Office of the Assistant Secretary of Defense: Research & Engineering (OASD(R&E))

TECHSIGHT SNAPSHOT REPORT

OCTOBER 2017

Office of Net Technical Assessments (ONTA)

Deep Learning

[1]

Executive Summary

Deep learning encompasses a branch of machine learning research wherein artificially intelligent systems, broadly inspired by biological nervous systems, discover discriminating features from subtly different examples. This approach has proven useful for tasks involving high complexity and large amounts of data, such as signal processing, game solving, and classification. Already finding numerous applications in the commercial sector, deep learning has great potential for addressing informational problems in the military, such as ISR exploitation, logistics optimization, target identification, force coordination, tactical decision aids, and enabling autonomous systems. Current research focuses on specific methods of constructing deep learning systems, along with specific applications for speech recognition. Meanwhile, in the patent space, deep learning approaches are being applied to a range of problems including audio/visual processing, medicine, and materials development. While comprising only a segment of the overall field of machine learning, deep learning constitutes a medium-sized research field in its own right, with a total of ~6000 publications and ~900 patents over the past 20 years, the bulk of which has been produced in the past 5 years. Given the nature of the field, it is difficult to assess the rate of technological transition, but both publications and patents have grown rapidly since the mid-2000s. In terms of international competition, the U.S. and China emerge as clear leaders in the field, with China slightly ahead in overall publication numbers and the U.S. ahead in patents. However, in terms of quality, the U.S. maintains a clear lead in citation rates, demonstrating the high impact of U.S. originated R&D.

[1] A graphical representation of a deep neural network. Derived from (Meade, 2017).

I. Introduction to Snapshot Reports

Snapshot reports provide a short overview of recent activity in emerging and potentially disruptive research areas using quantitative metrics generated from the statistical analysis of publication trends in the scientific and patent literature exclusively using ONTA's TechSight System. The aim of these reports is to generate questions for deeper investigations, and they are engineered to be produced monthly in a rapid, timely fashion with figures automatically generated by TechSight. Since these figures are inserted from a dynamic interface, readers are encouraged to access this data on TechSight for further exploration. TechSight is available to all DoD personnel and contractors (*see APPENDIX for access instructions*). Future snapshot reports will analyze top organizations and entities as system improvements to TechSight such as entity disambiguation is implemented.

II. What is Deep Learning?

- **Deep learning, also known as deep structured learning or hierarchical learning, describes a class of machine learning[2] approaches centered on the construction of artificial neural networks.** Deep learning systems commonly consist of two key features: "(1) models consisting of multiple layers or stages of nonlinear information processing; and (2) methods for supervised or unsupervised learning of feature representation at successively higher, more abstract layers" (Deng & Yu, 2014).

- **"Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction…higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations."** (LeCun, Bengio, & Hinton, 2015). In facial recognition, for example, a system may take a raw input of pixels and, in the first layer, extract the features of edges and shadows. In the second layer, it may group these features into simple shapes. In the third layer, it may extract complex structures, such as eyes, or noses. Finally, in the top layer, the system can discriminate a host of features to identify and

