UNIVERSITÀ
DEGLI STUDI
FIRENZE

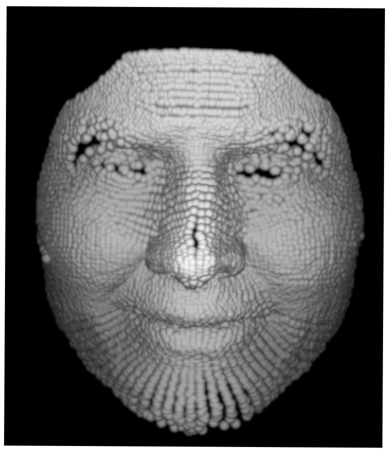
PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

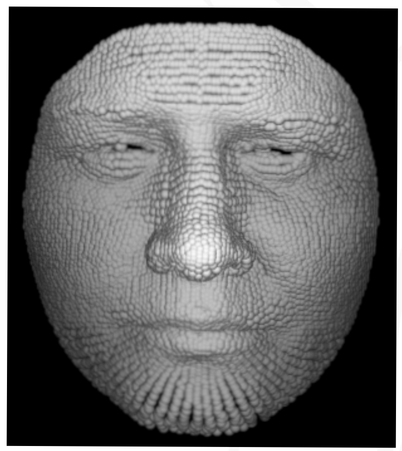
Identitiy, expression and unicode recognition

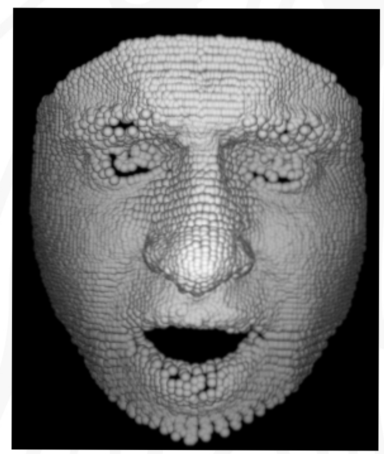
Prof:Stefano Berretti
Student: Fatemah Alhamdoosh

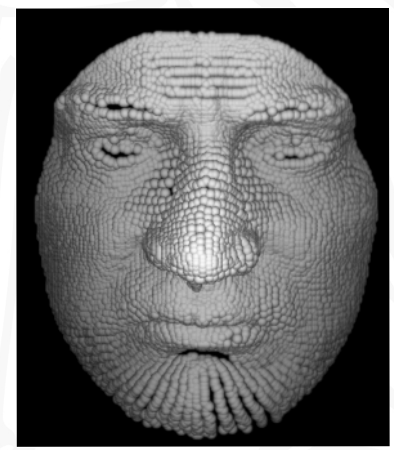
Problem: classification

- Classification of point cloud which basically categorizes a set of data into classes.


001_LFAU_12


000_E_SADNESS


033_E_FEAR


037_E_ANGER

Data: 3D Point Cloud

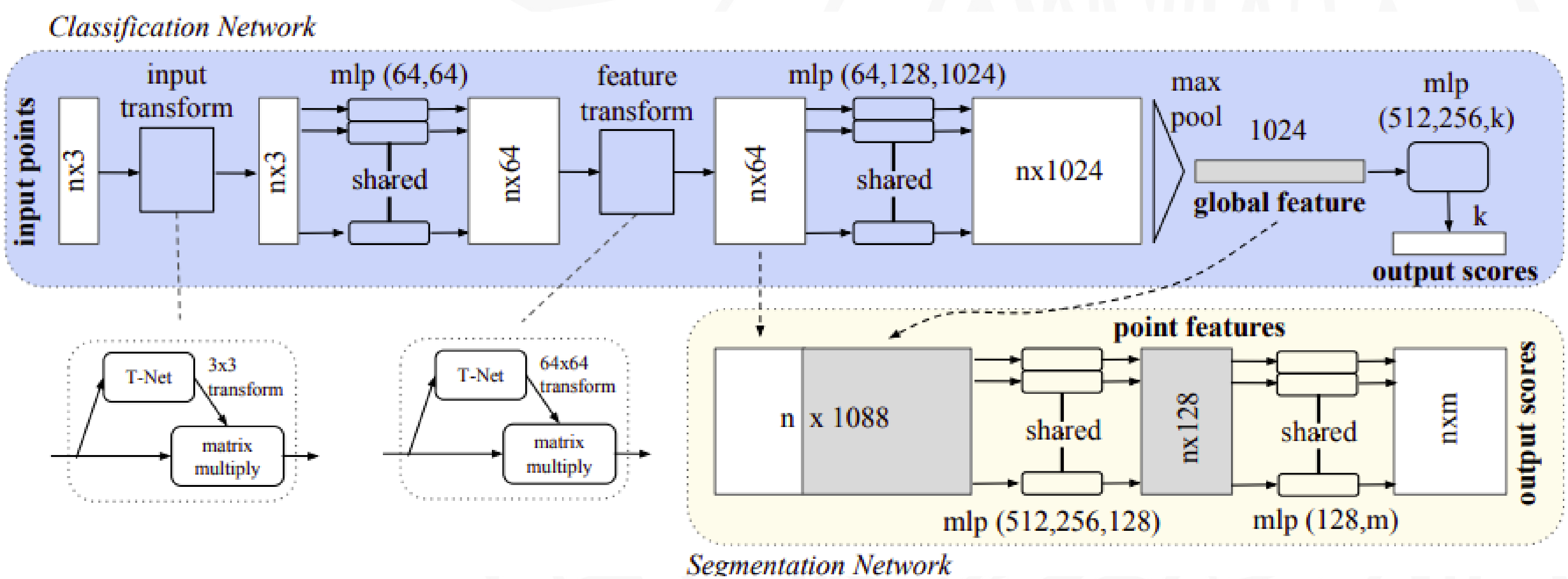
- **Description of our Data set:**
- Number of point Cloud: **2889** ,each with up to **6704** point.
- Number of classes: **105** classes for id recognition, **35** classes for expression and **5** classes for unicod recognition.
- **Point cloud:** is represented as a set of 3D points {Pi | i = 1, ..., n}, where each point Pi is a vector of its (x, y, z) coordinate plus extra feature channels such as color, normal etc.
- **Properties:**
- Point cloud is an unordered set of vectors
- Interaction among points: neighboring points form a meaningful subset
- Invariance under transformations: translation and rotation.
- **Point cloud features:**
- Intrinsic or extrinsic
- Local features and global features

