62nd ISI WORLD STATISTICS CONGRESS 2019
18 – 23 August 2019, Kuala Lumpur
Come | Connect | Create

**IPS 35**

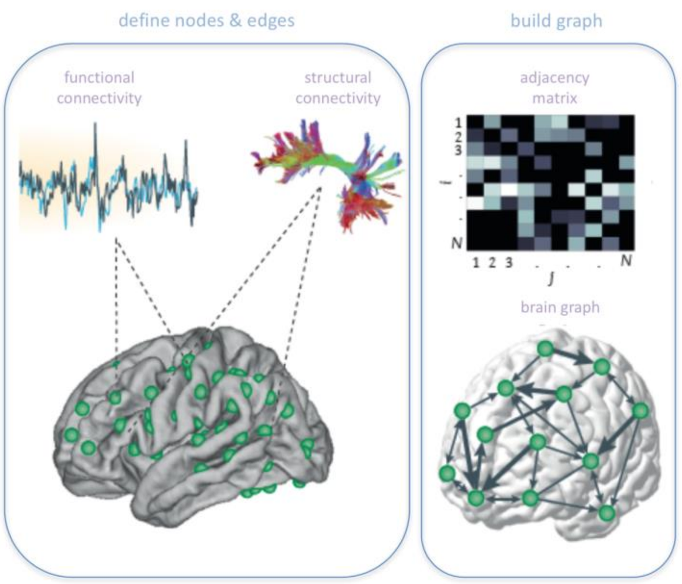
**Hypothesis Testing on Neuroimaging Brain Networks with Deep Generative Neural Networks**

Eric Tatt Wei Ho
Department of Electrical & Electronics Engineering
Universiti Teknologi PETRONAS, Malaysia
19th August 2019, 2.00 pm – 3.40 pm

# Outline/Content

1. Introduction: Brain networks
2. Problem Statement : high dimensionality & multiple comparisons
3. Generative Deep Neural Network (GDNN) Models
4. Model Formulation with GDNN
5. Hypothesis Testing with GDNN
6. Model Accuracy & Utility of Diverse GDNN Models
7. Challenges for Future Work

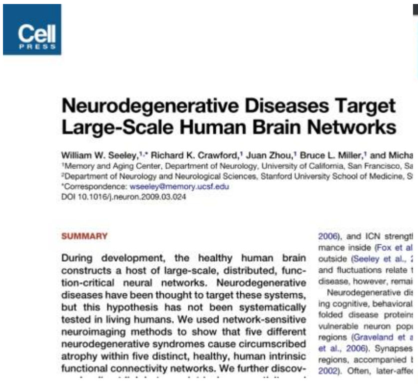
# Brain Networks




*Slide from Alex Fornito, Monash Univ*



# Brain Networks, Degeneration & Disease



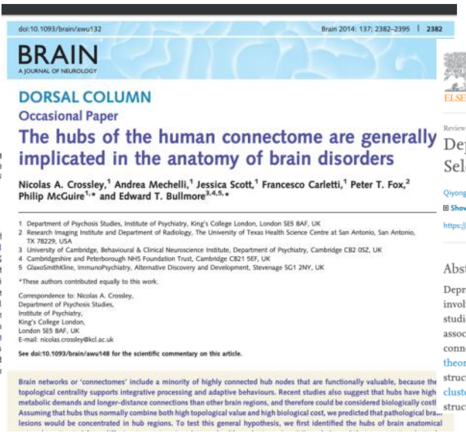
Neurodegenerative Diseases Target Large-Scale Human Brain Networks

William W. Seeley,[1,*] Richard K. Crawford,[1] Juan Zhou,[1] Bruce L. Miller,[1] and Micha
[1]Memory and Aging Center, Department of Neurology, University of California, San Francisco, Sa
[2]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, S
*Correspondence: wseeley@memory.ucsf.edu
DOI 10.1016/j.neuron.2009.03.024

SUMMARY

During development, the healthy human brain constructs a host of large-scale, distributed, function-critical neural networks. Neurodegenerative diseases have been thought to target these systems, but this hypothesis has not been systematically tested in living humans. We used network-sensitive neuroimaging methods to show that five different neurodegenerative syndromes cause circumscribed atrophy within five distinct, healthy, human intrinsic functional connectivity networks. We further discov-