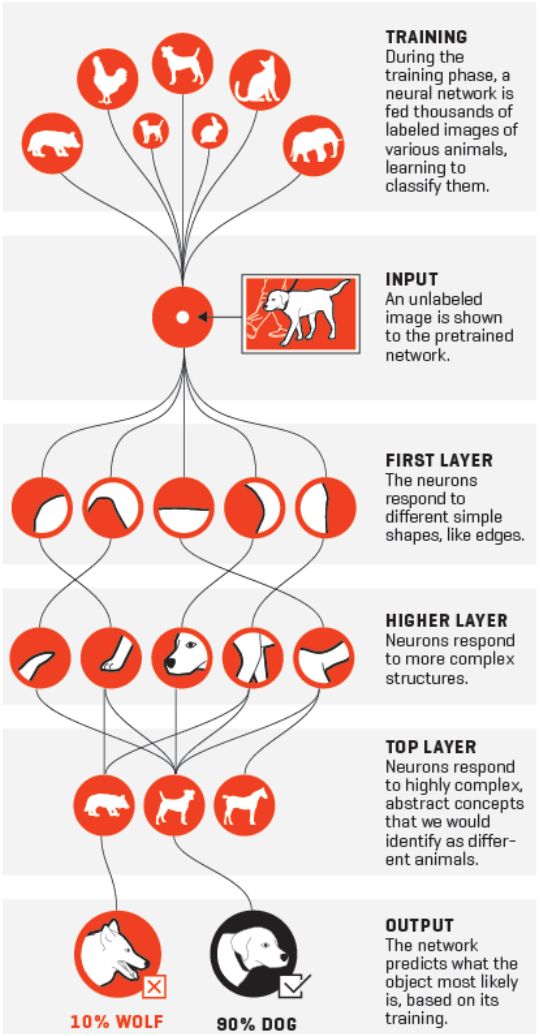


Figure 1: A fundamental explanation of how deep neural networks perform image recognition. Figure and description derived from **(Parloff, 2016)**.

[2] Machine learning is a branch of computer science that attempts to give computers the ability to learn tasks without the use of explicit programming instructions.

discriminate between specific individuals. This process is outlined above in Figure 1. The critical aspect of these systems, though, is that the features extracted are not human-defined, but constructed purely from the data (LeCun, Bengio, & Hinton, 2015).

- Rather than employ explicit, rule-based instructions, deep learning enables a machine to extract *features* from raw data to build higher-level *representations*, which then guide its task performance. With the ability to build representations from the ground up, deep learning systems excel at processing vast amounts of data, making them well suited for a variety of tasks, such as **signal processing, classification, optimization,** and **anomaly detection.** Due to its adaptability, researchers and companies have applied deep learning to a range of applications here such volumes of labeled data are abundant, including, but not limited to: object recognition, voice recognition, natural language processing, game playing, information retrieval, predictive analytics, autonomous driving, robotics, image generation, synthetic audio, real-time translation, industrial optimization, fraud detection, cybersecurity, and finance (Trost, 2016) (Allison, 2017) (Perez, 2016) (Hadad, 2017).

- **Deep learning is an old field with new breakthroughs.** The principles of deep learning extend back to the 1940s as researchers considered the explicit calculus of neuron activity (McCulloch & Pitts, 1943). The original artificial neuron, the perceptron, was proposed in 1958 (Rosenblatt, 1958). In the following four decades researchers developed new methods of training, such as backpropagation (Rumelhart, Hinton, & Williams, 1986), and applications, such as handwriting recognition (Lecun, et al., 1989). However, the field failed to achieve large-scale appeal due to the difficulty of training networks and the high resources required for their use. In the 2000s, deep learning grew rapidly out of the combination of two major breakthroughs. First, in 2006, researchers demonstrated a method of quickly and accurately training deep networks, a task that had previously been considered intractable (Hinton, Osindero, & Teh, 2006). Following that, in 2009, researchers began to utilize graphical processing units (GPUs) as the hardware base for neural networks, demonstrating a 70x boost in performance using low-cost, consumer-grade hardware (Raina, Madhavan, & Ng, 2009). From this point, researchers were able to build much deeper systems than any time prior, allowing them to attack more difficult and complex problems. The effects of this advancement can be seen in Figure 2, which details the rapid growth of network complexity beginning in the mid-2000s. Likewise, the shift to relatively low-cost hardware opened access to researchers with much more limited resources, expanding both the depth and breadth of the field.

