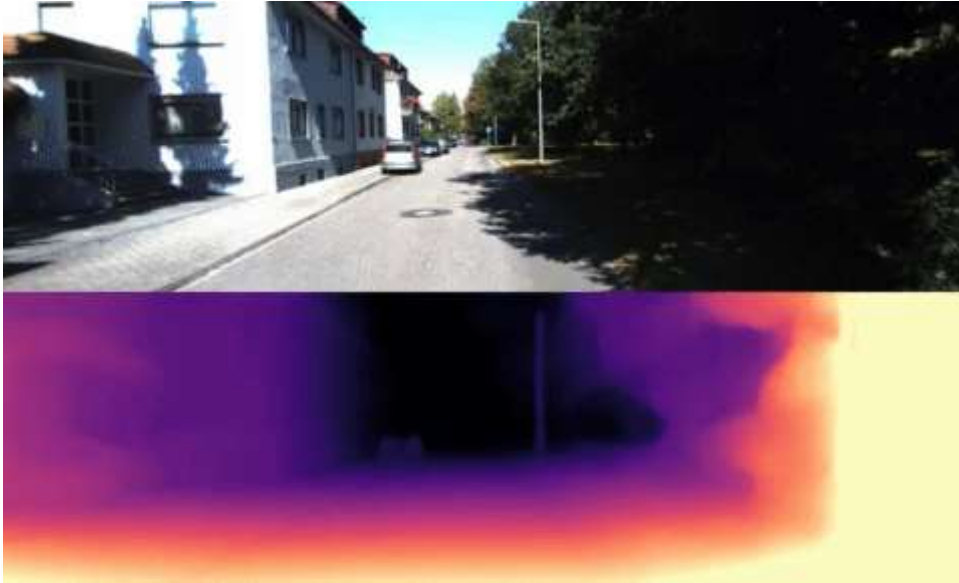


▼ 基于深度学习的视差估计方法

▼ Monodepth2

Monodepth2是一个用于单目深度估计的预训练模型和代码库，它基于深度学习技术。它可以从单个RGB图像中预测场景的深度信息。Monodepth2使用自监督学习的方法，通过使用成对的图像和视差地图进行训练。



模型代码: <https://github.com/mrharicot/monodepth>

```
!git clone https://github.com/nianticlabs/monodepth2.git
import os
os.chdir("/content/monodepth2")
!pwd

Cloning into 'monodepth2'...
remote: Enumerating objects: 180, done.
remote: Total 180 (delta 0), reused 0 (delta 0), pack-reused 180
Receiving objects: 100% (180/180), 10.27 MiB | 18.64 MiB/s, done.
Resolving deltas: 100% (89/89), done.
/content/monodepth2

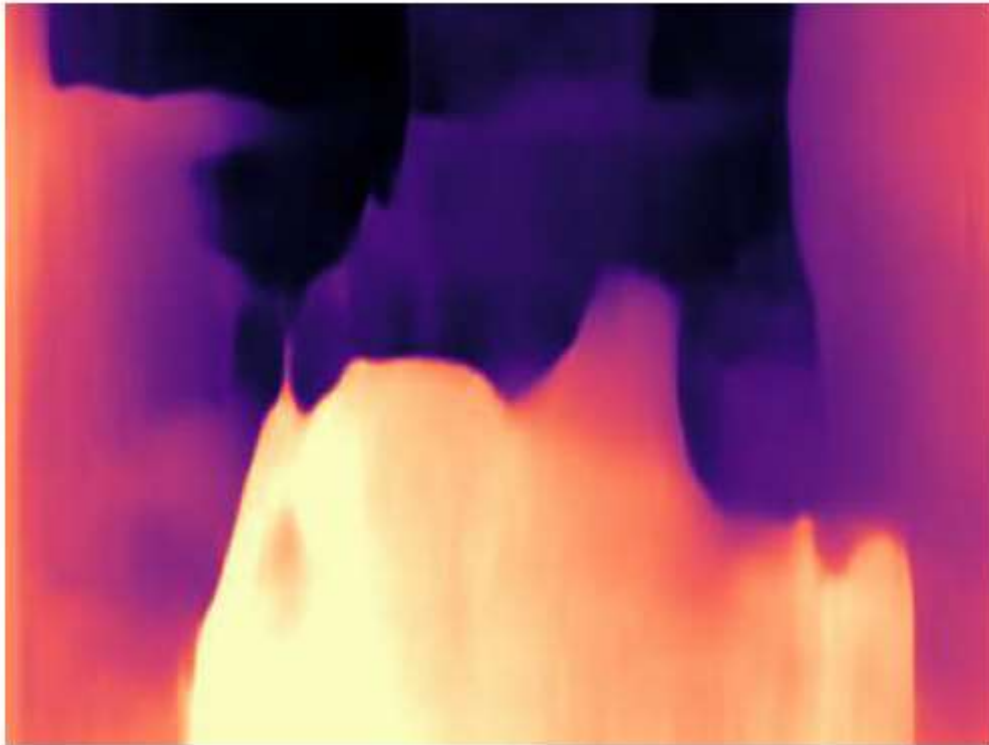
!python test_simple.py --image_path ./tsukuba_l.png --model_name mono+stereo_640x192

-> Loading model from models/mono+stereo_640x192
    Loading pretrained encoder
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:135: UserWarning: Using
warnings.warn(
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:223: UserWarning: Argum
warnings.warn(msg)
    Loading pretrained decoder
-> Predicting on 1 test images
    Processed 1 of 1 images - saved predictions to:
```

```
- ./tsukuba_l_disp.jpeg
- ./tsukuba_l_disp.npy
-> Done!
```

```
import matplotlib.pyplot as plt
from PIL import Image

image_path = './tsukuba_l_disp.jpeg' # 图像文件路径
image = Image.open(image_path)
plt.imshow(image)
plt.axis('off') # 关闭坐标轴
plt.show()
```

