

Applications of Machine Learning in Relativistic Heavy Ion Collisions

Huichao Song

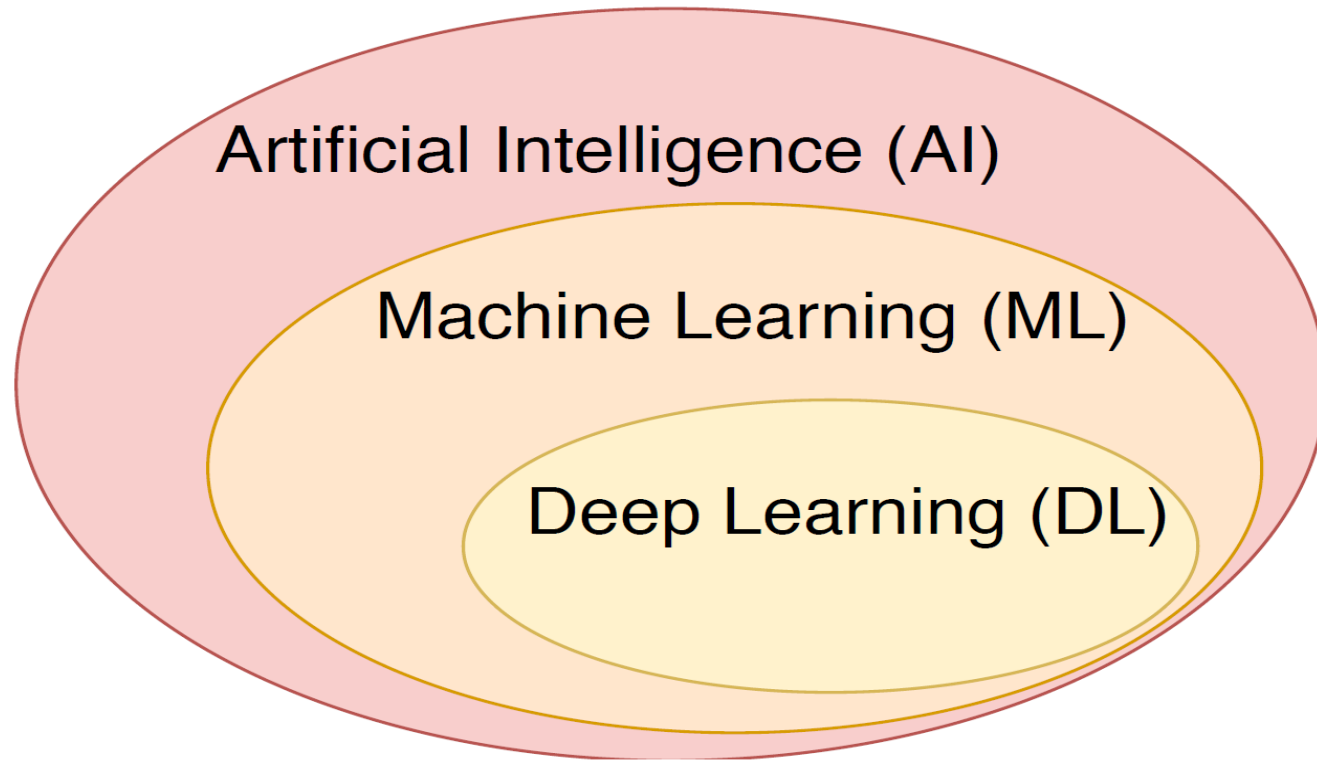
宋慧超

Peking University

**The 7th Asian Triangle Heavy-Ion Conference
(ATHIC-2018) Nov. 3-6 Hefei**

Nov 4, 2018

What is Machine Learning / Deep Learning?



AI : the broadest term, applying to any technique that enables computers to mimic human intelligence.

ML: A subset of AI aiming at optimizing a performance criterion using example data or past experience, but without explicit instruction.

DL: A subset of ML aiming at understanding high-level representations of data using a deeper structure of multiple processing layers

Broad Applications of Machine Learning

Computer vision

- Image identification
- Image style transition
- Image generation

... ..

Language processing

- Machine translation
- Speech recognition
- Chinese poetry generation

... ..

Playing Games

- AlphaGo (by Google DeepMind)

... ..

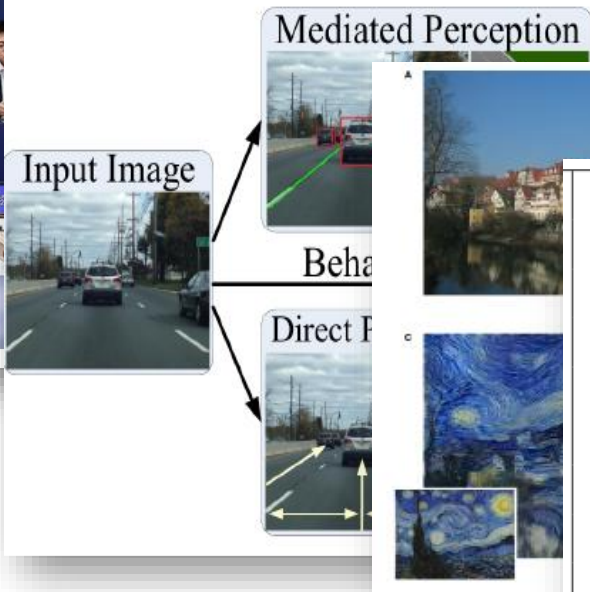
Autonomous Driving

... ..



秋夕湖上
By a Lake at Autumn Sunset
荻花风里桂花浮，
The wind blows reeds with osmanthus flying,
恨竹生云翠欲流。
And the bamboos under clouds are so green as if to flow down.
谁拂半湖新镜面，
The misty rain ripples the smooth surface of lake,
飞来烟雨暮天秋。





秋夕湖上
By a Lake at Autumn Sunset
荻花风里桂花浮，
The wind blows reeds with osmanthus flying,
恨竹生云翠欲流。
And the bamboos under clouds are so green as if to flow down.
谁拂半湖新镜面，
The misty rain ripples the smooth surface of lake,
飞来烟雨暮天愁。
And I feel blue at sunset .

Categories:

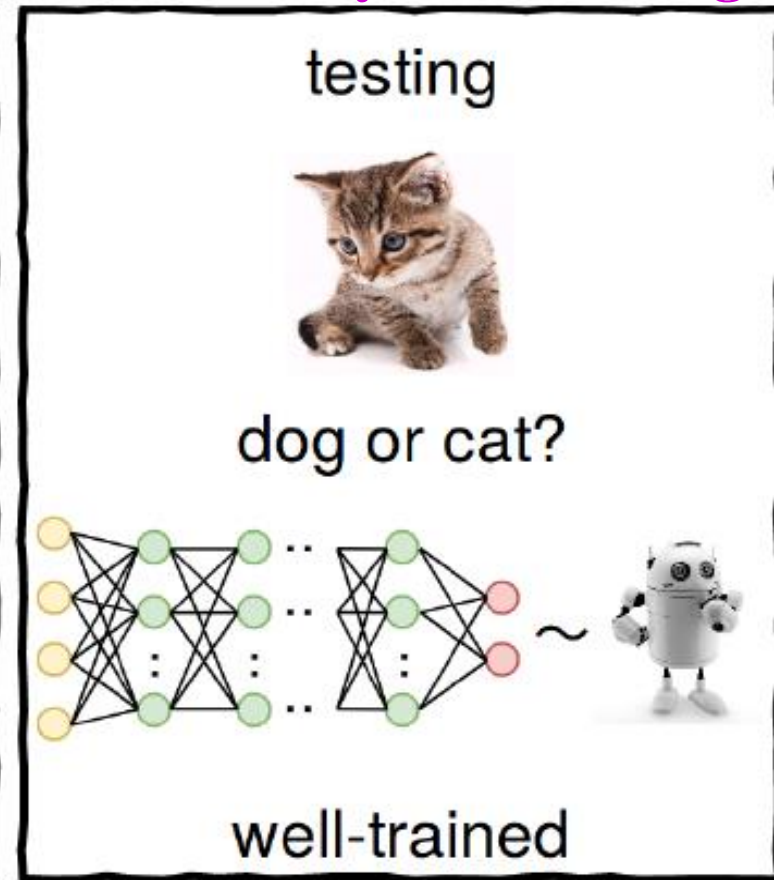
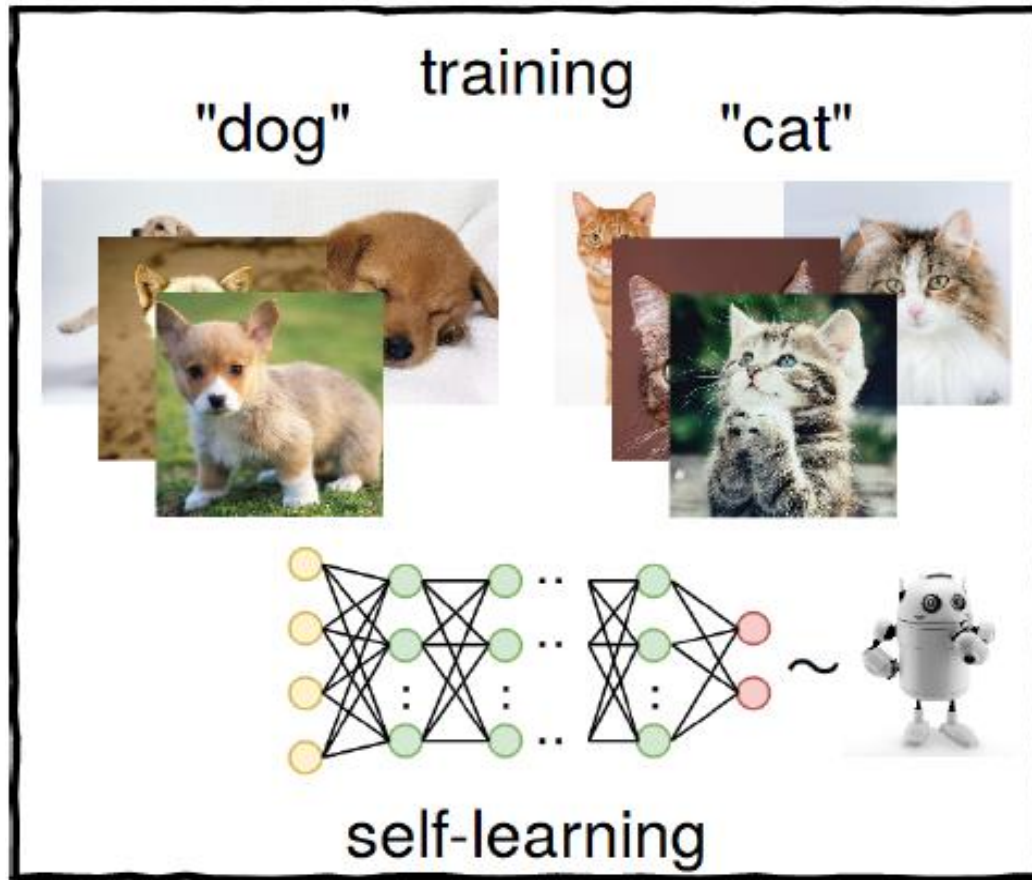
- Supervised learning
- Unsupervised learning
- Reinforcement learning

... ..

Ian Goodfellow, Yoshua Bengio, and
Aaron Courville,
<http://www.deeplearningbook.org>
MIT Press, 2016

An example of **Supervised Learning**

-Identify cats and dogs

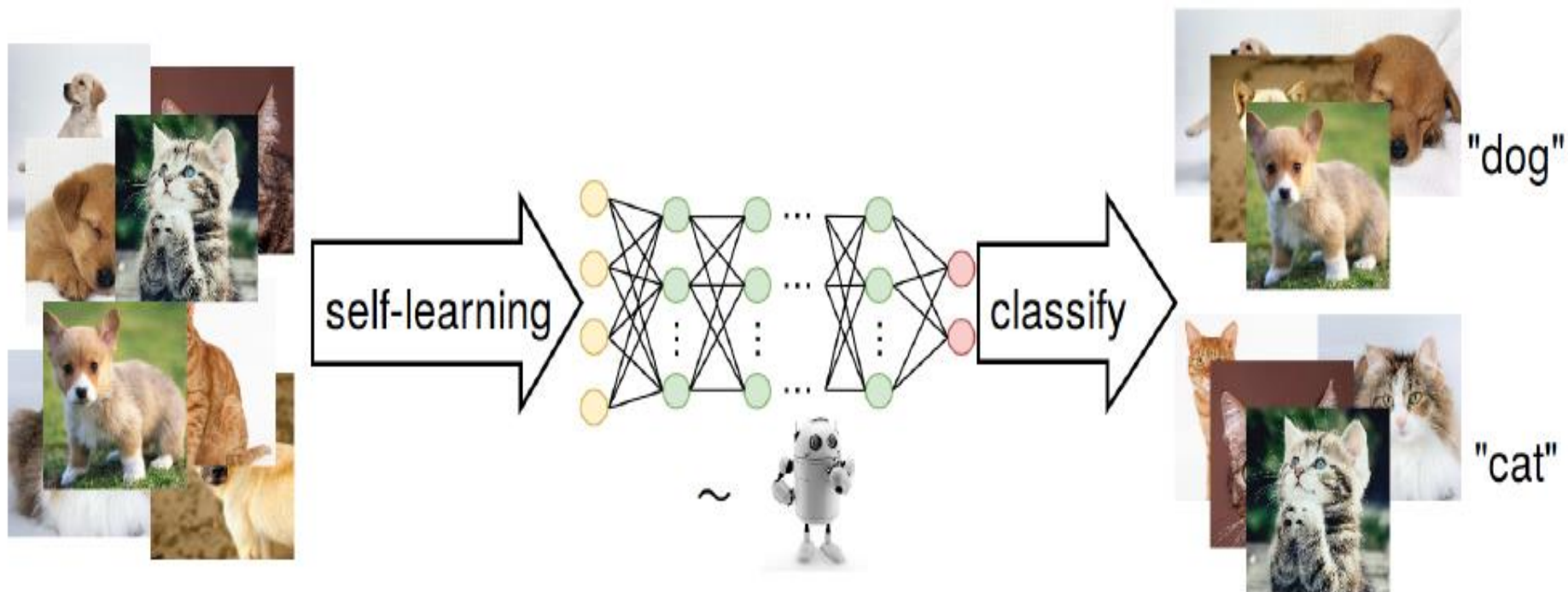


Supervised learning:

Training on a dataset contains many features and associated with a label or target.

