software programs, and end-users. It serves as a roadmap for development, ensuring effective communication between different components and optimizing the overall performance of the application..

3.1 Functional Requirements

A functional requirement process provides a system's functionality, and some of its subsystems. It also focuses on what behavior the system should offer; (What the system should do). The functional system requirements should also be explicitly defined in detail about the system services

1. **Image upload:** The Users should be able to upload images for Camouflaged military person segmentation.

2. **Feature Extraction:** The trained DDN model should be extracting the features of the uploaded image.

3. **Segmentation:** The model after feature Extraction should Segmentation of the Camouflaged Soldier in the image given.

3.2 Non-Functional Requirements

Non-Functional Requirements include:

1. **Performance:** The system should be resilient to network failures and avoid congestion. It should recover gracefully from path failures.
2. **Accuracy:** The reinforcement learning algorithms should adapt to changes in network conditions.
3. **Scalability:**Efficiently use network resources, such as Latency, without overloading any particular path.
4. **Portability:** Ensure low latency for data transmission and routing decisions to enhance network responsiveness.

3.3 Software Requirements

1. **Programming Language:** : Python
2. **Platform:** : Google Colab / kaggle

3.4 Hardware Requirements

1. **Processor** : 11th Gen Intel(R) Core(TM) i3-11154G4 @ 3.00GHz 2.19GHz
2. **RAM** : 16GB

Chapter 4

SOFTWARE DESIGN

4.1 Software Development Lifecycle: Agile model

The agile model was used for the development of the project. Agile life cycles are composed of several iterations or incremental steps towards the completion of a project. Figure 4.1 describes the Agile model.

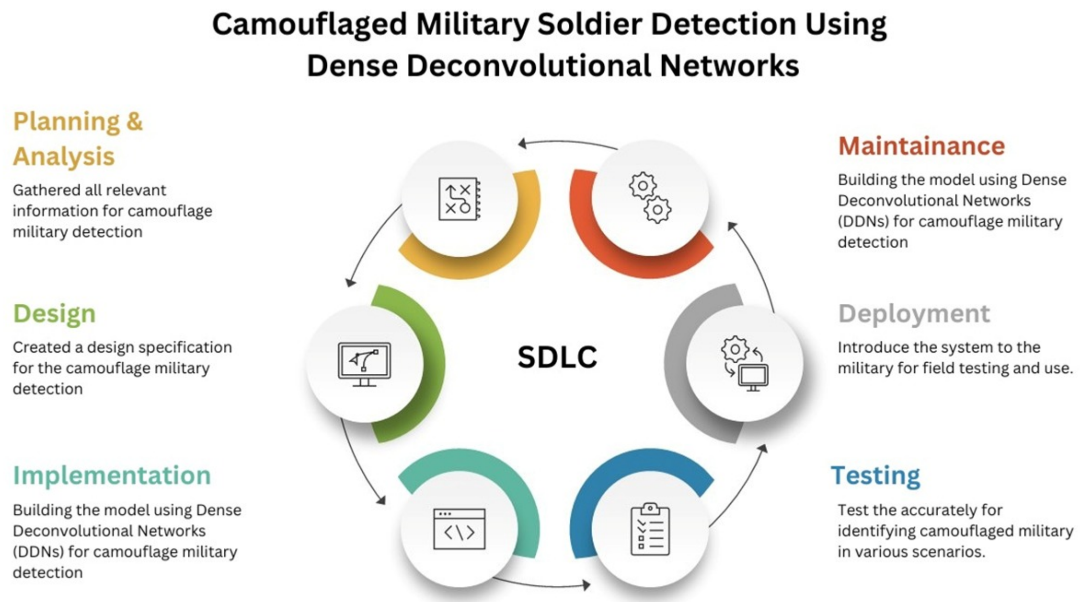

Figure 4.1: Software Development Lifecycle: Agile Model

The Agile model represents a collaborative approach to software development. It begins with the requirements phase, where user stories or product backlogs are gathered and refined to define the software's goals. Following this, analysis is performed on the requirements to refine them and verify their accuracy. Moving forward, the design phase focuses on crafting a flexible design capable of adapting to changing requirements. Subsequently, the Implementation phase involves iterative work in short bursts or sprints, where prioritized user stories are transformed into functional software components. Testing remains a continuous process throughout development, employing various methods to ensure quality and identify issues early. In the deployment phase, software is released incrementally, utilizing continuous integration and delivery practices for frequent updates. Finally, the evaluation phase assesses the system's performance, with feedback guiding further improvements.

4.2 UML Diagrams

4.2.1 Use case Diagram

A UML use case diagram serves as the main way to outline the requirements for a new software program. Use cases describe what the system should do, rather than how it should do it. These use cases can be represented both visually and in written form. One important aspect of use case modeling is that it allows us to design a system based on how users will interact with it. It's a helpful method for describing the system's behavior in terms that users can understand, detailing all the actions they'll see from the system. Figure 4.2 shows the use case diagram which consists of actors like source, ad Deep learning System. It contains use cases like collect dataset which includes the work of Collecting the dataset and preprocessing the images of required dimension and shape, Extract Features from images, Sample feature Map and Segment the Images are the parts of training the model Use case Which are done by DDN Architecture, and the final Use case is performance Evaluation which is used for evaluating the performance the model.

