Research Article

CrossMark

How can AI mend California wild fires?

Abstract

Wildfires are a natural phenomenon to rejuvenate soil with burning ashes. Nonetheless, the ashes can be replaced by recycle paper-boxes. We must reduce the collateral damage to the mankind and wild animals. For example, San Francisco Sierra Area discovered on Nov. 14 2018 as PG&E outage and damage to a transmission tower shortly before the fire started and killed 78 peoples, injured 5 firefighters, missing 699 peoples, affected ten thousands families; the air became unhealthy to breath halting Airline travels and Holiday tourist shopping at the loss of $100B. The purpose of this paper is response to Governor J. Brown $1B initiative to suggest three proposals: short-term, mid-term, and long-term R/D. (i) Short-Term: Design Drones flying over tree tops and using smart AI image processing to provide the early autonomous detection of wildfires (ii) Mid-Term: Commercial Off–The Shelf (COTS) Robots Team clean the dry leaf and bushes, as well as (iii) Long-Term: federal research of plant gnomon from the redwood to the other trees for enhancement of the fire retardation.

Keywords: explore macroscopic aero surveillance, early warning by AI image understanding, explore microscopic possible genetic mutation for fire retardation

Harold H Szu

Department of Biomedical Engineering, The Catholic University, USA

Correspondence: Harold Szu, Department of Biomedical Engineering, The Catholic University, Wash DC, USA, Tel 2404 8268 89; Email szuharaldh@gmail.com

Background: How to mend the California wildfires?

We wish to in (i) short term, to equip less costly drone technology with AI tool to enhance the surveillance image capability; (ii) midterm, to apply AI Robot to clear out dry bushes among forest; (iii) long term, invigorate the plant genome study to endow the trees other than redwood with the fire retardation property. For (i) we can carry the knowhow of micro-UAV drone, Silver Fox, to overcome the blur caused by the small airplane motion & jittering under the wind, before AI deep learning image understanding can be applied. For example, if we can identify three control points of a triangle for the correspondingpoints in two overlapping frames, we can solve three equations for two dimensions which determine six unknowns (a,b,c,d,e,f)needed to transform one frame to another. This is called the affine transform[1] which is equivalent to triangle area-preserving (e,g, two eyes and one mouth point) between a smiling face-to be register on a crying face. If one can more than three are identified for area-moment preserving which can furthermore improve the signal to noise ratio at the dusk. We define by the high project geometry the jitter to be sub-pixel or small-amplitude vibration to be compensated. Until we can overcome the intrinsic jitter blur, of small drone we cannot take the advantageof the modern AI algorithms. After passively deduce the jitters and reduce the, we can further afford active device infrared (IR) Lidar for better ground surveillance.

Since CA State Governor J. Brown is recommended $1B funding to alleviate CA wildfires and the US Presidnet D. Trunk wishes to clear bushes. We contribute our thoughts in three time line categories

a) Short-term to detect early smoke or fire on board of drones flying over the tree top in windy days, and

b) Mid term to design robots to sweep dry bushes among trees; as well as

c) Long term to apply modern plant genome technology enhancing the fire retardation of all trees besides the redwood.

Motivation of three timeline proposals

The global warning has increased over many decades about 3 degrees in Celsius (4.5 degrees in Fahrenheit) that has melted the North ice cap raising the sea level, and reversing the carbon cycles. These facts have confused to the general public. The temperature is related to a random thermal energy that is quadratic in variances that mean large up and down fluctuation amplitude, from very hot to very cold. This fact is known as the Einstein fluctuation and dissipation theorem thus, the weather can fluctuate from a very cold and dry to a very hot and wet in the same seasons. Recently, the wildfires in California Sierra foot-hold e.g. Paradise, etc on Nov. 8, 2018 due to the buried power line sparks, and spreading quickly in wind, which had so far killed 78 people, and injured five firefighters & 699 people still missing, damaged 14,033 structures, destroyed 10,623 homes in 15 towns. Furthermore the air pollution had spreader dark smokes all over local residences and stopped Thanksgiving and Christmas seasonal business and tourism. Thus, we recommend three time proposals short-term, mid-term, and long-term effects (Figures 1&2). California Governor Jerry Brown calls for $1B effort to control wildfire. President Donald Trump weighed in with a call to "rake" forest floors of debris, which we propose a Mid-Term applying AI Robots team to do this job systematically. Since the plant genome program discovered about the redwood fire, insect, fungus retardation for decade, we propose a long term R/D to apply the know-how to transfer to the other species if possible. Thus, we have three time-schedule approaches:

Short-Term R/D: Call for proposal for AI deep learning of aerial video overcoming the sub-pixel jittering of drone to determine the black stream of smoke cluttered fire points, as opposed to other motion blue of the imaging sequence itself (Figure 3-5).

We assume bright object(s) on the ground to be visible by many frames of video within their field of views (FOV).Such three visible control points are also required to compute the distortion affine transform Eq.(1)for a lossless mosaic painting of the landscape in time on the flight (Figure 6). Until we can overcome the intrinsic jitter blur,[2] we cannot take the advantage of the other change detection and ATR algorithms. To develop a correction of jitter, we cannot afford active device, except passively deduce the jitter.

Mid-Term R/D: AI Robot team to clear dry bushes in forests.

For example, visionary policy helps`--Pine Ridge School in Magalia (Butte County) survived this month's firestorm largely

because a local group had recently thinned 11 acres of thick, fire-prone forest near the campus. Fire experts say hotter, drier conditions fueled by climate change are undermining efforts to prevent and put out unruly flames and limit destruction (Figure 7). Temperatures in California have increased about 3 degrees Celsius over the past century while droughts have become longer and more severe, only ripening the conditions for uncontrollable burns. "If the weather gets too extreme, all bets are off," said North, with the U.S. Forest Service. The Butte County Fire Safe Council, which has its office in Paradise but will likely relocate to Chico, isn't letting the bleak outlook deter its work. A box of notices that the group had planned to send to local property owners, informing them of an upcoming forest-clearing project, is believed to have survived in the back of an SUV. Executive Director DeAnda said that once she can get into the burned-out community to retrieve the mailers, the effort will move forward. The project involves reinforcing a 10-mile fuel break around Paradise, which may or may not have slowed the spread of the current fire. DeAnda believed it was only a matter of time before the area would burn. "It's been my worst nightmare," she said. "But I'm not surprised. (*Rain to bring relief to Bay Area and fire zone, but one storm isn't a cure-all* (Figure 8). Likewise, one hundred and eleven years ago in 1906 the San Francisco disaster fire happened, again by morning stove fire, quickly spreading all over the city. San Francisco Earthquake in **1906** all struck the coast of Northern California at **5:12** a.m. when early risers cooking stoves were set the fire on April 18 with an estimated moment magnitude of **7.8**, along the end of St. Andrew Fall Line (Figure 9).

(A)

(B)

Figure 1

A. Passengers got strangered as 100 airline flights have been cancelled at San Francisco Internarional Airport. Hundreds Park Rangers together with thoudsands volunteer frefighters dug fire isolation lanes to achieve 65% confinement of fire;

B. hundreds of homes were evacuated in Sirerra Town, e.g. Morgan, Paradise, etc. towns, fire burning in SF Bay area with thick smoky filled with the air that made it difficult to breath withou wearing an air mask.
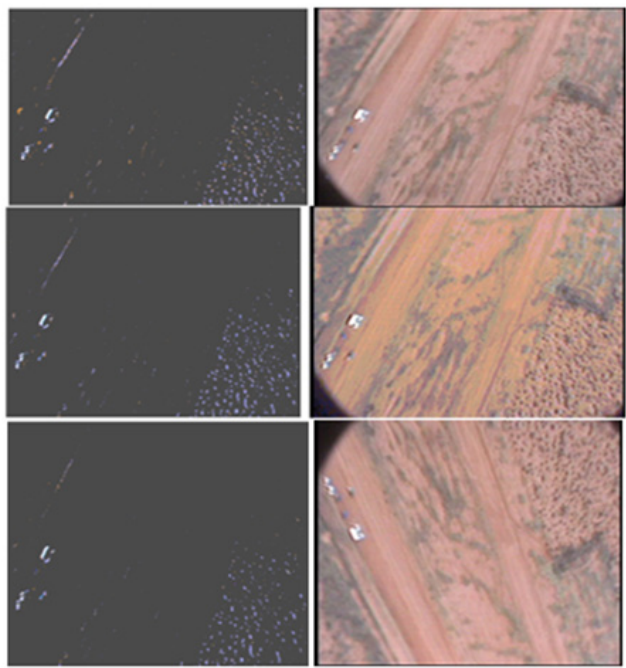

Figure 2

A. Fire Alarm in SF;near AT &T Bryon St,;

B. Wildfire may be caused by PG &E Powerline Utility downed wire in early morning 6:30 a.m. on Nov. 8, 2018. The fire might ignite by poiwerline but spread by wind blow to neighborhood bushes, and then the fire was spreading guickly beyond campers' control. The wildfires & smokes were all over the Northern California dry Sierra foot hold about dozen towns over 150,000 acreage of land;

C. None stop rescue helicopters of Park Rangers have evacuated and saved hundreds residence that some shall be equipped with AI Deep learning Video Image Understanding for early detection from smokes to fires. We woud recommend lesser expensive mini-Unmanned Aero Vehicle, namely Drone, to survey the forest field fpr wildfires.