An example of Unsupervised Learning

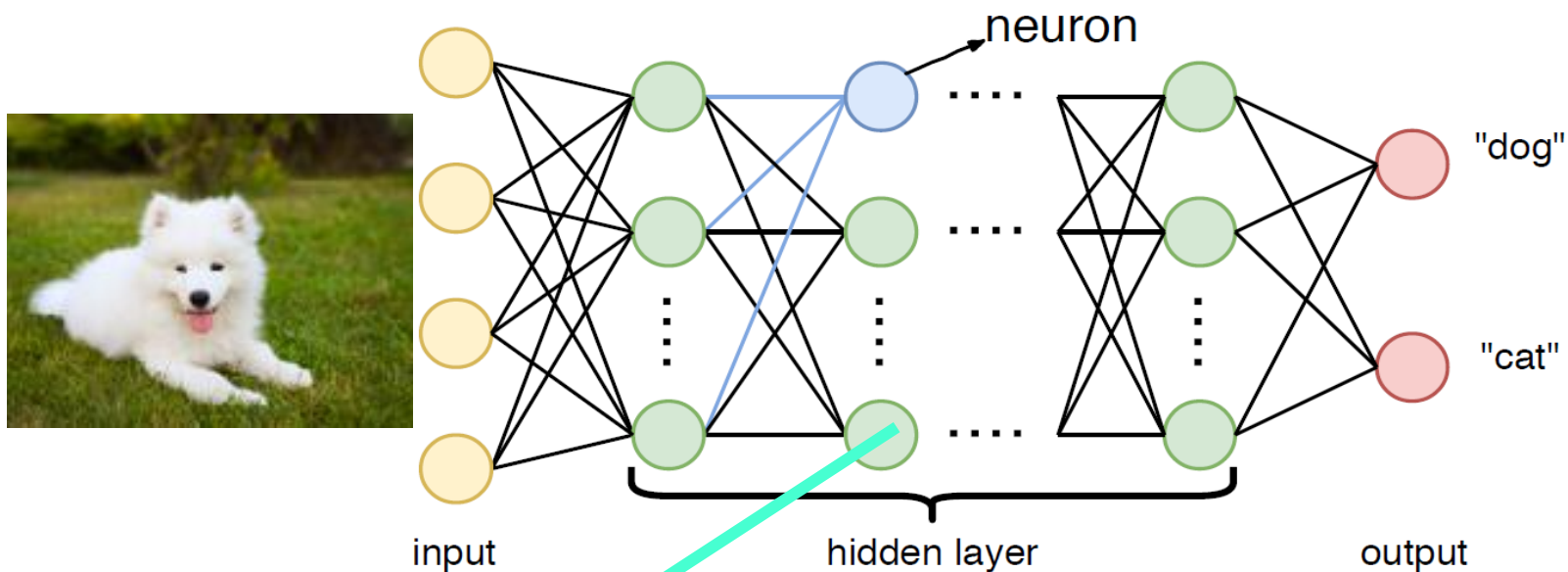
-Classify cats and dogs



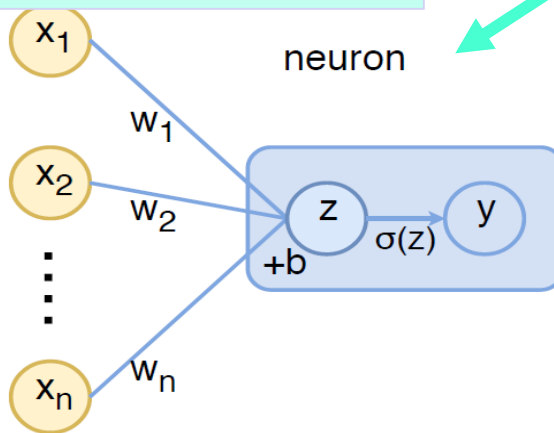
Unsupervised learning

-experience a dataset contains many features but **without labels**, and learn useful properties of the structure of this dataset.

Deep Neural Network



Neuron



Linear operation

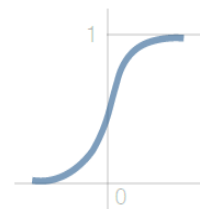
$$z_j = \sum_{i=1}^N x_i w_{ij} + b_j$$

scaling, rotating, boosting,
changing dimensions

Non-linear activation function

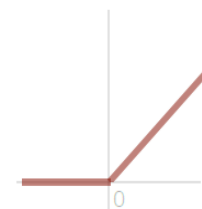
(a) Sigmoid

$$\sigma(z) = \frac{1}{1 + \exp(-z)}$$



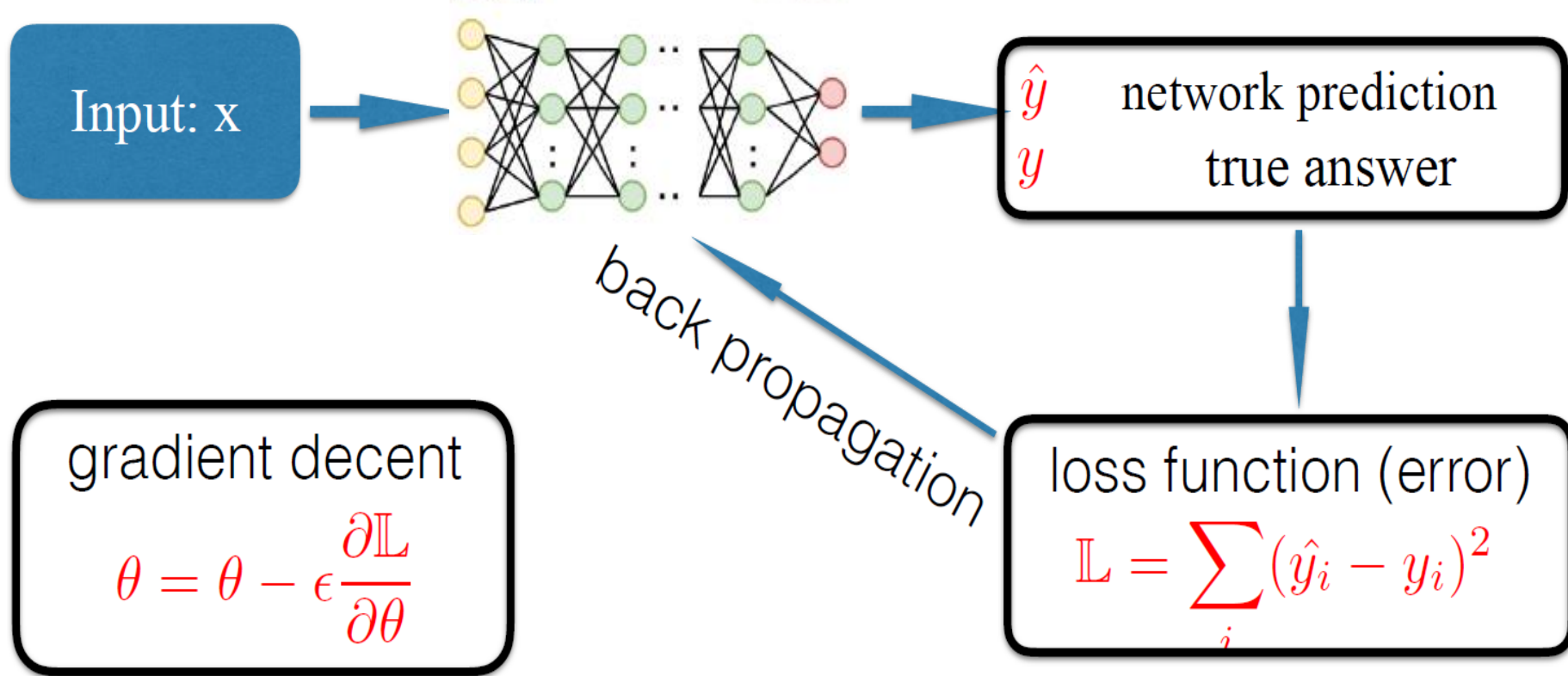
(b) ReLU

$$\sigma(z) = \begin{cases} z, & z > 0 \\ 0, & z \leq 0 \end{cases}$$



Deep Neural network

-Loss function, back propagation & gradient decent

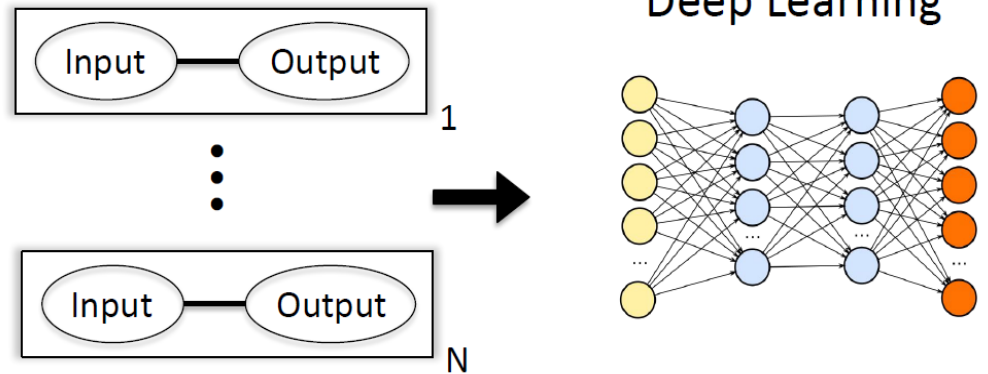


-Deep neural network can reduce fitting error by updating model parameters through back propagation and gradient decent.

Applications of Machine Learning in Physics

Why Deep Learning in Physics?

“Unlike earlier attempts ... Deep Learning systems can see patterns and spot anomalies in data sets far larger and messier than human beings can cope with.”



Can “**Black-box**” models learn patterns and models solely from data without relying on scientific knowledge?



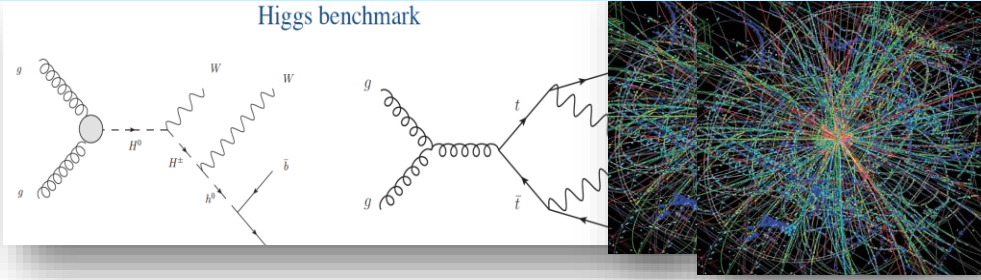
Applications of Deep Learning in Physics

- Y. D. Hezaveh, L. Perreault Levasseur and P. J. Marshall, Nature 548, 555 (2017)
- J. Carrasquilla and G. R. Melko, Nature Phys. 13, 431 (2017)
- Carleo et al., Science 355, 602-606 (2017)
- E. P. L. van Nieuwenburg, Y. H. Liu, S. Huber, Nature Phys. 13, 435 (2017)
- Pierre Baldi, Peter Sadowski, and Daniel Whiteson, Nature Commun. 5 (2014) 4308
- Luke de Oliveira, Michela Paganini, and Benjamin Nachman, Comput Softw Big Sci (2017) 1: 4
- Long-Gang Pang et al., Nature Commun. 9 (2018) no.1, 210
- . . . , . . . ,
- . . .



