**Image super-resolution** application is necessary - **ESRGAN+**

Advantage – Blur removal

**3D reconstruction –** VoxelMap, point cloud, Mesh, Occupancy network

Voxel representation is **simplest** but requires a lot of **processing for each Voxel**. it is a **rough representation** using it for prepossessing of **data generates** more accurate final output

Point Cloud is **fast** but **needs post processing**. Can generate mesh model for final texture prediction. helps to segment the semantic features and predict the **mesh more accurately**.

Mesh **makes unwanted overlaps**. It is better to use mesh model instead of voxels if the model is rich in silhouettes.

**Neural networks –** Autoencoder , RNN (recurrent Neural Network), CNN(convolutional Neural Network) , GAN(Generative Adversarial Neural Network)

**Autoencoders** are simplest and consisting of encoder and decoder. Since it primarily does only encoding and decoding and can act as a deep learning-based identity matrix that **could reconstruct the given input,** but **with desired resolution or output that could be generated from its lossy features.**

**RNN** performs very well when **sequential data is provided**, therefore it allows previous output to be used as current input. RNN has the **tendency to forget previous information**.

CNN - requires **convolving matrix** and **can predict 3D reconstruction** is CNN. are used in 3D reconstruction

GAN - a deep neural net which consists of 2 different neural net, **a generator, and a discriminator.** The **generator** is actually the **main algorithm** that will finally **produce the final output** but in order to evaluate if its output is fine or not, a **discriminator** is used that can **actually determine** if the output is **real or fake, and based on its judgment generator** changes its parameter to come close to a near accurate output.

Bootstrap method- **without disturbing the internal structure** of the neural net it is possible to improve its output. (self-supervised method - not need any new dataset in order to train it.)

Pose Estimation- contributes to **make a better 3D reconstruction.** can be estimated using neural network such as **CNN.**

**Methods**

**Photo geometric Autoencoder:** is **very accurate** when **detecting faces** because it utilizes the geometry of the object in the photogram. the results are best **when used in object with high symmetry**. And can **perform** well **in unsupervised condition**. . but The input image resolution needs to be high

**Super-resolution - The** main purpose of 3D reconstruction requires several other important tasks such as removal of blurs due to camera shakes. And the necessity to estimate pose to generate the 3D image. Face Deblurring method is needed.

**Face deblurring -** The method **requires convolution**. And it will generate 3D facial priors in presence of Up-convolution and Resnet. This method has the **limitation** that It runs slower that other various low-resolution frames. This method also did not focus on providing a high-quality 3D representation

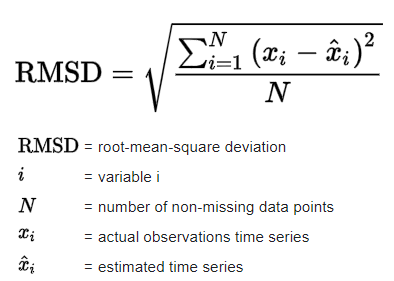
**Occupancy Network - overcomes** the missing features of the method describe above but also is a **much better approach than voxel map, point net, and mesh net.** Because **voxel map** needs more processing and is not accurate. And **point net is** fast but requires post processing and **mesh net** will create a lot of overlapping structures and none of these problems can be seen in this method. This method can work on single view, Multiview and voxel blocks. This method is a type of 3D representation therefore the issues of self-occlusion and **posed are not focused in it.**

**Deep autoencoder-** can estimate the poses of human body and upscale the body of a model. By which the limitations of the **current method can be resolved**.

**Experimental Results**

In order to **calculate the quality of an image** there are two ways. 1) Objective and 2) subjective evaluation methods. The **objective evaluation** is usually done by observers (experts in the field of study) which may give scorings (usually from 1-10 scale) to each image (original and filtered). The **subjective evaluation** is done by using metrics such as the PSNR, SNR, MSE, quality index, structural similarity index and many more

**RMSE (Root mean square Error)** - The root-mean-square deviation or root-mean-square error is a frequently used measure of the differences between values predicted by a model or an estimator and the values observed



**PI (Perceptual Index)**

**PSNR (Peak signal to noise ratio) -** This ratio is used as a quality **measurement between the original** and a **compressed** image. The **higher the PSNR**, the better the quality of the compressed, or reconstructed image.

**SSIM (structure Similarity)** - The structural similarity index measure is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. SSIM is used for measuring the similarity between two images.

**IoU (Intersection Over Union) -** an evaluation metric used to measure the accuracy of an object detector on a particular dataset.

**CD (Chamfer Distance)** - is a sum of positive **distances** and is defined for unsigned **distance** functions

**CE (Cross Entropy loss) - measures the performance of a classification model** whose output is a probability value between 0 and 1. **Cross**-**entropy loss** increases as the predicted probability diverges from the actual label.

**NME (Normalized Mean Error)** A widely used **air quality model performance index**.

For **2D result verification**, the SSIM has shown a much more reliable verification than other output such as PSNR or RMSE. In the case of **3D evaluation**, EMD, IoU, CD, are more commonly seen to be used for evaluations

Neural net such as **GAN(Generative Adversarial Network)** based ones can be have their generator network trained to the desired degree in order to predict a higher accuracy. The neural network needs a good optimizer such as ADAM for a better output. The loss functions, normalization also important in order to generate a **high accuracy output**. Apart from image upscaling image deblurring is also necessary.

Some of the other neural network that **can successfully generate a good output include CNN,** **Autoencoder and RNN. RNN will not suitable though if the training needs to go on for long term.**