CAE

convolutional autoencoder

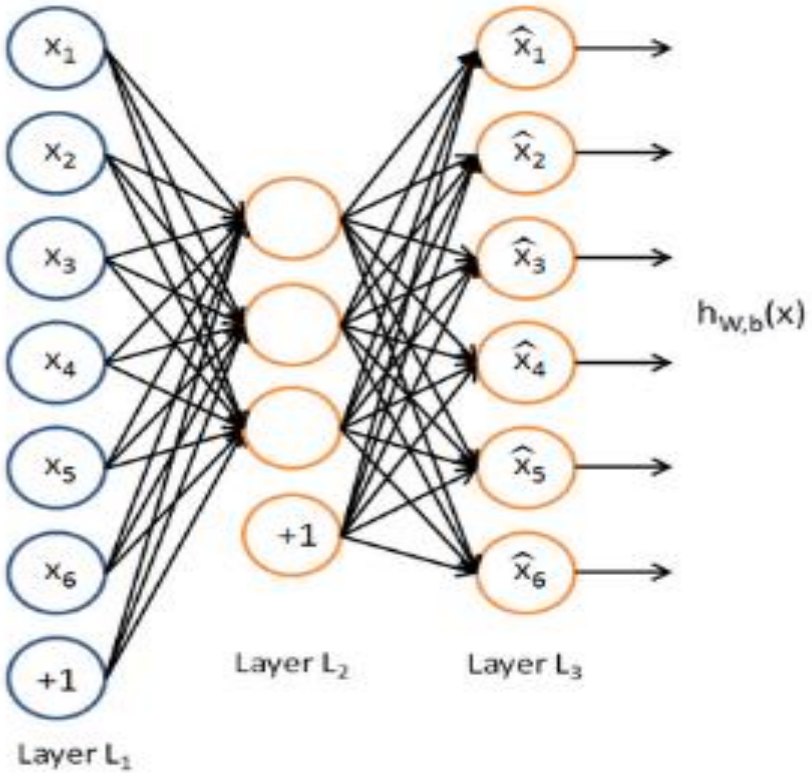
Vision@OUC
Wang Chao
Group of DL

Overview

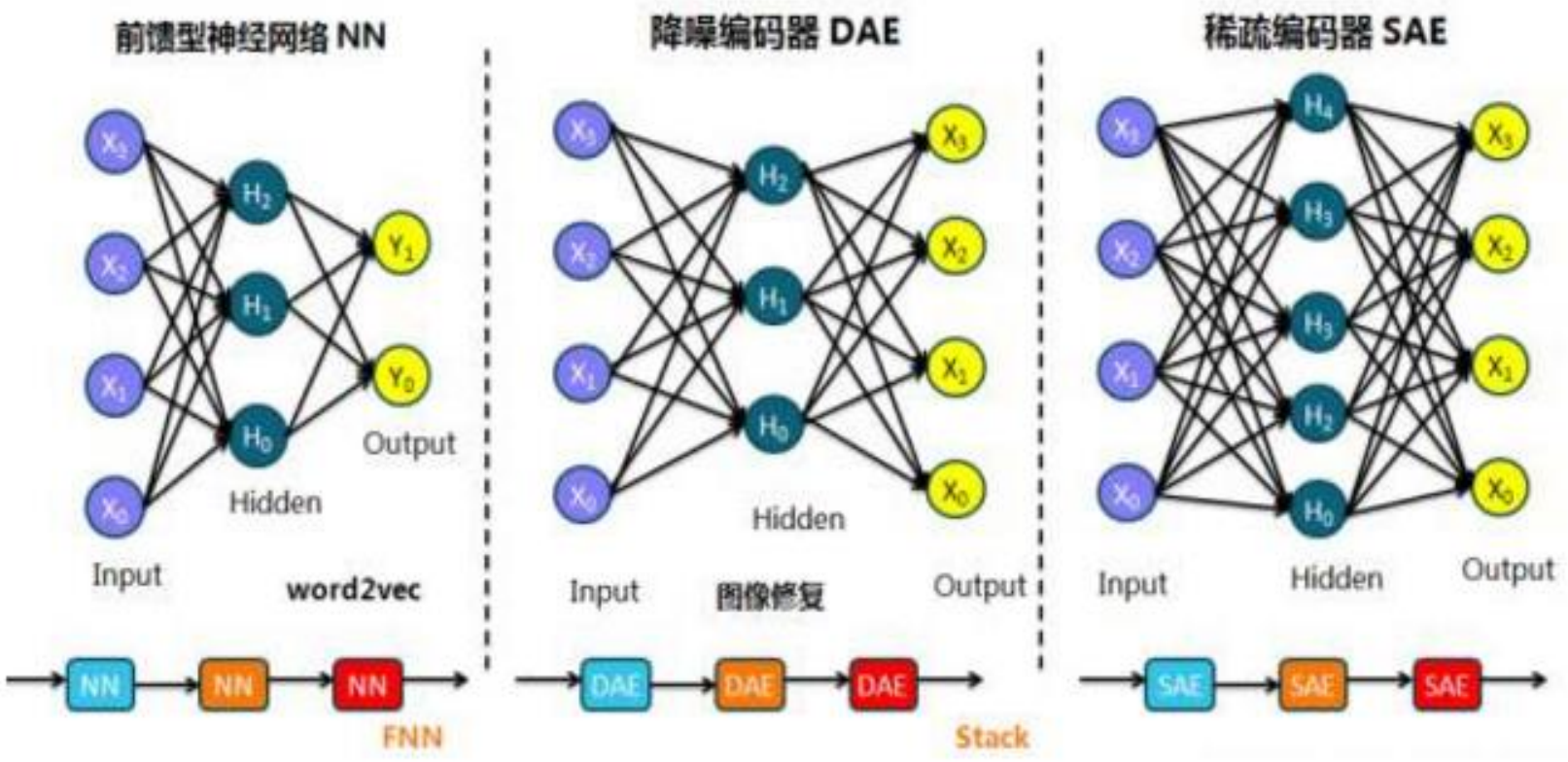
- Introduction
- Experiments
- What's Next
- Q&A

