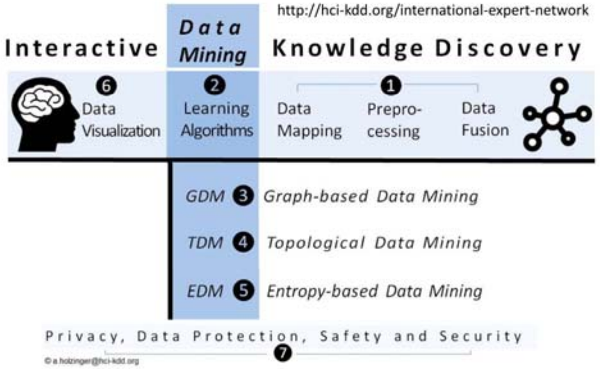
Andreas Holzinger
185.A83 Machine Learning for Health Informatics
2017S, VU, 2.0 h, 3.0 ECTS
Lecture 13 - Week 25 – 20.06.2017

Deep Learning Medical
Towards Deep Transfer Learning

a.holzinger@hci-kdd.org
http://hci-kdd.org/machine-learning-for-health-informatics-course

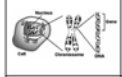
ML needs a concerted effort fostering integrated research



Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: **Cognitive Science meets Machine Learning.** IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

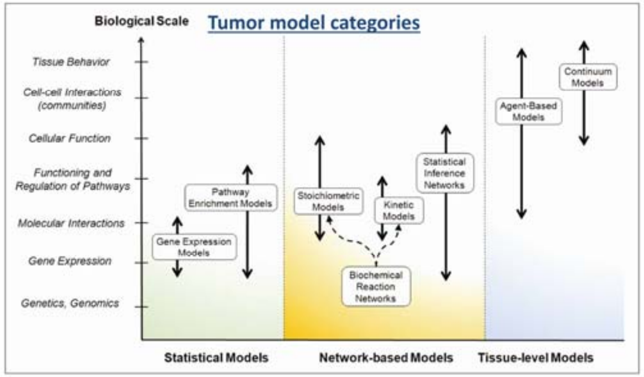
Red thread through this lecture (NN=Neural Networks)

Biological Universe vs. Computational Universe

NOTION	*BIOLOGICAL UNIVERSE*	*COMPUTATIONAL UNIVERSE*
Chromosome 		
Fitness 		
Gene 		
Generation		
Individual		
Population		

Holzinger, K., Palade, V., Rabadan, R. & Holzinger, A. 2014. Darwin or Lamarck? Future Challenges in Evolutionary Algorithms for Knowledge Discovery and Data Mining. *In: LNCS 8401.* Heidelberg, Berlin: Springer, pp. 35-56.

Tumor Growth Modeling



Edelman, L. B., Eddy, J. A. & Price, N. D. 2010. In silico models of cancer. Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 2, (4), 438-459, doi:10.1002/wsbm.75.

01 Fundamentals: from Neural Networks to Deep Learning

Advance Organizer (1)

- Deep Learning := ML method based on learning representations of data. An observation (e.g., an image) can be represented in many ways such as a vector of intensity values per pixel, or in a more abstract way as a set of edges, regions of particular shape, etc.
- Feature:= specific measurable property of a phenomenon being observed.
- Feature engineering:= using domain knowledge to create features useful for ML. ("Applied ML is basically feature engineering. Andrew Ng").
- Feature learning:= transformation of raw data input to a representation, which can be effectively exploited in ML.

Key Challenges

- High variety of data in the life sciences –
- a key to deal with high variety is data integration;
- What levels in deep learning architectures are appropriate for feature fusion with heterogeneous data [1]?

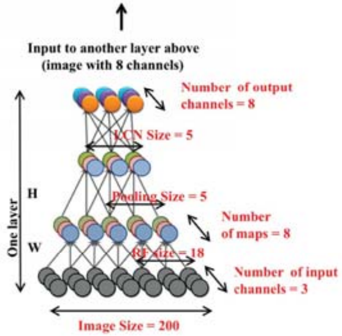
[1] Xue-Wen, C. & Xiaotong, L. 2014. Big Data Deep Learning: Challenges and Perspectives. *Access, IEEE,* 2, 514-525.

10 million 200x200 px images downloaded from Web ...

Le, Q. V. Building high-level features using large scale unsupervised learning. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013. IEEE, 8595-8598.

Perform a num. optimization to find optimal stimuli

Le, Q. V. Building high-level features using large scale unsupervised learning. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013. IEEE, 8595-8598.



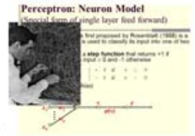
$$x^* = \arg\min_x f(x; W, H), \text{ subject to } ||x||_2 = 1.$$

Historical Issues of Deep Learning

Gauss, C. F. (1809). Theoria motus corporum coelestium in sectionibus conicis solem ambientium.

Gauss, C. F. (1821). Theoria combinationis observationum erroribus minimis obnoxiae (Theory of the combination of observations least subject to error)

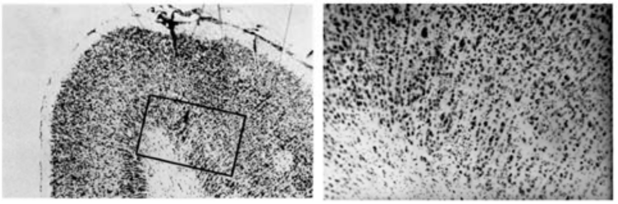
Hebb, D. O. 1949. The organization of behavior: A neuropsychological approach, John Wiley & Sons.

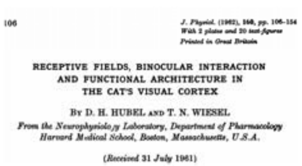


Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65, (6), 386.

Excellent Review Paper: Schmidhuber, J. 2015. Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

Cerebral cortex interconnections Hubel & Wiesel (1962)





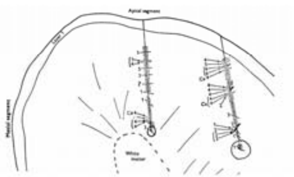
106 J. Physiol. (1962), 160, pp. 106-154
With 2 plates and 20 text-figures
Printed in Great Britain

RECEPTIVE FIELDS, BINOCULAR INTERACTION AND FUNCTIONAL ARCHITECTURE IN THE CAT'S VISUAL CORTEX

By D. H. HUBEL and T. N. WIESEL

From the Neurophysiology Laboratory, Department of Pharmacology Harvard Medical School, Boston, Massachusetts, U.S.A.

(Received 31 July 1961)



Hubel, D. H. & Wiesel, T. N. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology, 160, (1), 106-154.

Principles of human information processing ...