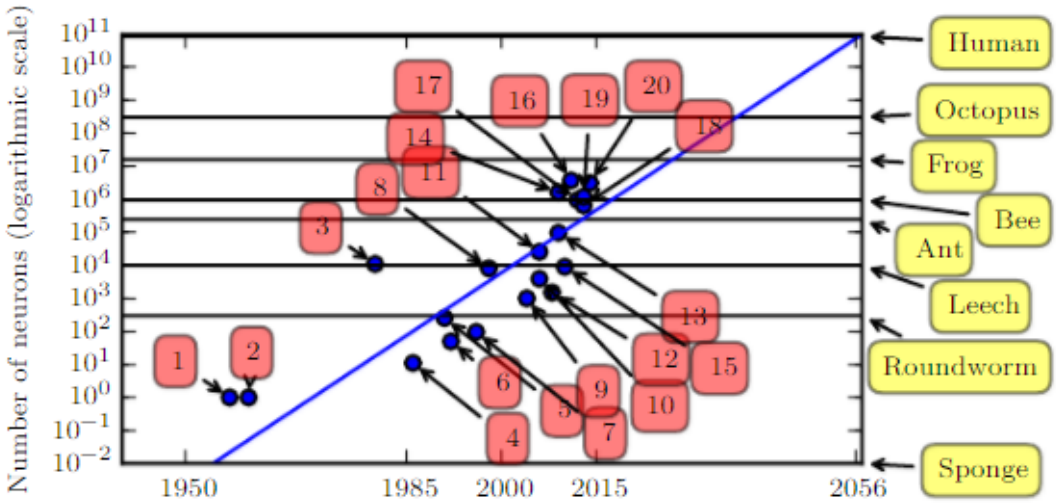


Figure 2: "Increasing neural network size over time. Since the introduction of hidden units, artificial neural networks have doubled in size roughly every 2.4 years." List of artificial neural networks represented: "1. Perceptron 2. Adaptive linear element 3. Neocognitron 4. Early back-propagation network 5. Recurrent neural network for speech recognition 6. Multilayer perceptron for speech recognition 7. Mean field sigmoid belief network 8. LeNet-5 9. Echo state network 10. Deep belief network 11. GPU-accelerated convolutional network 12. Deep Boltzmann machine 13. GPU-accelerated deep belief network 14. Unsupervised convolutional network 15. GPU-accelerated multilayer perceptron 16. OMP-1 network 17. Distributed autoencoder 18. Multi-GPU convolutional network 19. COTS HPC unsupervised convolutional network 20. GoogLeNet" Figure and description derived from **(Goodfellow, Bengio, & Courville, 2016)**.

- **Current approaches to deep learning face challenges in three areas: access to data, hardware performance, and generalizability.** At this point, the most common applications of deep learning, such as image recognition and natural language processing require rely on techniques of *supervised* learning, where humans label the training data and the machine learns how to most accurately recreate those labels on unseen sets. This approach requires vast amounts of training data, and, more problematically, massive human effort to label the training sets. As these systems become more complex, the ability to generate such data may be out of reach for most researchers. Likewise, even with the advancements in GPU processing, advanced deep learning research currently requires massive hardware investment. For example, AlphaGo Fan, Google's first iteration of a Go-playing system that beat a human grandmaster required 176 GPUs distributed over many machines (Silver, et al., 2017). Finally, as capable as these systems become, they often prove brittle in the face of changing data and display poor performance when operating outside of a few narrowly-defined tasks (Shrivastava, 2017).

III. What is the Research Landscape?

Top Research Disciplines: Electrical Engineering and Computer Science

- *Prevalence of engineering-related categories (**Electrical Engineering, Computer Science**) suggests that most of the innovations occur in deep learning applications, primarily in **imaging, acoustics** and **telecommunications.***

Top Research Topics: Deep Learning Techniques and Speech Recognition

- *Deep neural network models and techniques are highly ranked terms, particularly **Convolutional Neural Networks**, which are used in vision processing, **Restricted Boltzmann Machines**, used for extracting representations out of noisy data, and **Recurrent Neural Networks,** which have found use in machine translation and text analysis.*
- ***Speech Recognition** stands out as a topic that is rapidly gaining prominence, despite it not being referenced in the original query.*

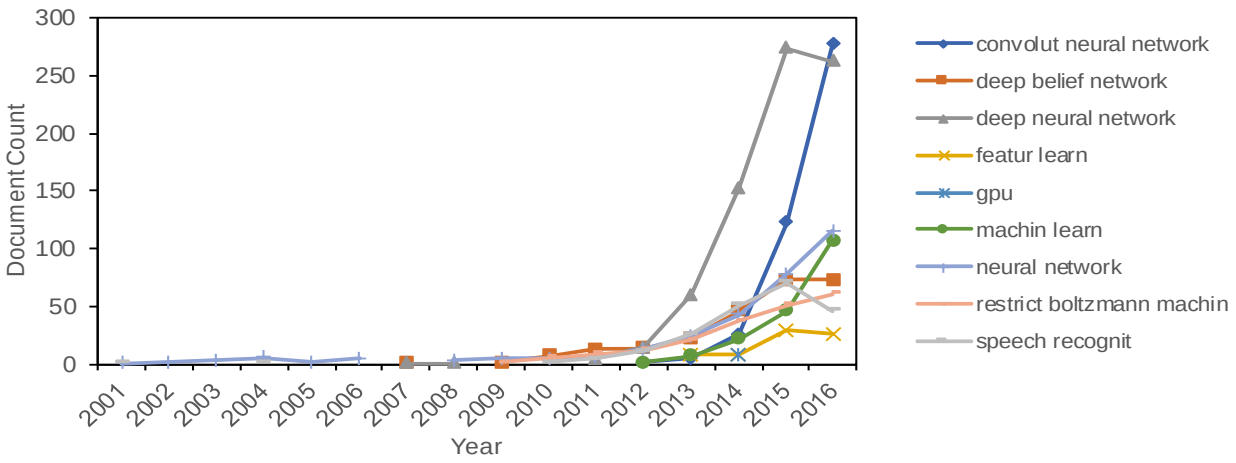
Top Application Areas: Signal Processing, Medicine, Materials Research