Introduction

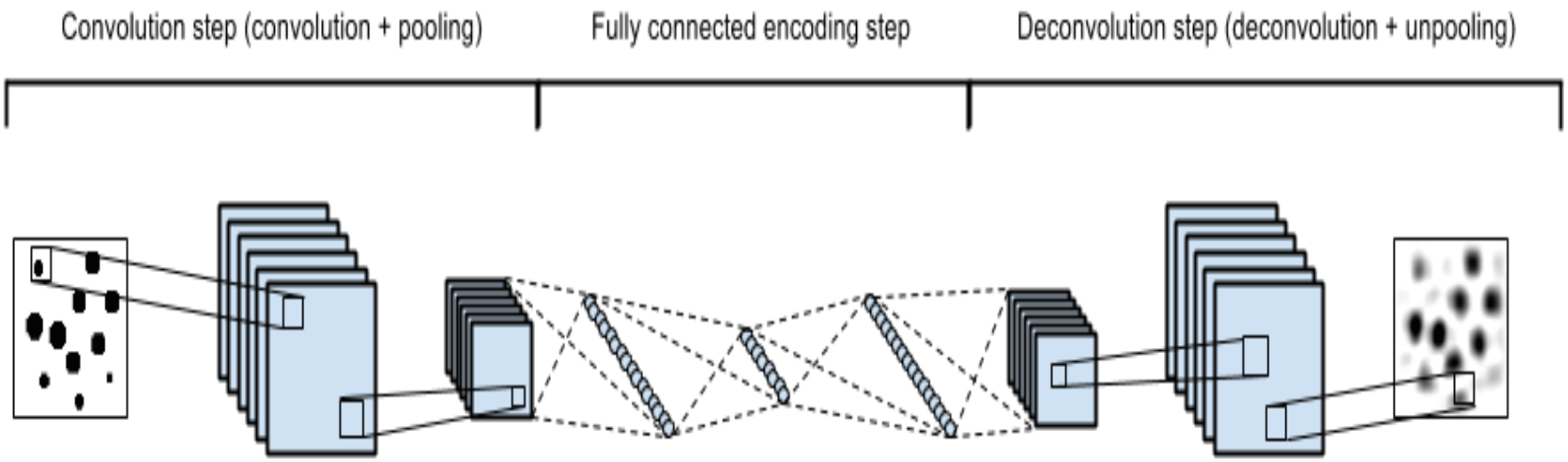
What is autoencoder ?



NN & AE & SAE


前馈型神经网络 NN
降噪编码器 DAE
稀疏编码器 SAE
Input
Hidden
Output
word2vec
图像修复
NN
DAE
SAE
FNN
Stack

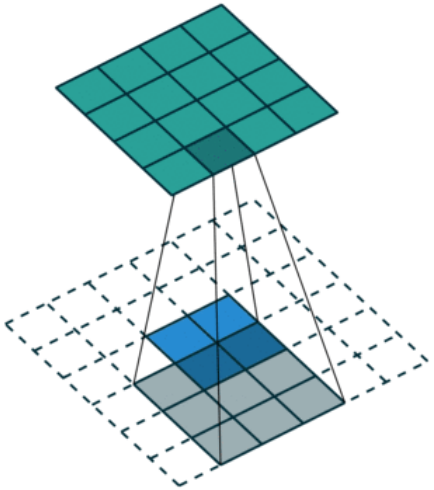
What is CAE?



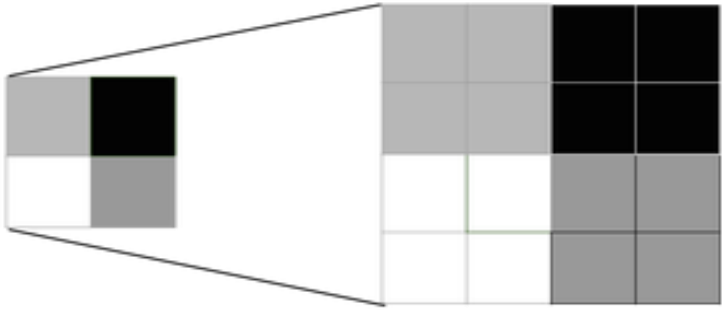
Masci, Jonathan, et al. "Stacked convolutional auto-encoders for hierarchical feature extraction." *International Conference on Artificial Neural Networks*. Springer Berlin Heidelberg, 2011.