2006), and ICN strengt
mance inside (Fox et al
outside (Seeley et al., 2
and fluctuations relate t
disease, however, remai
Neurodegenerative dis
ing cognitive, behavioral
folded disease proteins
vulnerable neuron popu
regions (Graveland et a
et al., 2006). Synapses
regions, accompanied b
2002). Often, later-affec



doi:10.1093/brain/awu132 Brain 2014: 137; 2382–2395 | 2382

BRAIN
A JOURNAL OF NEUROLOGY

DORSAL COLUMN
Occasional Paper

The hubs of the human connectome are generally implicated in the anatomy of brain disorders

Nicolas A. Crossley,[1] Andrea Mechelli,[1] Jessica Scott,[1] Francesco Carletti,[1] Peter T. Fox,[2] Philip McGuire[1,*] and Edward T. Bullmore[3,4,5,*]

1 Department of Psychosis Studies, Institute of Psychiatry, King's College London, London SE5 8AF, UK
2 Research Imaging Institute and Department of Radiology, The University of Texas Health Science Centre at San Antonio, San Antonio, TX 78229, USA
3 University of Cambridge, Behavioural & Clinical Neuroscience Institute, Department of Psychiatry, Cambridge CB2 0SZ, UK
4 Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge CB21 5EF, UK
5 GlaxoSmithKline, ImmunoPsychiatry, Alternative Discovery and Development, Stevenage SG1 2NY, UK

*These authors contributed equally to this work.

Correspondence to: Nicolas A. Crossley,
Department of Psychosis Studies,
Institute of Psychiatry,
King's College London,
London SE5 8AF, UK
E-mail: nicolas.crossley@kcl.ac.uk

See doi:10.1093/brain/awu148 for the scientific commentary on this article.

Brain networks or 'connectomes' include a minority of highly connected hub nodes that are functionally valuable, because the topological centrality supports integrative processing and adaptive behaviours. Recent studies also suggest that hubs have high metabolic demands and longer-distance connections than other brain regions, and therefore could be considered biologically costl Assuming that hubs thus normally combine both high topological value and high biological cost, we predicted that pathological bra... lesions would be concentrated in hub regions. To test this general hypothesis, we first identified the hubs of brain anatomical

Biological Psychiatry
Volume 77, Issue 3, 1 February 2015, Pages 223-235

Review

Depression, Neuroimaging and Connectomics: A Selective Overview

Qiyong Gong[a, b, c], Yong He[d, e]

Show more

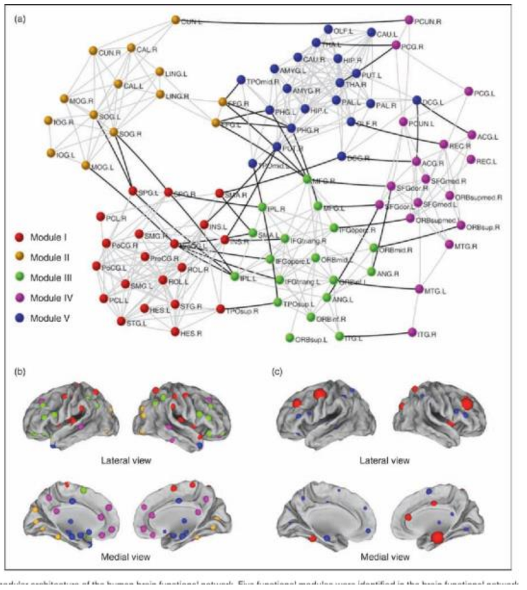
https://doi.org/10.1016/j.biopsych.2014.08.009 Get rights and content

Abstract

Depression is a multifactorial disorder with clinically heterogeneous features involving disturbances of mood and cognitive function. Noninvasive neuroimaging studies have provided rich evidence that these behavioral deficits in depression are associated with structural and functional abnormalities in specific regions and connections. Recent advances in brain connectomics through the use of graph theory highlight disrupted topological organization of large-scale functional and structural brain networks in depression, involving global topology (e.g., local clustering, shortest-path lengths, and global and local efficiencies), modular structure, and network hubs. These system-level disruptions show important

