Applications of Deep Learning in Relativistic Hydrodynamics

Hengfeng Huang
With Huichao Song

Peking University

May 16th, 2018

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

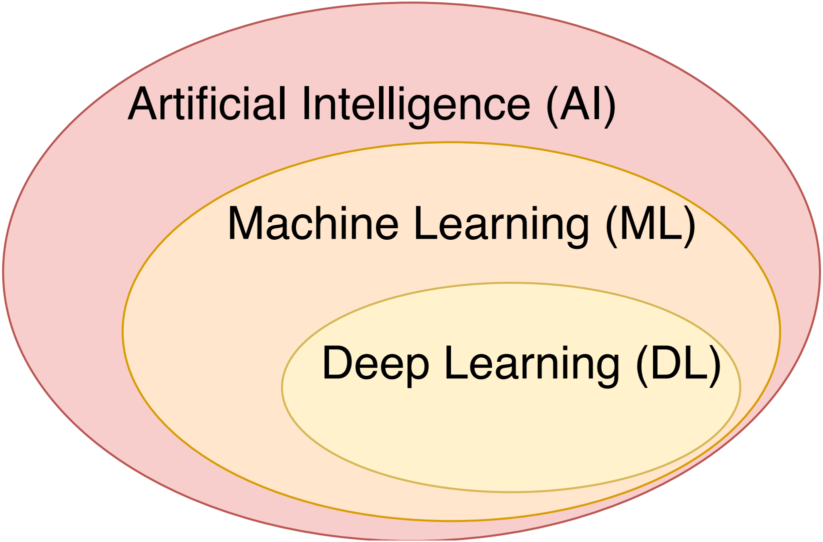
arXiv:1801.03334

Outline

First impression about deep learning

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

What is deep learning?



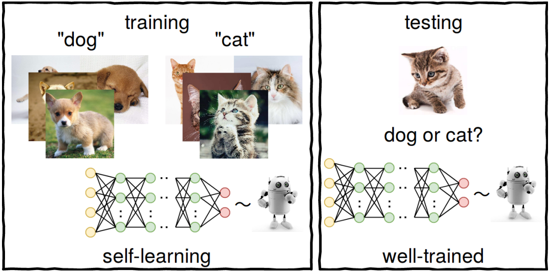
Categories of deep learning

- Supervised learning
- Unsupervised learning
- Reinforcement learning
- ...

Example of DL: dogs vs. cats

Supervised learning

experience a dataset contains many features, and each example is also associated with a label or target.

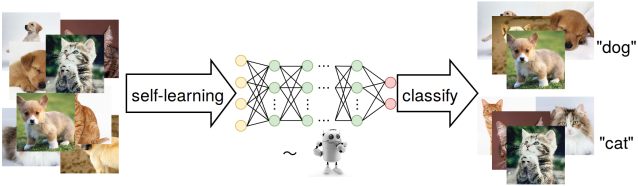


Ian Goodfellow, Yoshua Bengio, and Aaron Courville, http://www.deeplearningbook.org MIT Press, 2016

Example of DL: dogs vs. cats

Unsupervised learning

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

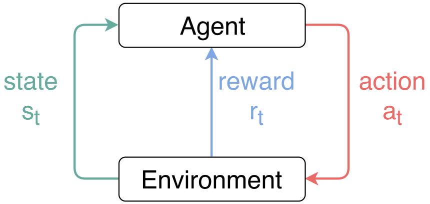


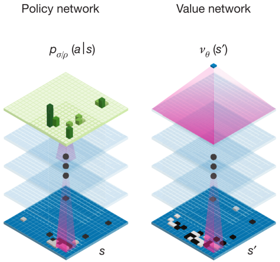
Ian Goodfellow, Yoshua Bengio, and Aaron Courville, http://www.deeplearningbook.org MIT Press, 2016

Reinforcement Learning

Reinforcement learning

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





WIKI:https://en.wikipedia.org/wiki/Reinforcement_learning

David Silver et al., Nature 529(2016) 484-489

What deep learning can do?

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

Language Processing

Machine translation

J. Lee, K. Cho, and T. Hofmann, TACL, (2017), arXiv: 1610.03017

(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 .