Keys to implement

- Deconvolution



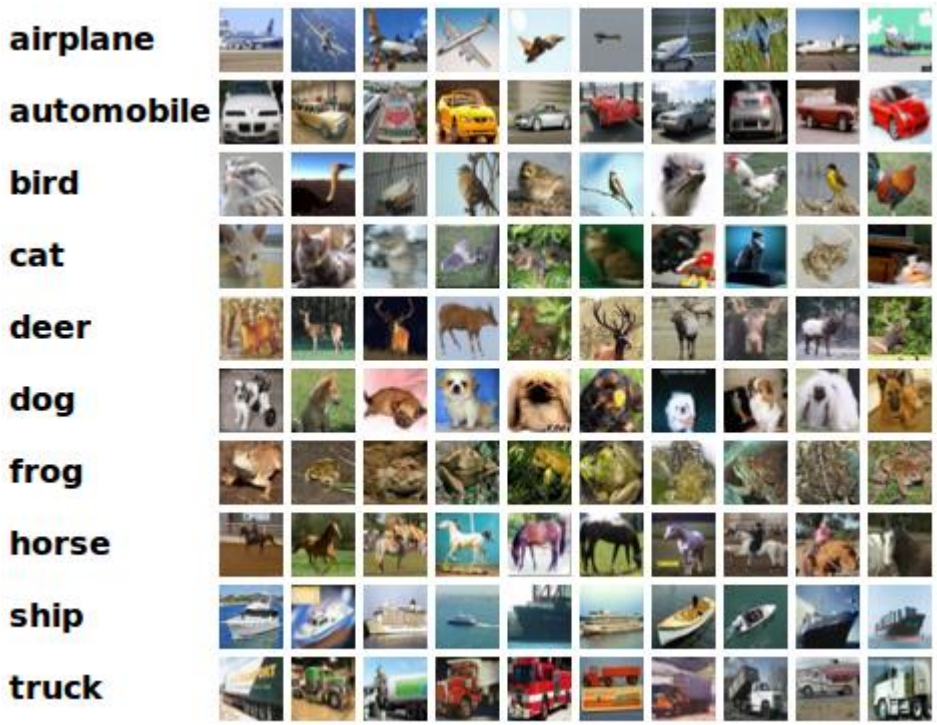
- Unpooling



Experiments

Toy examples

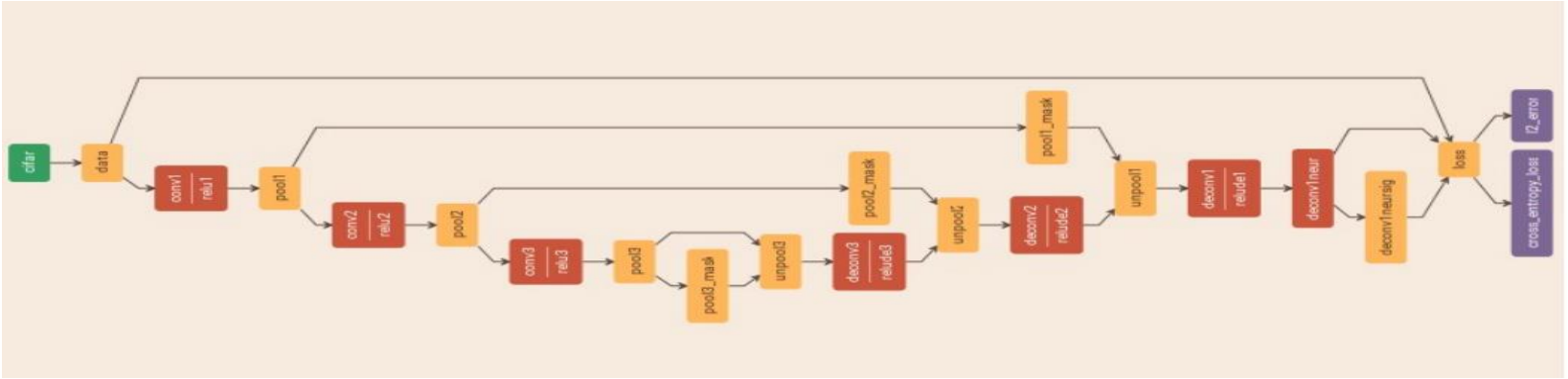
- Dataset: cifar10



R. A. Fisher (1936). "The use of multiple measurements in taxonomic problems". Annals of Eugenics.

Experiments details

- CAE architecture

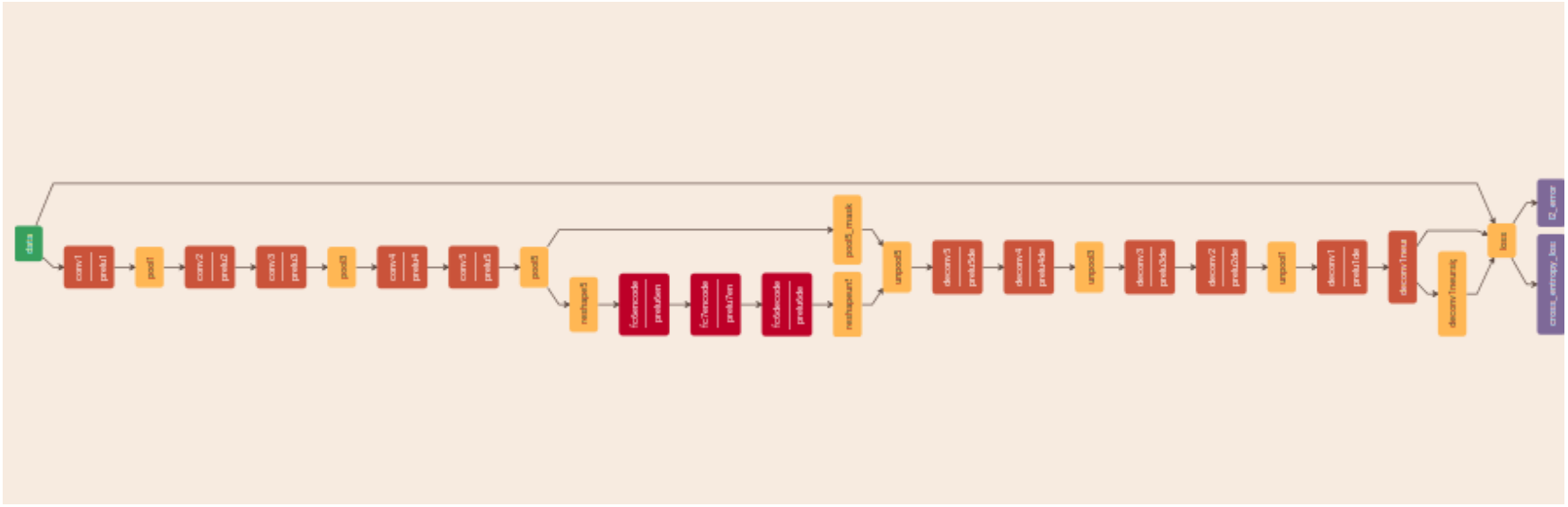


Encoder	Conv1	Pool1	Conv2	Pool2	Conv3	pool3
Fliter	64*5*5	2*2	32*3*3	2*2	16*3*3	2*2
Stride	1	2	1	2	1	2
Pad	2		1		1	
Decoder	Unpool3	Deconv3	Unpool2	Deconv2	Unpool1	Deconv1
Fliter	2*2	16*3*3	2*2	32*3*3	2*2	64*5*5
