Applications of Deep Learning in Relativistic Hydrodynamics

Huichao Song

宋慧超

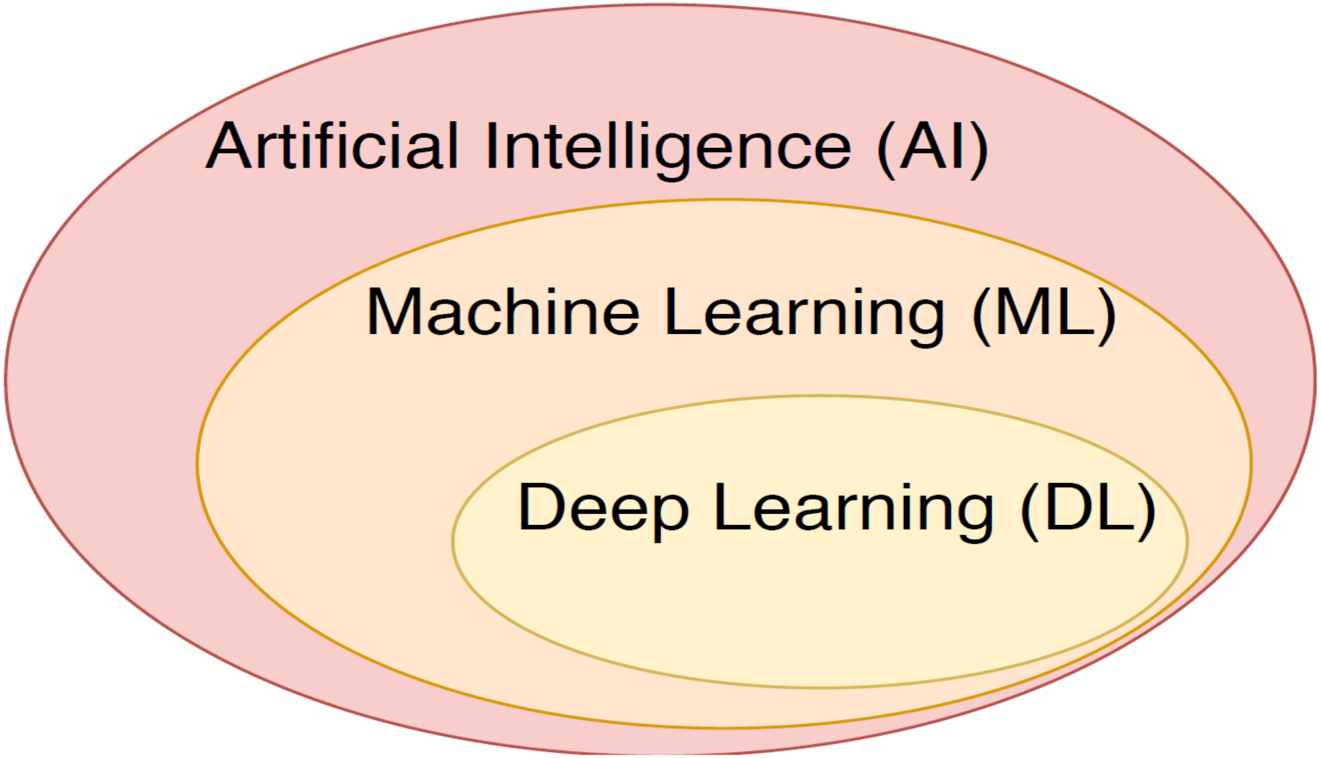
Peking University

Second international workshop on Collectivity in Small Collision Systems (CSCS 2018)

Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu and Huichao Song,arXiv:1801.03334

June. 15, 2018

What is 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 Deep Learning

AlphaGo (by Google DeepMind) beat human champion, March, 2016

Google DeepMind, London Nature 529, 484-489 (2016)

Broad Applications of Deep Learning

Language processing

-Machine translation

J. Lee, K. Cho, and T. Hofmann, TACL,

-Chinese poetry generation

Z. Wang et al., COLING, page 1051-1060,

ACL, (2016), arXiv: 1610.09889

-Speech recognition

W. Xiong et al., IEEE/ACM Transaction

Audio Speech & Language Processing, 20

秋夕湖上
By a Lake at Autumn Sunset
一夜秋凉雨湿衣，
A cold autumn rain wetted my clothes last night,
西窗独坐对夕晖。
And I sit alone by the window and enjoy the sunset.
W
A

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 .

(e) Multilingual

Multi src	Bei der Metropolitního výboru pro dopravu für das Gebiet der San Francisco Bay erklärten Beamte , der Kongress könne das Problem банкротство доверительного Фонда строительства шоссейных дорог einfach durch Erhöhung der Kraftstoffsteuer lösen .
EN ref	At the Metropolitan Transportation Commission in the San Francisco Bay Area , officials say Congress could very simply deal with the bankrupt Highway Trust Fund by raising gas taxes .
bpe2char	During the Metropolitan Committee on Transport for San Francisco Bay , officials declared that Congress could solve the problem of bankruptcy by increasing the fuel tax bankrupt .
char2char	At the Metropolitan Committee on Transport for the territory of San Francisco Bay , officials explained that the Congress could simply solve the problem of the bankruptcy of the Road Construction Fund by increasing the fuel tax .

Broad Applications of Deep Learning

Computer vision

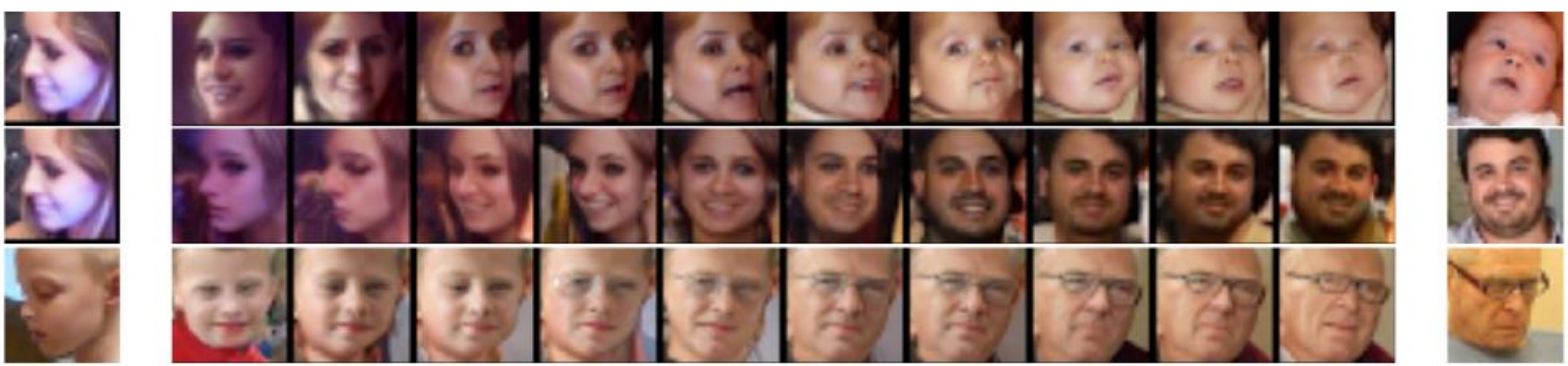
-Image style transition:

Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge, In: Computer Vision and Pattern Recognition. 2016, pp.2414-2423

-Image generation:

A. van den Oord et al., NIPS, (2016), arXiv: 1606.05328

… … … …



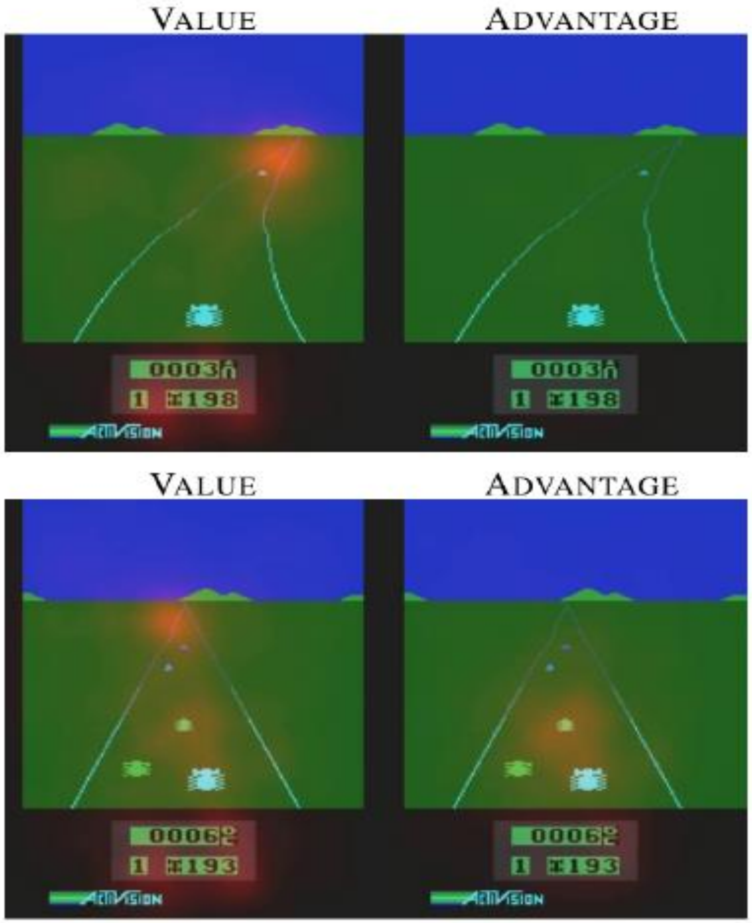
Broad Applications of Deep Learning

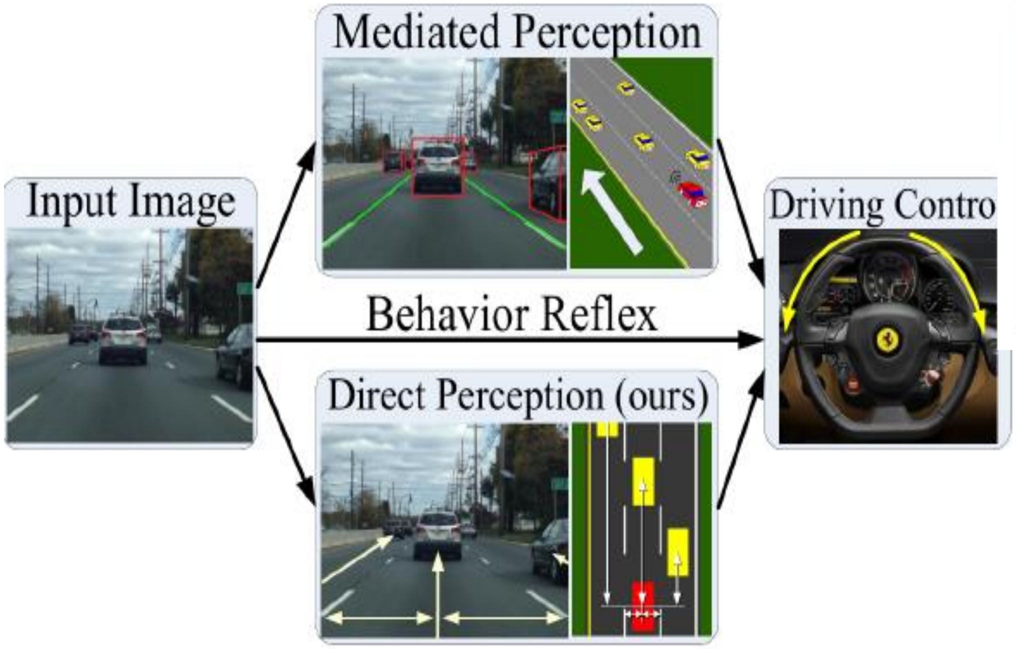
-Playing Games:

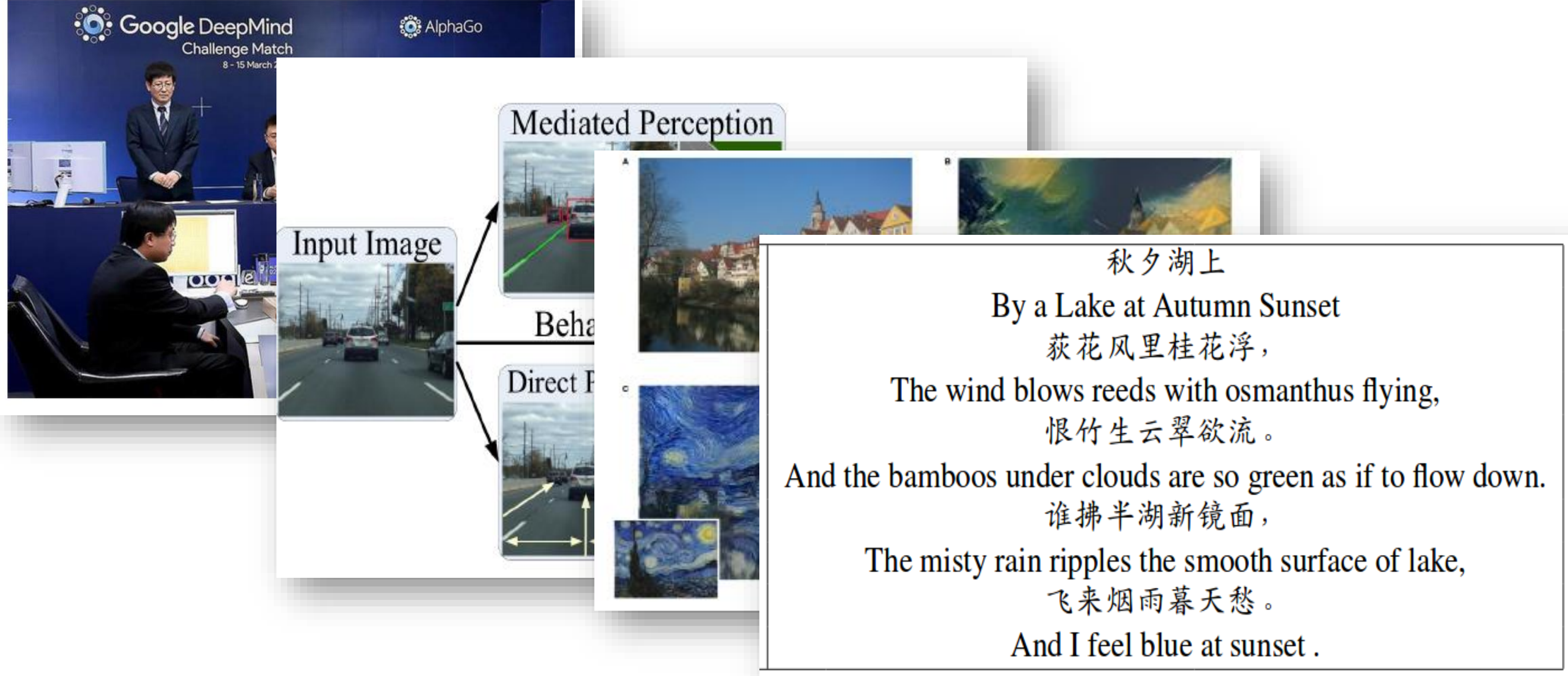
Z. Wang, T. et al., ICML,
page 1995-2003. JMLR.org, (2016),

-Autonomous Driving

C. Chen et al., ICCV, page 2722-2730.
IEEE Computer Society, (2015),







Categories of deep learning

-Supervised learning

-Unsupervised learning

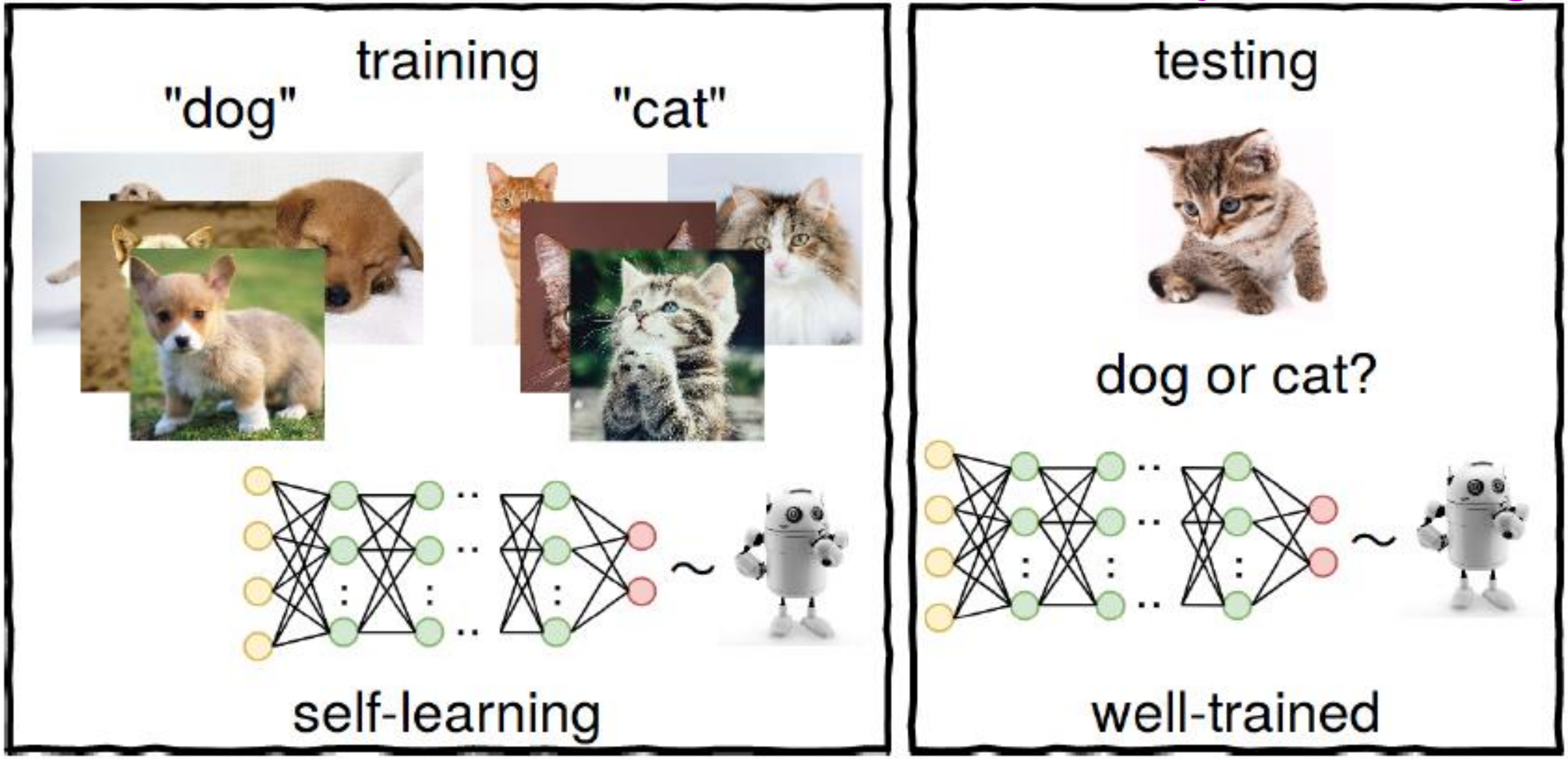
-Reinforcement learning

… … … …

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

An example of Supervised Learning

-Identify cats and dogs

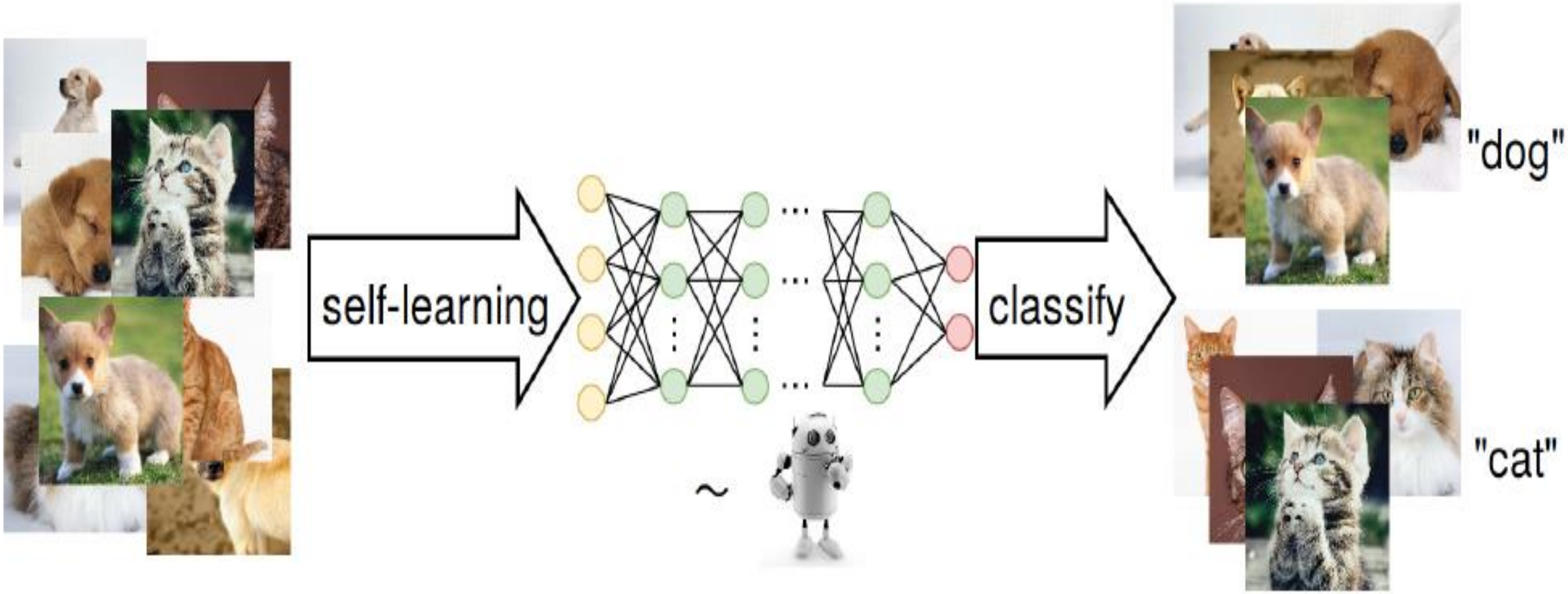


Supervised learning:

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

An example of Unsupervised Learning

-Classify cats and dogs

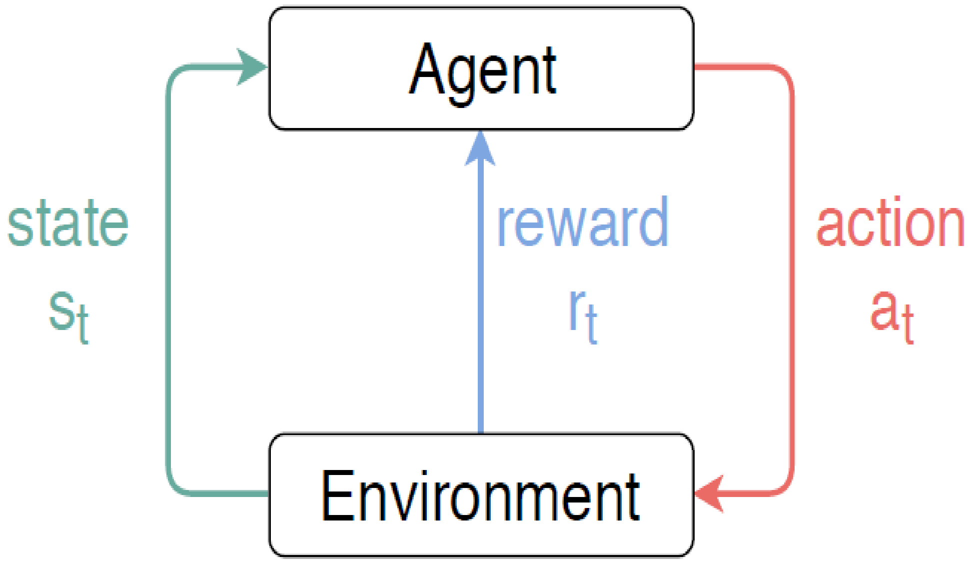


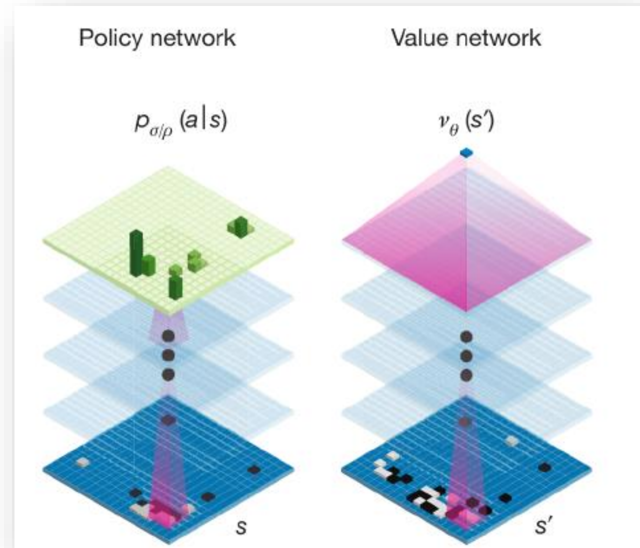
Unsupervised learning

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

An example of Reinforcement Learning

-play games





Reinforcement learning

concern with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward

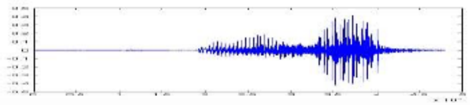
Deep Neural Network

-- more details

Deep learning tutorial for non-experts

Somewhat similar to look for a super complex function

- Speech Recognition

$f($  $)=$ "How are you"

- Image Recognition

$f($  $)=$ "Cat"

- Dialogue System

$f($ "Hi" $)=$ "Hello"

(what the user said) (system response)

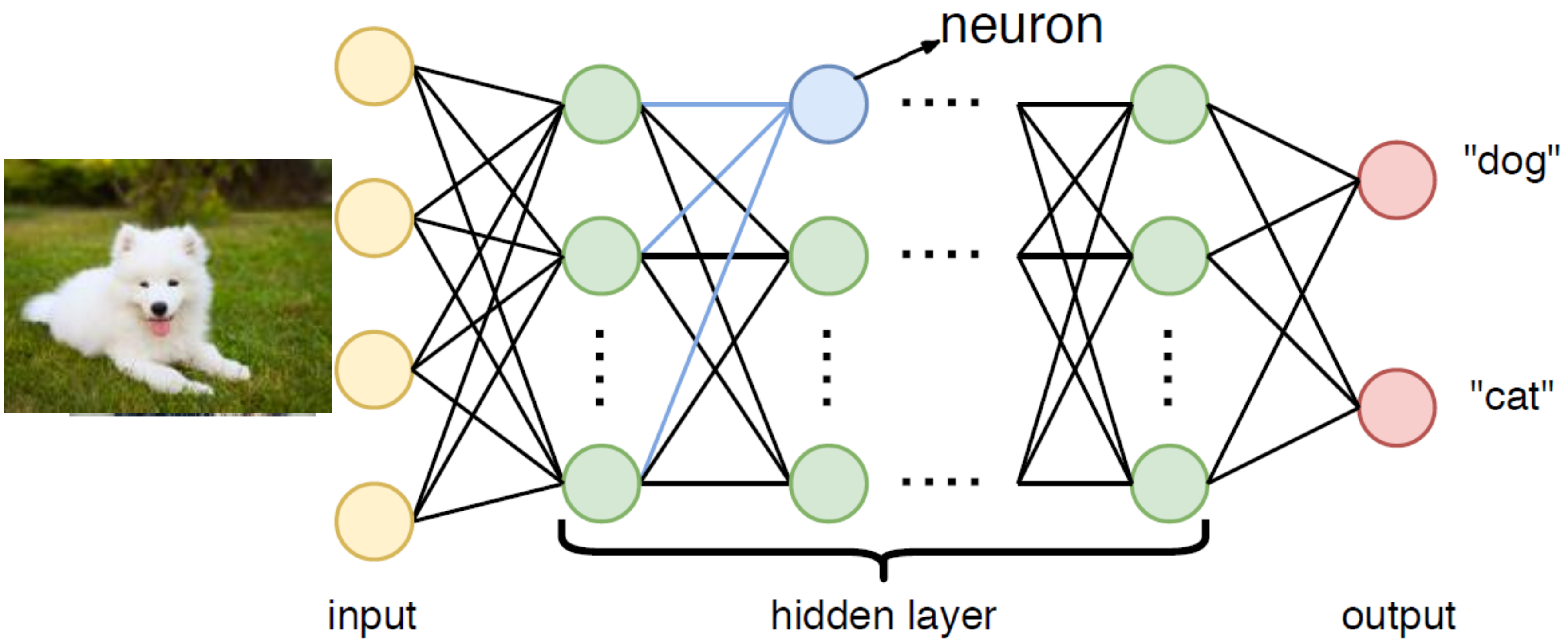
-Define a set of functions

-Evaluate each function

-Pick up the best

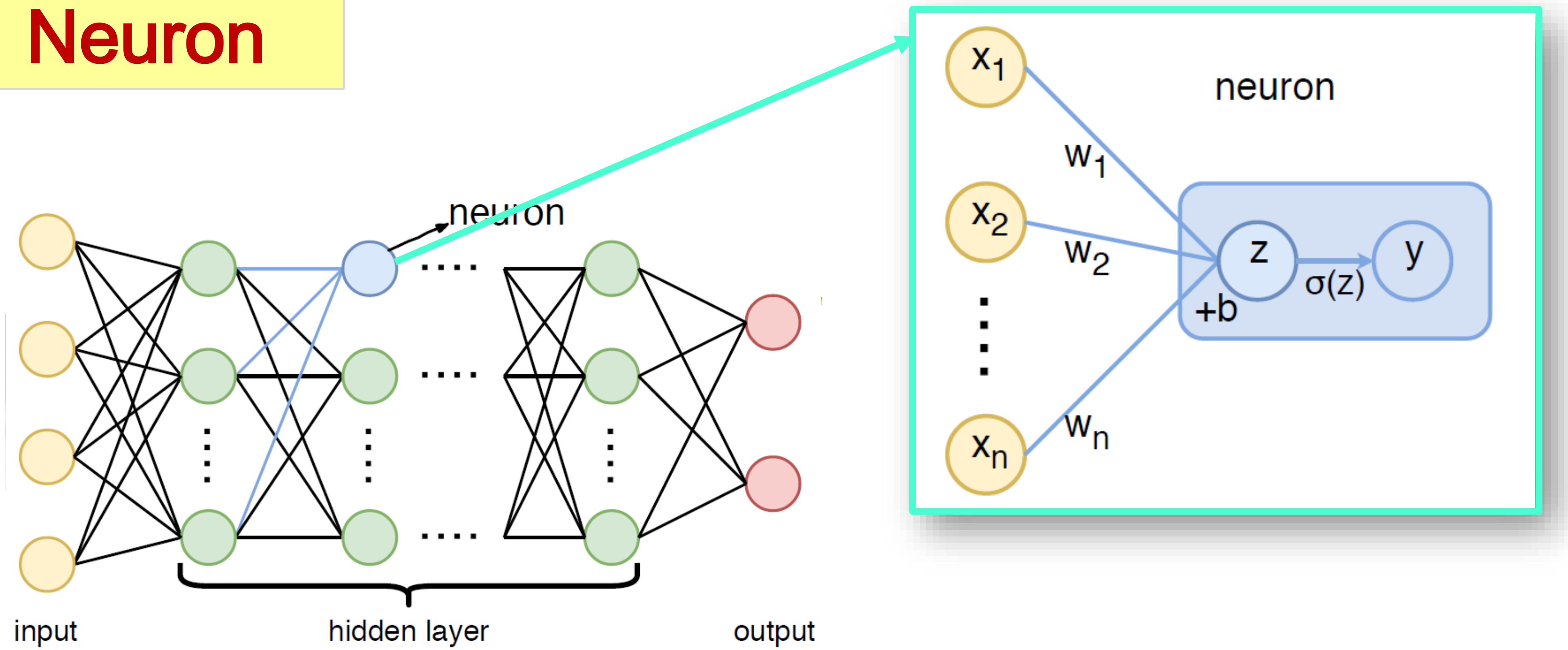
H. Lee, Deep Learning Tutorial,
https://www.slideshare.net/tw_dsconf/ss-62245351

Neural Network / Deep Neural Network



-In fact, there is not a “function” but to build up a (deep) neural network with huge tunable parameters to connect the inputs and outputs.

Neuron



Linear operation

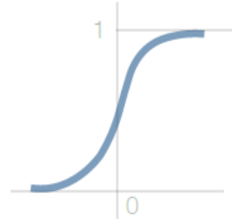
$$z_j = \sum_{i=1}^{N} x_i w_{ij} + b_j$$

scaling, rotating, boosting, changing dimensions

Non-linear activation function $h_j = \sigma(z_j)$

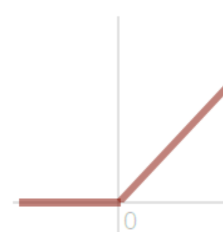
(a) Sigmoid

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



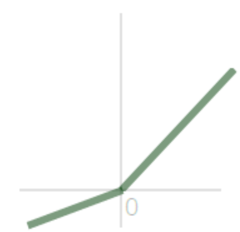
(b) ReLU

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

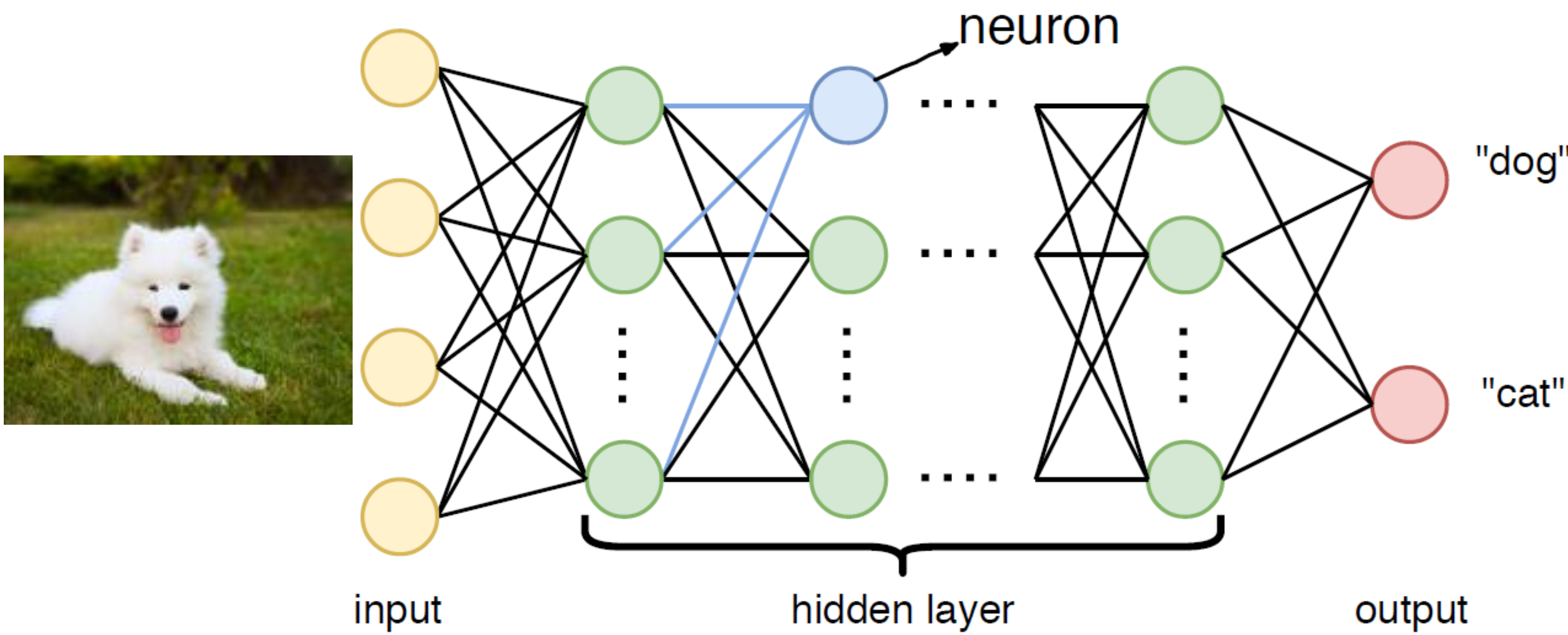


(c) PReLU

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



Deep Neural network-loss fuction



Loss function:

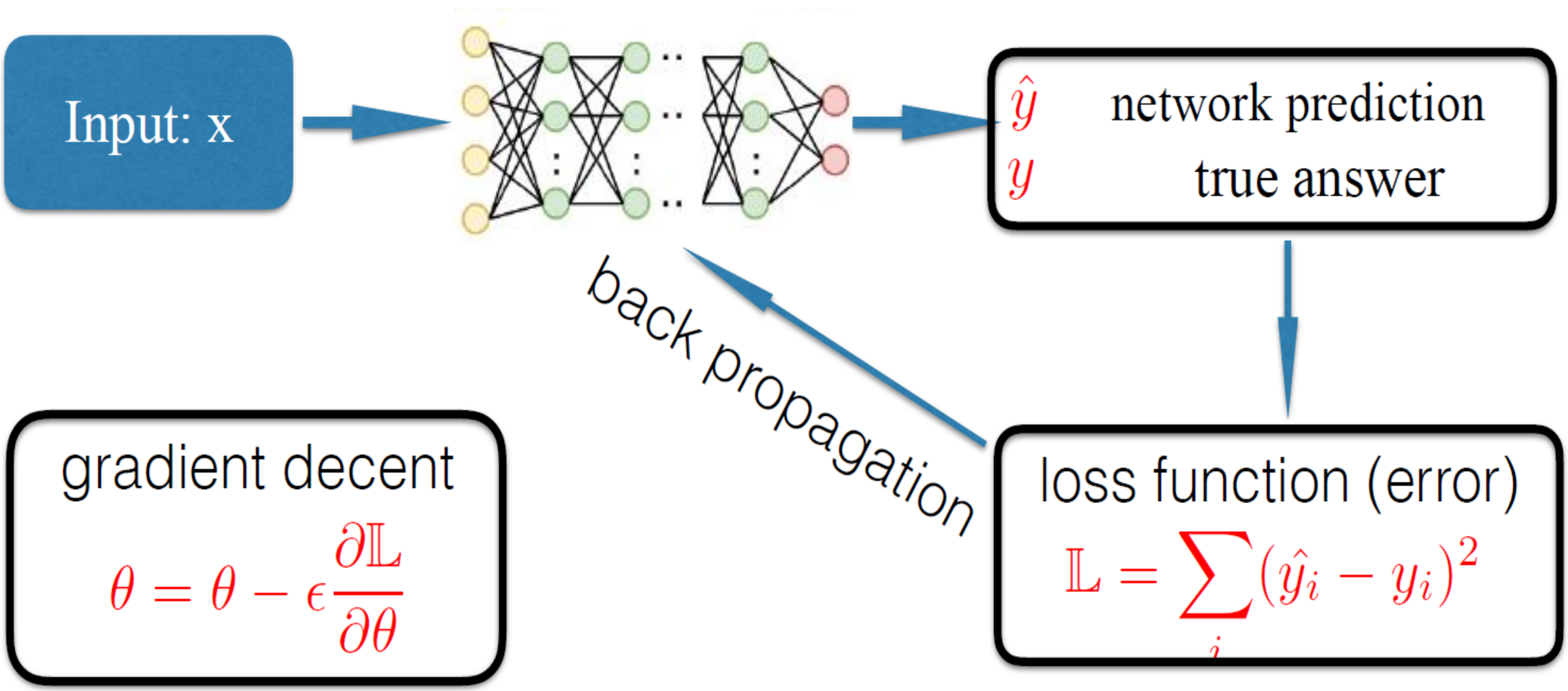
-To evaluate the difference between the network's outputs and learning targets.

- $\ell(\theta) = \frac{1}{2n} \sum_x [y(x) - \hat{y}(x)]^2$
- $\ell(\theta) = -\frac{1}{n} \sum_x [y(x) \ln \hat{y}(x) - (1 - y(x)) \ln(1 - \hat{y}(x))]$

$\hat{y}$ network prediction

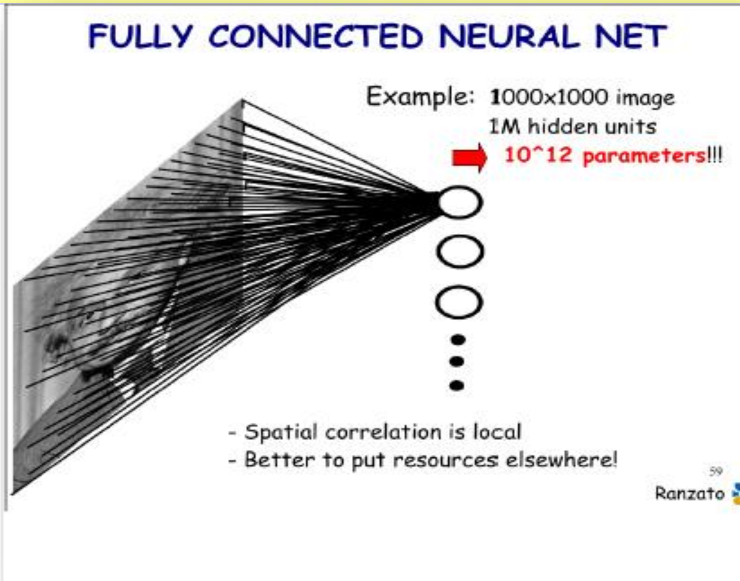
y true answer

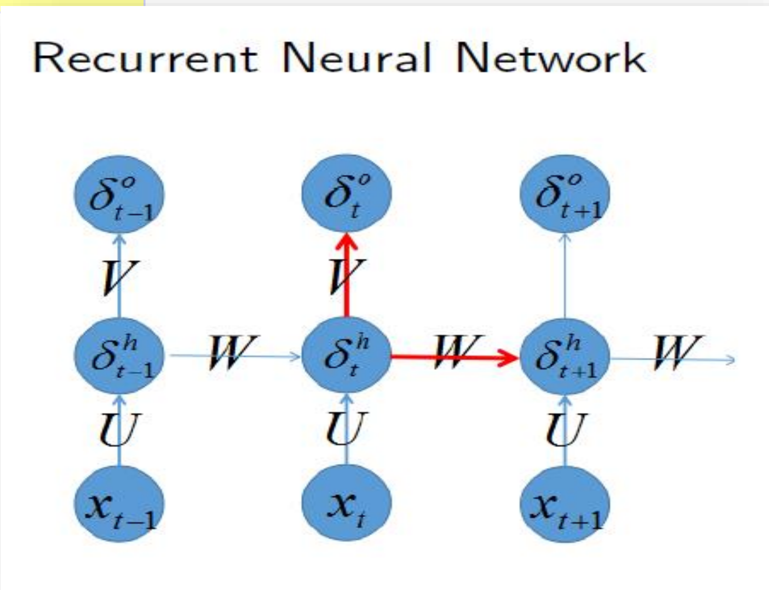
Deep Neural network-back propagation & gradient decent

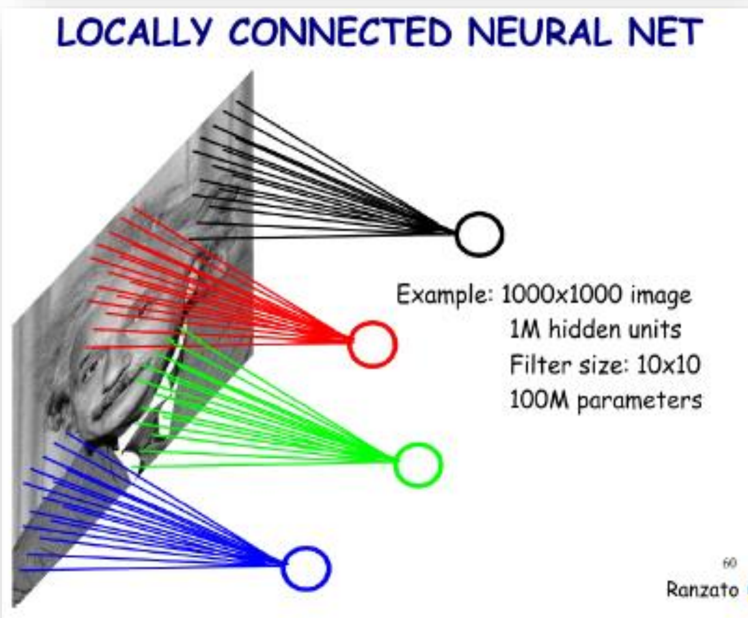


-Deep neural network can reduce fitting error by updating model parameters through back propagation and gradient decent.