Figure 3

a) ONR Silver Fox Drone;
b) COTS Cameras; and
c) Vibration Spectrum of different Airplane.

Figure 4 Test Fly over the ground in thousand meters shows three typical frames sequence A simple pair-wise neighborhood fields/frames produced the "agreement in the right and disagreement in the left.

Figure 5 Flying over sequence of pictures, using COTS On-board direct mounted Video camera and Silver Fox Model Airplane; Raw Video having sub-pixel jitter blur (LHS) vs. real-time improvement (RHS).

Figure 6 Affine Transform[2]: If we can identify three control points of a triangle for the corresponding points in two overlapping frames, we can solve three equations for two dimensions which determine six unknowns (a,b,c),(d,e,f) needed to transform one frame to another. This is called a perfect affine transform which is equivalent to triangle area-preserving (e,g, two eyes and one mouth point) between a smiling face-and a crying face/ If more than three are identified for better signal to noise reasons, then the geometry meaning has shown.[1]

AI Unsupervised learning at gradient descent at minimum helmholtz free energy

In this paper, we propose how to use cheaper drone for aero surveillance video benchmarked with i-Phone ground video taking in-situ real-time surveillance pictures along the road. We will load both with Artificial Intelligence (AI) & Artificial Neural Networks (ANN). In this paper in the interests of efficiency, we show a flying drone high over for possible detection of fire from smokes by correcting motion jittering using modern AI deep learning. Deep learning means multiple layers feature extraction, namely one layer per feature, in cases of multiple class Automatic Target Recognition (ATR). We assume all tall trees must be up-right against the gravity, and need only the Vertical Edge Filters keeping only 6 Pixels on Target (POT) on tree trunks reducing multi-billion pixels to multi-million pixels: Then, anything else are low growing bushes, and weeds. Furthermore,

A. Hand-on engineering to increase the relevance from the irrelevance, namely trimming tree top canopy & tree bottom bushes, keeping trunks.

B. Down-sampling the pixels resolution dynamic range keeping the same image size by Padding, & Max-Pooling;

C. Get rid of color R,G,B channels to grey-scale image.

Thus, by means of (a)(b)(c), we have increased the explainability of AI, (DARPA PM G. Manning XAI), we can improve the performance Receiver Operation Characteristics (ROC), by increasing the Probability of Detection (PD) versus the reducing the False Alarm Rate (FAR).

Through the history of CA, the "clean cuts" tax bill became "*selective logging,*" in reducing the density of *Redwood Forests*. Industrialization pollution caused the Atmosphere blanket heating, mixing thermal energy with turbulent heavy raining & winds storms causing the soil loss around loner *redwood* trees. Then exposed shallower roots of ERW's caused unfortunately, the up-rooting of redwoods trees. A quick fix by Forrest Board was planting rapid growth *Douglass-fir* surrounding loner *redwood's*. Yet, the rapid growth *Douglas-fir* trees inhibited the slow growth *ERW* trees. The dilemma calls for a healthy ratio between two species. We decide to hike and check the Ground Truth of sparse population better preserved healthy forests in Woodside of San Francisco Bay Area

All training exemplars of AI Deep Learning are image vectors by lexicographically scanned 2-D image into 1-D single vector $\vec{\mathbf{x}}_i$, where $i = 1, 2, \ldots m$ a single training matrix

$$\left[\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \bar{\mathbf{x}}_3, \ldots \right] = \left[\mathbf{X_{training}} \right]. \qquad (1)$$

Once we have replaced the temporal average with the spatial average, we can introduce the irreversible thermodynamics of the Maxwell-Boltzmann Canonic Probability called the Natural Intelligence (NI) which is based on two necessary and sufficient principles observed from the common physiology of all animal brains (Szu et al., circa 1990)

Homeostasis Principle: we observe

Thermodynamic equilibrium all animals roaming on the Earth have isothermal brains operated at a constant temperature T_o, (Homosapiens 37°C for the optimum elasticity of hemoglobin, Chicken 40°C for hatching eggs)

Power of Sensor Pairs: All isothermal brains have pairs of input sensors $\bar{X}_{pairs}$ for the co-incidence account to de-noise: "*agreed, the signals; disagreed, the noise,*" for instantaneously signal filtering processing.

Figure 7 The tool box includes Surveillance Drones Figure 4, AI Walking Robot (a) and (b); (C) AI Sophia Video Image understanding (d) AI Robot conversational understanding home alone senior. Beside operation recommendation namely one shall hire more firefighters, detect and remove the undergrounding spark-producing power lines, equipping homes with fire-resistant building materials, clearing overgrown vegetation and more. All of the tactics are pricey, and even pursued in tandem, there are no guarantees. That's why we call for smart solution from US wide to share the $1B grant.

Figure 8 Can PG&E survive the Camp Fire?.

Figure 9 Early morning on 5:12 a.m. stove fire ignited disaster fire during the Earthquake of 7.8 Richet Scale on April 18, 1906.

The thermodynamics governs the total entropy production ΔS_{tot} of the brain and its environment that led to the irreversible heat death, due to the never-vanish Kelvin temperature (the 3rd law of thermodynamics) there always an incessant collision mixing toward more uniformity and larger entropy value.

$$\Delta S_{tot} > 0 \qquad (2)$$

Unsupervised learning rule

$$\Delta H_{brain} = \Delta E_{brain} - T_o \Delta S_{brain} \leq 0 . \qquad (3)$$

This is Darwinian Fittest Survival Natural (selection) Intelligence (NI). The unsupervised learning has been shown by Szu et al. at Herman Helmholtz Minimum Free Energy (MFE) useful in the most intuitive decision for Aided Target Recognition (AiTR) for a Maximum Probability of Detection and Minimum the dangerous False Negative Rate (FNR) (while the False Positive Rate is nuisance) for Darwinian evolution reasons.

The survival NI is intuitively simple, flight or fight, Parasympathetic nerve system as an auto-pilot.

Maxwell-Boltzmann equilibrium probability is derived early in (4b) in terms of the exponential weighted Helmholtz Free Energy of brain:

$$H_{brain} = E_{brain} - T_o S_{brain} + const. \qquad (4)$$

Of which the sigmoid logic follows as the two state of biological Neural Net (BNN) (growing new neuron recruit or trim prune old neuron) probability normalization dropping the integration constant:

$$\exp\left(-\frac{H_{recruit}}{k_B T_o}\right) / \exp\left(-\frac{H_{prune}}{k_B T_o}\right) + \exp\left(-\frac{H_{recruit}}{k_B T_o}\right) = 1 / [\exp(\Delta H) + 1] = \sigma(\Delta H) = \begin{cases} 1, \Delta H \to \infty \\ 0, \Delta H \to -\infty \end{cases}$$

$$(5)$$

$$in\ unit\ of\ k_B T_o dimensionless\ \Delta H = H_{recruit} - H_{prune} \quad Q.E.D.$$

Note that Russian Mathematician G. Cybenko has proved "Approximation by Superposition of a Sigmoidal Functions," Math. Control Signals Sys. (1989)2:303-314. Similarly, A. N.