Pointnet: Deep Learning on Point Sets

- **Neural Network:** take directly point cloud.
- **Two keys:**
- **1- Max Pooling Symmetry Function for Unordered Input:** OrderMatters!

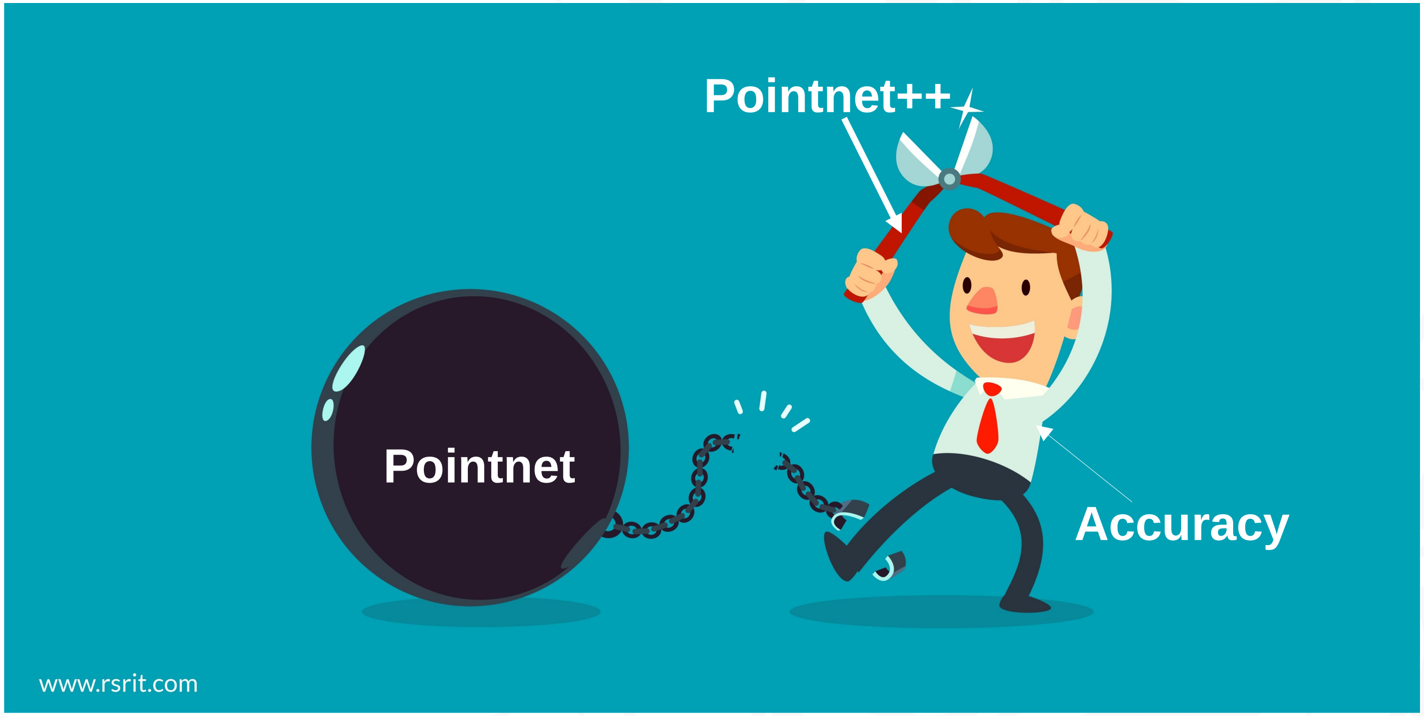
$$f\left(x_1, x_2, \ldots, x_n\right) \approx g\left(h\left(x_1\right), h\left(x_2\right), \ldots, h\left(x_n\right)\right) where\ f: 2^{R^N} \rightarrow R, h: R^N \rightarrow R^K, g: \underbrace{R^K \times R^K \times \ldots \times R^K}_{n} \rightarrow R$$

