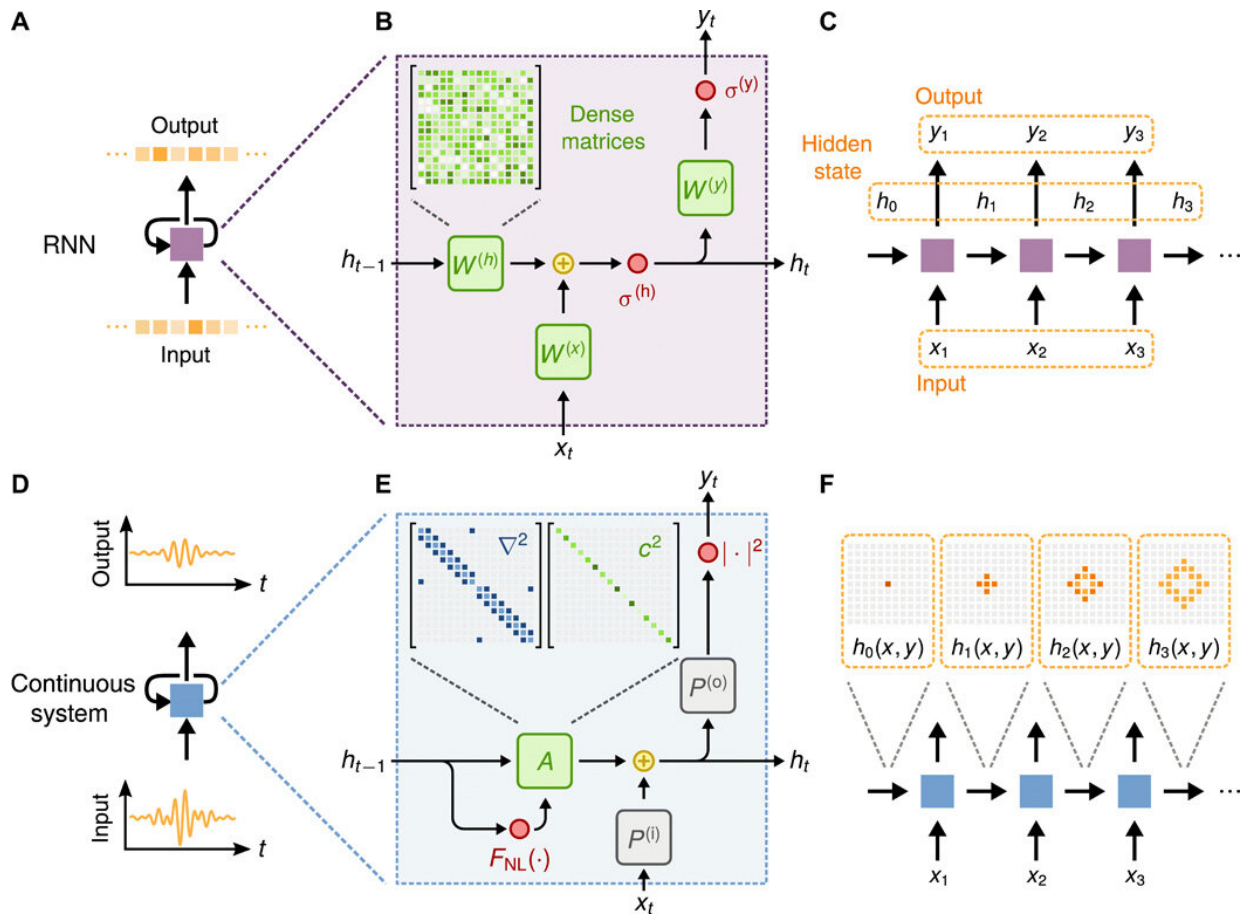


Wave physics as an analog recurrent neural network

January 10 2020, by Thamarasee Jeewandara



Conceptual comparison of a standard RNN and a wave-based physical system. (A) Diagram of an RNN cell operating on a discrete input sequence and producing a discrete output sequence. (B) Internal components of the RNN cell, consisting of trainable dense matrices $W^{(h)}$, $W^{(x)}$, and $W^{(y)}$. Activation functions for the hidden state and output are represented by $\sigma^{(h)}$ and $\sigma^{(y)}$, respectively. (C) Diagram of the directed graph of the RNN cell. (D) Diagram of