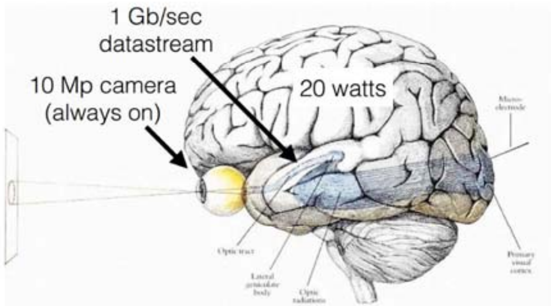
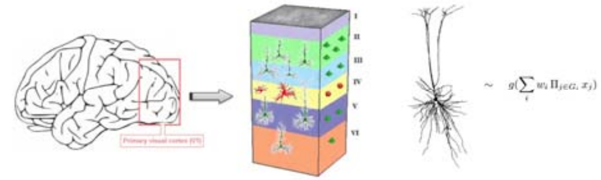
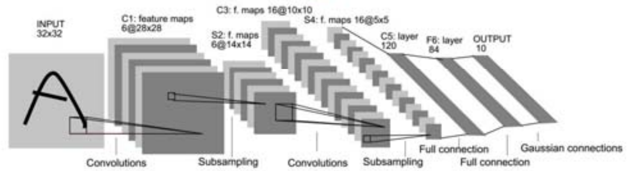


Image credit to Bruno Olshausen, Redwood Center for Theoretical Neuroscience, UC Berkeley

Olshausen BA (2014) Perception as an inference problem.
In: The Cognitive Neurosciences V,M. Gazzaniga, R. Mangun, Eds. MIT Press.

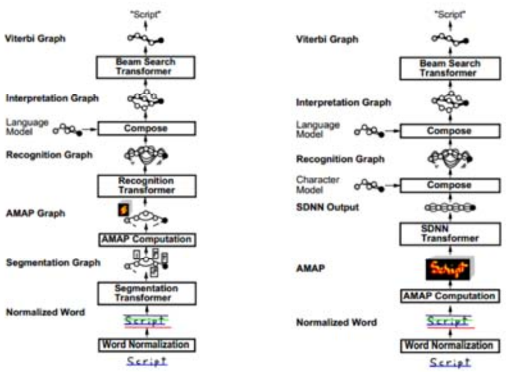
From Deep Neural Learning to Deep Learning





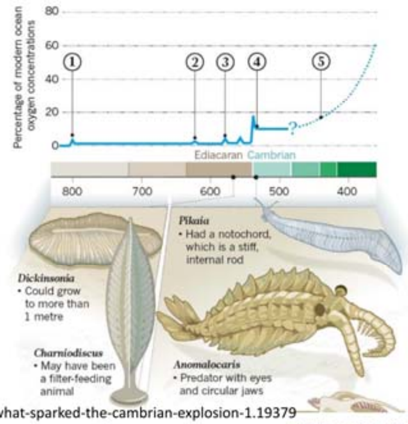
Le Cun, Y., Bottou, L., Bengio, Y. & Haffner, P. 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86, (11), 2278-2324.

Deep Learning 1998



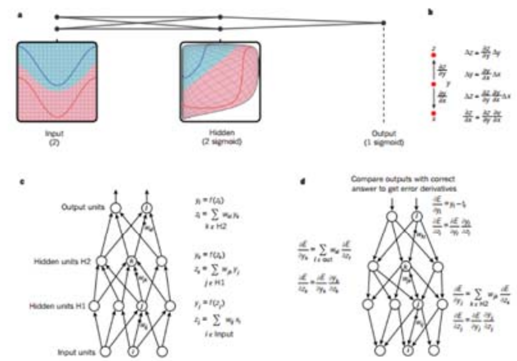
Le Cun, Y., Bottou, L., Bengio, Y. & Haffner, P. 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86, (11), 2278-2324.

15 years later ... a kind of Cambrian explosion



http://www.nature.com/news/what-sparked-the-cambrian-explosion-1.19379

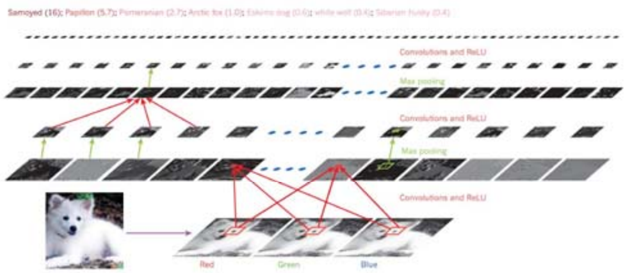
2015 ...



Le Cun, Y., Bengio, Y. & Hinton, G. 2015. Deep learning. Nature, 521, (7553), 436-444.

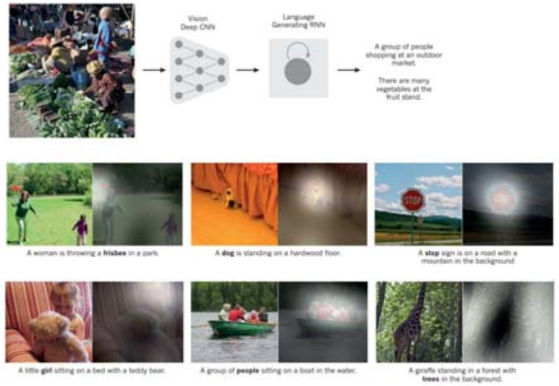
Convolutional Neural Network in computer vision

$$U(f) = (C_{W^{(K)}} \ldots \circ C_{W^{(2)}} \circ C_{W^{(1)}})(f)$$



Le Cun, Y., Bengio, Y. & Hinton, G. 2015. Deep learning. Nature, 521, (7553), 436-444.

Example



Le Cun, Y., Bengio, Y. & Hinton, G. 2015. Deep learning. Nature, 521, (7553), 436-444.

Features are key for learning and understanding

- **Initial features**
  - Bio data -> DNA, RNA, Biomarker, Sequences, Genes, etc.
  - Images -> pixels, contours, textures, etc.
  - Time series data -> trends, reversals, anomalies, ticks, etc.
  - Signal data -> spectrograms, samples, etc.
  - Text data -> words, grammar classes, relations, etc.
- **Combined features**
  - Combinations that linear system cannot represent, e.g.:
  - polynomial combinations, logical conjunctions, dec. trees.
  - Total number of features then grows **very quickly**.
- **Solutions**
  - Kernels
  - Feature selection

Limitations of Deep Learning approaches

- Computational resource intensive (supercomps, cloud CPUs, **federated learning,** …)
- Data intensive (needs often millions of training samples – **"big data"** is necessary!)
- Black-Box approaches – lack **transparency,** do not foster trust and acceptance among end-user, however, legal aspects make it difficult!
- **Non-convex:** difficult to set up, to train, to optimize, needs a lot of expertise, error prone
- Most of all: bad in dealing with **uncertainty** …

Features!

Image courtesy of Leon Bottou (2010)

Book Recommendations

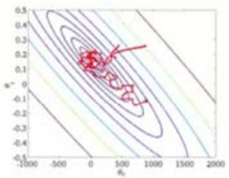
02 Representing and Dealing with Uncertainty

03 From Bayesian Networks to Gaussian Processes

04 Stochastic Gradient Descent

Minimizing an objective function

$$f(\boldsymbol{\theta}) = \mathbb{E}\left[f(\boldsymbol{\theta}, \mathbf{z})\right] \qquad \overline{\boldsymbol{\theta}}_k = \frac{1}{k}\sum_{t=1}^{k}\boldsymbol{\theta}_t$$



Nemirovskii, A. & Yudin, D. 1978. Cezare convergence of gradient method approximation of saddle points for convex-concave functions. Doklady Akademii Nauk SSSR, 239, 1056-1059.