Kolmogorov wrote (Dokl. Akad. Nauk, SSSR, 114(1957), 953-956) "On the representation of continuous functions of many variables by superposition of continuous function of one variable and addition"

Those activation of neurons are denoted as the column *vector $\bar{a}$* of thousands neurons denoted as lower case

$$\bar{a}^T = \left(a_1, a_2, \ldots \right)$$

after the squash binary sigmoid logic function, or bi-polar hyperbolic tangent logic function within the multiple layer deep learning, with the backward error propagation requires gradient descent derivatives: Massively Parallel Distributed Processing; superscript $l \in (1,2,\ldots) = R^1$ denotes $l-th$ layers. The 1K by 1K million pixels image spanned in the linear vector space of million orthogonal axes where the collective values of all neuron's activations $\vec{a}^{[l]}$ of the next l-th layer in the infinite dimensional Hilbert Space. The slope weight matrix $\left[W^{[l]} \right]$ and intercepts $\vec{\theta}^{[l]}$ will be adjusted based on the million inputs $\vec{X}^{[l-1]}$ of early layer. The threshold logic at the output will be Figure 5 (a) Do Away All Do loops using one-step MDP Algorithm within layers will be bi-polar hyperbolic tangent and (b) output layer bipolar sigmoid

$$\bar{a}^{[l]} = \sigma\left(\left[W^{[l]} \right] \bar{X}^{[l-1]} - \vec{\theta}^{[l]} \right),$$ (6)

We derive a nonlinear recursive iteration of the learning weight matrix as follows the long division of the inverse expansion:

$$\left[W^{[l]} \right] = \left[A^{[l]} \right]^{-1} = [[I] - \left([I] - \left[A^{[l]} \right] \right)]^{-1} \cong \left([I] - \left[A^{[l]} \right] \right) + \left([I] - \left[A^{[l]} \right] \right)^2 + \ldots$$ (7)
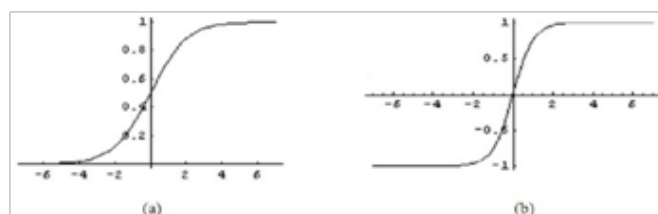
While Frank Rosenblatt developed ANN, Marvin Minsky challenged it and coined the name of Artificial Intelligence (AI) as the classical rule-based system. Steve Grossberg and Gail Carpenter of Boston Univ. developed Adaptive Resonance Theory (ART) model that have folded three layers down to itself as the top down & bottom up for local concurrency. Richard Lipmann of MITRE has given a succinct introduction of neural networks in IEEE ASSP Magazine 1984, where he proved that a single layer can do a linear classifier, and multiple layers give convex hull classifier to maximize the Probability of Detection (PD), and minimize the False Alarm Rate (FAR). Stanford Bernie Widrow; Harvard Paul Werbos, UCSD David Rumelhart, Carnegie-Mellon James McClelland, U. Torrente Geoffrey Hinton, UCSD Terence Sejnowski, have pioneered the Deep Learning multiple layers Models, Backward Error Propagation computational (backprop) model. The Output Performance could efficiently be the supervised learning at Least Mean Square (LMS) error cost function of the desired outputs versus the actual outputs. The Performance model could be more flexible by the relaxation process as unsupervised learning at Minimum Herman Helmholtz Free Energy: Brain Neural Networks (BNN) evolves from the Charles Darwinian fittest survival viewpoint he breakthrough came when he noted Lyell's suggestion that fossils found in rocks that the Galapagos Islands each supported its own variety of finch bird, a theory of evolution occurring by the process of Natural Selection or Natural Intelligence at the isothermal equilibrium thermodynamics due to[1] for a constant temperature brain (Homosapiens $37^o C$; Chicken $40^o C$) operated at a minimum isothermal Helmholtz free energy when the input power of pairs transient random disturbance of β —brainwaves may be represented by the degree of uniformity called the entropy S, as indicated by the random pixel histogram are relaxed to do the common sense work for the survival.

Healthy brain memory may be modeled as Biological Neural Networks (BNN) serving Massively Parallel and Distributed (MPD) commutation computing, and learning at synaptic weight junction level between j-th and i-th neurons that Donald Hebb introduced a learning model [$W_{j,i}$] 5 decades ago. The mathematical definition has been given by McCullough-Pitts and Von Neumann introduced the concept of neurons as binary logic element as follows (Figure 10):

$$0 \leq \bar{a} = \sigma\left(\bar{X}\right) \equiv \frac{1}{1+\exp\left(-\bar{X}\right)} \leq 1; \frac{d\sigma(x)}{dx} = \bar{a}\left(1-\bar{a}\right)$$ (8a)

$$-1 \leq \bar{a} = \tan\left(i\bar{X}\right) = \frac{e^{\bar{X}} - e^{-\bar{X}}}{e^{\bar{X}} + e^{-\bar{X}}} = \frac{\sinh\left(\bar{X}\right)}{\cosh\left(\bar{X}\right)} = \tanh\left(\bar{X}\right) \leq 1;$$

$$\frac{d\tanh(x)}{dx} = 1 - \tanh(x)^2$$ (8b)

Figure 10 Nonlinear threshold logics of activation firing rates
a)   For the output classifier,
b)   Hidden layers hyperbola tangent.

We have reviewed how to appreciate the BNN. Albert Einstein's Brain has been kept after his pass away. he has 10 billion neurons just like we do, but he has 100B Glial cells that is important for house cleaning servant function to minimize Dementia Alzheimer Disease (DAD), which might have made him different from some of us. These house-keeping smaller glial cells surrounded the each neuron output called Axon that can keep positive ions vesicle move forward in a pseudo-real time which are repulsion to one another in line, as one ion is pushed in from one end of the Axon, so that those conducting positive charge ion vesicles have no way escape but line up by those insulating Glial cells in their repulsive chain in about 100 Hz, 100 ions per second, no matter how long or short the axon is. The longest axon is about 1 meter longer from the neck to the toe for the instantaneously issue the order from HVS to run away from the tiger. The insulated fatty acids, Myelin sheath, are known to be Glial cells, among those 6 types of Glial Cells.

The Glial Cells (glue force) is derived for the first time when the internal energy $E_{int.}$ is expanded as the Taylor series of the internal representation $\bar{S}_i$ related by synaptic weight matrix $\left[W_{i,j} \right]$ to the Power of the Pairs $\bar{S}_i = \left[W_{i,j} \right] \bar{X}_{pair}$ of which the slope turns out to be biological Glial cells identified by Donald O. Hebb learning rule

$$< \bar{g}_j > = -\left\langle \frac{\partial H_{int.}}{\partial \bar{D}_j} \right\rangle,$$ (9)

where the j-th Dendrite tree sum $\bar{D}_j$ of all i-th neurons whose firing rates in proportional to the internal degree of firing rate S_i called the Entropy uniformity:

$$< \bar{D}_j > = < \Sigma_i \left[W_{i,j} \right] \bar{S}_i >$$ (10)

From which we have verified Donald O. Hebb learning rule, in the *Ergodicity* ensemble average sense, who has formulated six decades ago in the brain neurophysiology. Given a time asynchronous increment $= |\Delta t|$, the learning plasticity adjustment is proportional to the pre-synaptic firing rate S_i and the post synaptic glue force $\bar{g}_j$

Theorem Asynchronous robot team and their convergence

Proof:

If & only if there exists a global optimization scalar cost function $H_{int.}$ known as Helmholtz Free Energy at isothermal equilibrium to each robot member, then each follows asynchronously its own clock time in **Newton-like** dynamics at its own time frame t_j; $\neq å_j$ 1 time causality with respect to its own initial boundary conditions with respect to the global clock time "t"

$$\frac{d[W_{i,j}]}{dt_j} = -\frac{\partial H_{int.}}{\partial[W_{i,j}]} \; ; \qquad (11)$$

Prof: The overall system is always convergent guaranteed by a quadratic A.M. Lyaponov force function:

$$\frac{dH}{dt} = \Sigma_j \frac{\partial H}{\partial[W_{i,j}]} \varepsilon_j \frac{d[W_{i,j}]}{dt_j} = -\Sigma_j \varepsilon_j \left|\frac{\partial H}{\partial[W_{i,j}]}\right|^2 \leq 0; \varepsilon_j \geq 1 \textit{time causality} \quad \textbf{Q.E.D.}$$

$$\Delta[W_{i,j}] = \frac{\partial[W_{i,j}]}{\partial t_j}\eta = -\frac{\partial H}{\partial[W_{i,j}]}\eta = -\frac{\partial H}{\Delta \bar{D}_j}\left(\frac{\partial \bar{D}_j}{\partial[W_{i,j}]}\right)\eta \equiv \bar{g}_j \bar{S}_i \eta$$

(Bilinear Hebb Rule) (12)

This Hebb Learning Rule may be extended by **chain rule** for multiple layer "Backprop algorithm" between neurons and glial cells

$$< [W_{i,j}] > = < [W_{i,j}]^{old} > + \bar{g}_j \bar{S}_i \eta \qquad (13)$$

We can conceptually borrow from Albert Einstein the space-time equivalent special relativity to trade the individual time life experience with the spatially distributed experiences gathered by Asynchronously Massively Parallel Distributed (AMPD) Computing through Cloud Image Training Databases with variety initial and boundary conditions. Also, Einstein said that "Science has nothing to do with the truth (which is a domain of theology); but the consistency." That's how we can define the Glial cells for the first time consistently Eq(8).