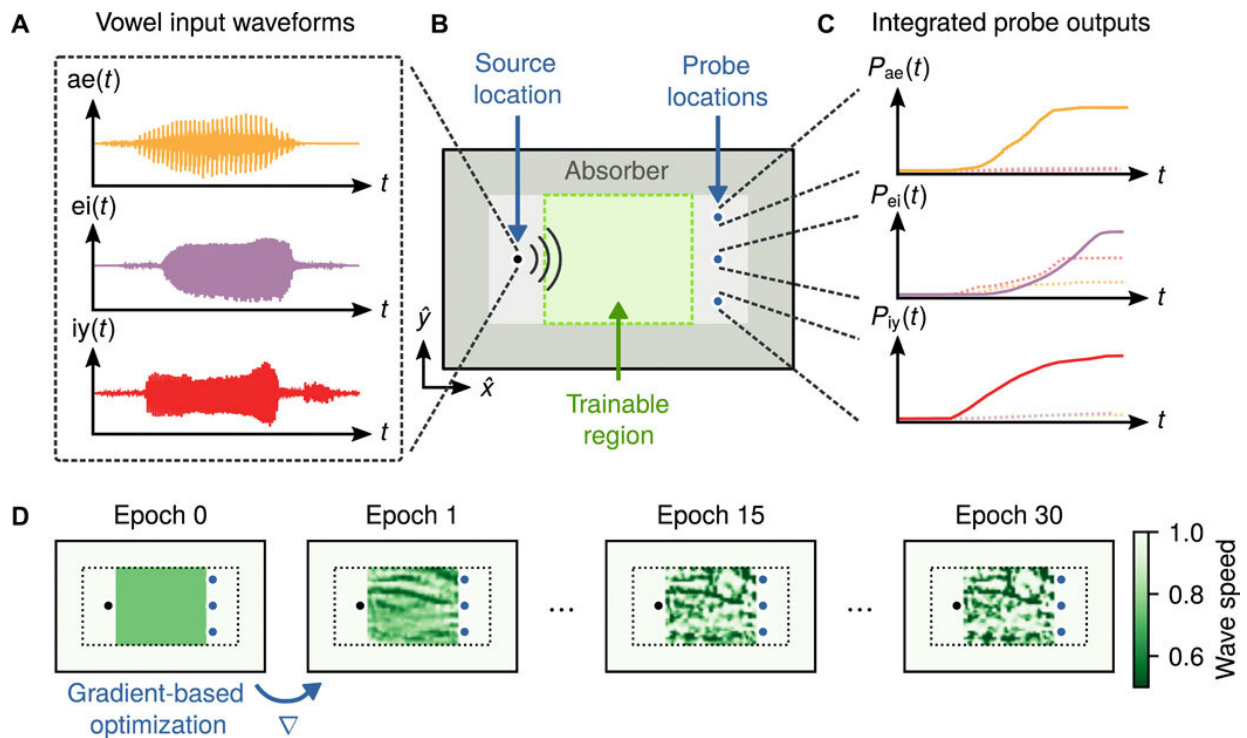
a recurrent representation of a continuous physical system operating on a continuous input sequence and producing a continuous output sequence. (E) Internal components of the recurrence relation for the wave equation when discretized using finite differences. (F) Diagram of the directed graph of discrete time steps of the continuous physical system and illustration of how a wave disturbance propagates within the domain. Credit: *Science Advances*, doi: 10.1126/sciadv.aay6946

Analog machine learning hardware offers [a promising alternative](#) to digital counterparts as a more energy efficient and faster platform. Wave physics [based on acoustics and optics](#) is a natural candidate to build analog processors for time-varying signals. In a new report on *Science Advances* Tyler W. Hughes and a research team in the departments of Applied Physics and Electrical Engineering at Stanford University, California, identified mapping between the dynamics of wave physics and computation in recurrent neural networks.

