

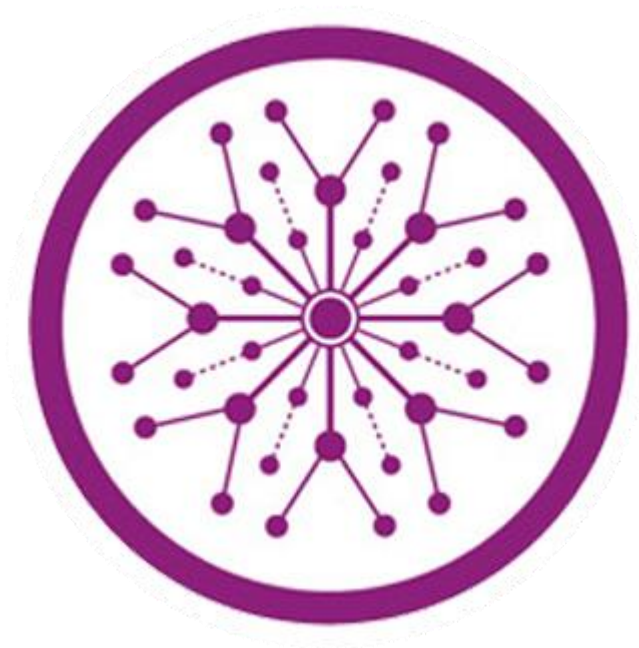
3D Reconstruction from Image

Final Year Project

Session 2019-2023

A project submitted in partial fulfillment of the degree of

BS in Data Science



Department of Software Engineering

Faculty of Computer Science & Information Technology

The Superior University, Lahore

Spring 2022

| | | | | |
|------------------------------|----------------------------------------------------------------------------------------------------------------|---------------|----------------------------------------------------------------------|------------|
| Type (Nature of project) | <input type="checkbox"/> Development <input type="checkbox"/> Research <input checked="" type="checkbox"/> R&D | | | |
| Area of specialization | Machine Learning / Deep learning | | | |
| FYP ID | BSDS-FYP1-F22-002 | | | |
| Project Group Members | | | | |
| Sr.# | Reg. # | Student Name | Email ID | *Signature |
| (i) | | Ali Amjad | alishah33n@icloud.com | |
| (ii) | | Hamza Mustafa | hamzamustafa7501@gmail.com | |
| (iii) | | Abdul Rafay | rafaymajeed66@gmail.com | |

*The candidates confirm that the work submitted is their own and appropriate credit has been given where reference has been made to the work of others

Plagiarism Free Certificate

This is to certify that, I Ali Amjad S/D of Amjad Rafique, group leader of FYP under registration no _____ at Software Engineering Department, The Superior College, Lahore. I declare that my FYP report is checked by my supervisor.

Date: **05-06-23** Name of Group Leader: **Ali Amjad** Signature: _____

Name of Supervisor: Ms. Arshi Iftikhar

Designation: Lecturer

Signature: _____

HOD: Dr. Tehreem Masood

Signature: _____

Project Report

3D Reconstruction from Image

Change Record

| Author(s) | Version | Date | Notes | Supervisor's Signature |
|-----------|---------|------|---------------------------------------------|------------------------|
| | 1.0 | | <Original Draft> | |
| | | | <Changes Based on Feedback from Supervisor> | |
| | | | <Changes Based on Feedback From Faculty> | |
| | | | <Added Project Plan> | |
| | | | <Changes Based on Feedback from Supervisor> | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

APPROVAL

PROJECT SUPERVISOR

Comments: _____

Name: _____

Date: _____

Signature: _____

PROJECT MANAGER

Comments: _____

Date: _____

Signature: _____

HEAD OF THE DEPARTMENT

Comments: _____

Date: _____

Signature: _____

Dedication

This modest effort is dedicated to our beloved parents who raised us to the level of excellence where we are studying today, looking for the most promising and brilliant future ahead for which they sacrificed most of their life, as well as to our respected and brilliant teachers who guided us throughout our academic careers! First and foremost, we dedicate our project to Allah Almighty and to Muhammad (Peace Be Upon Him), to whom the world owes its existence. Also, thank you to everyone who included us in their prayers!

Many thanks to all of my teachers!

Acknowledgments

I would like to extend my heartfelt gratitude to our esteemed teachers and final-year supervisor, Ma'am Arshi Iftikhar. Their invaluable recommendations and unwavering encouragement played a pivotal role in our project's successful completion, especially in the creation of this report. I am also immensely grateful to all my teachers who diligently assisted me in assembling the various components and provided invaluable guidance on numerous matters.

