

# Denoising Survey

---

## Table of Contents

- Denoising Survey
  - Problem
  - Dataset
  - Traditional methods
    - 单帧降噪
    - 多帧降噪
  - Deep Learning methods
    - DnCNN
      - Paper
      - Architecture
      - Code
    - FFDNet
      - Paper
      - Architecture
      - Code
    - CBDNet
      - Paper
      - Architecture
      - Code
    - PMRID
      - Paper
      - Architecture
      - Code
  - Some idea
    - 1.模型小型化
    - 2.变换域的探究
    - 3.数据集和损失函数
  - Referenced

## Problem

---

图像视频的采集、压缩、传输、显示等过程中都会发生不同层度的失真。我们ISP去噪关注的是图像采集过程中的一种普遍失真：噪声。ISP图像噪声主要有以下几种。

### 1. 光子散列噪声 (Photon shot noise)

当光子按照一定的概率分布撒到探测器上会因为量子涨落形成光子散列噪声，这种噪声无法通过改善相机设计来减少。光子散列噪声符合泊松分布。

### 2. 热噪声 (Thermal noise)

sensor中自由电子的布朗运动引起的噪声，服从高斯分布，又称高斯白噪声。

### 3. 读出噪声 (Read noise)

是读出电路前后电压随机波动的结果，服从高斯分布。

### 4. 暗电流噪声 (Dark current noise)

sensor由于热激发产生的电子形成电流被称为暗电流，温度越高暗电流也随之变高。暗电流噪声服从泊松分布。

---

## Dataset

---

[Waterloo Exploration Database](#)

[polyU-Real-World-Noisy-Images datasets](#)

---

## Traditional methods

### 单帧降噪

**速度快：**均值/中值/高斯滤波，小波滤波，Bilateral Filter，Guided Filter，Domain Transform Filter，Adaptive Manifold Filter等

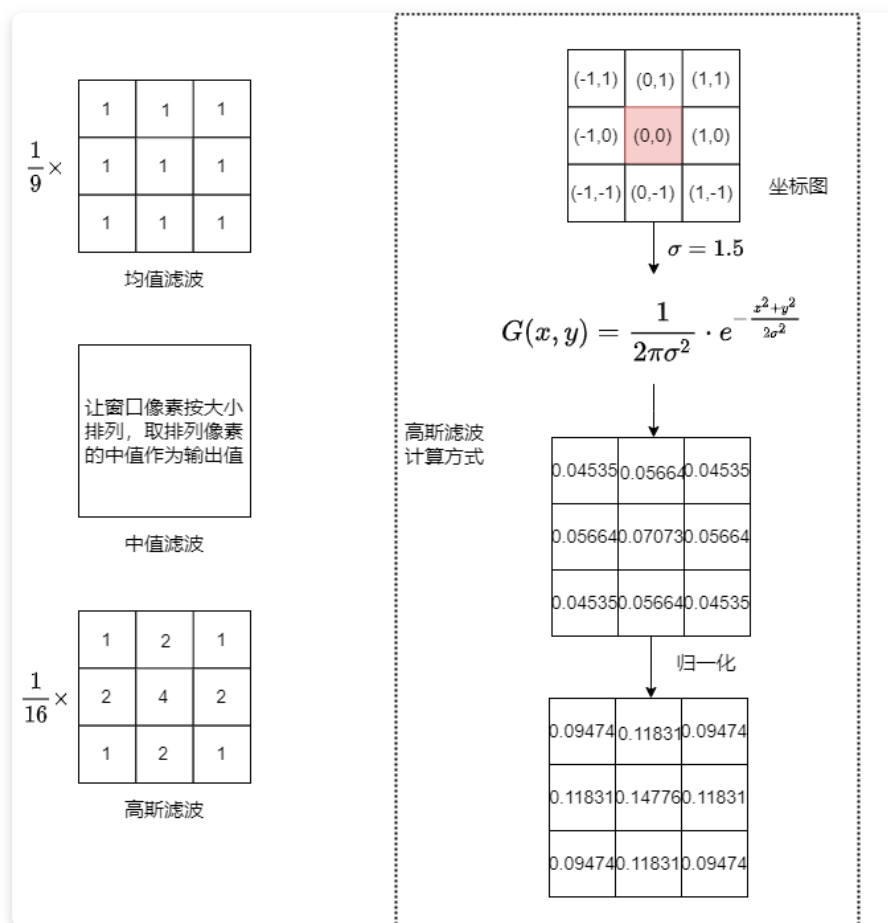


Figure 1. 均值/中值/高斯滤波

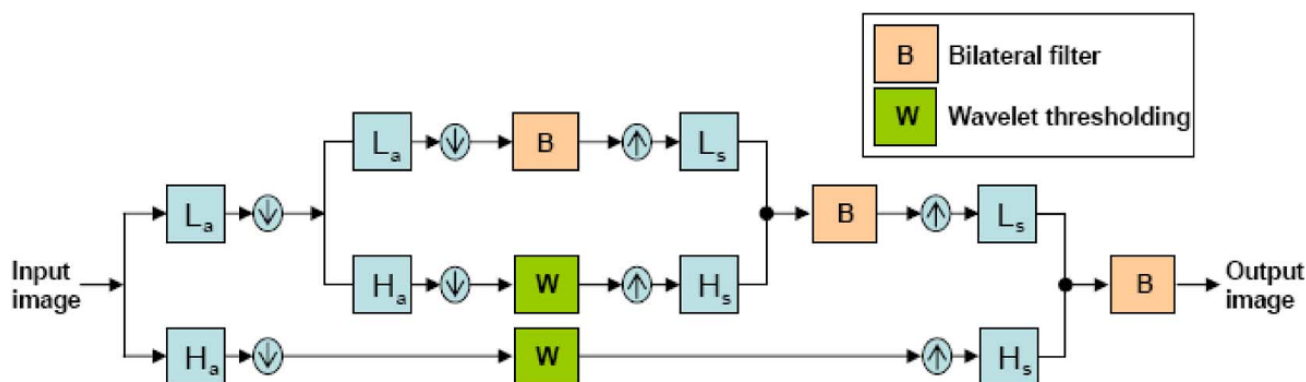
**双边滤波简介：**以高斯滤波来看，其intuition在于认为像素在空间域上变化缓慢，所以利用其像素点周围像素的距离赋予其权重。但是这种思路在图像边缘部分是不妥的，图像边缘部分像素值差异可能很大，继续按照空间距离来赋权会带来边缘模糊的问题。