Searching for Exotic Particles in High-Energy Physics

Higgs benchmark



Deep learning can improve the power for the collider search of exotic particles

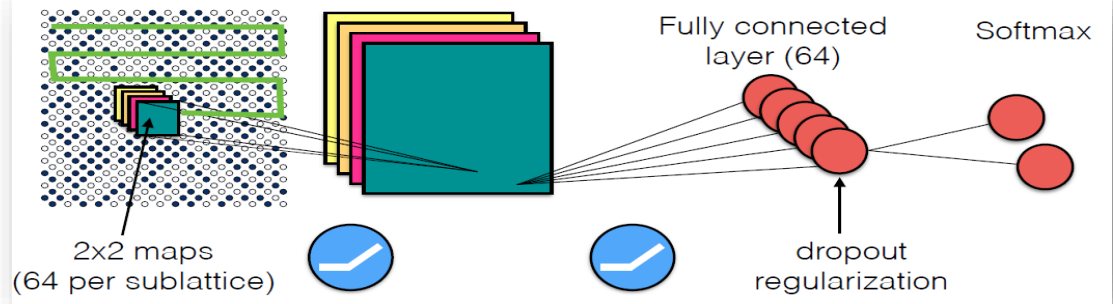
P.Baldi, P.Sadowski, & D. Whiteson *Nature Commun.* 5, 4308 (2014)

Classifying the Phase of Ising Model

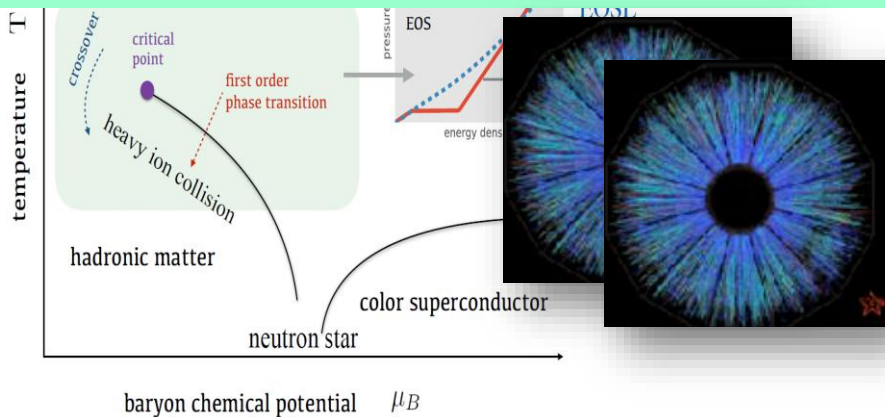
For the case of Ising gauge theory

$$H = -J \sum_p \prod_{i \in p} \sigma_i^z$$

J. Carrasquilla and R. G. Melko. *Nature Physics* 13, 431–434 (2017)



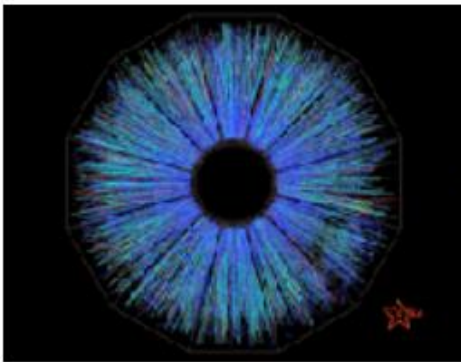
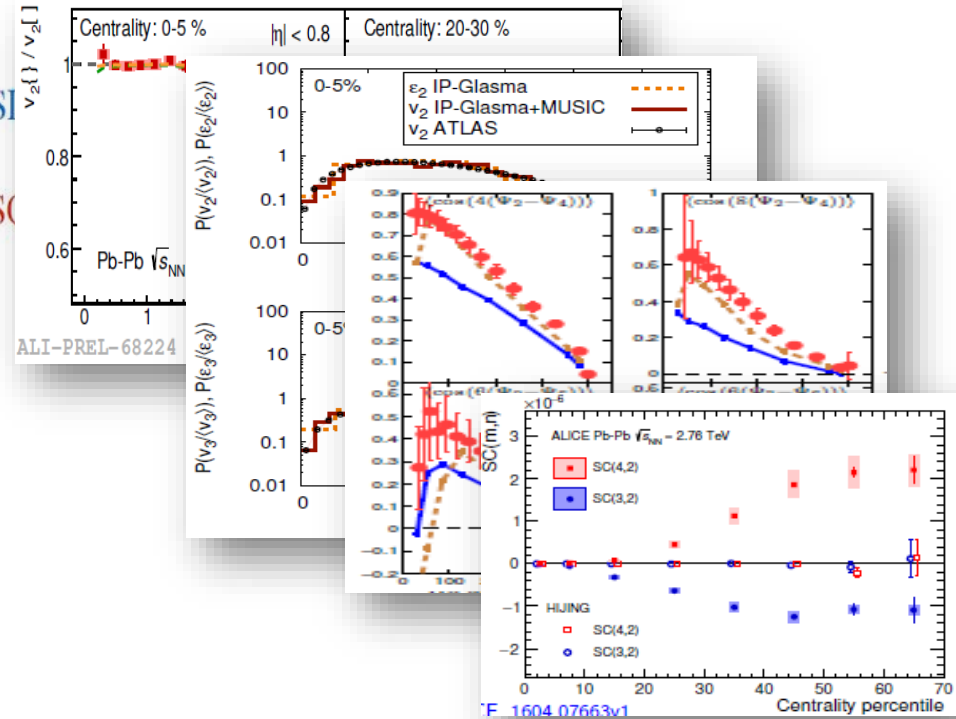
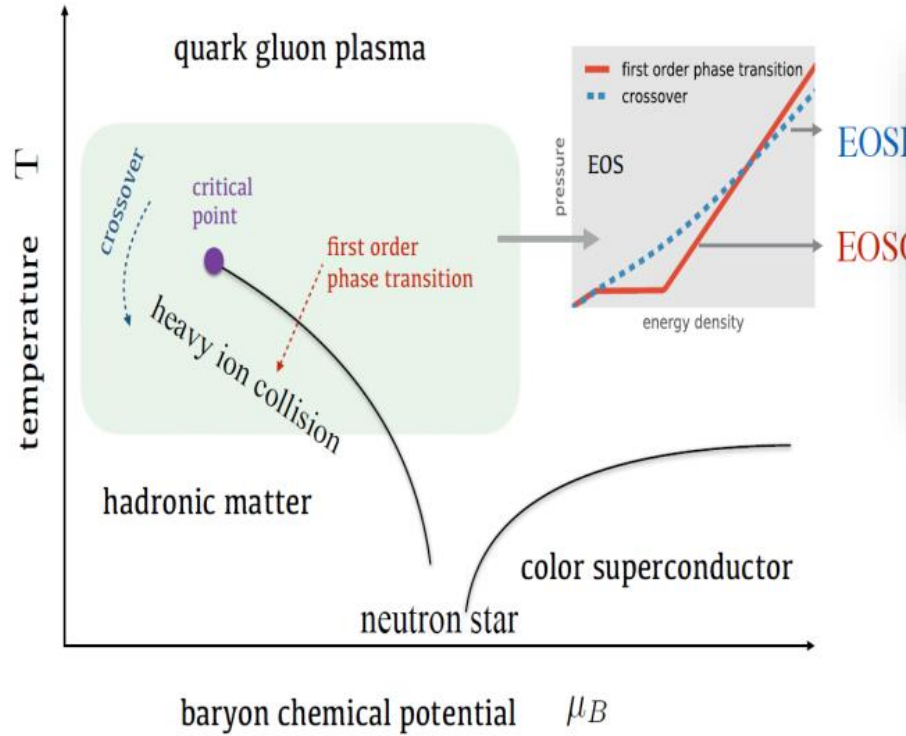
Identify QCD Phase Transition with Deep Learning



DNN efficiently decode the EOS information from the complex final particle info event by event

LG. Pang, K.Zhou, N.Su, H.Petersen, H. Stoecker, XN. Wang. *Nature Commun.* 9 (2018) no.1, 210

Identify QCD Phase Transition with Deep Learning



$$\rho(p_T, \Phi)$$

Motivation:

- Traditionally, the properties of the QCD matter are extracted from the event averaged observables
- Can deep learning identify different EoS from the raw data of heavy ion collisions?

LG. Pang, K.Zhou, N.Su, H.Petersen, H. Stoecker, XN. Wang. Nature Commun.9 (2018) no.1, 210

Identify QCD Phase Transition with Deep Learning

A) Generating training/testing data:

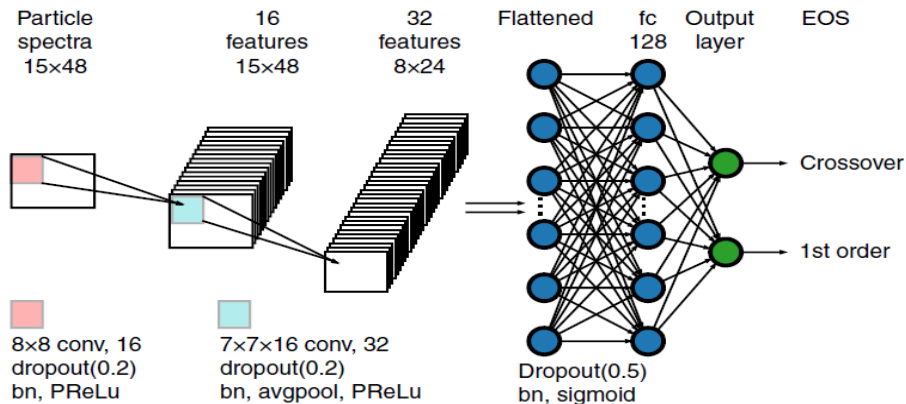
- Run Hydro with EOS L and EOS Q
- particle spectra - image (15*48 pixels)

$$\rho(p_T, \phi) \equiv \frac{dN_i}{dY_{p_T} dp_T d\phi} = g_i \int_{\sigma} p^{\mu} d\sigma_{\mu} f_i,$$

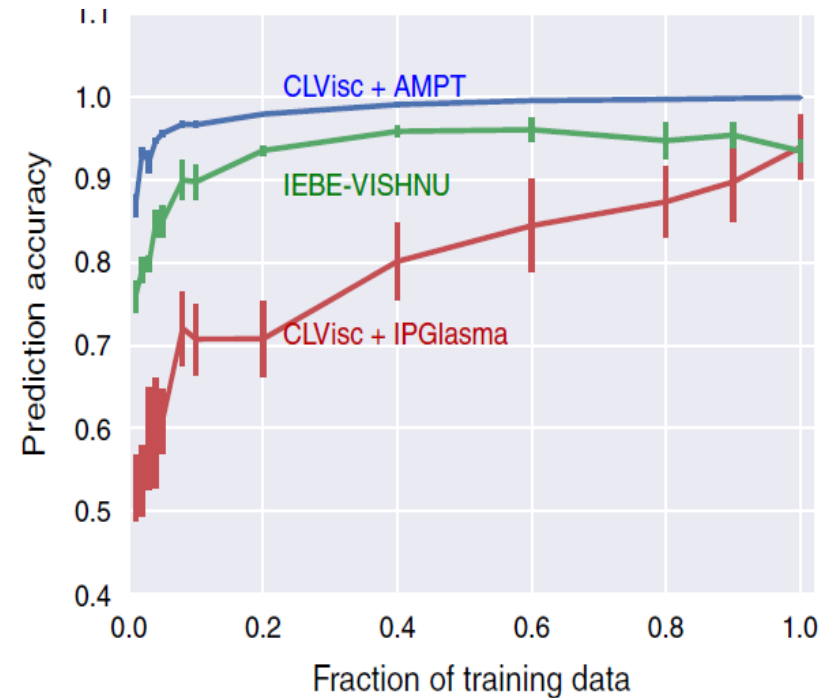
B) Training CNN

Table 1 The training data set **Hydro CLVis (AMPT)**

| Training data set | $\eta/s = 0$ | | $\eta/s = 0.08$ | |
|----------------------------------|--------------|------|-----------------|------|
| | EOSL | EOSQ | EOSL | EOSQ |
| Au-Au $\sqrt{s_{NN}} = 200$ GeV | 7435 | 5328 | 500 | 500 |
| Pb-Pb $\sqrt{s_{NN}} = 2.76$ TeV | 4967 | 2828 | 500 | 500 |



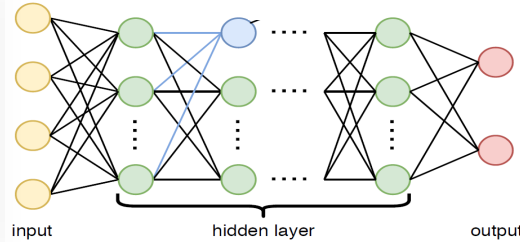
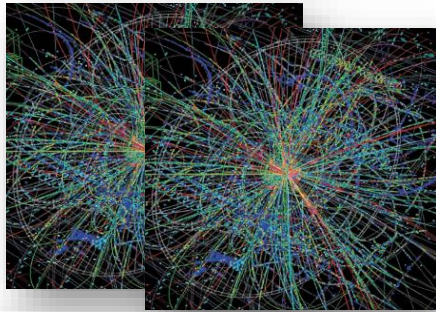
C) testing the trained net work



One can efficiently decode the EOS information from the complex final particle info event by event using deep learning

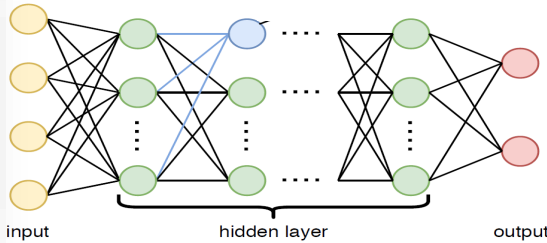
LG. Pang, K.Zhou, N.Su, H.Petersen, H. Stoecker, XN. Wang. Nature Commun.9 (2018) no.1, 210

Image identification



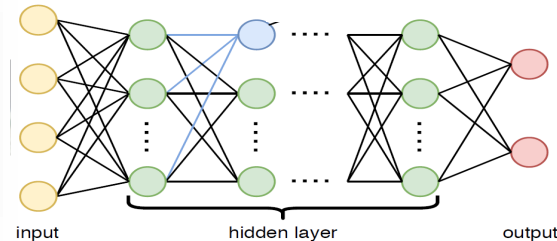
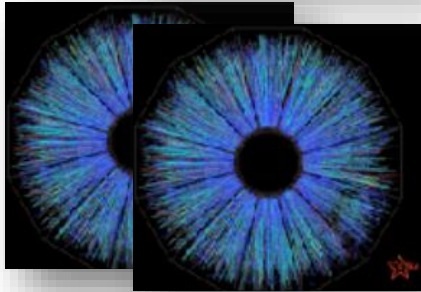
Higgs signal or background?

P.Baldi, et al, Nature Commun.(2014)



High temperature or low temperature phase?

