

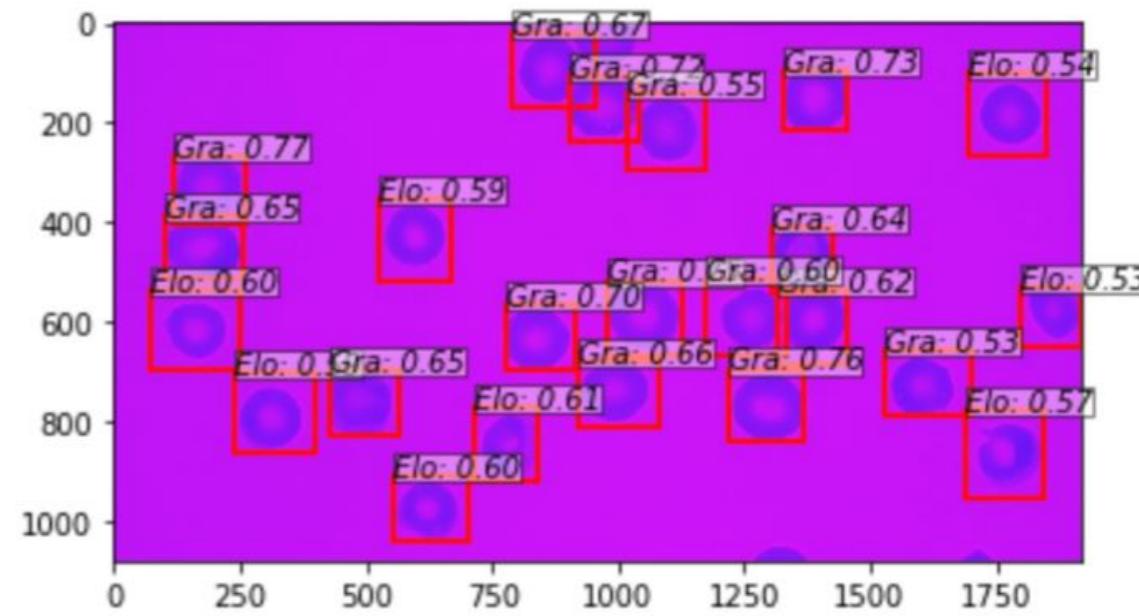
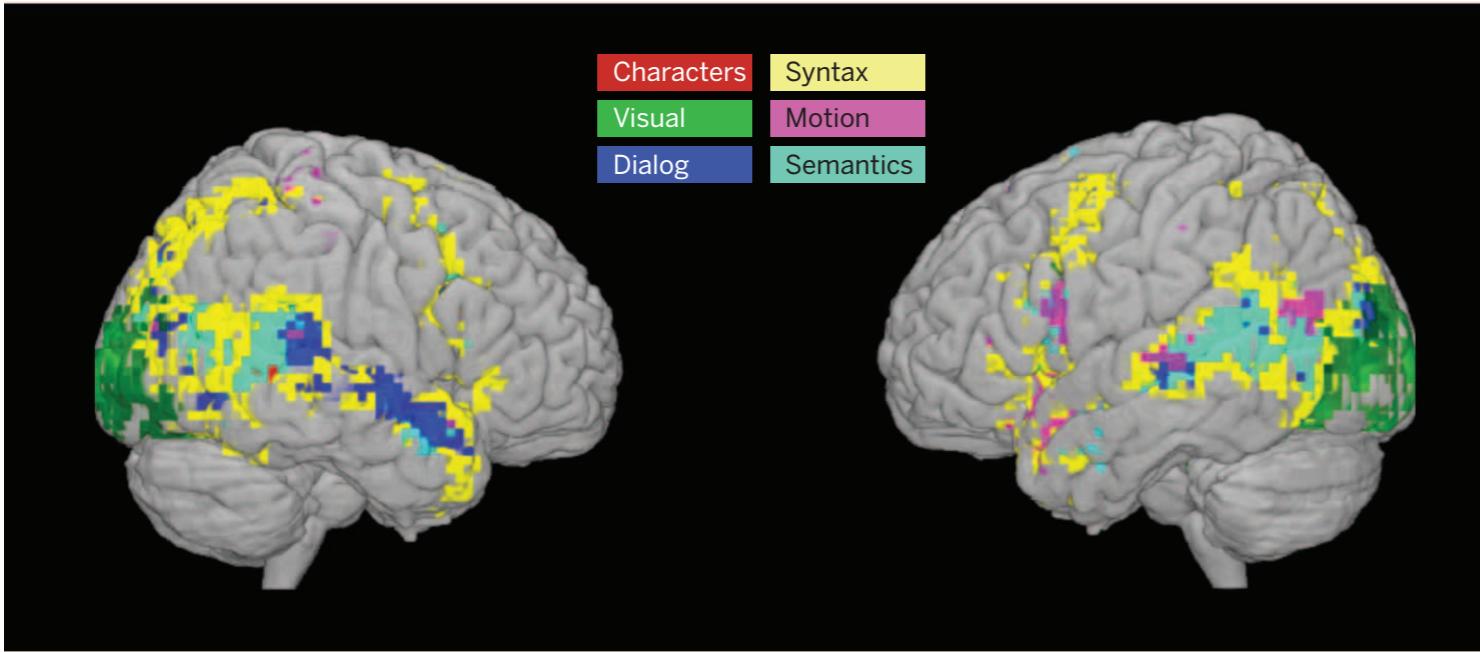
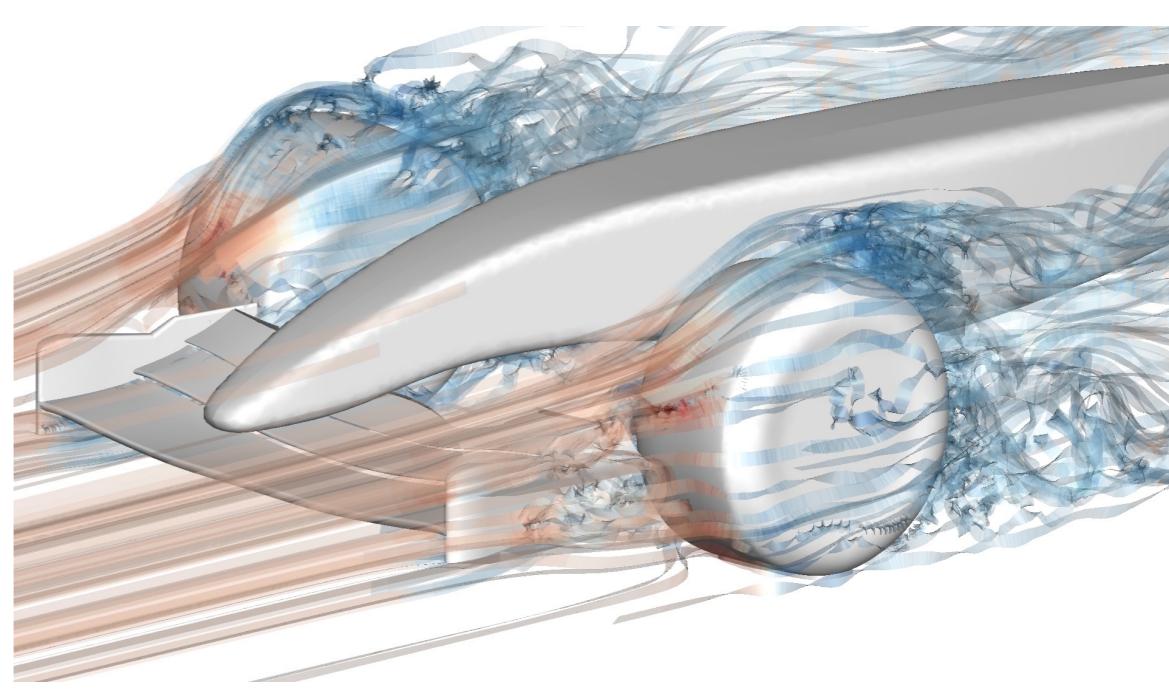
ENM 531: Introduction to Data-driven Modeling

Lecture #13: Highlights and Applications

Paris Perdikaris
March 2, 2020



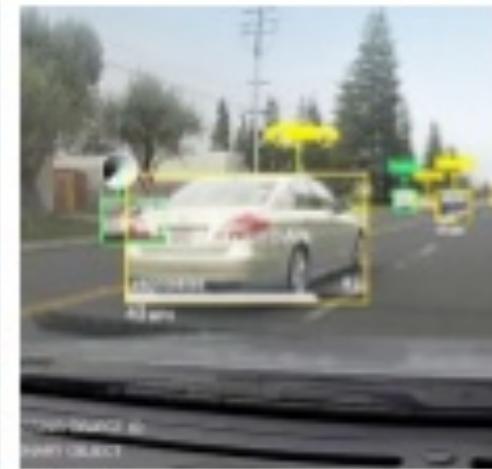
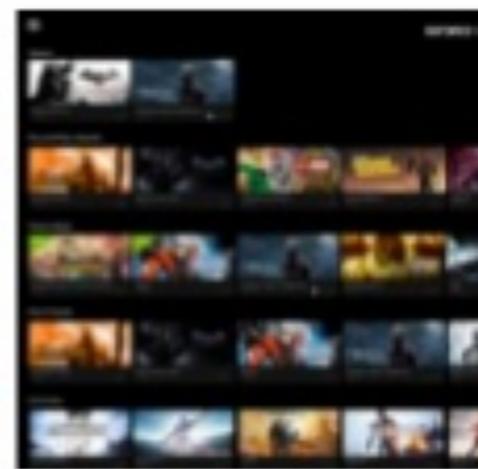
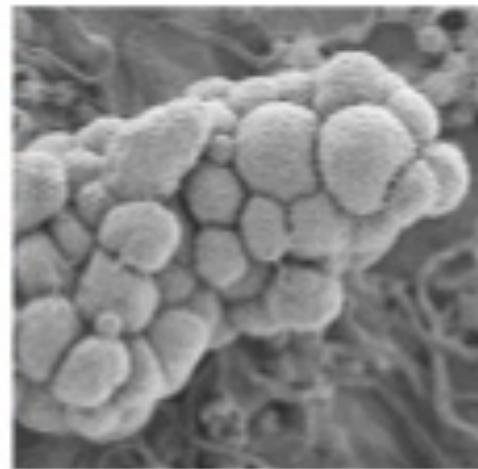
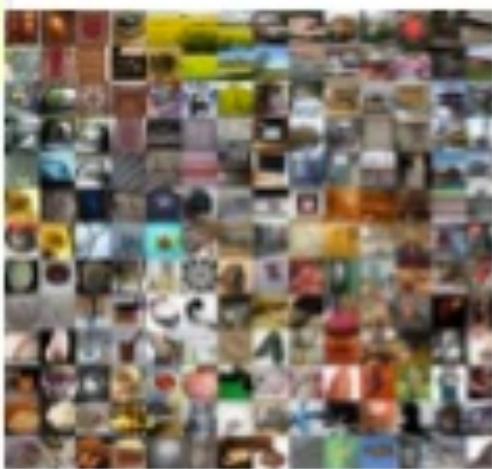
Deep learning: History and recent success



Applications of machine learning. Machine learning is having a substantial effect on many areas of technology and science; examples of recent applied success stories include robotics and autonomous vehicle control (top left), speech processing and natural language processing (top left), neuroscience research (middle), and applications in computer vision (top right).

Deep learning: History and recent success

DEEP LEARNING EVERYWHERE



INTERNET & CLOUD

- Image Classification
- Speech Recognition
- Language Translation
- Language Processing
- Sentiment Analysis
- Recommendation

MEDICINE & BIOLOGY

- Cancer Cell Detection
- Diabetic Grading
- Drug Discovery

MEDIA & ENTERTAINMENT

- Video Captioning
- Video Search
- Real Time Translation

SECURITY & DEFENSE

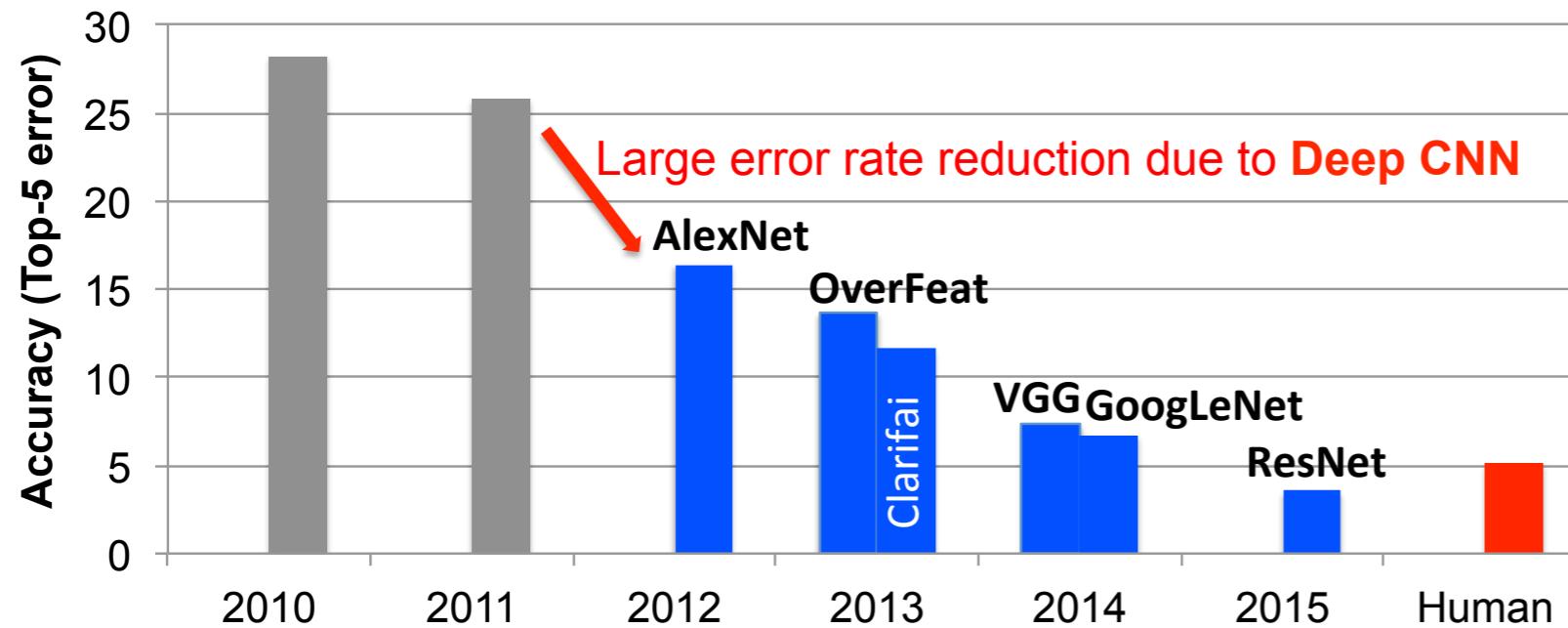
- Face Detection
- Video Surveillance
- Satellite Imagery

AUTONOMOUS MACHINES

- Pedestrian Detection
- Lane Tracking
- Recognize Traffic Sign

Deep learning: History and recent success

- 1940s - Neural networks were proposed
 - 1960s - Deep neural networks were proposed
 - 1989 - Neural networks for recognizing digits (LeNet)
 - 1990s - Hardware for shallow neural nets (Intel ETANN)
 - 2011 - Breakthrough DNN-based speech recognition (Microsoft)
 - 2012 - DNNs for vision start supplanting hand-crafted approaches (AlexNet)
 - 2014+ - Rise of DNN accelerator research (Neuflow, DianNao...)



In 2015, the ImageNet winning entry, ResNet, exceeded human-level accuracy with a top-5 error rate below 5%.

Since then, the error rate has dropped below 3% and more focus is now being placed on more challenging components of the competition, such as object detection and localization.