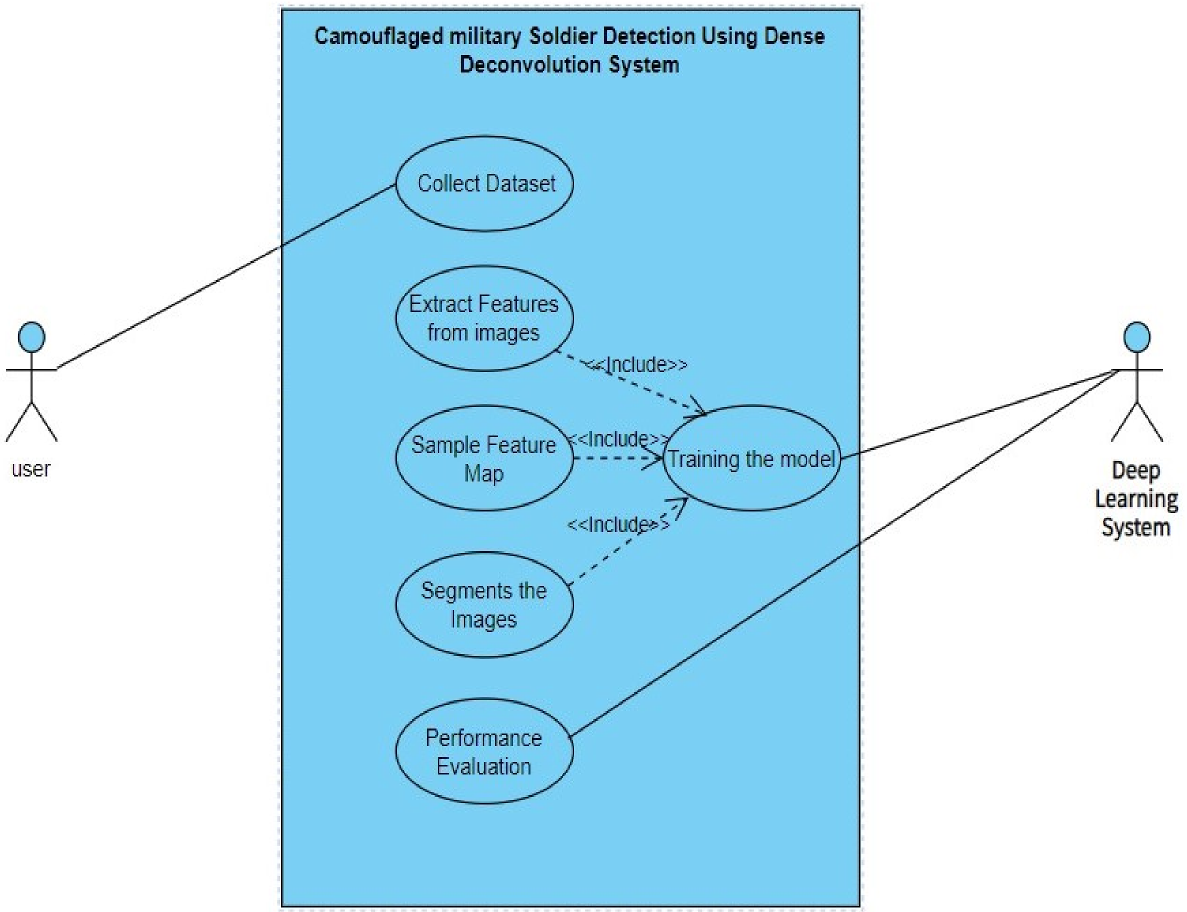

Figure 4.2: Use case Diagram

4.2.2 Sequence Diagram

Sequence Diagrams show how things happen step by step, like a story. They're like timelines but for interactions between different parts of a system. It's all about showing how different pieces of a system work together, whether it's a user talking to the system, the system talking to other systems, or parts of the system chatting with each other. The sequence diagram in Figure 4.3 consists of objects like the user, and deep learning system. The flow of interactions starts with the user by collecting Dataset and uploading the images to the deep learning system. The deep learning system will be extracting images from Feature Map and then Segments the Camouflaged Military Soldier with the DDN Architecture.Then the performance Evaluation is done for the model.