Unpool_ size	8		16		32	

Representation result

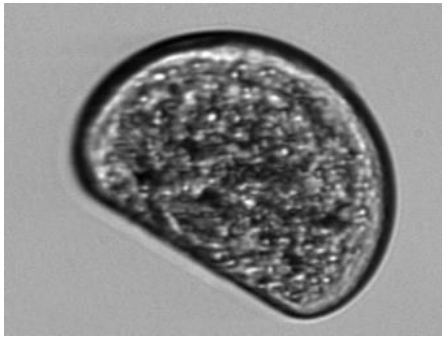
Experiments on WHOI

- Architecture

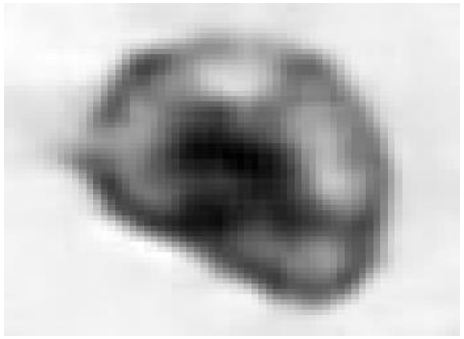


- Solver
  - learning rate：0.001； solver type：Adam； momentum：0.9 ；
    weight_decay:0.0005; loss function: euclideanloss, cross_entropy_loss.

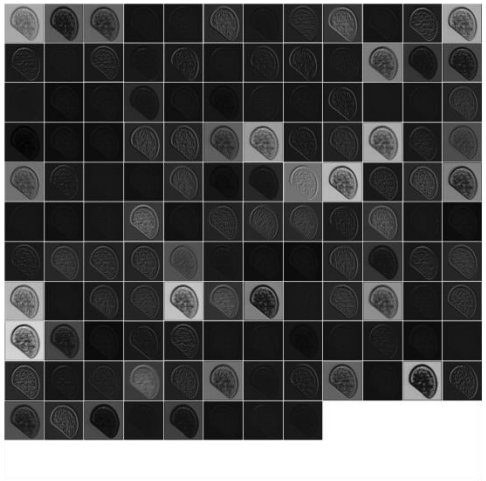
Results and analysis

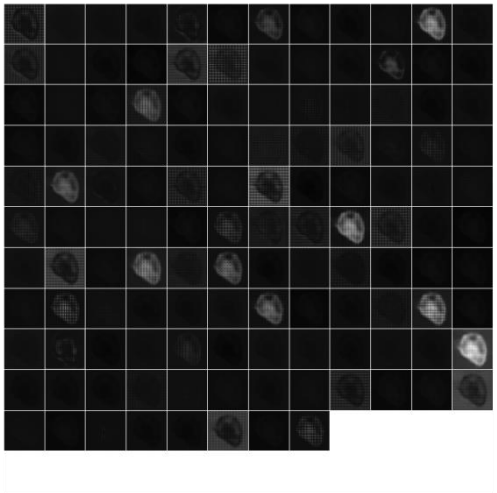


Origin image



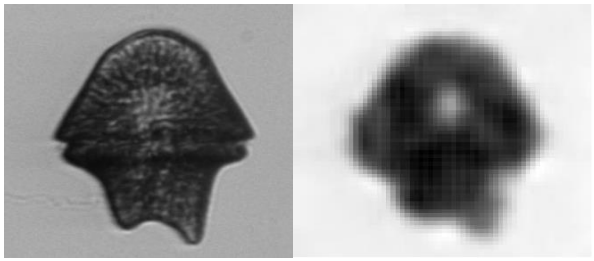
Reconstruction image





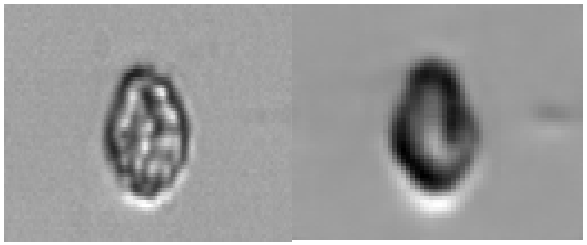
Insights

- On small class(Akashiwo)