Common Network Structures







Fully Connected Network

-recognize handwrite digits

… … … …

Convolutional Neural network

-image recognition

-image classification

… … … …

Recurrent Neural Network

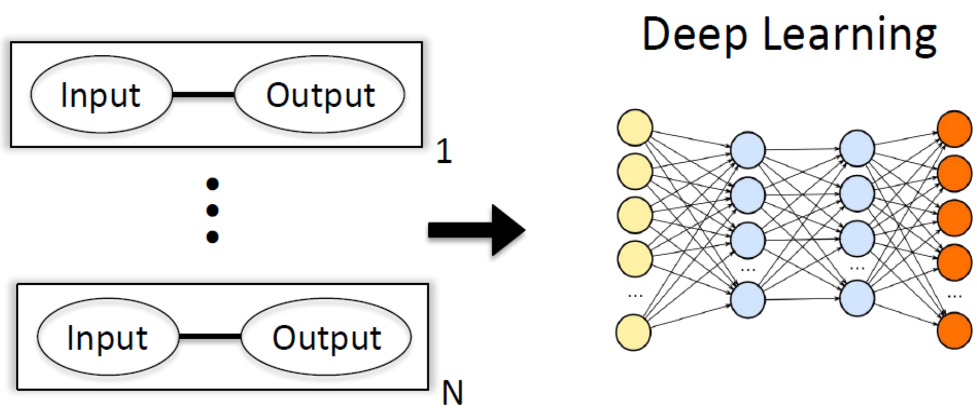
-speech recognition

… … … …

Applications of Deep 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."

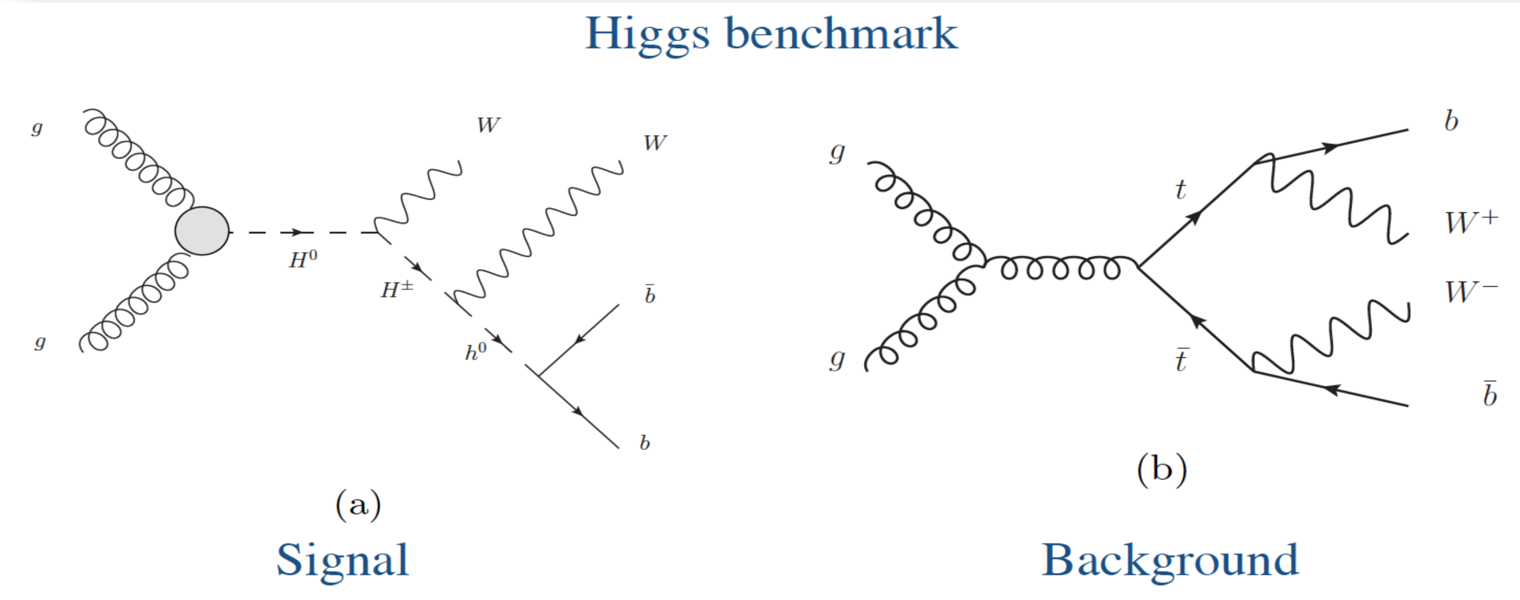


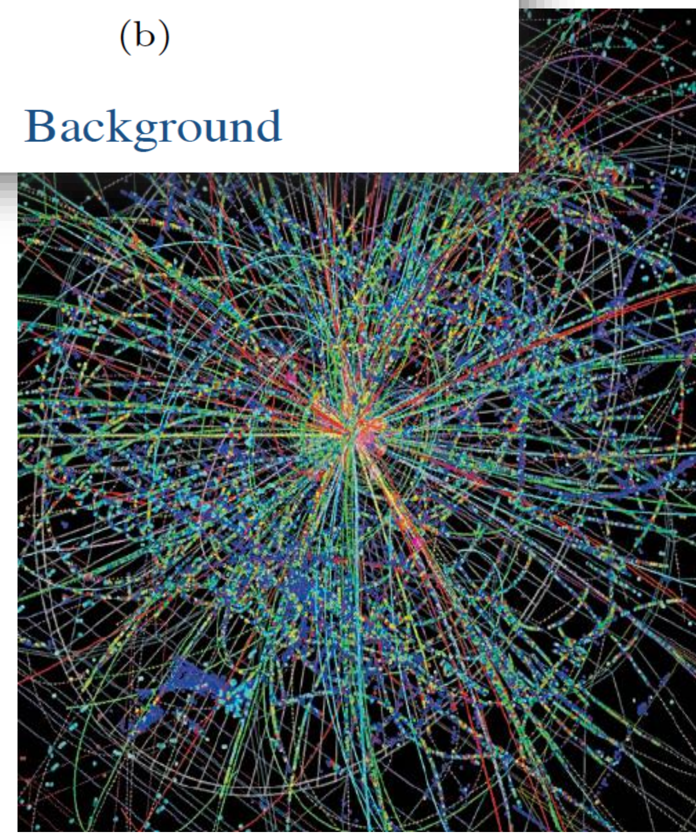
Can "**Black-box**" models learn patterns and models solely from data without relying on scientific knowledge?

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





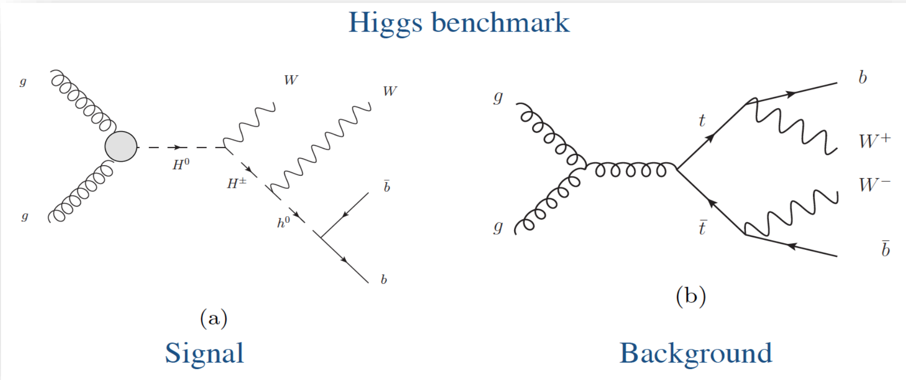
Motivation:

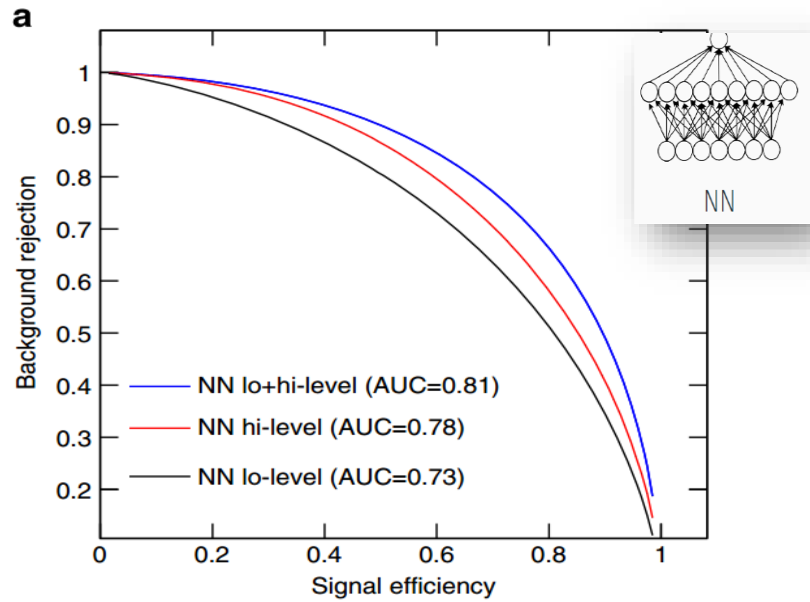
-Finding the rare particles in high-energy particle colliders requires a successful distinguish of the signal from the huge & messy background

-Traditional 'shallow' machine learning models that have a limited capacity for such task

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

Searching for Exotic Particles in High-Energy Physics





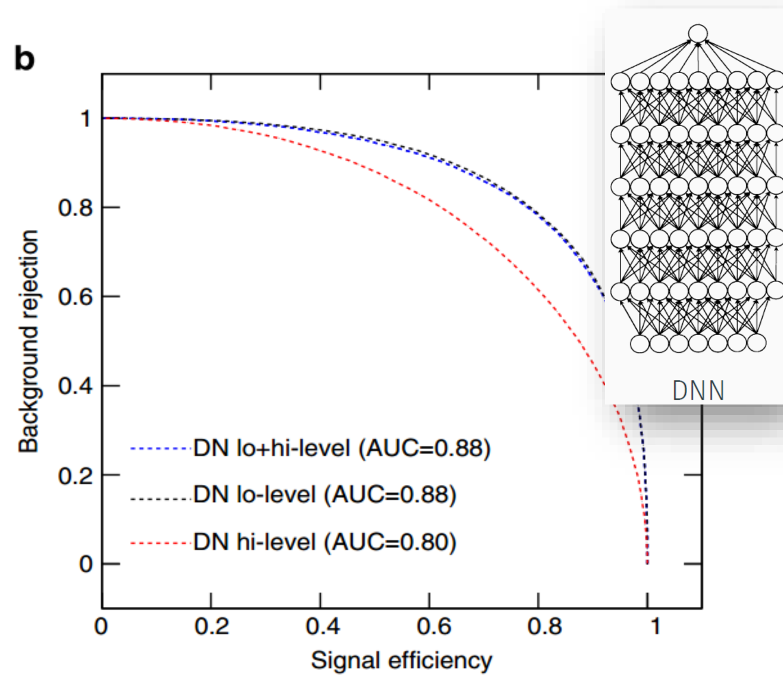
A) Generating training/testing data

MadGraph (collisions) +PYTHIA (showering & hadronization) +DELPHES (detector response)

B) Training DNN/NN: supervised learning with 11 million data (low level / high level)

C) Testing DNN/NN

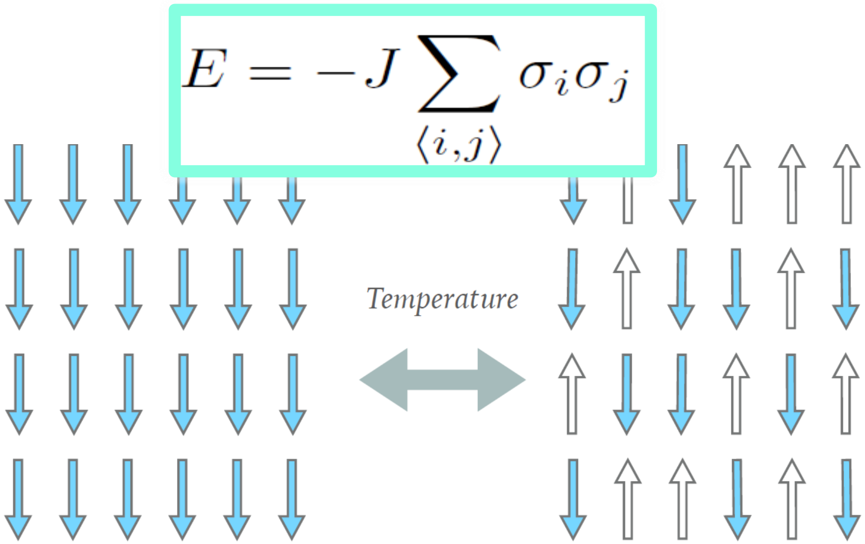
-DNN improves AUC by 8% compared to NN

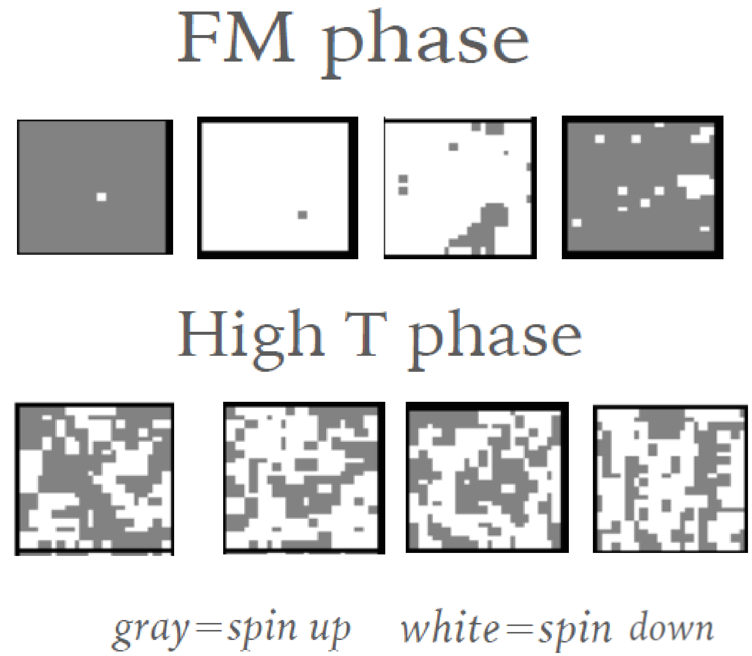


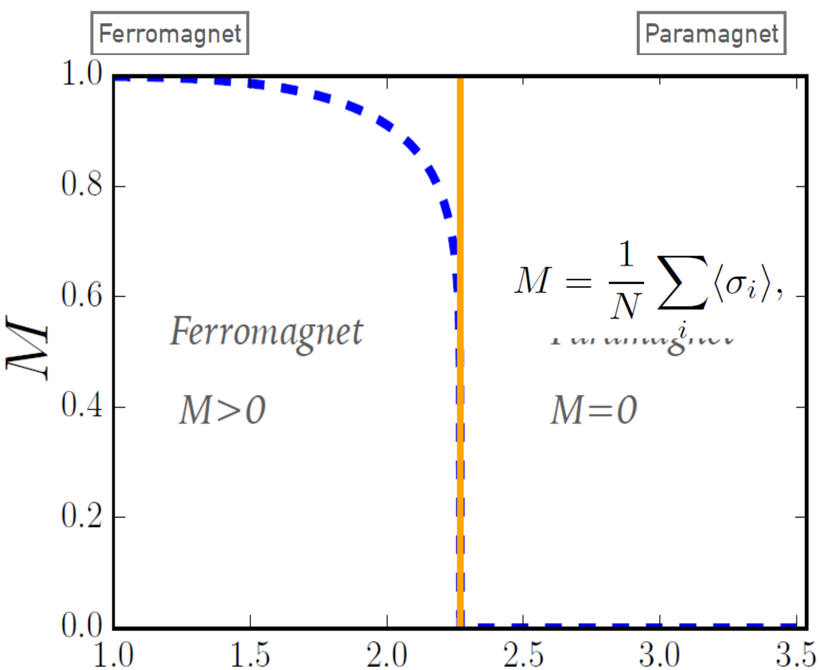
Deep learning can improve the power for the collider search of exotic particles

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

Classifying the Phase of Ising Model







Motivation:

-Traditionally, the study of phases transition of condense matter systems is to calculate the associated order parameter, measure specific heat, ect

-Can deep learning identify phases and phase transitions?

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

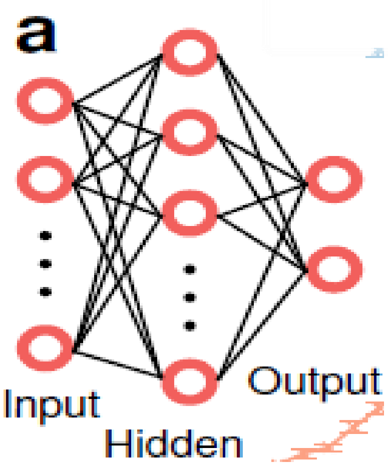
Classifying the Phase of Ising Model

A） Generating training/testing data:

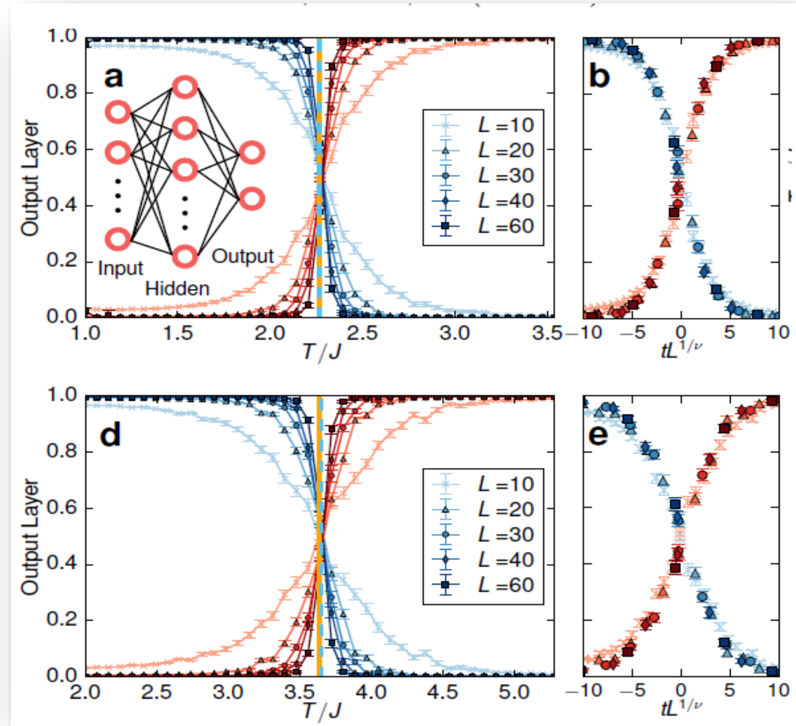
traditional MC method from the Boltzmann distribution $p(\sigma_1, \sigma_2, \ldots, \sigma_N) = \dfrac{e^{-\beta E(\sigma_1, \sigma_2, \ldots, \sigma_N)}}{Z(\beta)}$

B) training fully connected network:

with these raw configurations of **square-lattice Ising-model** $E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$



C) testing the trained net work



-The trained fully connected network also past the test of triangular-lattice Ising model, showing its abilities of generalize to task beyond their original design

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

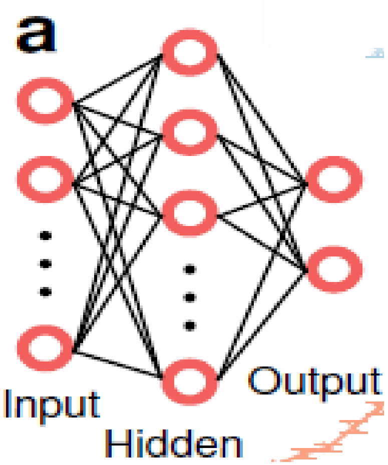
Classifying the Phase of Ising Model

A） Training/testing data:

traditional MC method from the Boltzmann distribution $p(\sigma_1, \sigma_2, ..., \sigma_N) = \dfrac{e^{-\beta E(\sigma_1, \sigma_2, ..., \sigma_N)}}{Z(\beta)}$

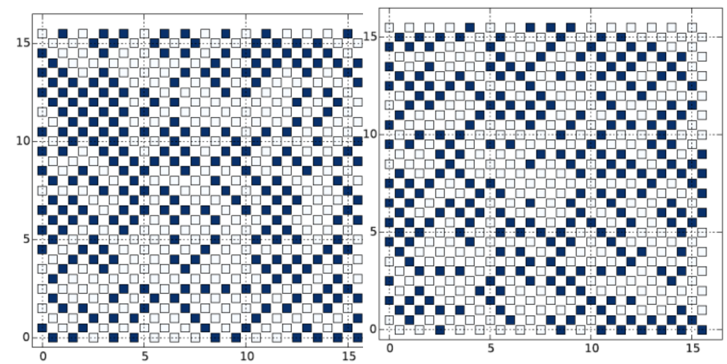
B) training fully connected network:

with these raw configurations of **square-lattice Ising-model** $E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$



Further test- Ising gauge theory

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



Fully connected network fails (50% accuracy)
-equal to simply guessing

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

Classifying the Phase of Ising Model

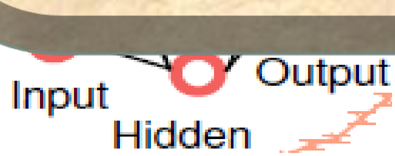
A) Training/testing data:

traditional MC method from the Boltzmann distrib

No free lunch theorem[1]

No machine learning algorithm is consistently better than another. In other words, there is no silver bullet, deep learning and neural networks not exempted. In fact, the most universal feedforward neural network does worse than tree based methods or SVM on many problems. Therefore, when adapting a model to new problems, one should be aware of model assumptions and ensures that they still holds.