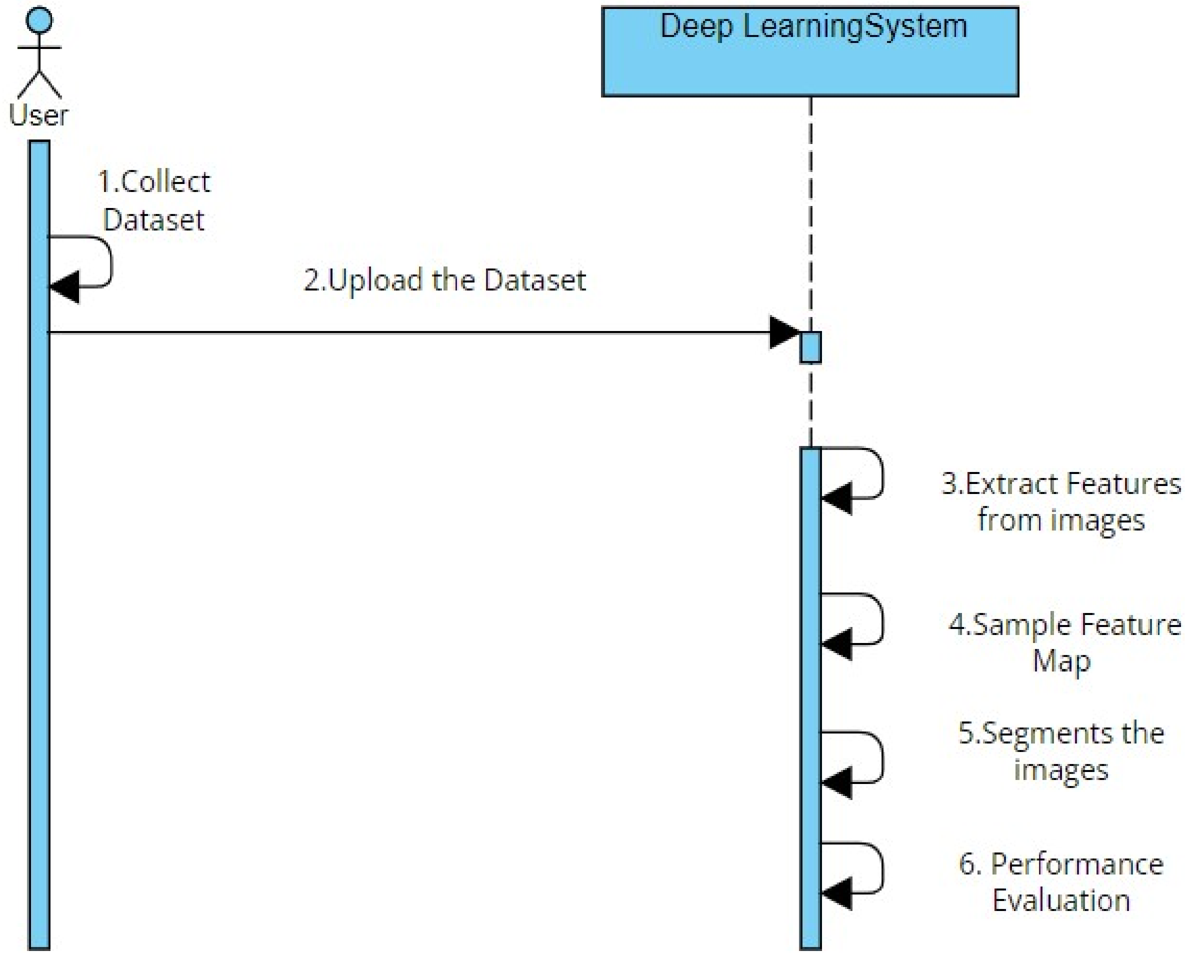

Figure 4.3: Sequence Diagram

4.2.3 Activity Diagram

Activity Diagrams show how different tasks work together to provide a service. They can describe activities at various levels of detail. Usually, they're used to explain how events are accomplished through a series of operations. This is especially useful when an operation needs to achieve multiple goals and needs coordination. They also help in understanding how events in a single scenario or task are connected, especially when these activities overlap and need to be coordinated. Figure 4.4 gives the activity diagram for the proposed work. It consists of objects like User ans Deep learning System. The interaction starts from the user by collecting the dataset and sending it to the deep learning system. After receiving the Dataset from the user preprocess the images and Extract features from the images and produces a feature map, which is then used for segmentation of the image through Deconvolution layers. Later the performance evaluation for the model is done through IoU evaluation metrics.