The map indicated the possibility of training physical wave systems to learn complex features in temporal data using standard training techniques used for [neural networks](#). As proof of principle, they demonstrated an inverse-designed, inhomogeneous medium to perform English vowel classification based on raw audio signals as their waveforms scattered and propagated through it. The scientists achieved performance comparable to a standard digital implementation of a recurrent neural network. The findings will pave the way for a new class of [analog machine learning](#) platforms for fast and efficient information processing within its native domain.

The [recurrent neural network](#) (RNN) is an important [machine learning model](#) widely used to perform tasks including [natural language processing](#) and [time series prediction](#). The team trained wave-based physical systems to function as an RNN and passively process signals

and information in their native domain without [analog-to-digital conversion](#). The work resulted in a substantial gain in speed and reduced power consumption. In the present framework, instead of implementing circuits to deliberately route signals back to the input, the [recurrence relationship](#) occurred naturally in the time dynamics of the physics itself. The device provided the memory capacity for information processing based on the waves as they propagated through space.



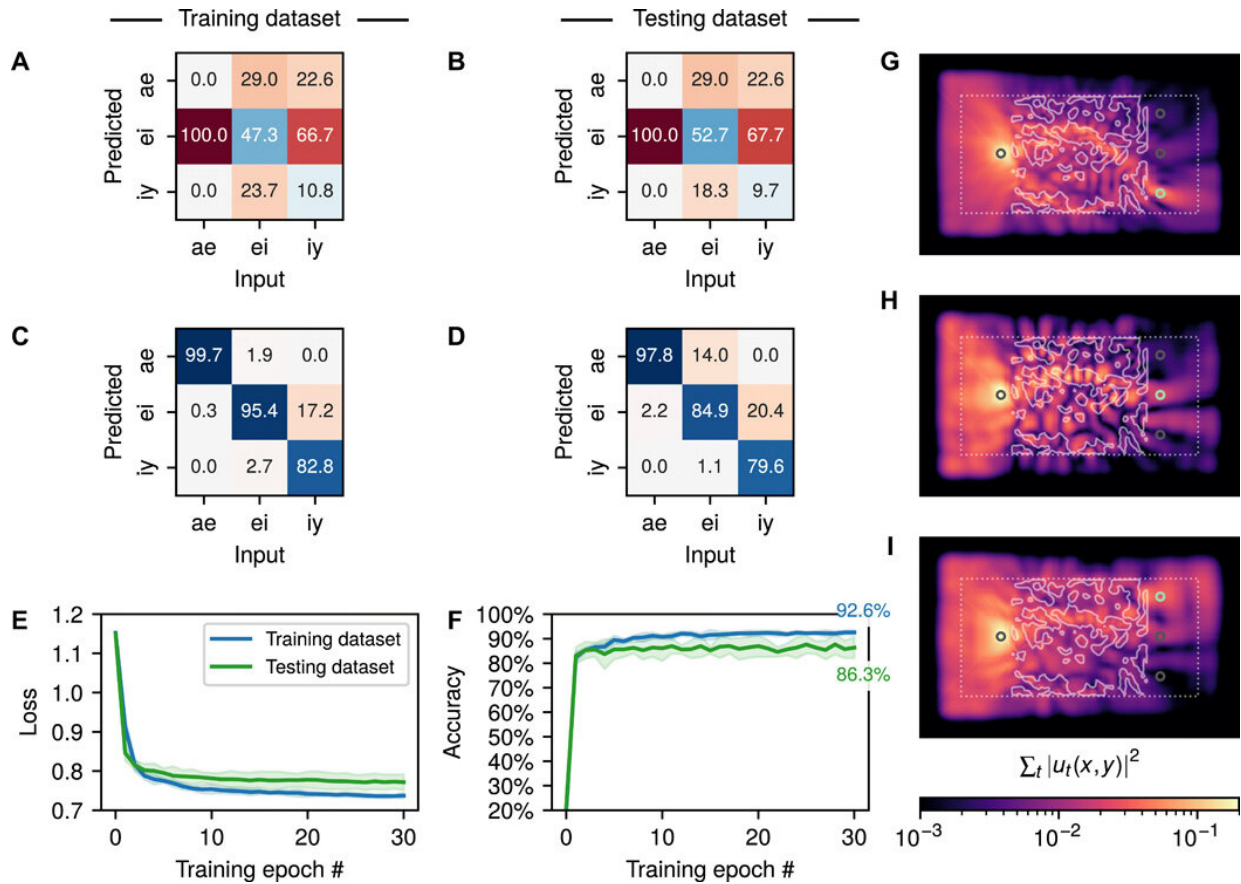
Schematic of the vowel recognition setup and the training procedure. (A) Raw audio waveforms of spoken vowel samples from three classes. (B) Layout of the vowel recognition system. Vowel samples are independently injected at the source, located at the left of the domain, and propagate through the center region, indicated in green, where a material distribution is optimized during training. The dark gray region represents an absorbing boundary layer. (C) For classification, the time-integrated power at each probe is measured and normalized to be interpreted as a probability distribution over the vowel classes. (D) Using automatic differentiation, the gradient of the loss function with