Chinese poetry generation

Z. Wang et al., COLING, page 1051-1060, ACL, (2016), arXiv: 1610.09889

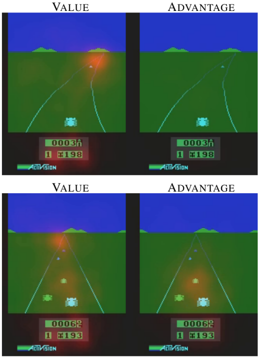
Speech recognition

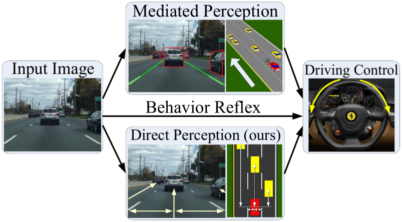
W. Xiong et al., IEEE/ACM Transactions on Audio Speech & Language Processing, 2016, PP(99), arXiv: 1610.05256

Playing Games and Autonomous Driving

Z. Wang, T. et al., ICML, volume 48 of JMLR Workshop and Conference Proceedings, page 1995-2003. JMLR.org, (2016), arXiv: 1511.06581

C. Chen et al., ICCV, page 2722-2730. IEEE Computer Society, (2015), arXiv: 1505.00256

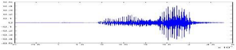




How does the deep learning work?

Similar to "Looking for a Function"

- Speech Recognition

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

- Image Recognition

$f($  $)=$ "Cat"

- Playing Go

$f($  $)=$ "5-5" (next move)

- Dialogue System

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

(what the user said) (system response)

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

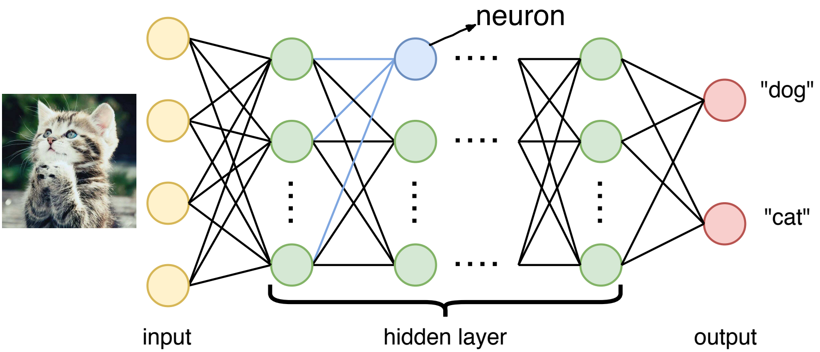
So, we should ...

1. define a set of function
2. evaluate each function
3. pick up the best

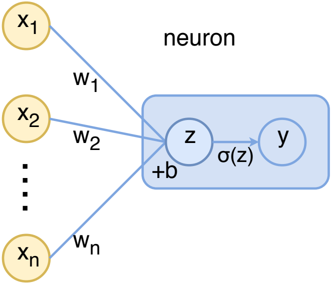
H. Lee, Deep Learning Tutorial,
https://www.slideshare.net/tw_dsconf/ss-62245351

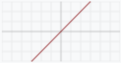
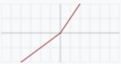
Neural Network

- In general, deep learning is to build up a neural network to connect the inputs and outputs.



Neuron



Identity		$f(x) = x$
Logistic (a.k.a. Sigmoid or Soft step)		$f(x) = \sigma(x) = \frac{1}{1+e^{-x}}$ [1]
Leaky rectified linear unit (Leaky ReLU)[11]		$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$

Neuron

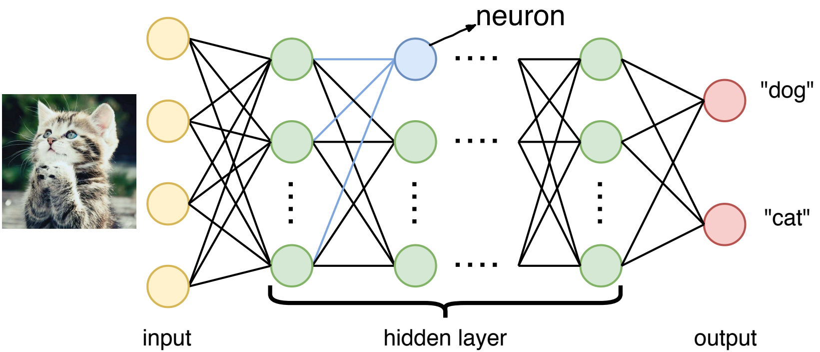
$$y = \sigma(z), \qquad z = w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + b$$