[1] Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural computation*, 8(7), 1341-1390.



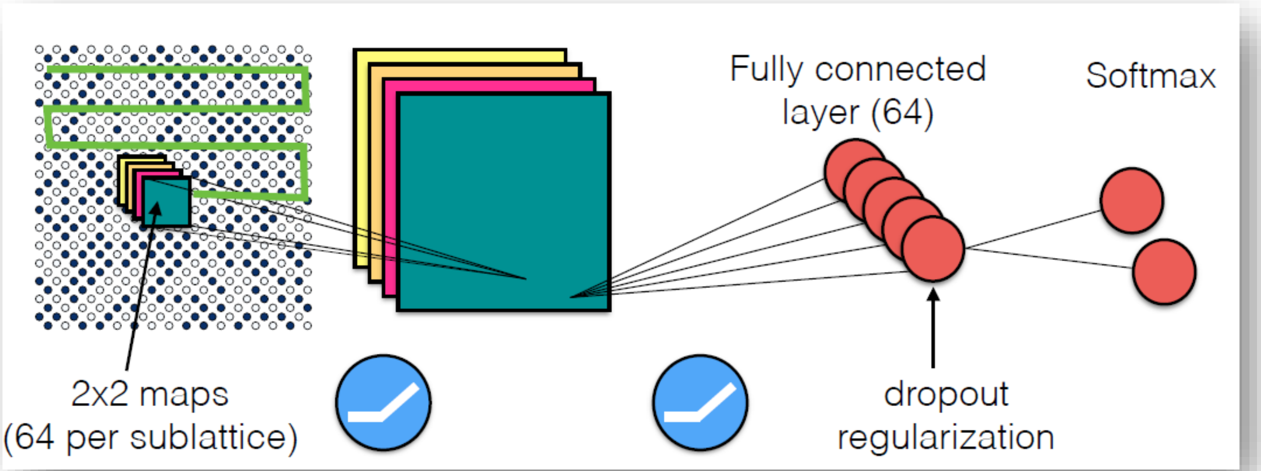
Nature Physics 13, 431–434 (2017)

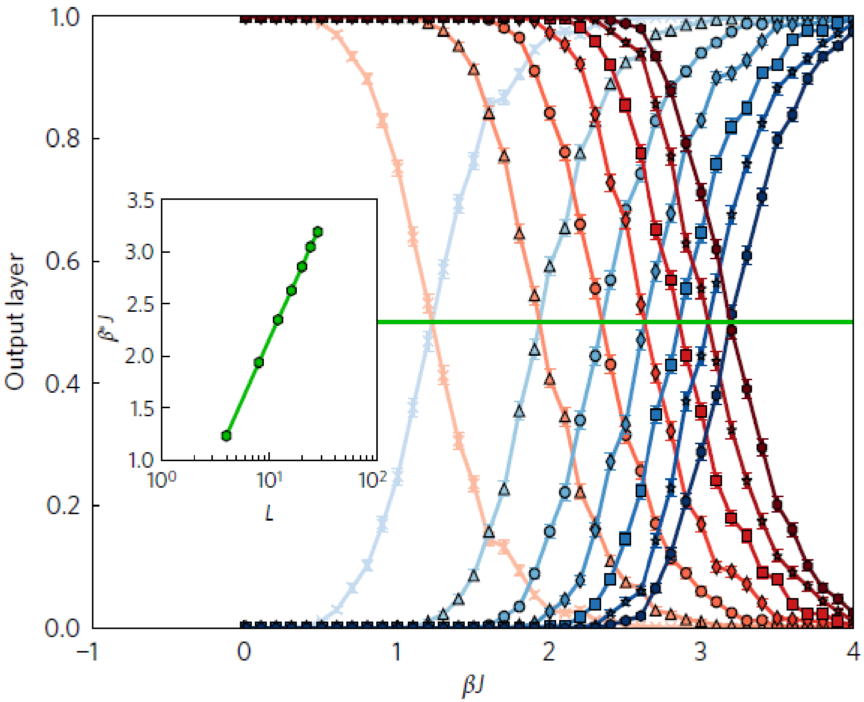
Classifying the Phase of Ising Model

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)

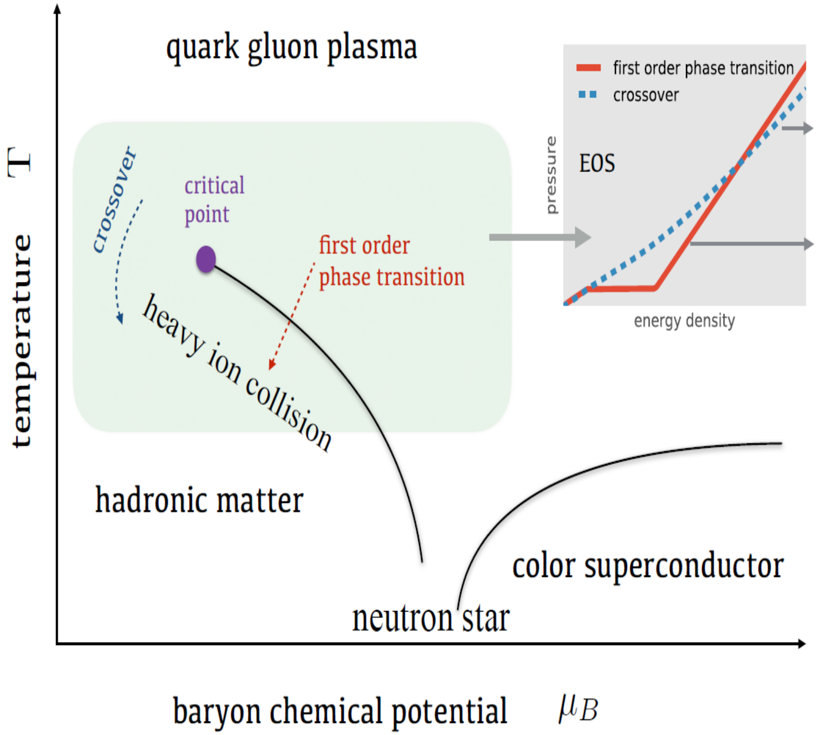


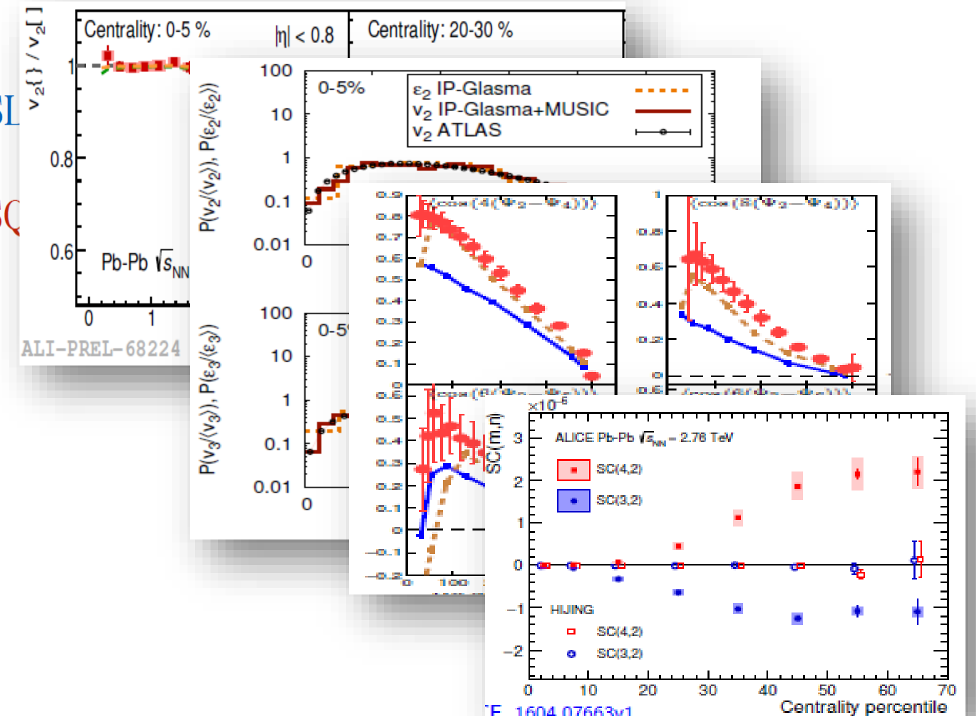


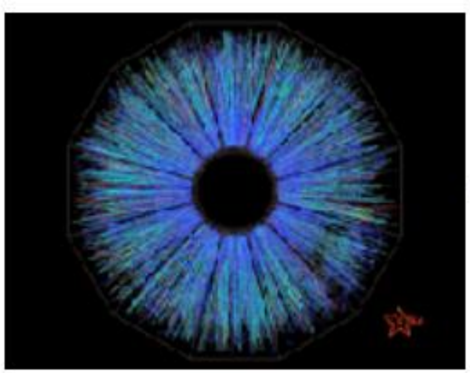
The trained CNN discriminates high-temperature from ground States with very high accuracy in spite of the lack of an order Parameter or qualitative differences in the spin-spin correlations

Neutral network can be used to encode phases of matter and discriminate phase transitions in correlated many-body systems.

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

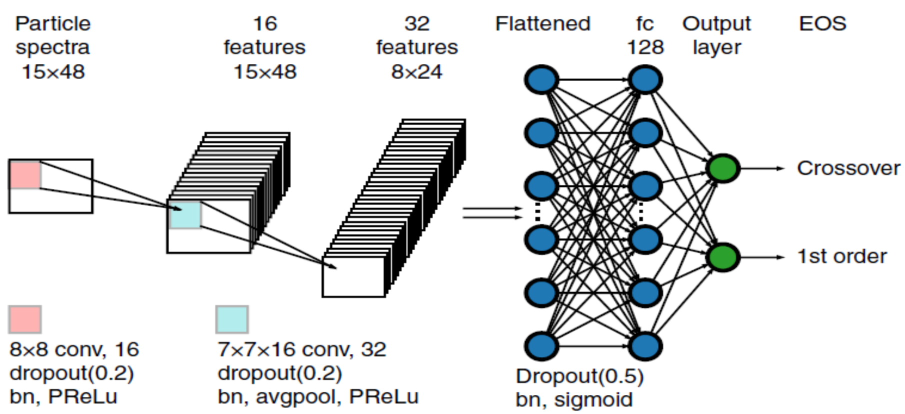
-particle spectra - image (15*48 pixels)

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

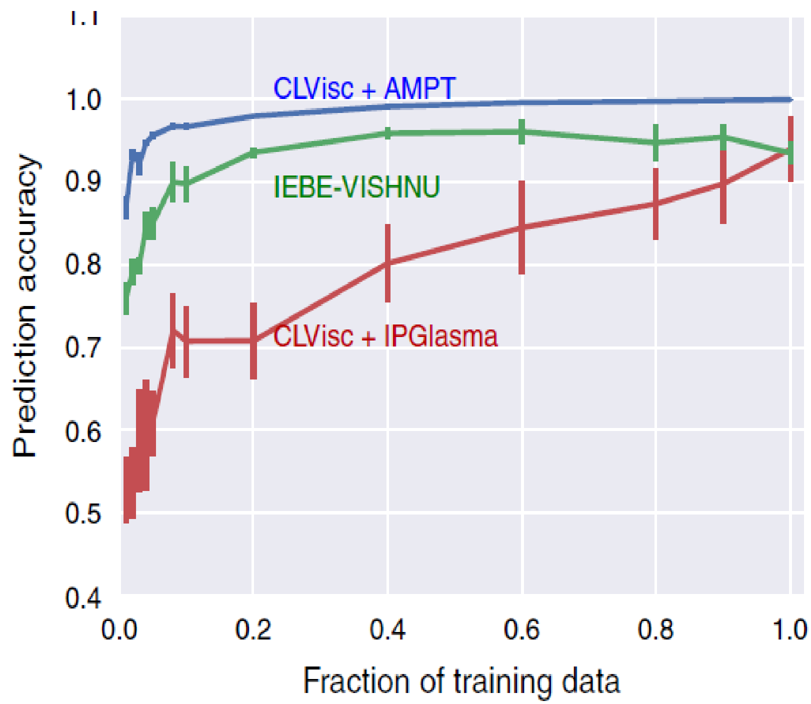
B) Training CNN

Table 1 The training data set Hydro CLVis (AMPT)

Training data set	$\eta/s=0$ EOSL	$\eta/s=0$ EOSQ	$\eta/s=0.08$ EOSL	$\eta/s=0.08$ 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

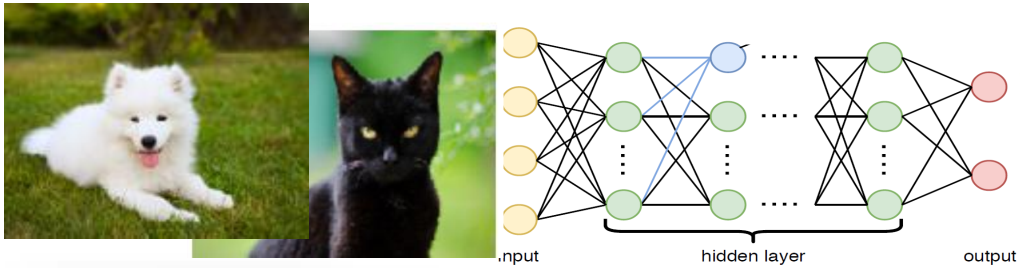


One can efficiently decode the EOS information from the complex final particle info event by event using deep learning

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

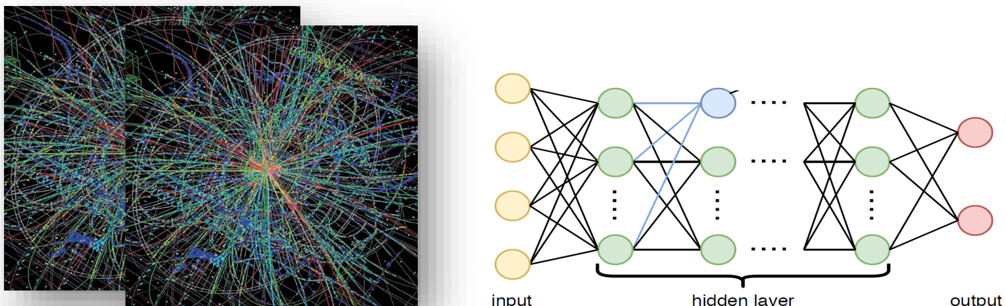
More Comments on several examples of supervised learning

Image classification



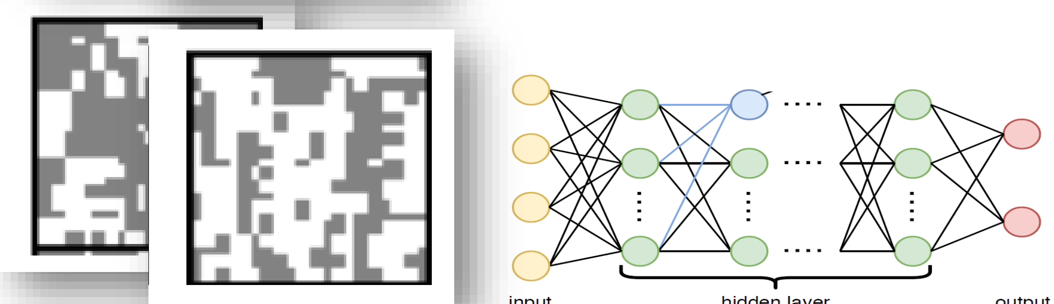
Dog or Cat ?

Yes or No ?



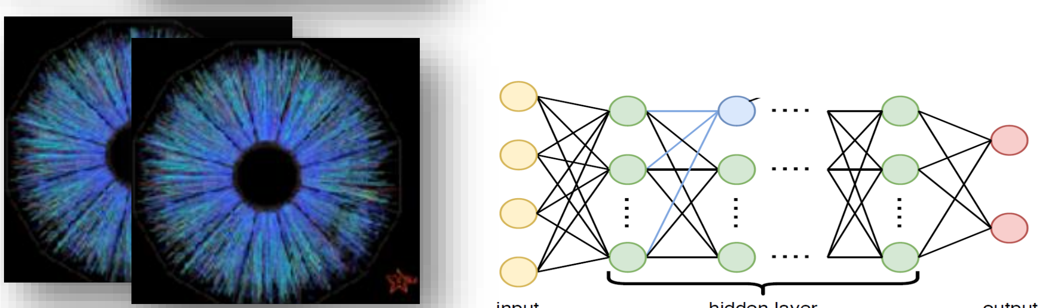
Higgs signal or background?

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



High temperature or low temperature phase?

Carrasquilla & Melko. Nature Physics (2017)

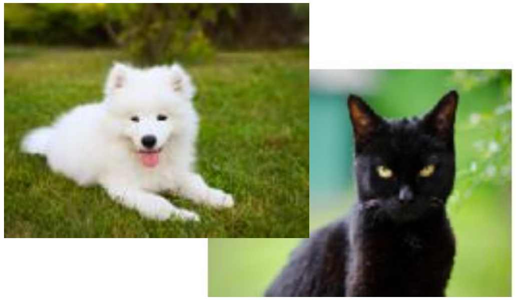


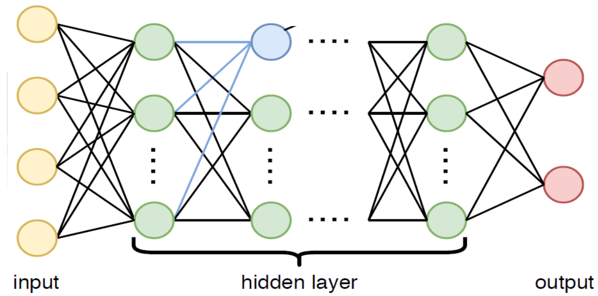
EoS L or EOSQ ?

Pang,et al Nature Commun.(2018)

Deep learning can do more … …

Image classification





Dog or Cat ?