Image super-resolution

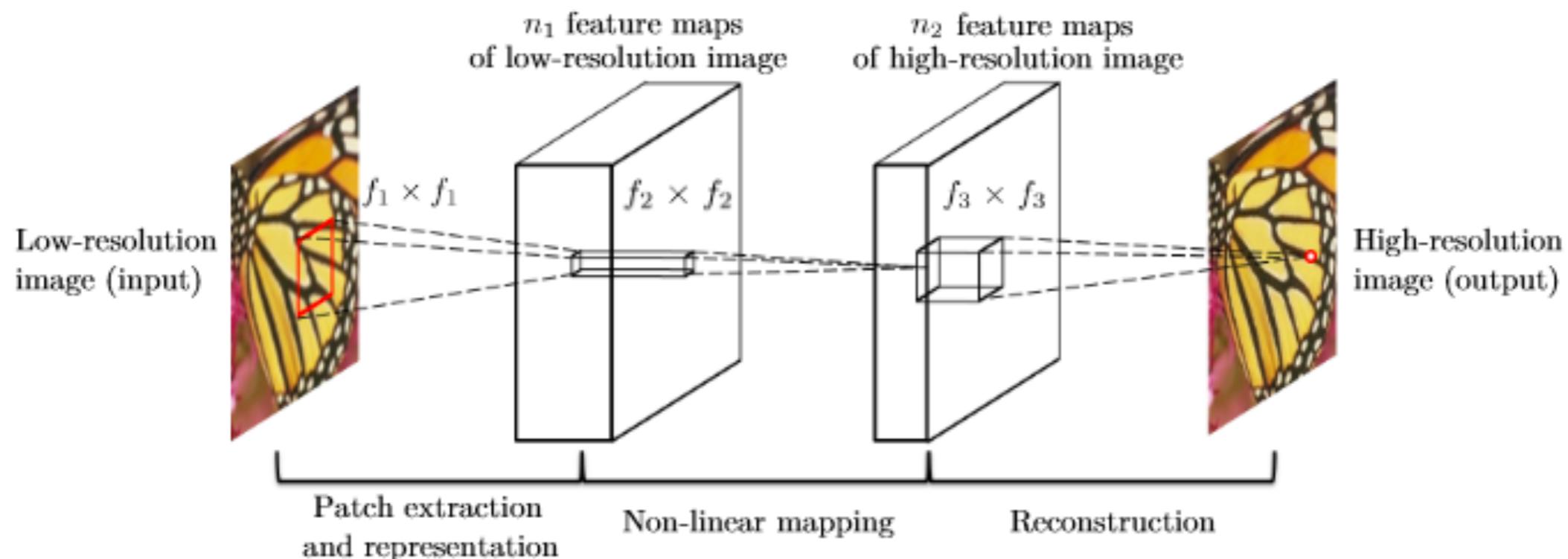
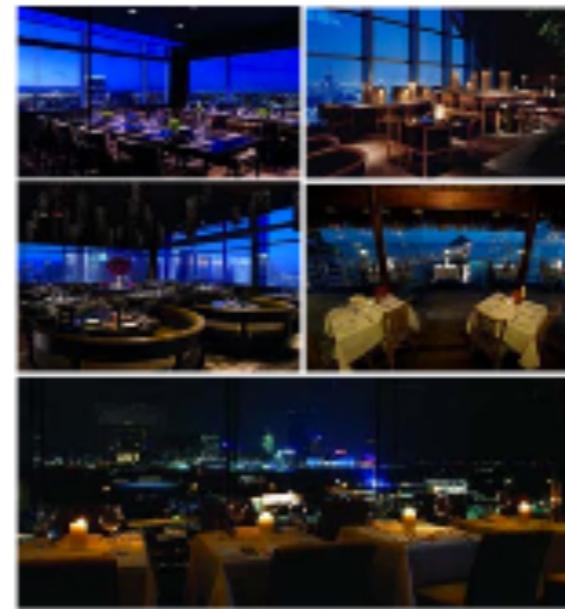


Fig. 2. Given a low-resolution image \mathbf{Y} , the first convolutional layer of the SRCNN extracts a set of feature maps. The second layer maps these feature maps nonlinearly to high-resolution patch representations. The last layer combines the predictions within a spatial neighbourhood to produce the final high-resolution image $F(\mathbf{Y})$.

Transfer learning



Source



Refs ("restaurant night")



Ours



Source



Refs ("building beach")



Ours

Modeling long-term dependencies in sequence data

For $\bigoplus_{n=1,\dots,m} \mathcal{L}_{m,n} = 0$, hence we can find a closed subset \mathcal{H} in \mathcal{H} and any sets \mathcal{F} on X , U is a closed immersion of S , then $U \rightarrow T$ is a separated algebraic space.

Proof. Proof of (1). It also start we get

$$S = \text{Spec}(R) = U \times_X U \times_X U$$

and the comparicoly in the fibre product covering we have to prove the lemma generated by $\coprod Z \times_U U \rightarrow V$. Consider the maps M along the set of points Sch_{fppf} and $U \rightarrow U$ is the fibre category of S in U in Section, ?? and the fact that any U affine, see Morphisms, Lemma ???. Hence we obtain a scheme S and any open subset $W \subset U$ in $\text{Sh}(G)$ such that $\text{Spec}(R') \rightarrow S$ is smooth or an

$$U = \bigcup U_i \times_{S_i} U_i$$

which has a nonzero morphism we may assume that f_i is of finite presentation over S . We claim that $\mathcal{O}_{X,x}$ is a scheme where $x, x', s'' \in S'$ such that $\mathcal{O}_{X,x'} \rightarrow \mathcal{O}'_{X',x'}$ is separated. By Algebra, Lemma ?? we can define a map of complexes $\text{GL}_{S'}(x'/S'')$ and we win. \square

To prove study we see that $\mathcal{F}|_U$ is a covering of \mathcal{X}' , and \mathcal{T}_i is an object of $\mathcal{F}_{X/S}$ for $i > 0$ and \mathcal{F}_p exists and let \mathcal{F}_i be a presheaf of \mathcal{O}_X -modules on \mathcal{C} as a \mathcal{F} -module. In particular $\mathcal{F} = U/\mathcal{F}$ we have to show that

$$\widetilde{M}^\bullet = \mathcal{I}^\bullet \otimes_{\text{Spec}(k)} \mathcal{O}_{S,s} - i_X^{-1} \mathcal{F}$$

is a unique morphism of algebraic stacks. Note that

$$\text{Arrows} = (\text{Sch}/S)^{\text{opp}}_{fppf}, (\text{Sch}/S)_{fppf}$$

and

$$V = \Gamma(S, \mathcal{O}) \longrightarrow (U, \text{Spec}(A))$$

is an open subset of X . Thus U is affine. This is a continuous map of X is the inverse, the groupoid scheme S .

Proof. See discussion of sheaves of sets. \square

The result for prove any open covering follows from the less of Example ???. It may replace S by $X_{\text{spaces},\text{étale}}$ which gives an open subspace of X and T equal to S_{Zar} , see Descent, Lemma ???. Namely, by Lemma ?? we see that R is geometrically regular over S .

Lemma 0.1. Assume (3) and (3) by the construction in the description.

Suppose $X = \lim |X|$ (by the formal open covering X and a single map $\underline{\text{Proj}}_X(\mathcal{A}) = \text{Spec}(B)$ over U compatible with the complex

$$\text{Set}(\mathcal{A}) = \Gamma(X, \mathcal{O}_{X,\mathcal{O}_X}).$$

When in this case of to show that $\mathcal{Q} \rightarrow \mathcal{C}_{Z/X}$ is stable under the following result in the second conditions of (1), and (3). This finishes the proof. By Definition ?? (without element is when the closed subschemes are catenary. If T is surjective we may assume that T is connected with residue fields of S . Moreover there exists a closed subspace $Z \subset X$ of X where U in X' is proper (some defining as a closed subset of the uniqueness it suffices to check the fact that the following theorem

(1) f is locally of finite type. Since $S = \text{Spec}(R)$ and $Y = \text{Spec}(R)$.

Proof. This is form all sheaves of sheaves on X . But given a scheme U and a surjective étale morphism $U \rightarrow X$. Let $U \cap U = \coprod_{i=1,\dots,n} U_i$ be the scheme X over S at the schemes $X_i \rightarrow X$ and $U = \lim_i X_i$. \square

The following lemma surjective restrocomposes of this implies that $\mathcal{F}_{x_0} = \mathcal{F}_{x_0} = \mathcal{F}_{x,\dots,x}$.

Lemma 0.2. Let X be a locally Noetherian scheme over S , $E = \mathcal{F}_{X/S}$. Set $\mathcal{I} = \mathcal{J}_1 \subset \mathcal{I}'_n$. Since $\mathcal{I}^n \subset \mathcal{I}^n$ are nonzero over $i_0 \leq p$ is a subset of $\mathcal{J}_{n,0} \circ \overline{A}_2$ works.

Lemma 0.3. In Situation ???. Hence we may assume $q' = 0$.

Proof. We will use the property we see that p is the next functor (??). On the other hand, by Lemma ?? we see that

$$D(\mathcal{O}_{X'}) = \mathcal{O}_X(D)$$

where K is an F -algebra where δ_{n+1} is a scheme over S . \square

Example: Text generation with LSTMs

Modeling long-term dependencies in sequence data

Proof. Omitted. \square

Lemma 0.1. Let \mathcal{C} be a set of the construction.

Let \mathcal{C} be a gerber covering. Let \mathcal{F} be a quasi-coherent sheaves of \mathcal{O} -modules. We have to show that

$$\mathcal{O}_{\mathcal{O}_X} = \mathcal{O}_X(\mathcal{L})$$

Proof. This is an algebraic space with the composition of sheaves \mathcal{F} on $X_{\text{étale}}$ we have

$$\mathcal{O}_X(\mathcal{F}) = \{\text{morph}_1 \times_{\mathcal{O}_X} (\mathcal{G}, \mathcal{F})\}$$

where \mathcal{G} defines an isomorphism $\mathcal{F} \rightarrow \mathcal{F}$ of \mathcal{O} -modules. \square

Lemma 0.2. This is an integer \mathcal{Z} is injective.

Proof. See Spaces, Lemma ??.

Lemma 0.3. Let S be a scheme. Let X be a scheme and X is an affine open covering. Let $\mathcal{U} \subset \mathcal{X}$ be a canonical and locally of finite type. Let X be a scheme. Let X be a scheme which is equal to the formal complex.

The following to the construction of the lemma follows.

Let X be a scheme. Let X be a scheme covering. Let

$$b : X \rightarrow Y' \rightarrow Y \rightarrow Y' \times_X Y \rightarrow X.$$

be a morphism of algebraic spaces over S and Y .

Proof. Let X be a nonzero scheme of X . Let X be an algebraic space. Let \mathcal{F} be a quasi-coherent sheaf of \mathcal{O}_X -modules. The following are equivalent

- (1) \mathcal{F} is an algebraic space over S .
- (2) If X is an affine open covering.

Consider a common structure on X and X the functor $\mathcal{O}_X(U)$ which is locally of finite type. \square

This since $\mathcal{F} \in \mathcal{F}$ and $x \in \mathcal{G}$ the diagram

$$\begin{array}{ccccc}
 S & \xrightarrow{\quad} & & & \\
 \downarrow & & & & \\
 \xi & \xrightarrow{\quad} & \mathcal{O}_{X'} & & \\
 \text{gor}_s & & \uparrow & \searrow & \\
 & & = \alpha' & \longrightarrow & \\
 & & \downarrow & & \\
 & & = \alpha' & \longrightarrow & \alpha \\
 & & & & \\
 \text{Spec}(K_\psi) & & \text{Mor}_{\text{Sets}} & & d(\mathcal{O}_{X_{/\kappa}}, \mathcal{G}) \\
 & & & & \downarrow \\
 & & & & X
 \end{array}$$

is a limit. Then \mathcal{G} is a finite type and assume S is a flat and \mathcal{F} and \mathcal{G} is a finite type f_* . This is of finite type diagrams, and

- the composition of \mathcal{G} is a regular sequence,
- $\mathcal{O}_{X'}$ is a sheaf of rings.

Proof. We have see that $X = \text{Spec}(R)$ and \mathcal{F} is a finite type representable by algebraic space. The property \mathcal{F} is a finite morphism of algebraic stacks. Then the cohomology of X is an open neighbourhood of U . \square

Proof. This is clear that \mathcal{G} is a finite presentation, see Lemmas ??.

A reduced above we conclude that U is an open covering of \mathcal{C} . The functor \mathcal{F} is a “field”

$$\mathcal{O}_{X,x} \rightarrow \mathcal{F}_x \dashv (\mathcal{O}_{X_{\text{étale}}}) \rightarrow \mathcal{O}_{X_\ell}^{-1} \mathcal{O}_{X_\lambda}(\mathcal{O}_{X_\eta}^\vee)$$

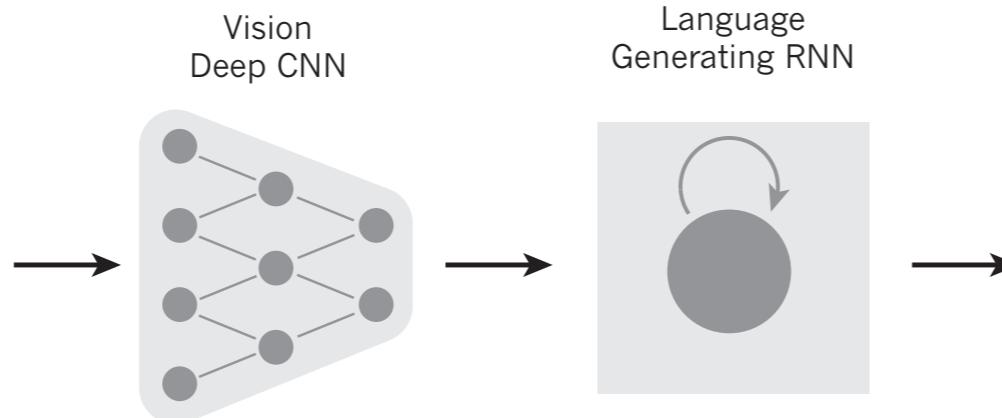
is an isomorphism of covering of \mathcal{O}_{X_i} . If \mathcal{F} is the unique element of \mathcal{F} such that X is an isomorphism.

The property \mathcal{F} is a disjoint union of Proposition ?? and we can filtered set of presentations of a scheme \mathcal{O}_X -algebra with \mathcal{F} are opens of finite type over S . If \mathcal{F} is a scheme theoretic image points. \square

If \mathcal{F} is a finite direct sum \mathcal{O}_{X_λ} is a closed immersion, see Lemma ???. This is a sequence of \mathcal{F} is a similar morphism.

Example: Text generation with LSTMs

Attention and captioning



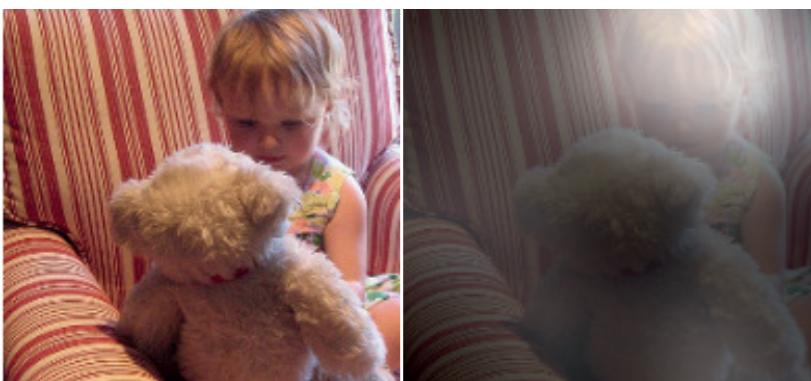
A woman is throwing a **frisbee** in a park.



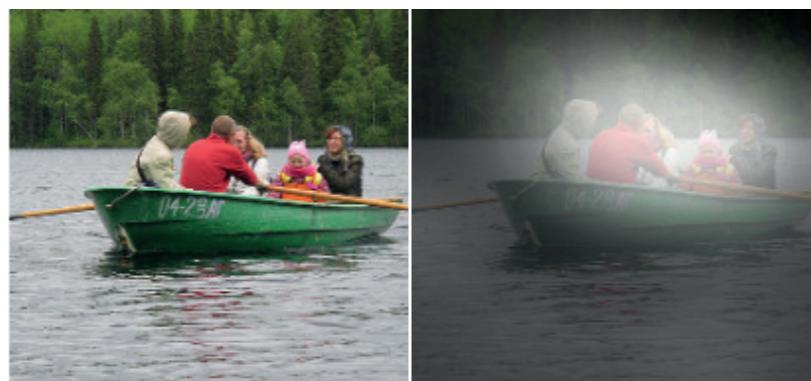
A **dog** is standing on a hardwood floor.



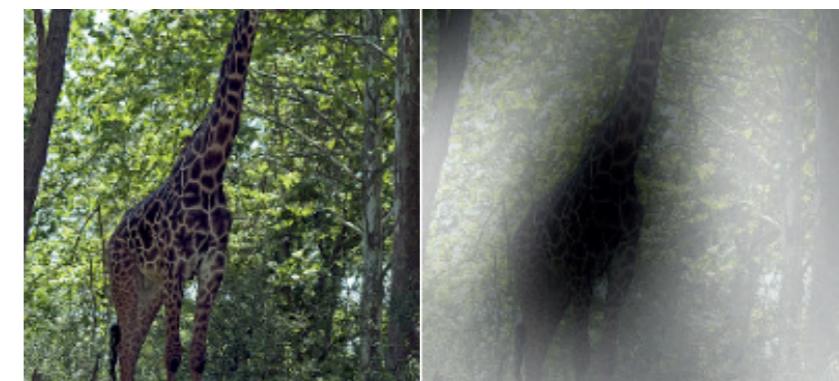
A **stop** sign is on a road with a mountain in the background



A little **girl** sitting on a bed with a teddy bear.