Executive Summary

3D reconstruction from 2D is a computer vision technique used to create 3D models from 2D images or videos. The process involves estimating the depth and shape of objects in an image and converting that information into a 3D representation. This is achieved through techniques such as structure from motion, stereo vision, or single-view depth estimation. The use of 3D reconstruction is widespread in various fields such as computer graphics, robotics, virtual reality, and medical imaging. The benefits of 3D reconstruction include the ability to create interactive 3D models, improve accuracy and efficiency in tasks such as object recognition, and support simulation and visualization in a wide range of applications. However, the accuracy of the reconstruction heavily depends on the quality of the input images, the algorithms used, and the desired level of detail.

Table of Contents

| | |
|---------------------------------------------------|------|
| 3D Reconstruction from Image..... | i |
| Final Year Project | i |
| Session 2019-2023 | i |
| BS in Data Science..... | i |
| Plagiarism Free Certificate | ii |
| Dedication | v |
| Acknowledgments..... | vi |
| Executive Summary | vii |
| Table of Contents | viii |
| List of Figures | x |
| List of Tables | xi |
| Chapter 1 | 1 |
| Introduction..... | 1 |
| 1.1 Background | 2 |
| 1.2 Motivations and Challenges | 2 |
| 1.3 Goals and Objectives..... | 3 |
| 1.4 Literature Review/Existing Solutions | 3 |
| 1.5 Gap Analysis | 4 |
| 1.6 Proposed Solution | 5 |
| 1.7 Project Plan | 5 |
| 1.7.1 Work Breakdown Structure | 6 |
| 1.7.2 Roles & Responsibility Matrix | 7 |
| 1.7.3 Gantt Chart..... | 9 |
| 1.8 Report Outline | 9 |
| 1.9 Empathy Map | 10 |
| Chapter 2..... | 11 |
| Data Collection | 11 |
| 2 Data Collection | 12 |
| 2.1 Introduction | 12 |
| 2.1.1 Sources of Data | 12 |
| 2.2 Data preparation | 12 |
| 2.3 Data Storage | 12 |
| 2.4 Data Validation | 12 |
| 2.5 Annotation (if necessary) | 12 |
| 2.6 Data sampling (if necessary) | 13 |
| 2.7 Data privacy and security | 13 |
| Chapter 3..... | 14 |
| Data Preprocessing..... | 14 |
| 3 Data Preprocessing | 15 |
| 3.1 Description of the data cleaning process..... | 15 |
| 3.2 Identification of data issues:..... | 15 |
| 3.3 Handling of missing values:..... | 16 |

| | | |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.4 | Handling of outliers:..... | 16 |
| 3.5 | Handling of inconsistencies: | 17 |
| | Handling inconsistencies in image-based datasets is important for ensuring the quality and reliability of the data. Inconsistencies can arise from a variety of sources, such as errors in data collection or data entry, or differences in data formatting. We remove those images that having errors or incorrect format. | 17 |
| 3.6 | Data normalization: | 17 |
| 3.7 | Data transformation:..... | 17 |
| 3.8 | Conclusion:..... | 17 |
| | Chapter 4..... | 18 |
| | Data Exploration | 18 |
| 4 | Data Exploration..... | 19 |
| 4.1 | Description of Dataset..... | 19 |
| 4.2 | Descriptive statistics..... | 19 |
| 4.3 | Visualizations | 20 |
| | Chapter 5..... | 22 |
| | Purposed Approach..... | 22 |
| 5 | Purposed Approach: | 23 |
| 5.1 | Data Model:..... | 23 |
| 1. | Single-View Images: | 23 |
| 2. | Ground Truth 3D Representations:..... | 23 |
| 3. | Correspondence between Images and 3D Representations: | 23 |
| 4. | Data Split:..... | 24 |
| 5.2 | Model Selection: | 24 |
| 5.3 | Architecture diagram:..... | 24 |
| 6 | References..... | 25 |

List of Figures

| | |
|-----------------------------------------|---|
| Figure 1 Work Breakdown Structure | 6 |
| Figure 2 Gantt Chart | 9 |

List of Tables

| | |
|------------------------------------------------|----|
| Table 1 Roles & Responsibility | 9 |
| Table 2 Empathy Map | 10 |
| Table 3 Grid Graph of sample dataset | 20 |
| Table 4 Heat map Graph of sample dataset | 21 |
| Table 5 Architecture Diagram | 24 |

Chapter 1

Introduction

Chapter 1: Introduction

Creating a 3D model from 2D images that is realistic as possible is one of the fundamental issues of image-based modeling and computer vision. Currently, there are several methods of 3D reconstruction from 2D images each algorithm has its own conditions of execution, its strengths as well as its weak points. In this we give a definition and the domains of uses of 3D reconstruction from 2D images. We will develop a system that can convert 2D image into 3D by applying some different deep learning algorithms.

1.1 Background

3D reconstruction from 2D images is a technique used to create a three-dimensional representation of an object from a series of two-dimensional images. The process involves using computer vision algorithms to analyze the images and extract information about the shape and structure of the object. This information is then used to create a 3D model of the object.