双边滤波的处理思路是将同时考虑像素点的空域（domain）信息和值域（range）信息，即如果相邻像素值差距很大，即使距离很近也会被赋予很低的权重，从而解决边缘模糊问题。

[双边滤波的python和c++实现](#)

[双边滤波的一种快速实现](#)

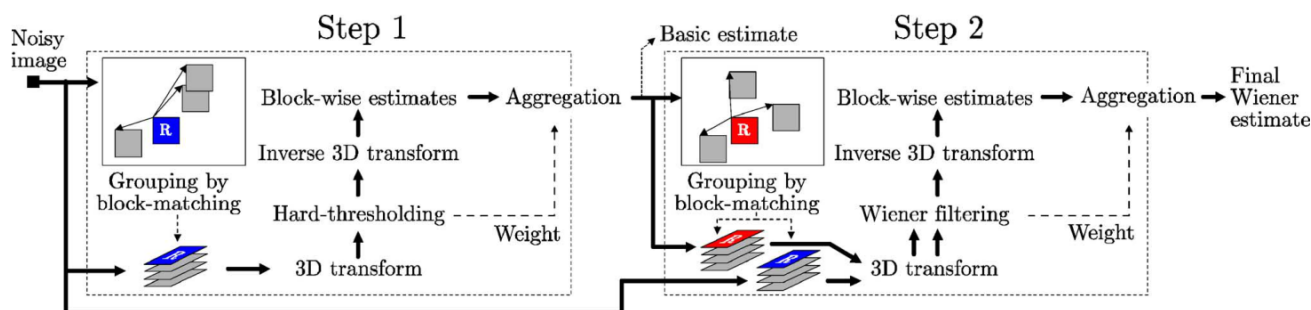
举例：[Multiresolution Bilateral Filtering for Image Denoising \(readpaper.com\)](#)



概述：双边滤波+小波(小波变换原理todo)

效果好：统计模型， Low Rank， 稀疏编码， 字典学习， 自相似， 自相似+变换域

举例：BM3D算法



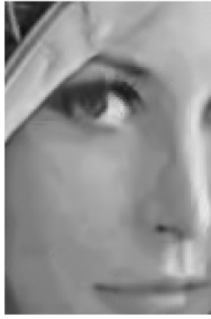
BM3D算法是对NLM（No local mean,非局部均值）算法的继承和发展。NLM算法的主要思想是，以目标区域的图片块（block）为基础，在图像中搜索跟目标区域相似的块，再根据相似度对这些块进行加权平均，求得降噪结果。其Intution主要基于图像中区域的自相似性。

BM3D算法Step1：首先也会寻找多个相似块，将这些相似块堆叠成3D数组。继而对这个3D数组的第三维即图块叠起来后，每个图块同一个位置的像素点构成的数组，进行DCT变换。接着采用硬阈值的方式将变换后的数组小于阈值的成分置为0。同时统计非零成分的数量作为后续权重的参考。然后将数据进行3D逆变换。最后这些图块逆变换后放回原位，利用非零成分数量统计叠加重，将叠放后的图除以每个点的权重就得到基础估计的图像。

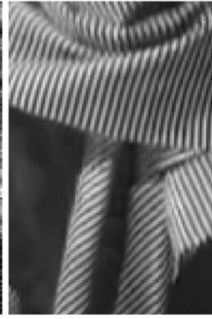
Step2：开始第二步中的聚合过程与第一步类似，不同的是，这次将会得到两个三维数组：噪声图形成的三维矩阵和基础估计结果的三维矩阵。接着两个3D矩阵都进行二维和一维变换。用维纳滤波将噪声图形成的3D矩阵进行系数放缩，该系数通过基础估计的3D矩阵的值以及噪声强度得出。最后与Step1中一样，只是此时加权的权重取决于维纳滤波的系数和噪声强度，得到最终结果。

论文中给出的效果如图：

(a) *Lena* (PSNR 32.08 dB)



(b) *Barbara* (PSNR 30.73 dB)



(c) *Camerman* (PSNR 29.45 dB)



(d) *Man* (PSNR 29.62 dB)



(e) *Boats* (PSNR 29.91 dB)



(f) *Couple* (PSNR 29.72 dB)



---

[pure Python implement of BM3D \(github.com\)](#)

## 多帧降噪

多帧降噪等同于增加曝光时间，提高信号量，提高信噪比。主要步骤包括多帧对齐+融合。

举例： [Google HDR+](#)

---

# Deep Learning methods

---

## DnCNN

Paper

[Link](#)