respect to the density of material in the green region is computed. The material density is updated iteratively, using gradient-based stochastic optimization techniques until convergence Credit: *Science Advances*, doi: 10.1126/sciadv.aay6946

Equivalence between wave dynamics and an RNN

To demonstrate the equivalence between wave dynamics and an RNN, Hughes et al. introduced the function of an RNN and its connection to wave dynamics. For example, an RNN can convert a sequence of inputs into a sequence of outputs by applying the same basic operation to each input sequence member in a stepwise process. The RNN's hidden state will then encode the memory of previous steps to update at each step. The hidden states could retain memory of past information and learn temporal structure and [long-range dependencies in data](#).

At a given step, as an example, the RNN can function on the current input vector in the sequence (x_t) and the hidden state vector from the previous step (h_{t-1}), to produce an output vector (y_t) and an updated hidden state (h_t). While many variations of RNNs exist, Hughes et al. implemented a [commonly incorporated strategy](#) in the present work. The research team observed a nonlinear response, which is typically encountered in a wide variety of wave physics including [shallow water waves](#), [nonlinear optical materials](#) (study of intense laser light with matter) and acoustically within [soft materials and bubbly fluids](#). When modeled numerically in discrete time, the wave equation defined an operation that mapped into that of an RNN.

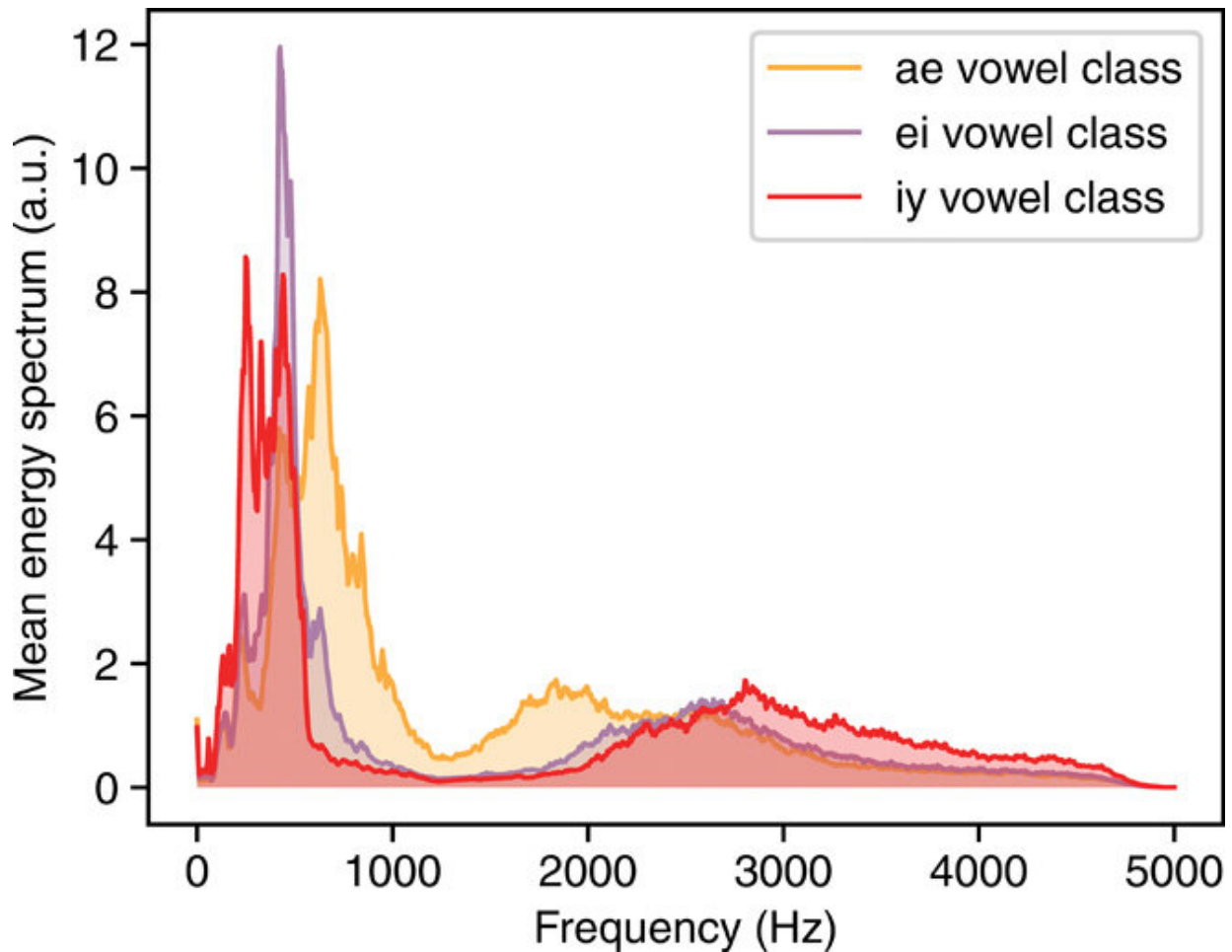