- *While the terminology within patents appears to be fairly general (e.g. "networks", "neural"), these terms appeared in a variety of classification categories not traditionally associated with computer science.*
- ***Audio/Video Recording Systems**, **Computer Peripheral Equipment,** and **Musical Instruments, Acoustics** are listed as top ranked patent classification areas. Likewise **Speech Recognition** is one of the most popular and fastest growing keywords in recent years.*
- ***(s) Electrical Medical** and **(p) Diagnosis Surgery** are also highly ranked patent classifications.*
- *The patent classification areas of **(b) Natural Products and Polymers** and **(d) Fermentation Industry** stand out as potential application areas for deep learning for data analysis.*

Research fields are comprised of different thrust categories. Self-organization of articles often occurs around these thrust categories, key questions and drivers in the field. The semantic content of the technical language, namely the technical terms, is typically a good indicator for tracking this activity. Similarly, how this research populates curated hierarchical subject categories can indicate what disciplines influence and dominate the field as well as other fields where this research has been influential. To extract the research field of interest, the following Boolean query was used:

"deep learning" OR "deep reinforcement learning" OR "deep structured network" OR "deep q-network" OR ("deep learning" AND "recurrent neural network") OR "deep recurrent neural network" OR "bidirectional recurrent neural network" OR "long short-term memory" OR "deep stacking network" OR ("stacked auto encoders" NOT "alzheimer disease" NOT "orthopedic" NOT "medical") OR "deep neural net" OR "deep neural network" OR (dnn NOT physics) OR "deep generative models" OR "deep belief network" OR "deep Boltzmann machine" OR "restricted Boltzmann machine" OR ("rbm" AND "restricted Boltzmann machine") OR "deep convolutional nets" OR "deep convolutional network" OR ("cnn" AND "convolutional neural nets" NOT "cellular neural network" NOT cell NOT media) NOT education NOT psychology

---



Top Publication Keywords

Rank	Keywords	Counts	Citations
1	deep learn	1,633	5,980
2	deep neural network	772	2,317
3	convolut neural network	433	933
4	neural network	311	2,766
5	deep belief network	252	2,393
6	speech recognit	216	1,576
7	restrict boltzmann machin	201	919
8	machin learn	185	491
9	recurr neural network	180	1,720
10	long short term memori	132	469

Top Patent Terms

Rank	Term	Counts	Citations
1	method	660	1,248
2	network	371	616
3	based	370	1,074
4	deep	320	336
5	learning	281	280
6	neural	255	449
7	image	229	278
8	data	221	801
9	process	209	287
10	model	198	224

Top Publication Subjects

Rank	Subject	Counts	Citations
1	engineering, electrical & electronic	2,352	11,902
2	computer science, artificial intelligence	2,205	14,137
3	computer science, theory & methods	1,197	2,583
4	computer science, information systems	742	1,111
5	acoustics	630	4,224
6	computer science, interdisciplinary	614	1,836
7	imaging science & photographic technology	522	1,444
8	computer science, software engineering	276	712
9	telecommunications	263	416
10	computer science, hardware & architecture	253	1,243

Top Patent Classifications

Rank	Classification	Counts	Citations
1	(t) digital computers	836	1,433
2	(w) audio/video recording and systems	334	463
3	(t) computer peripheral equipment	143	118
4	(w) telephone and data transmission systems	76	250
5	(s) electrical medical	49	643
6	(p) musical instruments, acoustics	29	235
7	(b) natural products and polymers	28	695
8	(p) diagnosis, surgery	28	510
9	(s) scientific instrumentation	25	104
10	(d) fermentation industry	18	160

Figure 2 These figures are derived from a search query performed to define the deep learning field. **(Top)** Shows popular, stemmed keywords used in Web of Science over time. **(Middle, Left)** Shows the popular keywords used in Web of Science. **(Middle, Right)** Shows the popular terms used in the title section of patents in the Derwent Patent Index. **(Bottom, Left)** Shows the subject categories used in Web of Science. **(Bottom, Right)** Shows the patent classification keywords used in the Derwent Patent Index.