σ is a simple non-linear activation function

$w_1, w_2, \ldots, w_n, b$ is the trainable parameters

https://en.wikipedia.org/wiki/Activation_function

Neural Network



- With multi-neuron and multilayer, it can define a set of function with huge size.

Evaluate each function

- The evaluation is to evaluate the difference between the network's outputs and learning targets. — loss function

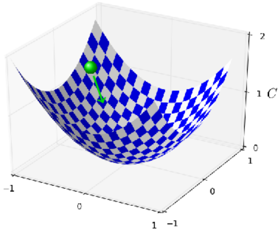
For supervised learning, since there has been a target $y(x)$, loss function can be defined as

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

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

Pick up the best

- In this way, pick up the best is the same to minimize the loss function.



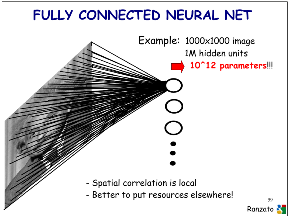
Stochastic Gradient Descent

$$\theta' = \theta - \epsilon \frac{\partial \ell(\theta)}{\partial \theta}$$

Michael A. Nielsen, "Neural Networks and Deep Learning", Determination Press, 2015

Common Network Structures

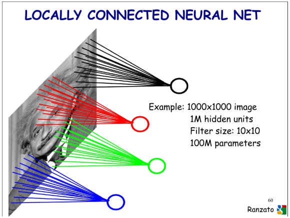
Fully Connected Network



recognize handwrite digits

. . .

Convolutional Neural Network

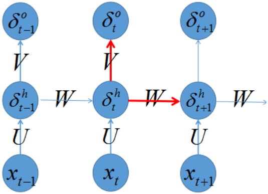


face recognition

image recognition

. . .

Recurrent Neural Network



speech recognition

Chinese poetry generation

. . .

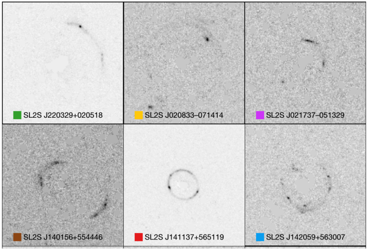
M. Ranzato, NEURAL NETS FOR VISION, https://cs.nyu.edu/~fergus/tutorials/deep_learning_cvpr12/tutorial_p2_nnets_ranzato_short.pdf

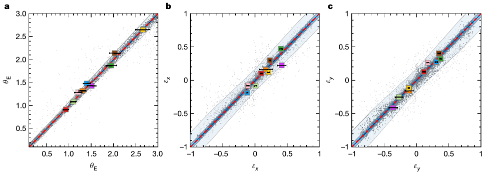
Applications of Deep Learning in Physics

- 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
- . . ., . . .,
- . . .

Fast automated analysis of strong gravitational lenses

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

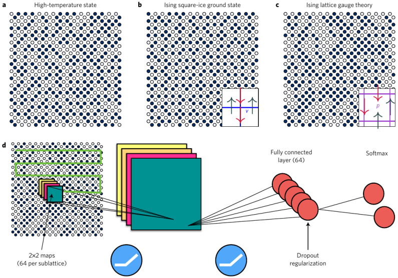


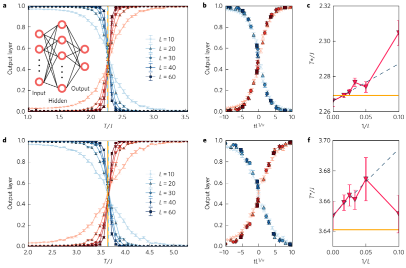


- Inputing the lensed galaxy images, CNN can estimate lensing parameters in an extremely fast and automated way.