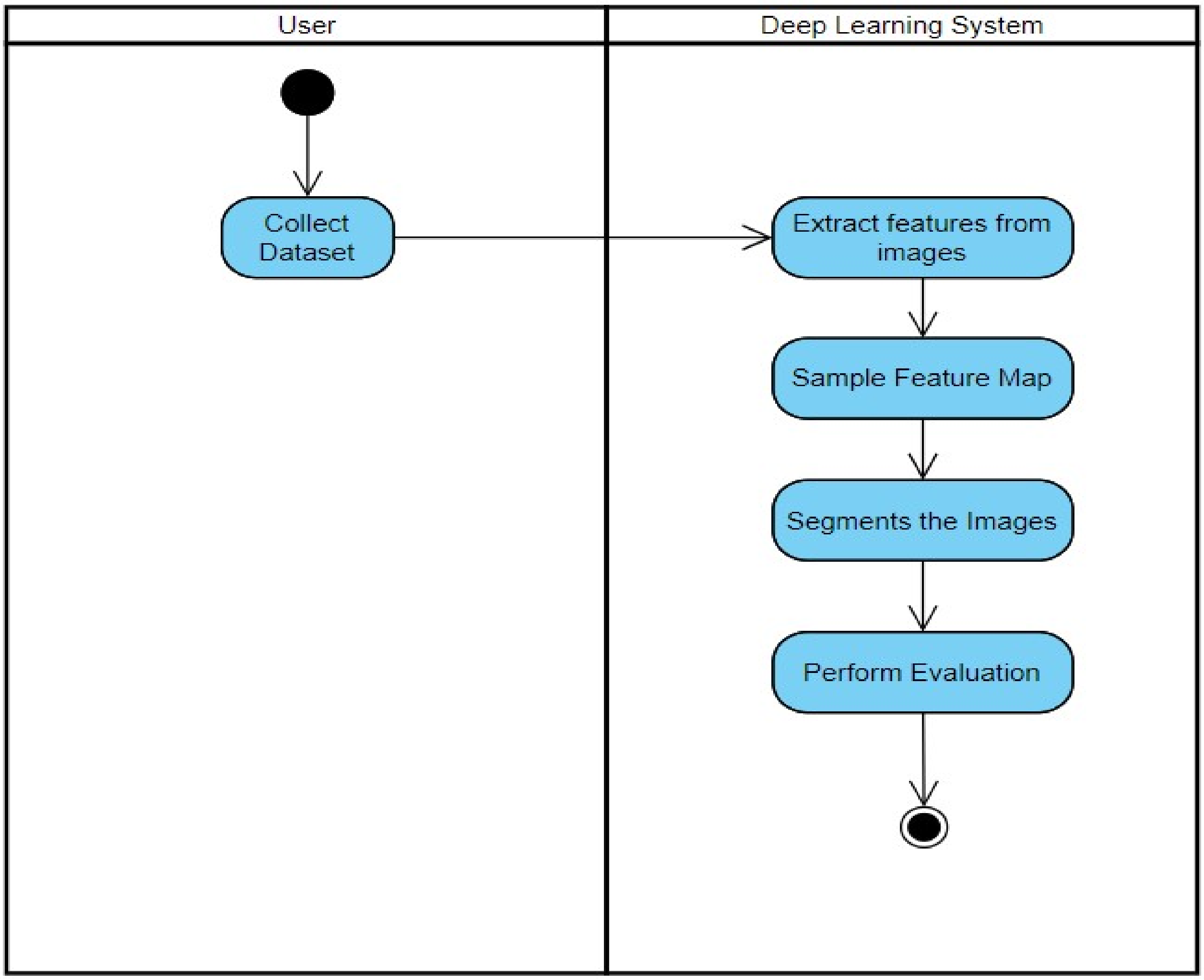

Figure 4.4: Activity Diagram

Chapter 5

PROPOSED SYSTEM

This chapter depicts a clear process and understanding of the proposed system and the project flow by including the module-wise description and illustration of the algorithm for each module.

5.1 Process Flow Diagram

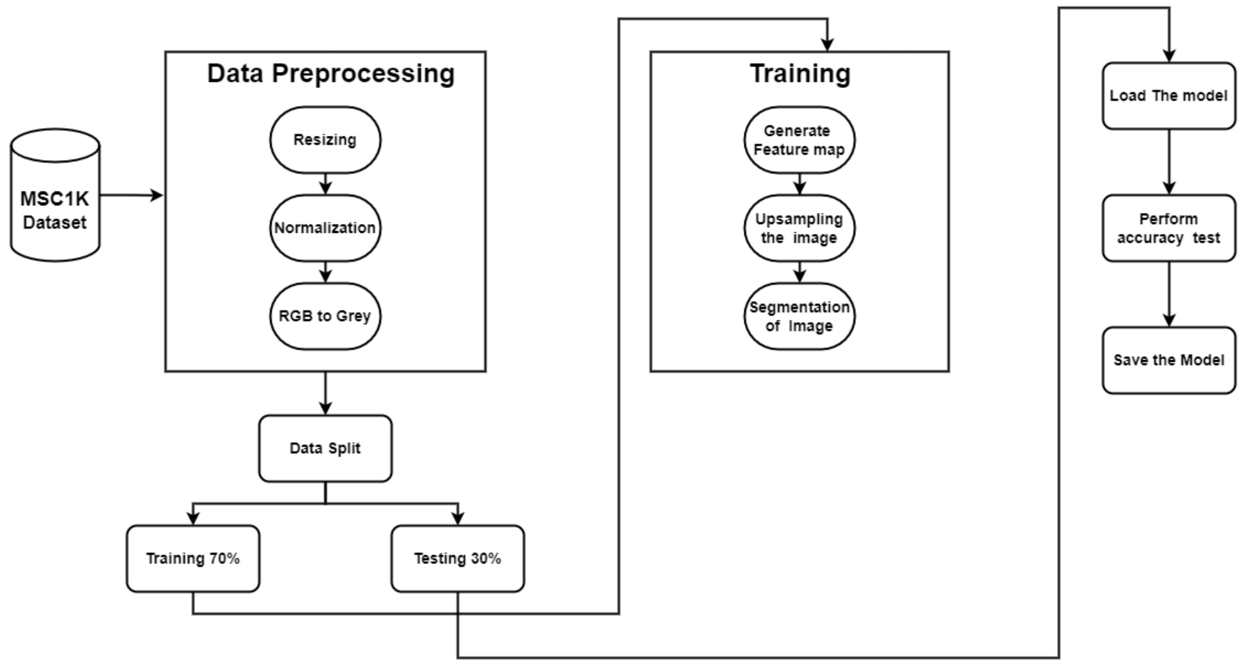

Figure 5.1: Process flow diagram

The figure 5.1 describes the proposed work flow for the model which segments the camouflaged military soldier. First, preprocess the data set. Give an image to generate the Feature map with the help of convolution layers and then the feature map is Upsampled and segmented using the Deconvolutional layers and a Feature matrix is generated. The feature matrix is then converted into Image again using PIL module of python, which contains the segmented image as output.

5.2 Architecture

5.2.1 Deconvolutional networks

Deconvolutional layers are used to Upsample feature maps, increasing their spatial resolution. These layers can be thought of as the opposite of convolutional layers, instead of reducing the spatial dimensions of the input, deconvolutional layers expand the spatial dimensions.

Upsampling

Deconvolutional networks inherently perform upsampling of feature maps to generate highresolution outputs. This process is essential for tasks like image super-resolution, where the goal is to generate a high-resolution image from a low-resolution.

5.2.2 Unet Architecture

The unet architecture as shown in figure 5.2 is commonly used for Image segmentation tasks.The architecture is characterized by its U-shaped design, which consists of a contracting path (encoder) and an expansive path (decoder), connected by a series of skip connections. This unique design allows the network to capture both local and global context information while preserving spatial information, making it particularly effective for tasks requiring precise segmentation.