The techniques used in 3D reconstruction from 2D images can be divided into two main categories: stereo reconstruction and structure from motion. Stereo reconstruction uses two or more images of the same scene taken from different viewpoints to calculate the depth of each point in the scene. Structure from motion, on the other hand, uses a series of images taken of an object or scene as it moves to create a 3D model.

3D reconstruction from 2D images has a wide range of applications, including in robotics, industrial inspection, film and game production, archaeology, and many more. With the development of modern computer vision algorithms and the increasing availability of powerful computational resources, the accuracy and quality of 3D reconstructions has greatly improved.

1.2 Motivations and Challenges

In recent years, 3D imaging has become important in industrial and consumer applications. Machine vision systems empowered with 3D imaging allow faster, more accurate inspection of components at manufacturing sites. In the consumer realm, 3D imaging provides greater image depth for media. 3D imaging can be used for a wide range of applications – analyzing, measuring,

and positioning parts are among the most important. To get the best results possible, however, it's crucial to design a system with the necessary performance and environmental constraints in mind.

1.3 Goals and Objectives

The goal of 3D image reconstruction is to create a three-dimensional representation of an object from a set of two-dimensional images. The objective is to recover the shape, surface texture, and other properties of the object as accurately as possible. This can be done using a variety of techniques, such as computed tomography (CT), magnetic resonance imaging (MRI), and stereo vision. The resulting 3D models can be used in a wide range of applications, including medical imaging, robotics, computer vision, and virtual reality.

1.4 Literature Review/Existing Solutions

A literature review of 3D reconstruction from 2D images would explore the various methods, algorithms, and techniques that have been proposed and used in this field. This would likely include an overview of the different types of 3D reconstruction methods, such as stereo reconstruction and structure from motion, as well as a discussion of the advantages and limitations of each approach. [1]

One key area of research in 3D reconstruction from 2D images is the development of algorithms that can accurately and efficiently extract depth information from images. This includes methods such as photometric stereo, which uses the shading of an object to calculate its shape, as well as feature-based methods, which use point correspondences between images to estimate depth. [2]

Another important area of research in 3D reconstruction from 2D images is the creation of dense 3D models. Dense models are more detailed and accurate than sparse models, which only contain a small number of points. Dense reconstruction can be achieved through techniques such as multi-view stereo, which uses multiple images of the same scene to create a dense point cloud, and surface reconstruction, which generates a smooth surface that approximates the object being imaged. [3]

Additionally, recent research has been focused on solving the problem of occlusions, which occur when part of an object is obscured in an image, making it difficult to infer its shape. Approaches such as Multi-View Stereo and Multi-view CNN have been proposed to handle that problem.

In recent years, with the advent of deep learning, there has been significant progress in the field of 3D reconstruction from 2D images. Many researchers have proposed deep learning-based methods that can generate 3D models directly from 2D images, often with high accuracy and low computational complexity. [4]

In conclusion, the field of 3D reconstruction from 2D images is an active area of research that has seen significant progress in recent years. Various techniques, such as stereo reconstruction, structure from motion, multi-view stereo, and deep learning have been proposed and applied to the problem, each with its own advantages and limitations. The literature review will likely touch on the evolution of techniques and how recent techniques have improved the accuracy and quality of 3D reconstructions. [5]

1.5 Gap Analysis

A gap analysis in the field of 3D image reconstruction refers to the process of identifying the differences between the current state of 3D imaging technology and the desired state, or the ideal capabilities and performance that researchers and engineers hope to achieve. The gap analysis can include evaluating current algorithms and techniques used in 3D image reconstruction, as well as identifying areas where further research and development is needed.

Some potential areas of focus in a gap analysis of 3D image reconstruction could include:

- High-resolution imaging: current techniques may struggle to produce 3D images with sufficient resolution, especially when working with small or detailed objects.
- Real-time imaging: many current 3D reconstruction techniques are computationally intensive and may not be able to produce images in real-time, which can limit their usefulness in certain applications.

- **Robustness to noise and errors:** existing algorithms may be sensitive to noise or errors in the input data, which can make it difficult to produce accurate 3D images in certain conditions.
- **Generalization to new domains:** current techniques may be tailored to specific types of imaging modalities or objects, and may not generalize well to new domains.
- **Complex scenes:** Some current techniques lack the ability to model the 3D structure of complex and dynamic scenes, including the handling of occlusions, and different levels of depth of field.

Overall, a gap analysis of 3D image reconstruction would aim to identify the current limitations of the technology and the areas where further research and development is needed to improve performance and capabilities.

1.6 Proposed Solution

We will develop a model that can reconstruct the image from 2D to 3D by applying the concept of deep learning. The first step in reconstructing 2d image into 3d is preprocessing, which focuses on make the image into grey scale and get only the object from an image. Which means ignoring rest of background to make process easy for the model. This will be done by using OpenCV and numpy. In the second stage, apply model pytorch3d for converting images from 2d-3d. In the third stage, which is for testing in which we will test the dataset for measuring the accuracy of the model.