Carrasquilla & Melko. Nature Physics (2017)

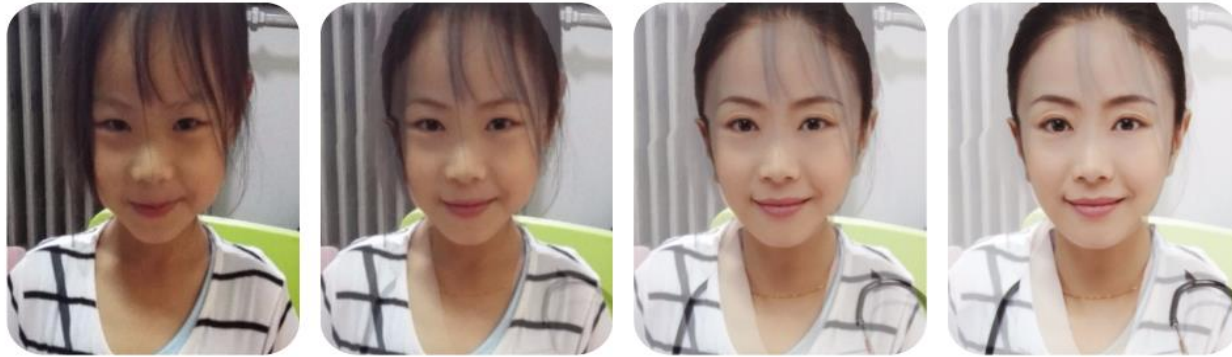


EoS L or EOSQ ?

Pang, et al Nature Commun.(2018)

“Unlike earlier attempts ... Deep Learning systems can see patterns and spot anomalies in data sets far larger and messier than human beings can cope with.”

Image generation



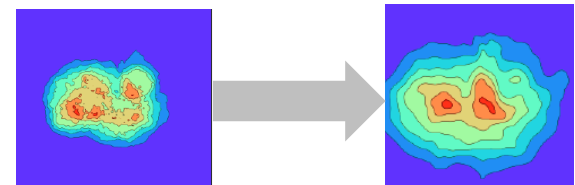
没毛病，比我还好看

提前看到2038年宝宝的样子

For hydrodynamics can we use deep learning to learn/predict the pattern transformation between initial and final profiles?

Initial energy density profiles

----- > final energy density velocity profiles



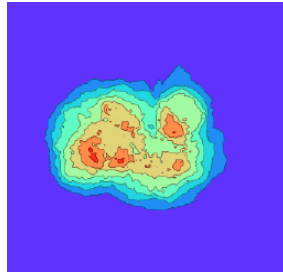
For the non-linear hydro system, can the **black-box** network could learn pattern transformations solely from data without relying on scientific knowledge?

(conservation laws)

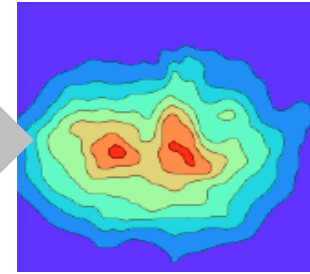
Applications of deep learning to relativistic hydrodynamics

**H.Huang, B.Xiao, H.Xiong, Z.Wu, Y. Mu and H.Song arXiv:
1801.03334; NPA2018**

Traditional hydrodynamics

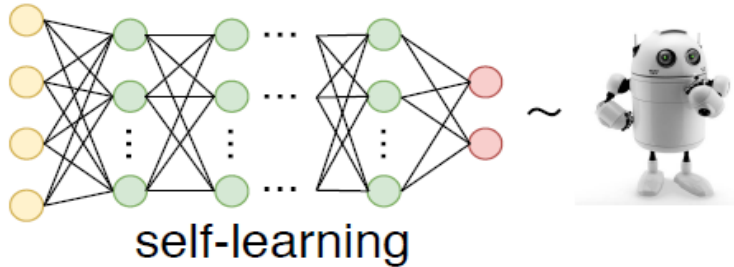
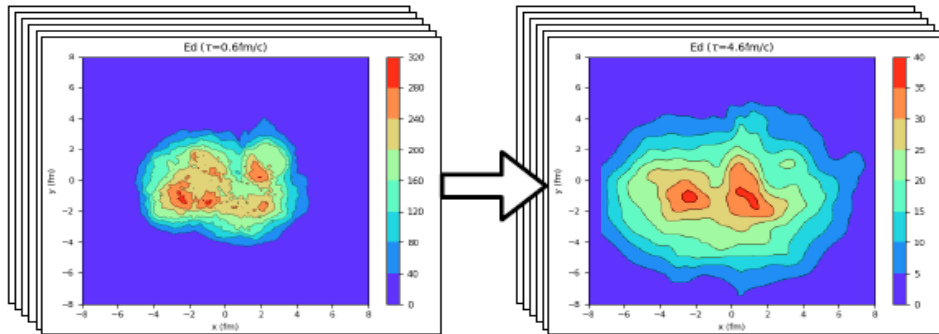


$$\partial_{\mu} T^{\mu\nu}(x) = 0$$

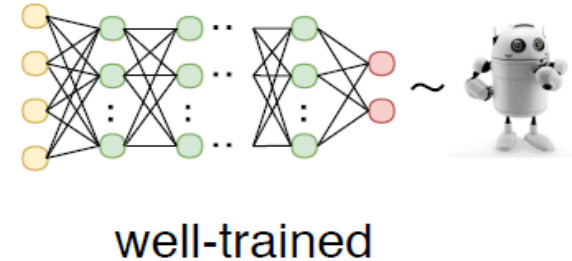
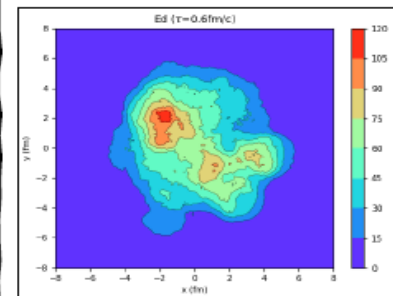


Deep Learning

training



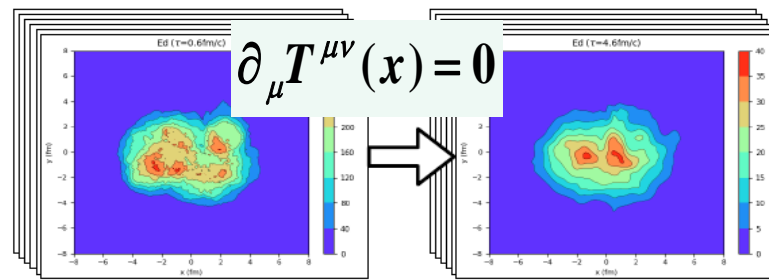
testing



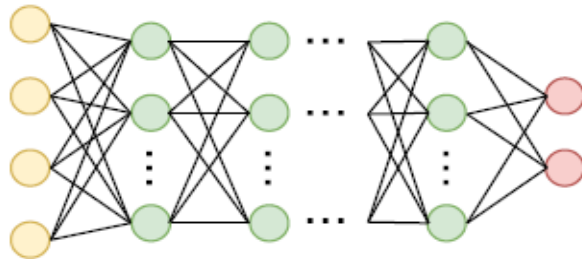
-Such deep learning systems do not need to be programmed with the hydro equation $\partial_{\mu} T^{\mu\nu}(x) = 0$ Instead, they learn on their own

Deep Learning

Step1) Generate the training/testing data sets from hydro

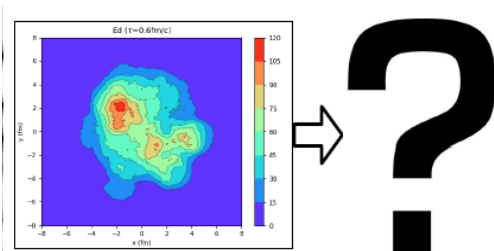


Step2) Design & train the deep neural network



| The Training Data Sets | |
|------------------------|-------|
| hydro VISH2+1 | MC-GI |
| | 10000 |

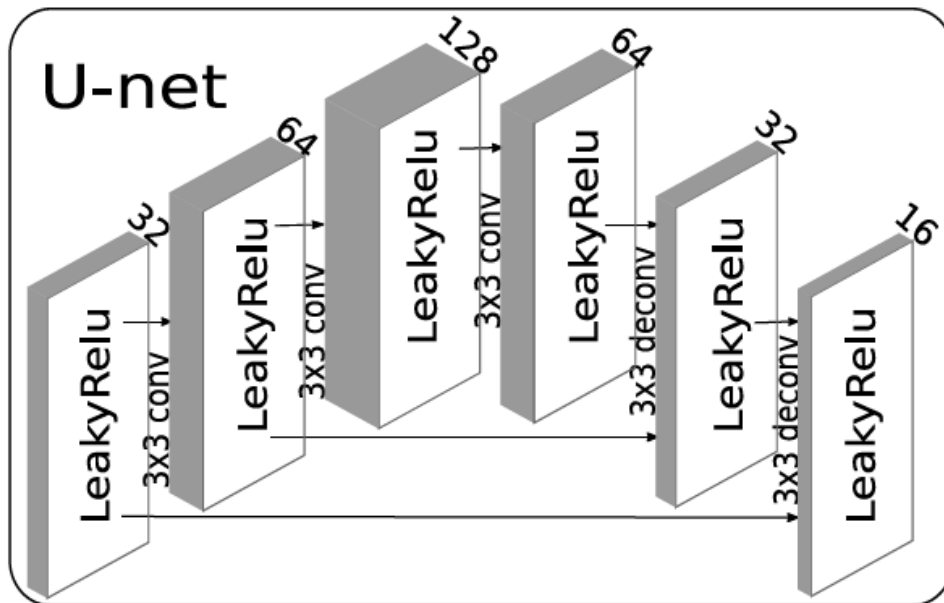
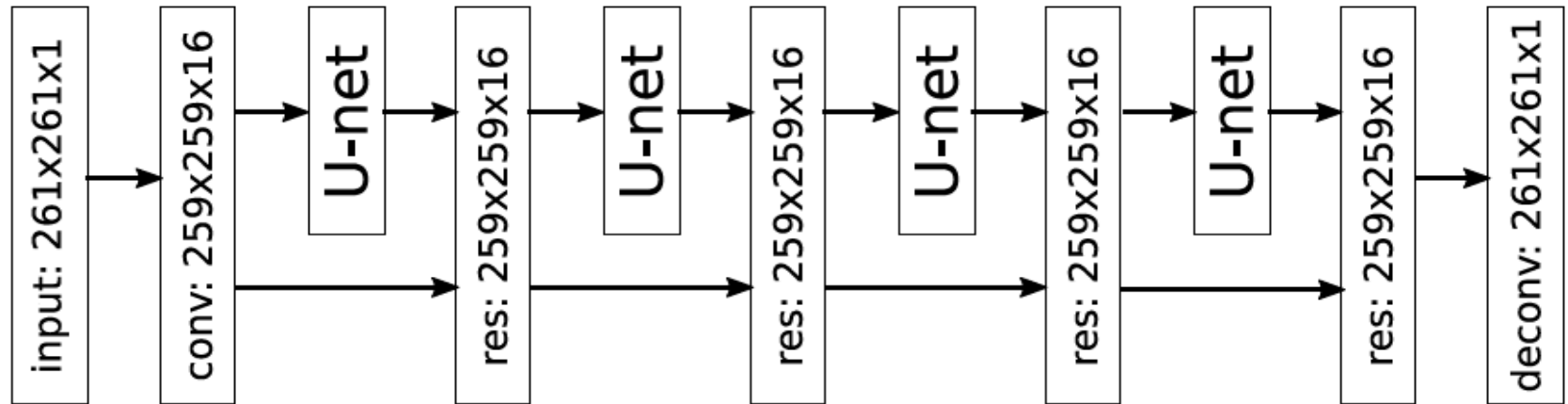
Step3) Test the deep neural network



| The Testing Data Sets | | | | |
|-----------------------|-------|--------|-------|--------|
| hydro VISH 2+1 | MC-GI | MC-KLN | AMPT | Trento |
| | 10000 | 10000 | 10000 | 10000 |

Stacked U-net for 2+1-d hydro

Stacked U-net



The activation function:

$$\text{Leaky ReLU } f(x) = \max\{x, 0.03x\}$$

The loss function:

normalized MAE loss $Loss = \frac{|y_1 - y_0|}{|y_0|}$

H.Huang, B.Xiao, H.Xiong, Z.Wu, Y. Mu and H.Song arXiv: 1801.03334, NPA 2018

Training / Testing data sets from 2+1-d hydro

$$T^{\tau\tau}_{,\tau} + (\bar{v}_x T^{\tau\tau})_{,x} + (\bar{v}_y T^{\tau\tau})_{,y} = -\frac{p+T^{\tau\tau}}{\tau} - (p\bar{v}_x)_{,x} - (p\bar{v}_y)_{,y}$$

$$T^{\tau x}_{,\tau} + (\bar{v}_x T^{\tau x})_{,x} + (\bar{v}_y T^{\tau x})_{,y} = -p_{,x} - \frac{T^{\tau x}}{\tau}$$

$$T^{\tau y}_{,\tau} + (\bar{v}_x T^{\tau y})_{,x} + (\bar{v}_y T^{\tau y})_{,y} = -p_{,y} - \frac{T^{\tau y}}{\tau}$$