Vowel recognition training results. Confusion matrix over the training and testing datasets for the initial structure (A and B) and final structure (C and D), indicating the percentage of correctly (diagonal) and incorrectly (off-diagonal) predicted vowels. Cross-validated training results showing the mean (solid line) and SD (shaded region) of the (E) cross-entropy loss and (F) prediction accuracy over 30 training epochs and five folds of the dataset, which consists of a total of 279 total vowel samples of male and female speakers. (G to I) The time-integrated intensity distribution for a randomly selected input (G) ae vowel, (H) ei vowel, and (I) iy vowel. Credit: Science Advances, doi: 10.1126/sciadv.aay6946

Training a physical system to classify vowels

The team then demonstrated how the wave equation dynamics could be trained to [classify vowels](#) by constructing an inhomogeneous material distribution. For this, they used a dataset of 930 raw audio recordings of 10 vowel classes from 45 different male speakers and 48 different female speakers. For the learning task, Hugh et al. selected a subset of 279 recordings corresponding to three vowel classes represented by the vowel sounds "ae," "ei" and "iy," relative to their use in the words "had," "hayed" and "heed." The physical layout of the vowel recognition system contained a two-dimensional domain in the x-y plane and infinitely extended in the z-direction. They injected the audio waveform of each vowel via a source at a single grid cell on the left side of the domain for emitting waveforms to propagate through a central region with a trainable distribution of the wave speed. They defined three probes on the right-hand side of the region and assigned each to one of the three vowel classes. Hugh et al. then measured the time-integrated power at each probe to determine the system's output.