### Architecture

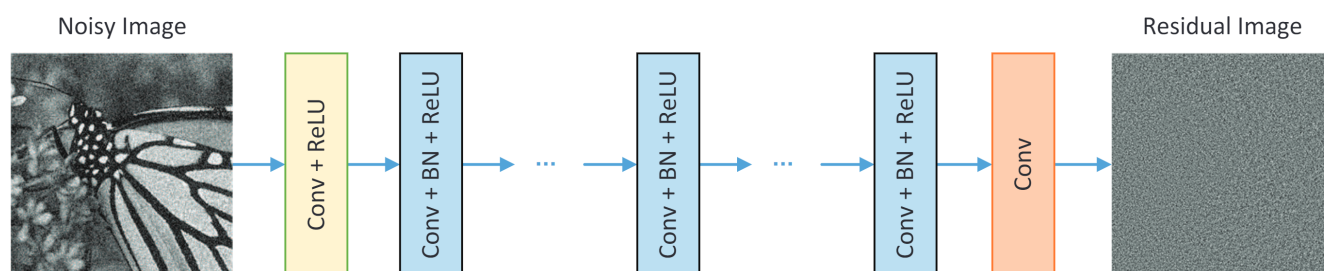


Fig. 1. The architecture of the proposed DnCNN network.

文章将噪声定义为 $y = x + v$ ，其中 $x$ 是无噪声图像， $v$ 是噪声， $y$ 为含噪图像。一般去噪的过程相当于输入 $y$ ，求 $x$ 。

DnCNN的网络结构相当于一个改版的vgg，其模型内部并没有skip connection。不过模型的输出却并不是 $x$ ，而是 $v$ ，相当于一个global的skip connection。文章采用MSE作为损失函数，数据集只有无噪声图像，人工增加加性高斯白噪声构建需要的输入和输出。

Code

[DnCNN-官方\(github.com\)](#)

[DnCNN-pytorch-第三方](#)

---

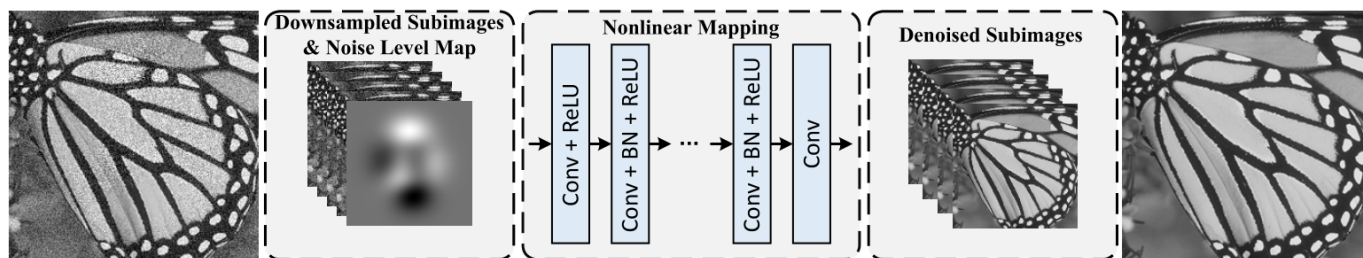
## FFDNet

Paper

[Link](#)

### Architecture





FFDNet是DnCNN基础上的一次升级，算法的backbone并没有改变，但是输入输出不同。FFDNet的输入为原始带噪图像降采样获得的四张子图以及一张由用户输入的参数 $\sigma$ 生成的噪声水平图像，输出为四张降噪后的子图，通过上采样获得最终的降噪图像。使用的损失函数仍然是MSE。

由原图生成4张子图过程就是pixel shuffle的逆操作，根据超参数 $\sigma$ 生成噪声水平图的过程尚需通过代码确认 (todo)。从设计上来看，该网络的处理能力更加灵活，似乎能够针对不同噪声水平来做不同层度的降噪处理，这点对于深度学习方法来说是值得**参考**的。

Code

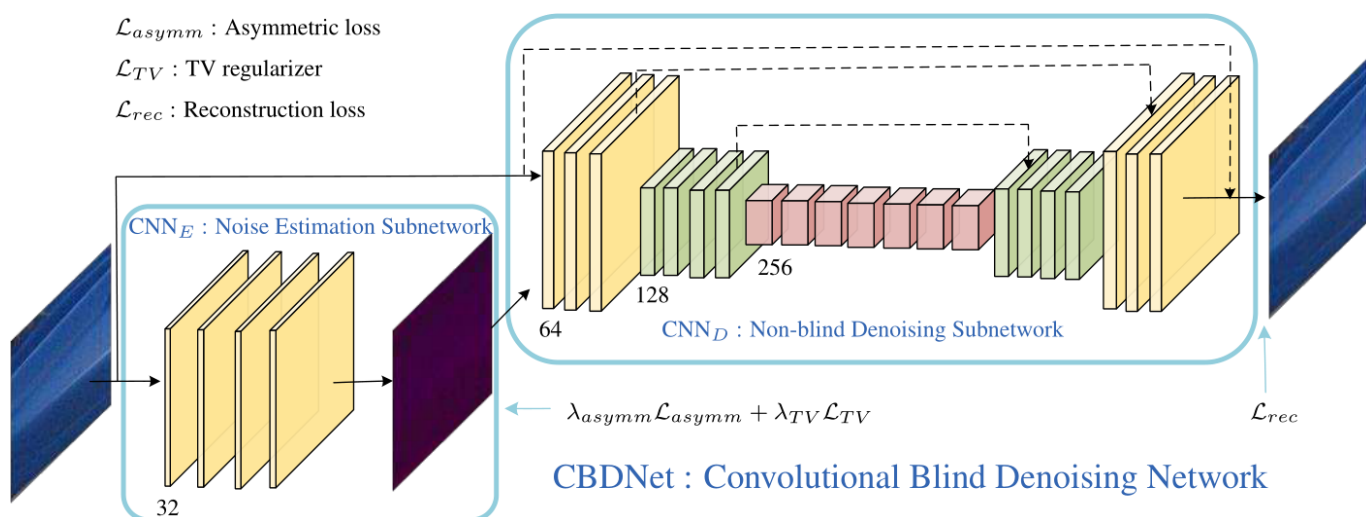
[FFDNet](#)