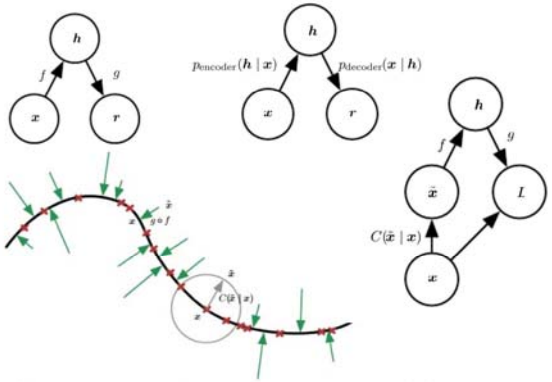
Algorithm 8.3: Stochastic gradient descent

```
1 Initialize θ, η;
2 repeat
3     Randomly permute data;
4     for i = 1 : N do
5         g = ∇f(θ, z_i);
6         θ ← proj_Θ(θ − ηg);
7         Update η;
8 until converged;
```

Murphy, K. P. 2012. Machine learning: a probabilistic perspective, Cambridge (MA), MIT press; Chapter 8.5.2

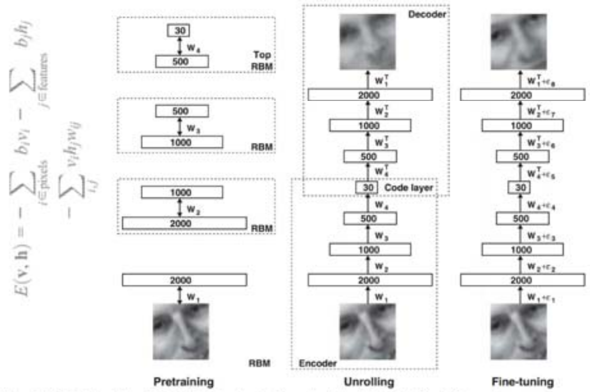
05 Deep Autoencoders (unsupervised NN)

General Structure of an Autoencoder



Goodfellow, I., Bengio, Y. & Courville, A. 2016. Deep Learning, Cambridge (MA), MIT Press, Chapter 14

Pretraining a stack of restricted Boltzmann Machines



Hinton, G. E. & Salakhutdinov, R. R. 2006. Reducing the Dimensionality of Data with Neural Networks. Science, 313, (5786), 504-507, doi:10.1126/science.1127647.

What makes a good representation? Trad. Autoencoders

- Encoder: Det. mapping f_θ that transforms an input vector x into a representation y

$$f_\theta(\mathbf{x}) = s(\mathbf{W}\mathbf{x} + \mathbf{b})$$

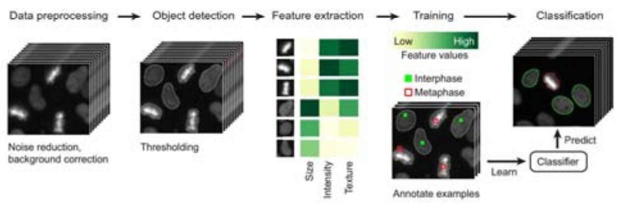
- Decoder: Resulting hidden representation y is then mapped back to a reconstructed d-dimensional vector z in input space, z = g θ′ (y). This mapping gθ ′ is called the decoder

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. & Manzagol, P.-A. 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. The Journal of Machine Learning Research, 11, 3371-3408.

$$g_{\theta'}(\mathbf{y}) = s(\mathbf{W}'\mathbf{y} + \mathbf{b}')$$

06 Deep Learning Applications in Biomedicine

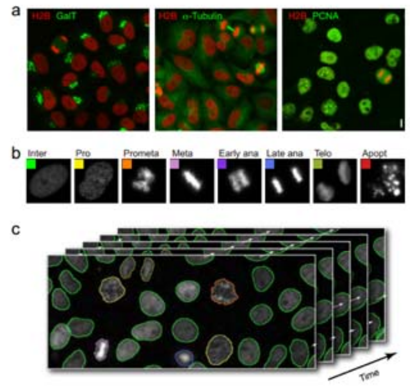
ML-Pipeline for Cell Phenotyping



Sommer, C. & Gerlich, D. W. 2013. Machine learning in cell biology–teaching computers to recognize phenotypes. Journal of Cell Science, 126, (24), 5529-5539.

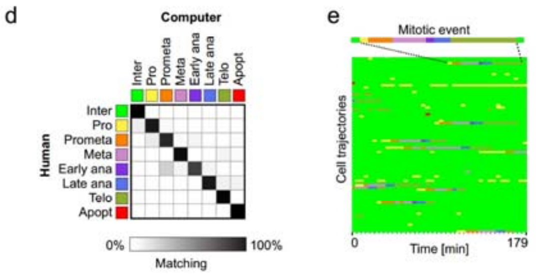
Supervised ML and classification of morphologies

Held, M., Schmitz, M. H., Fischer, B., Walter, T., Neumann, B., Olma, M. H., Peter, M., Ellenberg, J. & Gerlich, D. W. 2010. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. Nature methods, 7, (9), 747-754.



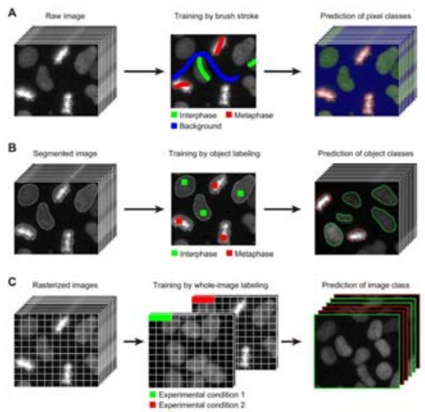
Classification

http://www.nature.com/nmeth/journal/v7/n9/extref/nmeth.1486-S3.mov



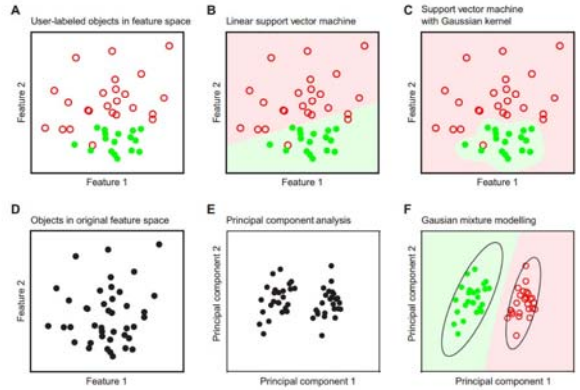
Held, M., Schmitz, M. H., Fischer, B., Walter, T., Neumann, B., Olma, M. H., Peter, M., Ellenberg, J. & Gerlich, D. W. 2010. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. Nature methods, 7, (9), 747-754.