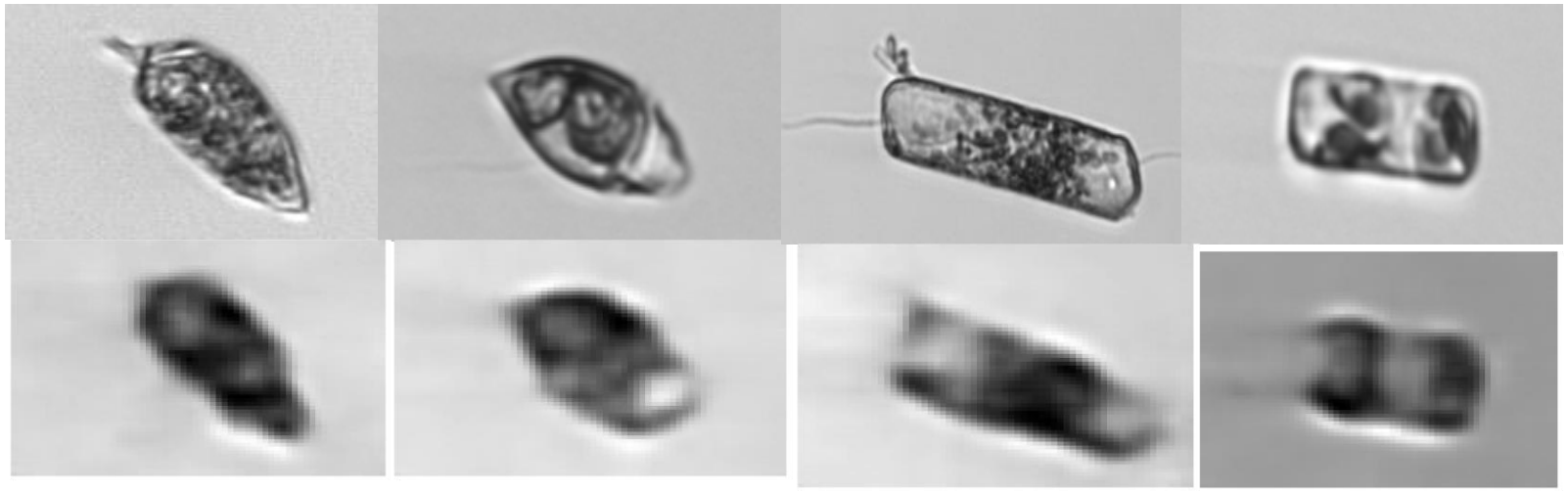
Only 4 images in training dataset

- Large class(dino30)