## CBDNet

Paper

[Link](#)

### Architecture



CBDNet是FFDNet网络基础上的又一次升级。对于FFDNet网络中的噪声水平为 $\sigma$ 的噪声图，CBDNet中采用一个五层的小网络估计出来，另一部分用去去噪的主要网络也由之前的VGG换成了一个16层的UNet。这种设计看起来有点自注意力机制的味道，跟SENet有点像，对自身的噪声水

平进行估计的方法比较符合人类解决问题的思路，从论文给出的结果看也是一种work的思路，可以参考。

此外论文设计了一种非对称的损失函数。（目前并未深究，但损失函数应该是个可探究的点，todo）

Code

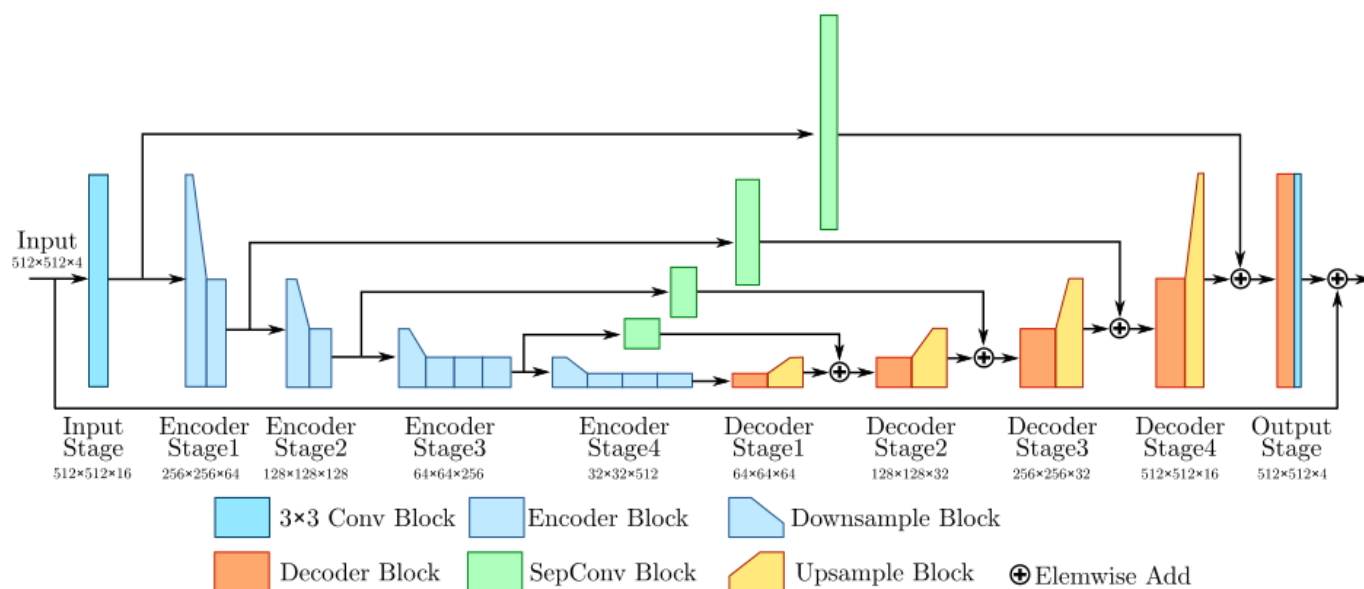
[CBDNet](#)

## PMRID

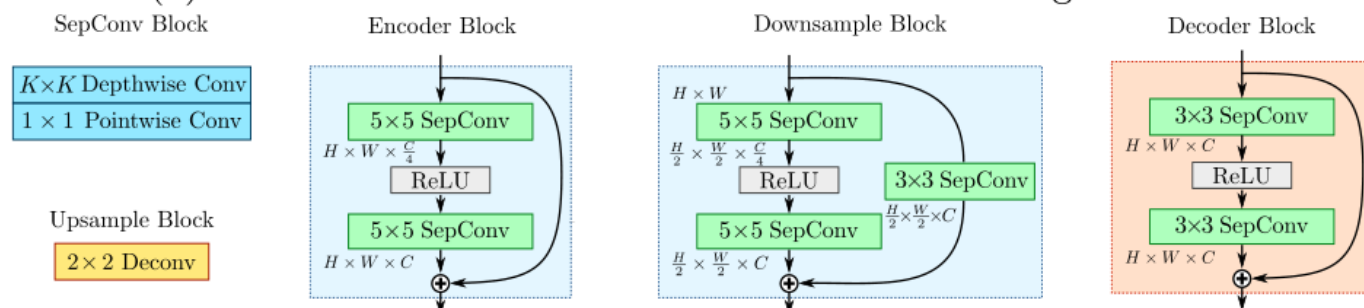
Paper

[Link](#)

## Architecture



(a) The U-Net-like overall structure of the denoising network.





