I. INTRODUCTION

Three-dimensional (3D) reconstruction is one of the most important topics to be discussed. 3D reconstruction is a technology that can reconstruct images into 3D objects using computer vision technology. The main goal of 3D reconstruction is to create accurate 3D digital models of real-world physical objects. These digital renderings have a broad range of applications, most notably in the field of augmented reality, where they perfectly integrate the virtual and real worlds, enabling a new age of immersive new experiences.

In general, this technology is used in various fields to improve or update building architecture, such as in building mapping or space planning [17], [18]. In addition, 3D reconstruction technology is also used in healthcare, such as threedimensional computer tomography imaging [19], in creating replicas of complex objects such as statues or historical objects that have been destroyed [20], and in creating 3D animations [21]. At its core, 3D modeling requires certain skills and is still manual in nature. To reconstruct 3D objects from 2D images automatically and in detail requires such expertise, hence the need for a method that can transform 2D images into 3D objects automatically. However, 3D reconstruction technology also has some challenges, such as limitations in recognizing complex or unusually shaped objects [3].

Machine learning as well as deep learning technologies have been used in the development and improvement of recent technology. The reconstruction of the entire human body, including the form of the head, hands, torso, and legs as well as the position and style of clothing, is one use of this technology. Human reconstruction can be implemented in various fields such as human movement simulation [4], [5], posture identification [14], virtual try-on [15], and even for medical purposes [19]. The solution proposed is to utilize Deep Learning technology to construct an accurate and more realistic 3D human model based on captured 2D images.

The utilization of deep learning was preferred as a possible solution due to its capability to learn complex patterns and create more accurate and realistic models. In fact, deep learning has proven successful in applying 3D human reconstruction in several studies. Performing a human 3D reconstruction is a challenging task where researchers try to produce a 3D rendering result that provides details that resemble the original object. This is a challenge because the human body shape is asymmetrical and there are some extreme poses and clothing shapes that are difficult to predict [1], [2], [4]–[9].

There have been several applications of deep learning in previous literature such as Pixel2Mesh which used a graphbased Network [3], Tex2Shape [6] uses a pix2pix convolutional neural network (CNN) by applying a skinned body model [10] as its human model and DeepHuman [7] which uses a CNN architecture consisting of three other modules. However, none of them have been able to overcome challenges such as extreme human pose estimation and extreme clothing shapes.

In this study, we propose several refinement methods that become the focus of the study. These refinements include increasing the input image resolution from 512x512 to a higher resolution of 1024x1024, reconfiguring the sampling points, and making adjustments to the network architecture.