A group of **people** sitting on a boat in the water.



A giraffe standing in a forest with **trees** in the background.

Detecting Depression with Audio/Text Sequence Modeling

Abstract

Medical professionals diagnose depression by interpreting the responses of individuals to a variety of questions, probing lifestyle changes and ongoing thoughts. Like professionals, an effective automated agent must understand that responses to queries have varying prognostic value. In this study we demonstrate an automated depression-detection algorithm that models interviews between an individual and agent and learns from sequences of questions and answers without the need to perform explicit topic modeling of the content. We utilized data of 142 individuals undergoing depression screening, and modeled the interactions with audio and text features in a Long-Short Term Memory (LSTM) neural network model to detect depression. Our results were comparable to methods that explicitly modeled the topics of the questions and answers which suggests that depression can be detected through sequential modeling of an interaction, with minimal information on the structure of the interview.

Index Terms: medical speech signal processing, depression, neural networks, computational paralinguistics, question answering

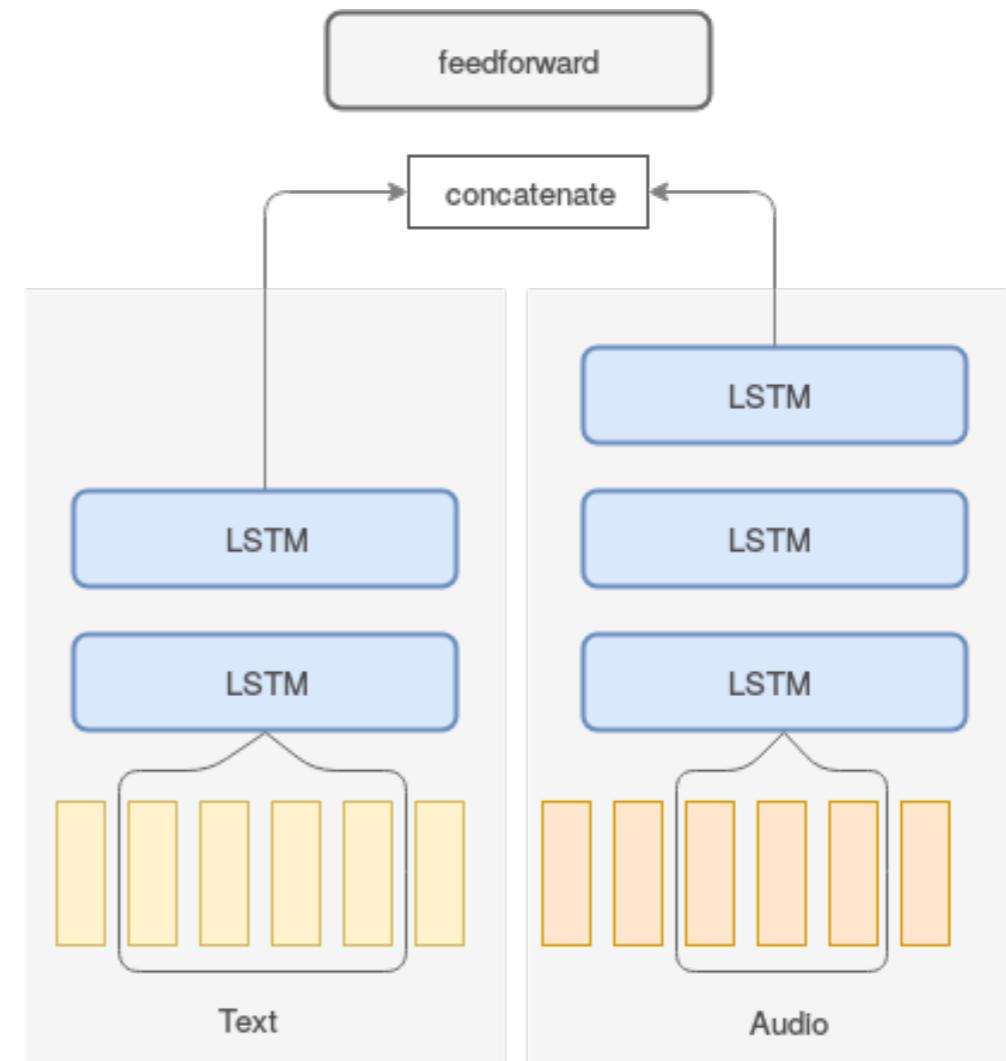
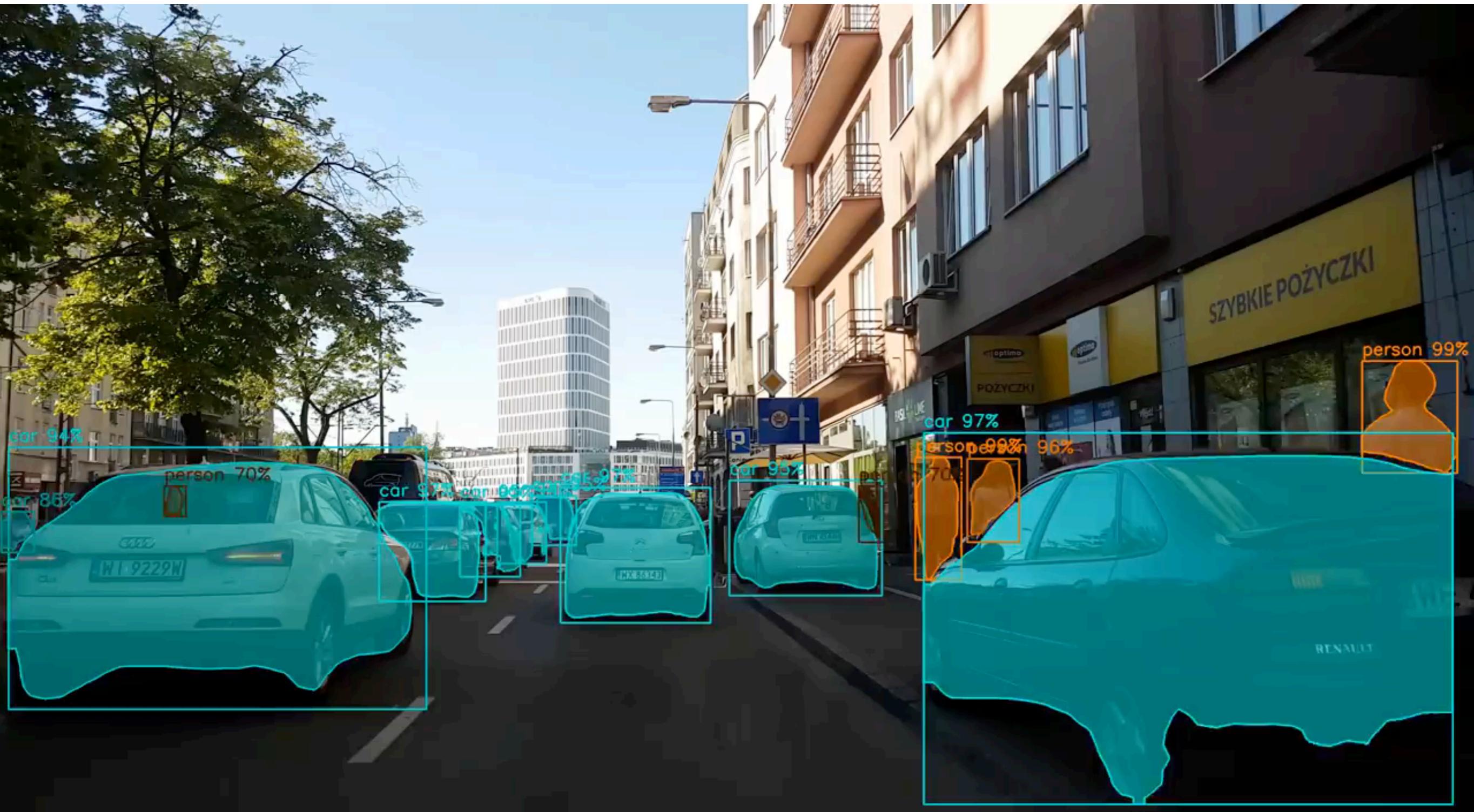


Figure 1: **Diagram of network topology.** Each modality (audio and text) were trained separately as bi-directional LSTMs with differing hyperparameters capturing the characteristics of each feature set. A multi-modal model that combined both audio and text was also trained through concatenation into feedforward layers.

Object segmentation and tracking



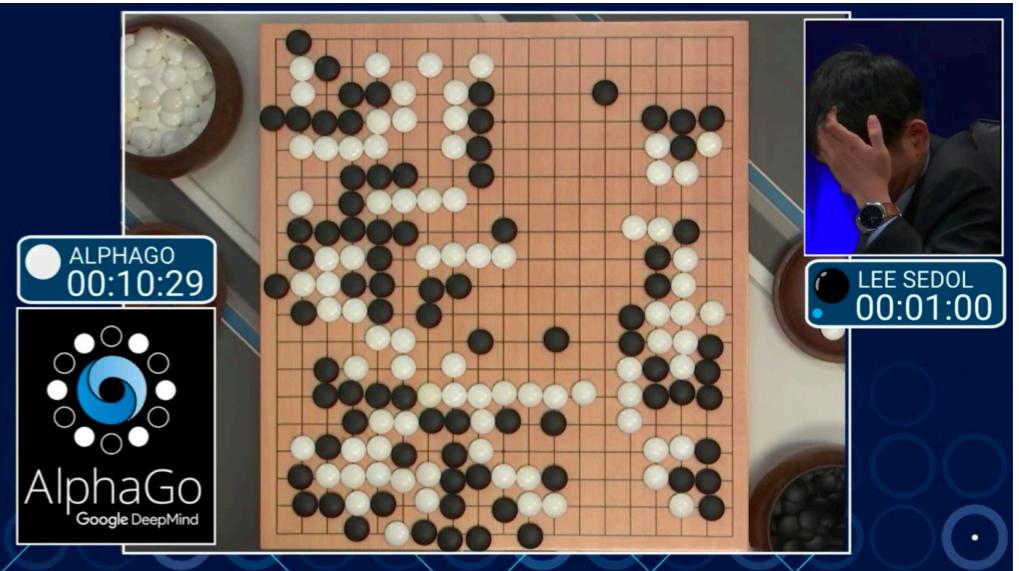
He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017, October). Mask r-cnn. 2017 IEEE International Conference on Computer Vision (ICCV).

Mastering games

Google DeepMind's Deep Q-learning

The algorithm will play Atari breakout.

The most important thing to know is that all the agent is given is sensory input (what you see on the screen) and it was ordered to maximize the score on the screen.



Emergent Tool Use from Multi-Agent Interaction



Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., & Mordatch, I. (2019). *Emergent tool use from multi-agent autocurricula*. arXiv preprint arXiv:1909.07528.

Forecasting chaotic systems

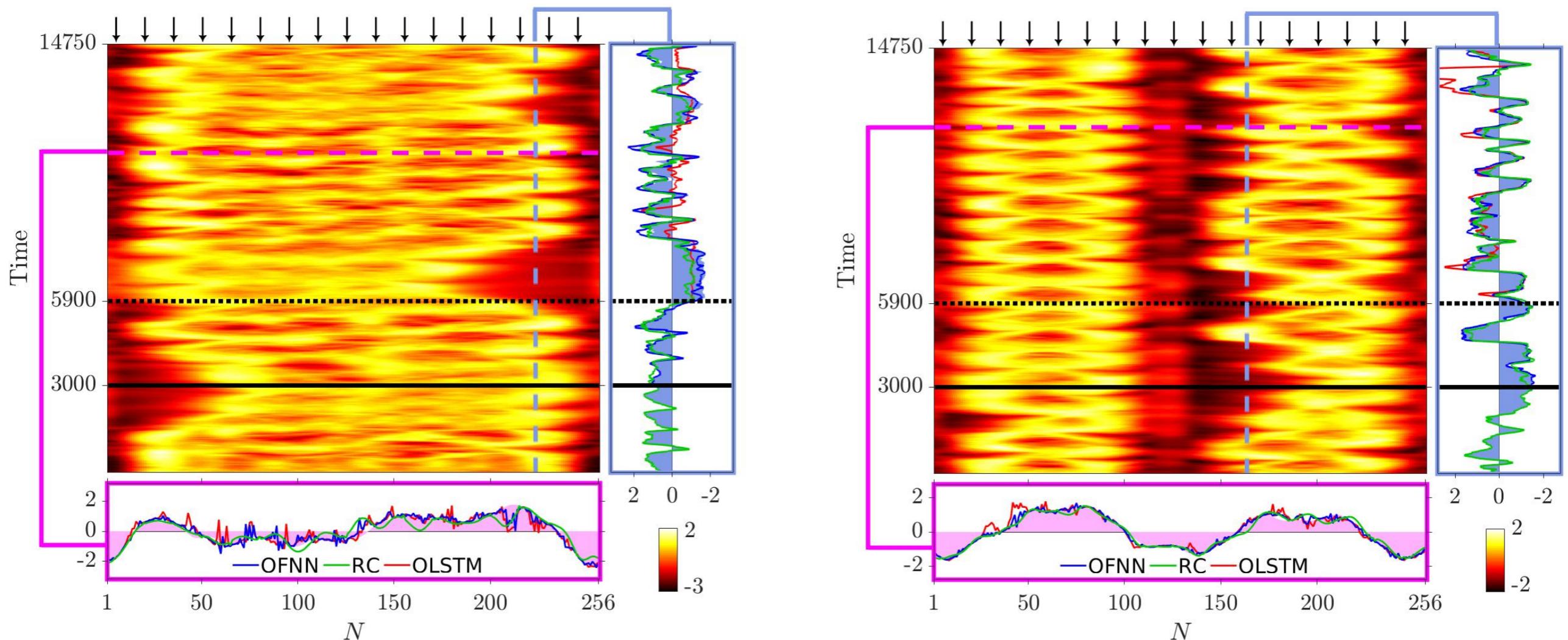


Figure 1. Spatiotemporal plots of single- and double-headed chimeras and predicted time series and fluxes. (A) Spatiotemporal plot of a single-headed chimera state, generated in the 1- dimensional SQUID array of Ref (19) Fig. 4e, depicting the evolution of the fluxes (values are color-coded), for the one-headed chimera observed in a large array of 256 coupled SQUIDs by dipole-dipole moments, which has been studied numerically for a nonlocal coupling scheme. In predicting the spatiotemporal evolution of this chimera, 17 “observers” have been placed in the positions marked by the tips of the arrows.

Neofotistos, George, Marios Mattheakis, Georgios D. Bamparis, Johanne Hizanidis, Giorgos Panagiotis Tsironis, and Efthimios Kaxiras. "Machine learning with observers predicts complex spatiotemporal behavior." *arXiv preprint arXiv: 1807.10758* (2018).