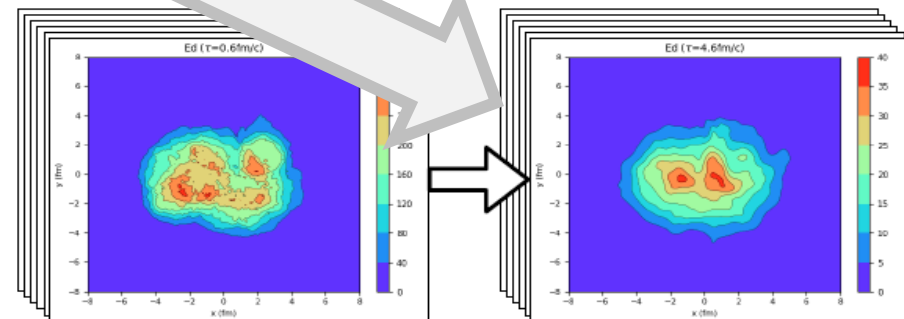
Initial conditions: MC-Glauber, MC-KLN, AMPT, Trento EoS: $p=e/3$,

hydro evolution time: $\tau - \tau_0 = 2.0, 4.0, 6.0$ fm/c

The Training Data Sets

2+1-d hydro
VISH2+1

MC-Glauber
10000 events



The Testing Data Sets

2+1-d hydro
VISH 2+1

MC-Glauber
10000 events

MC-KLN

10000 events

AMPT

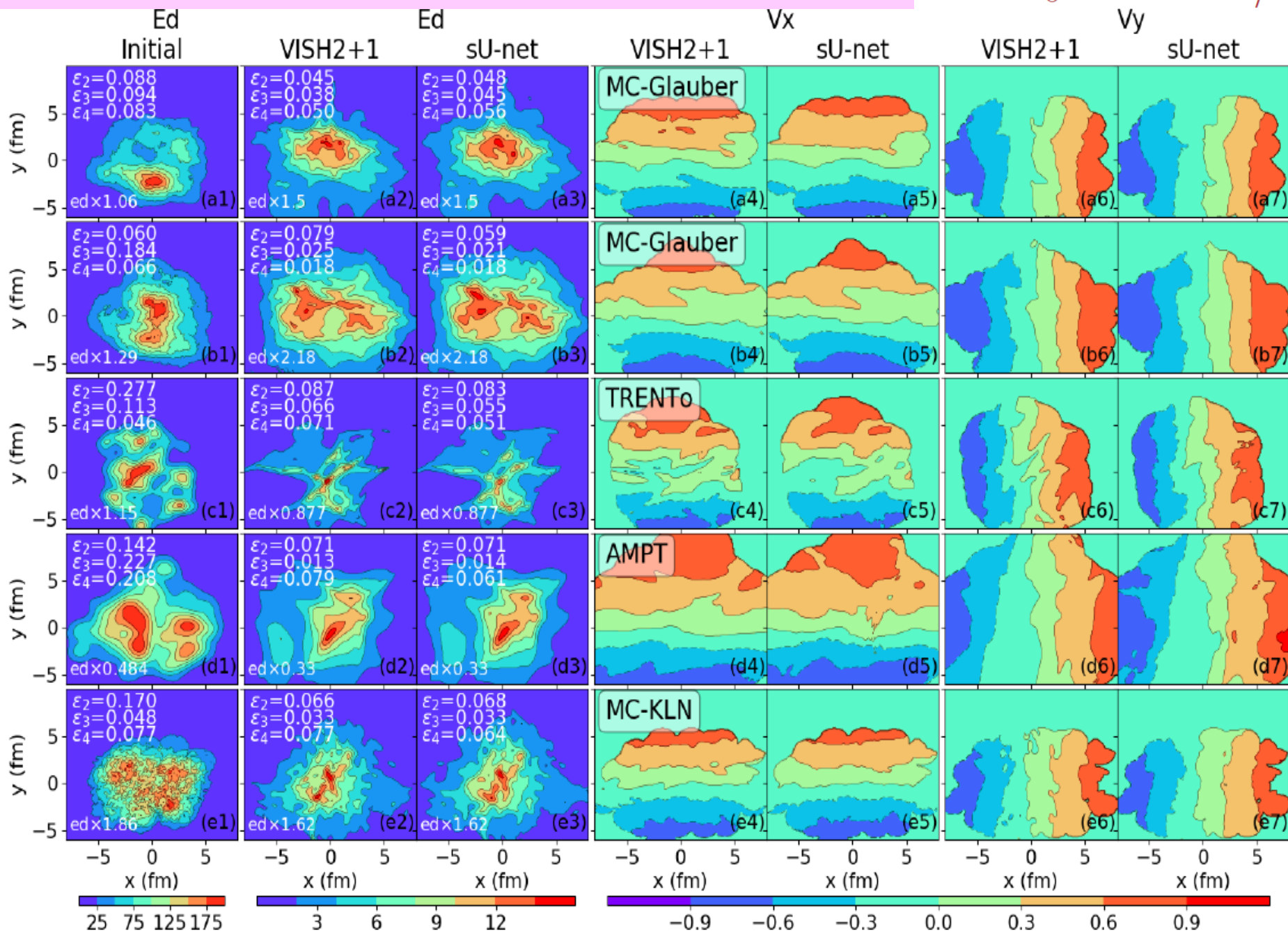
10000 events

Trento

10000 events

sUnet prediction vs. hydro simulations

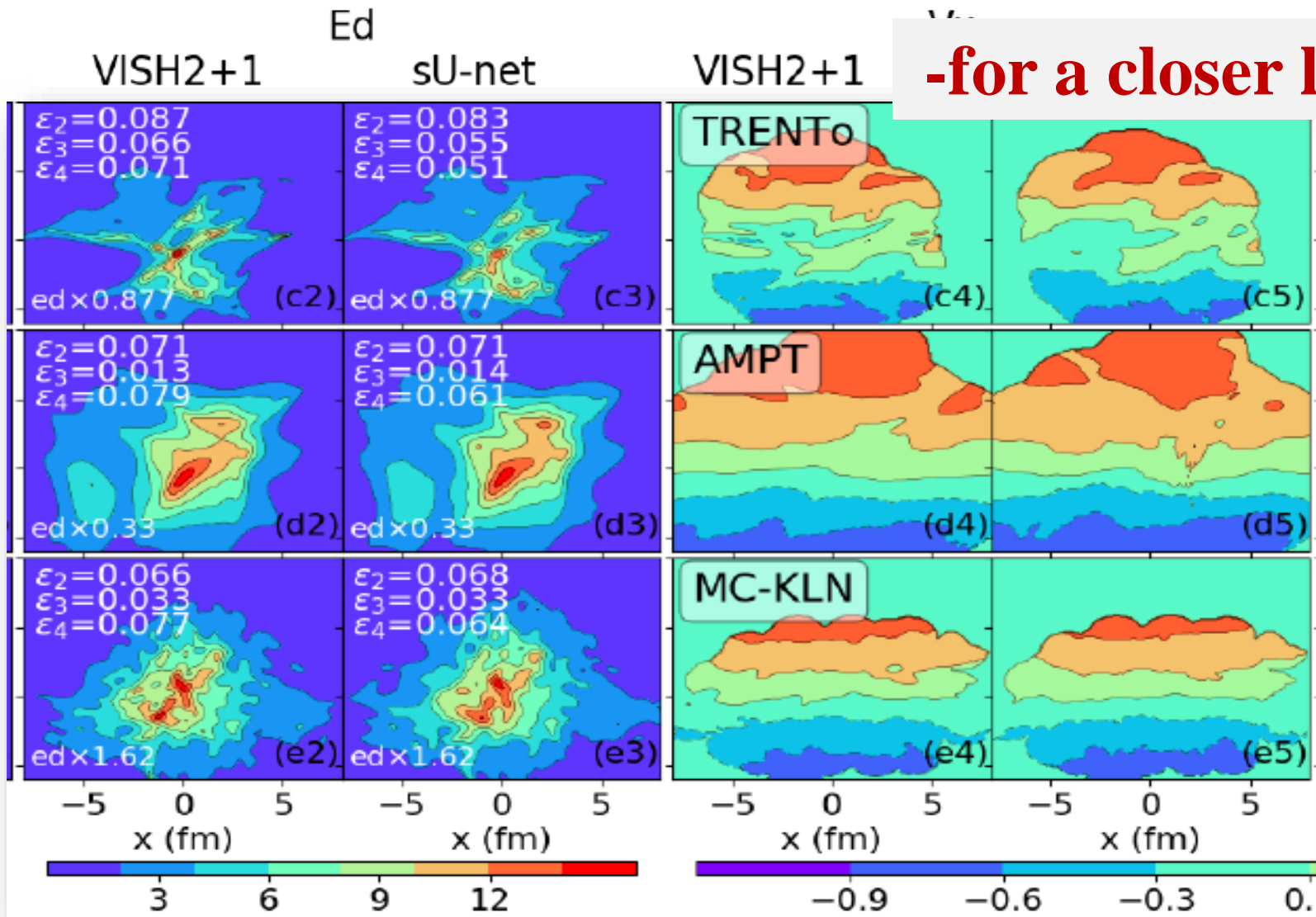
$$\tau - \tau_0 = 6.0 \text{ fm}/c$$



sUnet prediction vs. hydro simulations

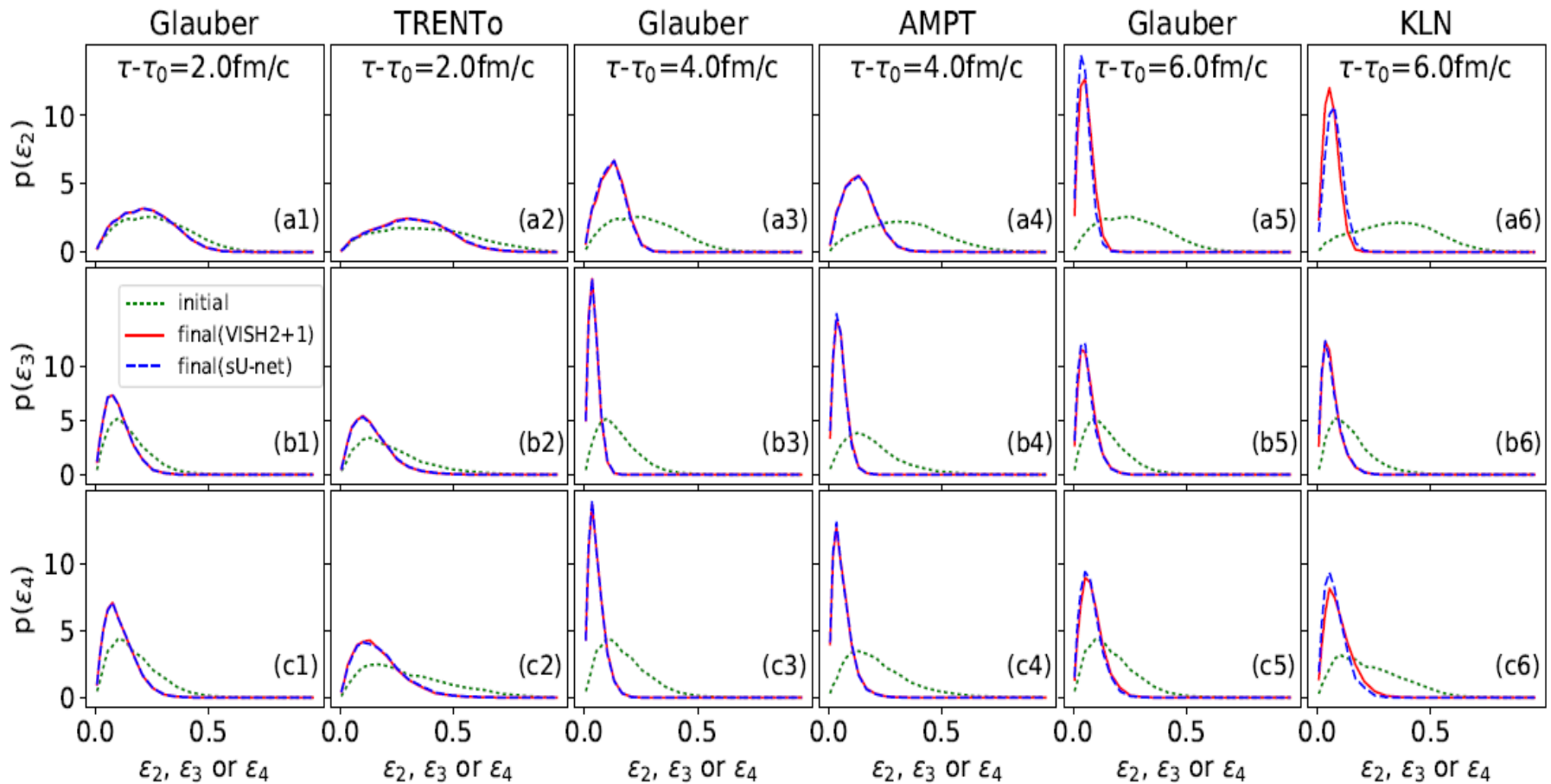
$$\tau - \tau_0 = 6.0 \text{ fm}/c$$

-for a closer look

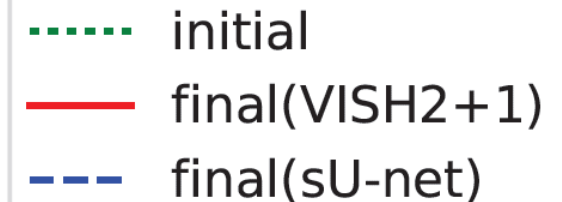


sUnet prediction vs. hydro simulations

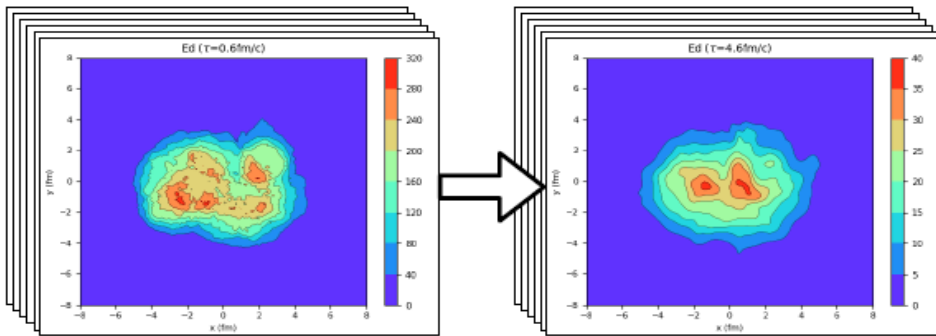
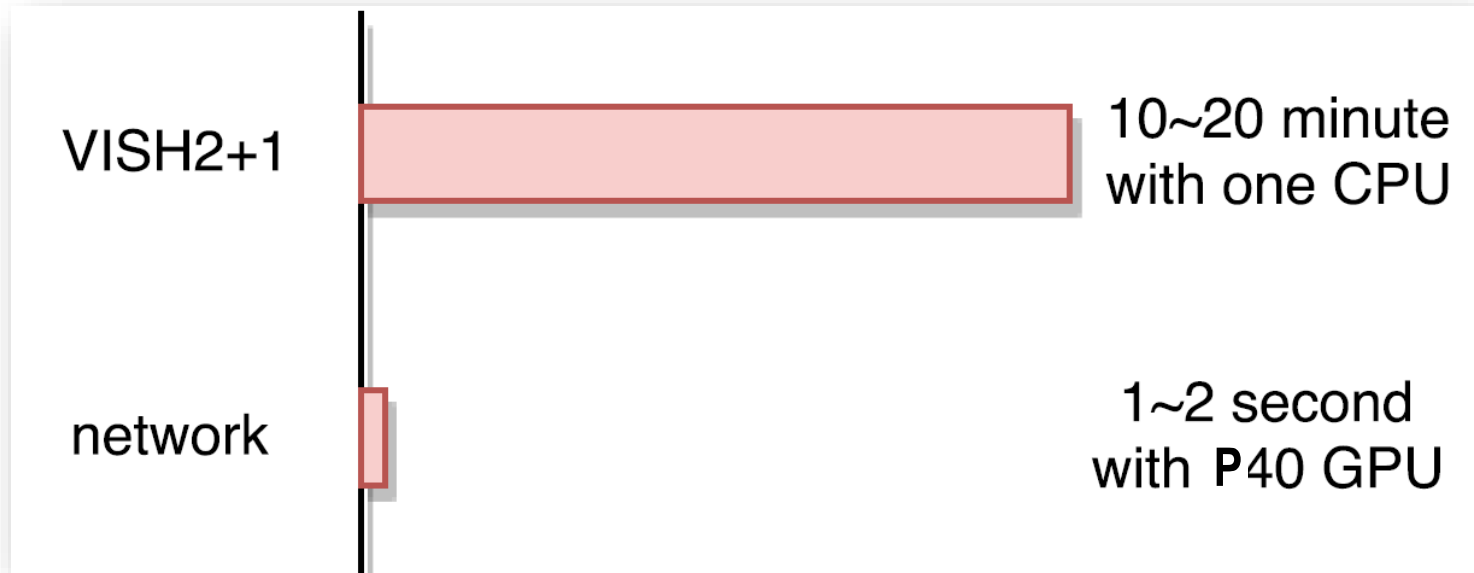
Eccentricity distributions:



$$\epsilon_n e^{in\Phi_n} = - \frac{\int dx dy r^2 e^{in\phi} e(x,y)}{\int dx dy r^2 e(x,y)}$$



Simulation time: sUnet vs. hydro



With the well trained network, the final state profiles can be quickly generated from the initial profiles. (5-10 times faster for GPU based calculations)

Principal Component Analysis

-- Can Machine Learning discover flow harmonics from complex data sets?

Z. Liu, W. Zhao, and H. Song, in preparation

What is Principal Component Analysis (PCA)