In this paper, we emphasized besie aerial drone surveillance, checked with counting trunks on the ground for the density ratio within the field of view of i-Phone. We could have applied **AI ANN** to do two simple tasks:

A. To recognize the face of standing redwood old man whose trunk skin is made of elongated & peal-able cells, about 3~5 inches long and 2~3 inches wide, in a twisted upward fashion.

B. To count the redwood trunks as "+1" & the other smoother Douglas fir trunk as "0+1" into each tally. As outliers, the machine had discovered of the fungus-infected trees in a clever strategy climbing up a ladder of a leaning sick trees to the slowly dying natural redwood forests in the San Francisco Bay Area.

Technically speaking, we collected three hundreds high-definition i-Phone color pictures of 3Mb each, we have a total of about 300x3Mb=0.9 billion parameters. It turns out to be those trunk textures that are twisted upward with elongated elliptic brown cells, not just the size of ERW. We must prone all those irrelevant and confused canopy and bush input info, and keep only the relevant tree truck info. In other words, if we could have fed the Machine Learning with garbage in, then we get the predictable garbage out reliably every time, of which the common sense must be derived by a lot of experience that may not necessarily be gained by an individual robot, automatically fact-find from i-Phone image database, say about Redwood forest in the secluded Woodside of San Francisco Bay Area. Furthermore, we challenge the machine learning to derive an empirical observation to count the symbiotic density ratio in a healthy forest among Endangered Redwood (ERW) as "class #1" and the other trees such as Comparable Douglas-firs (CDF) as "Class #0." Firstly, we downsize the training data by cutting the top canopies and the bottom bushes off as the confused and irrelevant 2/3 information, and kept only the middle trunks in the gray-scale non-color imageries as the relevant input. As a result, we discovered the symbiotic co-existence ratio **<X>** between **ERW** and **CDF** to be 20%(1:5)≤**<X>**≤33%(1:3).

This result is important because the original intension of forest managers has been planting the rapidly growing CDF to help holding down the exposed ERW roots after "selective logging;" the unexpected back fire has been the fast CDF which can inhibit those slow growth ERW furthermore slower. Moreover, we discovered an important outlier as follows: In spite of the evolutional advantages, the ERW have been able to against the fire, insects, and fungi e.g. ground mushroom. The outlier of the binary classifier has implicated the unreported strategy of the different kind of plant cell growth. Namely, the fungi utilized the weaken dead trees to be" leaning" laddering tree onto ERW for their constant and stealthy attack to propagate fungi.

Historical perspective may help reader understand about the endangering Redwood Forest. During April 18, 1906, San Francisco earthquake struck the coast of Northern California at 5:12a.m. with an estimated Richter scale of 7.8magnitude. To appreciate the calamity, we shall compare it with another earthquake on May 12, 2008 a Century later in Sichuan, China at the scale of 7.1 Richter (Figure 11).

Microscopically R/D of genetic and epigenetic of Redwood in order to transfer the lesson g to enhance the other trees for fire retardation if possible

Redwoods can last around 240 million years and some in CA for at least 20 million years, compared to about 200,000 years for "modern" humans. It has many virtues. Redwoods live so long–and are treasured by humans for building – because they are extremely resistant to insects, fire and rot. At one time, San Francisco's building codes required redwood lumber to be used in the foundations of new structures. A redwood's bark can be one foot thick, and it contains tannin which protects the tree from fire, insects, fungus and diseases.

Exemplar: It may be a well-known fact that the planting of those rapidly growth Comparable Douglas-fir (CDF), as in Figure 3, have protected the up-rooting of redwood trees from strong wind storms; but there as unexpected side effect, i.e. the CDF trees can also inhibit the growth of redwood trees. We shall separate the Irrelevant info from the Relevant info as follows (Figure 12):
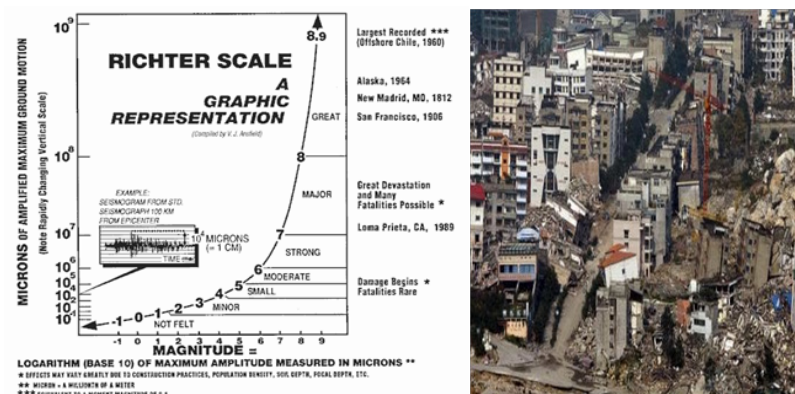
Figure 11

a) World Major Earthquakes; High intensity shaking was felt from Eureka on the North Coast to the Salinas Valley, an agricultural region to the South of the San Francisco Bay Area. Devastating fires soon broke out in the city and lasted for several days. As a result, up to 3,000 people died and over 80% of the city of San Francisco was destroyed.

b) three gorges on the Yangtze River" This disaster is nothing to do with relative primitive kitchens in April 18, 1906, because it should be compared against a hundred & two years later in May 12, 2008 at one of the richest Sichuan Province located in the middle of the, the third longest in the world, in the Hubei Province of China in the geologically stable Central Tectonic Plate. China suffered about the same Richet scale 8.0 earthquake. Despite of Sichuan is located in the Central Tectonic Plate, the land has been suffered by the heavy water weight of Three Gorge Dam (TGD) reservoir and generated the unstable Saddle Point Instability. Since TGD was built in1994 & its operation in full 14 years later. There were more than 1.2 million people have been displaced and the number of 13 cities and 140 towns and 1350 villages flooded, and length of reservoir is more than 600 kilometers. There were over 69,000 people lost their lives in the quake, including 68,636 in Sichuan province. 374,176 were reported injured, with 18,222 listed as missing in actions. The Sichuan Earthquake left about 4.8 million people homeless. Should there were i-Watches life of many people might be saved. At that time, San Francisco's building codes required redwood lumber to be used in the foundations of new structures.
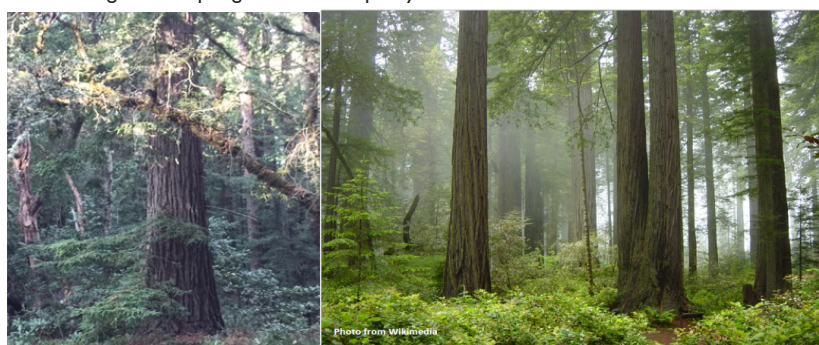

(A)   (B)   (C)

Applying AI Unsupervised Learning to monitpring the macroscopic dynamics of redwood forest

(D)   (E)

E) We first recognizes trunk faces, **Endangering Redwood** versus **Comparable Douglas-fir** (f) We counts each type the number of trees. (g) We compute their ratio, e.g. whether one **endangered redwood** and 3 or 5 **comparable doulas-fir** in the Field of View. $20\% << X_{ERW} > < 30\%$. The job is how to do all (e,f,g) steps by i-Phone and (h) conclude what is a healthy forest ratio $< X_{ERW} >$ automatically in years time.

Figure 12

A. To count the types of trees, we shall verify the redwood trucks, as if human face identification;

B. We shall take the hand engineering pre-processing by cutting off top canopy, and the bottom bush roots;

C. We then apply the well established MPD SIMD Python-3 AlexNet-2012 algorithm, which can increase the trust and thus satisfy DARPA XAI program.[7]

How can we conclude what is a healthy forest ratio $< X_{ERW} >$ automatically; (e) Only symbiotic evolution between ERW and CDF may be computed.