Classifying the Phases of Ising Model

J. Carrasquilla and G. R. Melko, Nature Phys. 13, 431 (2017)

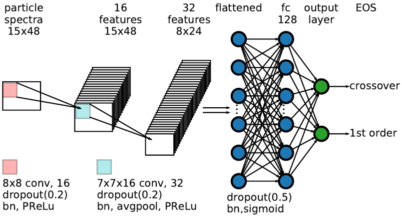


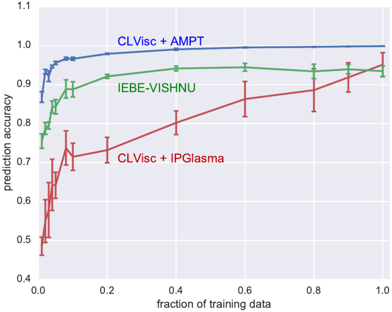


- By taking the spin configuration of Ising model with phase label as input, it can classisy the phase with high accuracy.

Identifying QCD transition

Long-Gang Pang et al., Nature Commun. 9 (2018) no.1, 210





- Processing the final state of heavy-ion collisions $\rho(p_T, \Phi)$ into images, the network can identify the QCD transition.

Applications of deep learning to relativistic hydrodynamics

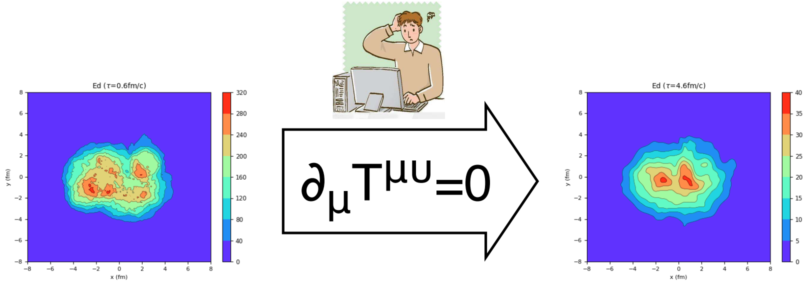
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu and Huichao Song
arXiv: 1801.03334

Traditional hydrodynamics

ideal hydro equations

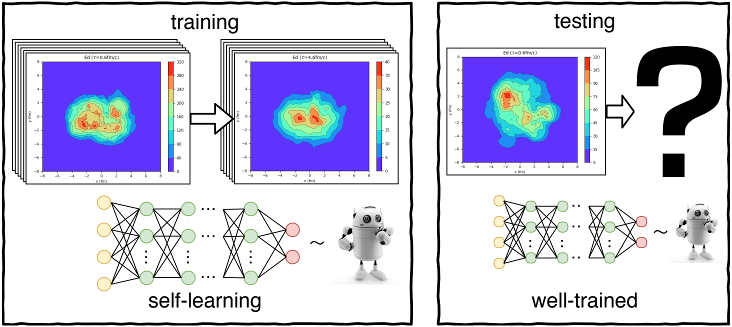
$$\partial_\mu T^{\mu\nu} = 0$$

where $T^{\mu\nu} = (e + p)u^\mu u^\nu - pg^{\mu\nu}$,e is the energy density, p is the pressuse, u^μ is the four velocity with $u^\mu u_\mu = 1$.



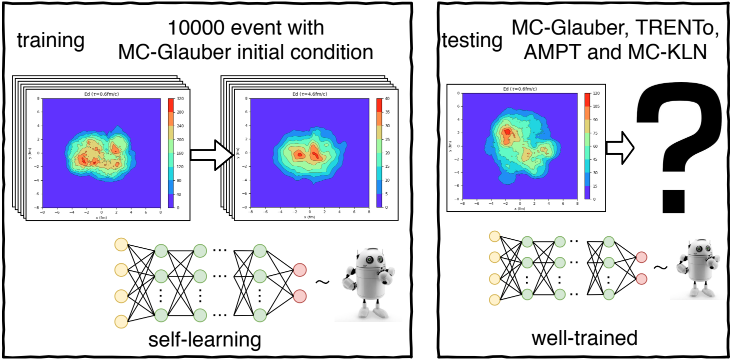
In the way of deep learning

- the evolution of hydrodynamics is processed as the pairs of initial and final profiles.