Exascale Deep Learning for Climate Analytics

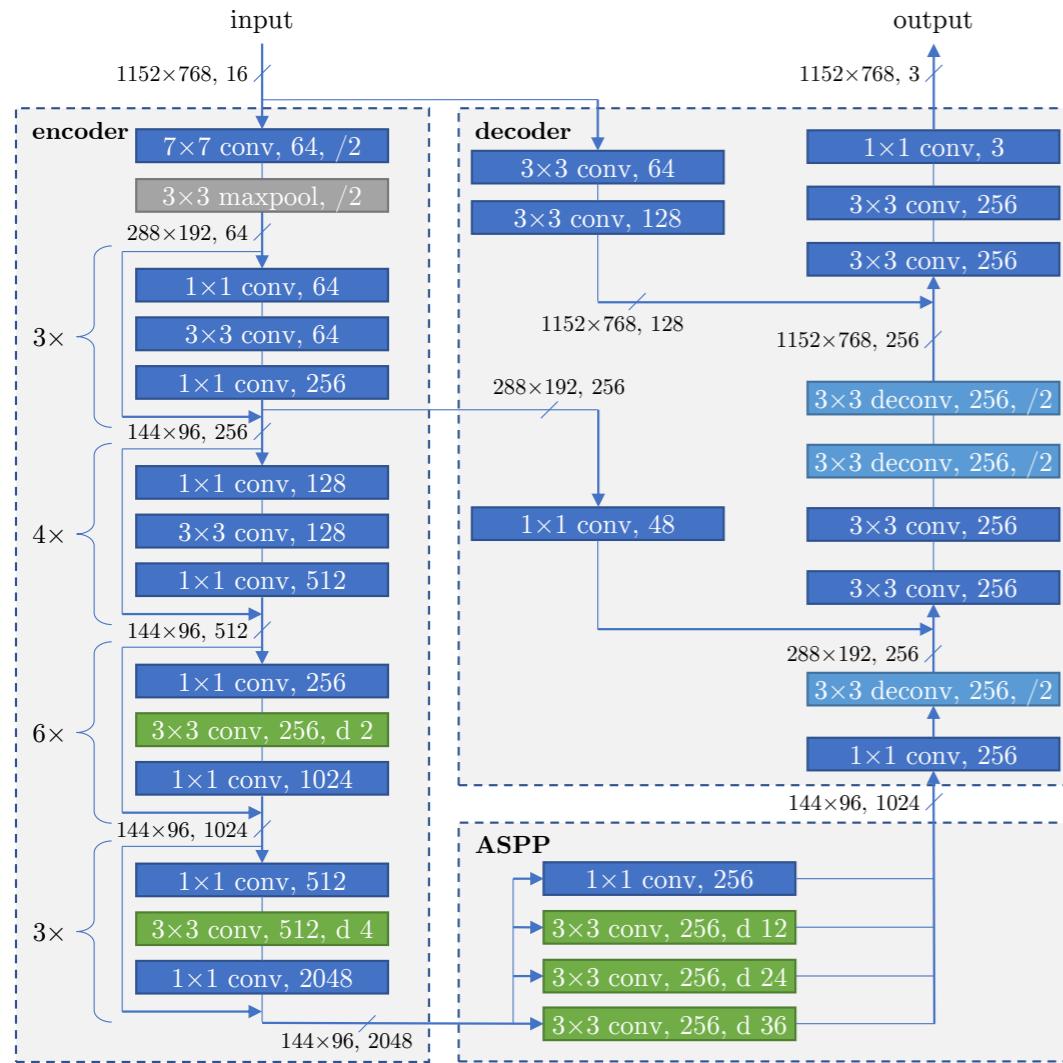
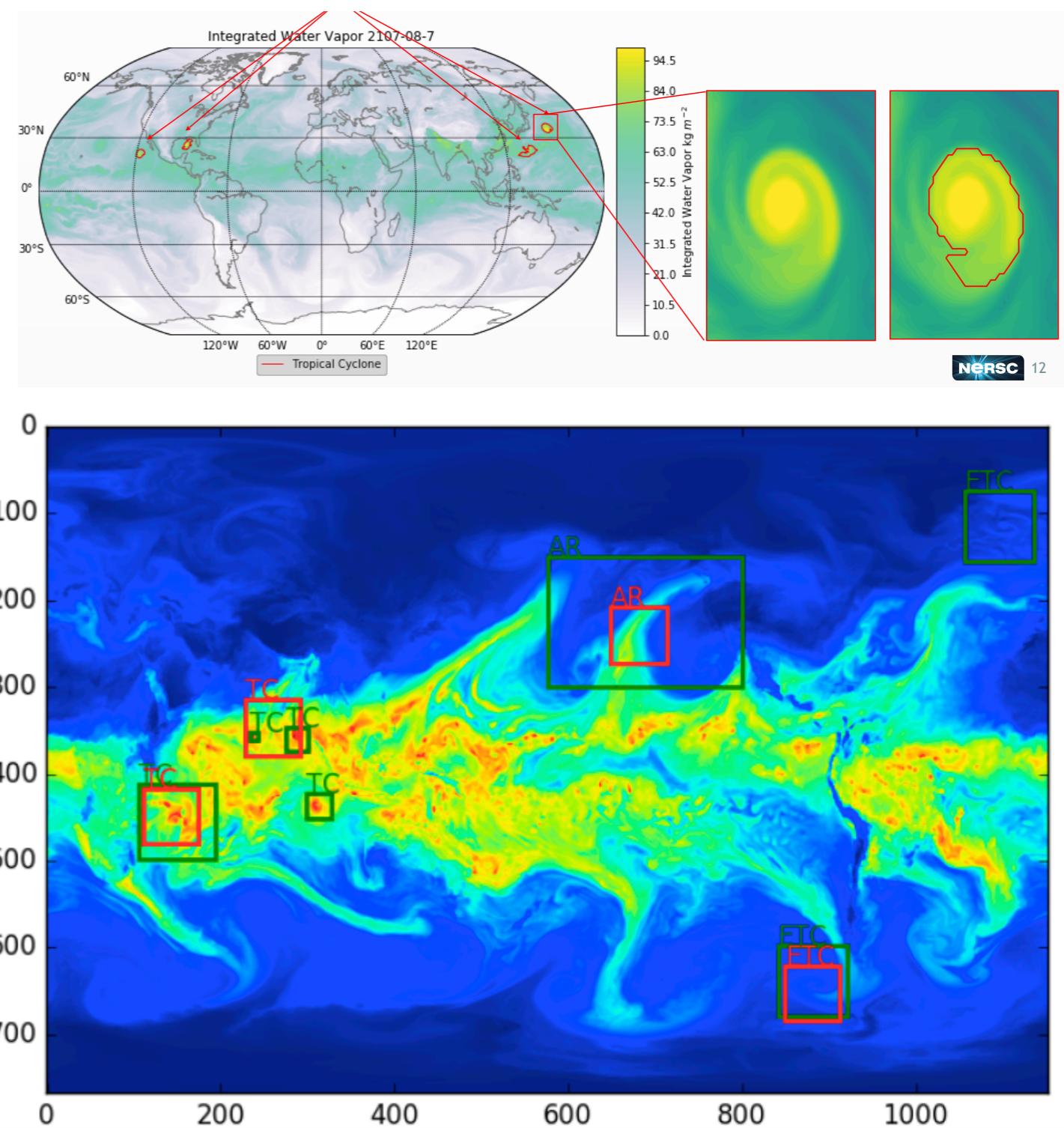
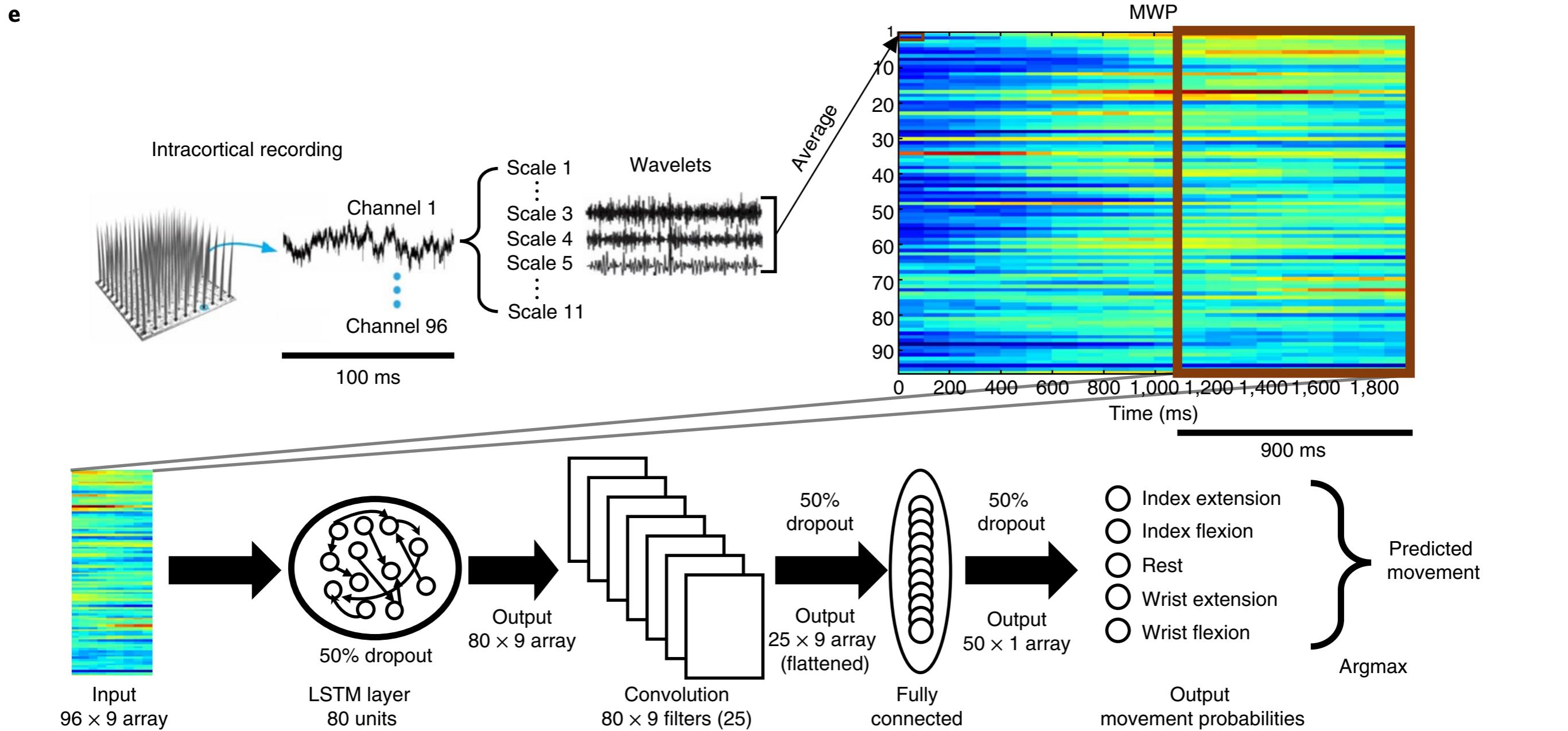


Fig. 1: Schematic of the modified DeepLabv3+ network used in this work. The encoder (which uses a ResNet-50 core) and atrous spatial pyramid pooling (ASPP) blocks are changed for the larger input resolution. The DeepLabv3+ decoder has been replaced with one that operates at full resolution to produce precise segmentation boundaries. Standard convolutions are in dark blue, and deconvolutional layers are light blue. Atrous convolution layers are in green and specify the *dilation* parameter used.



Brain-computer interfaces



How Neural Networks Can Read Thoughts and Restore Movement to Paralyzed Limbs

Schwemmer, Michael A., Nicholas D. Skomrock, Per B. Sederberg, Jordyn E. Ting, Gaurav Sharma, Marcia A. Bockbrader, and David A. Friedenberg. "Meeting brain–computer interface user performance expectations using a deep neural network decoding framework." *Nature medicine* 24, no. 11 (2018): 1669.

Systems identification (ODEs)

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t))$$

E.g. glycolytic oscillator model:

$$\frac{dS_1}{dt} = J_0 - \frac{k_1 S_1 S_6}{1 + (S_6/K_1)^q},$$

$$\frac{dS_2}{dt} = 2 \frac{k_1 S_1 S_6}{1 + (S_6/K_1)^q} - k_2 S_2(N - S_5) - k_6 S_2 S_5,$$

$$\frac{dS_3}{dt} = k_2 S_2(N - S_5) - k_3 S_3(N - S_6),$$

$$\frac{dS_4}{dt} = k_3 S_3(A - S_6) - k_4 S_4 S_5 - \kappa(S_4 - S_7),$$

$$\frac{dS_5}{dt} = k_2 S_2(N - S_5) - k_4 S_4 S_5 - k_6 S_2 S_5,$$

$$\frac{dS_6}{dt} = -2 \frac{k_1 S_1 S_6}{1 + (S_6/K_1)^q} + 2k_3 S_3(A - S_6) - k_5 S_6,$$

$$\frac{dS_7}{dt} = \psi \kappa(S_4 - S_7) - k S_7.$$

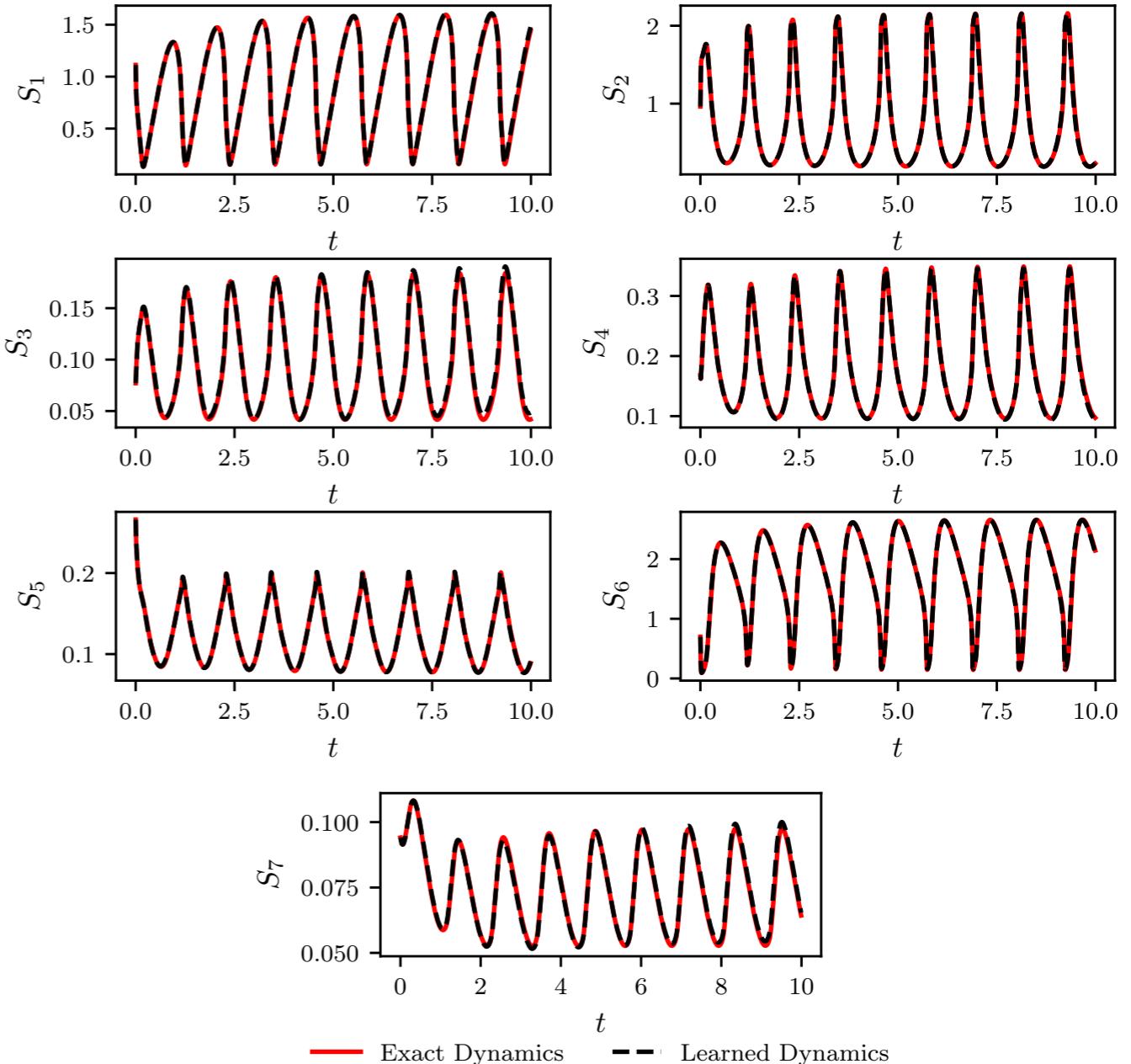


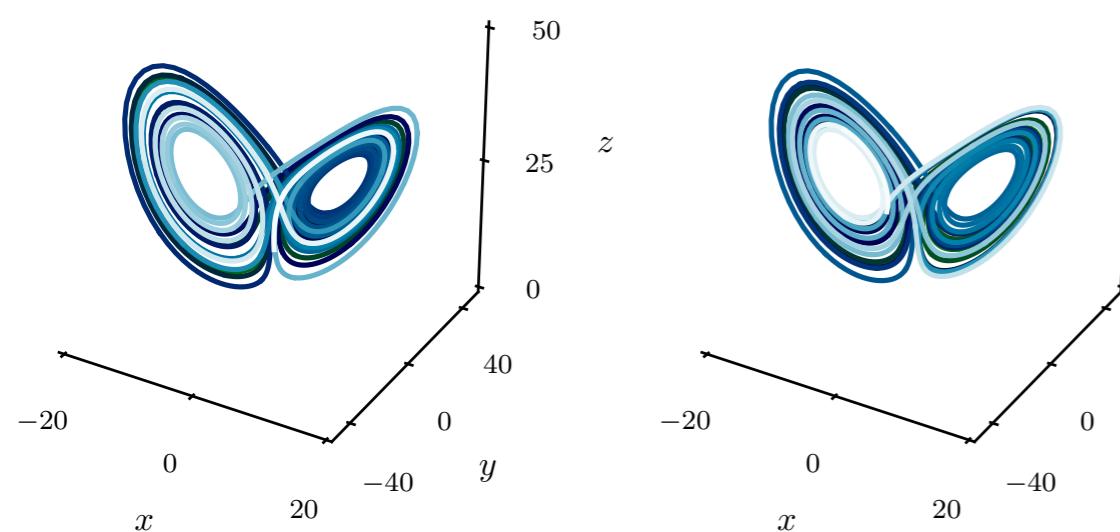
Figure 7: *Glycolytic oscillator*: Exact versus learned dynamics for random initial conditions chosen from the ranges provided in table 2 of [29].

