

Deep Learning Methods in Image Matting

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Lingtao Huang , Xipeng Liu , Xuelin Wang , Jiangqi Li , Benying Tan

Image matting is a fundamental technique used to extract a fine foreground image from a given image by estimating the opacity values of each pixel. It is one of the key techniques in image processing and has a wide range of applications in practical scenarios, such as in image and video editing.

image matting

trimap

image editing

medical imaging

cloud detection

1. Image Matting

Image matting ^[1] is a technique that is used for extracting the fine foreground of images, which is a computer vision task that has a wide range of application scenarios. The foreground of an image may include entities such as humans or animals, with delicate and complex edges such as hair as well as transparent objects such as glass, light bulbs, or water. These elements can be difficult for computers to accurately recognize. Image matting calculates the opacity of each pixel in an input image to obtain the alpha matte, which allows for the separation of the foreground from the background. The foreground can be composited with any background image to obtain a new image, as shown in **Figure 1**.

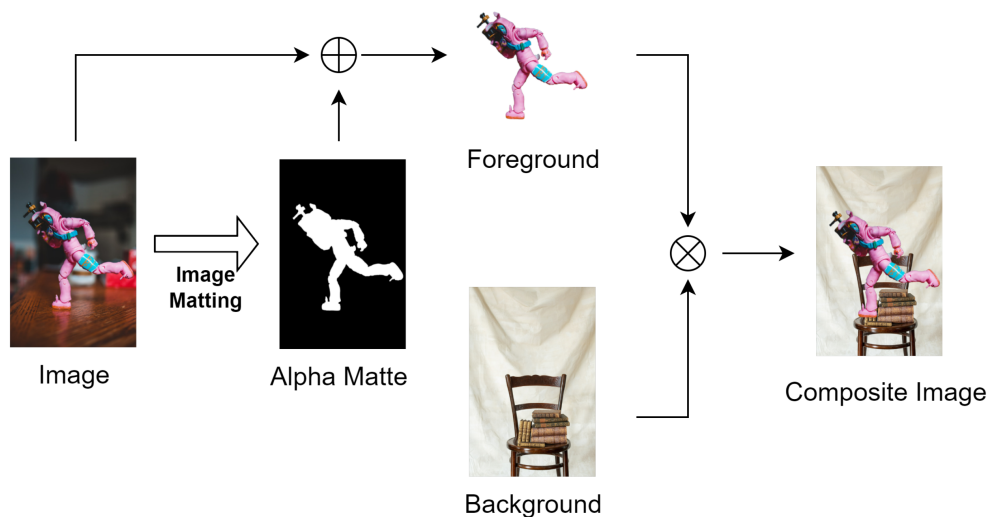


Figure 1. Specific process of image matting. \oplus indicates the foreground extraction operation, and \otimes indicates the image composition operation.

Chroma key matting ^[2] is a classic image matting technique that can be used to obtain the foreground from a solid background by adjusting the colors of pixels in the background to make them transparent. This technique requires

a special shooting environment, which is considerably limited in practical applications. Therefore, researchers have focused on extracting alpha mattes from images that have natural scenes as the background. However, image matting in natural scenes loosens the constraint of setting a solid color background, which leads to a decline in alpha matte accuracy. Mathematically, the problem of image matting can be expressed as follows:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i \quad (1)$$

where I_i denotes the RGB value at pixel i of the input image, α_i denotes the opacity value at pixel i ranging between 0 and 1, F_i denotes the RGB value at pixel i in the foreground, and B_i denotes the RGB value at pixel i in the background.

In image matting, the input image can be represented as a linear combination of foreground and background, with each pixel having only three known variables (RGB values) but seven unknown variables to be solved, as shown in Formula (1). Moreover, the definition of the foreground in an image is not precise and varies depending on the intended use. Consequently, image matting is a highly ill-posed problem that typically necessitates the use of auxiliary inputs such as trimaps to provide additional information.

2. Trimap

The trimap technique was proposed by Sun et al. [3] in 2004. A trimap is a mask that contains a foreground region, a background region, and an unknown region, and the regions are represented by α

values of 1, 0, and 0.5, respectively. **Figure 2** shows the image and corresponding trimap. The trimap can either be manually provided by users or automatically produced during network computations. When using the trimap as the auxiliary input, the known foreground and background regions provide considerable prior information, which helps to narrow down the solution space for unknown regions. Researchers can design relevant algorithms by using the information of these known regions.

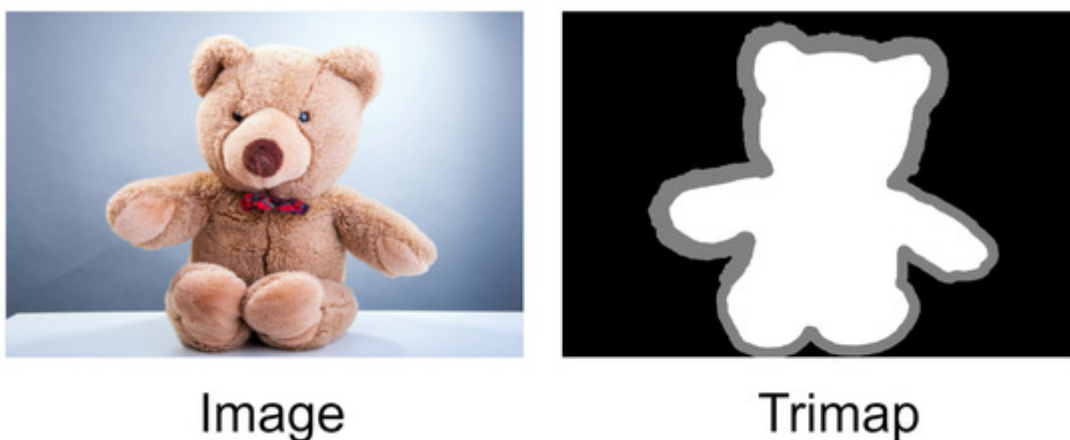


Figure 2. The image and its corresponding trimap are presented, where the foreground, background, and unknown regions of the trimap are denoted by white, black, and gray colors, respectively.