- **2- Joint Alignment Network:** predict an affine transformation matrix by a mini-network (T-net in Fig) and directly apply this transformation to the coordinates of input points



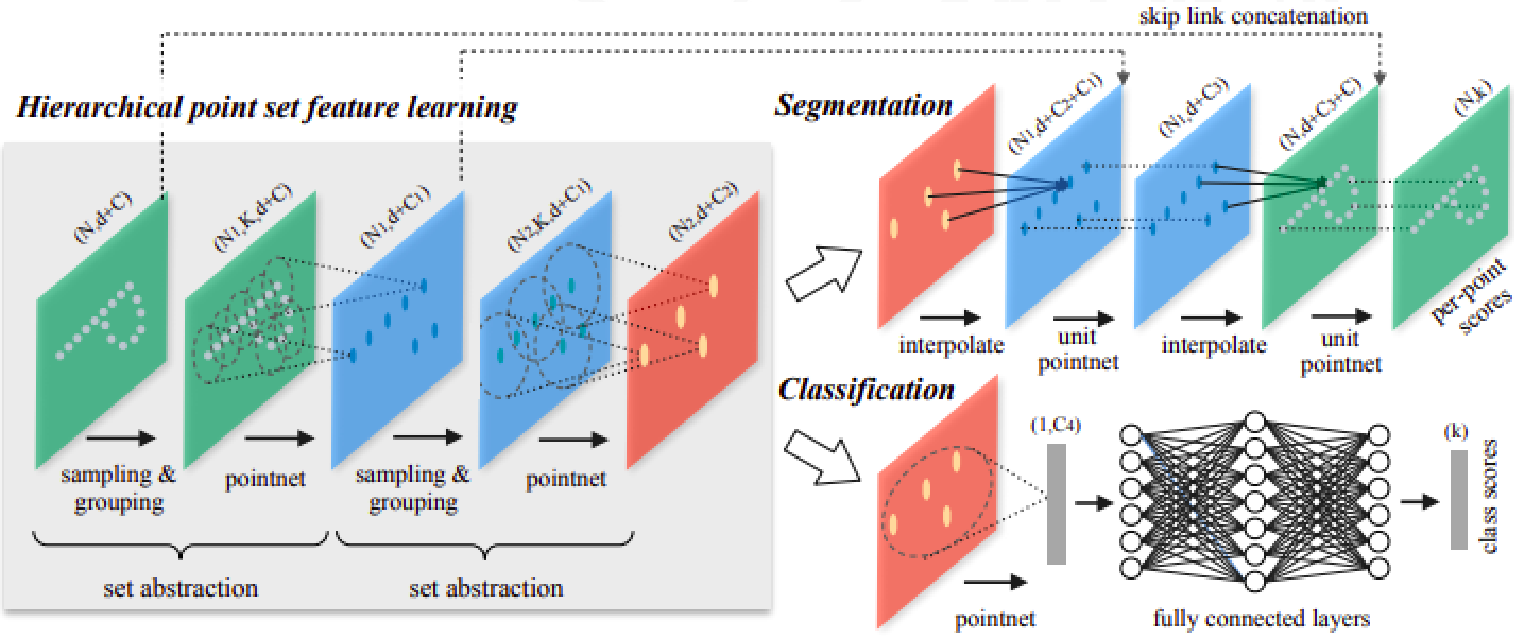
Limitation of Pointnet

- PointNet does not capture local structures induced by the metric space points live in, limiting its ability to recognize fine-grained patterns and generalizability to complex scenes



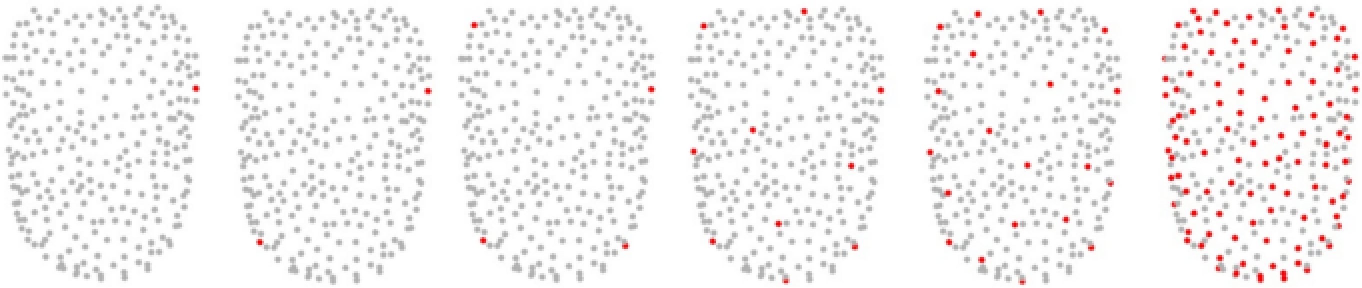
PointNet++: Hierarchical Feature Learning

- Pointnet++ introduce a hierarchical neural network that applies PointNet recursively on a nested partitioning of the input point set.



Pointnet++: How to generate overlapping partitioning of a point set?