The hand-on engineering is necessary to get rid of irrelevant and keep the relevant in the same size after zero-padding, as the pre-processing. Then, how to determine statistically by majority what is a healthy balance $\langle X_{ERW} \rangle$? The Allen Road can range from 3 **ERW's** per 7 **CDF's**: X=30% to 1 **ERW** per 9 CDF: X=20%. Trunks have no distinguished colors but texture from aged rugged twisted **ERW's** bucks to younger smoother but cracked **CDF's** bucks Thus, we chose at the well preserved by sparse populated 16 mansions along 1.5 mile Allen Road of Woodside from #200 beginning to #486 lion

Park to look for the evidence what will be a healthy density $\langle X_{ERW} \rangle$ between the redwood and Comparable Douglas-fir trees.

As a surprising result, we discovered the outlier of Automatic Pattern Recognition (APR) as few trees infested by fungi by a strategy not known before, fungi following a leaning ladder climbing up from the bushes to the tree trunk which would be slowly killing the redwood forest after the tannin has been used up by fungi in San Francisco Bay Area (Figure 13&14).
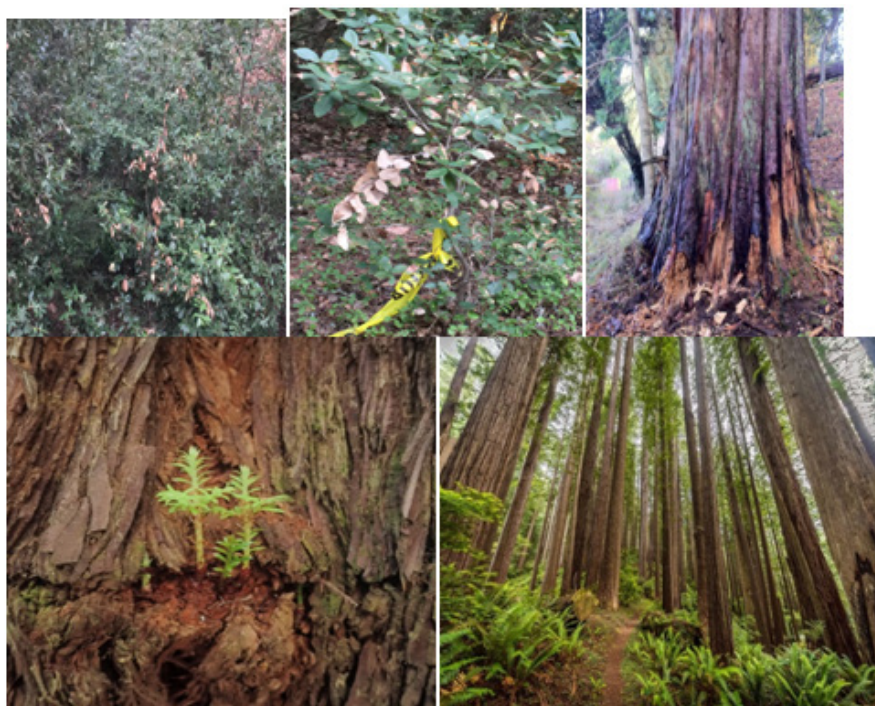

Figure 13

(a) Outliers from the i-Phone with AI ANN Deep Learning code solving the binary classifier problem, defined as "1" counting ERW's after Identification in the Field Of View, and "0" for Identification of CDF's.

(b) & (c) It turns out that

(d) The outliers are the fungi infected trees *ladder and leaning* against GF & ERW.

Figure14

a) A close-up view reveals the source might be those bottom bushes trespassing by dears et al.

b) Noticed by "Yellow Ribbon" tied by ranger's.

c) Fungi disease;

d) Stump re-growth;

e) Healthy ERW Nat'l State Park; but not the private reserve Benjamin Allen Road about 1.5 mile long where 12 Mansions occupied about 200 Acres Redwood Forest from the Bear Gulch Road to Le Honda Creek wild mountain reserve.

What is fungus and are fungi, like mushroom, and why they, as parasites, are so harmful? We quote first "(1) Virus: With the exception of newly discovered prions, viruses are the smallest agents of infectious disease. Most viruses are exceedingly small (about 20-200nanometers in diameter) and essentially round in shape. that cannot replicate outside of a host cell. Because viruses do not eat food - instead they seize materials and energy from host cells by hijacking cellular machinery; (2) Bacteria are ten to 100 times larger than viruses. They are typically 1 to 3 microns in length and take the shape of a sphere or rod. Most bacteria consist of a ring of DNA surrounded by cellular machinery, all contained within a fatty membrane. They acquire energy from the same essential sources as humans, including sugars, proteins, and fats. Some bacteria live and multiply in the environment while others are adapted to life within human or animal hosts; (3) Parasites are part of a large group of organisms called eukaryotes. Parasites are different from bacteria or viruses because their cells share many features with human cells including a defined nucleus;(4) Fungi are diverse in terms of their shape, size and means of infecting humans. Fungi are eukaryotes, meaning that like parasites, their cells have a true nucleus and complex internal structures. They are most commonly found as environmentally resistant spores and molds, but can cause disease in humans in the form of yeasts. Fungi most often cause skin infections and pneumonia. Fungal diseases are particularly dangerous to immuno-compromised people, such as those suffering from AIDS."

The well-established literature about Fungi and Trees; "(1) Spot Anthracnose that impacts tree leaves and blossoms, and;(2) the more harmful canker versions that disrupt a tree's vascular system. When Spot Anthracnose initially emerges, small light brown spots of dead tissue emerge on the leaves and blossoms in the late spring and summer. The spots develop during the cool, wet humid spring weather. Often the brown dieback is located along the leaf's vein system. The disease is caused by a fungus that over-winters on the bark of the tree or on fallen leaves. In the spring, the fungus is spread by rain and wind, which transports the fungus to newly developing healthy leaves and blossoms. When the tree is severely infected over several seasons the fungus will infect and kill branches." Then, we pointed out in this paper, we discovered another mechanism of transportation depicted in Figure 10. It turns out that the outliers are the fungi infected trees ladder and leaning against GF & ERW so that can concentrate on the contact, and penetration to wear out tannin chemical responsible for th

Microscopically genetic and epigenetic approach to passover redwood genome[4,5]

IT'S BEEN NEARLY two decades since high-throughput sequencing escaped the labs and went commercial. Today, you can get your whole genome for just $1,000 (quite a steal compared to the $1.5 million it cost to sequence James Watson's in 2008). But the data produced by today's machines still only produce incomplete, patchy, and glitch-riddled genomes. Errors can get introduced at each step of the process, and that makes it difficult for scientists to distinguish the natural mutations that make you as you from random artifacts, especially in repetitive sections of a genome. See, most modern sequencing technologies work by taking a sample of your DNA, chopping it up into millions of short snippets, and then using fluorescently-tagged nucleotides to produce reads—the list of As, Ts, Cs, and Gs that correspond to each snippet. Then those millions of reads have to be grouped into abutting sequences and aligned with

a reference genome. From there they can go on to variant calling—identifying where an individual's genes differ from the reference.1 A number of software programs exist to help do that. Free-Bayes, Var-Dict, Sam-tools, and the most well-used, GATK, depend on sophisticated statistical approaches to spot mutations and filter out errors. Each tool has strengths and weaknesses, and scientists often wind up having to use them in conjunction.

No one knows the limitations of the existing technology better than Mark DePristo and Ryan Poplin. They spent five years creating GATK from whole cloth. This was 2008: no tools, no bioinformatics formats, no standards. "We didn't even know what we were trying to compute!" says DePristo. But they had a north star: an exciting paper that had just come out, written by a Silicon Valley celebrity named Jeff Dean. As one of Google's earliest engineers, Dean had helped design and build the fundamental computing systems that underpin the tech titan's vast online empire. DePristo and Poplin used some of those ideas to build GATK, which became the field's gold standard.

But by 2013, the work had plateau. "We tried almost every standard statistical approach under the sun, but we never found an effective way to move the needle," says DePristo. "It was unclear after five years whether it was even possible to do better." DePristo left to pursue a Google Ventures-backed start-up called SynapDx that was developing a blood test for autism. When that folded two years later, one of its board members, Andrew Conrad (of Google X, then Google Life Sciences, then Verily) convinced DePristo to join the Google/Alphabet fold. He was reunited with Poplin, who had joined up the month before.