Systems identification (ODEs)

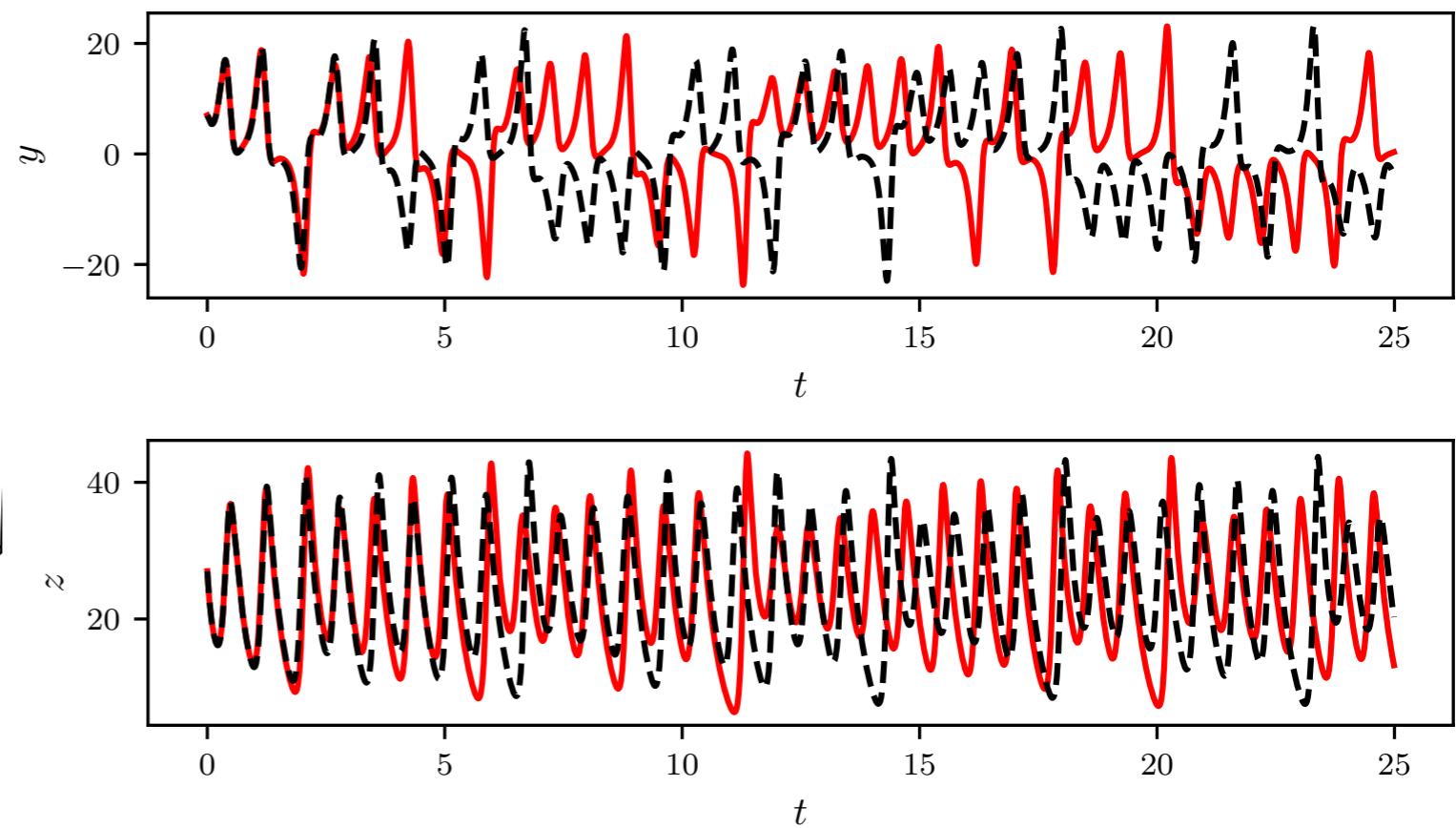
Lorenz System:

$$\begin{aligned}\dot{x} &= 10(y - x), \\ \dot{y} &= x(28 - z) - y, \\ \dot{z} &= xy - (8/3)z.\end{aligned}$$

Exact Dynamics



Learned Dynamics



— Exact Dynamics - - - Learned Dynamics

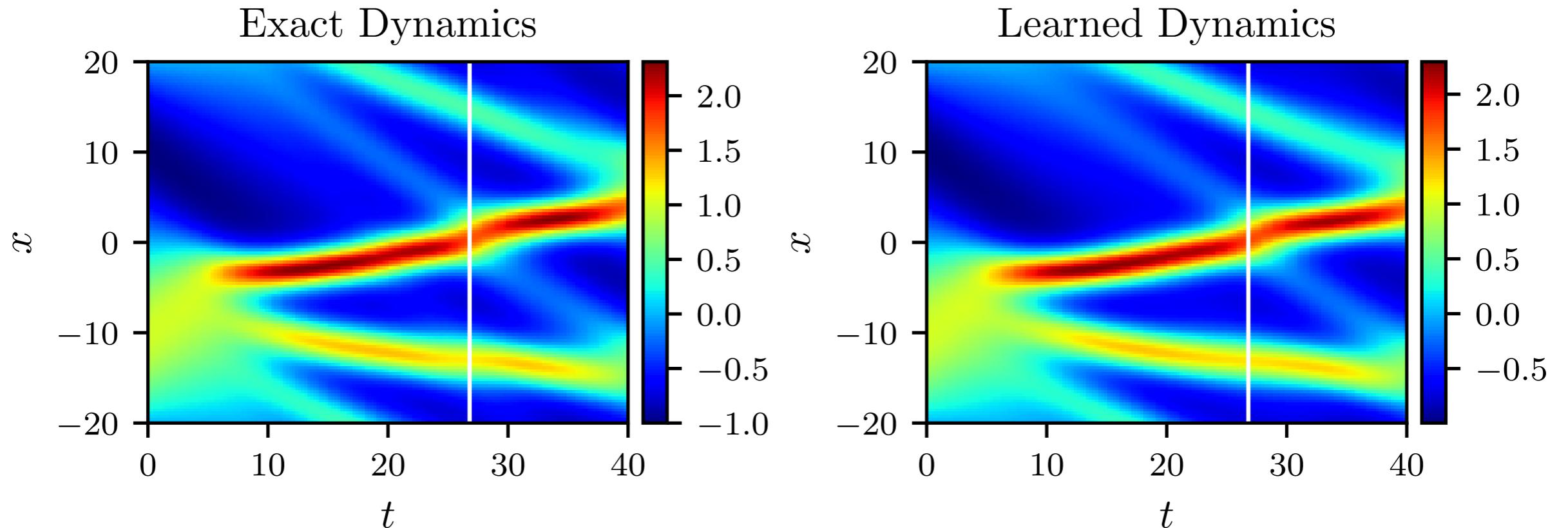
Systems identification (PDEs)

$$u_t = \mathcal{N}(t, x, u, u_x, u_{xx}, \dots)$$

Example:

The KdV equation:

$$u_t = -uu_x - u_{xxx}$$



Systems identification (PDEs)

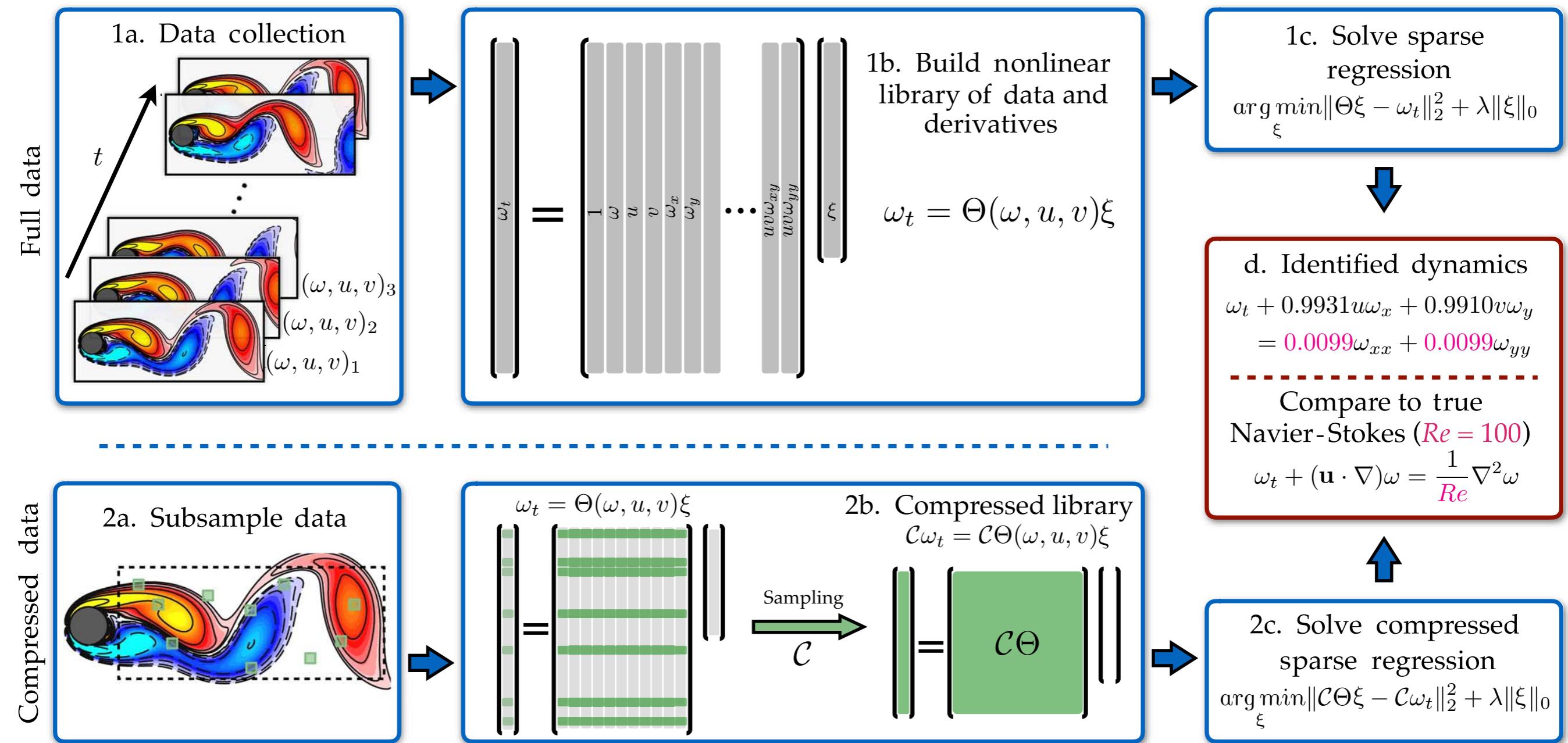
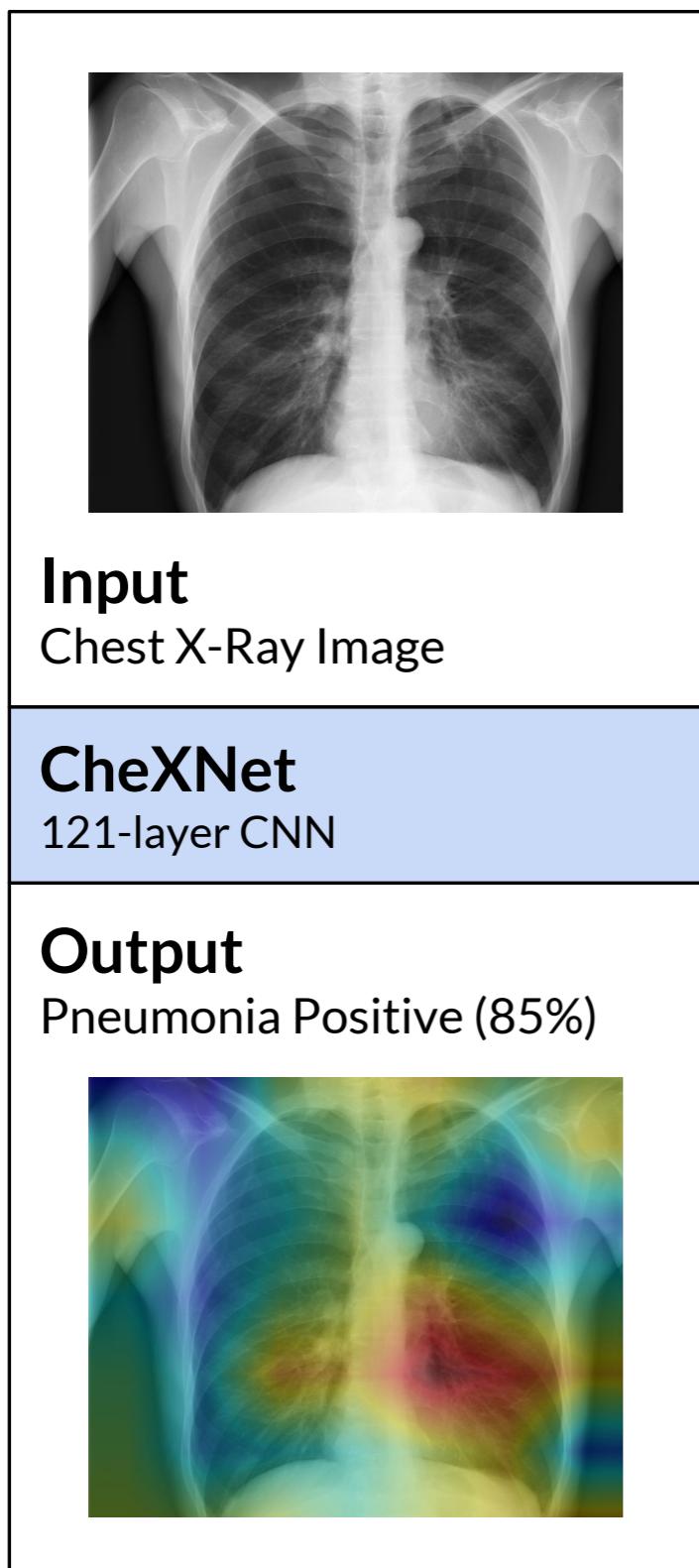


Fig. 1. Steps in the PDE functional identification of nonlinear dynamics (PDE-FIND) algorithm, applied to infer the Navier-Stokes equations from data. (1a) Data are collected as snapshots of a solution to a PDE. (1b) Numerical derivatives are taken, and data are compiled into a large matrix Θ , incorporating candidate terms for the PDE. (1c) Sparse regressions are used to identify active terms in the PDE. (2a) For large data sets, sparse sampling may be used to reduce the size of the problem. (2b) Subsampling the data set is equivalent to taking a subset of rows from the linear system in Eq. 2. (2c) An identical sparse regression problem is formed but with fewer rows. (d) Active terms in ξ are synthesized into a PDE.