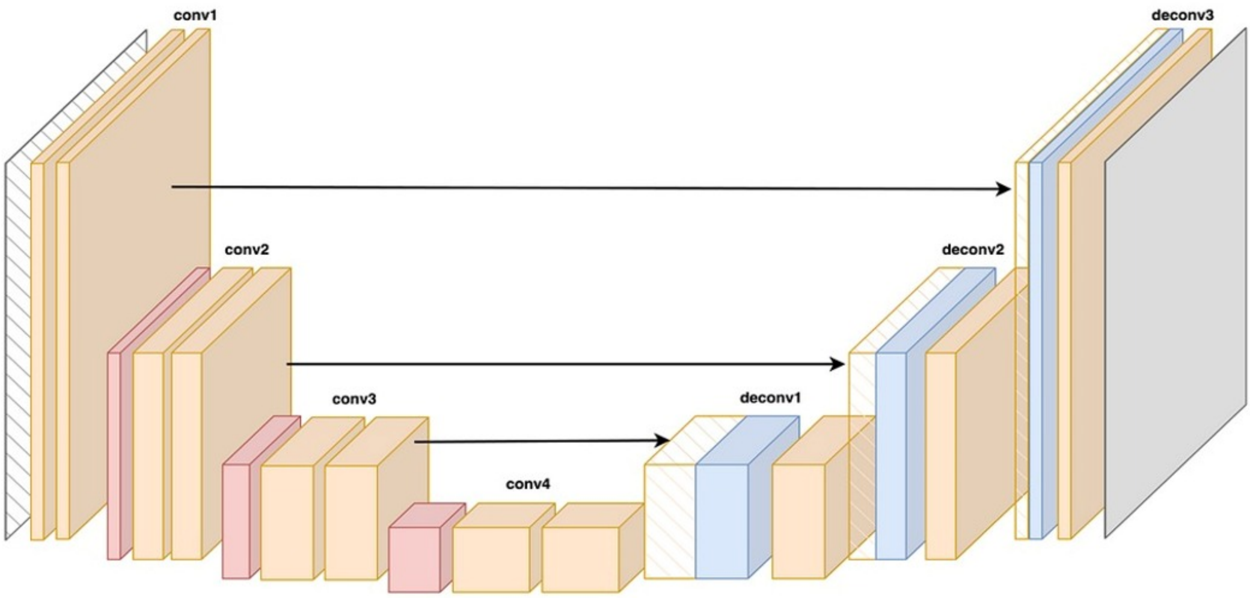

Figure 5.2: Architecture Diagram of Proposed Work [1]

5.3 Methodology

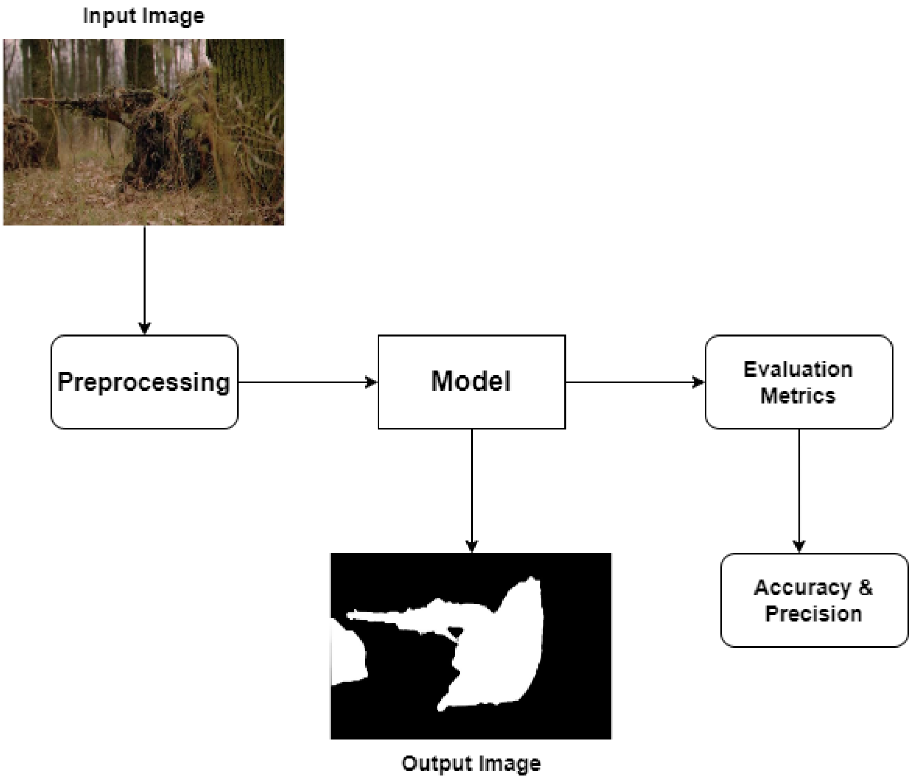

Figure 5.3: Methodology Diagram

The above Figure 5.3 represents the methodology of the project. In which first the image is preprocessed and sent to the model for segmentation of Camouflaged Soldier. The model produces an output image as shown in the Figure and later the model is Evaluated with the Evaluation Metrics.

5.3.1 Dataset Preprocessing module

Image resizing : Image Resizing is the process of adjusting the dimensions of an image, either reducing or enlarging it, while maintaining its aspect ratio and quality. It involves manipulating pixels to fit specific dimensions or aspect ratios.
for our model the Images are to be preprocessed into dimensions of (512,512,1) of Gray scale type before giving as input to the Model.

5.3.2 Training Module

Ensure you have the MSC1K dataset prepared, including images and their segmented ground truth instances. now after preprocessing the images ad resizing into the required dimension (512,512) load the image as Binary image and map it to it GT instance and prepare a numpy array of input and output images. The numpy arrays will be serving as inputs to our DDN model. Make sure the layers produce feature map and send to the Deconvolutional layers which segments the image by feature extraction through Upsampling. Use the Appropriate Activation functions to avoid loss. Now train the model using the dataset and evaluate the performance regularly to avoid Overfit.

5.3.3 Model Evaluation

The performance metric accuracy is considered to assess the effectiveness of the model. The metrics provided valuable insights into the model's ability.Accuracy is useful metric for evaluating the overall performance of a model.
Step 1: Evaluate the model with test image and generate output.
Step 2: calculate the precision and Accuracy by IOU technique.