-a statistical procedure that uses an orthogonal transformation to convert a set of observations into a set of values of linearly uncorrelated variables called principal components.

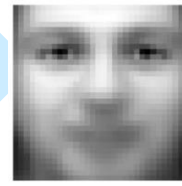
PCA for FACE analysis

Data sets: many many faces



PCA

mean μ



top eigenvectors: $\mu_1, \mu_2, \mu_3 \dots$



With PCA, each face is decomposed into superposition of eigenfaces.



=



+



$$\hat{X} = \mu + w_1 u_1 + w_2 u_2 + w_3 u_3 + w_4 u_4 + \dots$$

PCA for FACE analysis

Data sets: many many faces

top eigenvectors: $\mu_1, \mu_2, \mu_3 \dots$

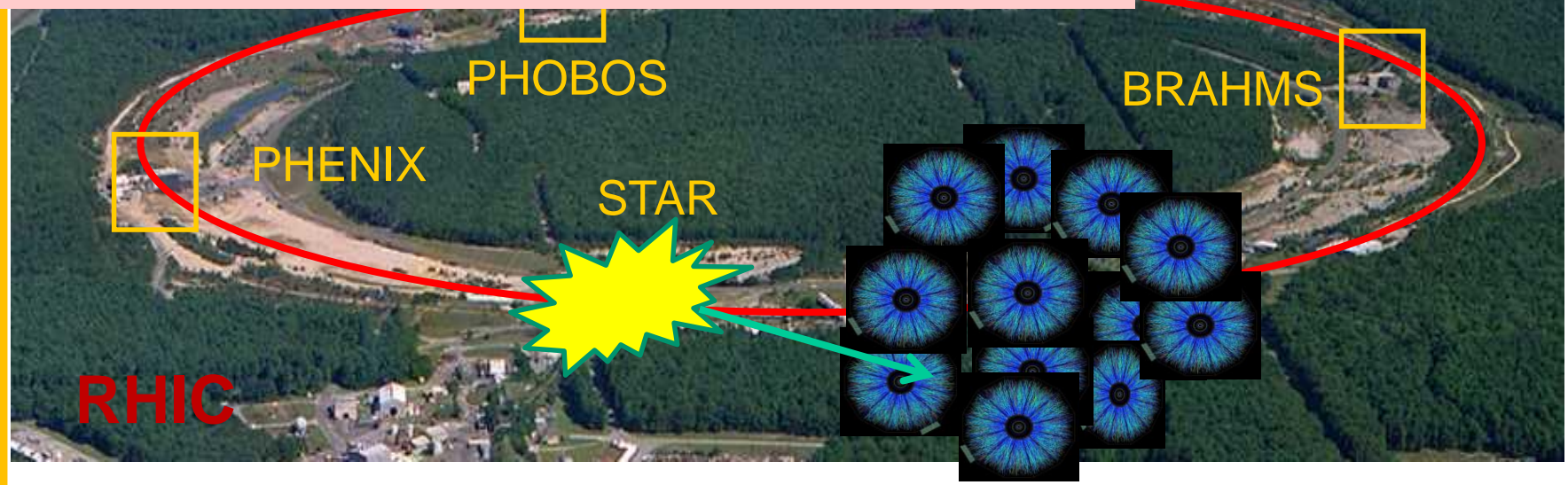


PCA

mean μ



PCA for Relativistic Heavy Ion Collisions?



Can PCA (machine) directly identify the different configurations behind the massive heavy ion data?

PCA for FACE analysis

Data sets: many many faces



PCA

mean μ

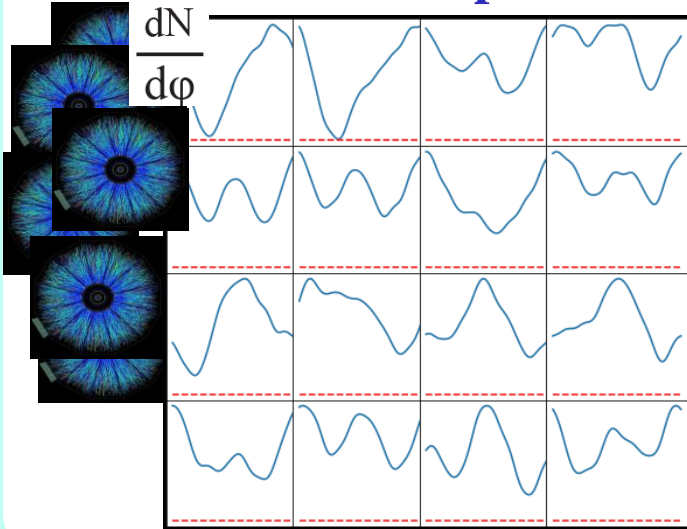


top eigenvectors: $\mu_1, \mu_2, \mu_3 \dots$



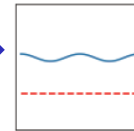
PCA for flow analysis –basic idea

Data sets: 1000 eve particle distr.

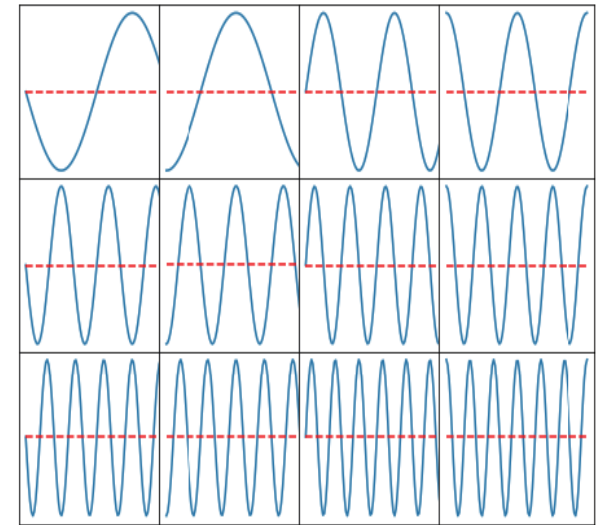


PCA

Mean μ

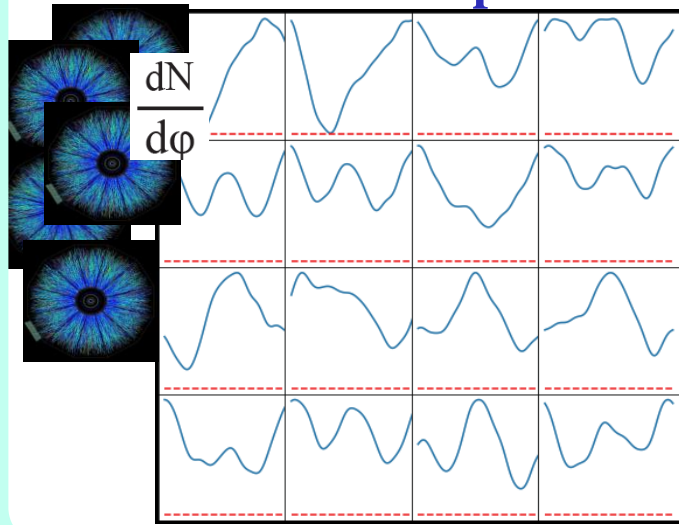


top eigenvectors: $\sigma_1, \sigma_2, \sigma_3 \dots$



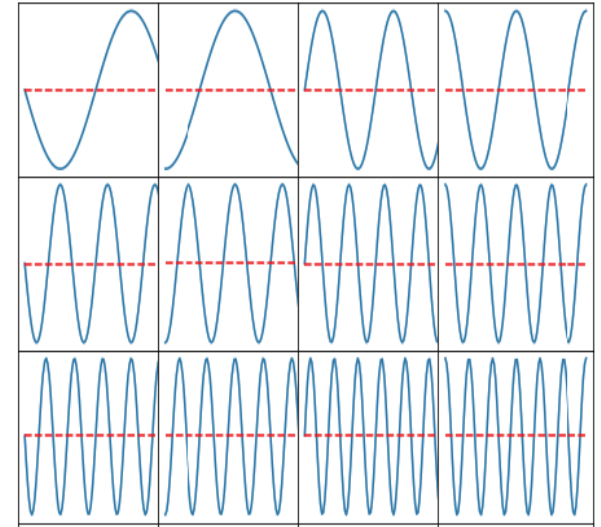
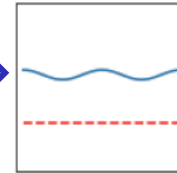
PCA for flow analysis –basic idea

Data sets: 1000 eve particle distr.