Diagnosing Pneumonia at Radiologist-Level Accuracy

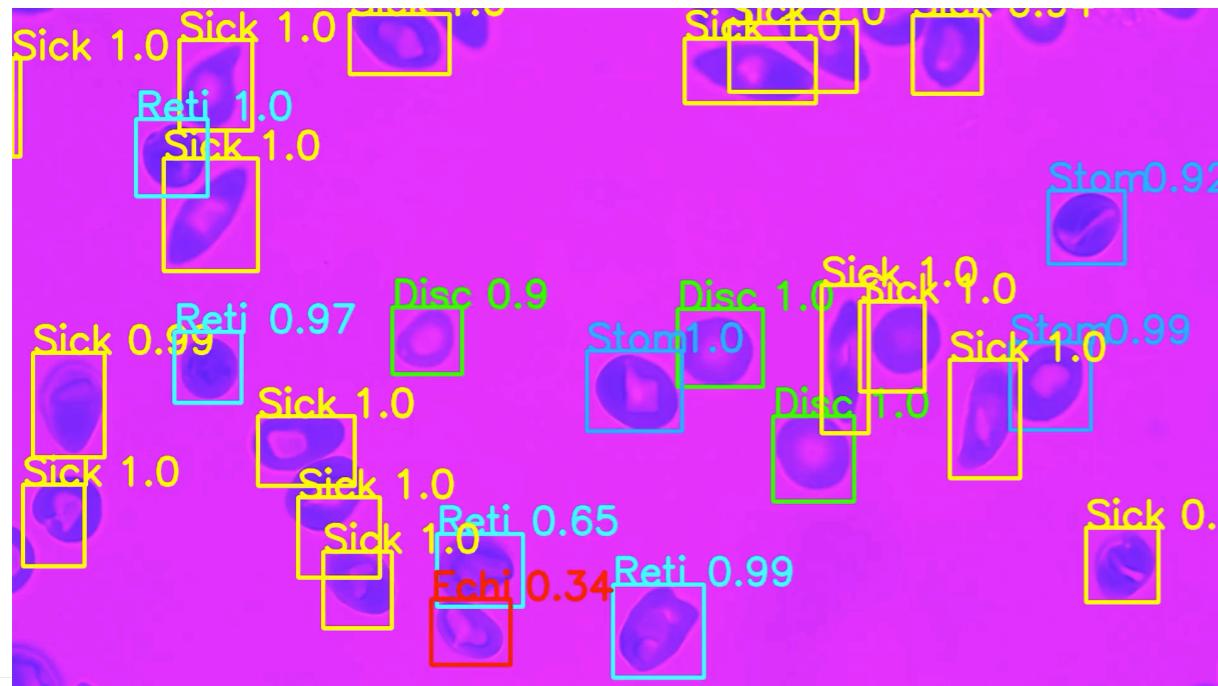


Abstract

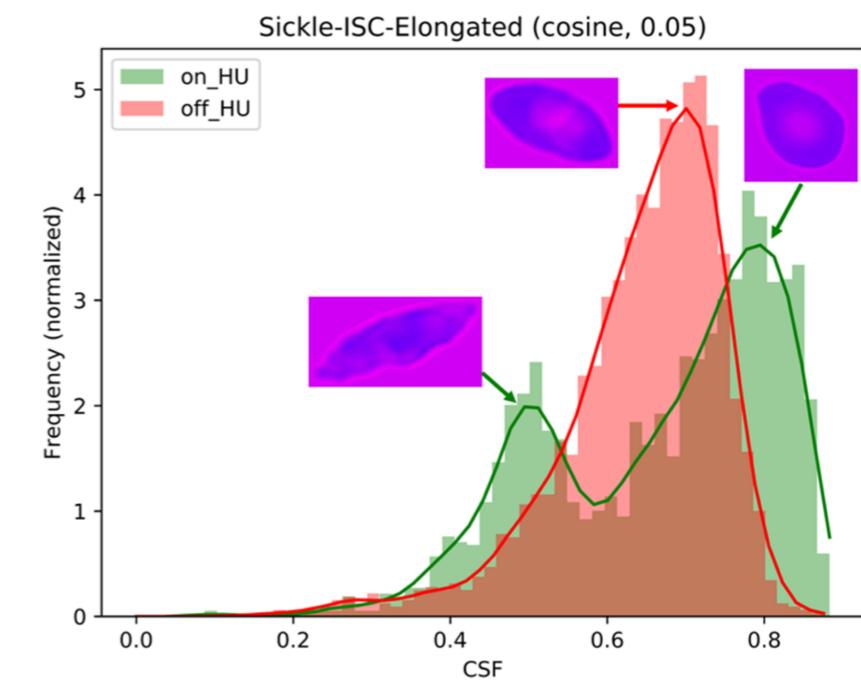
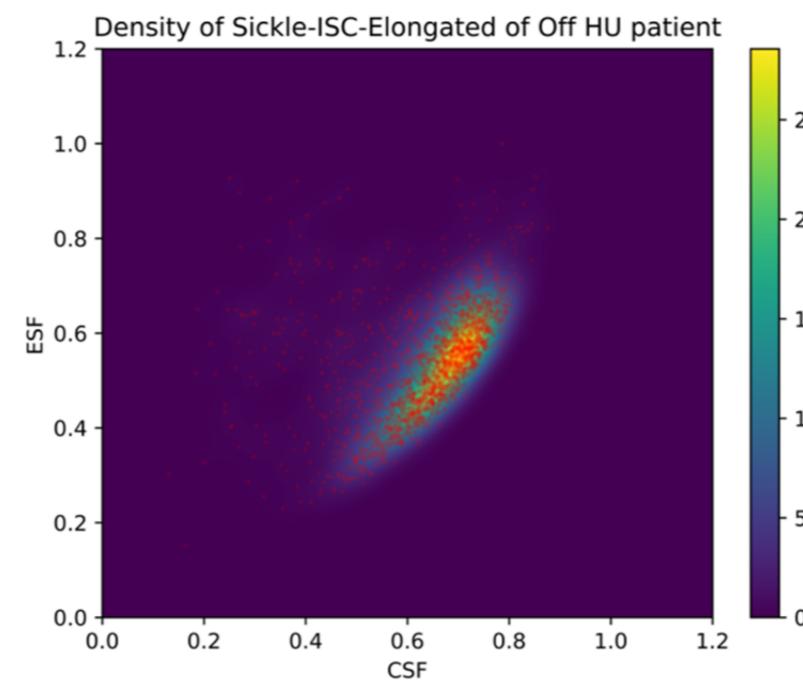
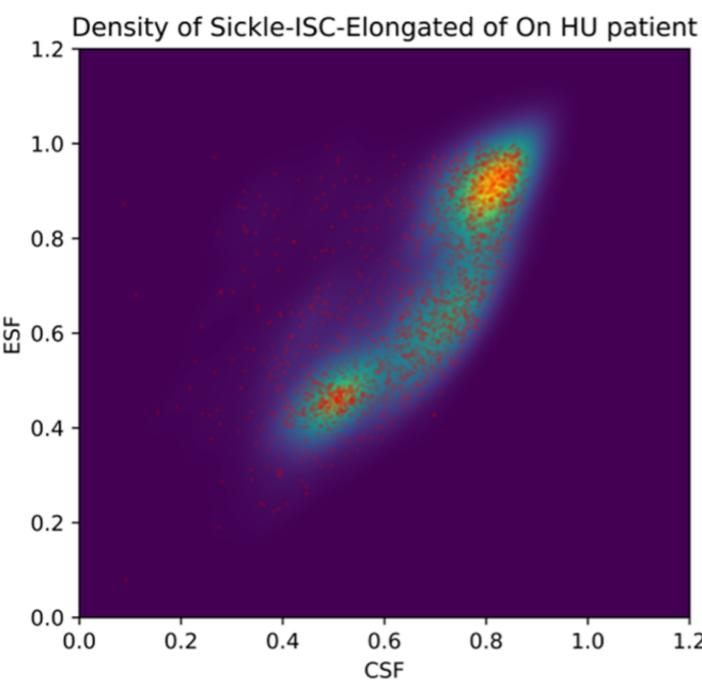
We develop an algorithm that can detect pneumonia from chest X-rays at a level exceeding practicing radiologists. Our algorithm, CheXNet, is a 121-layer convolutional neural network trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset, containing over 100,000 frontal-view X-ray images with 14 diseases. Four practicing academic radiologists annotate a test set, on which we compare the performance of CheXNet to that of radiologists. We find that CheXNet exceeds average radiologist performance on the F1 metric. We extend CheXNet to detect all 14 diseases in ChestX-ray14 and achieve state of the art results on all 14 diseases.

	F1 Score (95% CI)
Radiologist 1	0.383 (0.309, 0.453)
Radiologist 2	0.356 (0.282, 0.428)
Radiologist 3	0.365 (0.291, 0.435)
Radiologist 4	0.442 (0.390, 0.492)
Radiologist Avg.	0.387 (0.330, 0.442)
CheXNet	0.435 (0.387, 0.481)

A deep learning system for blood drop analysis



	On Drug		Off Drug	
	Count	% of Total	Count	% of Total
Discocytes_Oval	2412	27.29	637	13.31
Echinocytes-II-III	124	1.40	162	3.38
Granular-Echino-I	116	1.31	49	1.02
Reticulocytes	40	0.45	211	4.41
Sickle-ISC-Elongated	2316	26.20	2144	44.80
Stomatocytes	3831	43.34	1583	33.08
Total	8839	100	4786	100



Arrhythmia Detection with ConvNets

Abstract

We develop an algorithm which exceeds the performance of board certified cardiologists in detecting a wide range of heart arrhythmias from electrocardiograms recorded with a single-lead wearable monitor. We build a dataset with more than 500 times the number of unique patients than previously studied corpora. On this dataset, we train a 34-layer convolutional neural network which maps a sequence of ECG samples to a sequence of rhythm classes. Committees of board-certified cardiologists annotate a gold standard test set on which we compare the performance of our model to that of 6 other individual cardiologists. We exceed the average cardiologist performance in both recall (sensitivity) and precision (positive predictive value).

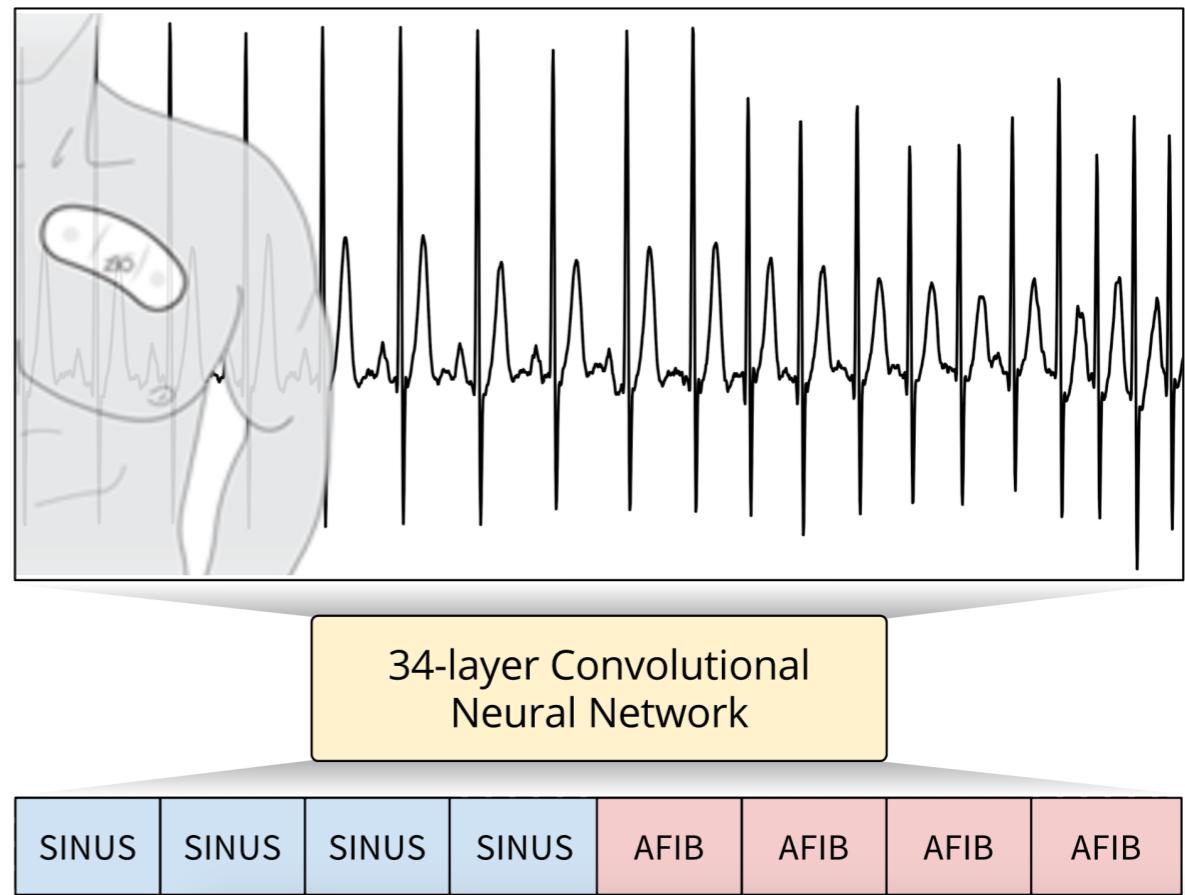
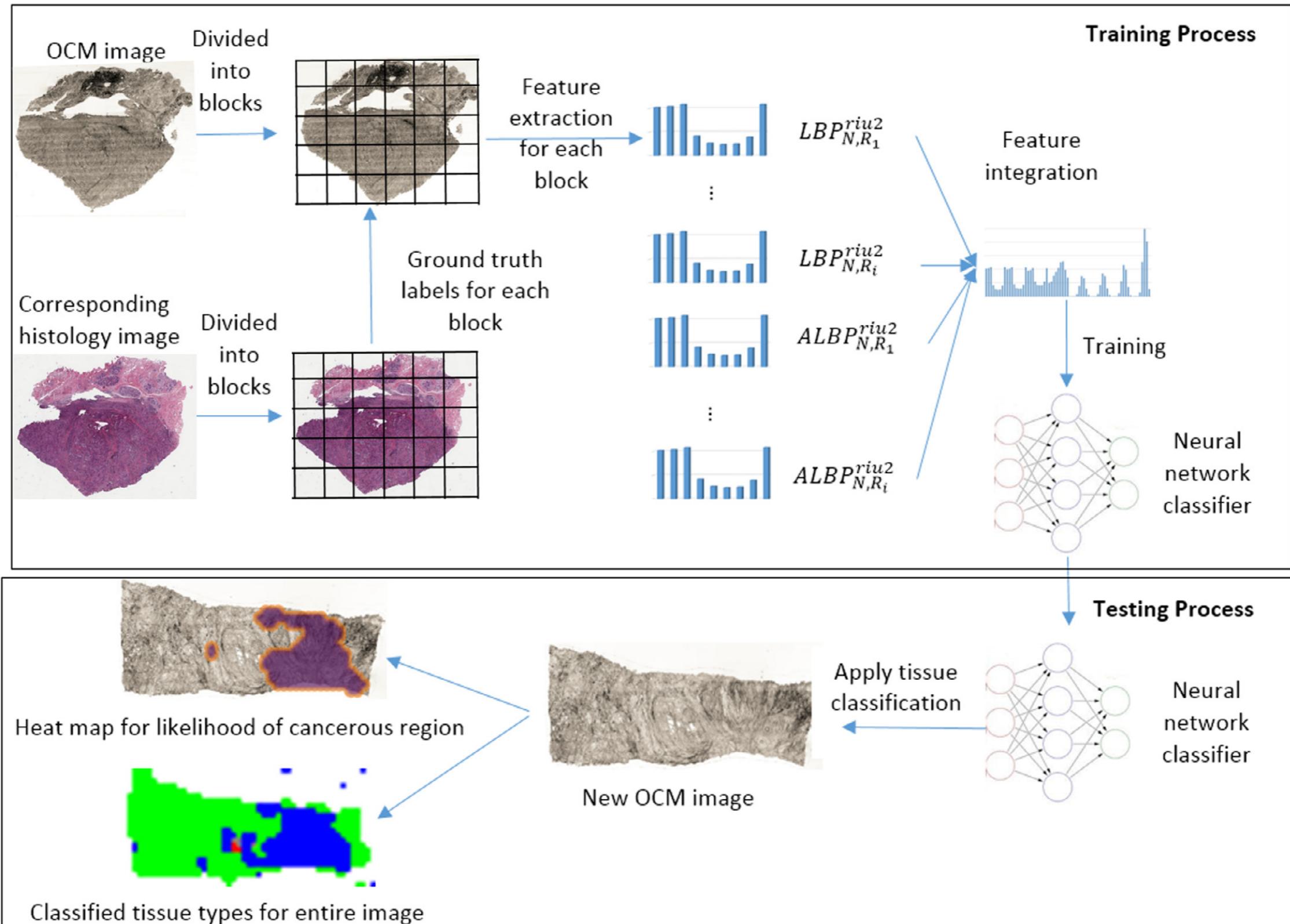
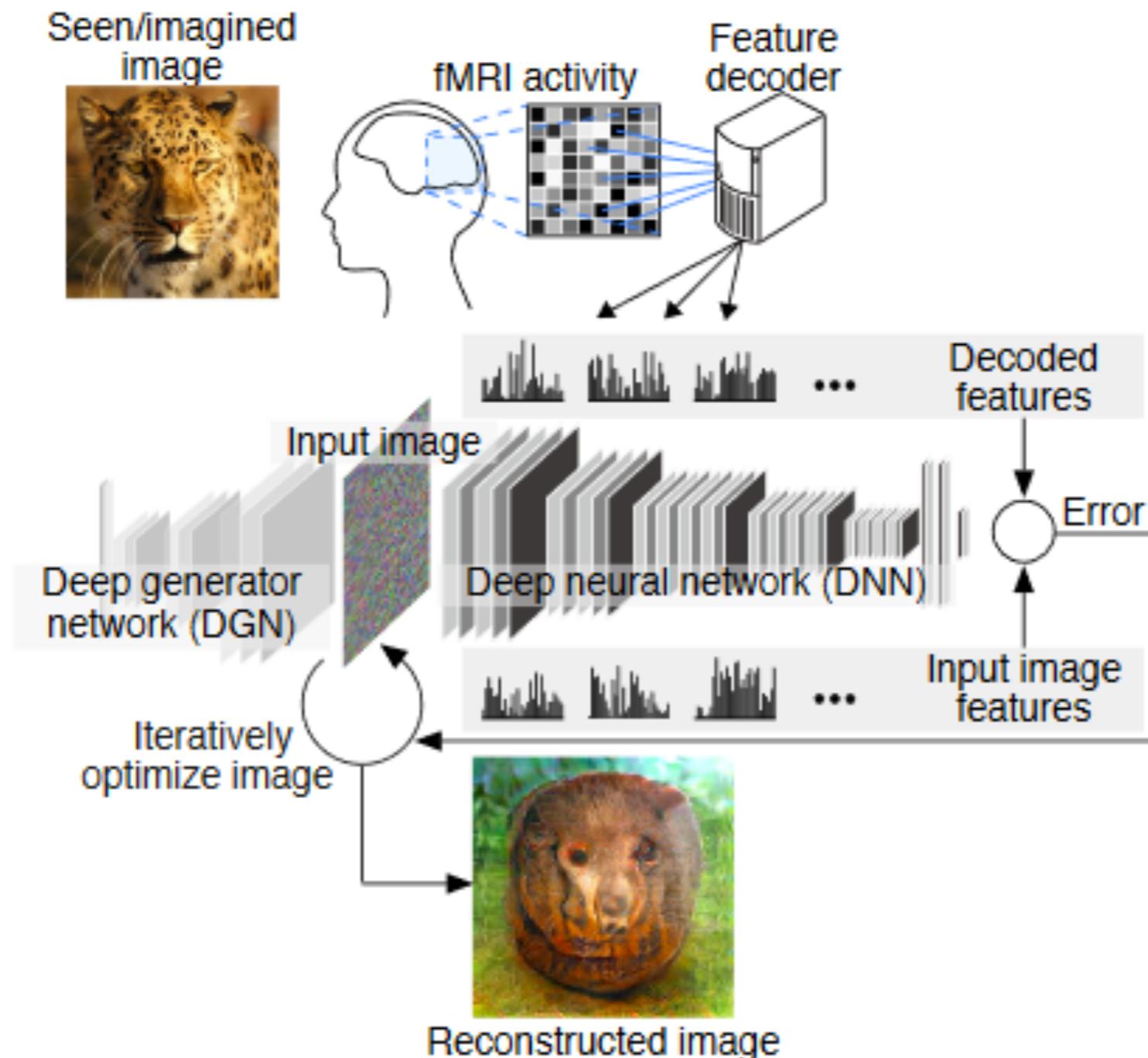


Figure 1. Our trained convolutional neural network correctly detecting the sinus rhythm (SINUS) and Atrial Fibrillation (AFIB) from this ECG recorded with a single-lead wearable heart monitor.