- **Farthest Point Sampling (FPS) in Euclidean space:** Sample k points (Centroids) from N points

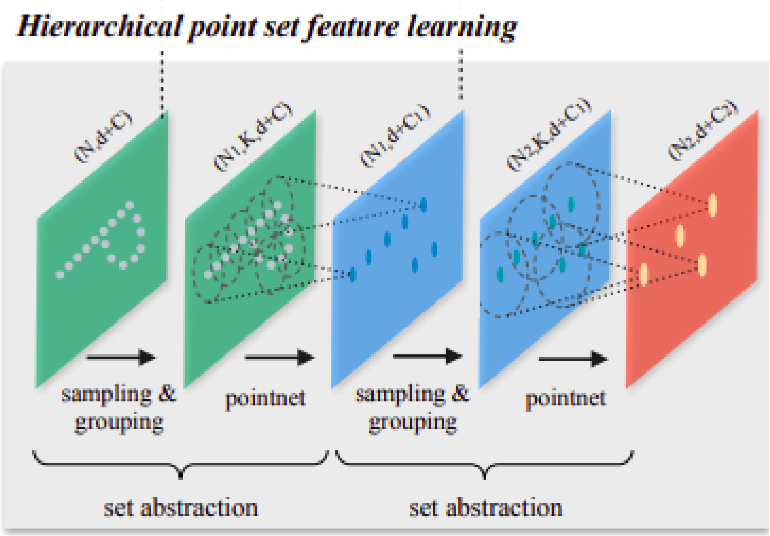


- **Query ball point :** Find neighborhood points of centroid within certain radius

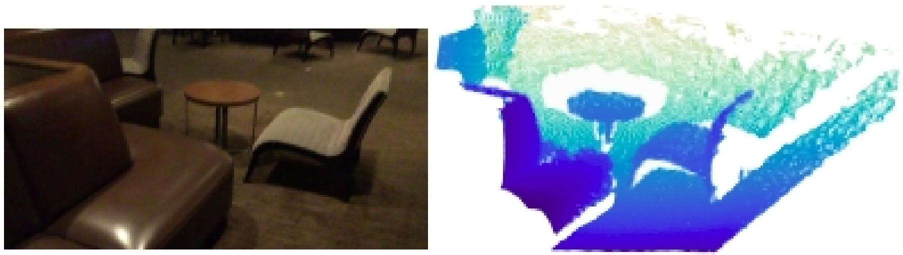
- **GPU Implementaion.**

Pointnet++: Single Scale Grouping (SSG)

- **Limitation:**
- **Variable densities at different areas:** With further observation that point sets are usually sampled with varying densities, which results in greatly decreased performance for networks trained on uniform densities, we propose novel set learning layers to adaptively combine features from multiple scales.