Image Classification via supervised ML



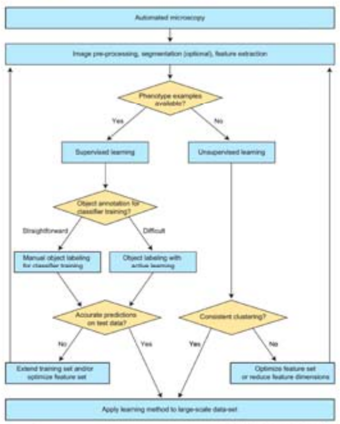
Sommer, C. & Gerlich, D. W. 2013. Machine learning in cell biology–teaching computers to recognize phenotypes. Journal of Cell Science, 126, (24), 5529-5539.

Object Representation in the feature space



Sommer, C. & Gerlich, D. W. 2013. Machine learning in cell biology–teaching computers to recognize phenotypes. Journal of Cell Science, 126, (24), 5529-5539.

Example for ML-pipeline in image-based screening

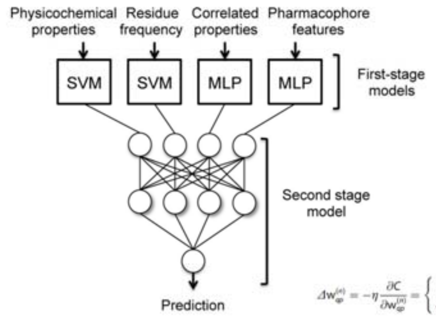


Sommer, C. & Gerlich, D. W. 2013. Machine learning in cell biology–teaching computers to recognize phenotypes. Journal of Cell Science, 126, (24), 5529-5539.

Examples of Open Source ML software for cell biologists

- Cell Profiler and Cell Profiler Analyst(Carpenter et al., 2006; Jones et al., 2008; Kamentsky et al., 2011) (http://www.cellprofiler.org).
  - Incl. modular workflow design, which enables rapid development of analysis assays. It provides a multi-class active learning interface based on boosting. CellProfiler runs on all major operating systems and supports computing on clusters for largescale screening.
- Cell Cognition (Held et al., 2010) (http://www.cellcognition.org/)
  - has been optimized for time-resolved imaging applications. It comprises a complete machine-learning pipeline from cell segmentation and feature extraction to supervised and unsupervised learning. Cell Cognition runs on all major operating systems and supports computing on clusters for large-scale screening.
- Ilastik (Sommer et al., 2011) (http://www.ilastik.org)
  - is an interactive segmentation tool based on pixel classification, which facilitates more complex image-segmentation tasks and provides real-time feedback.
- Bioconductor image HTS and EBImage (Gentleman et al., 2004; Pau et al., 2010; Pau et al., 2013) (http://www.bioconductor.org)
  - provides a versatile toolbox for statistical data analysis and image processing in the programming language R.

Example: Deep Learning in Drug Discovery

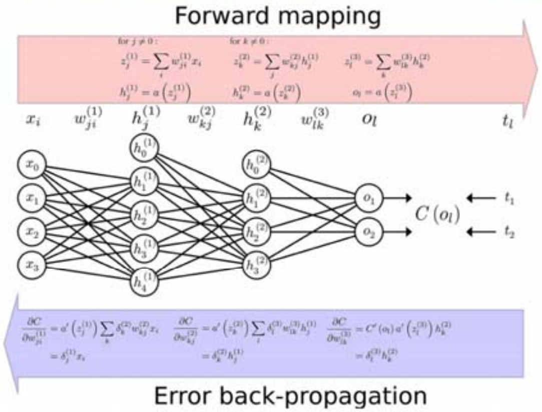


$$\Delta w_{qp}^{(n)} = -\eta \frac{\partial C}{\partial w_{qp}^{(n)}} = \begin{cases} -\eta \delta_q^{(n)} x_p, n = 1 \\ -\eta \delta_q^{(n)} h_p^{(n-1)}, n \neq 1 \end{cases}, \text{ with}$$

$$\delta_q^{(n)} = \frac{\partial C}{\partial z_q^{(n)}} \begin{cases} E'(o_q)\sigma'\left(z_q^{(n)}\right), n = N \\ \sigma'\left(z_q^{(n+1)}\right)\sum_r \delta_r^{(n+1)} w_{rq}^{(n+1)}, n \neq N \end{cases}$$

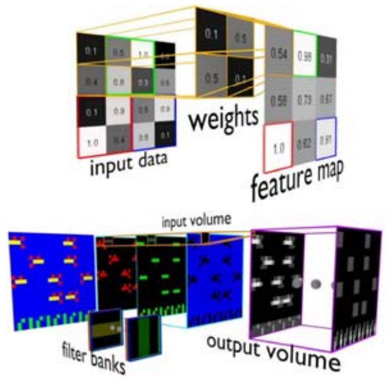
Gawehn, E., Hiss, J. A. & Schneider, G. 2016. Deep Learning in Drug Discovery. *Molecular Informatics*, 35, 3-14.

Example: Deep Learning in Drug Discovery



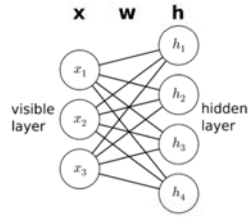
Gawehn, E., Hiss, J. A. & Schneider, G. 2016. Deep Learning in Drug Discovery. *Molecular Informatics*, 35, 3-14.

Example: Deep Learning in Drug Discovery

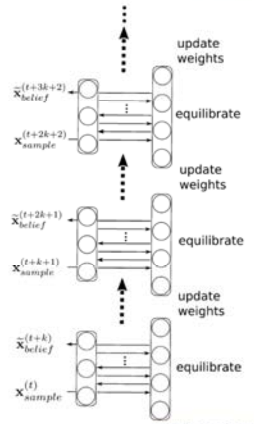


Gawehn, E., Hiss, J. A. & Schneider, G. 2016. Deep Learning in Drug Discovery. *Molecular Informatics*, 35, 3-14.

Restricted Boltzmann Machine

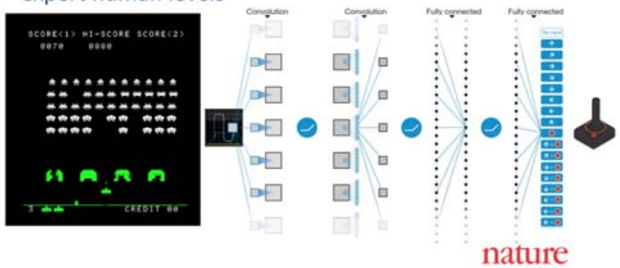


Gawehn, E., Hiss, J. A. & Schneider, G. 2016. Deep Learning in Drug Discovery. *Molecular Informatics*, 35, 3-14.