Reducing the Risk of Second Breast Cancer Surgery



Deep image reconstruction from human brain activity



Overview of deep image reconstruction. The pixels' values of the input image are optimized so that the DNN features of the image are similar to those decoded from fMRI activity. A deep generator network (DGN) is optionally combined with the DNN to produce natural-looking images, in which optimization is performed at the input space of the DGN.

Shen, G., Horikawa, T., Majima, K., & Kamitani, Y. (2017). Deep image reconstruction from human brain activity. *bioRxiv*, 240317.

Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning

Babak Alipanahi^{1,2,6}, Andrew Delong^{1,6}, Matthew T Weirauch^{3–5} & Brendan J Frey^{1–3}

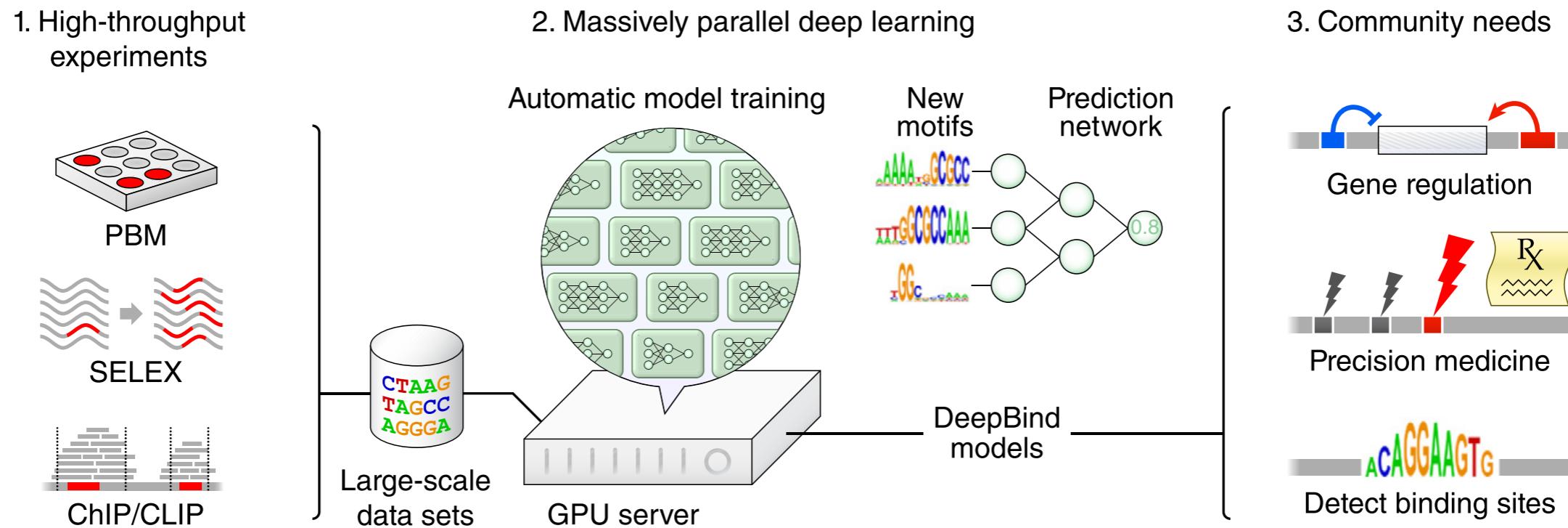
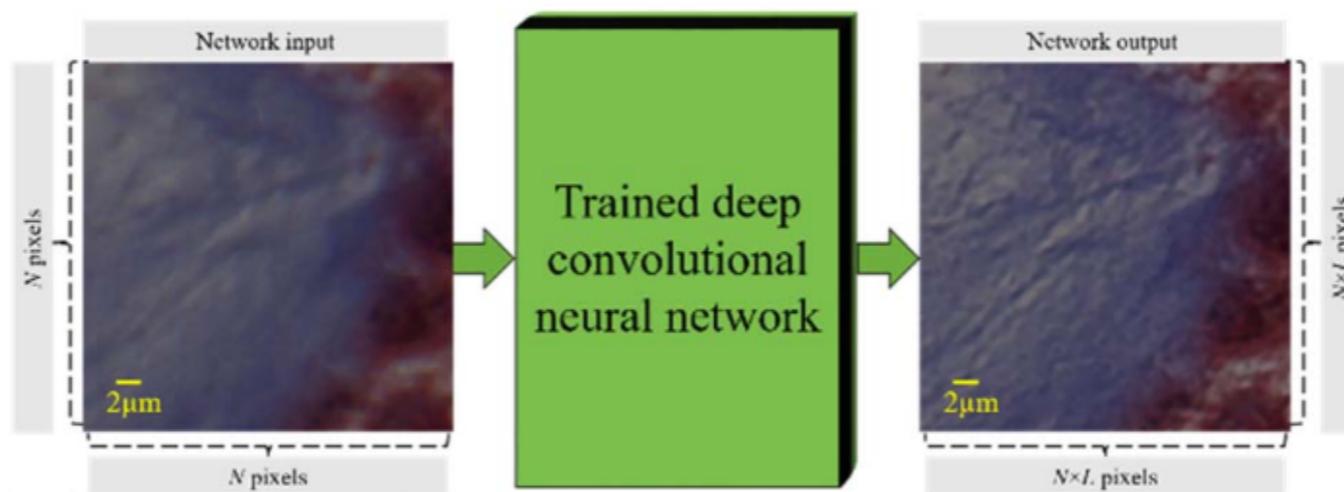
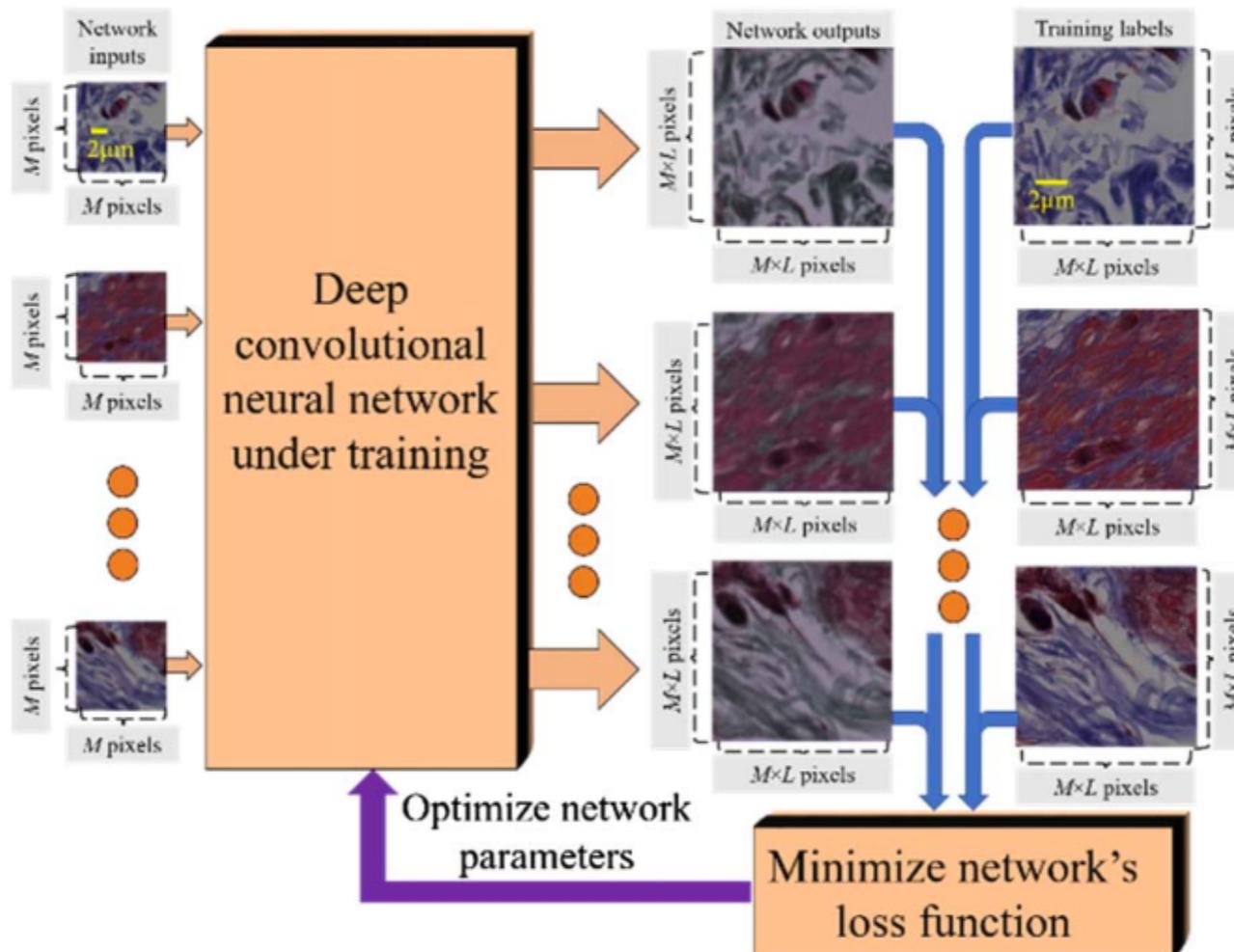


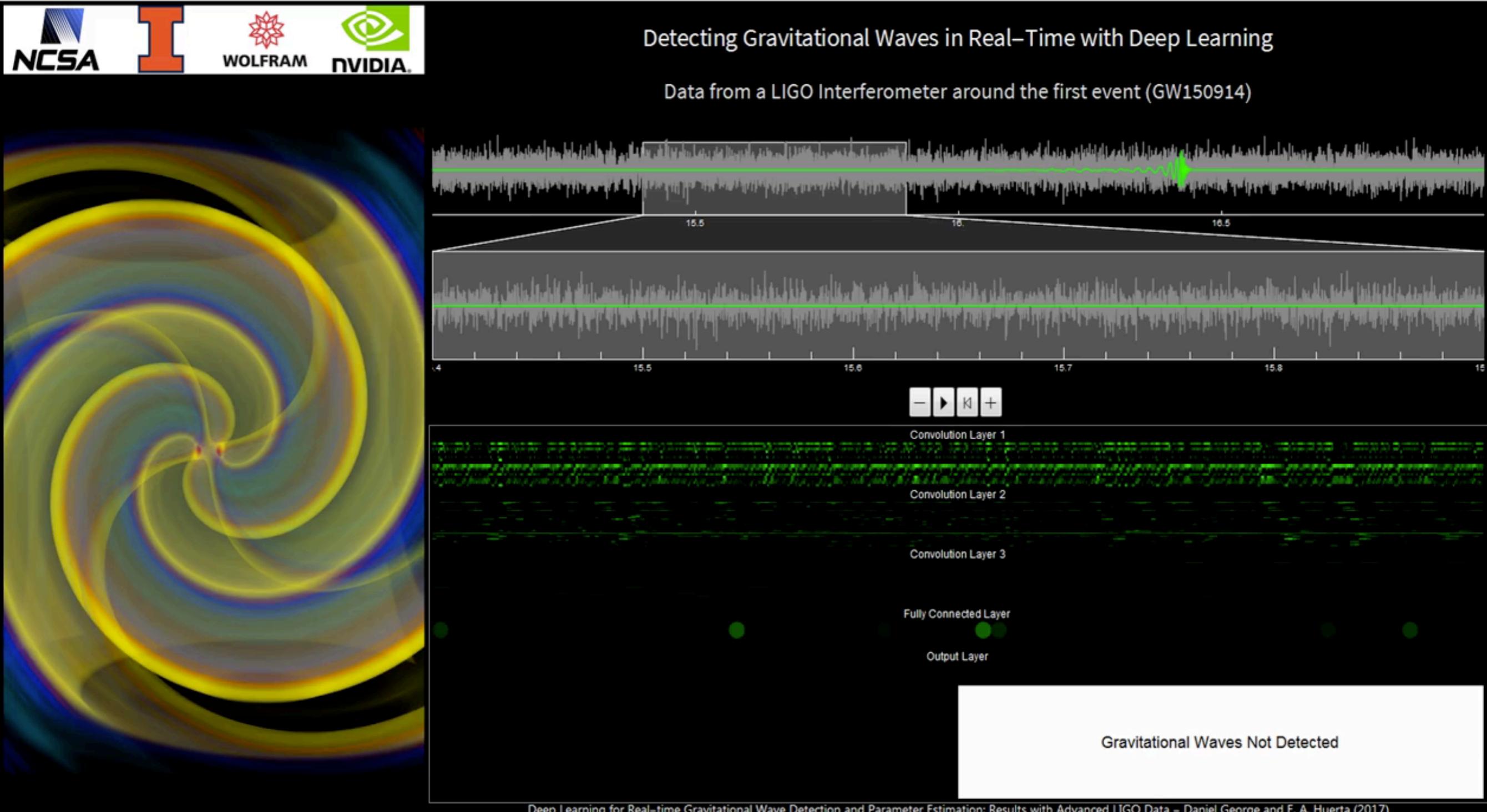
Figure 1 DeepBind's input data, training procedure and applications. 1. The sequence specificities of DNA- and RNA-binding proteins can now be measured by several types of high-throughput assay, including PBM, SELEX, and ChIP- and CLIP-seq techniques. 2. DeepBind captures these binding specificities from raw sequence data by jointly discovering new sequence motifs along with rules for combining them into a predictive binding score. Graphics processing units (GPUs) are used to automatically train high-quality models, with expert tuning allowed but not required. 3. The resulting DeepBind models can then be used to identify binding sites in test sequences and to score the effects of novel mutations.

Improving optical microscopy



Rivenson, Y., Zhang, Y., Gunaydin, H., Teng, D., & Ozcan, A. (2017). Phase recovery and holographic image reconstruction using deep learning in neural networks. *arXiv preprint arXiv:1705.04286*.