5.4 Dataset Details

Name of Dataset	: Military soldier camouflage dataset
Dataset ID	: MSC1K
Number of Images	: 2104 images
Train set size	: 1496 images
Test set size	: 608 images

Chapter 6

IMPLEMENTATION AND RESULTS

6.1 Output & Results

The implementation of the proposed work is implemented using DDN model [1]. which takes the input image of Camouflaged military soldier and gives segmented image as output.

6.1.1 Model Outputs

During the training phase, the model was trained on the MSC1K dataset using DDN of Unet architecture. The CNN layers were utilized as the backbone to the Model by producing the feature map of the input by extract image features. The Deconvolutional network was employed to Segment the image of camouflaged soldiers by Upsampling using Deconvolution layers based on the extracted image features.

Figure 6.1: Test Image with its Output

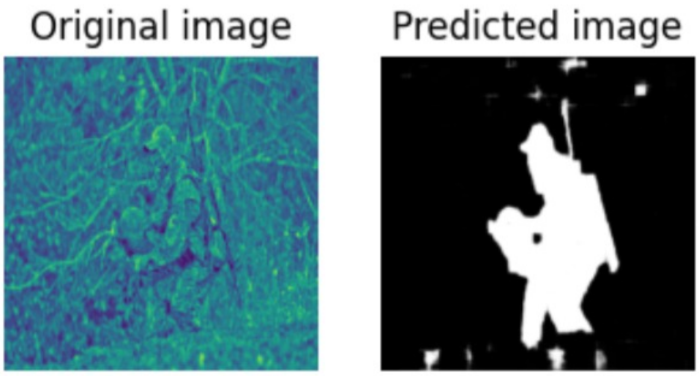

Figure 6.2: Segmented Output for given Image

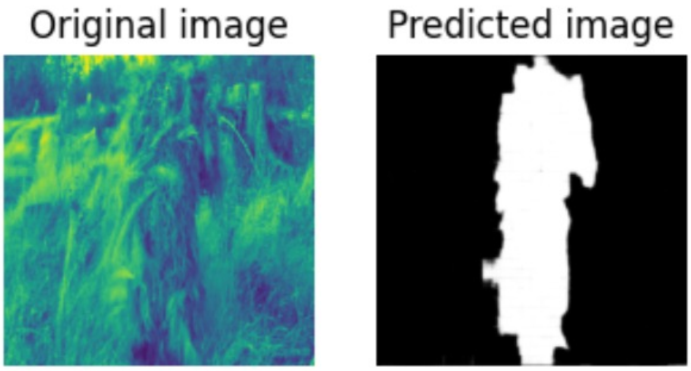

Figure 6.3: Segmented Output for given Image

The above Figures 6.1, 6.2, 6.3 shows the output for the example image and we can see that the person is segmented and we can compare the predicted image with the Ground Truth (GT) image.

6.1.2 Performance metrics

Intersection of Union (IoU)

Intersection over Union (IoU) : IoU is a common evaluation metric used in binary image segmentation tasks to measure the accuracy of the segmentation results. It quantifies the degree of overlap between the predicted segmentation (usually denoted as P) and the ground truth segmentation (usually denoted as GT). The IoU is calculated by dividing the area of overlap between the predicted and ground truth segmentation masks by the area of their union. In binary segmentation tasks, each pixel in the segmentation masks is typically classified as either belonging to the object of interest (foreground) or not (background). The Iou can be calculated from Equation 6.1.

$$IoU = \frac{|P \cap GT|}{|P \cup GT|} \tag{6.1}$$

where:

- The numerator denotes the number of pixels that are both in the predicted segmentation mask P and the ground truth segmentation mask GT.
- The denominator denotes the total number of pixels that are in either the predicted segmentation mask P or the ground truth segmentation mask GT.

Pixel Accuracy

Pixel accuracy measures the proportion of correctly classified pixels in the segmentation mask compared to the total number of pixels. It can be calculated from Equation 6.2.

$$\text{Pixel Accuracy} = \frac{\text{Number of correctly classified pixels}}{\text{Total number of pixels}} \tag{6.2}$$

Precision

In the context of binary segmentation, precision measures the proportion of correctly classified foreground pixels (true positives) out of all pixels classified as foreground (both true positives and false positives)

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{True Negatives}} \tag{6.3}$$

Where:

- True Positives (TP) are the number of correctly classified foreground pixels.
- False Positives (FP) are the number of pixels classified as foreground but are actually background.

The metrics Values we got for our model are,

- Overall Pixel Accuracy: 0.89166107177734
- Overall Precision: 0.86212550742585
- Overall IoU: 0.85165103565562

6.2 User Interface

6.2.1 Gradio

Gradio is a Python library for quickly creating customizable UI components for machine learning models. It allows developers to build interactive interfaces for model deployment without extensive web development experience. Gradio supports various input types, including text, images, and audio, making it versatile for a range of applications.

Our model (DDN model) is saved as a pickle file. With the python module **Gradio** the Graphical User Interface (GUI) is done by deploying the Pickle file in it. In which there will be a URL generated for 72 hours to use the interface. The User interface will be taking the image type as input and Produces Segmented Image as output as shown in Figure 6.4.