# Problem Statement

- High dimensionality of graph edges
- Challenging to acquire sufficient brain datasets
- Fisherian hypothesis test leads to multiple comparisons and inflated false positives / negatives



YJ He, AC Evans, Curr Op Neurology 2010



# Statistical Model of Brain Networks

- Brain nodes represented by Joint probability of k random processes, $n_1, n_2, \ldots, n_k$

$$P(\mathbf{n}) = P(n_1, n_2, \ldots, n_k)$$

- Brain network are random variables, $w_{ij}$ describing pairwise relationship between random processes which is fully described by joint distribution

$$P(\mathbf{w}) = P(w_{11}, w_{12}, \ldots, w_{1k}, w_{21}, w_{22}, \ldots, w_{2k}, \ldots, w_{kk})$$

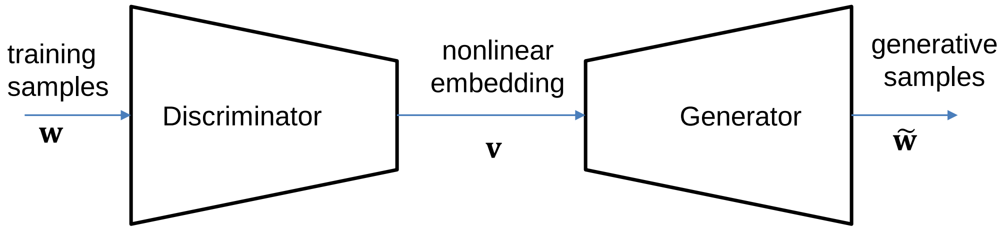
- Some examples of statistical pairwise relationship

$$w_{ij} = n_i * n_j$$
$$w_{ij} = \log P(n_i, n_j) - \log P(n_i)P(n_j)$$

- Correlation
- Mutual Entropy

# Generative Deep Neural Network Models

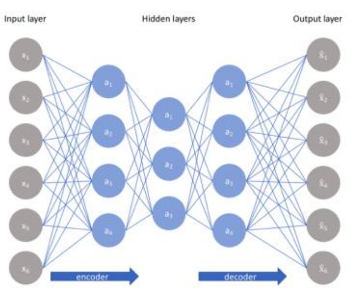
- Deep neural networks that learn a statistical model of the underlying dataset
- Purpose is to generate “novel” sample from that distribution