PMRID是旷世推出的一个面向移动端设备的去噪模型。其网络模型是一个大体上是一个改版的UNet。模型主要特点有：1.使用 $5 \times 5$ 的separable conv增大感受野降低计算量；2.使用stride 为2的卷积来进行下采样；使用 $2 \times 2$ 的deconv来进行上采样；3.使用 $3 \times 3$ 的separable conv来进行skip connection时的通道匹配；

虽然其面向移动端的网络结构优化工作很棒，但这篇文章网络结构之外的变换部分非常值得**参考**。在硬件实际取景过程中，ISO是动态变化的，而不同的ISO带来的噪音强度是不一样的。所以文章采取了一种k-sigma变换的方式将噪声强度归一化，从而避免不同噪声强度对去噪算法带来的负担。

k-sigma变换公式推导：

(todo)

Code

PMRID

---

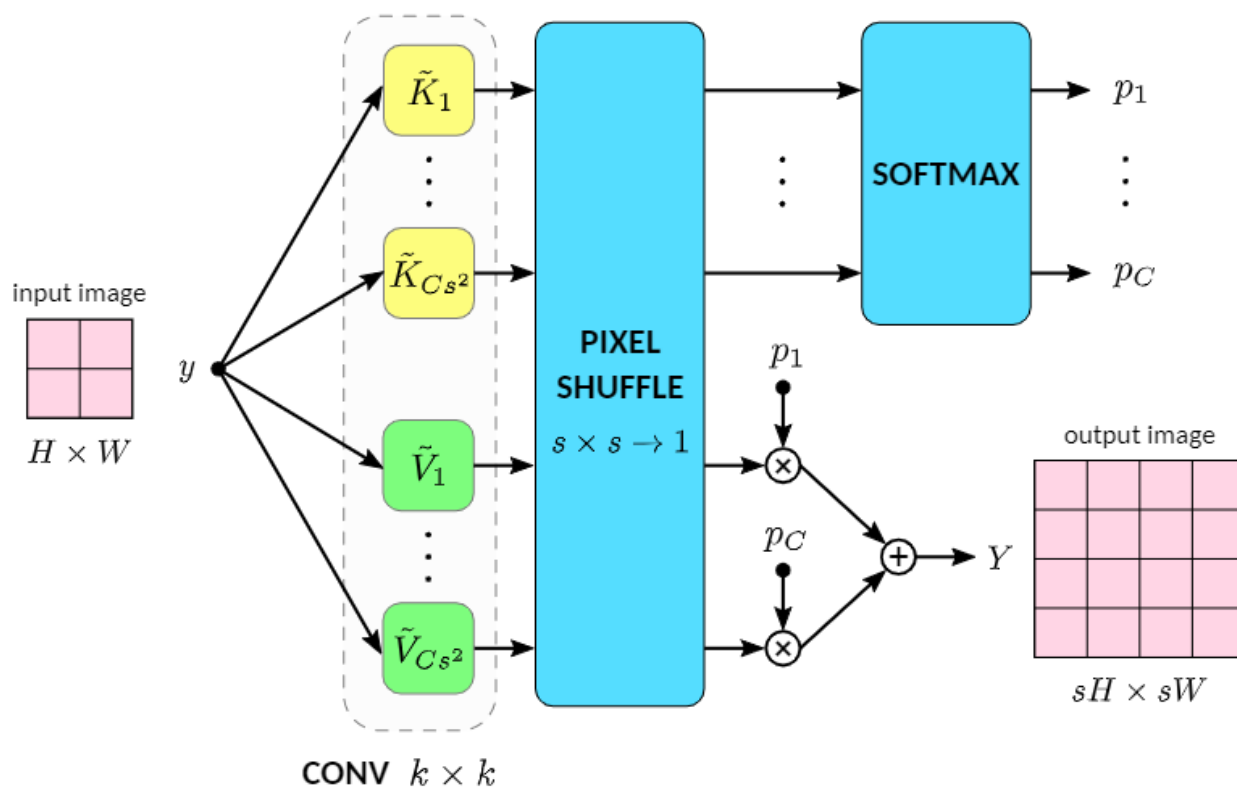