1.7 Project Plan

A project plan for a 3D reconstruction of an image typically includes the following steps:

1. **Data collection:** Acquire a set of images of the object or scene to be reconstructed. This can be done using a camera or a laser scanner.
2. **Image processing:** Process the images to correct for distortion and align them in a common coordinate system. This can involve techniques such as camera calibration, feature matching, and bundle adjustment.

3. **3D reconstruction:** Use the aligned images to create a 3D model of the object or scene. This can be done using methods such as structure from motion, multi-view stereo, or dense image matching.
4. **Mesh generation:** Convert the 3D model into a triangle mesh, which is a more suitable representation for 3D printing or rendering.
5. **Texturing:** Map the 2D images onto the 3D mesh to create a realistic and detailed representation of the object or scene.
6. **Quality assessment:** Evaluate the quality of the 3D reconstruction using metrics such as accuracy, completeness, and visual realism.
7. **Output:** Output the 3D reconstruction in a format that is suitable for the intended use (e.g., STL for 3D printing or PLY for rendering).
8. **Maintenance:** Keep the data updated and do regular maintenance on the 3D model.

1.7.1 Work Breakdown Structure

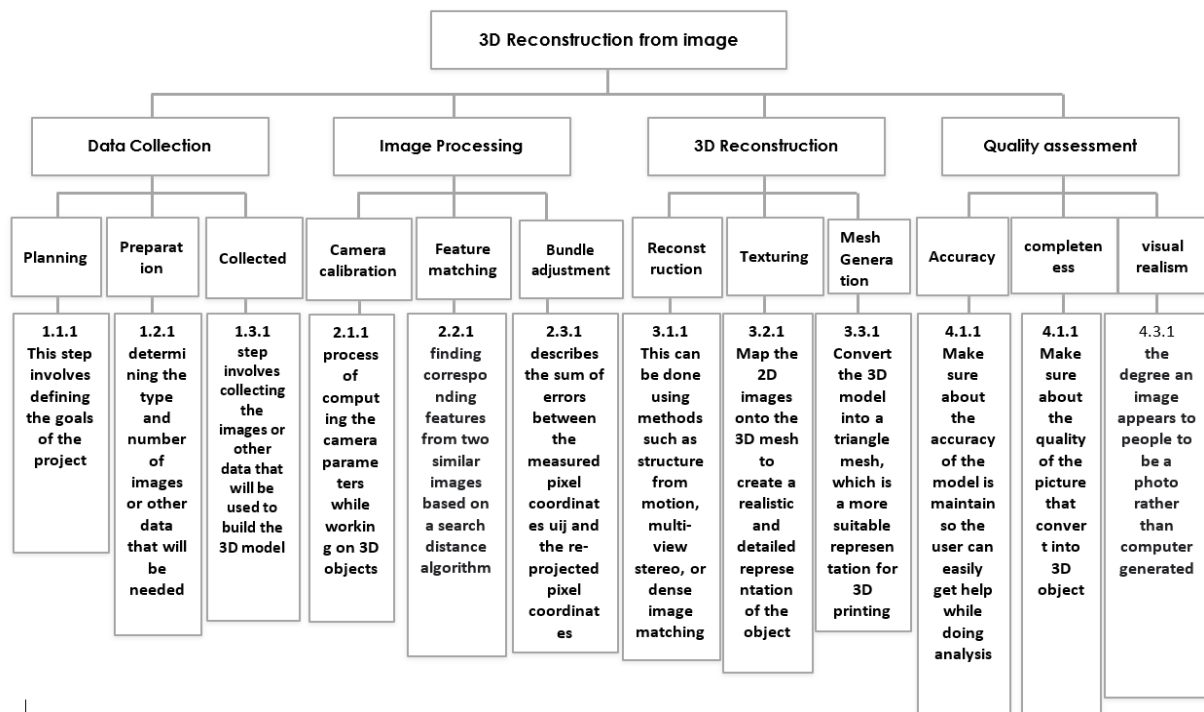


Figure 1 Work Breakdown Structure

1.7.2 Roles & Responsibility Matrix

| WBS # | WBS Deliverable | Activity # | Activity to Complete the Deliverable | Responsible Team Member(s) & Role(s) |
|--------------|------------------------|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|
| 1 | Data collection | 1 | Acquire a set of images of the object or scene to be reconstructed. This can be done using a camera or a laser scanner. | Ali Amjad Hamza Mustafa Abdul Rafay |
| 2 | Image processing | 2 | Process the images to correct for distortion and align them in a common coordinate system. This can involve techniques such as camera calibration, feature matching, and bundle adjustment. | Ali Amjad Hamza Mustafa |
| 3 | 3D reconstruction | 3 | Use the aligned images to create a 3D model of the object or scene. This can be done using methods such as structure from motion, multi-view | Ali Amjad Hamza Mustafa |