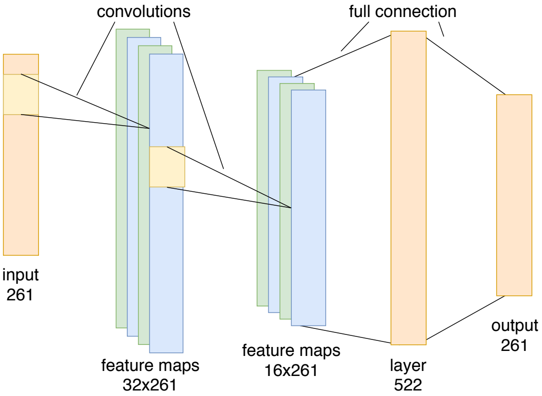
In the way of deep learning

- the evolution of hydrodynamics is processed as the pairs of initial and final profiles.



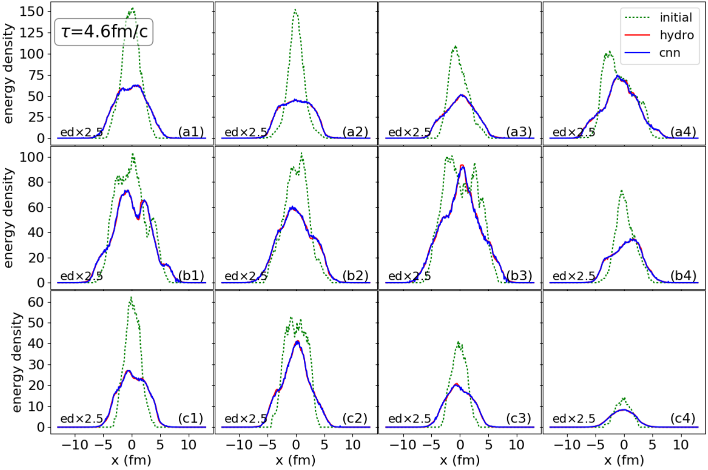
A simple case: 1+1d hydrodynamics

- By using a simple CNN, it can really works.

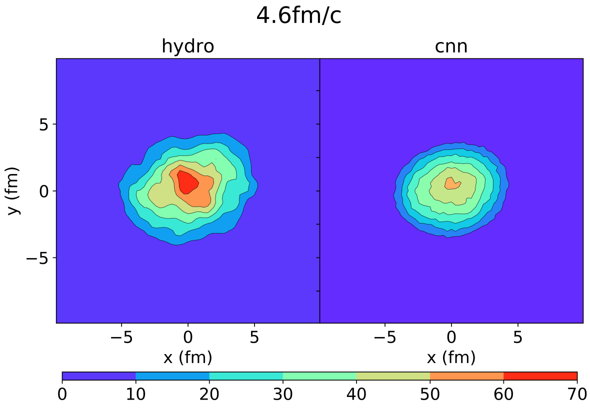


A simple case: 1+1d hydrodynamics

- It shows that deep learning can capture the nonlinear evolution of 1+1d hydrodynamics.



From 1+1d to 2+1d hydrodynamics

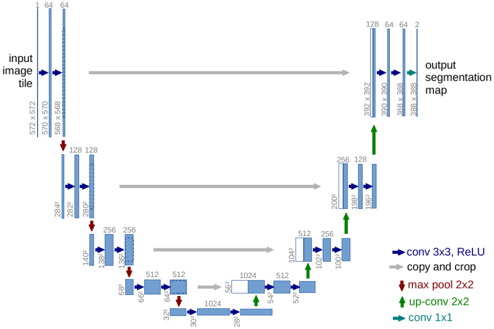


- When we turn to 2+1d hydro, CNN doesn't work, so does other network structures, such as locally connected layer.
- The main reason is the pixel of initial and final profiles is changed from 261 to 261×261.