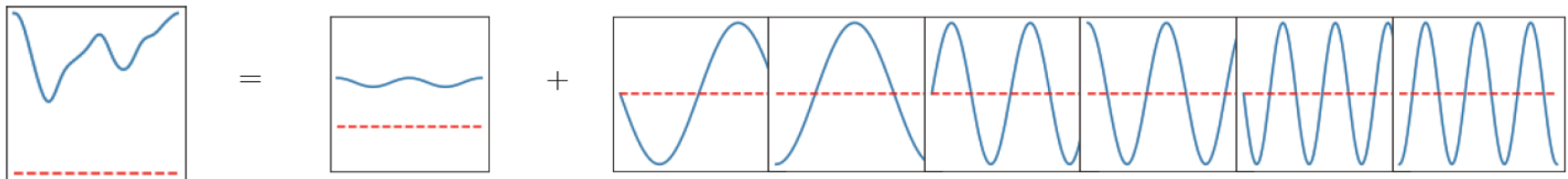


top eigenvectors: $\sigma_1, \sigma_2, \sigma_3 \dots$

Mean μ



With PCA, particle distributions in each events also decomposed into superpositions of eigenmodes



$$\frac{dN}{d\phi} = \mu + x_1 Z_1 + x_2 Z_2 + x_3 Z_3 + \dots$$

In the next few slides, I will SHOW

-PCA define its own flow harmonics (eigenmodes)

-PCA could analyze flow with event average/ event-by-event

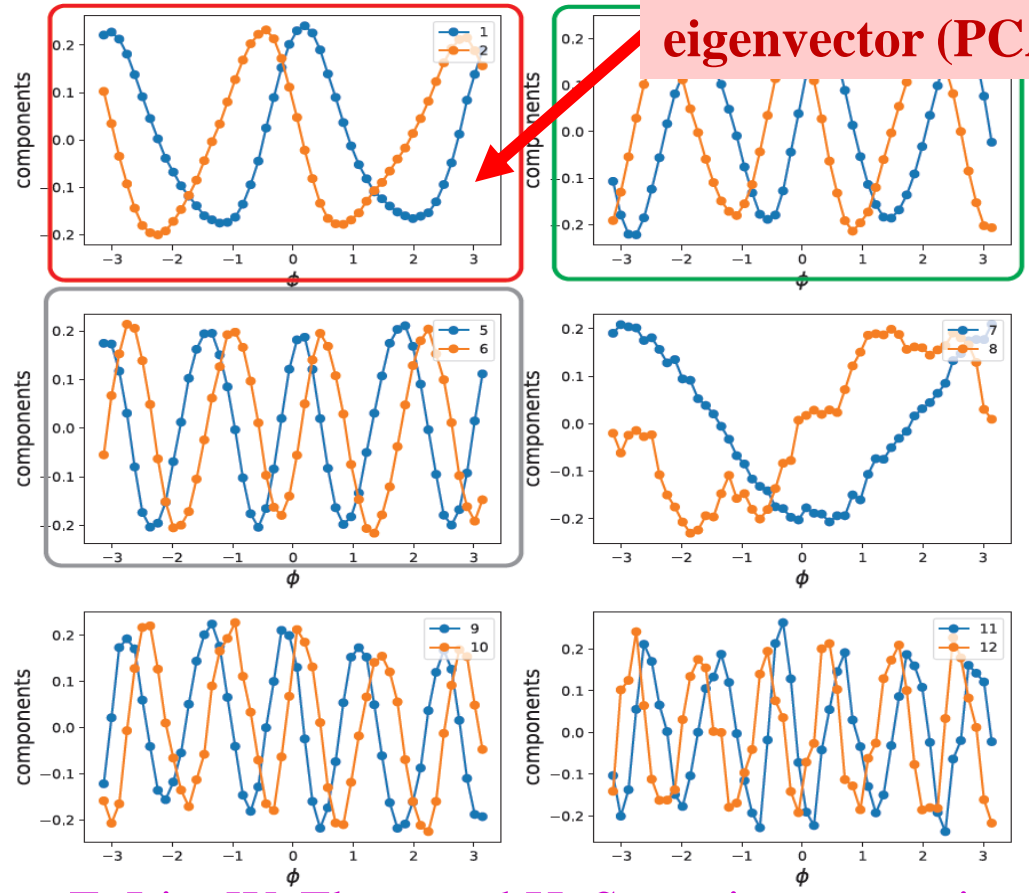
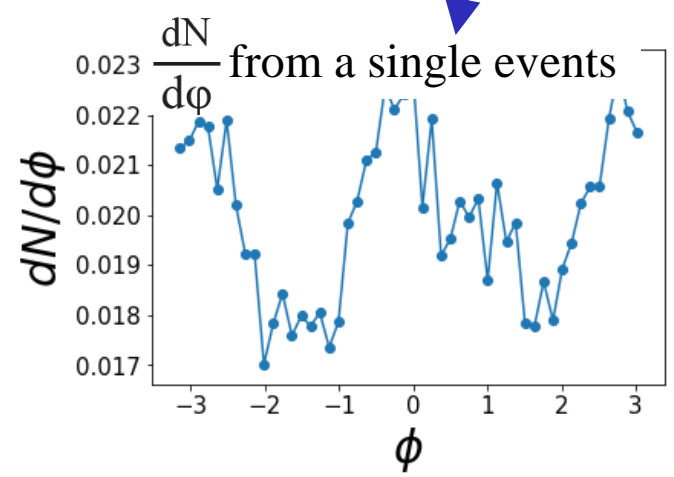
PCA for flow analysis –results(I)

events

$$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & \dots & a_m^{(1)} \\ a_1^{(2)} & a_2^{(2)} & \dots & a_m^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(n)} & a_2^{(n)} & \dots & a_m^{(n)} \end{pmatrix} =$$

Components

$$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_k^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n)} & x_2^{(n)} & \dots & x_k^{(n)} \end{pmatrix} * \begin{pmatrix} 1 \\ 2 \\ \vdots \\ k \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{pmatrix}$$

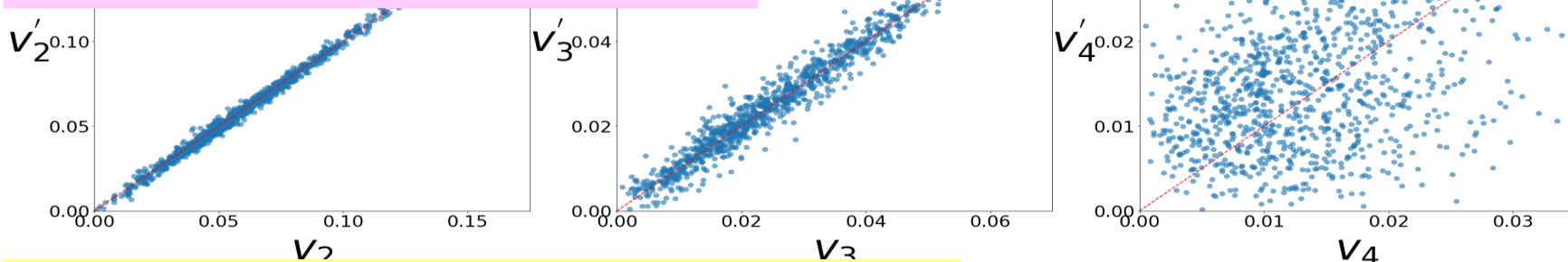


eigenvector (PCA)

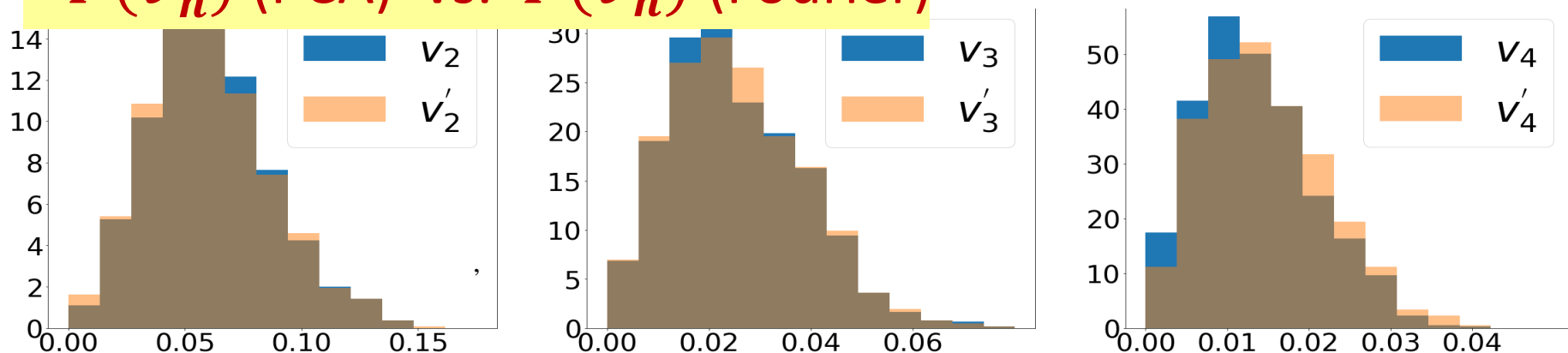
The eigenvector (PCA) is similar to the Fourier ones

$$\frac{dN}{dyd\phi} = \frac{dN}{dy} (1 + v_1 \cos\phi + v_2 \cos 2\phi + v_3 \cos 3\phi \dots)$$

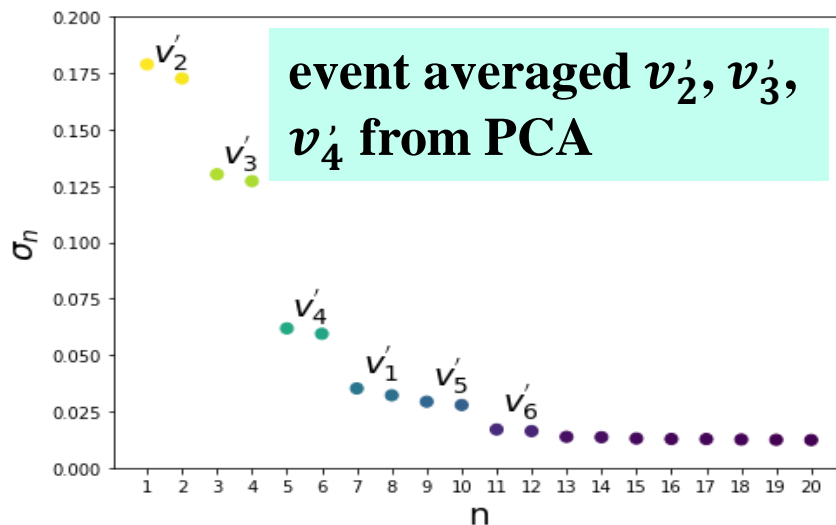
v'_n (PCA) vs. v_n (Fourier)



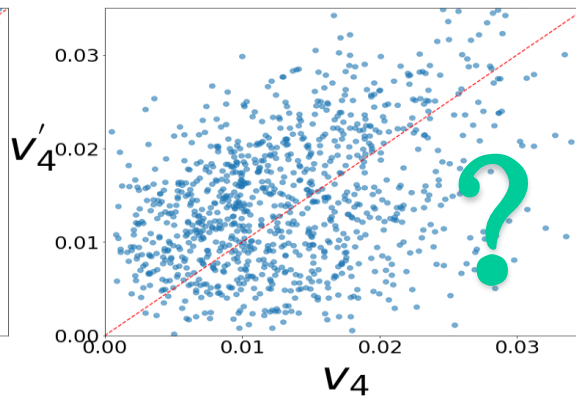
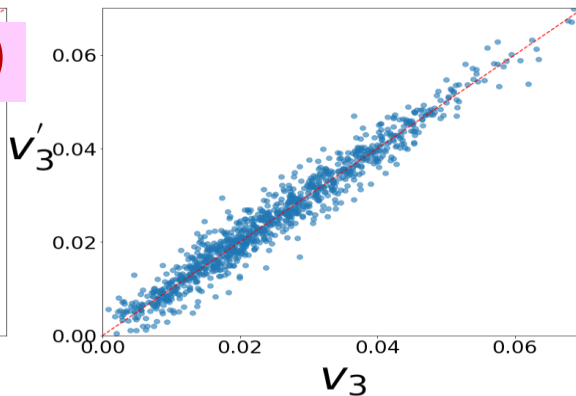
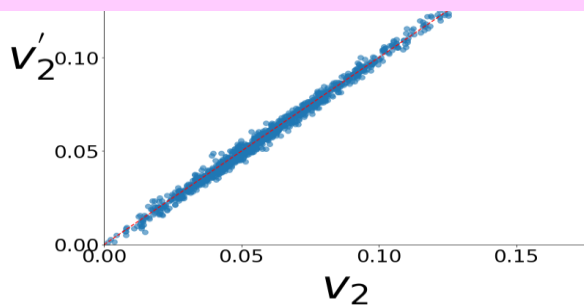
$P(v'_n)$ (PCA) vs. $P(v_n)$ (Fourier)



-PCA also gives the event averaged flow harmonics $v'_2, v'_3, v'_4 \dots$ and the event-by-event v'_n distributions, results are similar to the ones from traditional Fourier transform, except for v_4



v'_n (PCA) vs. v_n (Fourier)



Traditional Fourier Transform

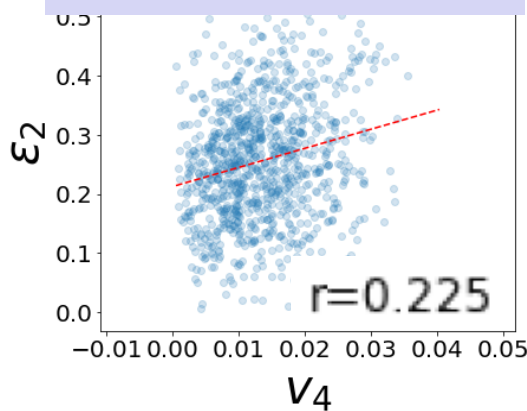
- Strong mode couplings between v_4 and v_2
- interoperated as highly non-linear hydro evolution that mix v_4 and ε_2^2

PCA:

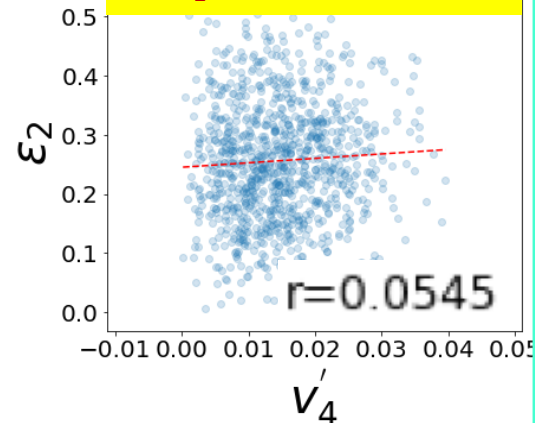
- Reduce the correlations between v'_4 and ε_2
- increase correlations between v'_4 and ε_4

Z. Liu, W. Zhao, and H. Song,
in preparation

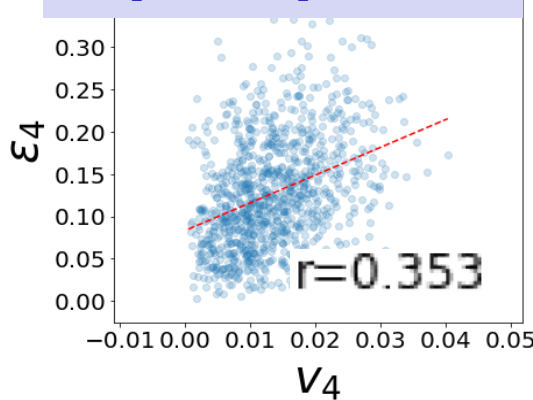
v_4 vs. ε_2 (Fourier)