| | | | | |
|---|--------------------|---|------------------------------------------------------------------------------------------------------------------|------------------------------|
| | | | stereo, or dense image matching. | |
| 4 | Mesh generation | 4 | Convert the 3D model into a triangle mesh, which is a more suitable representation for 3D printing or rendering. | Hamza Mustafa Abdul Rafay |
| 5 | Texturing | 5 | Map the 2D images onto the 3D mesh to create a realistic and detailed representation of the object or scene. | Ali Amjad Abdul Rafay |
| 6 | Quality assessment | 6 | Evaluate the quality of the 3D reconstruction using metrics such as accuracy, completeness, and visual realism. | Ali Amjad Abdul Rafay |
| 7 | Output | 7 | Output the 3D reconstruction in a format that is suitable for the intended use | Hamza Mustafa |

| | | | | |
|--|--|--|---------------------------------------------------|-------------|
| | | | (e.g., STL for 3D printing or PLY for rendering). | Abdul Rafay |
|--|--|--|---------------------------------------------------|-------------|

Table 1 Roles & Responsibility

1.7.3 Gantt Chart

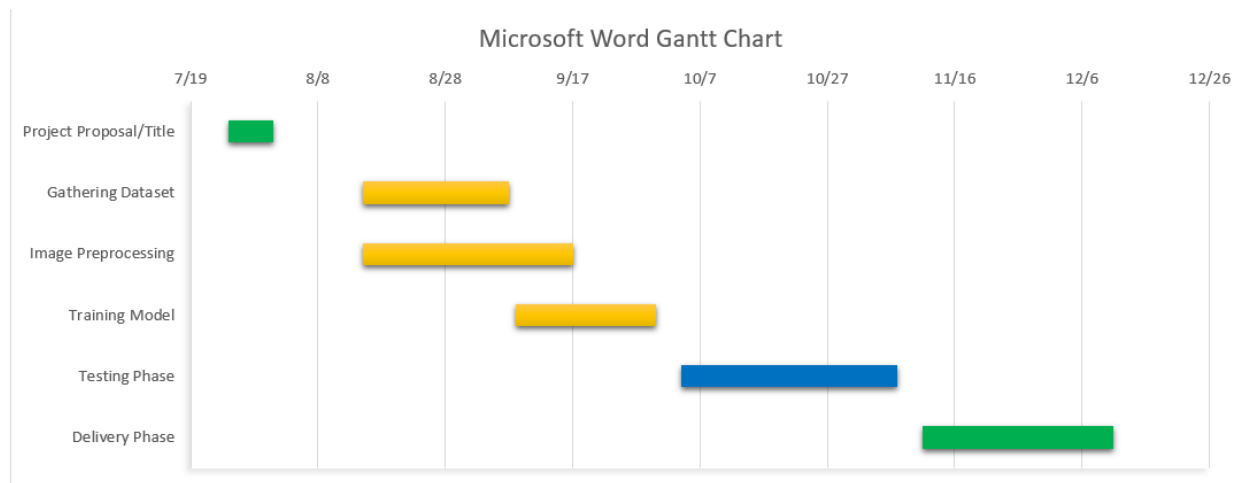


Figure 2 Gantt Chart

1.8 Report Outline

The outline of a report on 3D reconstruction from images would likely include the following sections:

1. **Introduction:** This section should provide an overview of the project, including its goals and objectives, as well as a brief summary of the methods and techniques used.
2. **Data Collection:** This section should describe the methods used to acquire the images, including the type of camera and lens used, the lighting conditions, and the techniques used to capture the images. It should also include details on how the images were pre-processed, such as cropping, resizing, and color correction.
3. **Image Processing:** This section should describe the steps used to prepare the images for reconstruction. This may include camera calibration, image rectification, stereo

correspondence, and other steps that are used to extract accurate depth information from the images.

4. **Reconstruction:** This section should describe the algorithms and methods used to create the 3D model, including details on the type of reconstruction (i.e. stereo, multiple view, depth from single image etc.), the parameters used, and the limitations and assumptions that were made.
5. **Conclusion:** This section should summarize the main findings of the project and its contributions to the field. It should also include suggestions for future work and improvements to the methods used.

1.9 Empathy Map

| | |
|---------------------------------------------------------------------------|------------------------------------------------------------------------|
| SAYS | THINKS |
| Where Should I Start? What Do You Think? I Want Something Reliable. | How it will Works? Why is This So Hard? I Want Something Awesome |
| DOES | FEELS |
| More Research Makes Small Decisions List Pros/Cons | Excited Nervous |

Table 2 Empathy Map

Chapter 2

Data Collection

2 Data Collection

2.1 Introduction

2.1.1 Sources of Data

We scrape the data from the given website istockphoto.com. The dataset is based on images and contains more than 150 images. We scrape the data using the beautiful Soup and Request module in Python.