IV. How Mature is the Field?

ONTA estimates that Deep Learning is rapidly maturing with a moderate footprint in both research and innovation. It bases this judgment on the factors below.

Size: Medium

- *The field is moderately sized with ~6,000 publications and ~1,000 patents.*

Growth: Rapid

- *While deep learning was near stagnant from 1997 to 2012, the field has grown rapidly in the last 4 years; this is likely attributable to the breakthrough in using graphical processing units (GPUs) for timely and less expensive machine learning processing.*
- *The amount of articles published in 2014 is 2.5 times that published in 2015 (~750 vs. ~1800); 2015 to 2016 saw 140% growth in publications.*
- *Patents also experienced high growth since 2012, quadrupling in the last three years.*

Influence: Moderate

- *Publications in this field average about 5 citations per paper.*
- *Patents average about 3 citations per filing.*

Maturity: Developing

- *The ratio of patents to publications is ~1:6, indicating a high level of transition from basic research to real-world applications.*
- *However, because the field is moderately sized and both publications and patents are experiencing strong growth, the field and associated technologies are still developing.*

	Publications	Patents
Document Counts	6,180	902
Citations	28,619	2,351
Authors/Inventors	13,954	1,926

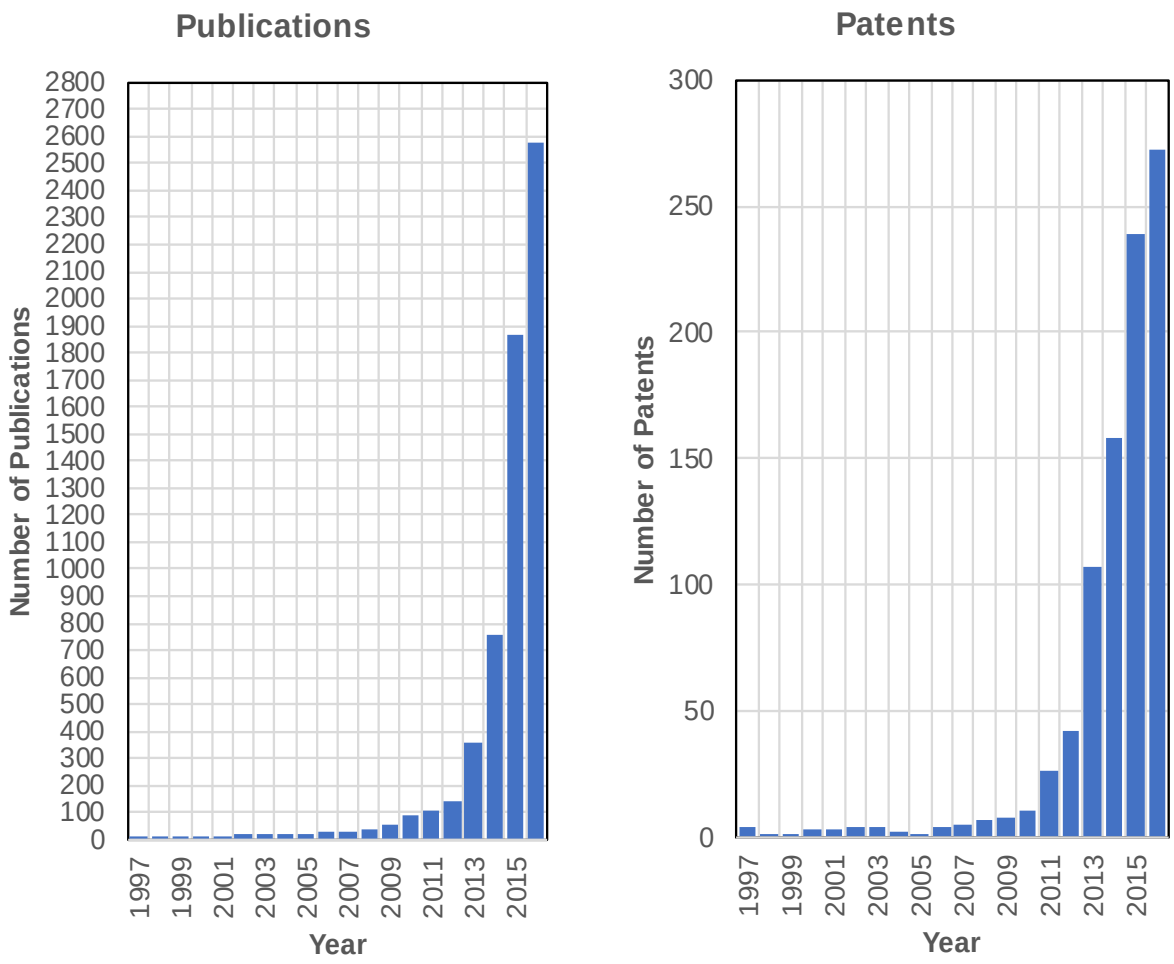