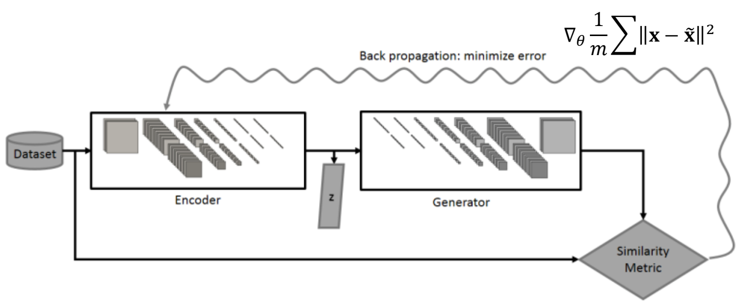





# Generative Deep Neural Network Models

## Example 1 : Autoencoders

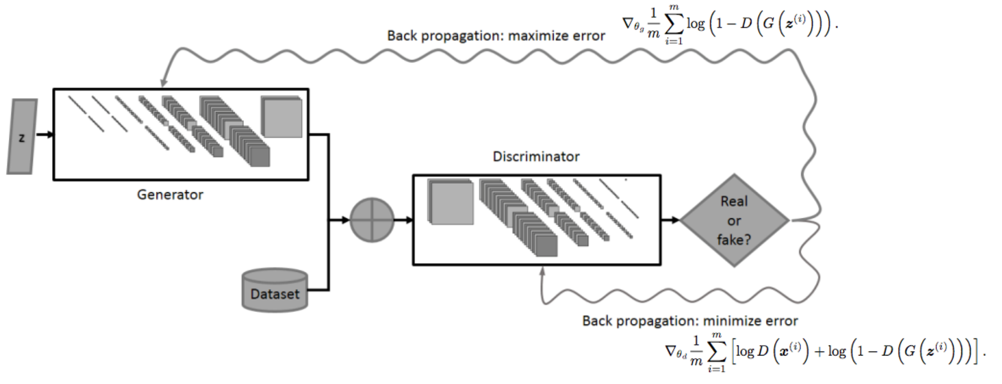



- Generator = Discriminator

# Generative Deep Neural Network Models

## Example 2 : Generative Adversarial Networks

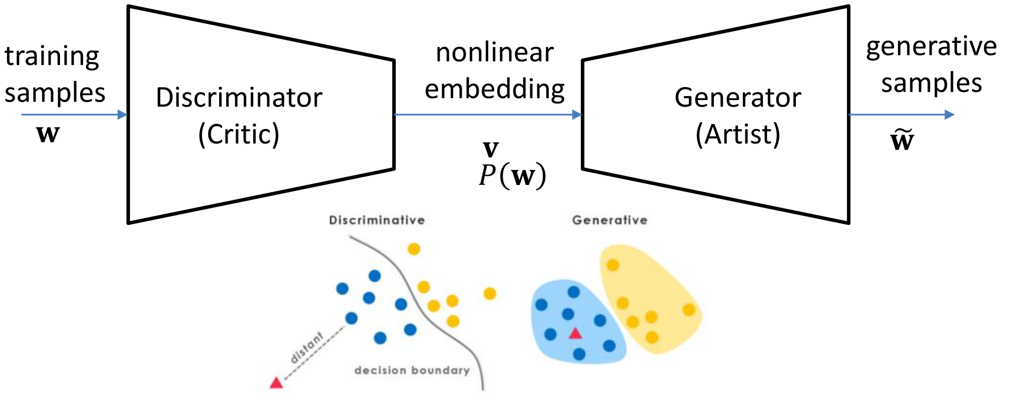



- Generator distinct from discriminator
- Learning is a competitive game between discriminator & generator optimization



# Generative Deep Neural Network Models

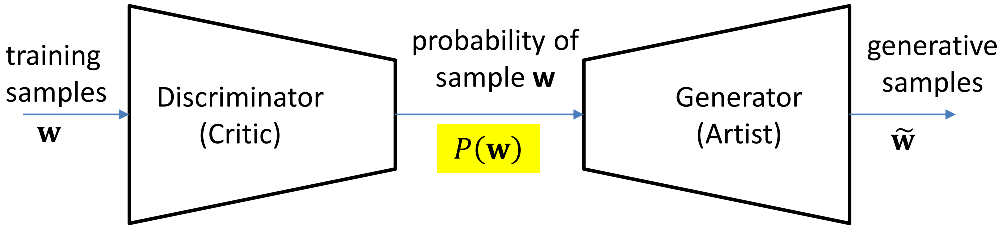
**Insight 1**
Training a GDNN is a model formulation on the dataset



## Model Formulation with GDNN

**Insight 2**
Compare models, not data points to avoid multiple comparisons

- In general, the non linear embedding is a vector, **v** with each element $\mathrm{v}_i \in [0,1]$
- Then the learnt probability distribution on the brain network dataset, $\tilde{P}(\mathbf{w}) = \prod_i \mathrm{v}_\mathrm{i}$



## Benefits of GDNN Hypothesis Testing

- Hypothesis tested by comparing models of data rather than data points. Avoid false positives arising from random test in each dimension of data.
- With appropriate distance measure, test can rank degree of similarity
- Generative models allow deeper investigation into loci of differences – not data constrained.

## Hypothesis Testing with GDNN

- $\tilde{P}(\mathbf{w})$is a model on the brain network dataset, **w** learnt by DGNN
- Hypothesis testing becomes a comparison between $\tilde{P}_d\ (\mathbf{w})$ and a reference distribution $\tilde{P}_0(\mathbf{w})$ which is also learnt

$\mathbf{w_d}$ brain networks from diseased cohort
$\mathbf{w_0}$ brain networks from healthy control cohort

H0 : $\mathbf{w_d}$ is not different from $\mathbf{w_0}$

Reject H0 if sufficient difference in the hypothesized probabilities

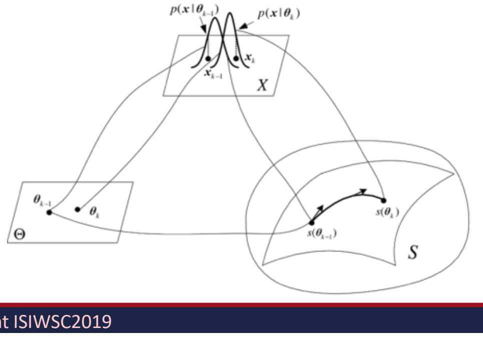
$$\tilde{P}_d(\mathbf{w}_d) \neq \tilde{P}_0(\mathbf{w}_0)$$