2.2 Data preparation

The method of three dimensional mesh reconstruction is used to supplementing the 3D reconstruction or 3D scanning method, which only provide with point cloud in 3D model. However, the 3D models in mesh mode illustrate clearer appearance and can be used to obtain more information.

In data preprocessing we create meshes for converting it into 3d.

2.3 Data Storage

Since the dataset is not very large so we stored it in a folder. When we publish the website then we will store the images in firebase database as a backend whom end-user upload to our website.

2.4 Data Validation

The dataset is based on images so there is no outliers or missing values in it.

2.5 Annotation (if necessary)

Our project is to convert the 2d image into 3d. We used Annotation in detection or classification type of projects in which we will predict labels. Since our project is different so there is no annotation concept in it.

2.6 Data sampling (if necessary)

In data sampling we divide the data into 90-10 ratio. 90% for training and 10% for testing, because the dataset is not big that's why we divide this ratio and we get better results by applying these sampling ratios.

2.7 Data privacy and security

When the website is published data will be restricted to only authorized persons. Because data is stored in database so it will have kept confidential. This is how we safe our data.

Chapter 3

Data Preprocessing

3 Data Preprocessing

Data preprocessing for images involves a series of steps to prepare the images for use in computer vision applications or machine learning models. This can be used for several reasons: Improving model performance, Reducing complexity, object recognition, classification, segmentation, etc. Overall, data preprocessing is an important step in the machine learning pipeline that helps to ensure that the data is ready for use and that the model can perform effectively.

3.1 Description of the data cleaning process

The image-based dataset cleaning process involves several steps to ensure that the data is consistent, free of errors, and suitable for use in a machine-learning model:

- Data quality check
- Duplicate removal
- Data balancing
- Missing data imputation
- Outlier detection
- Data annotation
- Normalization

3.2 Identification of data issues:

Duplicate images in the dataset that can lead to overfitting and reduced model performance. Missing images or missing labels for a specific class can affect the performance of the model and the accuracy of predictions. Outliers in the data that are not representative of the underlying distribution can affect the performance of the model and lead to incorrect predictions. These are major issues during this process because we manually scratch data from different websites.

3.3 Handling of missing values:

Removal: We can simply remove some images that have missing data like (blur image, incorrect format) but it may result in loss of information or dataset can be reduced.

Imputation: We replace the missing values with estimated values. There are several imputation methods that can be used, including:

- **Mean imputation:** Replacing missing values with the mean value of the feature for all other samples.
- **Median imputation:** Replacing missing values with the median value of the feature for all other samples.
- **Mode imputation:** Replacing missing values with the most common value of the feature for all other samples.
- **Predictive imputation:** Using machine learning models to predict missing values based on the values of other features in the same sample.

Interpolation: Estimating missing values based on the values of other samples in the dataset. This can be done using linear interpolation, polynomial interpolation, or spline interpolation.

3.4 Handling of outliers:

Outliers are data points that are significantly different from other points in the dataset and can affect the performance of machine learning models.

- **Removal:** We remove the samples that contain outliers. This can be done by setting a threshold for the values of each feature and removing samples that exceed this threshold.
- **Transformation:** We transforming the data to reduce the impact of outliers. This can be done using logarithmic or other non-linear transformations that reduce the impact of outliers on the distribution of the data.

3.5 Handling of inconsistencies:

Handling inconsistencies in image-based datasets is important for ensuring the quality and reliability of the data. Inconsistencies can arise from a variety of sources, such as errors in data collection or data entry, or differences in data formatting. We remove those images that having errors or incorrect format.

3.6 Data normalization:

Data normalization is the process of converting data into a standard format to ensure consistency and eliminate the impact of scale differences between different features in an image-based dataset.

- **Min-Max Normalization:** We scales the pixel values in the image to a specific range, such as 0 to 255. This helps to ensure that the data is consistent across images and eliminates the impact of scale differences between different images.

3.7 Data transformation:

Transformation is an important step in the preprocessing of image-based datasets, as it can help to correct inconsistencies, improve the performance of machine learning models, and ensure the quality and consistency of the data. Here are some common transformations that are applied in image-based datasets:

- **Resizing:** This involves changing the size of the image to a standard size to ensure that images are consistent and the model is not affected by the differences in size between images.
- **Data Augmentation:** This involves generating additional data samples by applying transformations such as rotation, scaling, and flipping to the original images. This can help to reduce overfitting and improve the robustness of the model.

3.8 Conclusion:

The key findings from the data cleaning process identification of missing or incorrect data, Identification of outliers, Normalization of data, Transformation of Data. We face many challenges during maintain the quality of data also during dataset collection it is very difficult because we create dataset manually.