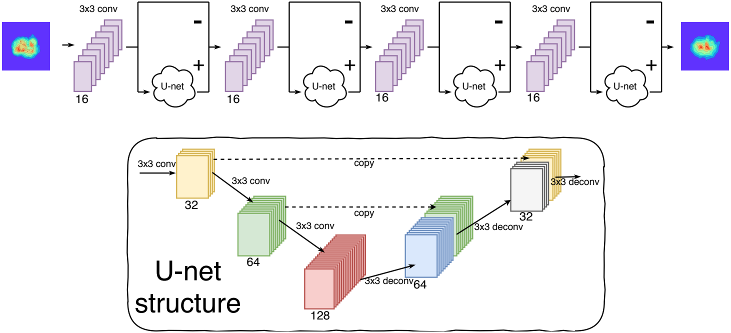
From 1+1d to 2+1d hydrodynamics

- Inspired by the biomedical image segmentation, we update our network to the stacked U-net.

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



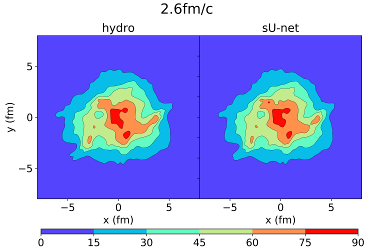
Network Structure — Stacked U-net

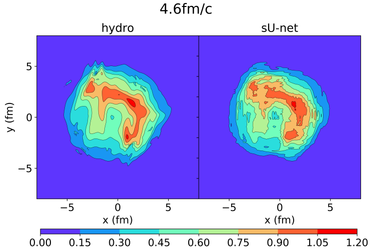


- Activation Function: LeakyRelu with $\alpha = 0.03$
- Loss Function: $L(\theta) = \frac{1}{2}\frac{\sum|\hat{y}-y|}{\sum|y|}$

2+1d hydrodynamics

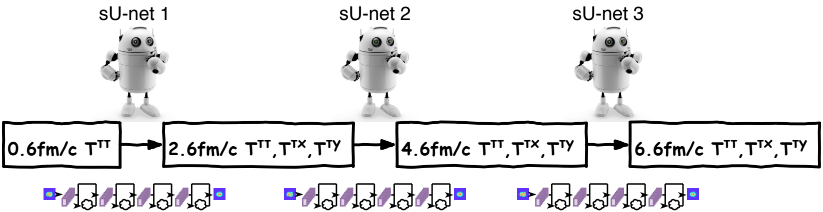
- With the help of stacked U-net, we can predict the profiles at $\tau = 2.6 fm/c$ very well, but gradually lose its precision at $\tau = 4.6 fm/c$. Probably we can improve it by increasing the number of U-net.





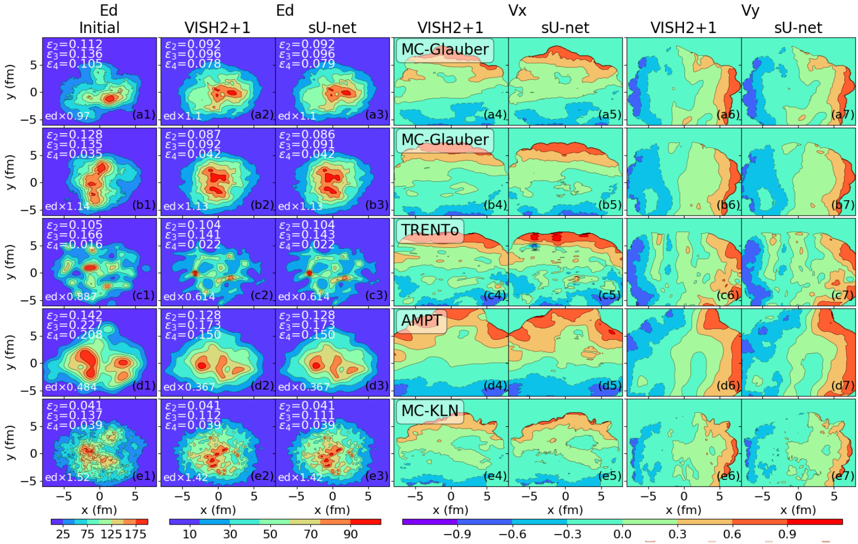
Calculation Procedure

- 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 $\Delta\tau$: $0.6 fm/c - 2.6 fm/c$, $2.6 fm/c - 4.6 fm/c$, $4.6 fm/c - 6.6 fm/c$.