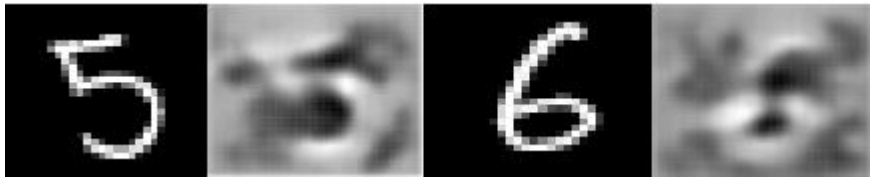


43903 images in training dataset

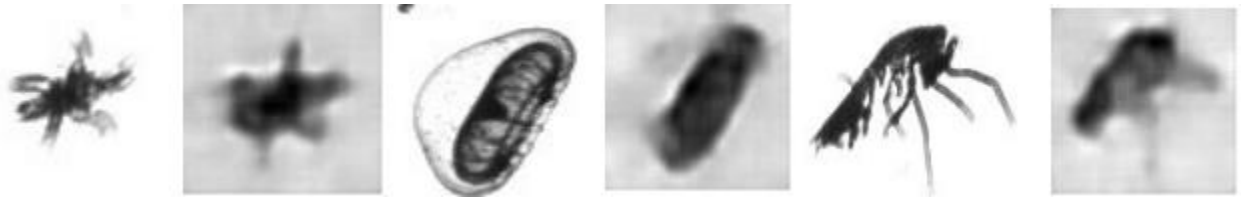
Shape alike classes



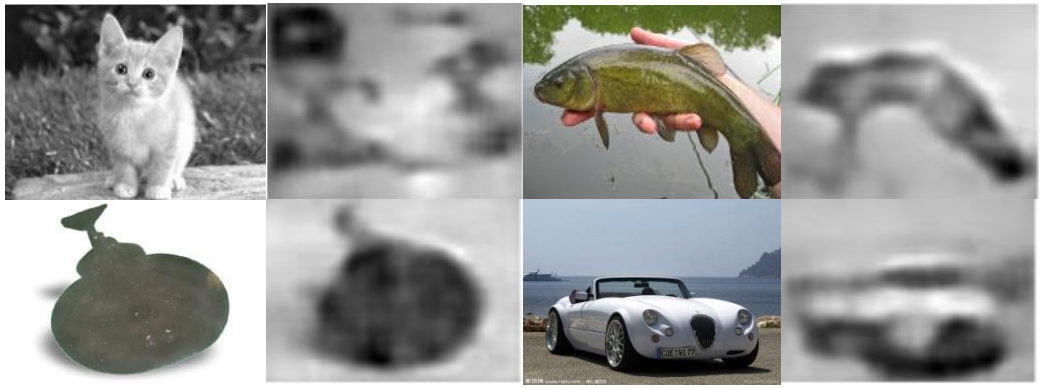
- Transferable
  - On mnist dataset



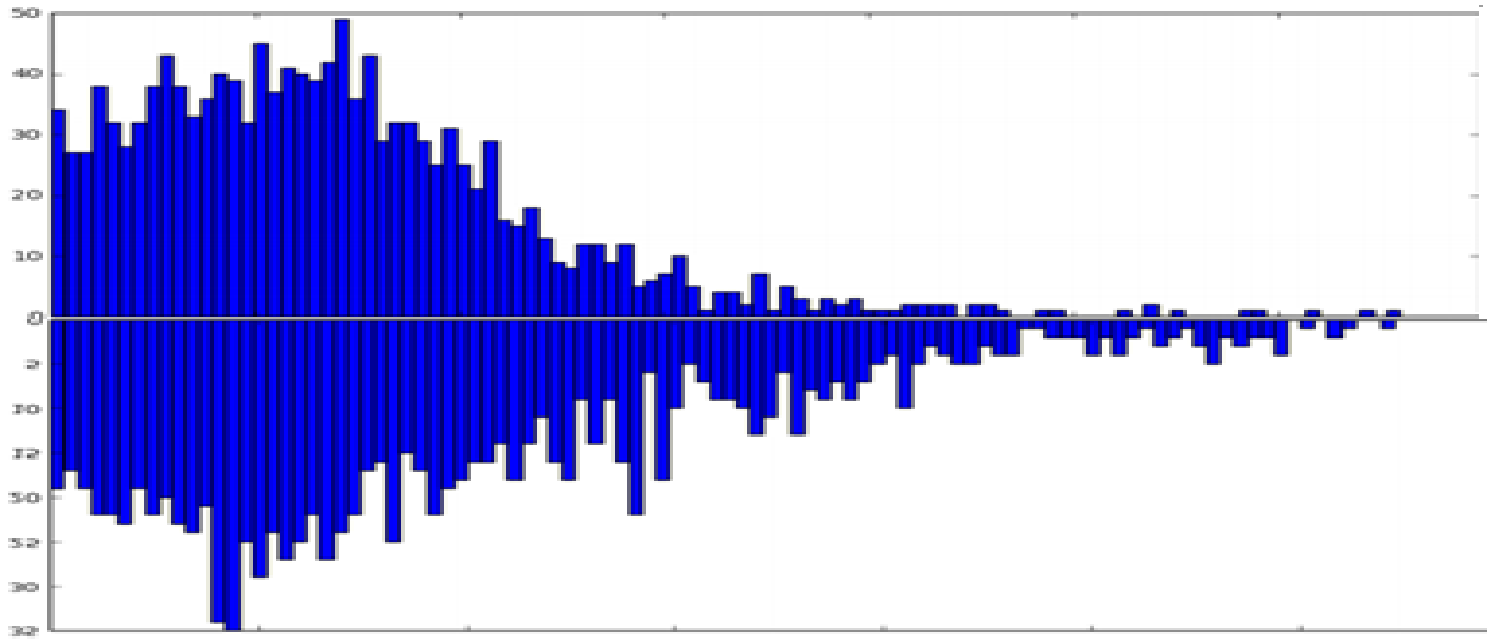
  - On Kaggle

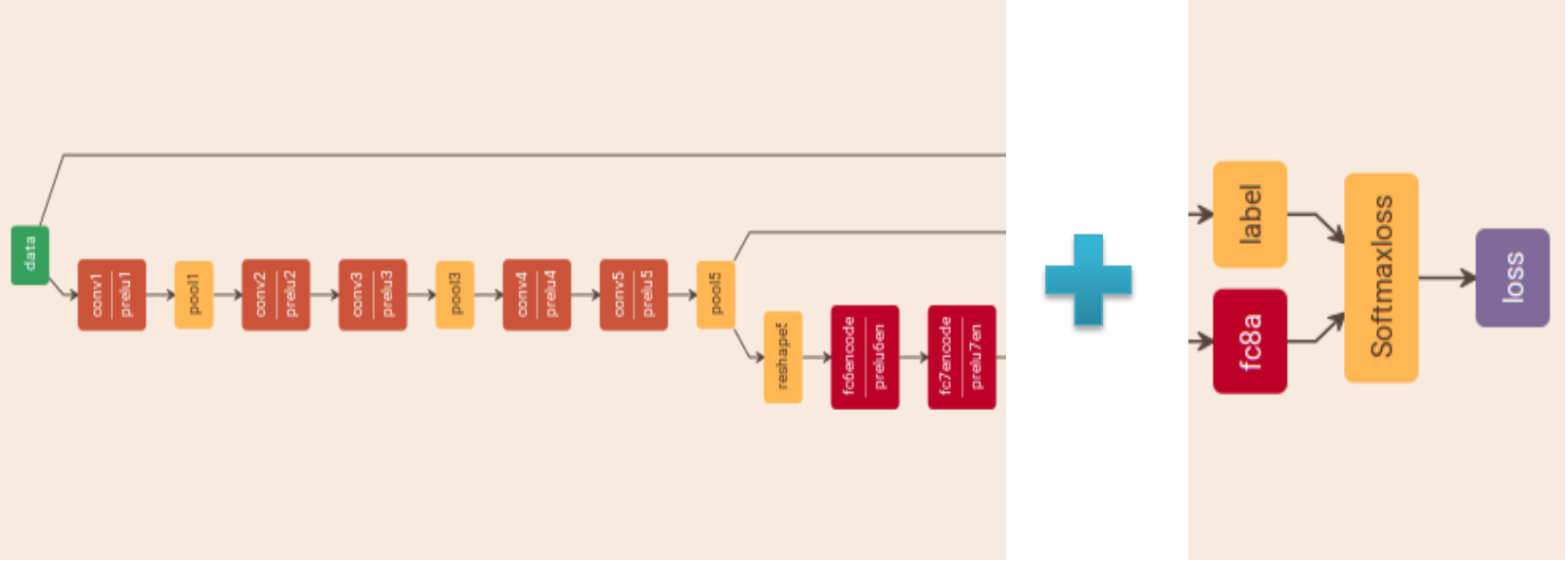


  - On imagenet



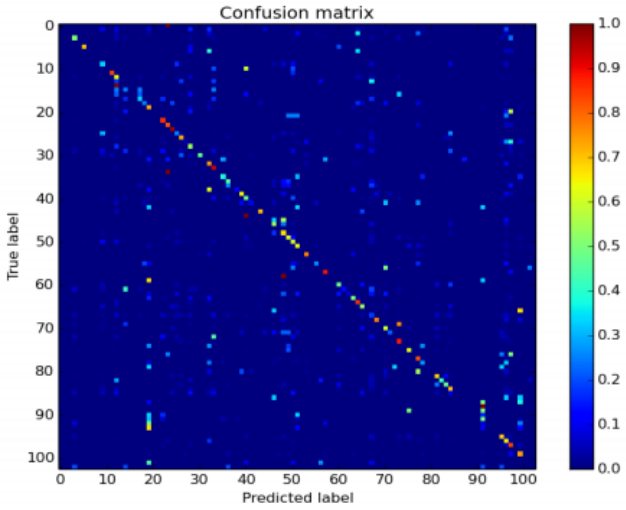
Fc encoder & decoder