## Some idea

### 1.模型小型化

模型小型化对运算速度和硬件成本有着最直接的影响，如何在确保效果的情况下把模型做小是工作的重要方向。从其他论文中借鉴了一些不错的小型化模型结构供参考：

#### 1.edge-SR中的eSR-TM模型

##### a) eSR-TM



#### 2.dh\_isp

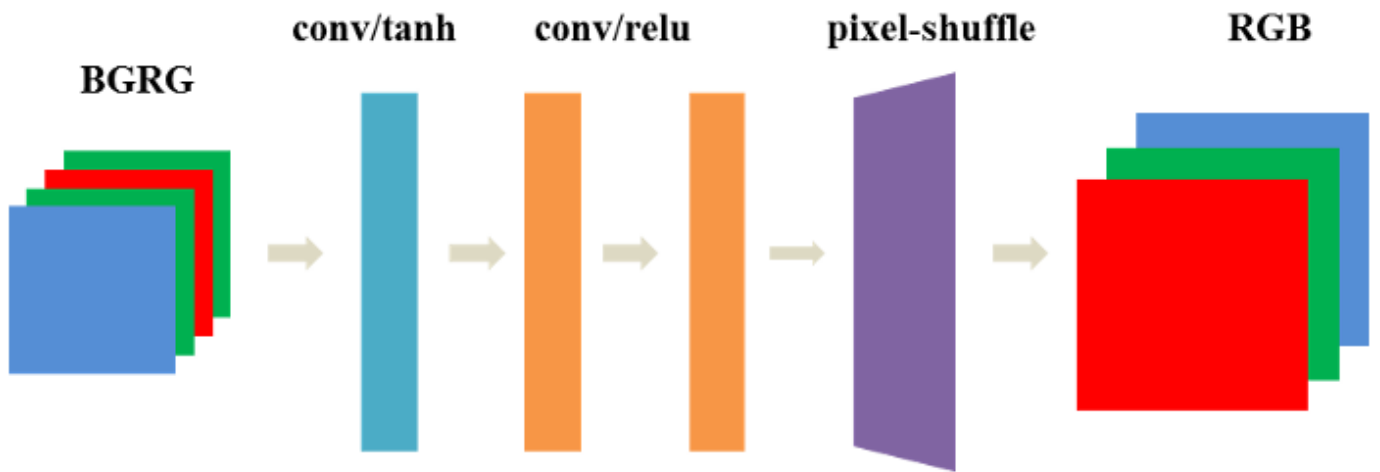


Figure 3. Smallnet architecture proposed by team dh\_isp.

### 3.AIISP

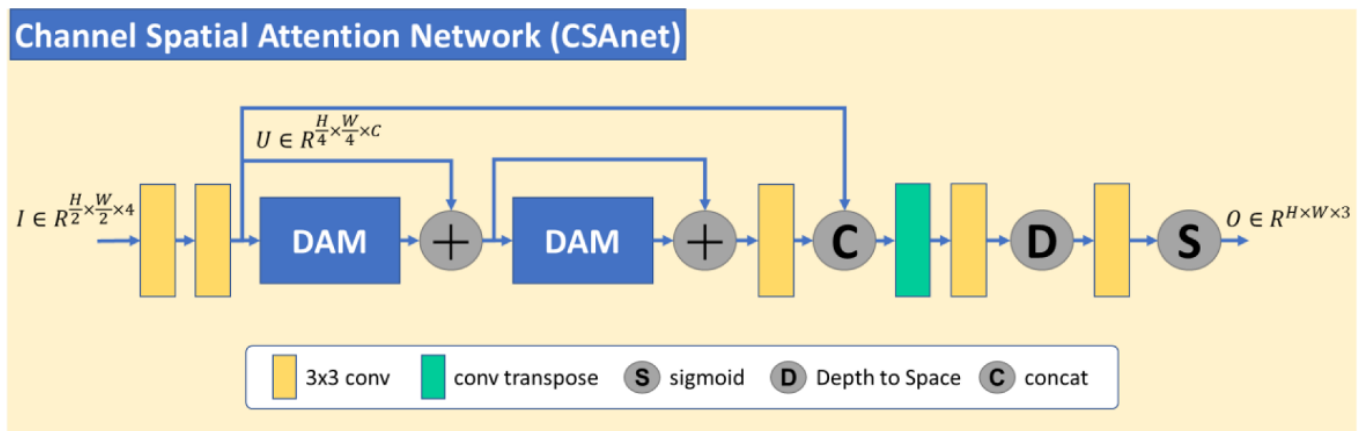
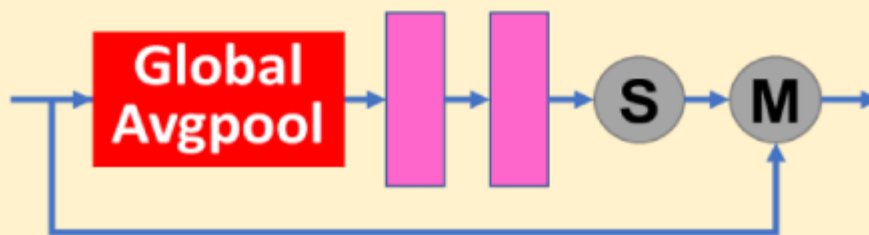


Figure 4. CSANet model from team AIISP with several attention blocks.