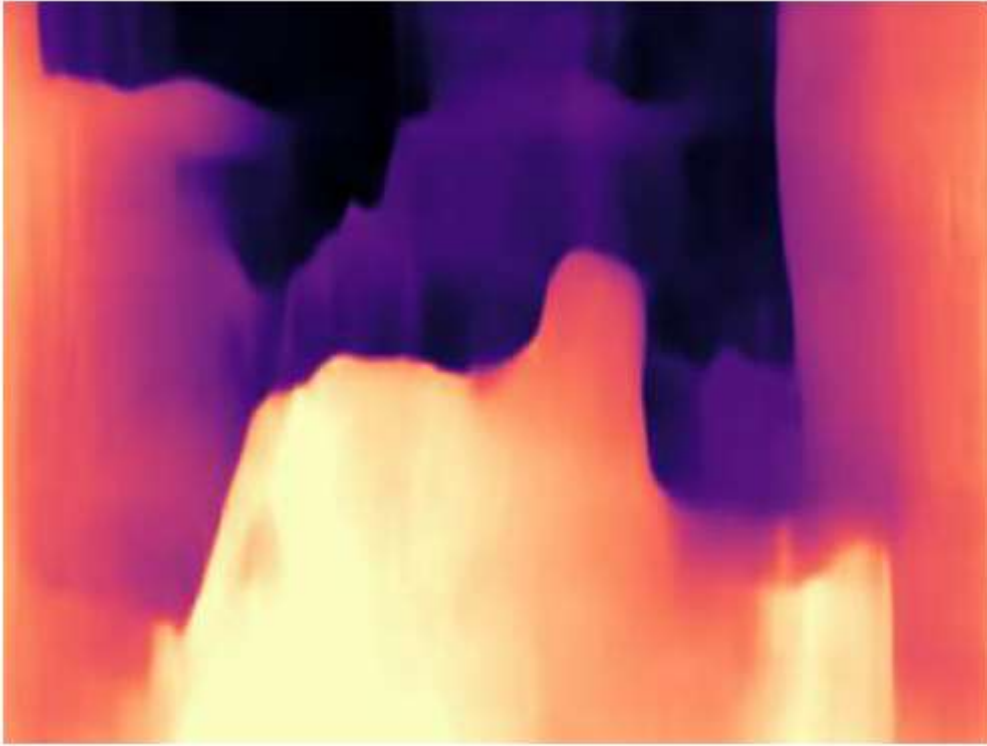


```
!python test_simple.py --image_path ./tsukuba_r.png --model_name mono+stereo_640x192 --pred_1
```

```
-> Loading model from models/mono+stereo_640x192
    Loading pretrained encoder
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:135: UserWarning: Using
warnings.warn(
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:223: UserWarning: Argum
warnings.warn(msg)
    Loading pretrained decoder
-> Predicting on 1 test images
    Processed 1 of 1 images - saved predictions to:
    - ./tsukuba_r_disp.jpeg
    - ./tsukuba_r_depth.npy
-> Done!
```

```
image_path = './tsukuba_r_disp.jpeg' # 图像文件路径
image = Image.open(image_path)
```