Figure 3 (Top, Left) Shows the *document counts, author counts and citation counts* for Deep Learning in the Web of Science. These fields correlate to *accumulated knowledge, workforce size and influence/quality*, respectively. **(Top Right)** Shows the *patent counts and citation counts* for the field of Deep Learning in the Derwent Patent Index. **(Bottom Left)** Shows the number of articles and conference proceedings mentioning Deep learning in the title or abstract per year of publication. **(Bottom Right)** Shows the number of patents mentioning Deep Learning in the title or abstract per year of publication. The current year has been omitted since not all the publications for this year have been indexed.

V. What are the Leading Countries?

Top Researcher: China & U.S.

- ***China*** *is the top producer of academic research, with the* ***U.S.*** *lagging by approximately 10%.*
- *Both the* ***U.S.*** *and* ***China*** *have high growth rates in this area, though* ***China's*** *growth rate is slightly higher.*
- *The* ***U.S.*** *has the most research citations, with 1.5 times greater than its closest competitor, China.*
- *The majority of growth in the field has occurred has occurred in the last 4 years and is almost entirely attributable to* ***China*** *and the* ***U.S.***

Top Innovator: U.S. & China

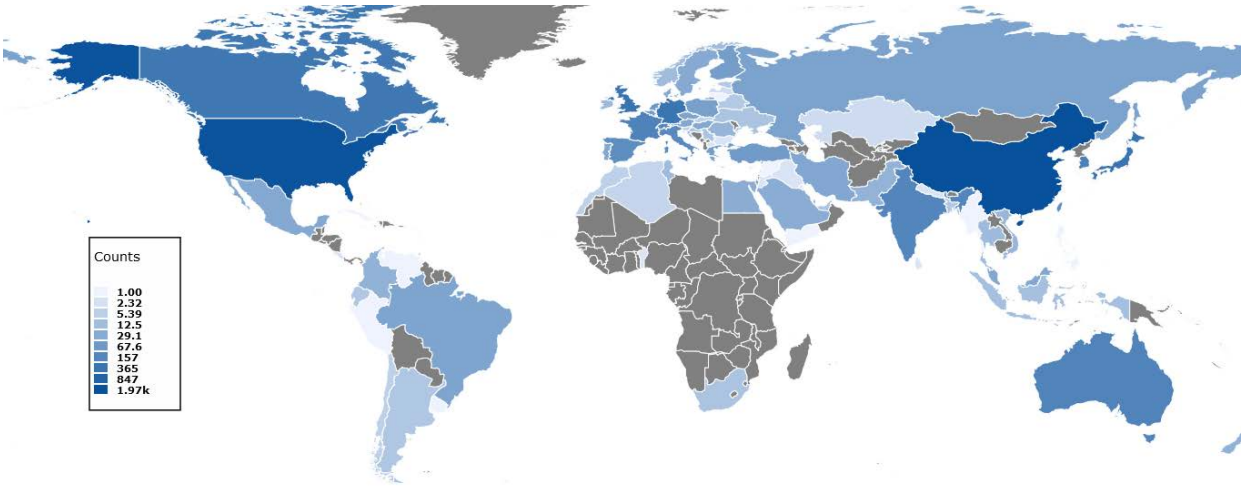
- *The* ***U.S.*** *and* ***China*** *produce a near equal number of patents.*
- *The* ***U.S.*** *has the most number of patent citations, with nearly 10 times that of its closest competitor,* ***China.***