3. Distinguishing Image Matting from Image Semantic Segmentation

The results generated by image matting may appear similar to those of semantic segmentation; however, in reality, they are fundamentally different techniques. Semantic segmentation is a classification task that extracts the semantic information in the input image and then classifies the pixels individually to obtain the semantic mask of the input image. When semantic segmentation only segments the foreground and background, the binary nature of segmentation leads to a strict boundary near the foreground edge. Image matting is a regression task that involves estimating the opacity of each pixel in an input image, which results in the extraction of the foreground via the alpha matte. A comparison of the results of semantic segmentation and image matting is shown in **Figure 3**.

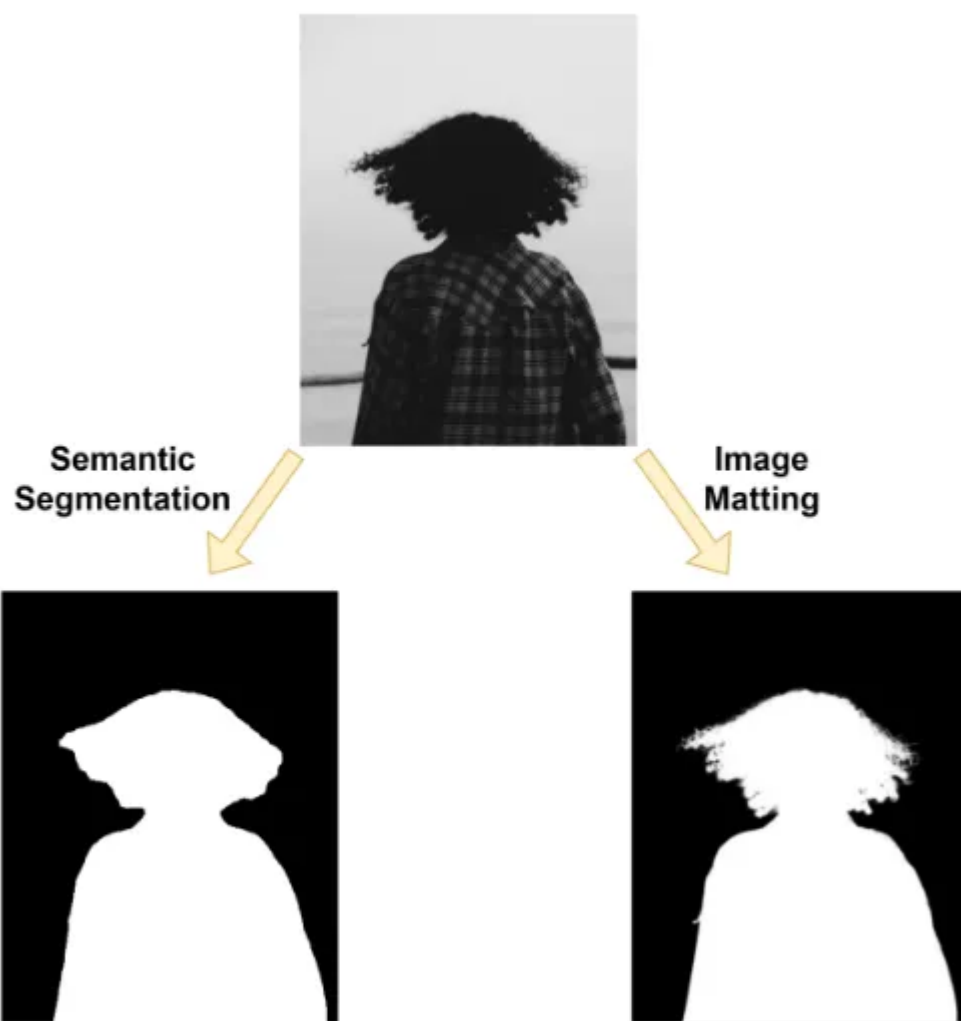


Figure 3. Comparison of the image semantic segmentation and image matting results.

4. Classification of Image Matting Methods

Over the years of research and development, researchers have designed a series of effective algorithms for various application scenarios of image matting; these algorithms can be categorized into three groups: sampling-

based, propagation-based, and learning-based methods. Sampling-based algorithms predict the opacity of the unknown region by collecting a series of pixels or pixel blocks from the known regions of the trimap. Propagation-based algorithms typically establish connections between adjacent pixels and then use an optimization strategy to propagate the opacity information of the known regions to the unknown regions in order to predict the opacity of each pixel in the unknown region. Learning-based algorithms learn the features of the image by using a considerable amount of data and use these features to predict the opacity. As deep learning algorithms have already been applied to various visual tasks and have completely surpassed traditional learning-based algorithms, they are gradually being introduced into image matting.

References

1. Smith, A.R.; Blinn, J.F. Blue screen matting. In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, LA, USA, 4–9 August 1996; pp. 259–268.
 2. Mishima, Y. Soft Edge Chroma-Key Generation Based Upon Hexoctahedral Color Space. US Patent 5,355,174, 11 October 1994.
 3. Sun, J.; Jia, J.; Tang, C.K.; Shum, H.Y. Poisson matting. In ACM SIGGRAPH 2004 Papers; ACM: New York, NY, USA, 2004; pp. 315–321.
-

Retrieved from <https://encyclopedia.pub/entry/history/show/102950>