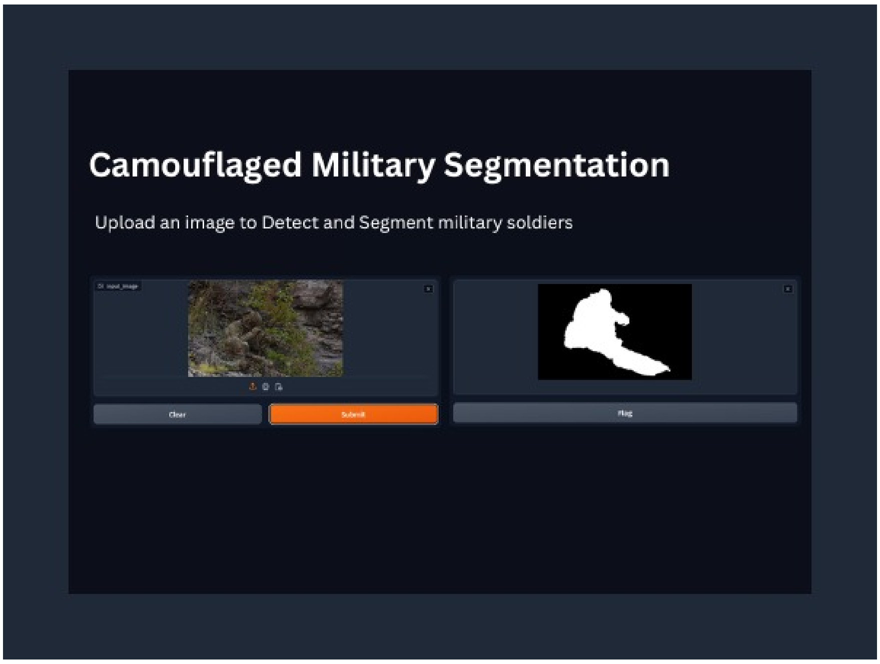

Figure 6.4: User Interface

6.3 Comparison with other models

The below Table 6.1 compares Accuracy of various models. The comparison table showcases the performance metrics of different models in the related work. Zheng's DDN achieved an accuracy of 0.76, Zhang's PRNet attained 0.82 accuracy, and Li's ZSCOD demonstrated an accuracy of 0.81. In contrast, our model, DDN (unet), outperformed these benchmarks with an accuracy of 0.85. This highlights the effectiveness of our model in achieving higher accuracy compared to existing approaches in the field.

Related work	model	Performance-metrics
Zheng, Y. [1]	DDN	Accuracy: 0.76
Zhang, D. [2]	PRNet	Accuracy: 0.82
Li, H. [3]	ZSCOD	Accuracy: 0.81
Our model	DDN (unet)	Accuracy: 0.85

Table 6.1: Camparision table

6.4 Results & Analysis

we employed Dense Deconvolutional Network model to effectively address the challenging task of identifying and isolating camouflaged soldiers within complex environments. Through extensive training on a comprehensive dataset, our model demonstrated satisfactory performance on the Dataset MSC1K , with an Overall Pixel Accuracy of 0.89166107177734, Overall Precision of 0.86212550742585 an Overall Intersection over Union (IoU) score of 0.85165103565562. This performance evaluation was meticulously conducted using rigorous metrics, ensuring robustness and reliability in our segmentation approach. Furthermore, to validate the accuracy of our model's predictions, we utilized sophisticated evaluation techniques, such as comparing the predicted segmentations with ground truth annotations. Once validated, our model seamlessly integrated into real-world scenarios, facilitating swift and accurate segmentation of camouflaged soldiers in real-time operations. This integration was optimized for efficiency, enabling rapid processing and analysis of images in dynamic environments, thus enhancing situational awareness and operational effectiveness.

Chapter 7

CONCLUSION AND FUTURE WORK

Our project focuses on utilizing Dense Deconvolutional Networks (DDNs) to address the critical challenge of segmenting camouflaged military soldiers in images. We have developed a sophisticated model that effectively identifies and separates camouflaged soldiers from their surroundings. By employing Upsampled feature maps, our model enhances the segmentation process, thereby reducing the threat posed by these concealed individuals. The utilization of Rectified Linear Unit (ReLU) and Sigmoid activation functions further refines the segmentation accuracy. Notably, our model is specifically designed to operate within snow and forest environments, where camouflage tactics are commonly employed. This targeted approach ensures optimal performance in scenarios relevant to military operations within these specific terrains.

While our current project marks a significant advancement in camouflaged object segmentation, there are several avenues for future exploration and enhancement. Firstly, extending the model's applicability to diverse environments beyond snow and forest landscapes could broaden its utility in various military scenarios. Additionally, incorporating attention mechanisms or reinforcement learning techniques may further improve segmentation accuracy and robustness in challenging conditions. Furthermore, optimizing the computational efficiency of the model to enable real-time deployment on resource-constrained platforms would enhance its practical usability in field operations. Exploring transfer learning approaches to adapt the model to new environments or refining the dataset to include a wider range of camouflage patterns could also contribute to performance improvements. Overall, continued research and development efforts in these directions promise to enhance the effectiveness and versatility of camouflaged object detection systems, thereby bolstering military situational awareness and operational effectiveness.

REFERENCES

[1] Zheng, Y., Zhang, X., Wang, F., Cao, T., Sun, M., Wang, X. (2018). Detection of people with camouflage pattern via dense deconvolution network. IEEE Signal Processing Letters, 26(1), 29-3.

[2] Zhang, D., Wang, C., Fu, Q. (2023). Efficient Camouflaged Object Detection via Progressive Refinement Network. IEEE Signal Processing Letters.

[3] Li, H., Feng, C. M., Xu, Y., Zhou, T., Yao, L., Chang, X. (2023). Zero-shot camouflaged object detection. IEEE Transactions on Image Processing.

[4] Chen, G., Liu, S. J., Sun, Y. J., Ji, G. P., Wu, Y. F., Zhou, T. (2022). Camouflaged object detection via context-aware cross-level fusion. IEEE Transactions on Circuits and Systems for Video Technology, 32(10), 6981-6993.