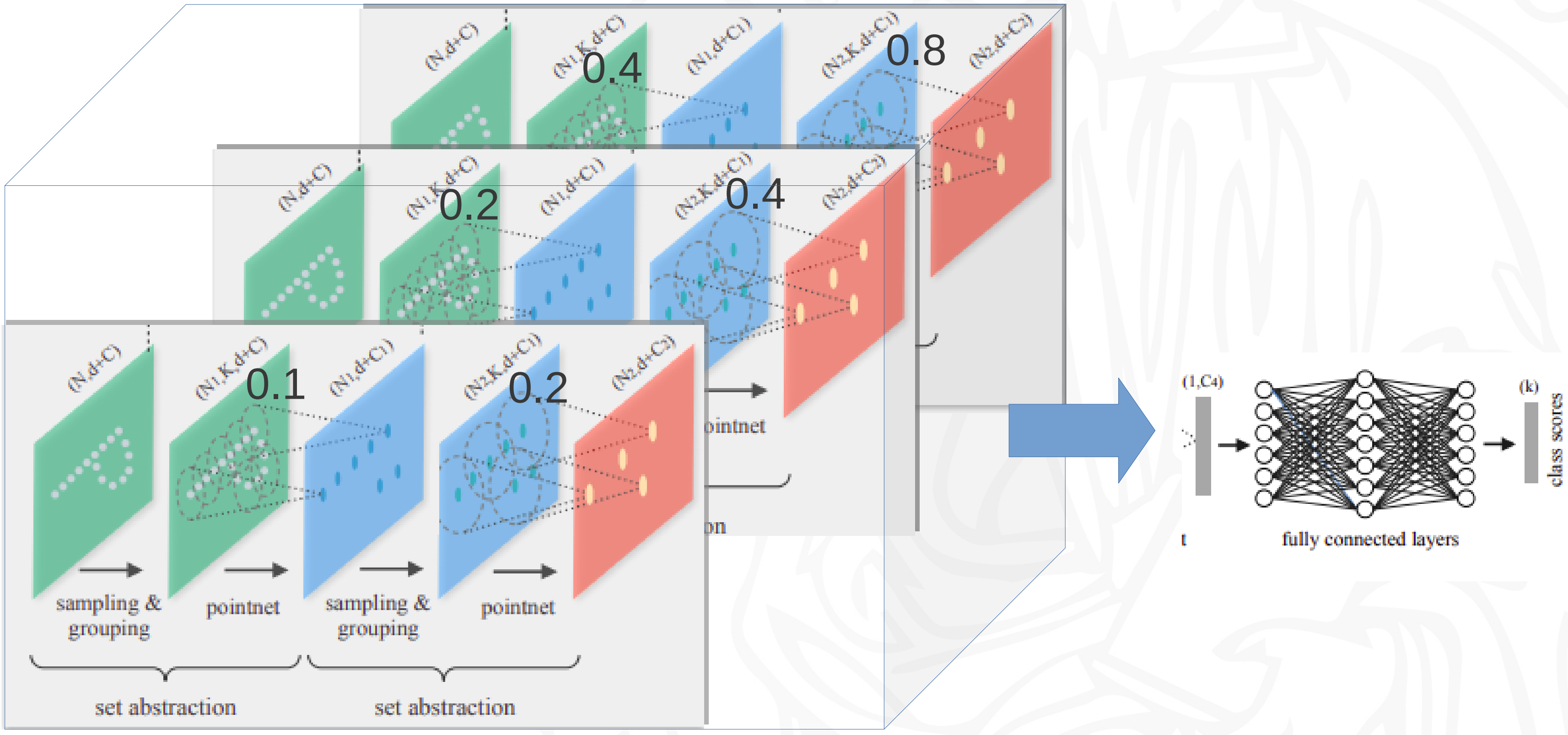


Scan captured from a Structure Sensor



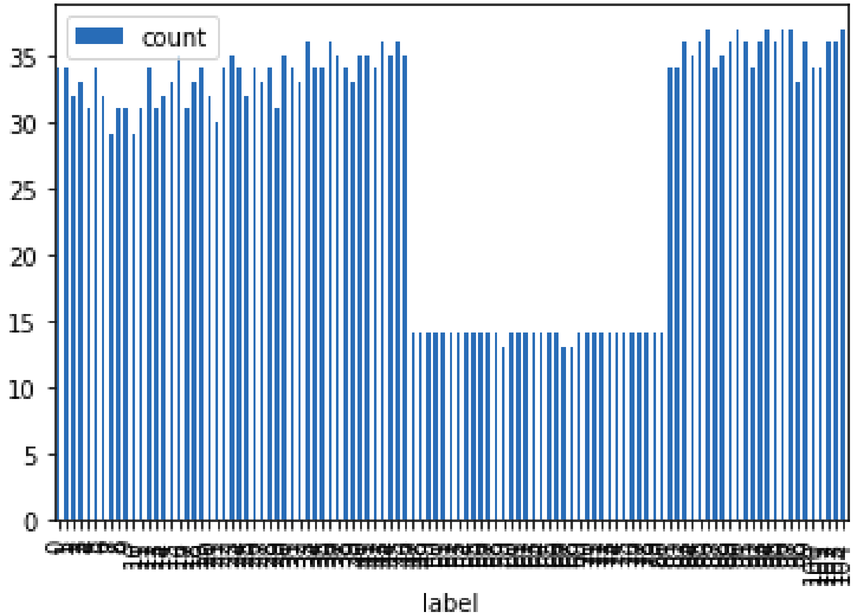
Pointnet++: Multi Scale Grouping (MSG)

Multi Radius: R1=[0.1, 0.2, 0.4] R2=[0.2, 0.4, 0.8]
Take more time for training



Experiments: Face identity classification

- Goal:classification of 2889 point cloud into 105 class.
- Divide data set into 70% training set and 30%test.



Distribution of classes over data set

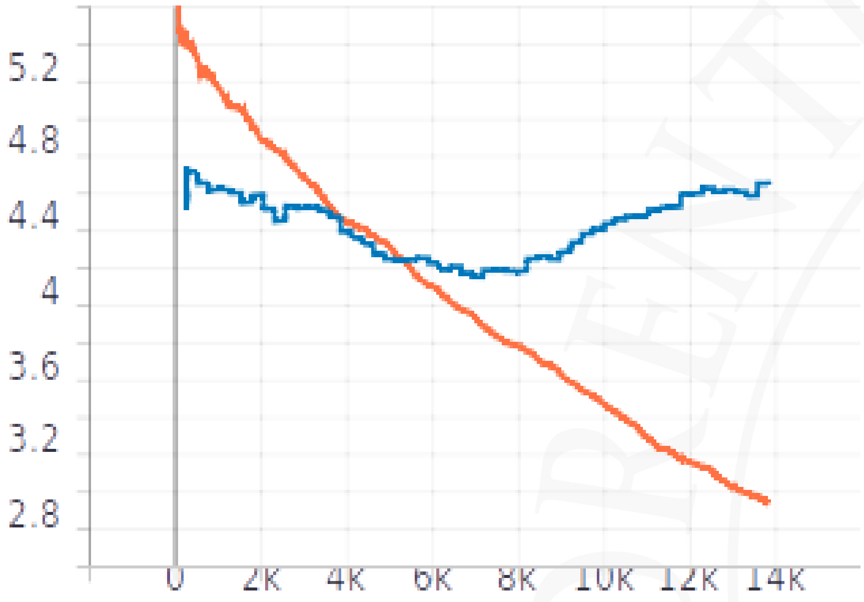
Experiments: Face identity classification