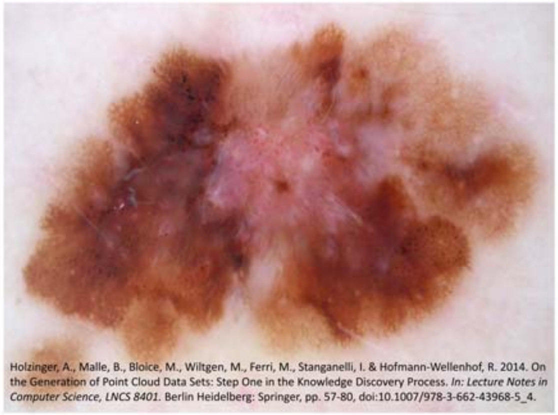
2015 – the year of reinforcement learning ☺

Deep Q-networks (Q-Learning is a model-free RL approach) have successfully played Atari 2600 games at expert human levels



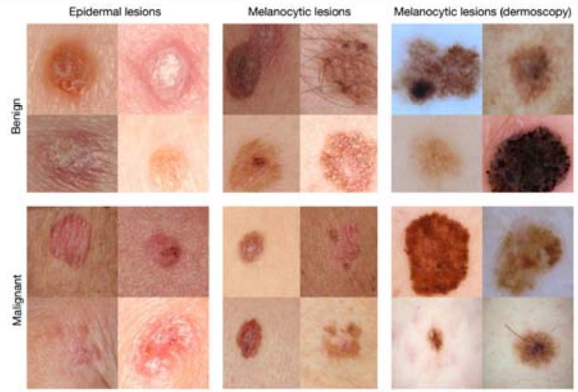
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. & Hassabis, D. 2015. Human-level control through deep reinforcement learning. Nature, 518, (7540), 529-533, doi:10.1038/nature14236

To learn e.g. structural anomalies from such data …



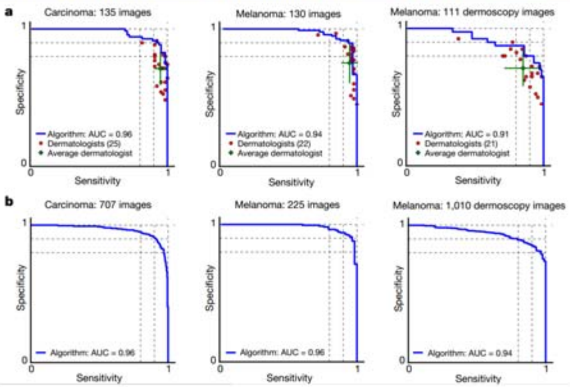
Holzinger, A., Malle, B., Bloice, M., Wiltgen, M., Ferri, M., Stanganelli, I. & Hofmann-Wellenhof, R. 2014. On the Generation of Point Cloud Data Sets: Step One in the Knowledge Discovery Process. *In: Lecture Notes in Computer Science, LNCS 8401*. Berlin Heidelberg: Springer, pp. 57-80, doi:10.1007/978-3-662-43968-5_4.

Dermatologist-level classification of skin cancer with DL



Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.

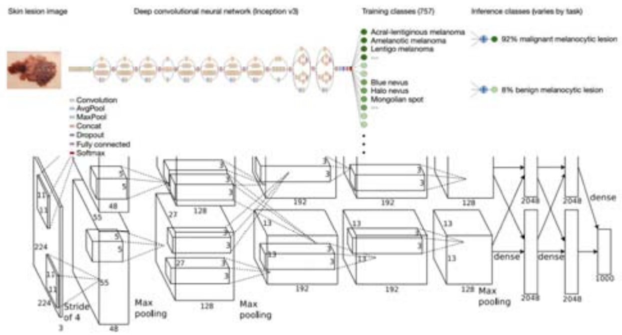
Skin cancer classification performance: Human vs. CNN



Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.

Deep Convolutional Neural Network Pipeline

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.



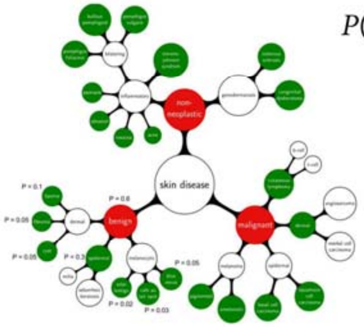
Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q., eds. Advances in neural information processing systems (NIPS 2012), 2012 Lake Tahoe. 1097-1105.

Tree-structured taxonomy of skin cancer

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.

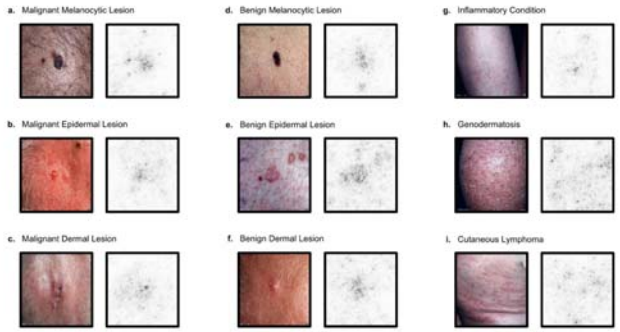
Probabilities over the taxonomy

$$P(u) = \sum_{v \in C(u)} P(v)$$



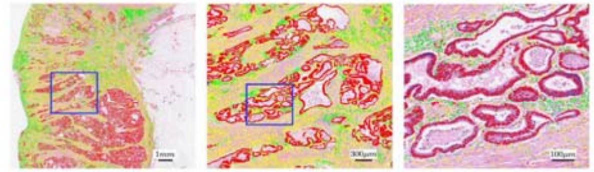
Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.

Saliency maps for nine examples

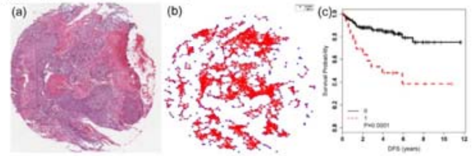


Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.

Handcrafted features for disease characterization

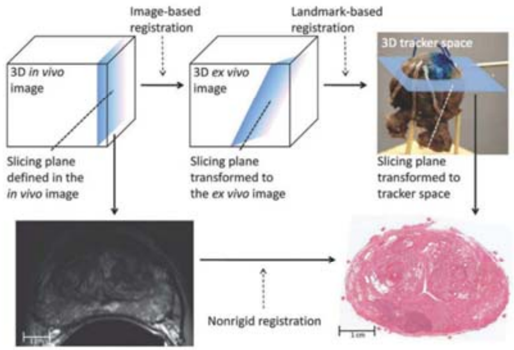


Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D. R. J., Cree, I. A. & Rajpoot, N. M. 2016. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Transactions on Medical Imaging, 35, (5), 1196-1206, doi:10.1109/TMI.2016.2525803.



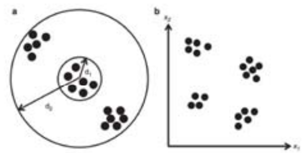
Madabhushi, A. & Lee, G. 2016. Image analysis and machine learning in digital pathology: Challenges and opportunities. Medical Image Analysis, 33, 170-175, doi:10.1016/j.media.2016.06.037.

Image Guided Slicing Process - Towards Digital Pathology

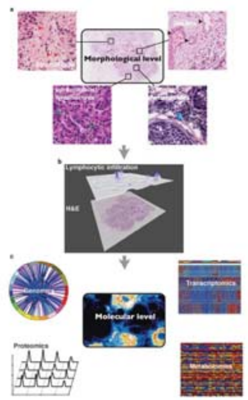


Ward, A. D., Crukley, C., Mckenzie, C. A., Montreuil, J., Gibson, E., Romagnoli, C., Gomez, J. A., Moussa, M., Chin, J., Bauman, G. & Fenster, A. 2012. Prostate: Registration of Digital Histopathologic Images to in Vivo MR Images Acquired by Using Endorectal Receive Coil. Radiology, 263, (3), 856-864, doi:10.1148/radiol.12102294.