And this time, Dean wasn't just a citation; he was their boss. As the head of Google Brain, Dean is the man behind the explosion of neural nets that now prop up all the ways you search and tweet and snap and shop. With his help, DePristo and Poplin wanted to see if they could teach one of these neural nets to piece together a genome more accurately than their baby, GATK. The network wasted no time in making them feel obsolete. After training it on benchmark datasets of just seven human genomes, Deep Variant was able to accurately identify those single nucleotide swaps 99.9587 percent of the time. "It was shocking to see how fast the deep learning models outperformed our old tools," says DePristo. Their team submitted the results to the Precision FDA Truth Challenge last summer, where it won a top performance award. In December, they shared them in a paper published on bio Rxiv. DEEPVARIANT WORKS BY transforming the task of variant calling—figuring out which base pairs actually belong to you and not to an error or other processing artifact—into an image classification problem. It takes layers of data and turns them into channels, like the colors on your television set. In the first working model they used three channels: The first was the actual bases, the second was a quality score defined by the sequencer the reads came off of, the third contained other metadata. By compressing all that data into an image file of sorts, and training the model on tens of millions of these multi-channel "images," Deep Variant began to be able to figure out the likelihood that any given A or T or C or G either matched the reference genome completely, varied by one copy, or varied by both. But they didn't stop there. After the FDA contest they transitioned the model to Tensor Flow, Google's artificial intelligence engine, and continued tweaking its parameters by changing the three compressed data channels into seven raw data channels. That allowed them to reduce the error rate by a further 50 percent. In an independent analysis conducted this week by genomics computing platform, DNA

nexus, Deep Variant vastly outperformed GATK, Freebayes, and Sam tools, sometimes reducing errors by as much as 10-fold. "That shows that this technology really has an important future in the processing of bio-informatic data," says DNAnexus CEO, Richard Daly. "But it's only the opening chapter in a book that has 100 chapters." Daly says he expects this kind of AI to one day actually find the mutations that cause disease. His company received a beta version of Deep Variant, and is now testing the current model with a limited number of its clients—including pharma firms, big health care providers, and medical diagnostic companies. The coast redwoods, ancient and threatened, mix with towering Douglas firs and opportunistic tanoaks throughout this restoration project on mountain top just miles from the sea. As California climate changes to one of the extremes and human continue to harvest, the only coast redwoods on the planet are in peril.

Recently, Dr. Emily Burns, UC Berkeley, Dir. Of Sci. Save the Redwoods League, IRS 301C,[6] is funding (at the tone of 1 millions) the (initial) effort by Prof. David Neale of UC Davis and Prof. Steve Salzberg of Johns Hopkins to map the massive genetic code of coast redwoods--a genetic code 38 billions 12 times larger than that of a human being 3 billion base pairs of DNA on its chromosomes. Prof. David Neale, dbneale@ucdavis.edu, primary research interest is in the discovery and understanding of function of genes in forest trees, especially those controlling complex traits, through genetic mapping and genomic science technologies. In the very near future, nearly all the genes in pine will have been cloned and their DNA sequences determined.[7]

The cloned genes will serve as tools for functional analysis which will lead to understanding of complex patterns of coordinated expression of genes leading to phenotype. Initially, we have focused on traits of practical value such as wood quality, growth, and disease resistance but we expect to expand our studies to identifying genes determining patterns of adaptation and response to environmental stresses.

The bigger the genetic code, the more can go wrong, and much of what the genome contains is unnecessary. The same is true of humans. But redwoods do live a fantastically long life, and it would be fascinating to discover why. No other tree comes close to absorbing more carbon than the redwood, making these forests invaluable in reducing greenhouse gases. (Just 5 percent of the redwoods that stood before 1849 are still alive) Saving them seems like a better investment than ever. (The genetic code of a single 1,300-year-old redwood and of a same-age sequoia will help understand how to [defend] forests against man-made threats. Scientists will [eventually] have the genetic fingerprints of hundreds of redwoods, enough to determine which trees best withstand increased moisture or drought, heat increases or temperature drops). What by the end of this year, scientists hope to have mapped the complete genome of the coast redwood" Decades later Bay Area Building Ordinance and the larger California Building Code (CBC) recommended the use of Redwood containing reddish tannin (poly-galloyl glucoses) against fire, insects and fungi. Thus, the tree lived the longest life (millions years since dinosaurs), and grew as the tallest (300 feet extending their intertwine roots more than 50 feet from the trunk) on the Earth. Thus Pacific Coast RW went through a massive "clean cuts" (remove stumps for tax saving) in decades. Environmentalists began serious concern about use of Redwoods e.g. notably Sempervirens Fund (SF) (Figure 15).[3]

Figure 15 The characteristics of ERW trunk reveals reddish "tannin" chemical. "Clear-cutting of redwood forest was generally the preferred method, for the state tax laws encouraged the cutting of 70% or more of the trees in a stand so as to remove standing timber from the tax rolls. The increased soil disturbance and lack of successful growth of redwoods from seeds sometimes encouraged the growth of other associated species such as Douglas-fir and hemlock, so companies sometimes supplemented the anticipated regeneration from stump sprouting by planting redwood seedlings. The logging industry was extremely important. in the redwood region in the 1950s and1960s, particularly in Del Norte and Humboldt Counties. In 1953, for example, the six largest timber operators paid over 40% of the taxes in Del Norte County, while about half of Humboldt County private-sector employees who had unemployment insurance in the 1950s worked in the timber/lumber industry. In the 1950s, moreover, the per-capita income in Humboldt County was comparable to the rest of California."

Likewise, wildfire can happen in the global warning dry seasons in California Sierra. It can be caused by thunderstorms or human errors by camp fires or household "BBQ" fire should be outlawed as well down power line which shall be clear cut surrounding the power-line pathway, Furthermore, the park rangers must built fire isolation lanes cross-cut all the forests, and more helicopter landing ports, as well as equipped with modern & scientific approach Artificial Intelligence (AI) to survey quickly and fundamentally prevent any of the future forest fire For example, we wish to load Helicopter with surveillance video camera operated to flyover with AI machine learning image processing & image understanding to surveillance quickly and efficiently high over the redwood forest; The redwood forest protection becomes a renewed effort to help catch potential camp fire early on using the standard deep learning novelty detection of big data analysis (BDA) (Figure 16). The goal is to enable scientists to maintain forest resiliency and genetic diversity by choosing the most robust, adaptable genes when planting or doing regeneration or habitat protection work.

"The cloned genes will serve as tools for functional analysis which will lead to understanding of complex patterns of coordinated expression of genes leading to phenotype. Initially, we have focused on traits of practical value such as wood quality, growth, and disease resistance but we expect to expand our studies to identifying genes determining patterns of adaptation and response to environmental stresses. The bigger the genetic code, the more can go wrong, and much of what the genome contains is unnecessary. The same is true of humans.

But redwoods do live a fantastically long life, and it would be fascinating to discover why" (cf. Appendix A)

We conjectured that the decade large size might be due to decade longer life of redwood. Nature might take a piece-meal bandage

approach, when more and more genetic code are added on to survive the stress of Nature.

Figure 16 Redwood Genome Program ($1M), Prof. David Neale of UC Davis; Prof. Steve Salzberg of Johns Hopkins. Coast redwood tree. (photo: Ann Filmer/UC Davis) (d) *Can Genome Save the Forest?* The five-year, $2.6 million Redwood Genome Project is the most intensive scientific study ever done on the state's famous primeval forests.

Why DNA of Redwood is 10 times larger than Homo sapiens? So it seems to be that the redwood life spans are also decade longer. One thing is for sure that the trees cannot move but to stand against the adversary conditions, e.g. they develop the most absorbent of carbon dioxides. They evolve the genome to support epiphenomena for survival the fittest. A genetic code 38 billions 12times larger than that of a human being 3billion base pairs of DNA on its chromosomes.

A. Hybrid of two species redwood and giant sequoia

B. Bandage genome approach over million years evolution in order to secrete the red tarnish chemical **tannin** (poly-galloyl glucoses) (Figure 17).

The redwood is one of three living species, each in its own genus, in the subfamily Sequoioideae. Molecular studies have shown the three to be each other's closest relatives, generally with the redwood and giant sequoia (Sequoiadendron giganteum) as each other's closest relatives. However, Yang and colleagues in 2010 queried the polyploid state of the redwood and speculate that it may have arisen as an ancient hybrid between ancestors of the giant sequoia and dawn redwood (Metasequoia). Using two different single copy nuclear genes, LFY and NLY, to generate phylogenetic trees, they found that Sequoia was clustered with Metasequoia in the tree generated using the LFY gene, but with Sequoiadendron in the tree generated with the NLY gene. Further analysis strongly supported the hypothesis that Sequoia was the result of a hybridization event involving Metasequoia and Sequoiadendron.
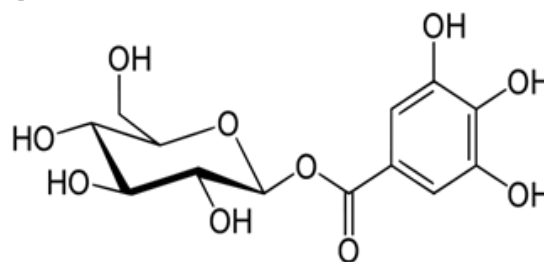
Figure 17 Tanin ($C_{76}H_{52}O_{46}$ polygalloyl glucoses) that can incidentally prevent influenza A entering the cell, and help to stand against fire, insect, fungus, etc.. The redwood contains the natural acidity chemical tannin ($C_{76}H_{52}O_{46}$ polygalloyl glucoses) for evolutionally advantages against fire, insect, & fungi.