## DAM



## Channel Attention



## Spatial Attention



3x3 conv 1x1 conv C concat S sigmoid M mul

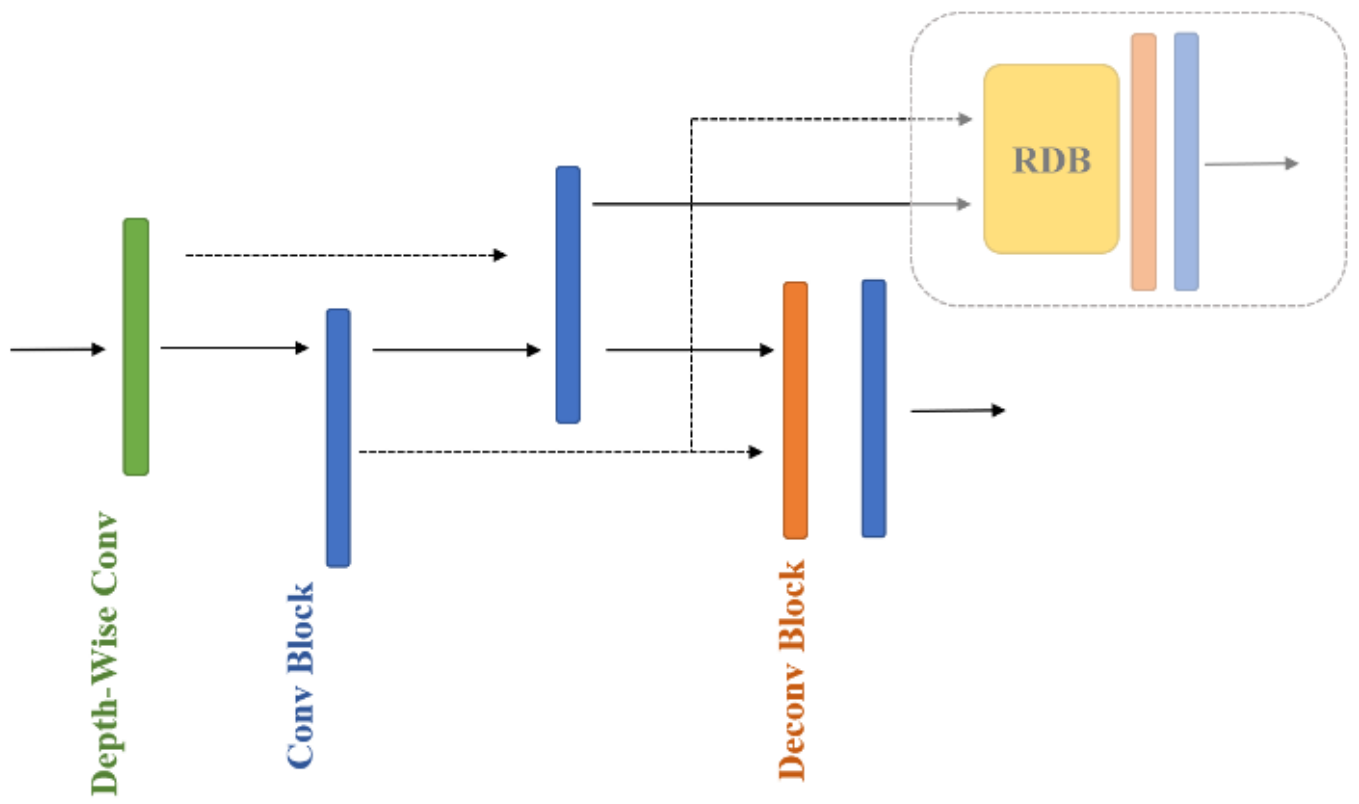


Figure 6 The model architecture proposed by ENED7 AI Research

5.isp\_forever

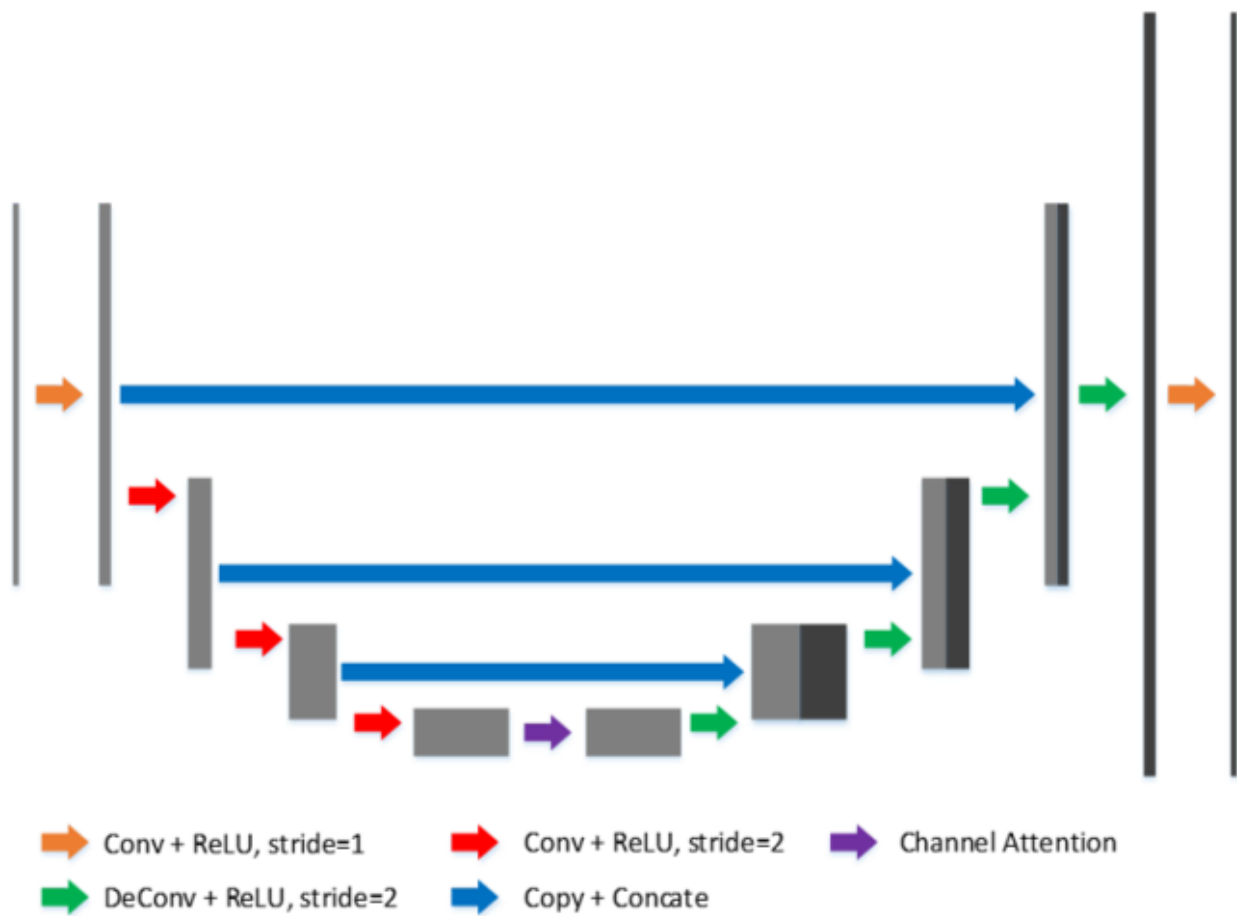


Figure 6.10: U-Net architecture for image segmentation. The network is trained using a combination of supervised and unsupervised learning.