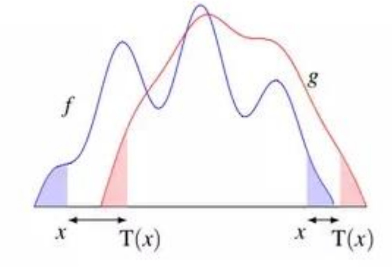
## Hypothesis Testing with GDNN

- Accept or reject hypothesis based on magnitude of a distance metric
- What is a suitable distance metric between distributions?
Two metrics known from Information Geometry & Optimal Transport
Active research ongoing to unify both metrics

Fisher-Rao distance




Wasserstein / Earth Mover distance



# Hypothesis Testing with GDNN

- Fisher-Rao distance, $|\mathrm{d}\mathbf{m}|$

Riemannian metric on statistical manifold given by $|\mathrm{d}\mathbf{m}|^2 = \sum_i \sum_j g_{ij} d\theta_i d\theta_j$

where $\theta_i$, $\theta_j$ are weights and biases in network 0 and network d respectively
$g_{ij}$ are elements in the **Fisher Information Matrix** defined by

$$g_{ij} = E\left[\frac{\partial \log \tilde{P}_0(\mathbf{w})}{\partial \theta_i} \frac{\partial \log \tilde{P}_d(\mathbf{w})}{\partial \theta_j}\right]$$

Choose direction of weights of each network with shortest geodesic distance between $P_0$ and $P_d$. In matrix formulation

$$|d\mathbf{m}|^2 = \boldsymbol{\theta}_0{}^{\mathrm{T}} \mathbf{G} \boldsymbol{\theta}_d$$




# Generalizability of Fisher-Rao distance

- Invariance of the manifold (key result from Information Geometry)

1. Invariance to reparameterization

$$p_\theta(\boldsymbol{x}) = \bar{p}_\xi(\boldsymbol{x}) \quad \text{and} \quad \sum \theta^2 \neq \sum \xi^2$$

2. Invariance under different representation

$$y = f(x) \quad \text{and} \quad \int \|p_{\theta 1}(x) - p_{\theta 2}(x)\|^2 dx \neq \int \|p_{\theta 1}(y) - p_{\theta 2}(y)\|^2 dy$$

- Implication: Fisher Rao distance is invariant if we change GDNN models (different number of weights, biases, layers, channels)

- Intuition: Distance metric between GDNN models defined on probability measures. Therefore parameterization of the probability measure should not matter

## Accuracy of the GDNN model?

- If we are comparing GDNN models on data, how representative are the models on the underlying distribution of dataset?
- Bias-Variance-model complexity on the underlying dataset
- Will diversity of models help?

## Utility of Diverse GDNN models

- Each GDNN model represents a point in the manifold of probability measures
- Multiple GDNN models on the same brain network dataset may be generated with
  (i) different training instances
  (ii) different learning algorithms
  (iii) architectures
  (iv) model complexity
- Any advantage for diversity of GDNN models on manifold of distributions?
  A Central Limit Theorem on Fréchet statistics ?
  Hypothesize that multiple GDNN models cluster around the true distribution
  Clusters may be identified by Fréchet mean and variance.
- How to handle bias (underfit) and variance (overfit)?

## Challenges for further work

- Choice of Fisher-Rao vs Wasserstein distance. Any benefits due to parameterization over GDNN parameters vs over brain variables
- Disentangling GDNN model bias-variance from true dataset distribution
- How degree of GDNN model diversity can ameliorate sample size requirement? Increase sample size improves the estimate of true probability distribution, can diversity help?
- Accuracy of comparisons at outliers/low probability events




62nd ISI World Statistics Congress
ISI 2019 Kuala Lumpur
18 - 23 August 2019 | www.isi2019.org

**THANK YOU**

hotattwei@utp.edu.my