- **Configurations:**
- nsample: Number of points at each local partition, 32 lead to overfitting
- Momentum vs Adam: optimizer adam lead to overfitting from the first epoch
- With normal vs without normal: without normal lead to overfitting.
- SSG VS MSG

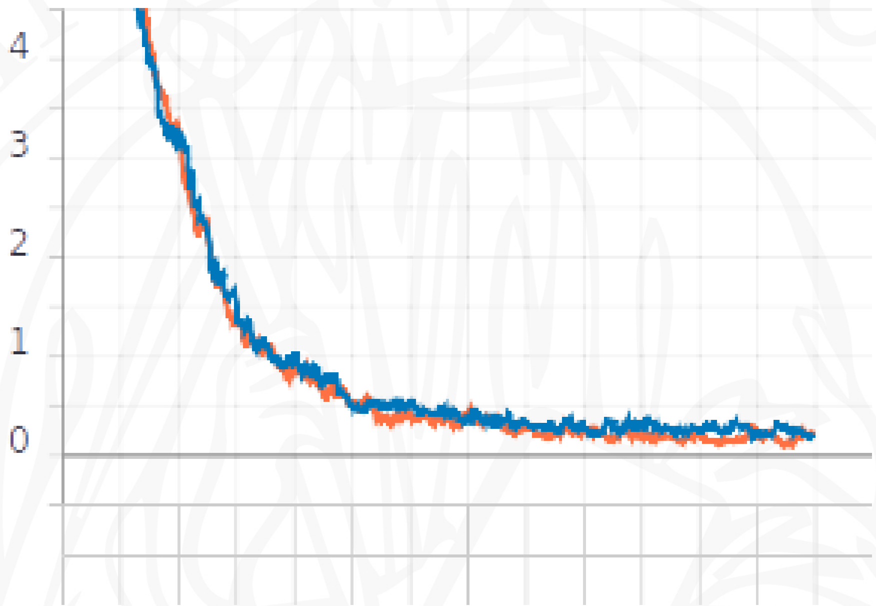
Model	npoint	nsample	DecayStep	Normal	acc	avg-acc	loss
SSG	6704	128	80000	True	**0.94**	**0.90**	**0.26**
SSG	6704	**32**	80000	True	0.01	0.002	3.45
SSG	6704	128	80000	**False**	0.011	0.11	6.65
MSG	6704	128	80000	True	0.17	0.16	3.93

Experiments: Face identity classification

- MSG vs SSG



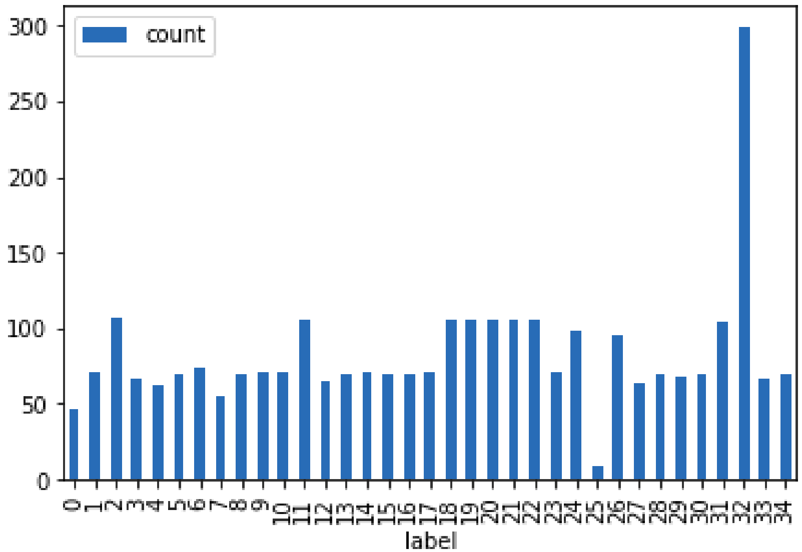
Loss curve using MSG



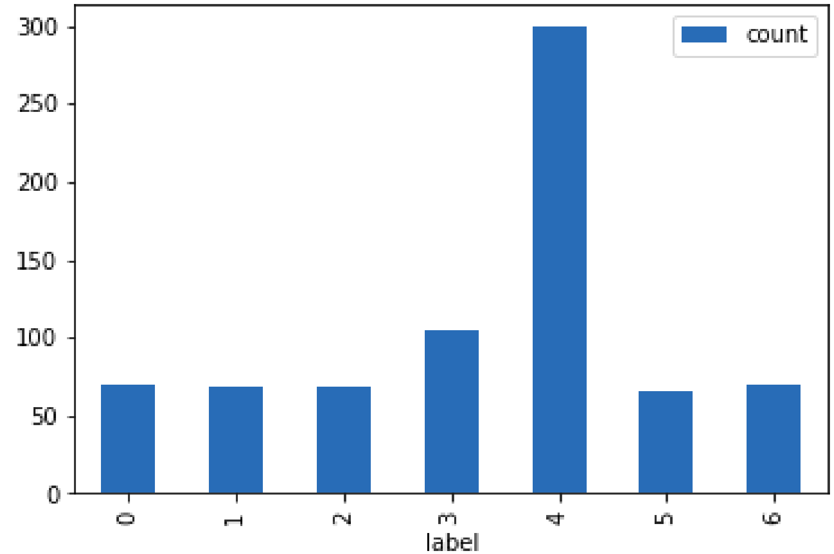
Loss curve using SSG

Experiments:Face Expression classification

- Imbalanced dataset
- Class 32: 10% of total datset
- Clas 25: 0.3% of total dataset.
- Repeated the same experiments of id classification, best results is:
- Loss: , Accuracy : , avg accuracy for class:



Distribution of 35 classes



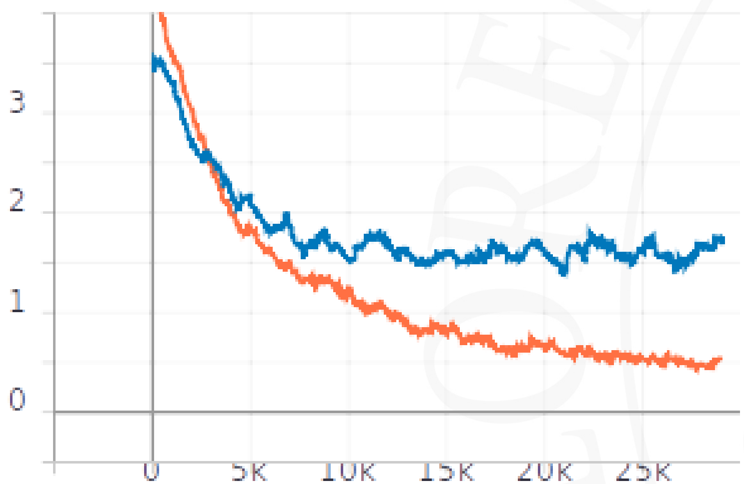
Distribution of 7 classes

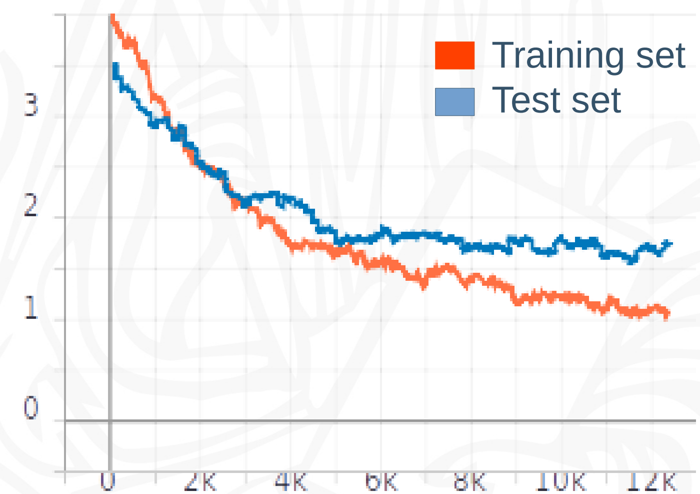
Experiments:Face Expression classification

- <u>**Things that dose not help:**</u>

- Using smaller Learning rate, decay step, nsample or npoints
- Removing the minority class: dose not help
- Removing both minority and majority classes:
- Using weighted loss cross entropy: try to weight up the loss of minority class and weight down the loss of majority class

$$weight[i] = \frac{1}{num_{classes}} * \left(\frac{Total\ number\ of\ point\ clouds}{number\ of\ point\ cloud \in class[i]} \right)$$



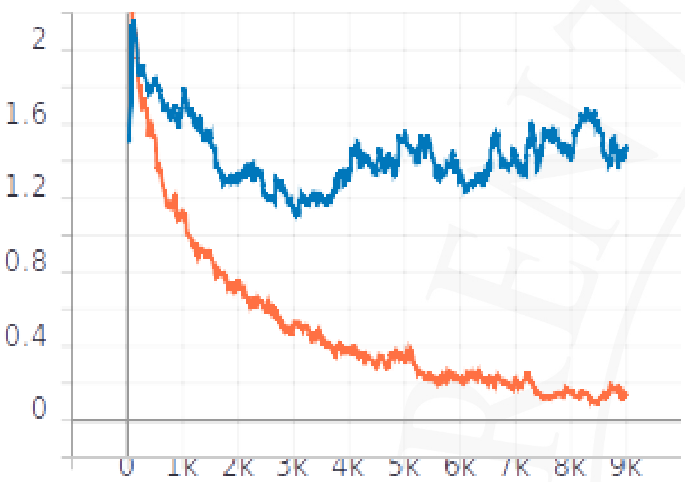


Loss curve with defualt parameters

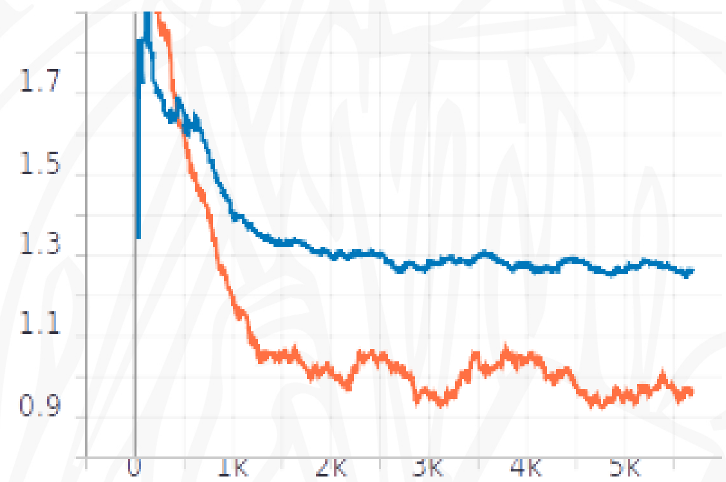
Loss curve using weighted loss

Experiments:Face Expression classification

- **Experiments with just 7 classes:**
- **Slightly better with smaller decay step and decay rate.**