[5] Y., Zhang, J., Dai, Y., Li, A., Barnes, N., Fan, D. P. (2023). Towards deeper understanding of camouflaged object detection. IEEE Transactions on Circuits and Systems for Video Technology.

[6] Yang, Y., Zhang, Q. (2023). Finding Camouflaged Objects along the Camouflage Mechanisms. IEEE Transactions on Circuits and Systems for Video Technology.

[7] Zhai, W., Cao, Y., Xie, H., Zha, Z. J. (2022). Deep texton-coherence network for camouflaged object detection. IEEE Transactions on Multimedia.

[8] Ren, J., Hu, X., Zhu, L., Xu, X., Xu, Y., Wang, W., ... Heng, P. A. (2021). Deep texture-aware features for camouflaged object detection. IEEE Transactions on Circuits and Systems for Video Technology.

[9] B. Alexe, T. Deselaers and V. Ferrari, “Measuring the objectness of image windows,”IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2189–2202, Nov. 2012.

[10] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A.W. M. Smeulders, “Selective search for object recognition,” Int. J. Comput. Vis., vol. 104, no. 2, pp. 154–171, 2013.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Region-based convolutional networks for accurate object detection and segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 1, pp. 142–158, Jan. 2016.

[12] S. Q. Ren, K. M. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017..

[13] Ranran Feng and Balakrishnan Prabhakaran, “Facilitating fashion camouflage art,” in Proceedings of the 21st ACM nternational conference on Multimedia, 2013, pp. 793–802.

[14] Ajoy Mondal, “Camouflaged object detection and tracking: A survey,” International Journal of Image and Graphics, vol. 20, no. 04, pp. 2050028, 2020.

[15] Yunqiu Lv, Jing Zhang, Yuchao Dai, Aixuan Li, Bowen Liu, Nick Barnes, and DengPing Fan, “Simultaneously localize, segment and rank the camouflaged objects,” in Proceedings of the IEEE/CVF Conference

[16] eng-Ping Fan, Ge-Peng Ji, Ming-Ming Cheng, and Ling Shao, “Concealed object detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 10, pp. 6024–6042, 2021.

[17] Fan Yang, Qiang Zhai, Xin Li, Rui Huang, Ao Luo, Hong Cheng, and Deng-Ping Fan, “Uncertainty-guided transformer reasoning for camouflaged object detection,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 4146–4155.

[18] A. Arnab, C. Sun, and C. Schmid, “Unified graph structured models for video understanding,” in Proc. IEEE/CVF Int. Conf. Comput. Vis.(ICCV), Oct. 2021, pp. 8117–8126.

[19] X. Li, Z. Zhong, J. Wu, Y. Yang, Z. Lin, and H. Liu, “Expectation maximization attention networks for semantic segmentation,” in Proc.IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 9167–9176.

[20] Y. Chen, M. Rohrbach, Z. Yan, Y. Shuicheng, J. Feng, and Y. Kalantidis, “Graph-based global reasoning networks,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.

Appendix A
Report Plagiarism

return project doc.pdf

ORIGINALITY REPORT

19%	11%	15%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Yang Yang, Qiang Zhang. "Finding Camouflaged Objects along the Camouflage Mechanisms", IEEE Transactions on Circuits and Systems for Video Technology, 2024 Publication	1%
2	www.coursehero.com Internet Source	1%
3	Imran Ahmed, Misbah Ahmad, Gwangill Jeon, Abdellah Chehri. "An Internet-of-Things and AI-Powered Framework for Long-Term Flood Risk Evaluation", IEEE Internet of Things Journal, 2023 Publication	1%
4	www.arxiv-vanity.com Internet Source	1%
5	Zheng YunFei, Xiongwei Zhang, Wang Feng, Tiieyong Cao, Meng Sun, Wang Xiaobing. "Detection of People with Camouflage Pattern Via Dense Deconvolution Network", IEEE Signal Processing Letters, 2018 Publication	1%

6	Dongdong Zhang, Chunping Wang, Qiang Fu. "Efficient Camouflaged Object Detection via Progressive Refinement Network", IEEE Signal Processing Letters, 2023 Publication	1%
7	arxiv.org Internet Source	1%
8	"Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2024 Publication	1%
9	www.researchgate.net Internet Source	1%
10	Haoran Li, Chun-Mei Feng, Yong Xu, Tao Zhou, Lina Yao, Xiaojun Chang. "Zero-shot Camouflaged Object Detection", IEEE Transactions on Image Processing, 2023 Publication	1%
11	cse.anits.edu.in Internet Source	1%
12	export.arxiv.org Internet Source	<1%
13	Submitted to University of Greenwich Student Paper	<1%
14	www.mdpi.com Internet Source	<1%

Appendix B
Code Plagiarism

Code for plagarism check.pdf

ORIGINALITY REPORT

28%	24%	18%	23%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	programtalk.com Internet Source	4%
2	Submitted to Kotebe University Student Paper	3%
3	haystar.tistory.com Internet Source	3%
4	github.com Internet Source	3%
5	Submitted to Deakin University Student Paper	2%
6	www.programcreek.com Internet Source	2%
7	blog.faradars.org Internet Source	2%
8	huggingface.co Internet Source	2%
9	developer.aliyun.com Internet Source	1%

10	blog.csdn.net Internet Source	1 %
11	Monie, Dileep Daniel. "Systems Biology for Engineering Regenerative Immunotherapies in Precision Neuro-Oncology", College of Medicine - Mayo Clinic, 2024 Publication	1 %
12	Submitted to University of Sunderland Student Paper	1 %
13	machinelearningmastery.com Internet Source	1 %
14	edoc.pub Internet Source	1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		