Results — $\tau - \tau_0 = 2.0 fm/c$ hydro vs. sU-net

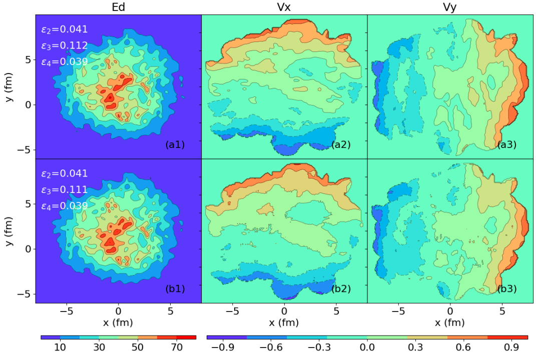
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu and Huichao Song, arXiv: 1801.03334



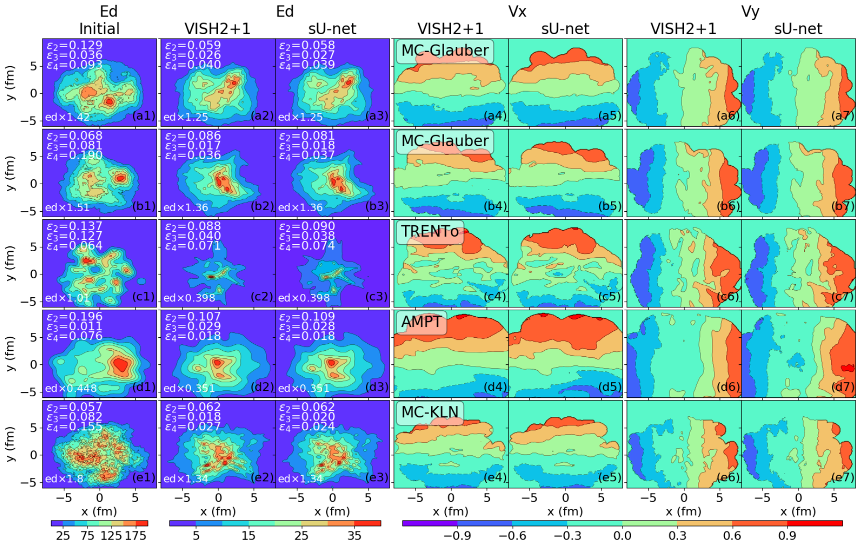
Results — $\tau - \tau_0 = 2.0 fm/c$ hydro vs. sU-net

MC-KLN initial for example:

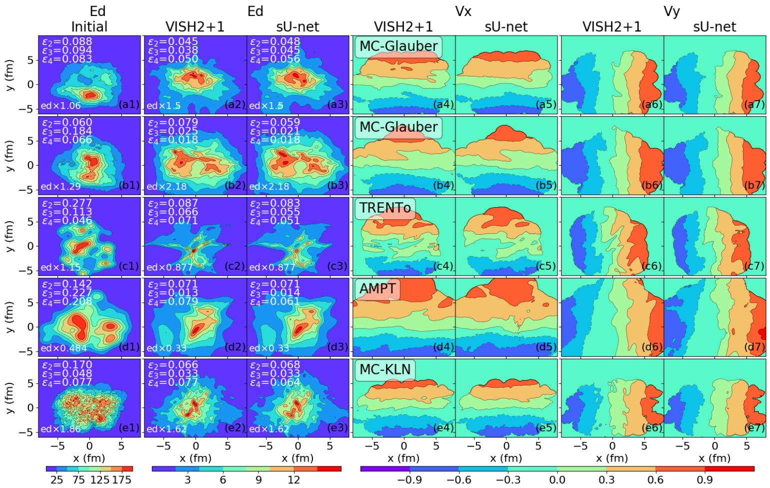
- hydro results
- network results



Results — $\tau - \tau_0 = 4.0 fm/c$ hydro vs. sU-net



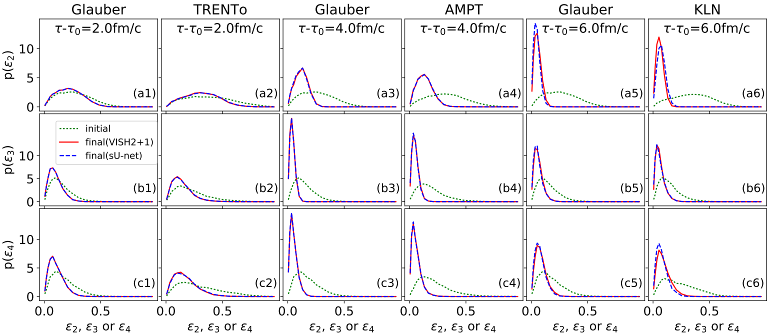
Results — $\tau - \tau_0 = 6.0 fm/c$ hydro vs. sU-net



- Obviously, we can really capture the non-linear response from initial to final profiles.

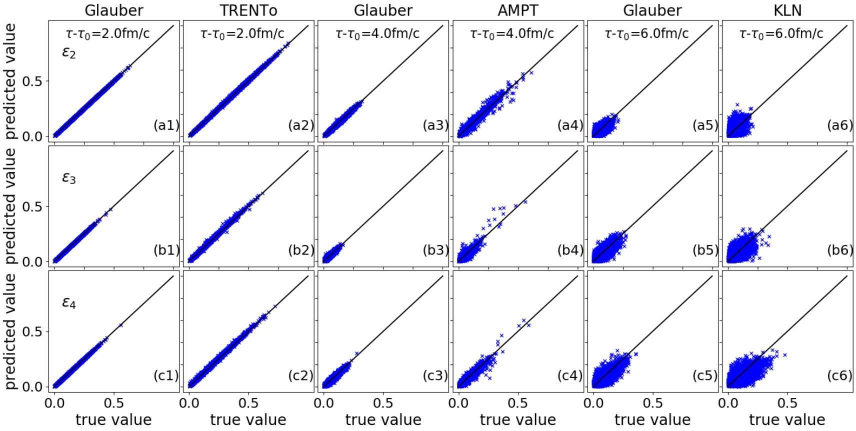
Eccentricity Distribution hydro vs. sU-net

- Besides, we calculate the distribution of harmonic eccentricity coefficients (ε_n) in the datasets at different time steps.
- $\varepsilon_n e^{in\Phi_n} = -\frac{\int dxdy r^2 e^{in\phi} e(x,y)}{\int dxdy r^2 e(x,y)}$



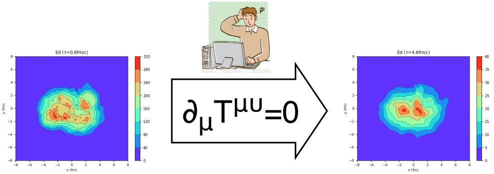
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu and Huichao Song, arXiv: 1801.03334

Histograms of ε_n hydro vs. sU-net

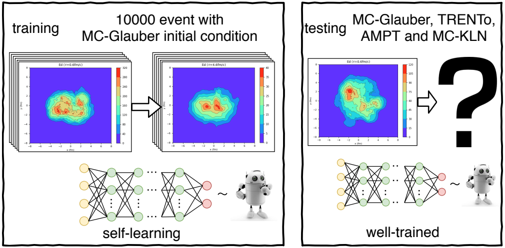


Summary and Outlook

- traditional hydrodynamics



- deep learning



Summary and Outlook

- Using 10000 initial and final profiles generated from VISH2+1 with MC-Glauber initial condition, we train the network called sU-net.
- After that, we use the well-train network to predict the final profiles associated with various initial conditions, including MC-Glauber, TRENTo, AMPT and MC-KLN. Out results shows 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 $P(\varepsilon_n)$.
- It shows that deep learning can capture the main features of the non-linear evolution of hydrodynamics.
- Deep learning is a new tool for heavy ion collision, and there will be lots of things we can do.

Thank you for your attention!