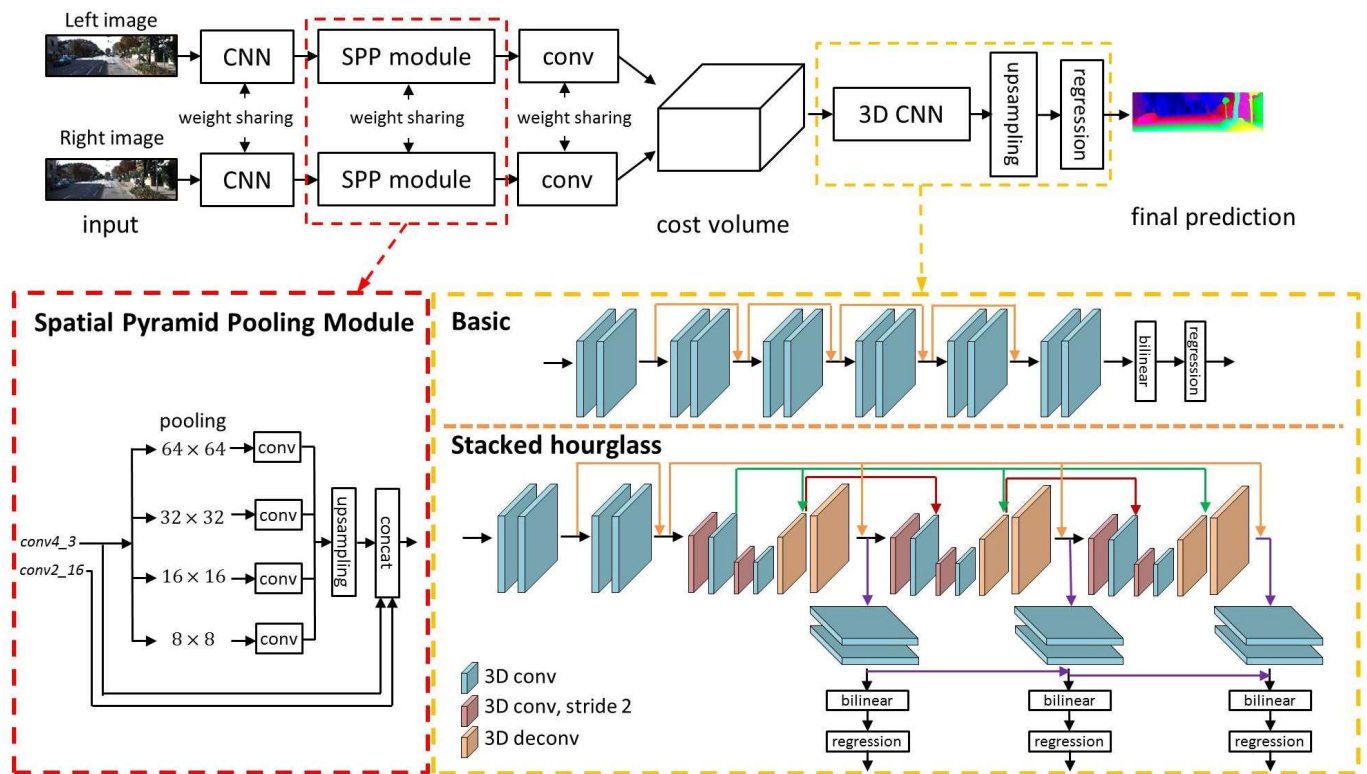
```
plt.imshow(image)
plt.axis('off')    # 关闭坐标轴
plt.show()
```



▼ PSMNet (Pyramid Stereo Matching Network):

这是一种基于金字塔结构的立体匹配网络，用于双目视差估计任务。它具有多尺度处理和全局代价聚合的能力。

论文：Chang, J. R., Chen, Y. S., & Hung, Y. P. (2018). Pyramid Stereo Matching Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 模型代码：<https://github.com/JiaRenChang/PSMNet>



然而，以上模型不支持更新版本的Pytorch，无法进行实验。

我们还搜索到了：

- GC-Net (Global Convolutional Network)：这是一种基于全局共形网络的双目视差估计模型，结合了卷积神经网络和图割算法。GC-Net通过最小化能量函数来估计视差图，具有较好的视差准确性和全局一致性。
- DispNet：这是一种经典的双目视差估计模型，包括 DispNetS（简化版本）和 DispNetC（完整版本）。它们使用卷积神经网络进行特征提取和视差预测，可以通过预训练模型进行使用。

但是并没有找到开源的代码仓库，就不在此进行尝试了。

总结

基于深度学习的视差估计方法是使用深度神经网络来学习从图像中预测视差信息的技术。这些方法通过训练神经网络来学习从输入图像对中提取特征，并预测每个像素点的视差值。以下是该方法的一些优点和缺点：

优点：

1. 高度准确：深度学习方法能够学习复杂的特征表示，并通过大量的训练数据来提高视差估计的准确性。
2. 端到端训练：深度学习方法可以通过端到端的训练方式直接从原始输入图像到视差图之间进行学习，简化了传统流程中的手工特征提取和处理步骤。
3. 对复杂场景的适应能力：深度学习方法可以处理复杂的场景和图像特征，能够在复杂背景、遮挡和纹理缺失等情况下进行鲁棒的视差估计。

缺点:

1. 需要大量的训练数据：深度学习方法通常需要大量的带有视差标注的训练数据来进行有效的训练，这可能限制了其在特定任务或应用中的使用。
2. 计算资源需求高：深度学习模型通常需要大量的计算资源和高性能硬件来进行训练和推断，这对于一些资源有限的设备或环境可能存在挑战。
3. 模型复杂性：深度学习模型往往较为复杂，包含大量的参数和层级结构，因此在部署和使用时需要考虑模型大小和计算复杂度。

✓ 1s completed at 1:22 AM