Mapping spatial heterogeneity tumor: Digital Pathology



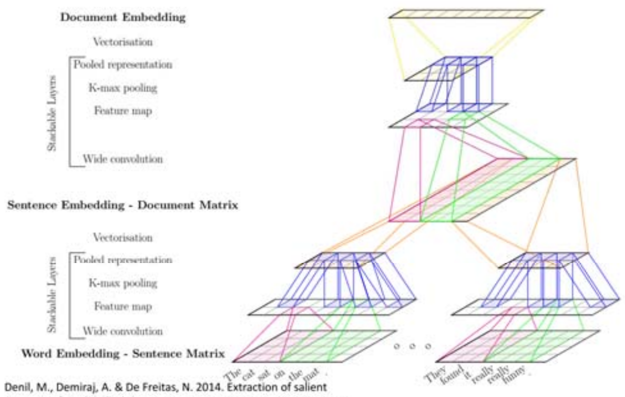
Heindl, A., Nawaz, S. & Yuan, Y. 2015. Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. Lab Invest, 95, (4), 377-384, doi:10.1038/labinvest.2014.155.



07 Future Challenges and Extravaganza *) Topics

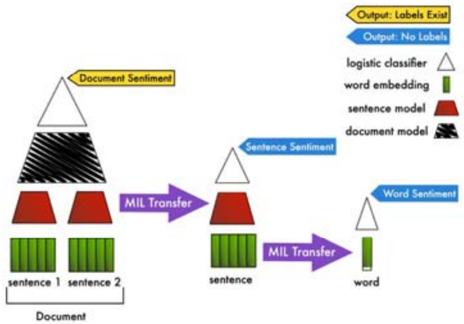
*) Holzinger, A. 2014. Extravaganza Tutorial on Hot Ideas for Interactive Knowledge Discovery and Data Mining in Biomedical Informatics. In: Slezak, D., Tan, A.-H., Peters, J. F. & Schwabe, L. (eds.) Brain Informatics and Health, BIH 2014, Lecture Notes in Artificial Intelligence, LNAI 8609. Heidelberg, Berlin: Springer, pp. 502-515, doi:10.1007/978-3-319-09891-3_46.

Knowledge Extraction via Deep Learning



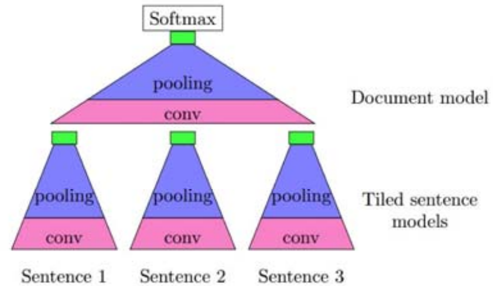
Denil, M., Demiraj, A. & De Freitas, N. 2014. Extraction of salient sentences from labelled documents. arXiv preprint arXiv:1412.6815.

Deep Transfer Learning



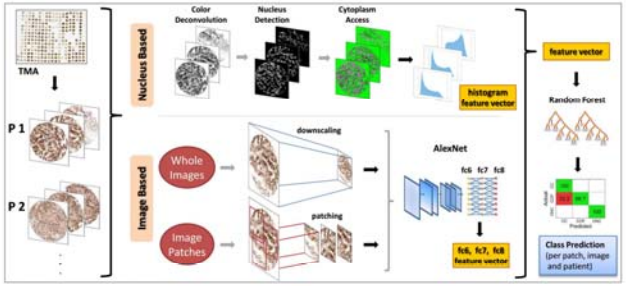
Kotzias, D., Denil, M., Blunsom, P. & De Freitas, N. 2014. Deep multi-instance transfer learning. *arXiv preprint arXiv:1411.3128.*

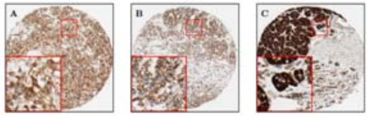
Deep Multi-Instance Learning



Kotzias, D., Denil, M., Blunsom, P. & De Freitas, N. 2014. Deep multi-instance transfer learning. *arXiv preprint arXiv:1411.3128.*

Learning from Deep vs. Flat Feature Representations





Schüffler, P. J., Sarungbam, J., Muhammad, H., Reznik, E., Tickoo, S. & Fuchs, T. 2016. Mitochondria-based Renal Cell Carcinoma Subtyping: Learning from Deep vs. Flat Feature Representations. In: Finale, D.-V., Jim, F., David, K., Byron, W. & Jenna, W. (eds.) Proceedings of the 1st Machine Learning for Healthcare Conference. Proceedings of Machine Learning Research: PMLR. 191--208.

The grand question of cognitive science

15b

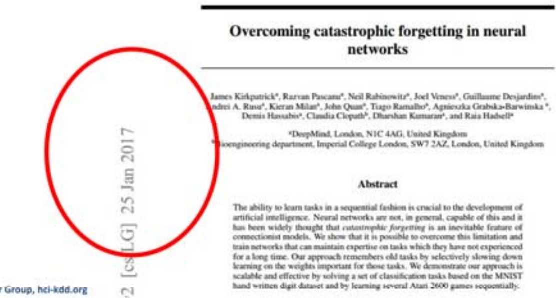
- **"How do humans generalize from very few examples?"**
- They transfer knowledge from previous learning:
  - Representation learning (features!)
  - Explanatory factors
  - Previous learning from unlabeled data and labels for other tasks
- Prior: shared underlying explanatory factors, in particular between $P(x)$ and $P(Y|X)$, with a causal link between $Y \rightarrow X$

Bengio, Y., Courville, A. & Vincent, P. 2013. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35, (8), 1798-1828, doi:10.1109/TPAMI.2013.50.

Catastrophic Forgetting – cutting edge research

20

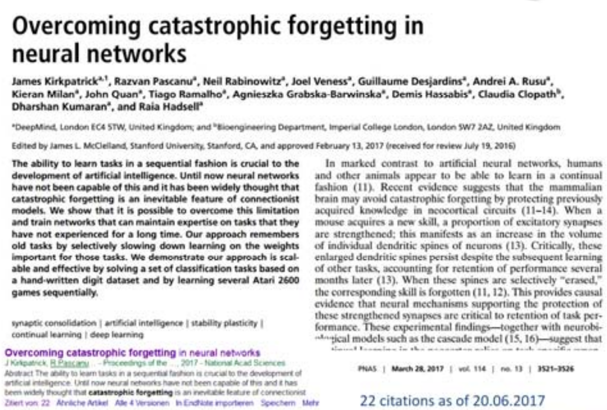
- When trained on one task, then trained on a 2nd task, many machine learning models ("deep learning"!) forget how to perform the first task.