Image generation

A. van den Oord et al., NIPS, (2016), arXiv: 1606.05328

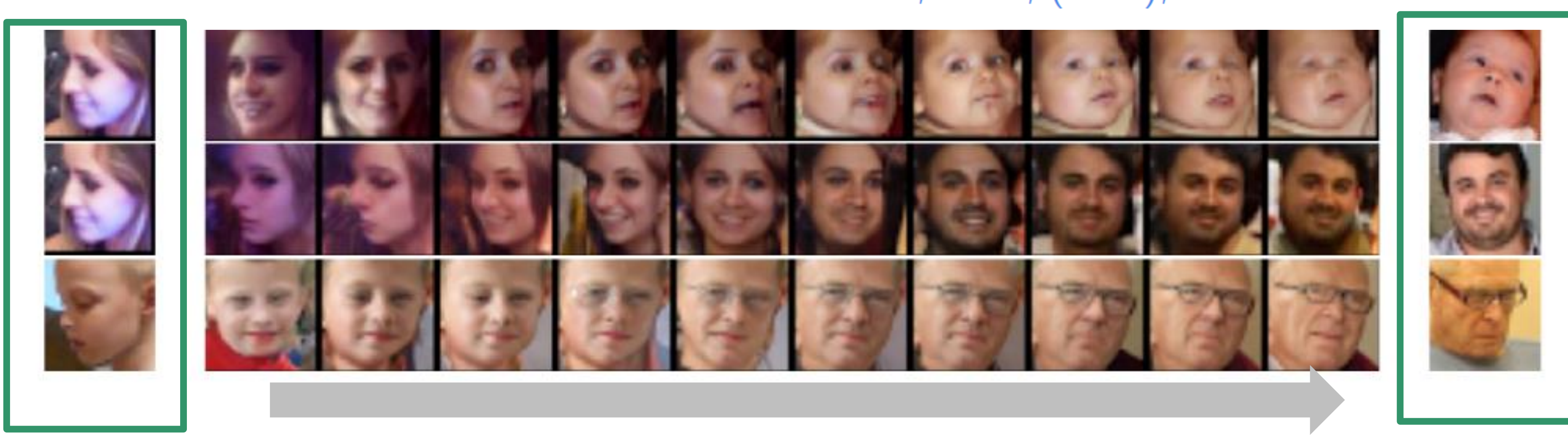
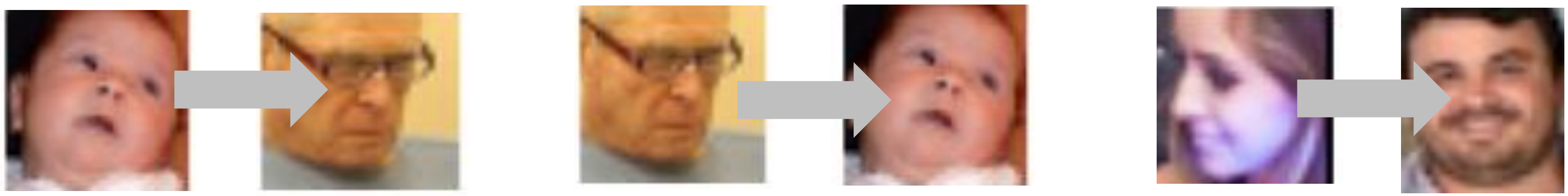


Image generation



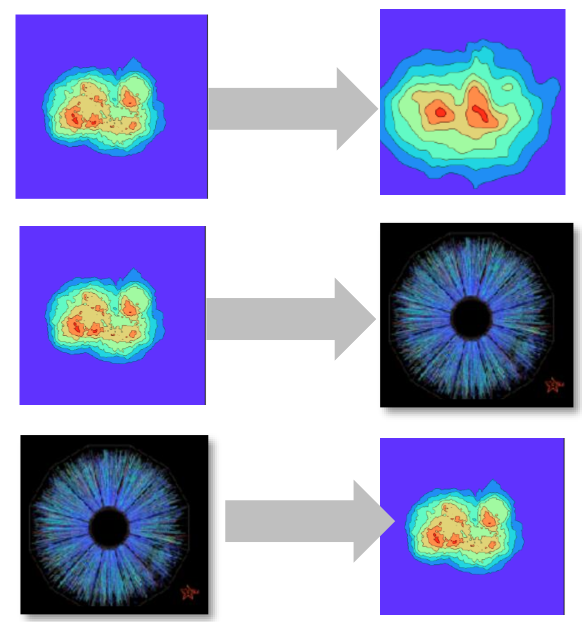
For hydrodynamics

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

Initial energy density profiles
-------- > final energy density velocity profiles

Initial energy density profiles
-------- > final particle profiles

Final particle profiles
-------- > Initial energy density profiles

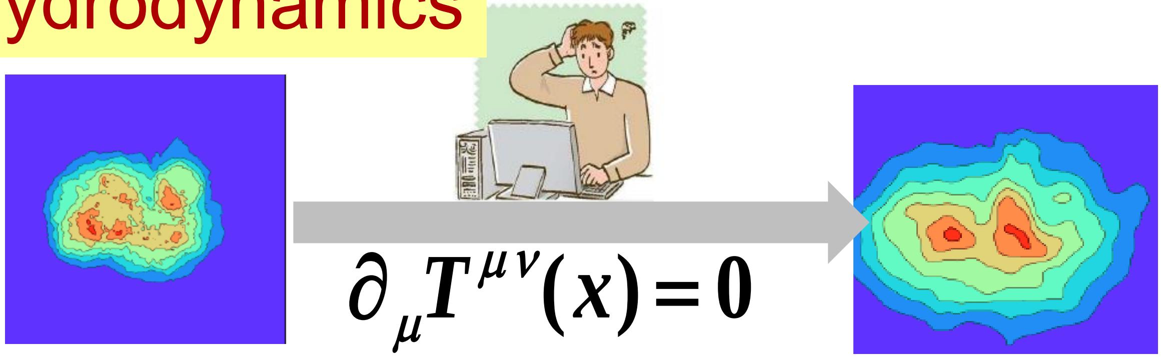


For the non-linear hydro system, can the "**Black-box**" network could learn patterns solely from data without relying on scientific knowledge? (conservation laws)

Applications of deep learning to relativistic hydrodynamics

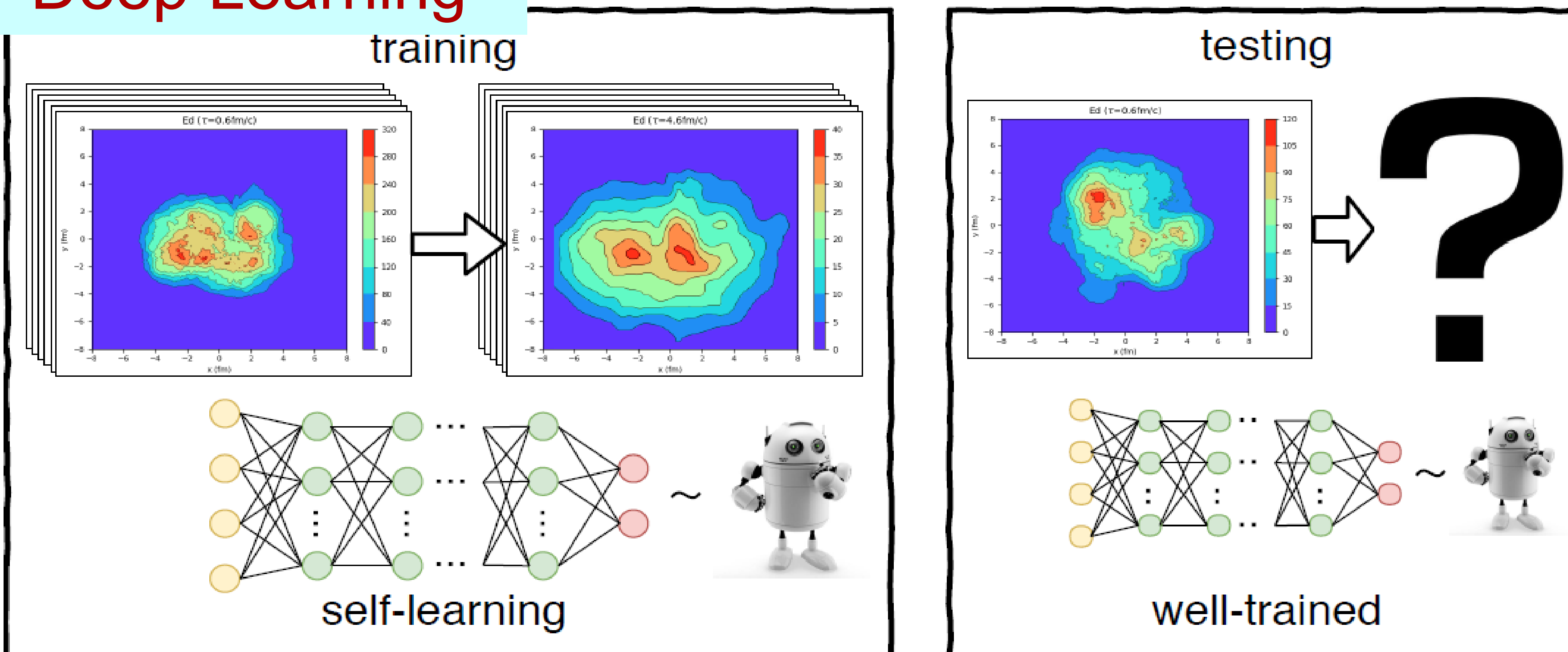
H.Huang, B.Xiao, H.Xiong, Z.Wu, Y. Mu and H.Song
arXiv: 1801.03334

Traditional hydrodynamics



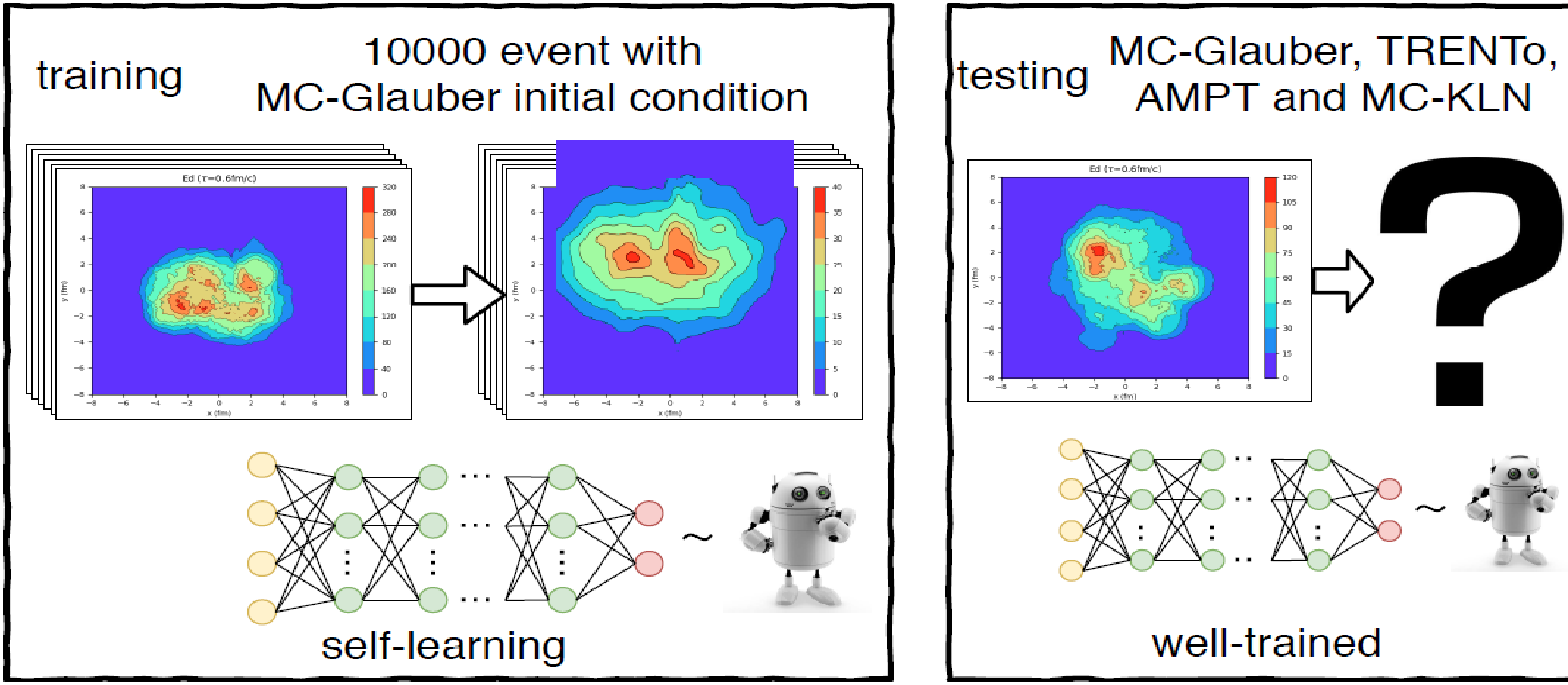
$$\partial_\mu T^{\mu\nu}(x) = 0$$

Deep Learning



-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 (VISH2+1)

Initial & final energy momentum tensor profiles ----> initial & final image sets

Step2) Design & train the deep neural network

Training sets: initial & final profiles from hydro with MC-Glauber initial conditions

Step3) Test the deep neural network

Compare DNN predictions with hydro results for different testing initial conditions (MC-Glauber, MC-KLN, AMPT Trento)

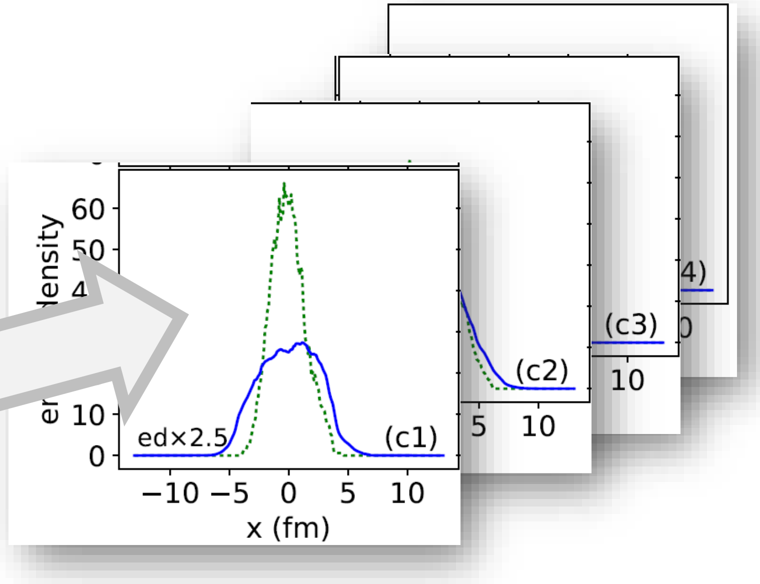
Deep Learning for 1+1-hydro

1) Generate the training/testing data sets from 1+1-d hydro

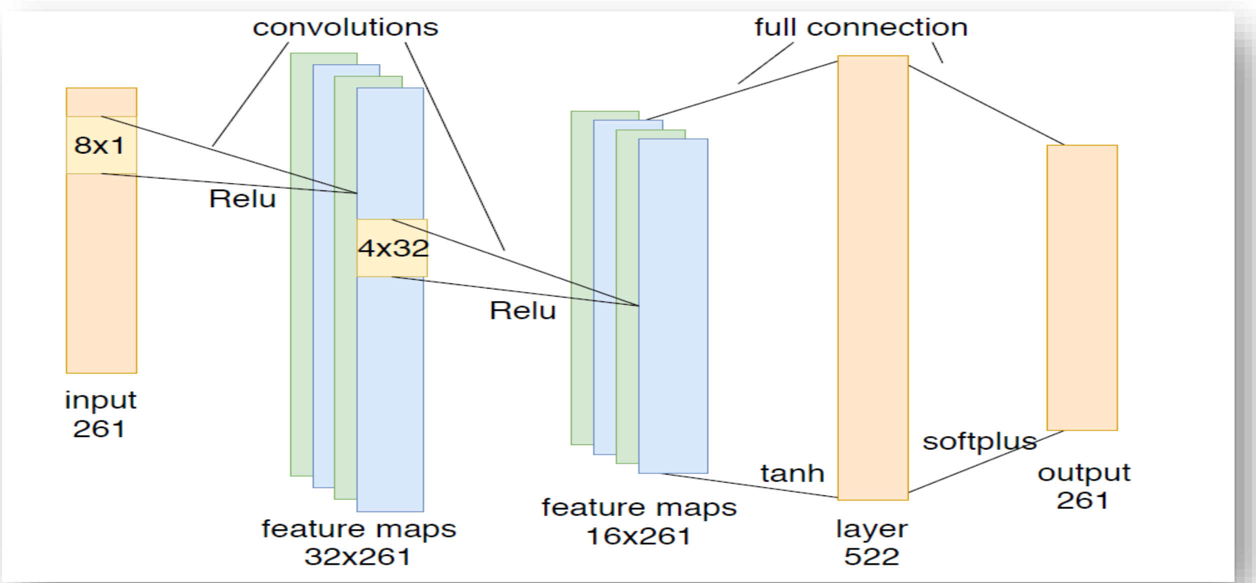
$$T^{\tau\tau}{}_{,\tau} + (\bar{v}_x T^{\tau\tau})_{,x} = -\frac{p+T^{\tau\tau}}{\tau} - (p\,\bar{v}_x)_{,x}$$

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

EoS: p=e/3, MC-Glauber/MC-KLN initial condtions $\tau - \tau_0 = 2.0,\ 4.0,\ 6.0\ \mathrm{fm/c}$



2) Design / train neural network (CNN)

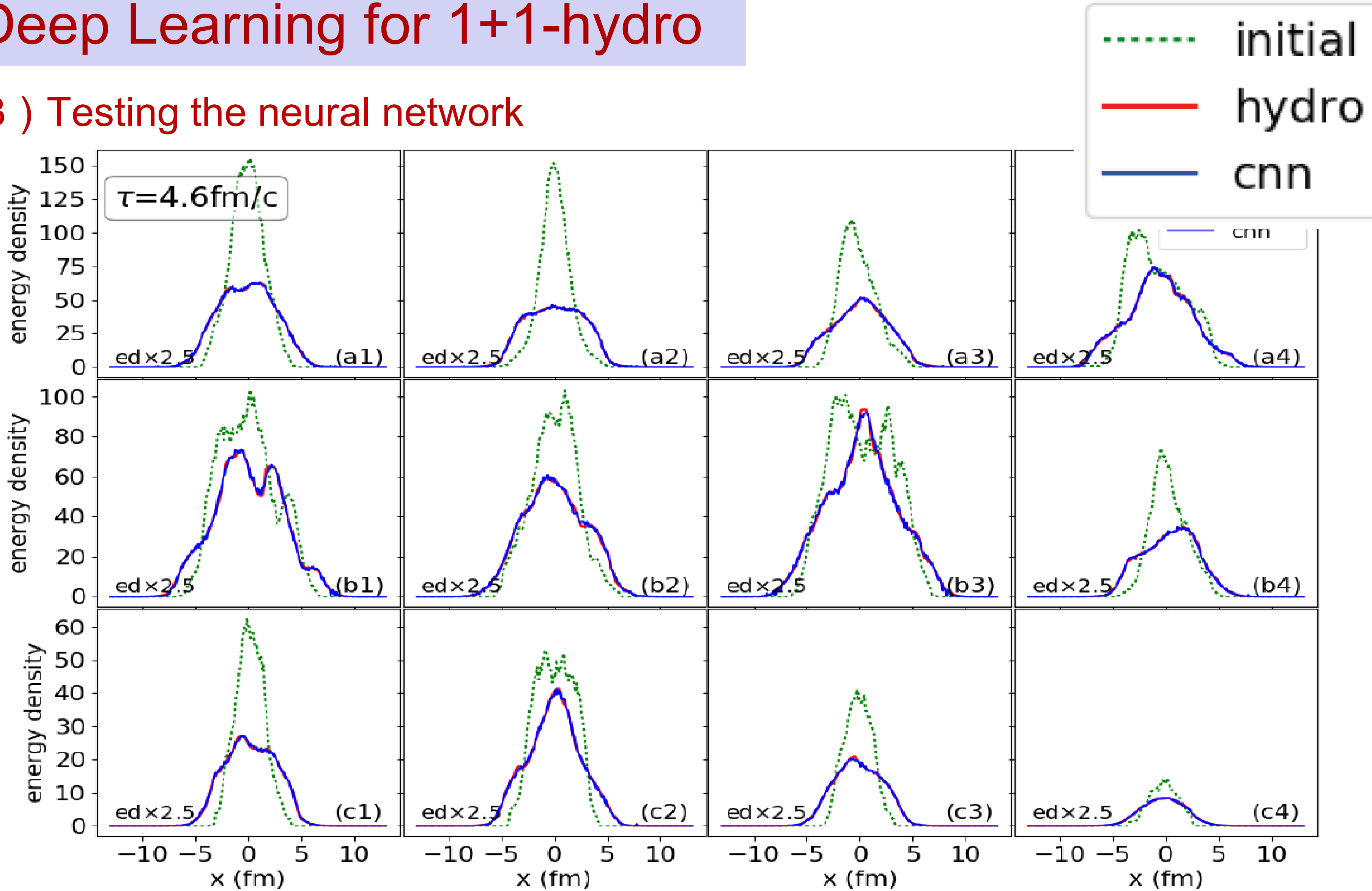


Loss Function $L(\theta) = \frac{1}{2}\frac{\sum(\hat{y}-y)^2}{\max\{y\}}$

Huang & Song unpublished notes

Deep Learning for 1+1-hydro

3) Testing the neural network



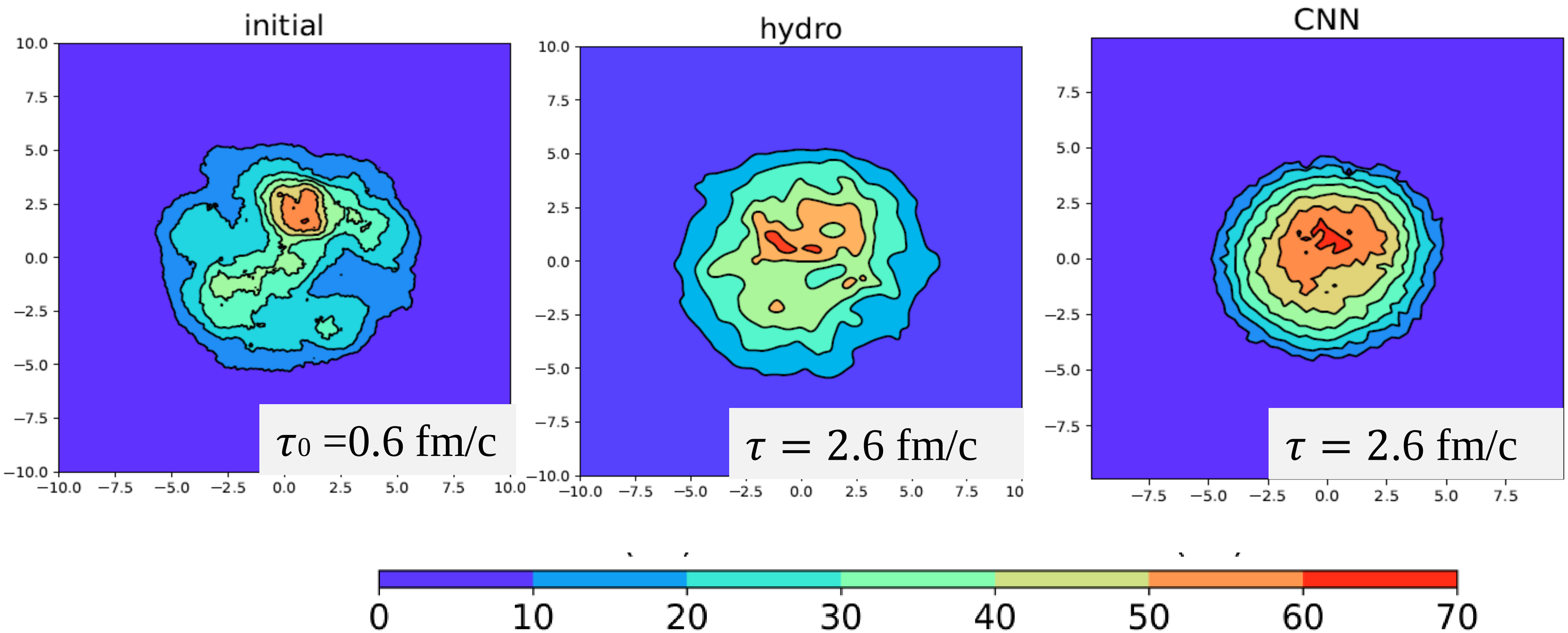
-CNN predictions vs. hydro simulation: CNN works for 1+1-d hydrodynamics