Chapter 4

Data Exploration

4 Data Exploration

The dataset used for 3D reconstruction from images comprises a collection of up to 200 images, all depicting subjects in a standing position. These images have been carefully curated and selected to capture various aspects necessary for the reconstruction process. Each image represents an individual observation within the dataset.

4.1 Description of Dataset

The dataset used for 3D reconstruction from images comprises a collection of up to 200 images, all depicting subjects in a standing position. These images have been carefully curated and selected to capture various aspects necessary for the reconstruction process. Each image represents an individual observation within the dataset.

4.2 Descriptive statistics

Since the dataset primarily consists of image data, traditional descriptive statistics may not be directly applicable. However, there are several metadata attributes that can be analyzed to provide insights into the technical characteristics of the dataset. Specifically, we can explore the resolution, aspect ratio, and file size of each image to gain a deeper understanding of the dataset's composition.

Examining the resolution of the images allows us to understand the level of detail captured in each photograph. Higher resolutions often indicate finer details, while lower resolutions may imply a more generalized representation. Analyzing the aspect ratio provides information on the relative proportions of the images, helping to assess if any distortions or variations are present within the dataset.

Additionally, investigating the file size of each image offers insights into the overall storage requirements and resource allocation necessary to handle the dataset. Given that each image ranges from 2 to 3MB, the total folder size of 100+ MB indicates a substantial collection of images, requiring adequate storage and processing capabilities.

By exploring these metadata attributes, we can develop a better understanding of the technical aspects of the dataset, laying the groundwork for subsequent analyses and 3D reconstruction efforts.

4.3 Visualizations

In the case of image datasets, visualizations become particularly important in understanding the underlying data distribution. Visualizations can include displaying sample images from the dataset, providing a glimpse into the subjects' standing positions.

Here is the grid graph of sample images.