data
conv1
prelu1
pool1
conv2
prelu2
conv3
prelu3
pool3
conv4
prelu4
conv5
prelu5
pool5
reshape1
fc6encode
prelu6en
fc7encode
prelu7en
label
fc8a
Softmaxloss
loss

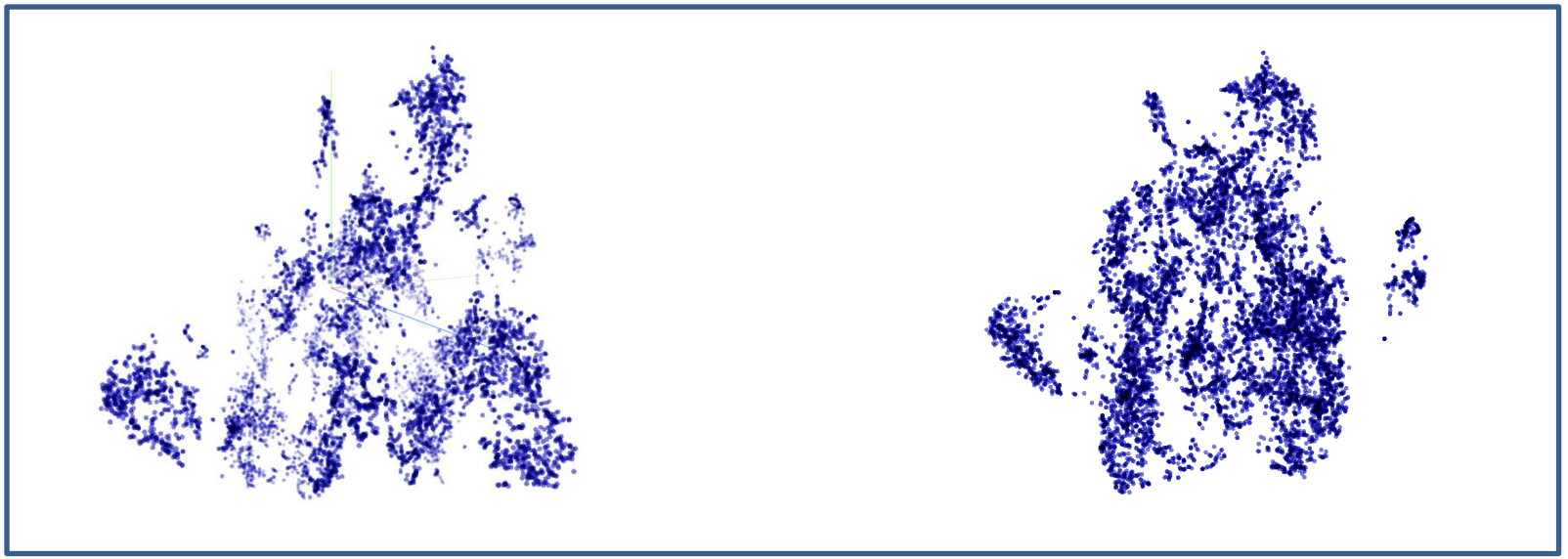
Classification result

Database	Model	Accuracy	F1 score
Full+Sample	**CAE(encode)**	**0.8530**	**0.2128**
Sample	Cifar10 net	0.8271	0.3086
Sample	Alexnet	0.8807	0.4211
Sample	Vgg16	0.8972	0.4919