-A first hint that a well designed network could capture the non-linear evolution of hydrodynamics

Huang & Song unpublished notes

Deep Learning (CNN): an extension to 2+1-hydro

Initial condition | **Hydro results** | **CNN prediction**

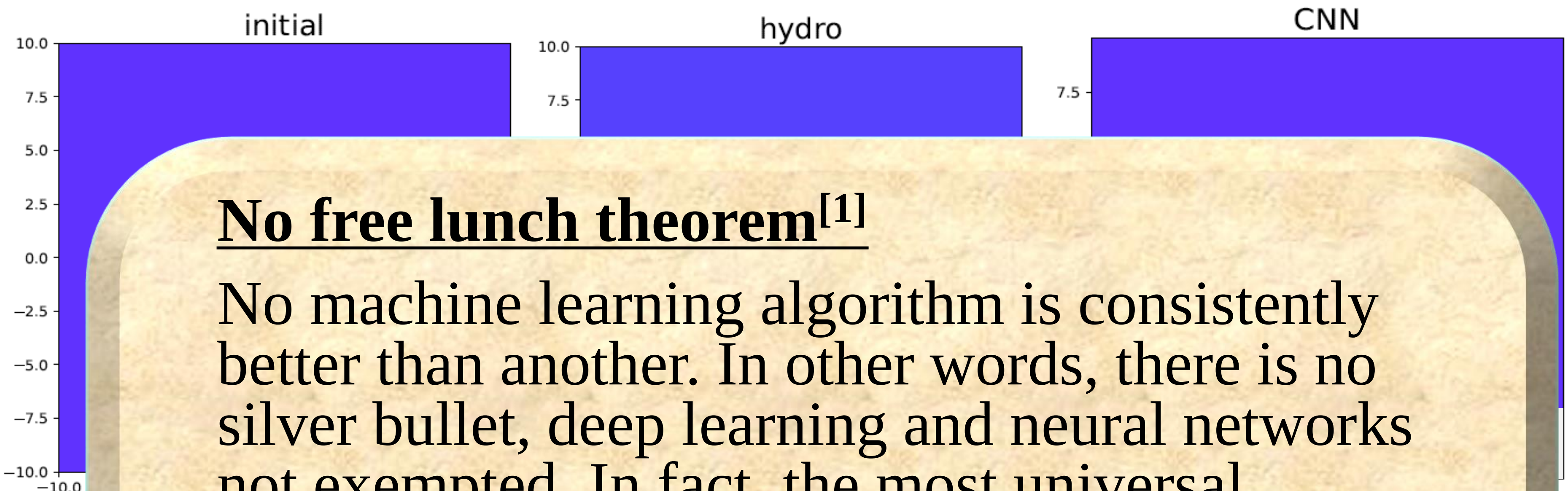


-For 2+1-d hydro, CNN does not work, so does not other common network, such as local connected layer

-From 1+1-d hydro to 2+1-d hydro, the pixel of the imagine sets increased from 200 to 40000 (200*200)

Deep Learning (CNN): an extension to 2+1-hydro

Initial condition **Hydro results** **CNN prediction**

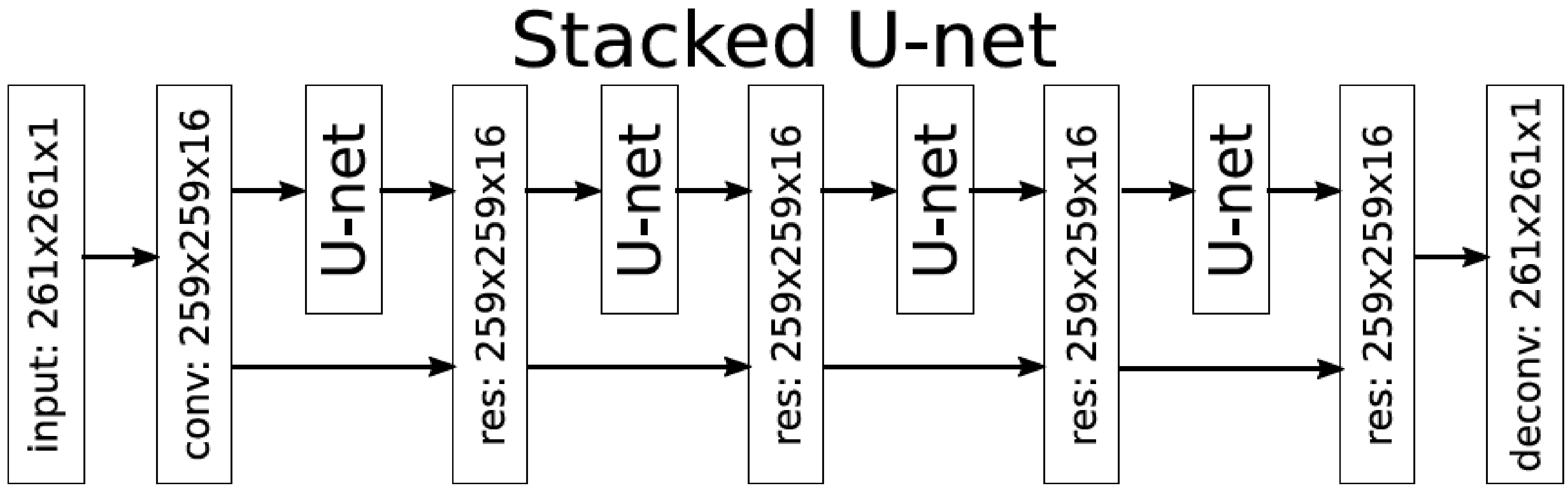


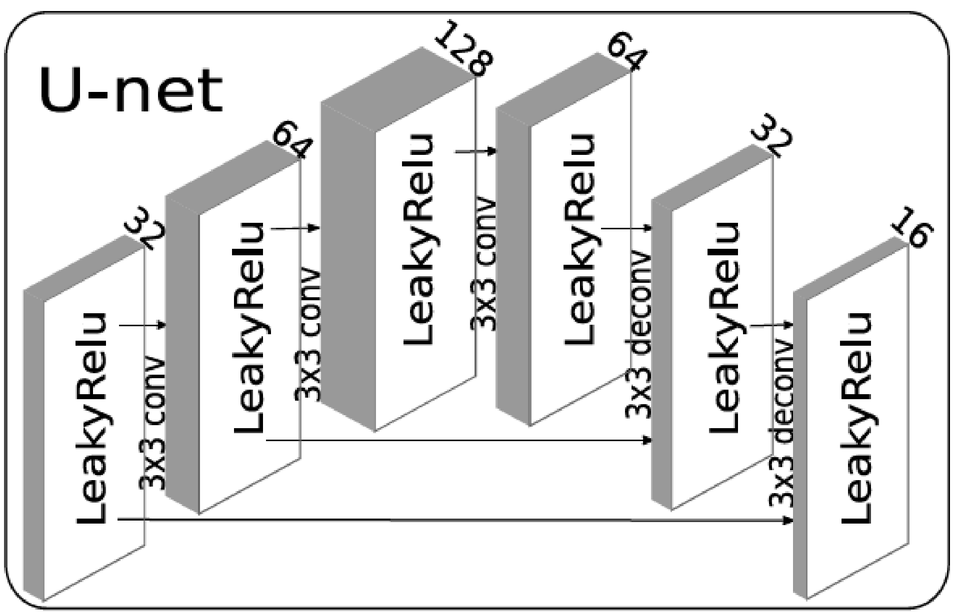
No free lunch theorem[1]

No machine learning algorithm is consistently better than another. In other words, there is no silver bullet, deep learning and neural networks not exempted. In fact, the most universal feedforward neural network does worse than tree based methods or SVM on many problems. Therefore, when adapting a model to new problems, one should be aware of model assumptions and ensures that they still holds.

[1] Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural computation*, 8(7), 1341-1390.

Stacked U-net for 2+1-d hydro





The activation function:

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

The loss function:

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

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

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

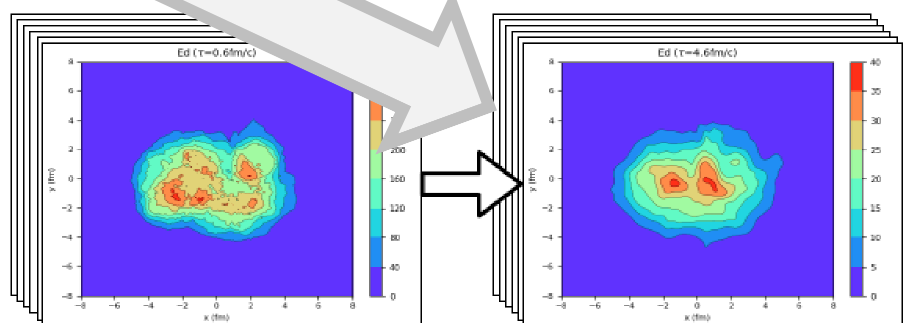
$$T^{\tau\tau}{}_{,\tau} + (\bar{v}_x T^{\tau\tau})_{,x} + (\bar{v}_y T^{\tau\tau}) = -\frac{p+T^{\tau\tau}}{\tau} - (p\,\bar{v}_x)_{,x} - (p\,\bar{v}_y)_{,y}$$

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

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

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

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



The Training Data Sets

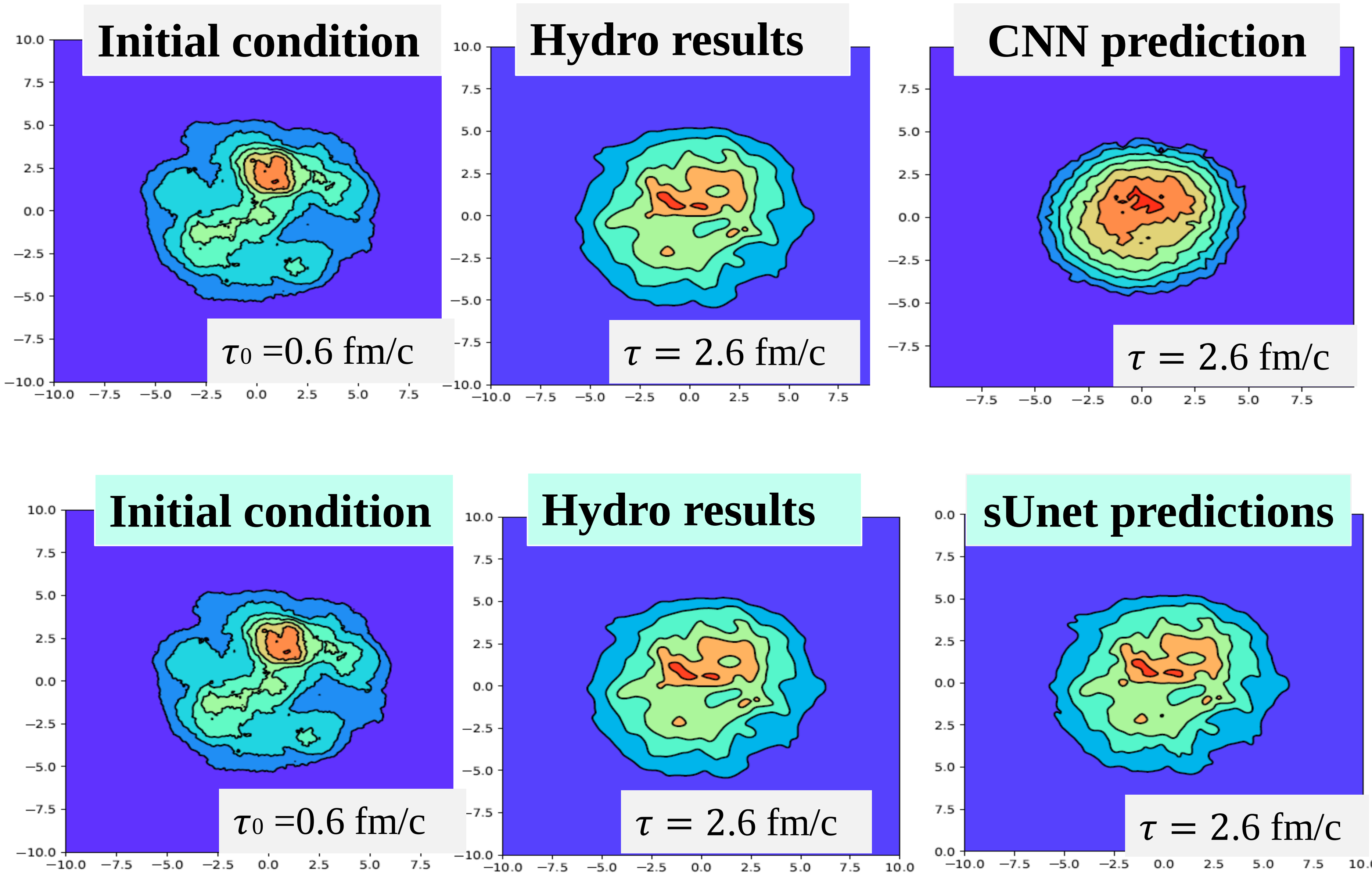
2+1-d hydro VISH2+1	MC-Glauber
	10000 events

The Testing Data Sets

2+1-d hydro VISH 2+1	MC-Glauber	MC-KLN	AMPT	Trento
	10000 events	10000 events	10000 events	*10000 events*

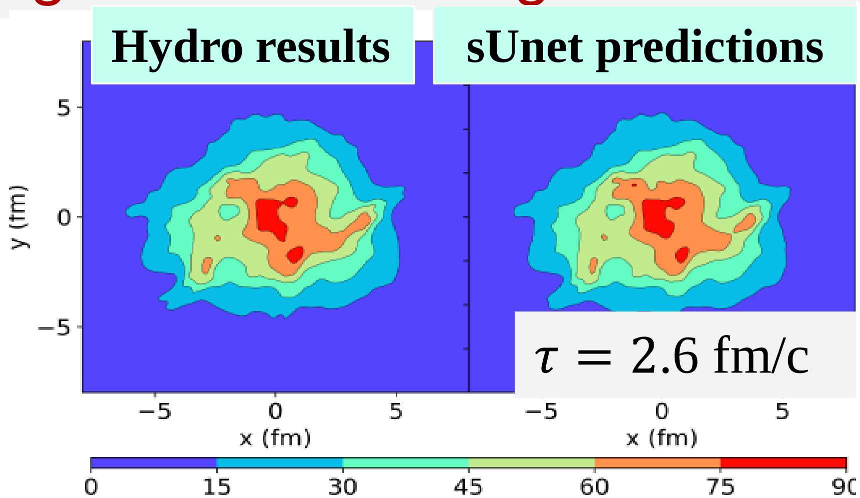
H.Huang, B.Xiao, H.Xiong, Z.Wu, Y. Mu and H.Song arXiv: 1801.03334

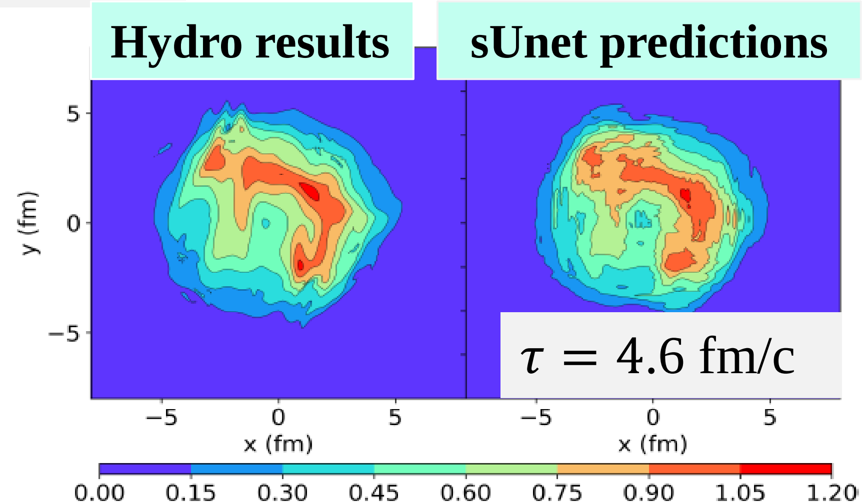
Predictions: Stacked U-net vs. CNN



-sUnet is the proper directions that works

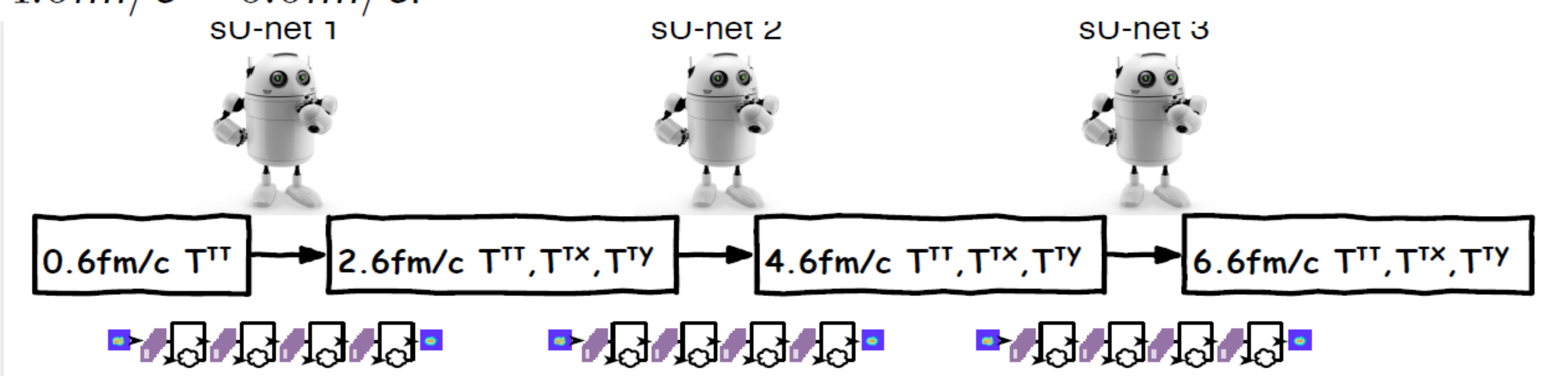
Single sUnet for longer time evolution





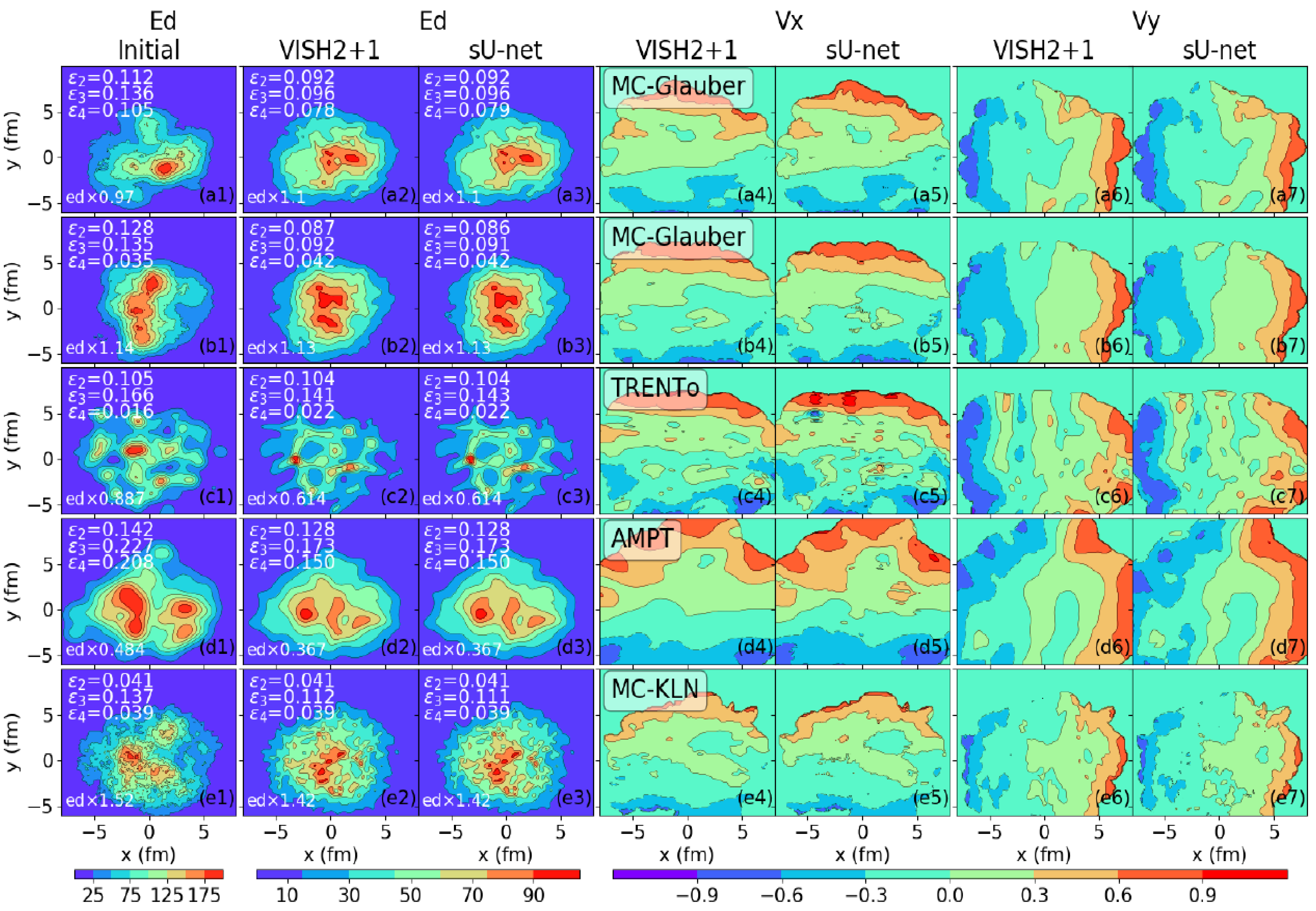
Combined sUnet for long time evolution

- Due to the size of GPU memory, it is costly to increase the number of U-net, so we divide the whole evolution time $\tau_0 - \tau$ into 3 parts with equal time interval $\Delta\tau$: $0.6fm/c - 2.6fm/c$, $2.6fm/c - 4.6fm/c$, $4.6fm/c - 6.6fm/c$.