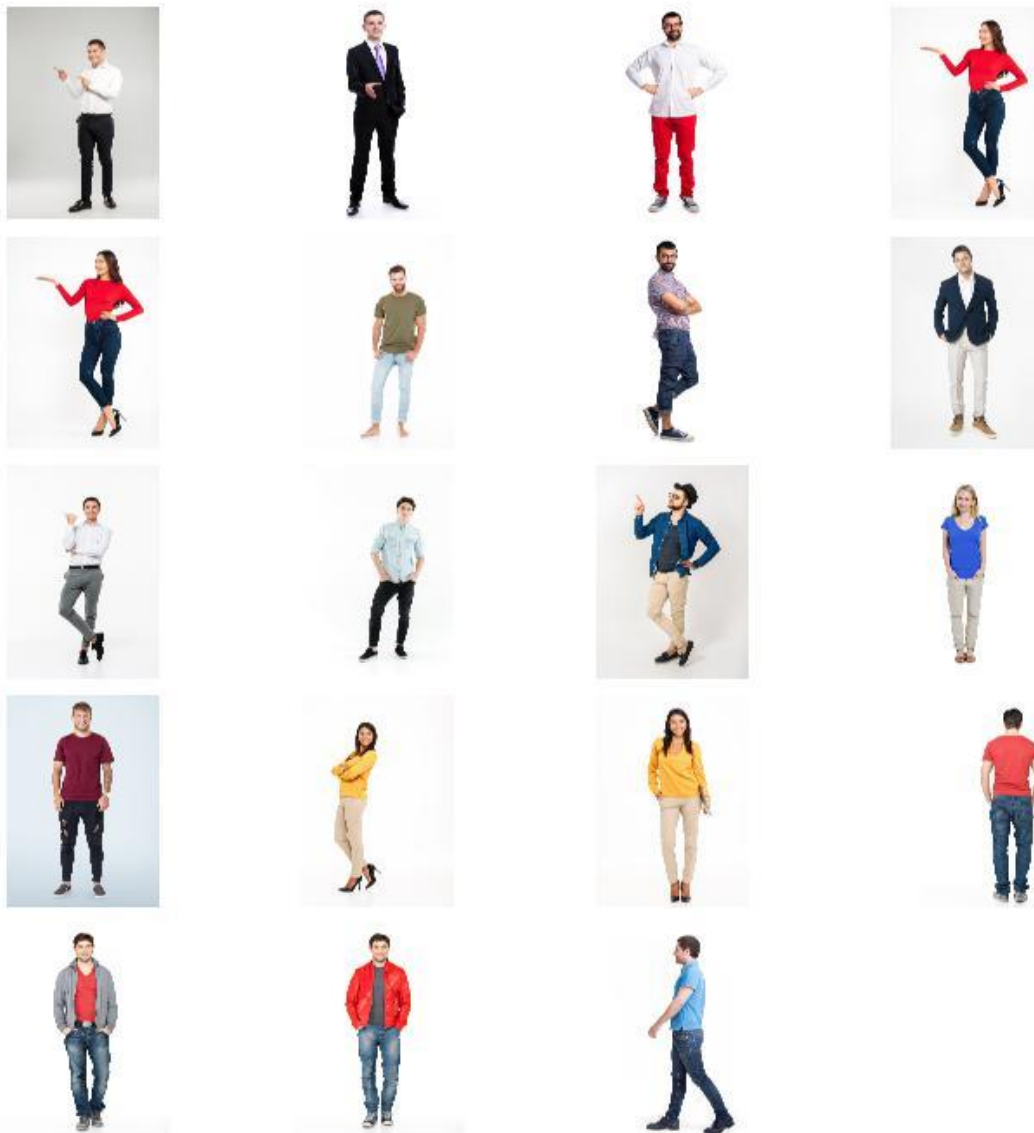


Table 3 Grid Graph of sample dataset

Here is the heatmap graph of sample images.



Table 4 Heat map Graph of sample dataset

Chapter 5

Purposed Approach

5 Purposed Approach:

This Chapter includes brief descriptions of the methods you are going to use for the project

- Data Model
- Which model you are going to use and why
- Architecture diagram of your project

5.1 Data Model:

The data model for this project involves the dataset we'll use to train and evaluate your PIFuHD-based model. In the context of PIFuHD, we have a dataset that contains single-view images of humans along with their corresponding ground truth 3D representations.

Here's a breakdown of the data model:

1. **Single-View Images:** We have a collection of images where each image represents a single view of a human subject. These images can be captured from different angles or viewpoints, showcasing various poses and appearances. The images should have a consistent format and resolution.
2. **Ground Truth 3D Representations:** Along with the single-view images, our dataset should provide ground truth 3D representations of the humans depicted in those images. These ground truth representations serve as the reference or target 3D shapes that you want your PIFuHD-based model to learn and reconstruct accurately.
3. **Correspondence between Images and 3D Representations:** It's crucial to establish a clear correspondence between the single-view images and their corresponding ground truth 3D representations. Each image in the dataset should be paired with the correct 3D representation, ensuring that the model learns to associate the visual information from the images with the corresponding 3D structure.

4. **Data Split:** To evaluate the performance of our model, it is common to divide the dataset into separate subsets for training, validation, and testing. The training set is used to train the model, the validation set is used to fine-tune hyper parameters and monitor performance during training, and the testing set is used to assess the model's generalization and accuracy on unseen data.

5.2 Model Selection:

PIFuHD (Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization) is a specific model architecture. It is designed to address the limitations of low-resolution reconstructions in the original PIFu method. PIFuHD utilizes a multi-level framework that includes a coarse reconstruction step and a fine reconstruction step to generate high-resolution 3D human models with detailed surface geometry. That's why we use this model in our project so we can perform the reconstruction of images efficiently.

5.3 Architecture diagram:

Here is the architecture diagram of our project.

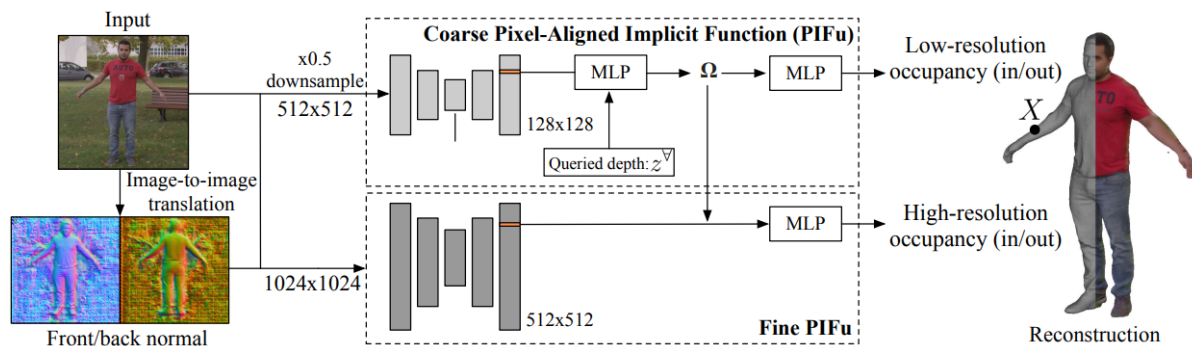


Table 5 Architecture Diagram

6 References

- [1] J. Konrad, "2D-to-3D Image Conversion by Learning Depth.," 16 July 2012..
- [2] R. MAMDOUH, "Converting 2D-Medical Image Files “DICOM” into 3D- Models, based," Februray 2020. .
- [3] A. Yuniarti, "A Review ofDeep Learning Techniques for 3D," 2019.
- [4] D. Yu, "Automatic 3D building reconstruction from multi-view aerial images," 13 November 2020. .
- [5] X.-F. Han, "Image-based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era," 1 May 2021.