No other tree comes close to absorbing more carbon than the redwood, making these forests invaluable in reducing greenhouse gases. (Just 5 percent of the redwoods that stood before 1849 are still alive) Saving them seems like a better investment than ever. (The genetic code of a single 1,300-year-old redwood and of a same-age sequoia will help understand how to [defend] forests against man-made threats. Scientists will [eventually] have the genetic fingerprints of hundreds of redwoods, enough to determine which trees best withstand increased moisture or drought, heat increases or temperature drops.)…What by the end of this year, scientists hope to have mapped the complete genome of the coast redwood…" TODAY, A TEASPOON of spit and a hundred bucks is all you need to get a snapshot of your DNA. But getting the full picture—all 3 billion base pairs of your genome—requires a much more laborious process. One that even with the aid of sophisticated statistics scientists still struggle over. It's exactly the kind of problem that makes sense to outsource to artificial intelligence.

On Monday, Google released a tool called Deep Variant that uses deep learning—the machine learning technique that now dominates AI—to identify all the mutations that an individual inherits from their parents.1 Modeled loosely on the networks of neurons in the human brain, these massive mathematical models have learned how to do things like identify faces posted to your Facebook news feed, transcribe your inane requests to Siri, and even fight internet trolls. And now, engineers at Google Brain and Verily (Alphabet's life sciences spin-off) have taught one to take raw sequencing data and line up the billions of As, Ts, Cs, and Gs that make you as you.

And oh yeah, it's more accurate than all the existing methods out there. Last year, Deep Variant took first prize in an FDA contest promoting improvements in genetic sequencing. The open source version the Google Brain/Verily team introduced to the world Monday reduced the error rates even further—by more than 50 percent. Looks

like grandmaster Ke Jie isn't be the only one getting bested by Google's AI neural networks this year.

Deep Variant arrives at a time when healthcare providers, pharma firms, and medical diagnostic manufacturers are all racing to capture as much genomic information as they can. To meet the need, Google rivals like IBM and Microsoft are all moving into the healthcare AI space, with speculation about whether Apple and Amazon will follow suit. While Deep Variant's code comes at no cost, that isn't true, of the computing power required to run it. Scientists say that expense is going to prevent it from becoming the standard anytime soon, especially for large-scale projects. But Deep Variant is just the front end of a much wider deployment; genomics is about to go deep learning. And once you go deep learning you don't go back. You Can Get Your Whole Genome Sequenced. But Should You? (SARAH ZHANG) Cheap DNA Sequencing Is Here. Writing DNA Is Next. Helix's Bold Plan to Be Your One Stop Personal Genomics Shop. To run Deep Variant effectively for these customers, DNAnexus has had to invest in newer generation GPUs to support its platform. The same is true for Canadian competitor, DNA Stack, which plans to offer two different versions of Deep Variant—one tuned for low cost and one tuned for speed. Google's Cloud Platform already supports the tool, and the company is exploring using the TPUs (tensor processing units) that connect things like Google Search, Street View, and Translate to accelerate the genomics calculations as well. Deep Variant's code is open-source so anyone can run it, but to do so at scale will likely require paying for a cloud computing platform. And it's this cost—computationally and in terms of actual dollars—that have researchers hedging on Deep Variant's utility. "It's a promising first step, but it isn't currently scalable to a very large number of samples because it's just too computationally expensive," says Daniel MacArthur, a Broad/Harvard human geneticist who has built one of the largest libraries of human DNA to date. For projects like his, which deal in tens of thousands of genomes, Deep Variant is just too costly. And, just like current statistical models, it can only work with the limited reads produced by today's sequencers. Still, he thinks deep learning is here to stay. "It's just a matter of figuring out how to combine better quality data with better algorithms and eventually we'll converge on something pretty close to perfect," says MacArthur. But even then, it'll still just be a list of letters. At least for the foreseeable future, we'll still need talented humans to tell us what it all means (Figure 18).
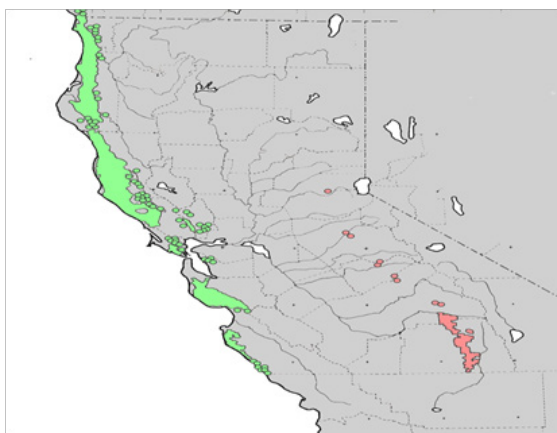
Figure 18 Distribution map for Sequoia sempervirens (coast redwood) and Sequoiadendron giganteum (giant sequoia) until human bulldoze massive deforestation.

Conclusion

The Machine learning can not find $\langle X_{redwood} \rangle$ unless we hand-engineering cutting off the top canopy and the bottom bushes as in Fig. The redwood distribution is given in Figure 17. Today, a teaspoon of spit and a hundred bucks is all you need to get a snapshot of your DNA. But getting the full picture—all 3 billion base pairs of your genome—requires a much more laborious process. With the aid of sophisticated statistics scientists still struggle over. It's exactly the kind of problem that makes sense to outsource to artificial intelligence. On Monday, Google released a tool called Deep Variant that uses deep learning on—the machine learning technique that now dominates AI—to identify all the mutations that an individual inherits from their parents.[1] Modeled loosely on the networks of neurons in the human brain, these massive mathematical models have learned how to do things like identify faces posted to your Facebook news feed, transcribe your inane requests to Siri, and even fight internet trolls. And now, engineers at Google Brain and Verily (Alphabet's life sciences spin-off) have taught one to take raw sequencing data and line up the billions of As, Ts, Cs, and Gs that make you as you. And oh yeah, it's more accurate than all the existing methods out there. Last year, Deep-Variant took first prize in an FDA contest promoting improvements in genetic sequencing. The open source version the Google Brain/Verily team introduced to the world Monday reduced the error rates even further—by more than 50 percent. Looks like grandmaster Ke Jie isn't be the only one getting bested by Google's AI neural networks this year.

Deep-Variant arrives at a time when healthcare providers, pharma firms, and medical diagnostic manufacturers are all racing to capture as much genomic information as they can. To meet the need, Google rivals like IBM and Microsoft are all moving into the healthcare AI space, with speculation about whether Apple and Amazon will follow suit. While Deep Variant's code comes at no cost, that isn't true, of the computing power required to run it. Scientists say that expense is going to prevent it from becoming the standard anytime soon, especially for large-scale projects. But Deep Variant is just the front end of a much wider deployment; genomics is about to go deep learning. And once you go deep learning you don't go back. It's been nearly two decades since high-throughput sequencing escaped the labs and went commercial. Today, you can get your whole genome for just $1,000 (quite a steal compared to the $1.5 million it cost to sequence James Watson's in 2008). But the data produced by today's machines still only produce incomplete, patchy, and glitch-riddled genomes. Errors can get introduced at each step of the process, and that makes it difficult for scientists to distinguish the natural mutations that make you as you from random artifacts, especially in repetitive sections of a genome. See, most modern sequencing technologies work by taking a sample of your DNA, chopping it up into millions of short snippets, and then using fluorescently-tagged nucleotides to produce reads—the list of As, Ts, Cs, and Gs that correspond to each snippet. Then those millions of reads have to be grouped into abutting sequences and aligned with a reference genome. From there they can go on to variant calling—identifying where an individual's genes differ from the reference.[1] A number of software programs exist to help do that. Free-Bayes, Var-Dict, Sam-tools, and the most well-used, GATK, depend on sophisticated statistical approaches to spot mutations and filter out errors. Each tool has strengths and weaknesses, and scientists often wind up having to use them in conjunction. No one knows the limitations of the existing technology better than Mark DePristo and

Ryan Poplin. They spent five years creating GATK from whole cloth. This was 2008: no tools, no bioinformatics formats, no standards. "We didn't even know what we were trying to compute!" says DePristo. But they had a north star: an exciting paper that had just come out, written by a Silicon Valley celebrity named Jeff Dean. As one of Google's earliest engineers, Dean had helped design and build the fundamental computing systems that underpin the tech titan's vast online empire. DePristo and Poplin used some of those ideas to build GATK, which became the field's gold standard.