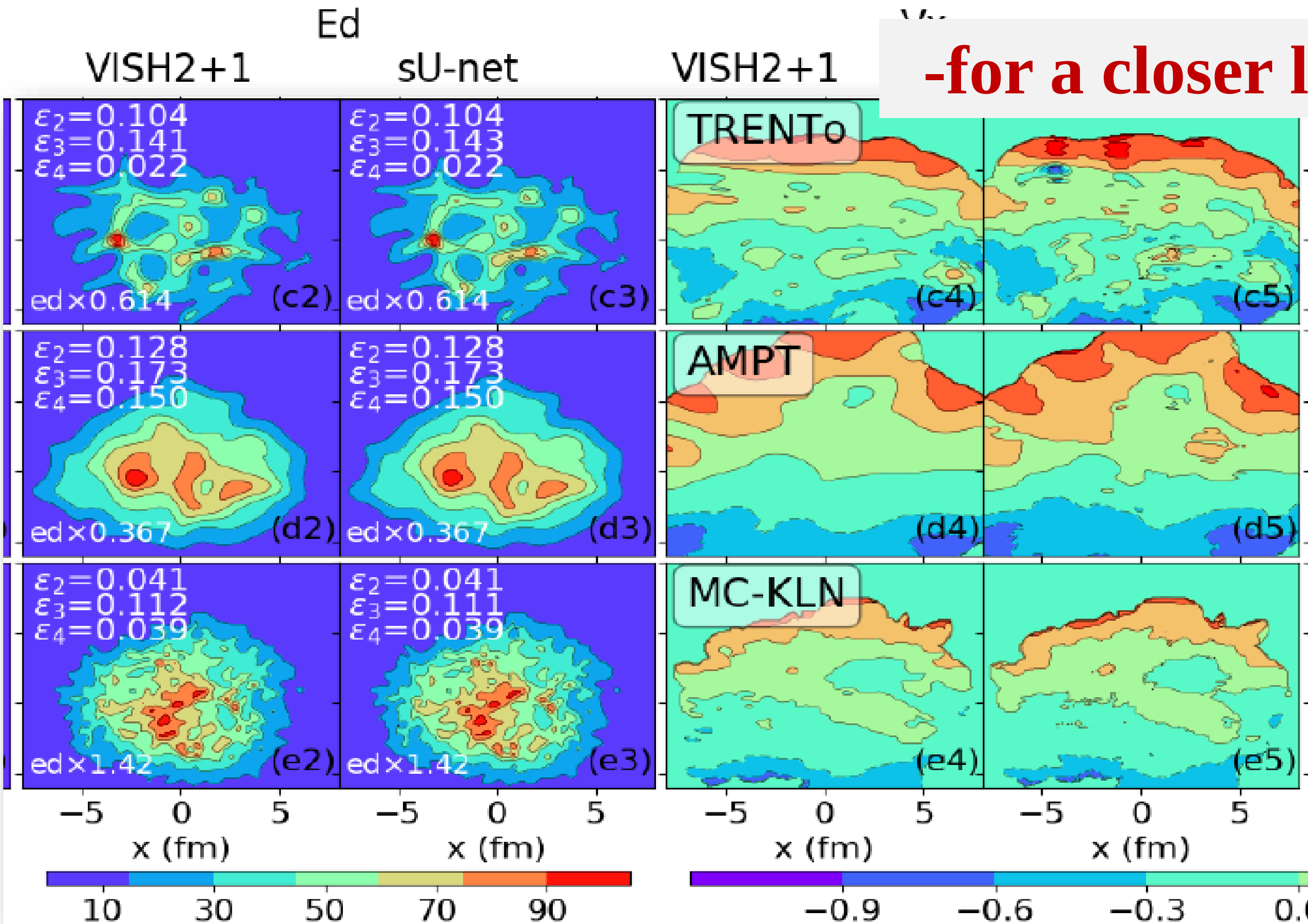
sUnet prediction vs. hydro simulations

$$\tau - \tau_0 = 2.0 fm/c$$



sUnet prediction vs. hydro simulations

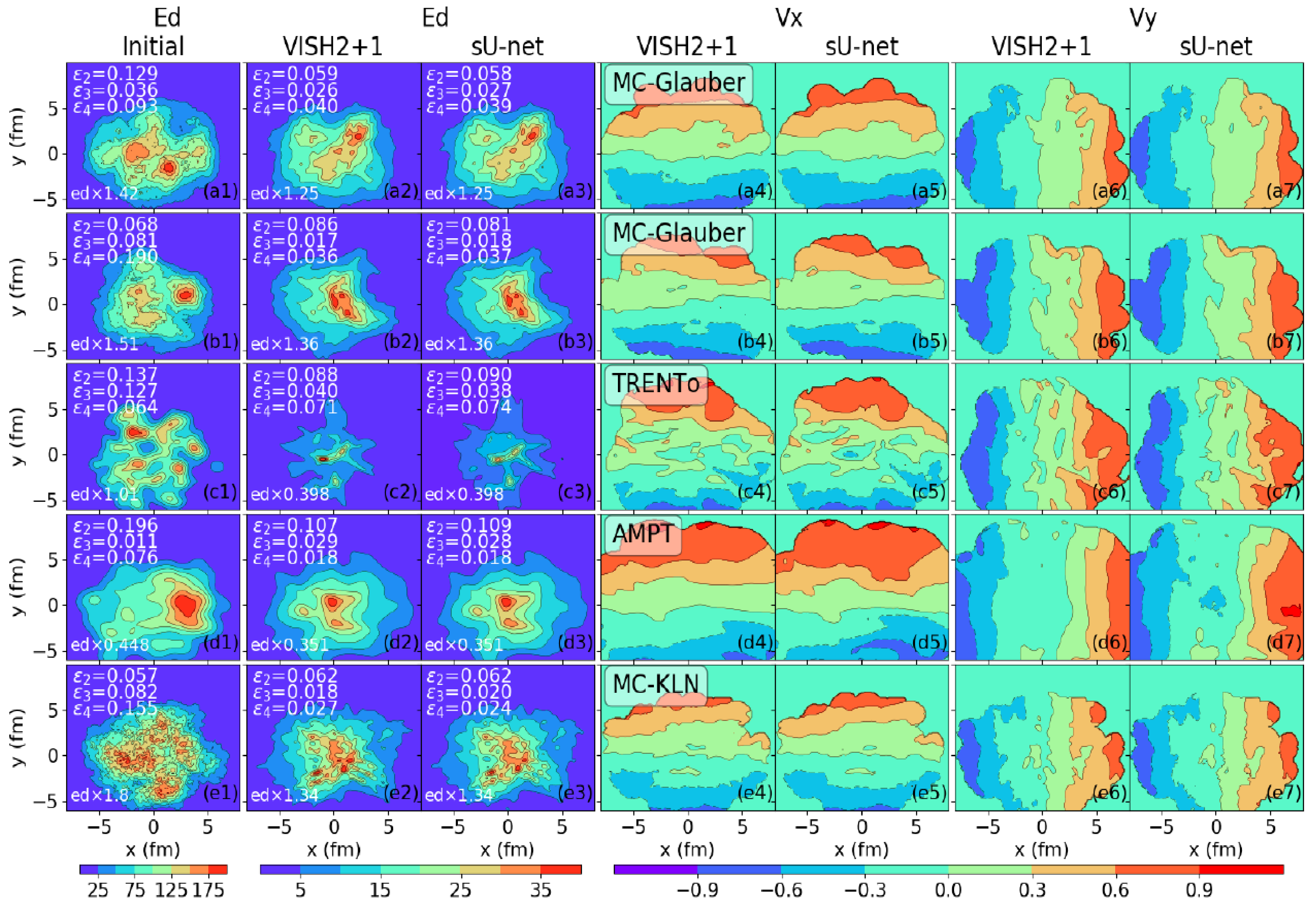
$\tau - \tau_0 = 2.0fm/c$



-for a closer look

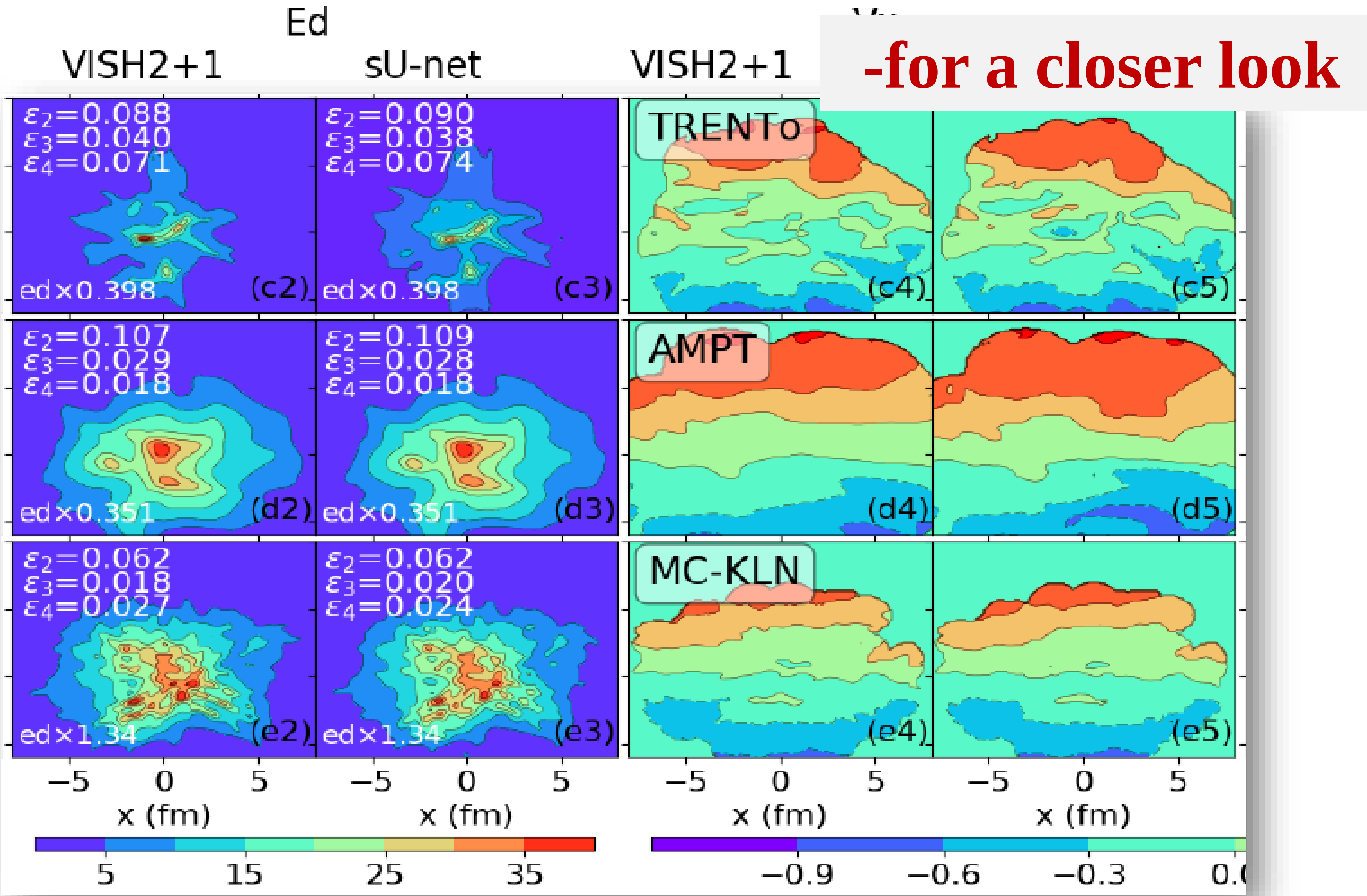
sUnet prediction vs. hydro simulations

$$\tau - \tau_0 = 4.0 fm/c$$



sUnet prediction vs. hydro simulations

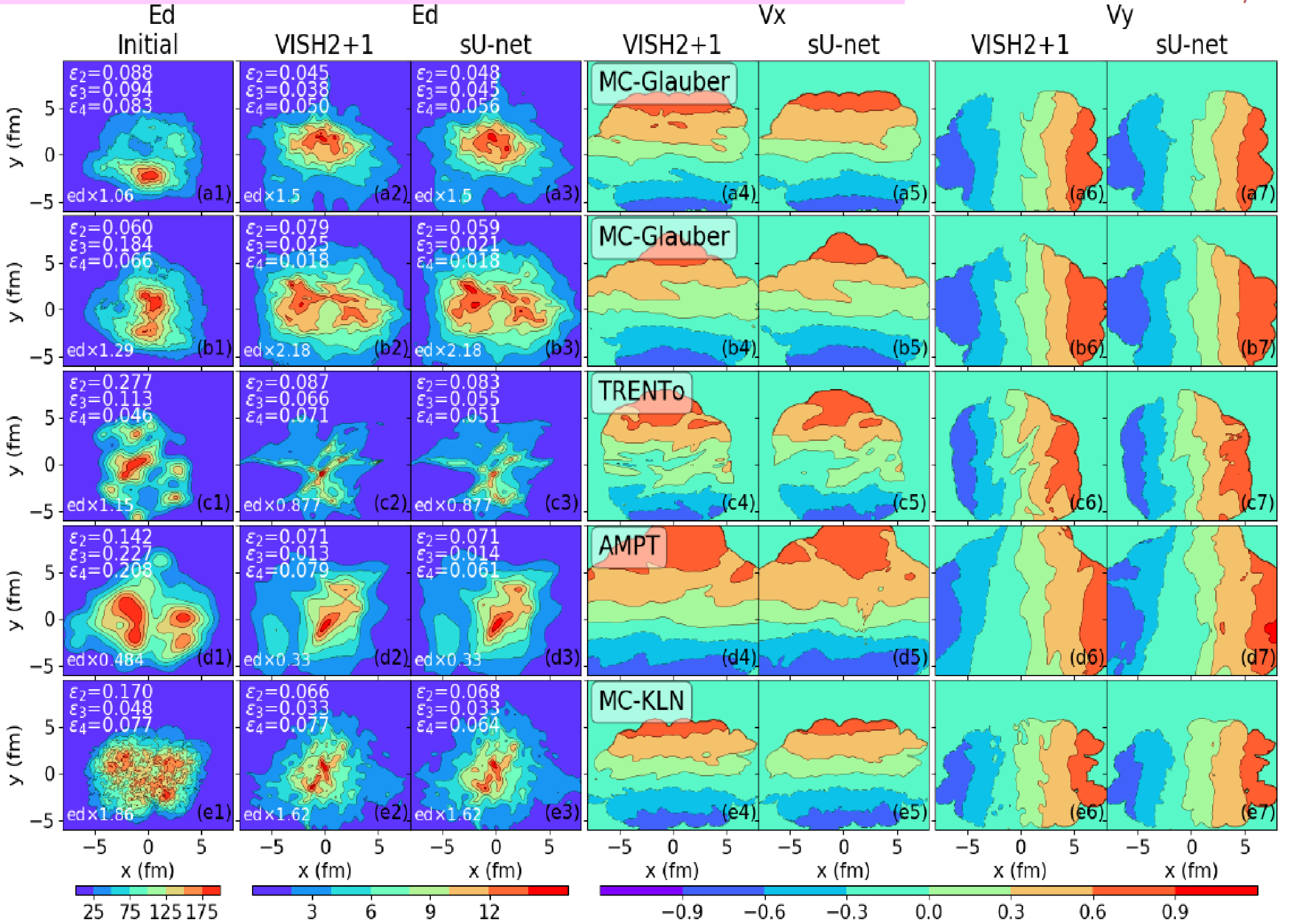
$\tau - \tau_0 = 4.0 fm/c$



-for a closer look

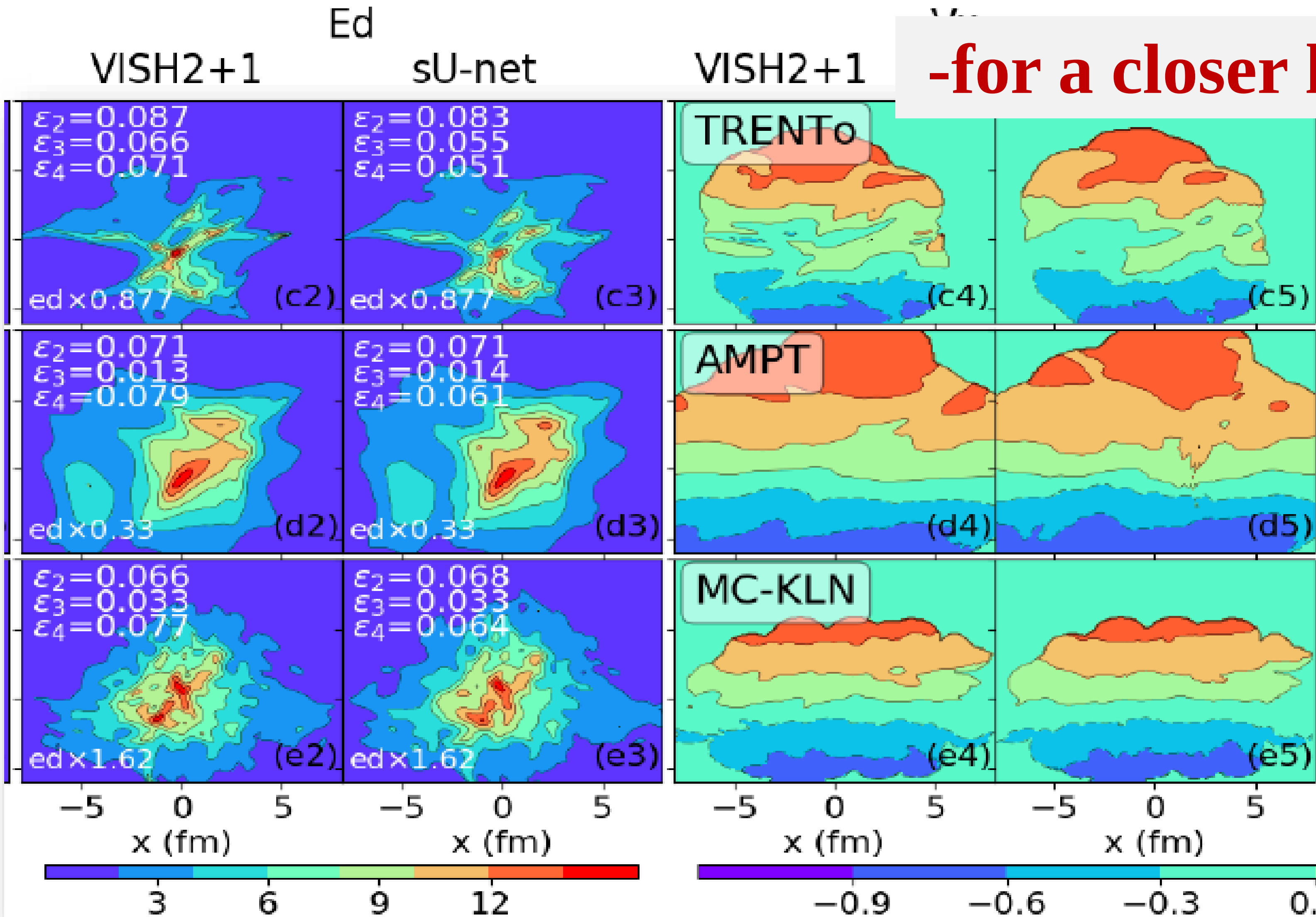
sUnet prediction vs. hydro simulations

$\tau - \tau_0 = 6.0 fm/c$



sUnet prediction vs. hydro simulations

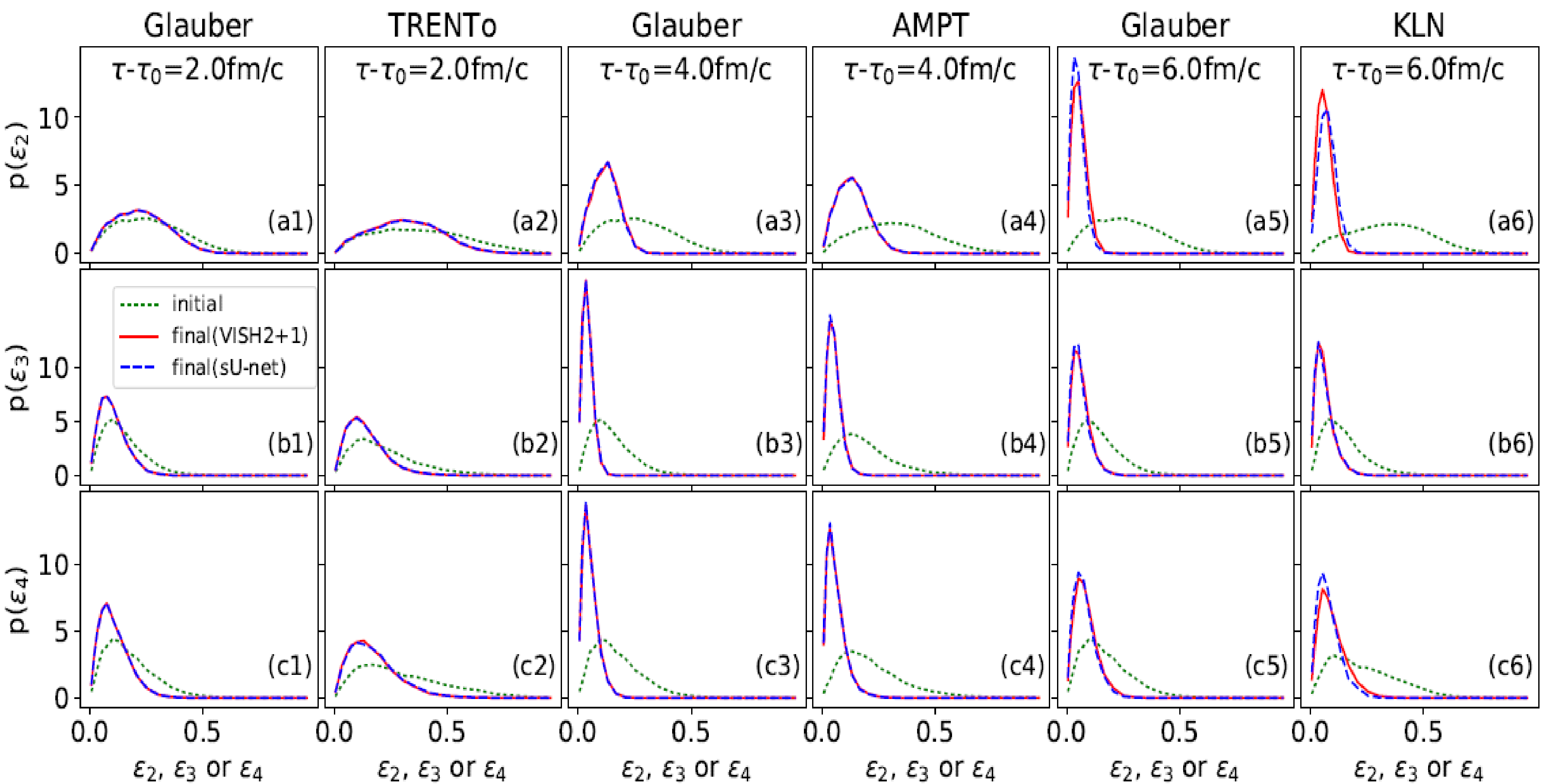
$\tau - \tau_0 = 6.0 fm/c$



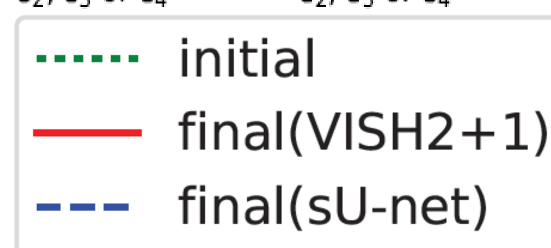
-for a closer look

sUnet prediction vs. hydro simulations

Eccentricity distributions:



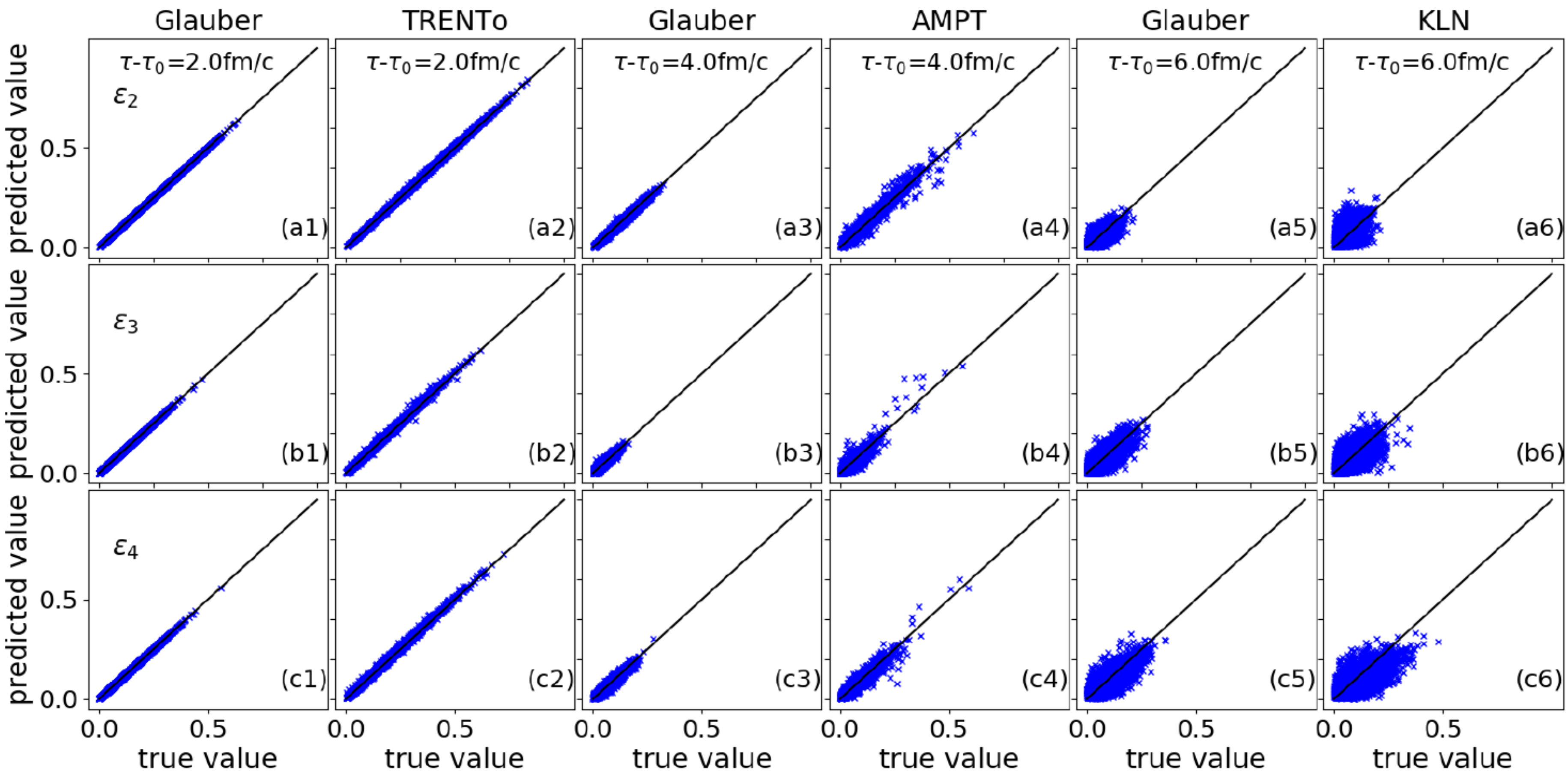
$$\varepsilon_n e^{in\Phi_n} = -\frac{\int dxdyr^2 e^{in\phi} e(x,y)}{\int dxdyr^2 e(x,y)}$$



sUnet prediction vs. hydro simulations

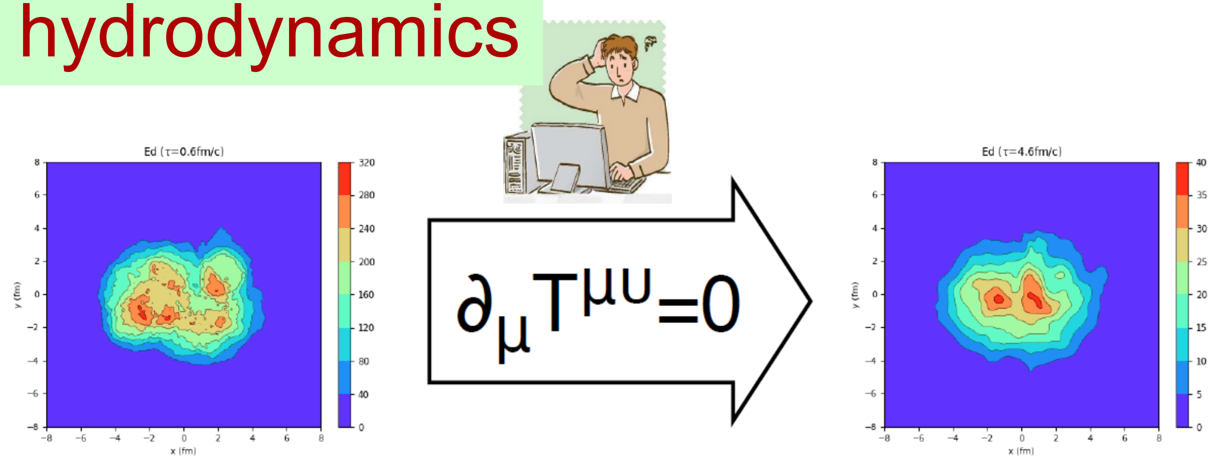
Histograms of ε_n

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

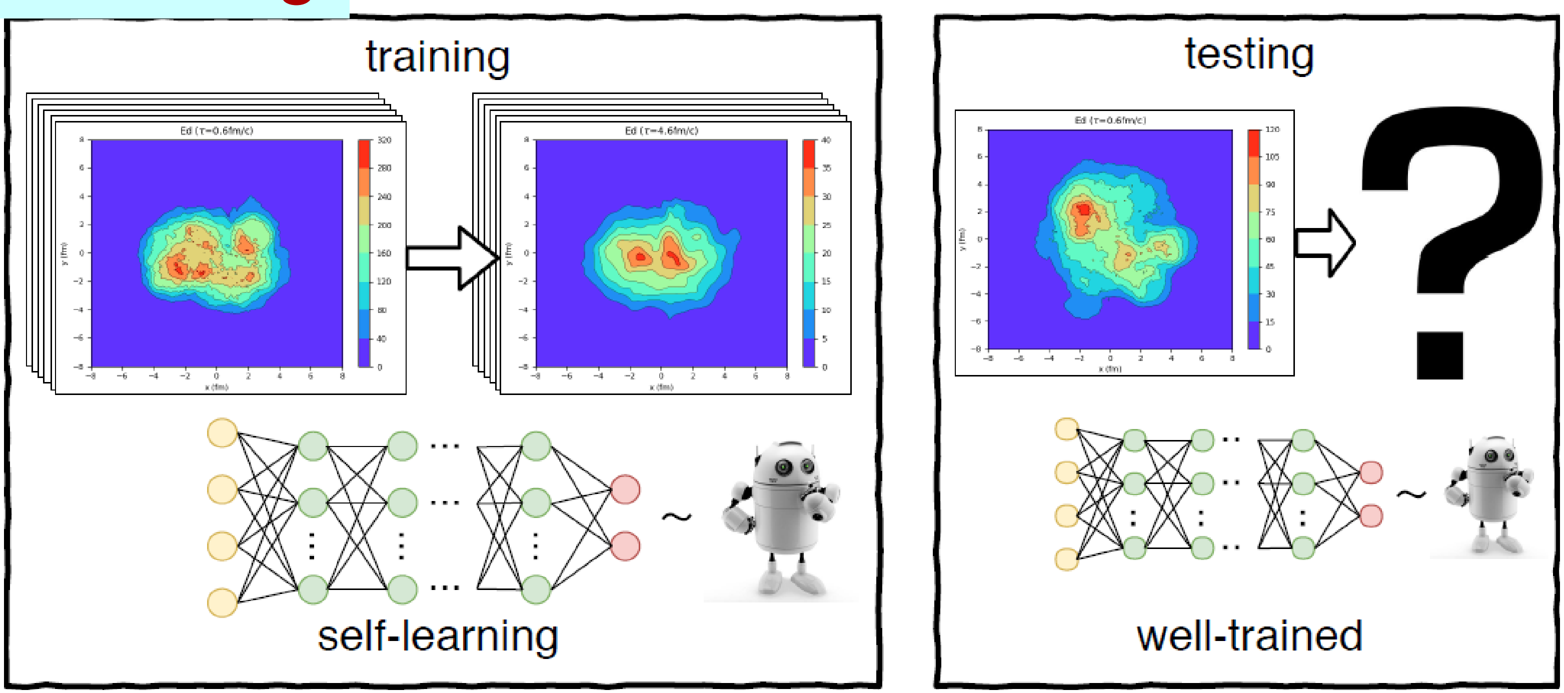


Summary & outlook

Traditional hydrodynamics

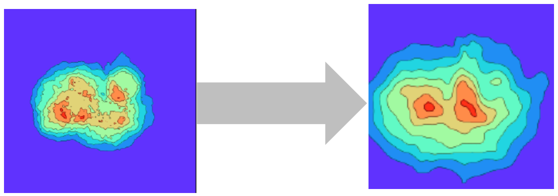


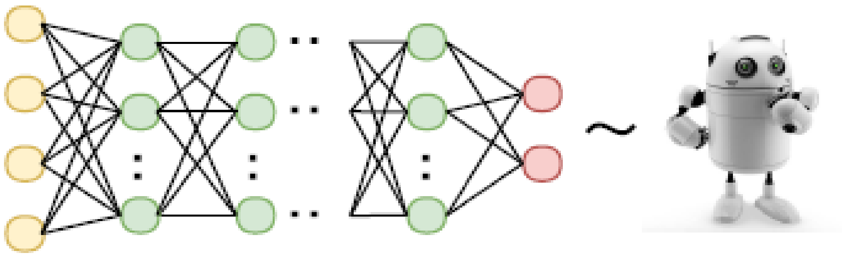
Deep Learning



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

Initial energy density profiles
-------- > final energy density velocity profiles





Summary

-Using 10000 initial and final profiles generated from VISH2+1 with MC-Glauber initial condition, we train the network called sU-net.

-We use the well-train network to predict the final profiles from with different initial conditions, including MC-Glauber, TRENTo, AMPT and MC-KLN.

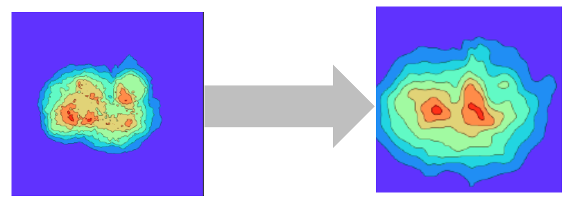
-Out results show that deep learning can predict the magnitude and inhomogeneous structures of the final energy density and flow velocity, which can also describe the related eccentricity distribution $P(\varepsilon_n)$.

-Deep learning can capture the main features of the non-linear evolution of hydrodynamics.

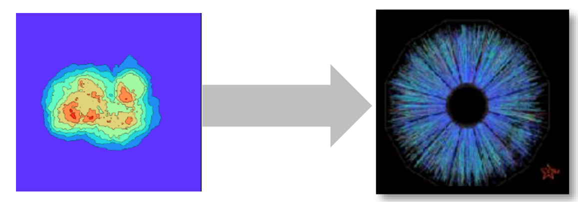
Outlook

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

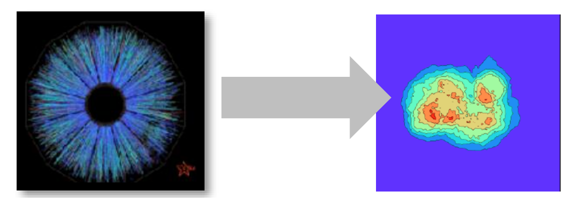
Initial energy density profiles
-------- > final energy density velocity profiles



Initial energy density profiles
-------- > final particle profiles

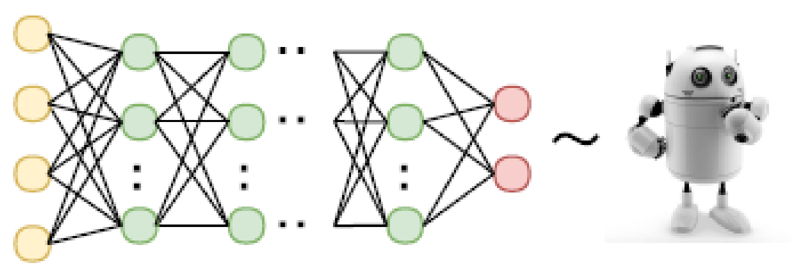


Final particle profiles
-------- > Initial energy density profiles



For flow in large and small systems

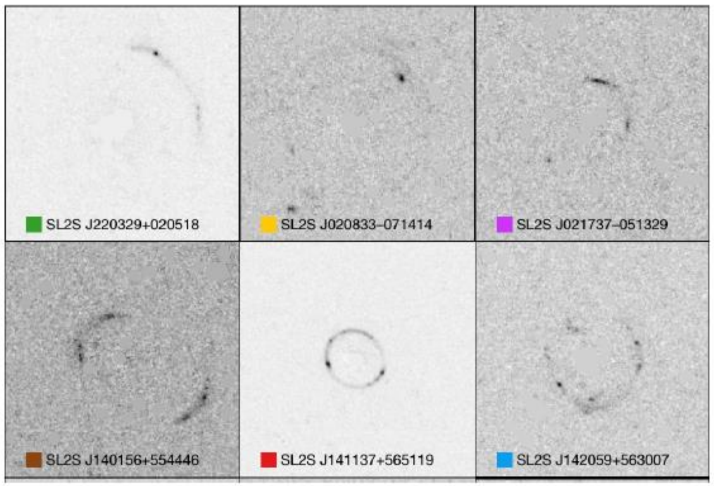
open for discussions

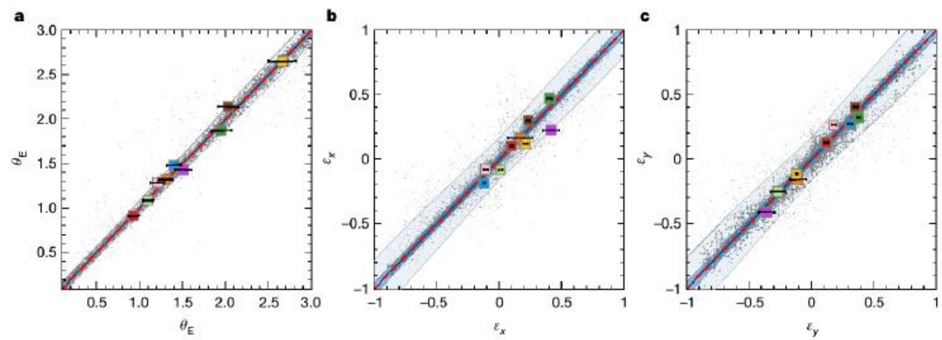


Thank You

Fast automated analysis of strong gravitational lenses

Y. D. Hezaveh, L. Perreault Levasseur and P. J. Marshall, Nature 548, 555 (2017)

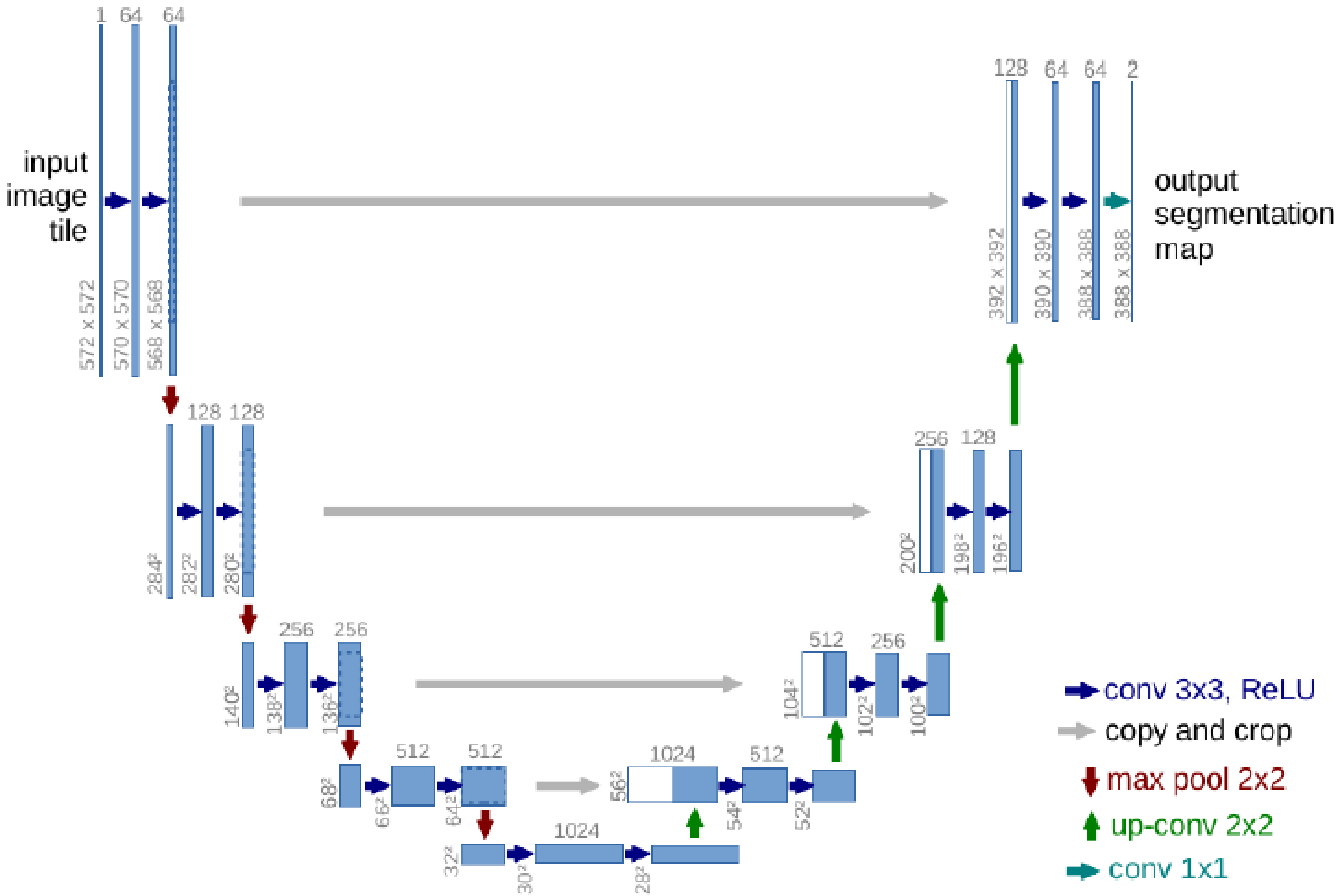




- Inputing the lensed galaxy images, CNN can estimate lensing parameters in an extremely fast and automated way.

Unet & Stacked Unet

-Inspiration from biomedical image segmentation



O. Ronneberger, P. Fisher, and T. Brox, MICCAI (3), volume 9351 of Lecture Notes in Computer Science, page 234-241. Springer, (2015), arXiv:1505.04597