Overcoming catastrophic forgetting in neural networks

[cs.LG] 25 Jan 2017

Abstract

Overcoming catastrophic forgetting published in PNAS


Overcoming catastrophic forgetting in neural networks

James Kirkpatrick[a,1], Razvan Pascanu[a], Neil Rabinowitz[a], Joel Veness[a], Guillaume Desjardins[a], Andrei A. Rusu[a], Kieran Milan[a], John Quan[a], Tiago Ramalho[a], Agnieszka Grabska-Barwinska[a], Demis Hassabis[a], Claudia Clopath[b], Dharshan Kumaran[a], and Raia Hadsell[a]

[a]DeepMind, London EC4 5TW, United Kingdom; and [b]Bioengineering Department, Imperial College London, London SW7 2AZ, United Kingdom

Edited by James L. McClelland, Stanford University, Stanford, CA, and approved February 13, 2017 (received for review July 19, 2016)

The ability to learn tasks in a sequential fashion is crucial to the development of artificial intelligence. Until now neural networks have not been capable of this and it has been widely thought that catastrophic forgetting is an inevitable feature of connectionist models. We show that it is possible to overcome this limitation and train networks that can maintain expertise on tasks that they have not experienced for a long time. Our approach remembers old tasks by selectively slowing down learning on the weights important for those tasks. We demonstrate our approach is scalable and effective by solving a set of classification tasks based on a hand-written digit dataset and by learning several Atari 2600 games sequentially.

synaptic consolidation | artificial intelligence | stability plasticity | continual learning | deep learning

In marked contrast to artificial neural networks, humans and other animals appear to be able to learn in a continual fashion (11). Recent evidence suggests that the mammalian brain may avoid catastrophic forgetting by protecting previously acquired knowledge in neocortical circuits (11–14). When a mouse acquires a new skill, a proportion of excitatory synapses are strengthened; this manifests as an increase in the volume of individual dendritic spines of neurons (13). Critically, these enlarged dendritic spines persist despite the subsequent learning of other tasks, accounting for retention of performance several months later (13). When these spines are selectively "erased," the corresponding skill is forgotten (11, 12). This provides causal evidence that neural mechanisms supporting the protection of these strengthened synapses are critical to retention of task performance. These experimental findings—together with neurobiological models such as the cascade model (15, 16)—suggest that

PNAS | March 28, 2017 | vol. 114 | no. 13 | 3521–3526

Overcoming catastrophic forgetting in neural networks

J Kirkpatrick, R Pascanu, ... - Proceedings of the ..., 2017 - National Acad Sciences

Abstract The ability to learn tasks in a sequential fashion is crucial to the development of artificial intelligence. Until now neural networks have not been capable of this and it has been widely thought that catastrophic forgetting is an inevitable feature of connectionist

Zitiert von: 22 Ähnliche Artikel Alle 4 Versionen In EndNote importieren Speichern Mehr

22 citations as of 20.06.2017

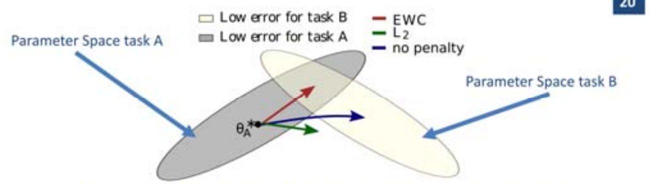
Elastic Weight Consolidation (EWC)

- Kirkpatrick et al. (2017) demonstrate that task-specific synaptic consolidation offers a unique solution to the continual-learning problem for artificial intelligence.
- Developed an algorithm analogous to synaptic consolidation for artificial neural networks,
- Elastic Weight Consolidation (EWC).
- This algorithm slows down learning on certain weights based on how **important they are to previously seen tasks.**
- They show how EWC can be used in supervised learning and reinforcement learning problems to train several tasks sequentially without forgetting older ones, in marked contrast to previous deep-learning techniques.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D. & Hadsell, R. 2017. Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114, (13), 3521-3526, doi:10.1073/pnas.1611835114.

Overcoming Catastrophic Forgetting: Deep Learning Bayes

20



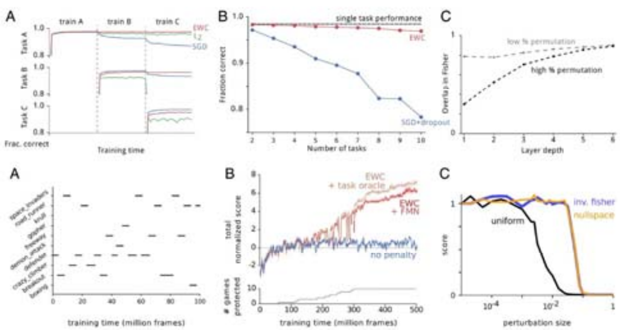
$$\log p(\theta|\mathcal{D}) = \log p(\mathcal{D}|\theta) + \log p(\theta) - \log p(\mathcal{D})$$

$$\log p(\theta|\mathcal{D}) = \log p(\mathcal{D}_B|\theta) + \log p(\theta|\mathcal{D}_A) - \log p(\mathcal{D}_B)$$

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i(\theta_i - \theta^*_{A,i})^2$$

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D. & Hadsell, R. 2016. Overcoming catastrophic forgetting in neural networks. arXiv preprint arXiv:1612.00796.

Results on MNIST Data Sets



Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D. & Hadsell, R. 2017. Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114, (13), 3521-3526, doi:10.1073/pnas.1611835114.

Machine Learning inspires Human Learning and v.v.


Trends in Cognitive Sciences

Volume 21, Issue 6, June 2017, Pages 407–408

Spotlight

Avoiding Catastrophic Forgetting

Michael E. Hasselmo[1]

[1] Center for Systems Neuroscience, Boston University, 2 Cummington Mall, Boston, MA 02215, USA

Available online 23 April 2017

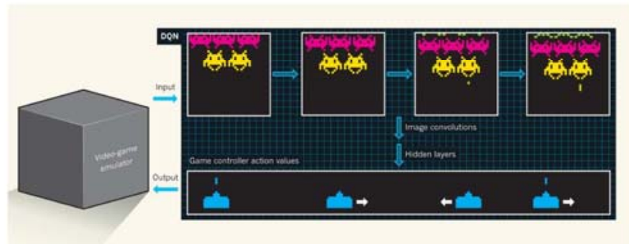
Show less

https://doi.org/10.1016/j.tics.2017.04.001 Get rights and content

Humans regularly perform new learning without losing memory for previous information, but neural network models suffer from the phenomenon of catastrophic forgetting in which new learning impairs prior function. A recent article presents an algorithm that spares learning at synapses important for previously learned function, reducing catastrophic forgetting.

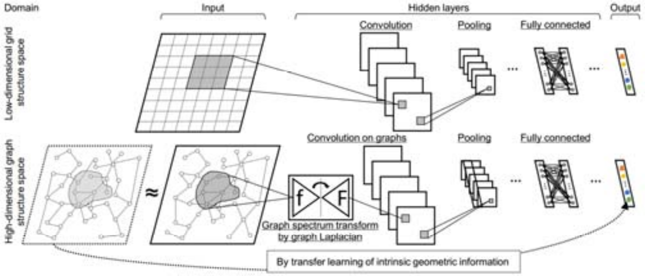
Hasselmo, M. E. 2017. Avoiding Catastrophic Forgetting. Trends in Cognitive Sciences, 21, (6), 407-408.

This experiment was done with Games ...



Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. & Hassabis, D. 2015. Human-level control through deep reinforcement learning. Nature, 518, (7540), 529-533, doi:10.1038/nature14236

Transfer Learning for Deep Learning on Graphs



Lee, J., Kim, H., Lee, J. & Yoon, S. 2016. Intrinsic Geometric Information Transfer Learning on Multiple Graph-Structured Datasets. arXiv:1611.04687.

Extravaganza: Going beyond Euclidean Data

Geometric deep learning: going beyond Euclidean data

Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, Pierre Vandergheynst

arXiv:1611.08097v1 [cs.CV] 24 Nov 2016

Many signal processing problems involve data whose underlying structure is non-Euclidean, but may be modeled as a manifold or (combinatorial) graph. For instance, in social networks, the characteristics of users can be modeled as signals on the vertices of the social graph [1]. Sensor networks are graph models of distributed interconnected sensors, whose readings are modelled as time-dependent signals on the vertices. In genetics, gene expression data are modeled as signals defined on the regulatory network [2]. In neuroscience, graph models are used to represent anatomical and functional structures of the brain. In computer graphics and vision, 3D objects are modeled as Riemannian manifolds (surfaces) endowed with properties such as color texture. Even more complex examples include networks of operators, e.g., functional correspondences [3] or difference operators [4] in a collection of 3D shapes, or orientations of overlapping cameras in multi-view vision ("structure from motion") problems [5].

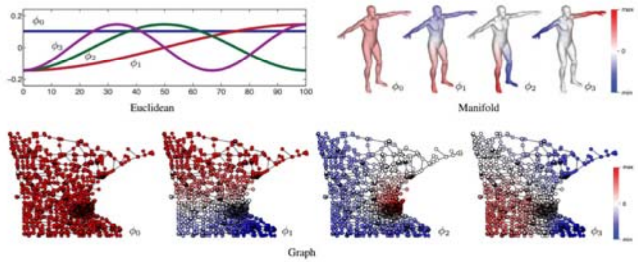
The complexity of geometric data and the availability of very large datasets (in the case of social networks, on the scale of billions) suggest the use of machine learning techniques. In particular, deep learning has recently proven to be a powerful tool for problems with large datasets with underlying Euclidean structure.

The purpose of this paper is to overview the problems arising in relation to geometric deep learning and present

Constructions that leverage the statistical properties of the data, in particular stationarity and compositionality through local statistics, which are present in natural images, video, and speech [18], [19], are one of the key reasons for the success of deep neural networks in these domains. These statistical properties have been related to physics [20] and formalized in specific classes of convolutional neural networks (CNNs) [21], [22], [23]. For example, one can think of images as functions on the Euclidean space (plane), sampled on a grid. In this setting, stationarity is owed to shift-invariance, locality is due to the local connectivity, and compositionality stems from the multi-resolution structure of the grid. These properties are exploited by convolutional architectures [24], which are built of alternating convolutional and downsampling (pooling) layers. The use of convolutions has a two-fold effect. First, it allows extracting local features that are shared across the image domain and greatly reduces the number of parameters in the network with respect to generic deep architectures (and thus also the risk of overfitting), without sacrificing the expressive capacity of the network. Second, as we will show in the following, the convolutional architecture itself imposes some priors about the data, which appear very suitable especially for natural images [25], [22].

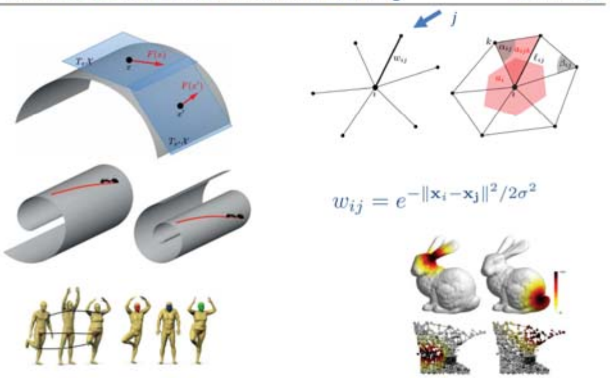
While deep learning models have been particularly successful when dealing with signals such as speech, images, or video,

Extravaganza: Going beyond Euclidean Data ...



Bronstein, M. M., Bruna, J., Lecun, Y., Szlam, A. & Vandergheynst, P. 2016. Geometric deep learning: going beyond Euclidean data. arXiv preprint arXiv:1611.08097.

Isometric Deformations in Manifold Learning



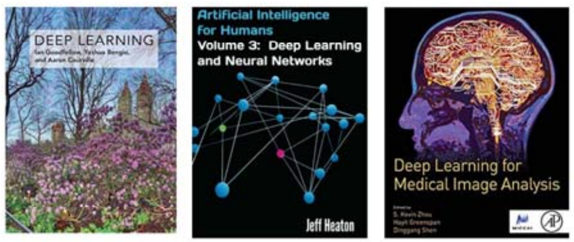
$$w_{ij} = e^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2}$$

Bronstein, M. M., Bruna, J., Lecun, Y., Szlam, A. & Vandergheynst, P. 2016. Geometric deep learning: going beyond Euclidean data. arXiv preprint arXiv:1611.08097.

Thank you!

Appendix

Book recommendations

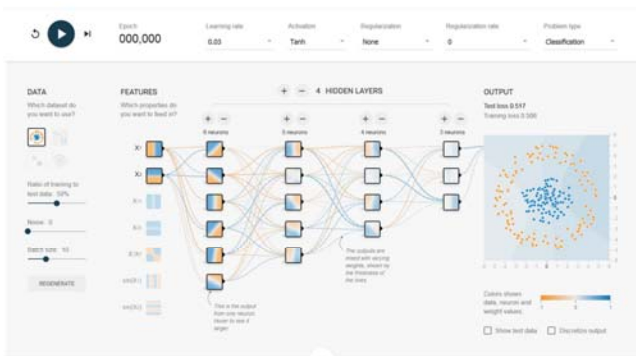


Goodfellow, I., Bengio, Y. & Courville, A. 2016. Deep Learning, Cambridge (MA), MIT Press.

https://mitpress.mit.edu/books/deep-learning

More books on Deep learning and related topics

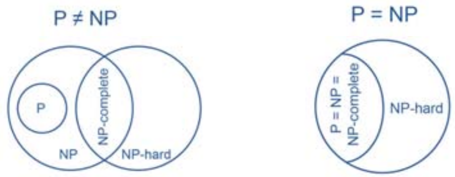
Tensorflow Playground



http://playground.tensorflow.org

Remember: Many problems in health informatics are hard

- **P:** algorithm can solve the problem in polynomial time (worst-case running-time for problem size n is less than F(n))
- **NP:** problem can be solved and any solution can be verified within polynomial time (P ⊆ NP)
- **NP-complete:** problem belongs to class NP and any other problem in NP can be reduced to this problem
- **NP-hard:** problem is at least as hard as any other problem in NP-complete but solution cannot necessarily be verified within polynomial time



Artificial Intelligence > Machine Learning > Deep Learning

Goodfellow, I., Bengio, Y. & Courville, A. 2016. Deep Learning, Cambridge (MA), MIT Press, Chapter 1, p.9

Open Problem: How to avoid negative transfer?