But by 2013, the work had plateau. "We tried almost every standard statistical approach under the sun, but we never found an effective way to move the needle," says DePristo. "It was unclear after five years whether it was even possible to do better." DePristo left to pursue a Google Ventures-backed start-up called SynapDx that was developing a blood test for autism. When that folded two years later, one of its board members, Andrew Conrad (of Google X, then Google Life Sciences, then Verily) convinced DePristo to join the Google/ Alphabet fold. He was reunited with Poplin, who had joined up the month before. And this time, Dean wasn't just a citation; he was their boss. As the head of Google Brain, Dean is the man behind the explosion of neural nets that now prop up all the ways you search and tweet and snap and shop. With his help, DePristo and Poplin wanted to see if they could teach one of these neural nets to piece together a genome more accurately than their baby, GATK. The network wasted no time in making them feel obsolete. After training it on benchmark datasets of just seven human genomes, Deep Variant was able to accurately identify those single nucleotide swaps 99.9587 percent of the time. "It was shocking to see how fast the deep learning models outperformed our old tools," says DePristo. Their team submitted the results to the Precision FDA Truth Challenge last summer, where it won a top performance award. In December, they shared them in a paper published on bioRxiv.

DEEPVARIANT WORKS BY transforming the task of variant calling—figuring out which base pairs actually belong to you and not to an error or other processing artifact—into an image classification problem. It takes layers of data and turns them into channels, like the colors on your television set. In the first working model they used three channels: The first was the actual bases, the second was a quality score defined by the sequencer the reads came off of, the third contained other metadata. By compressing all that data into an image file of sorts, and training the model on tens of millions of these multi-channel "images," Deep Variant began to be able to figure out the likelihood that any given A or T or C or G either matched the reference genome completely, varied by one copy, or varied by both. But they

didn't stop there. After the FDA contest they transitioned the model to Tensor Flow, Google's artificial intelligence engine, and continued tweaking its parameters by changing the three compressed data channels into seven raw data channels. That allowed them to reduce the error rate by a further 50 percent. In an independent analysis conducted this week by genomics computing platform, DNAnexus, Deep Variant vastly outperformed GATK, Freebayes, and Sam tools, sometimes reducing errors by as much as 10-fold (Figure 19).

Results showed that the density $\langle X_{redwood} \rangle$ is ranging from 1 to 3, or 1 to 5 ratio of the Redwood along the Woodside ridge called the Allen Road The ratio of redwood trees among Douglas-fir and Oaks varies 3:1 to 5:1 from the North East Le Honda Creek Park the end 420 Allen Road toward the South West of 200 Allen Road (from Right Hand downhill Side) to 201 Allen Road (Left Hand uphill side). There are mansions on both sides in increment 20 numbers, occupied about 5~10 acres woody lands. The younger trees less than 5'~10' in diameters are not counted, because when the pixel on target (POT) of i-Phone photo is less than 10 pixels, the "ConvNet" edge filter is overlapping and texture identification under variable sun light is not accurate enough for both sides of the tree. The important feature of tree truck are the burning & cracking dark red balk of redwood versus Douglas-fir balk dark brown with more smooth integrity. The ruminant pine fir leaves by the tree truck are an important indicator of separate the Douglas-fir versus elliptic-shape Oaks leave (Figure 20).
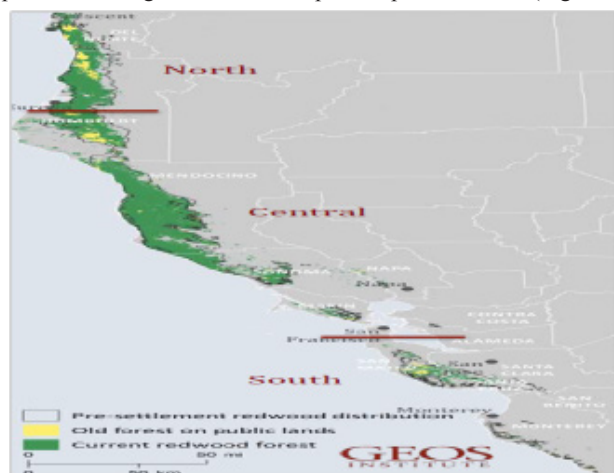
Figure19 According GEOS, the current locations of redwood forests in Pacific Coast. Note that on the scale, Woodside near San Francisco Bay area is barely discernible. Woodside overlooking Pacific Bay area healthy forest, but fungi using climbing ladder strategy will might take over as a dead forest.
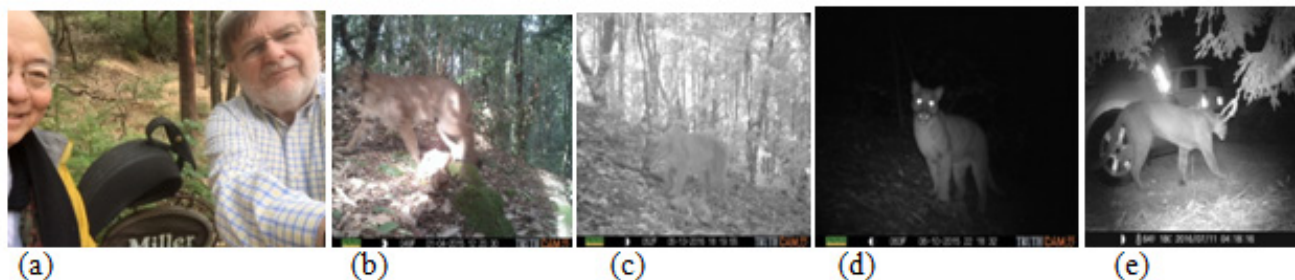
Figure 20 (a) A local friendly consultant Mr. David Miller and (b) (c) (d) & (e) are his kept record of day & night movements of mountain lions.

Unfortunately, the upcoming air pollution of industrialization blanketed the solar heat within the Atmosphere, and the accumulated moisture can generate the global warming "heavy raining and big mixing windy storms" that exasperated the soil loss and uproot loner ERW tree. In order to compensate the "clean-cut" phenomena, the management pass the ordinance of "selective logging" (Figure 21).

Figure 21

a)   Industrialization chimney; and

b)   the nickname of US National flower might be chosen "Car-nation," flower to represent the United States in the Inter State High Ways;

c)   We cannot help suffering from Carbon monoxides air pollution, causing heavy raining soil erosion and frequent windstorms up-rooted the tall ***ERW's***.

d)   The complex 50 feet ***ERW*** root is illustrated. "Over 137,000 acre so fold-growth forest remain with 89,000 acres being held in protected status with parks and preserves. Growth of redwood is estimated to be 658 million board feet per year during the period 1984 to1986. Harvest of redwood from Redwood Region forest save raged 719 million board feet per year during the period 1983-1987. While the knot free, fine grained old-growth redwood is prized in the market place, products from young-growth trees are rapidly replacing the finding ready acceptance among redwood users."

Figure 22 The special waist watch can help identify the location and health status of a wearer through the watch back-panel for heart beats measuring sensor. i-Cloud link the same data accessible from Wi Fi Hot spot all Apple/Samsung/LG devices, esp. i-Phone camera inputs.

Conflict of interest

Authors declare that there is no conflict of interest.

References

1.   Szu Harold. MO imagery techniques using arrays of large aperture telescopes. *Optics Communications*. 1980;32(2):229–234.

2.   Harold H Szu, James R Buss, Joseph P Garcia, et al. Subpixel jitter video restoration on board of micro-UAV, SPIE Proceeding 5439, Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks II.

3.   Wikipedia Sempervirens Fund.

4.   Scott Wilson. "Decoding the redwood" The Washington Post, Thursday, Feb 8, 2018, Sect. B8 cf. Appendix A.

5.   David Neale of UC Davis, dbneale@ucdavis.edu, Prof. Steve Salzberg of Johns Hopkins.

6.   Dr. Emily Burns, UC Berkeley, Dir. Of Sci. Save the Redwoods League

7.   California Wildfires. "Fixing state's fire problem: Costly, complex, next to impossible," Kurtis Alexander San Francisco Chronicle. 2018.