The simulation evolved for the full duration of the vowel recording and the team included an absorbing boundary region represented by a dark gray region to prevent energy build up within the computational domain. The wave speeds could be modified to correspond to different materials in practice. In an acoustic setting, for instance, if the material distribution consisted of air, the sound speed was 331 m/s, while porous silicone rubber constituted a [sound speed of 150 m/s](#). The choice of starting structure allowed them to shift the optimizer toward either of the two materials, to produce a [binarized structure](#) containing only one of the two materials. Hughes et al. trained the system by performing back-propagation through the model of the wave equation, in an approach mathematically equivalent to [the adjoint method](#) widely used for [inverse design](#). Using this design information, they updated the material density via the [Adam optimization algorithm](#), repeating until convergence on a final structure.



Frequency content of the vowel classes. The plotted quantity is the mean energy spectrum for the ae, ei, and iy vowel classes. a.u., arbitrary units. Credit: Science Advances, doi: 10.1126/sciadv.aay6946

Visualizing the performance

The scientists used a [confusion matrix](#) to visualize the performance across the training and testing data sets for the starting structures, averaged across five cross-validated training runs. The confusion matrix defined the percentage of correctly predicted vowels along its diagonal

entries and the percentage of incorrectly predicted vowels for each class in its off-diagonal entries. The diagonally dominant trained confusion matrices indicated the structure could indeed perform vowel recognition. Hughes et al. noted the cross-entropy loss value and the prediction accuracy as a function of the training epoch on the testing and training datasets.

The first epoch resulted in the largest reduction of the loss function and the largest gain in prediction accuracy, with a mean accuracy of 92.6 percent on the training dataset and a mean accuracy of 86.3 percent on the testing dataset. The team observed the system to obtain near-perfect prediction performance on the "ae" vowel alongside the ability to differentiate the "iy" vowel from the "ei" vowel—but with lesser accuracy within the unseen samples from the testing datasets. In this way, the team provided visual confirmation on the optimization procedure to route most of the signal energy to the correct probe. As a performance benchmark, they trained a conventional RNN on the same task to achieve classification accuracy comparable to the wave equation. However, they required a large number of free parameters for the task.

In this way, Tyler W. Hughes and colleagues presented a wave-based RNN with a number of favorable qualities to form a promising candidate to process temporally encoded information. The use of physics [to perform computation](#) may inspire a new platform for analog machine learning devices in order to perform computation far more naturally and efficiently than its digital counterparts. The research team determined the size of the analog RNN's hidden state and its memory capacity using the size of the propagation medium. They showed the dynamics of the wave equation to be conceptually equivalent to those of an RNN. The conceptual connection will pave the way for a new class of analog hardware platforms, wherein the evolving time dynamics will play a major role in both the physics and the dataset.

More information: Tyler W. Hughes et al. Wave physics as an analog recurrent neural network, *Science Advances* (2019). [DOI: 10.1126/sciadv.aay6946](https://doi.org/10.1126/sciadv.aay6946)

A. Silva et al. Performing Mathematical Operations with Metamaterials, *Science* (2014). [DOI: 10.1126/science.1242818](https://doi.org/10.1126/science.1242818)

Miguel Romera et al. Vowel recognition with four coupled spin-torque nano-oscillators, *Nature* (2018). [DOI: 10.1038/s41586-018-0632-y](https://doi.org/10.1038/s41586-018-0632-y)

© 2020 Science X Network

Citation: Wave physics as an analog recurrent neural network (2020, January 10) retrieved 18 April 2024 from <https://phys.org/news/2020-01-physics-analog-recurrent-neural-network.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.