6.CVML

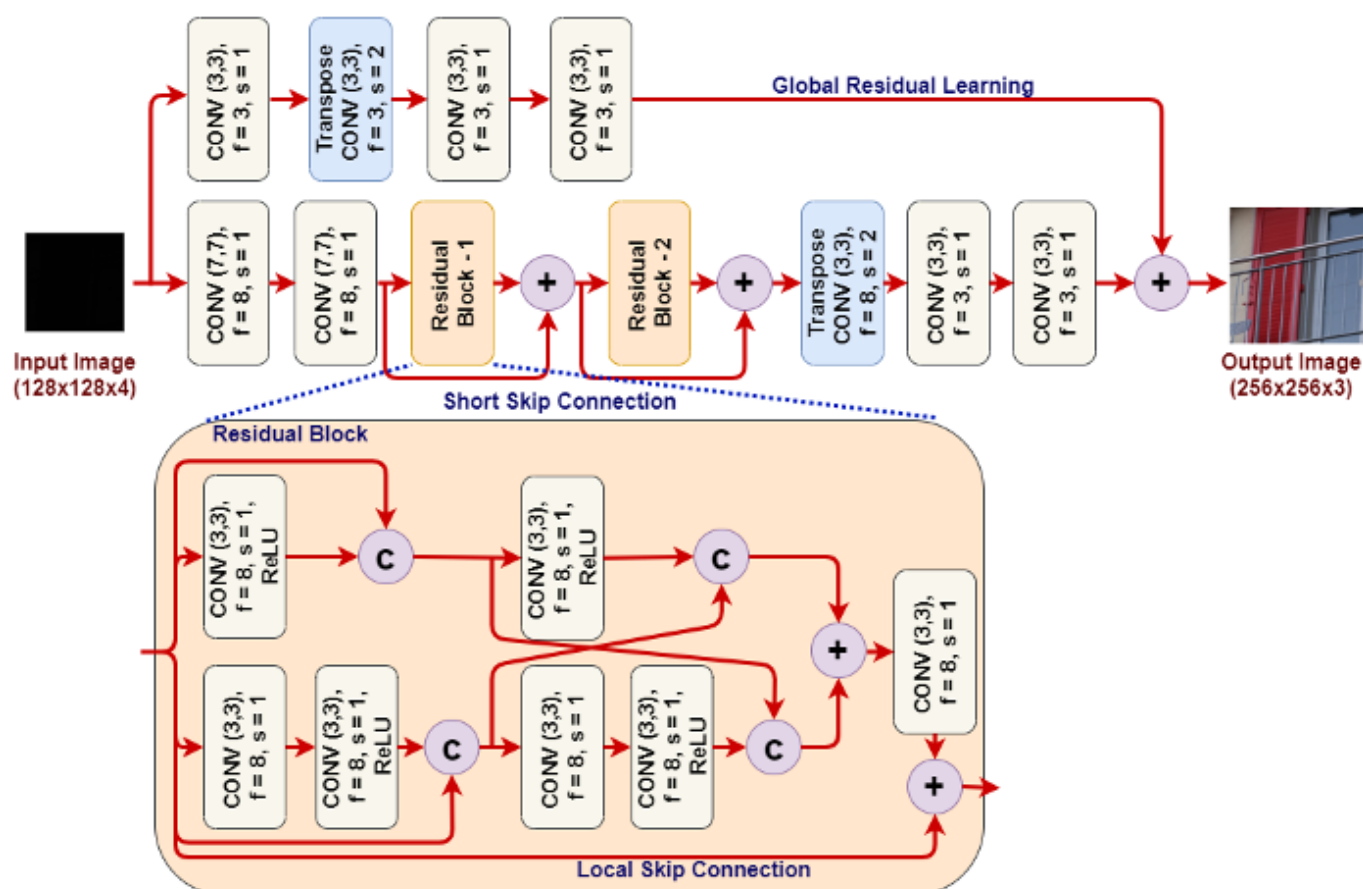


Figure 10. CNN architecture proposed by CVML team.

## 2.变换域的探究

传统方法中将图像从空间域转到其他域进行去噪是有效且常见的。旷世的那篇论文在深度学习方法中也采用了k-sigma变换这种类似的域变换方法。能否将深度学习和域变换两种有效方法耦合起来是接下来工作中可以探究的一个点。

## 3.数据集和损失函数

跟目标检测这种任务不一样的是，去噪的数据集的label并不是很统一。有效的数据集对模型效果应该会有很重要的影响。常见的去噪数据集制作方法有：1.图片加人工噪声（如加性高斯白噪声）；2.高ISO长曝光/低ISO短曝光，制作数据集。如何制作出最贴近ISP数据中噪声数据集也是可探究的点。

上述论文中采用的损失函数多为L1 loss和MSE loss。而PSNR和SSIM作为评价指标这些年也备受争议，因为实验中发现L1 loss有助于提高PSNR表现，但是PSNR高的结果并不一定符合人类观感中对图像“高质量”的评价。如何用某种指标量化人类视觉中的“高质量”感受很重要，据说这也是荣耀手机的AIISP团队主要工作方向。

此外有文章提到tanh比relu在小型网络中表现更好，可以作为实验点探究。



## Refenenced

---

1. [Deep Learning on Image Denoising: An overview \(readpaper.com\)](#)
2. [图像视频降噪的现在与未来——从经典方法到深度学习 - 知乎 \(zhihu.com\)](#)