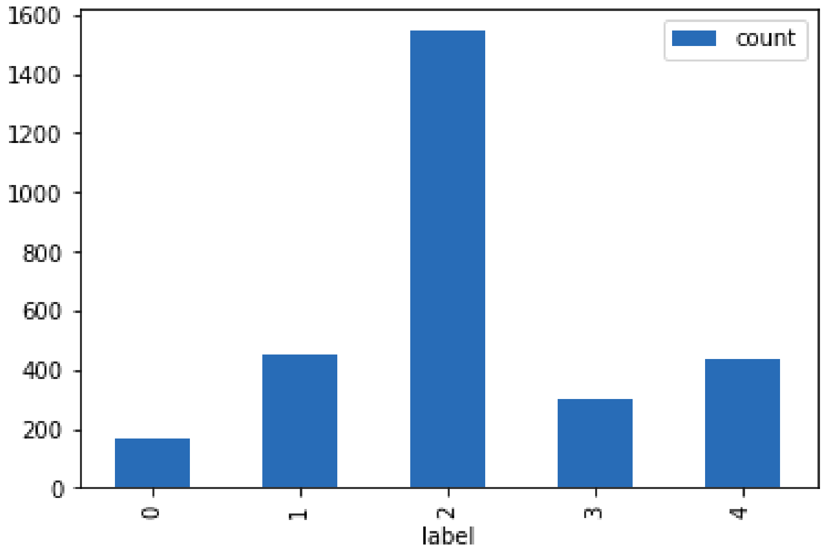
Loss curve with default parametrs



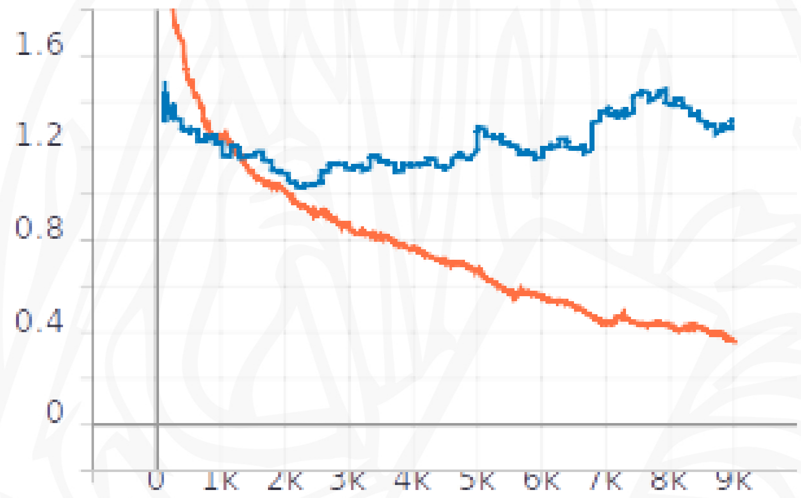
Loss curve wit decay step :10000 and decay rate 0.1

Experiments:Face unicod classification

- Similare to expression experiment..
- Imbalanced dataset: class 2 with up to 50% of total dataset
- Repeated the same experiments of id and expression.



Distribution of classes over data set



Loss curve with default parametrs

Experiments:Face unicod classification

- Worse than expression classification.
- Number of nsampl and npoint affect highly the result.
- Smaller learning rate lead to later convergence

Model	npoint	nsample	DecayStep	Normal	acc	avg-acc	loss
SSG	6704	128	40000	True	**0.64**	**0.57**	**1.12**
SSG	**2048**	**64**	40000	True	0.57	0.34	1.6
SSG	6704	128	40000 Lr=0.0001	**True**	0.58	0.48	1.11

Future work:

- Use Transfer learning technech : from identity model to expression and unicod models
- Fast marching Farthest point sampling
- Learned Wieghted FBS
- Multi resolution grouping

Conclusion:

- Pointnet ++ is a powerful neural network.
- SSG: very good results and performance.
- MSG : is more useful when data densities is variable.
- Too much parameters as like as the majority of neural networks: Tuning parameters is more challengeable.
- **Knowledge & Skills Gained:**
- Tensowerflow v1,operations, gpu ,tensorbord…
- Command line skills improvement.
- New algorithms: FPS, query ball.
- Working with 3d data , unordered point sets.
- Important neural network architecture : max pooling, rigid transformation as neural network.

UNIVERSITÀ
DEGLI STUDI
FIRENZE

Thank you
for your attention!