Encode visualization(T-SNE)

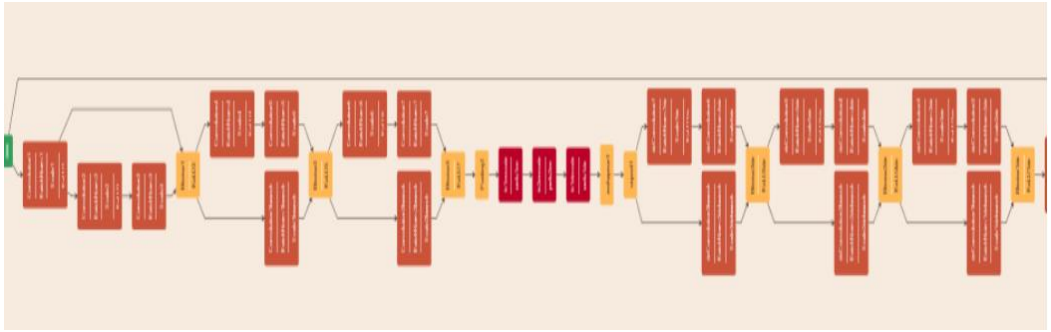
- 3D 2D



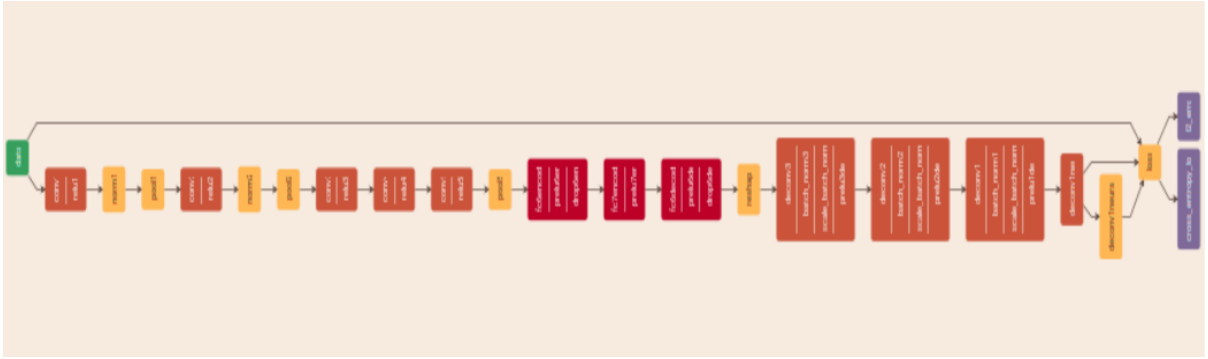
- Tools: http://projector.tensorflow.org/

CAE && beyond

- Program1（based on resnet）



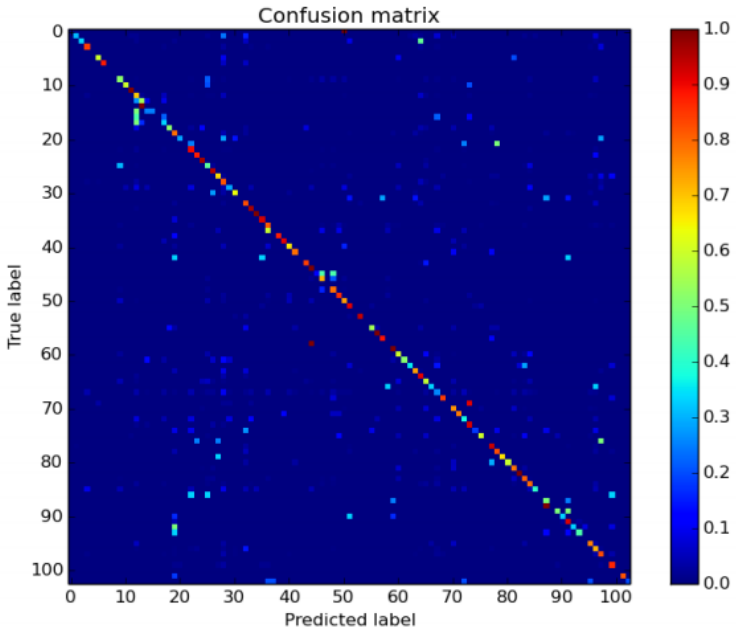
- Program2 （general encode decode）



Experiment result

- General encode & decode（alexnet + decode）

dataset	accuracy	precision	recall	F1 score
sample	0.886216	0.3765	0.5961	0.4110



Why not work well？

- Mix classes
- Fine-grained challenge
- Unsupervised learning（clustering）

What's Next

- Shallow CAE + general CNN
- CAE（Another architecture）
- Data enhancement（Use DCGAN）

Q & A