Real-time detection of gravitational waves



Design of efficient molecular organic light-emitting diodes

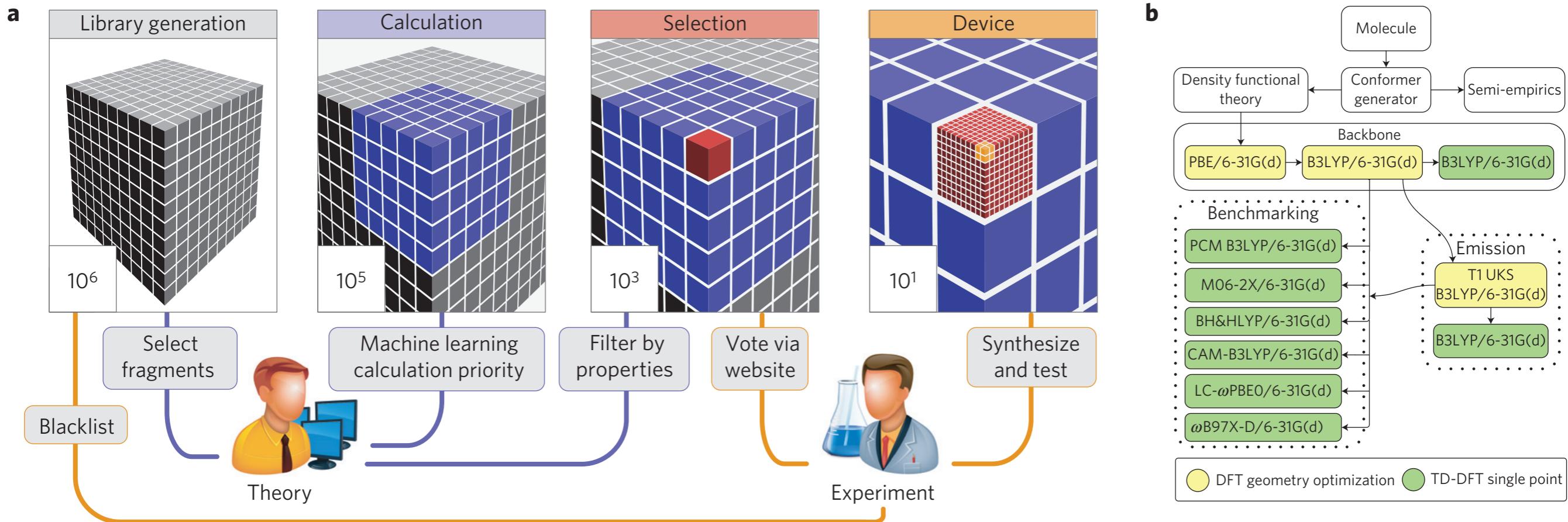


Figure 1 | Discovery pipeline. **a**, Diagram of the collaborative discovery approach: the search space decreases by over five orders of magnitude as the screening progresses. The cubes represent the size of the chemical space considered at any given stage of the process. The distinct screening stages, from left to right, involve different theoretical and computational approaches as well as experimental input and testing. **b**, Dependency tree for the quantum chemistry calculations employed in this study. The calculations labelled as backbone were performed for all analysed molecules, leading compounds were also characterized using the methods labelled emission, and the benchmarking calculations were used to assess predictive power.

Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., Hirzel, T. D., Duvenaud, D., Maclaurin, D., Blood-Forsythe, M. A., ... & Markopoulos, G. (2016). Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nature materials*, 15(10), 1120.

Molecule design

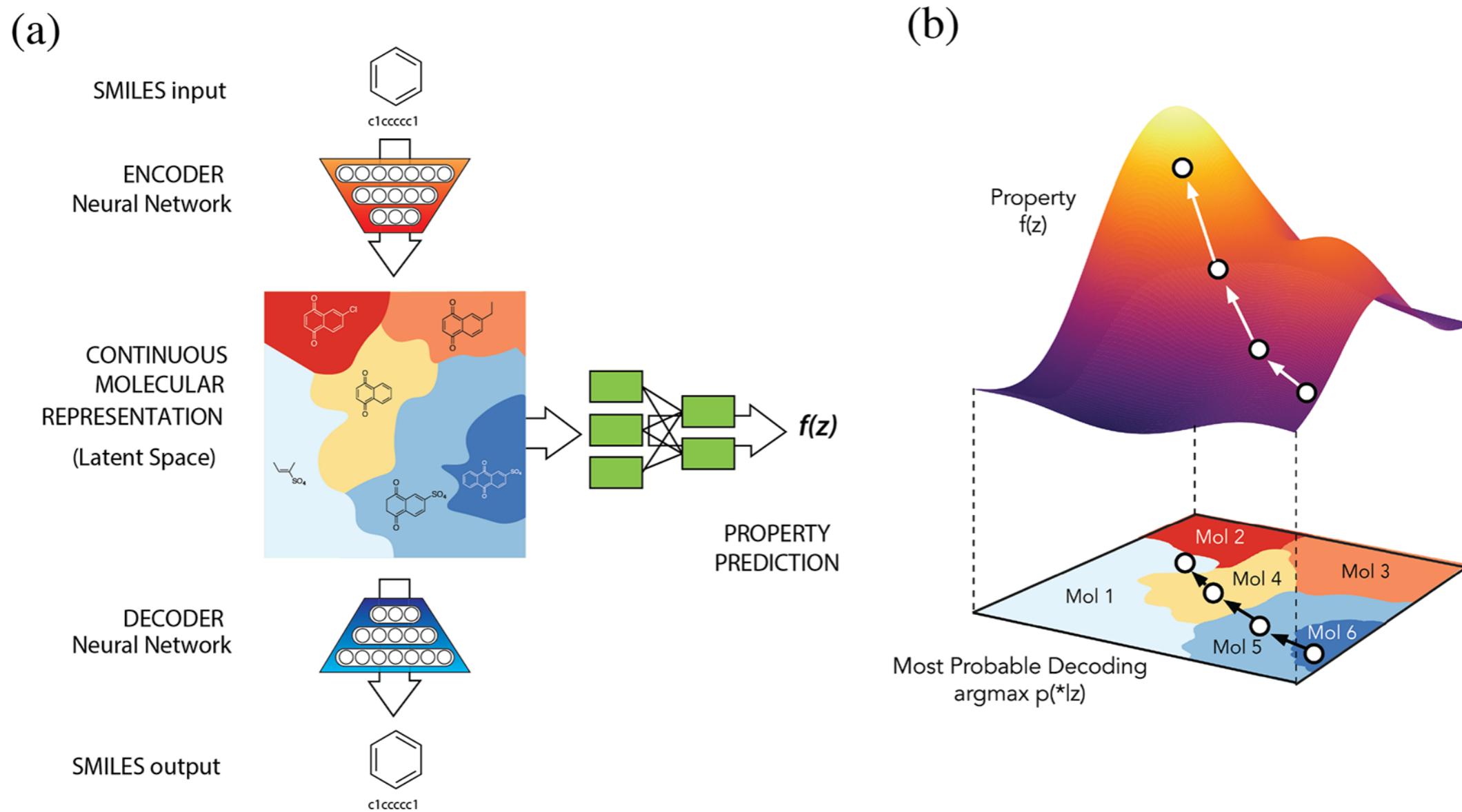


Figure 1. (a) A diagram of the autoencoder used for molecular design, including the joint property prediction model. Starting from a discrete molecular representation, such as a SMILES string, the encoder network converts each molecule into a vector in the latent space, which is effectively a continuous molecular representation. Given a point in the latent space, the decoder network produces a corresponding SMILES string. A multilayer perceptron network estimates the value of target properties associated with each molecule. (b) Gradient-based optimization in continuous latent space. After training a surrogate model $f(z)$ to predict the properties of molecules based on their latent representation z , we can optimize $f(z)$ with respect to z to find new latent representations expected to have high values of desired properties. These new latent representations can then be decoded into SMILES strings, at which point their properties can be tested empirically.

Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., ... & Aspuru-Guzik, A. (2016). Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*.

Auto-tuning code for high-performance computing

```
def avgpool(float(B, C, H, W) input) -> (output) {{
    output(b, c, h, w) += input(b, c, h * {sH} + kh, w * {sW} + kw)
        where kh in 0:{kH}, kw in 0:{kW}
}}
```



Tensor Comprehension for 2D Average Pooling

...and the list goes on and on

Shifting Ground: AI Analyzes Volcanoes for Signs of Eruption

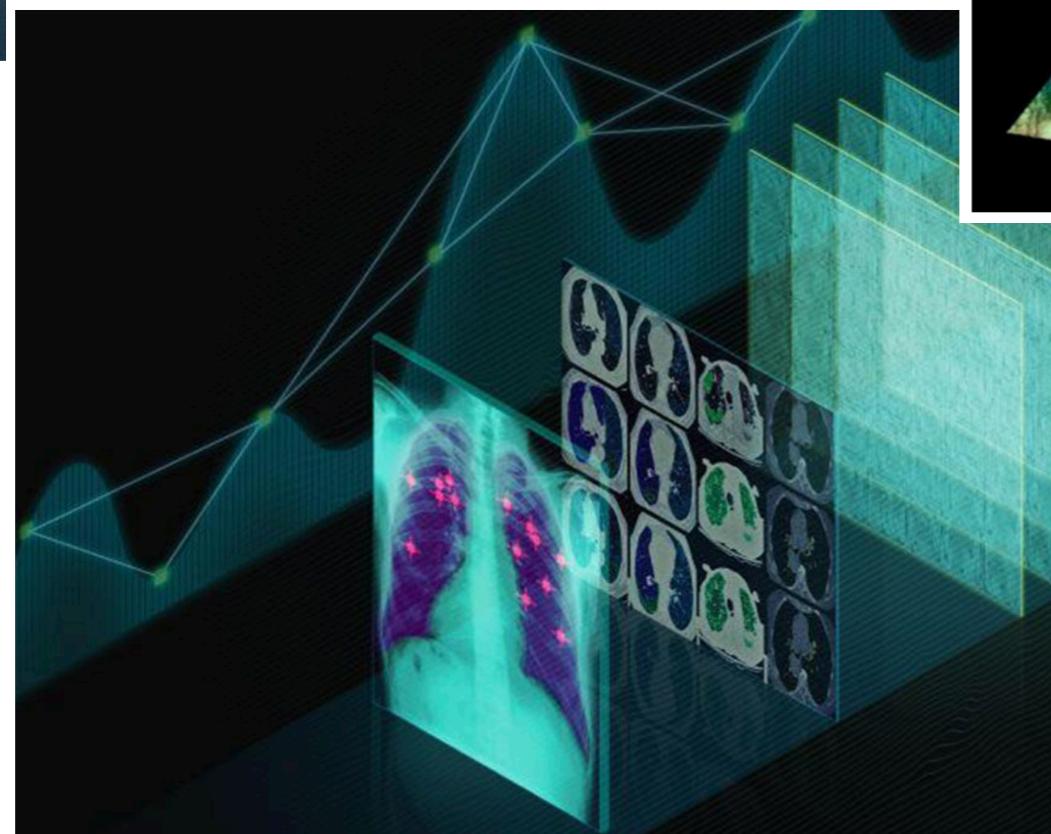
March 5, 2019 by ISHA SALIAN



How AI Is Changing Medical Imaging

Neural networks are analyzing medical imaging data, transforming the f

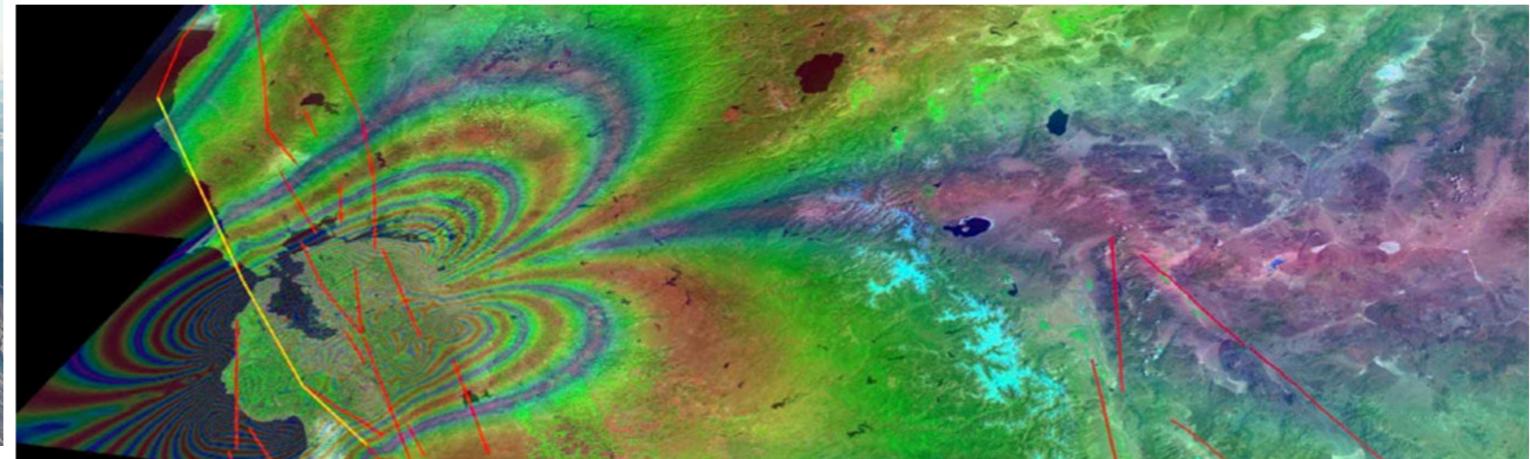
March 4, 2019 by ISHA SALIAN



Deep Learning Shakes Up Seismology with Quake Early Warning System

When these scientists say “earth-shaking deep learning innovation,” they mean it.

February 28, 2019 by ISHA SALIAN



Cell by Cell: Deep Learning Powers Drug Discovery Effort for Hundreds of Rare Diseases

January 14, 2019 by ISHA SALIAN



Alchemy or science?



Deep Learning: Alchemy or Science?

Deep learning has led to rapid progress in open problems of artificial intelligence—recognizing images, playing Go, driving cars, automating translation between languages—and has triggered a new gold rush in the tech sector. But some scientists raise worries about slippage in scientific practices and rigor, likening the process to “alchemy.” How accurate is this perception? And what should the field do to combine rapid innovation with solid science and engineering? A distinguished panel of experts convened by Visiting Professor Sanjeev Arora will discuss these issues.

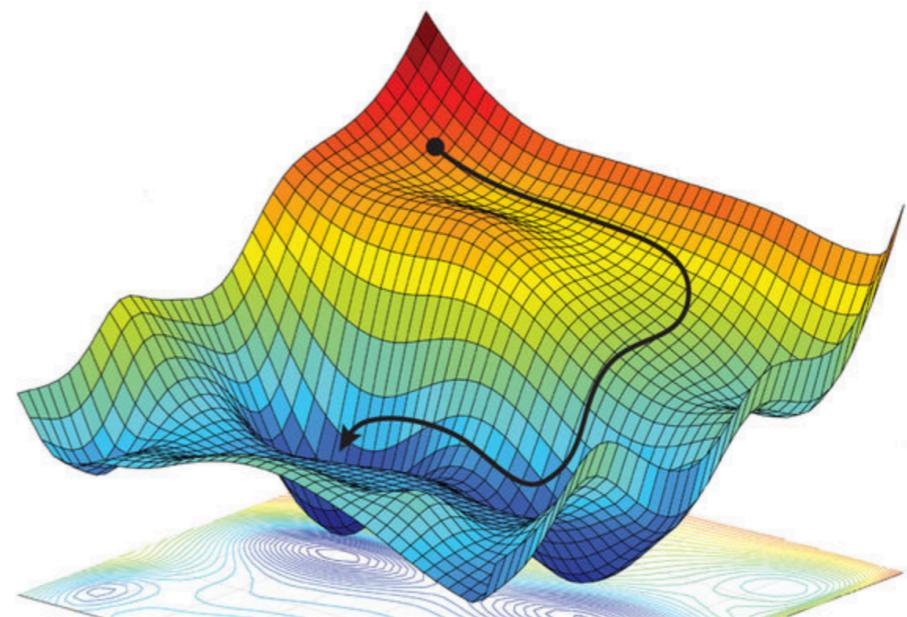
Speakers Include:

Michael Collins, Google Research, Columbia University
Yann LeCun, Facebook, New York University
Joelle Pineau, Facebook, McGill University
Zachary Lipton, Carnegie Mellon University
Shai Shalev-Shwartz, Hebrew University of Jerusalem

IAS INSTITUTE FOR ADVANCED STUDY

Casper What napaholics are saying: I will never leave my bed again. Caryn from California Learn more Log in | My account

SHARE



Gradient descent relies on trial and error to optimize an algorithm, aiming for minima in a 3D landscape.
ALEXANDER AMINI, DANIELA RUS. MASSACHUSETTS INSTITUTE OF TECHNOLOGY, ADAPTED BY M. ATAROD/SCIENCE

AI researchers allege that machine learning is alchemy

By Matthew Hutson | May 3, 2018 , 11:15 AM

Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, took a swipe at his field last December—and received a 40-second ovation for it. Speaking at an AI conference, Rahimi charged that machine learning algorithms, in which computers learn through trial and error, **have become a form of “alchemy.”** Researchers, he said, do not know why some algorithms work and others don’t, nor do they have rigorous criteria for choosing one AI architecture over another. Now, in a paper presented on 30 April at the International Conference on Learning Representations in Vancouver, Canada, Rahimi and his collaborators **document examples** of what they see as the alchemy problem and offer prescriptions for bolstering AI’s rigor.

How (un)certain are you about your machine learning model predictions?