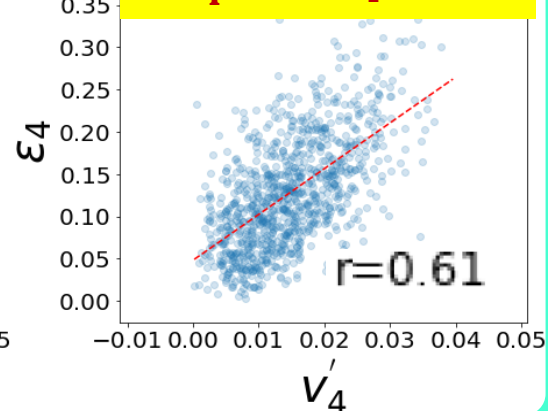
v'_4 vs. ε_2 (PCA)



v_4 vs. ε_4 (Fourier)

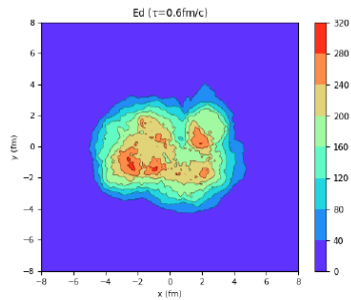


v'_4 vs. ε_4 (PCA)

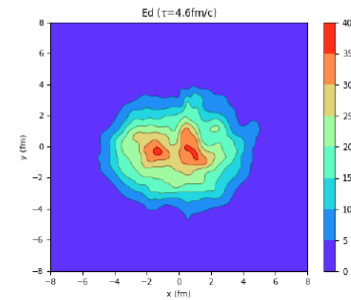


Summary & outlook

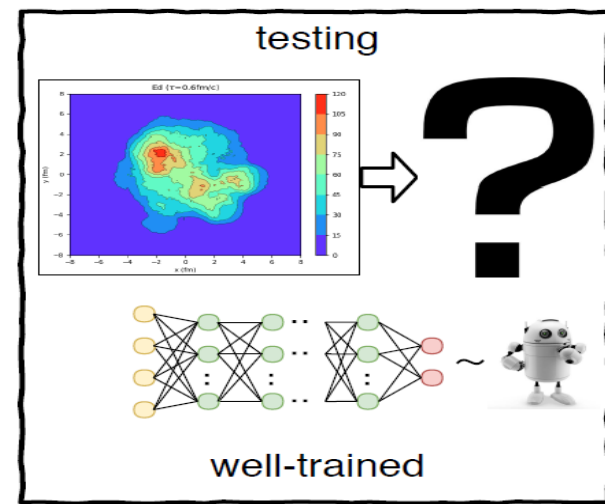
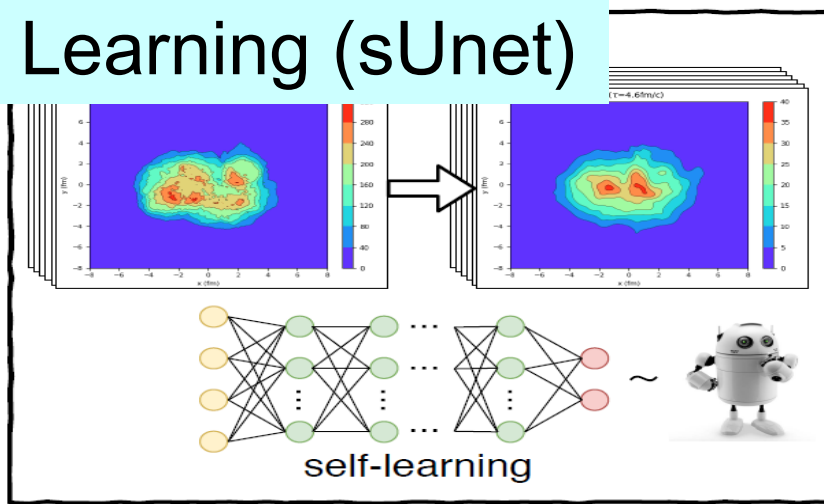
Traditional hydrodynamics



$$\partial_\mu T^{\mu\nu} = 0$$



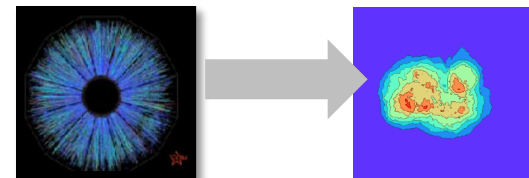
Deep Learning (sUnet)



Outlook

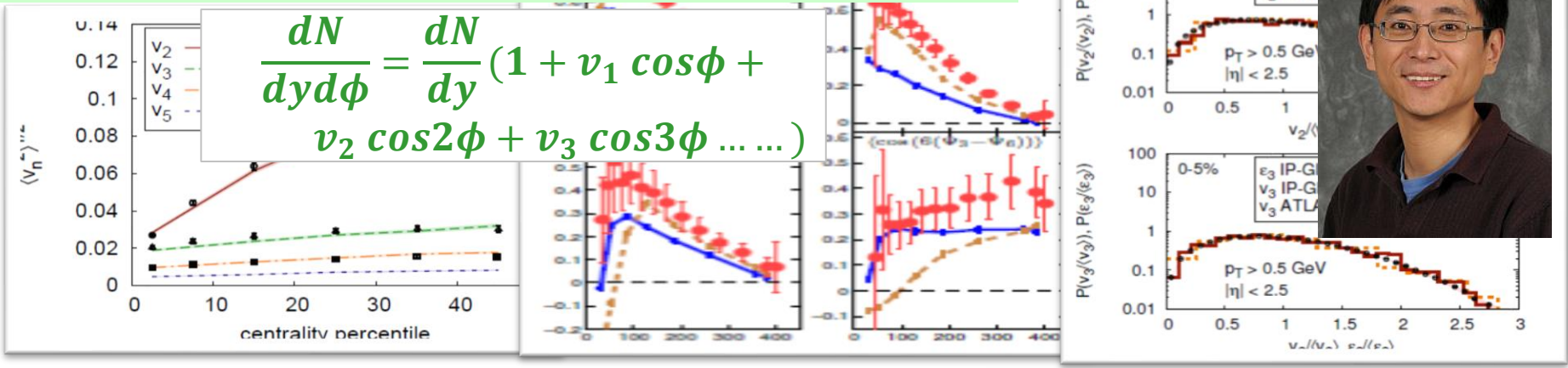
Final particle profiles

----- > Initial energy density profiles



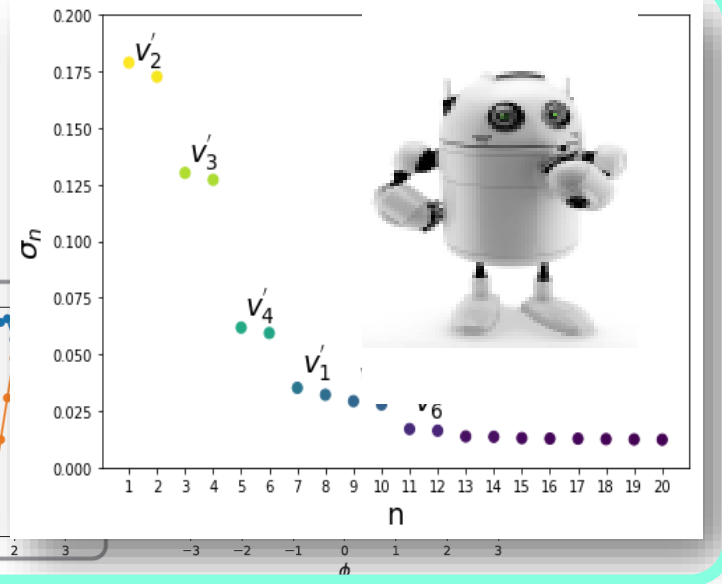
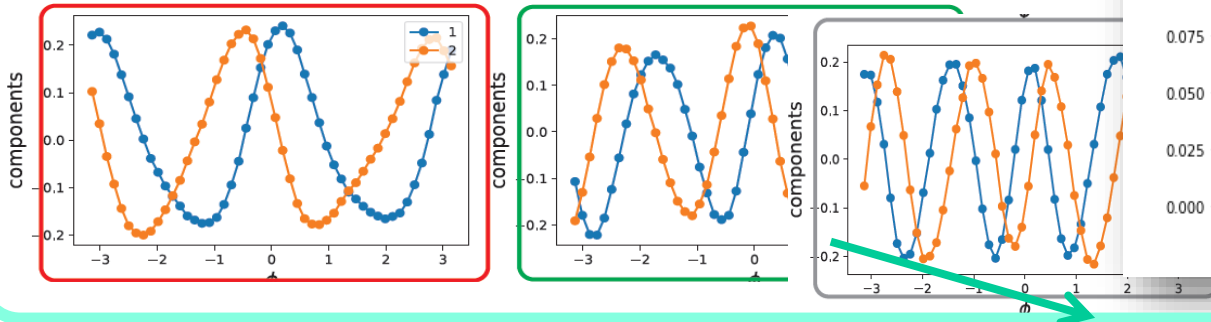
Can deep learning discover knowledge (conservation laws) from the massive data generated from hydrodynamics?

Flow: traditional Fourier Transform



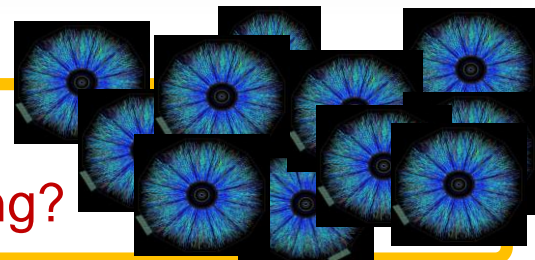
Unsupervised Learning (PCA)

It independently discovered the flow harmonics without explicit instructions from human being!



Outlook

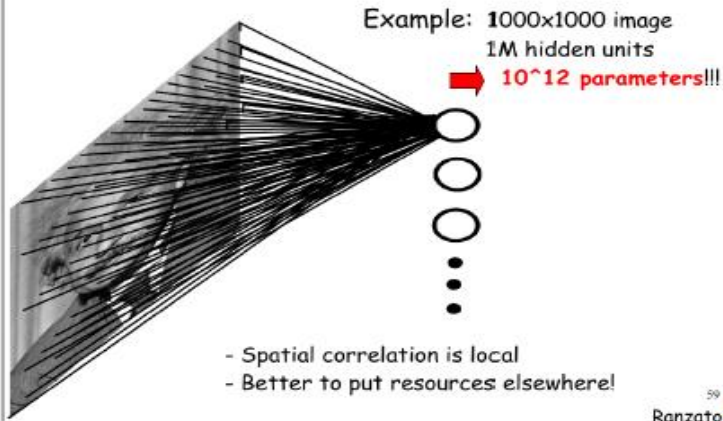
Can PCA detect modes or structures from the massive data that is not realized or easily defined by human being?



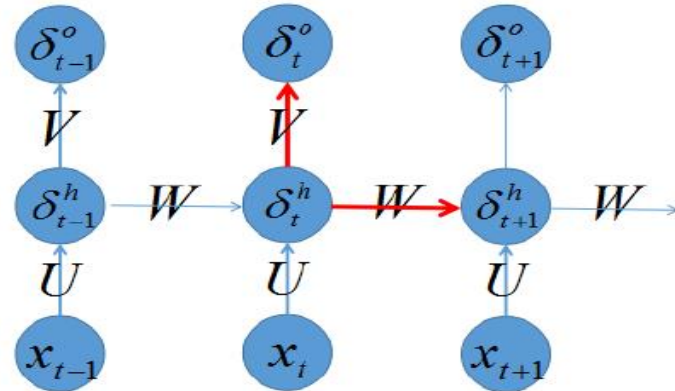
Thank You

Common Network Structures

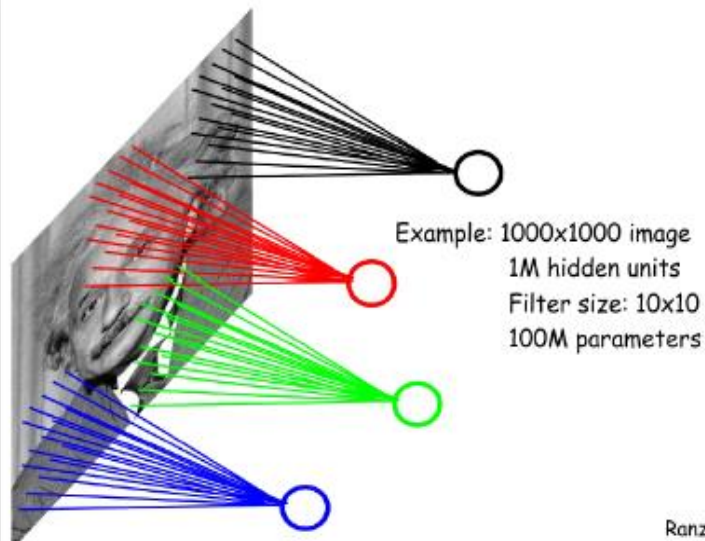
FULLY CONNECTED NEURAL NET



Recurrent Neural Network



LOCALLY CONNECTED NEURAL NET



Fully Connected Network

-recognize handwrite digits

... ..

Convolutional Neural network

-image recognition

-image classification

... ..

Recurrent Neural Network

-speech recognition

... ..