1. Problem Statement

本文试图解决的研究问题是什么?这个问题重要吗?为什么?

当前的技术状态(即其他竞争方法/解决方案)是什么?

这篇论文的动机是什么?这篇论文的定位是什么?

如果有必要，你可以用图片来更好地说明研究问题。

这篇论文研究的问题是允许用户在没有绿幕等特殊背景的情况下实时的进行视频的背景替换，以提高娱乐性或增强隐私性。

背景替换可以通过分割或消光来实现。基于分割的方法运行很快，但是在分辨率和细节方面有欠缺。传统的消光方法虽然有更好的表现，但是无法在高分辨率下实时运行，并且经常需要用户手动输入。无需任何外部输入的消光方法有很多限制。基于已知的自然背景进行消光的方法效果较好但是无法做到实时和高分辨率。

而本文提出的方法可以实时的高分辨率的进行背景替换。

2. Summarize the paper’s main contributions

这篇论文声称有哪些贡献?

这篇论文有什么新内容吗?这篇论文有什么创新之处?

作者是否夸大了他们的贡献，如果是，又是如何夸大的?

本文提出的方法能够在高分辨率下实时提供高质量的背景替换，并且不需要用户手动输入。该方法不是传统的分割或消光，而是使用两个神经网络以及一张额外捕获的背景图片。

然而，达成实时运行并获得好的效果需要硬件支持。可以通过降低消光质量来平衡硬件的限制。

3. Method and Experiment

他们提出了什么方法/想法/见解? 解释他们方法的关键(关键思想/洞察力/聪明才智)。

主要观点/论点是什么?

作者如何证实他们的主张?进行了哪些实验?这些实验的主要结果或发现是什么?

该方法的训练基于多个数据集，并爬取了网络上的图片用以生成数据集。

本文提出的方法基于两个神经网络。第一个是基础网络，用以对整个图像提供低分辨率的消光。这是一个全卷积编解码器网络（fully-convolutional encoder-decoder network），有Backbone, ASPP和Decoder三个模块，估计alpha matte和foreground residual。第二个Refinement Network对基础网络的结果中预测误差较高的前k个部分进行高分辨率的细节处理，从而提供较高的质量并提高运行速度。

在训练时，作者们使用了多种数据增强技术以避免过拟合和适应现实情况，并在每个小批中随机裁剪以支持任意分辨率和宽高比。而在验证时使用了合成数据集和真实拍摄的图片。

作者通过实验发现，使用AIM数据集由于低分辨率和样本数量少会使结果恶化。在基础网络中，使用ResNet-50作为encoder backbone较为高效。

4. Critical Analysis

**4.1. Are the paper’s contributions** **significant?**

贡献/改进是微不足道的、增量的吗?

为什么之前的努力失败了?

我认为本文的贡献是显著的。针对传统的基于分割或trimap-based matting以及其他方法在抠图质量、速度上存在不足，或需要额外输入，无法实时的在高分辨率下工作。针对这些限制，本文中的方法改变思路，使用神经网络进行消光，从而达到了实时的、高分辨率的、高质量的背景替换。

另外，获取大量的高质量的高分辨率的抠图数据集是很困难的。他们整理并公布了许多数据集，这对以后研究这一课题的人帮助很大。

**4.2. Are the authors’ main claims valid?**

他们是否令人信服地证实了自己的主要观点?

他们的论点，推导，实验有漏洞吗?

作者们令人信服地证实了自己的主要观点。

他们详细介绍了他们用的神经网络的结构以及使用的数据集，描述了实验的过程，并且从展示的结果图片来看，对比其他方法，他们提出的方法效果好于实时matting方法以及部分非实时的方法。他们也做了用户研究，结果显示大部分人认为他们提出的方法效果更优。

**4.3. Limitation and weaknesses**

他们的方法有什么限制/弱点吗?可以做些什么来改进这项工作?

你会如何解决/克服他们的弱点?

他们的方法需要捕获额外的背景照片，并且需要保证背景只有小幅度的移动。并且要做到实时的高分辨率的高质量的背景消光需要较好的硬件支持，这是受限于当前的硬件，可以通过降低消光质量来进行平衡。另外，当主体在背景上大量投影或者颜色与背景相近，以及背景有强烈纹理时，该方法可能会失效。

**4.4. Extension and future work**

你会建议作者做哪些额外的实验来强化结果?

你能想到论文中提出的方法/想法(假设有效)的其他可能应用吗?

未来可能的作品是什么?

我认为额外的工作主要针对于上面提出的限制和弱点。对于需要额外的背景图片以及背景不能大幅度移动的问题，由于主体和背景的相对运动可以让被遮挡的部分背景在不同的时刻显示出来，我认为可以尝试通过对比相邻的几帧图片，运用图像相似度匹配方法和识别算法（如人体识别）辅助识别出大概的主体和背景位置，从而对背景进行拼接并对背景消光。而对于其他的在某些背景下算法可能失效的问题，可能需要更多更好的训练样本和进行额外的训练。

**4.5. Conclusion and personal reflection**

首先，对本文进行总结。

那么，如果让你来解决这个研究问题，你会有什么不同的做法?还有别的解决方案吗?

最后，用一句话总结你从阅读本文中学到的东西。

这篇论文是令人兴奋的。作者们提出了解决背景替换的新方向，即使用二层的神经网络，从而做到实时的高分辨率的高质量的处理，这是一个很大的进步。

这个神经网络的第一层对整个图像进行粗略的处理，第二层对第一层的结果中预测误差较大的部分进行细节处理，从而提高质量和速度，这对我是一个启发。

对于这个问题，我可能会在神经网络的基础上结合图像相似性算法和人体识别算法对背景进行识别和拼接，从而提高对于移动背景的适应性。