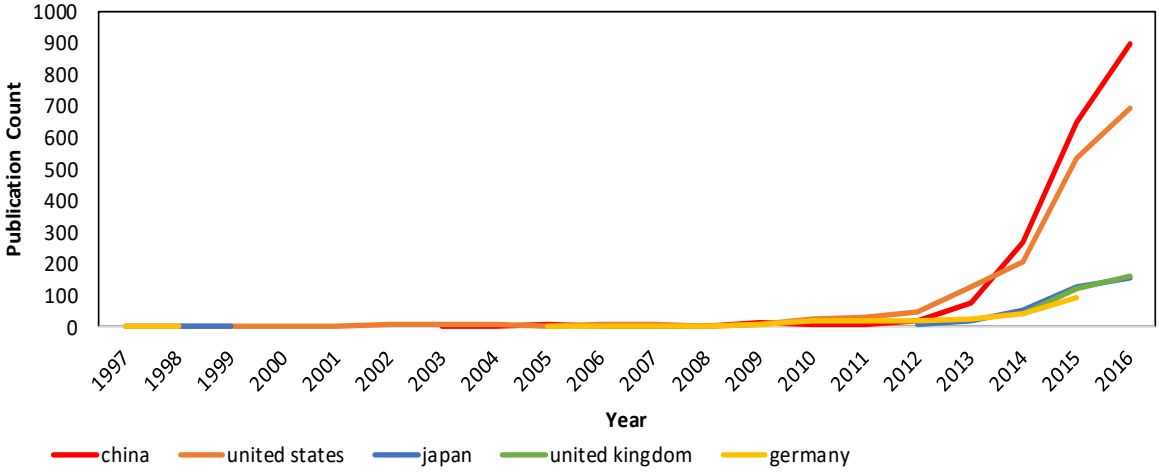
Top Market for Innovation: U.S. & China

- *In terms of citations, the* ***U.S. Patent Office*** *hosts the highest quality patents, outperforming its closest competitor,* ***China****, by a factor of 2.*

China currently has a slight lead over the U.S. in the total volume of research produced, but China's high growth rate in this area could increase their lead in the near future. While China is prolific in both publication and patent spaces, they are not as highly cited as those from the U.S, indicating lower impact research and innovation. The U.S. remains highly competitive, leading in innovation. Other strong countries include Japan (#4 in articles, #4 in patents) and Canada (#5 in articles, #6 in patents), who, despite their lower publication and patent rate, are the #3 most cited in publications.

The country affiliation of articles and patents are inferred through the mailing address. These broadly indicate a country's contribution and expertise. Additionally, we can also analyze the patent-granting authority, which is often associated with a country and indicates where an innovation area has the most protection and coverage.





Top Countries by Publication

Rank	Country	Articles	Citations
1	china	1,967	7,036
2	united states	1,721	11,205
3	united kingdom	385	1,982
4	japan	379	1,534
5	germany	353	2,754
6	canada	309	6,967
7	singapore	222	827
8	south korea	201	311
9	france	187	899
10	australia	177	700

Top Countries by Patents

Country	Patents	Citations
united states	382	1,943
china	377	225
south korea	44	4
japan	39	704
germany	28	138
canada	19	96
united kingdom	13	70
france	5	3
spain	5	15
belgium	4	2

Top Patent-Granting Offices

Office	Patents	Citations
united states	460	2,192
china	432	1,038
world intellectual property organization	203	1,208
european patent office	107	1,080
south korea	67	651
japan	58	811
india	27	636
canada	23	671
australia	19	649
germany	19	194

Figure 4 Shows the top countries in Deep Learning based on: **(Left)** the country affiliation of the authors in the Web of Science, **(Middle)** the country affiliation of the assignee from the Derwent Patent Index, **(Right)** the patent granting authority (typically a country) from the Derwent Patent Index.

VI. Questions for Further Study

Snapshot reports are meant to be quick scans of S&T and to ultimately stimulate interesting questions using only statistical data from the S&T literature. Answering these questions requires other methods, including interviewing stakeholders and experts.

International Competition

- *How are different countries building up human capital for the next generation of deep learning research?*
- *While China is still behind in the U.S. in terms of quality, its citation counts are rising; in which research or application areas is China beginning to excel?*
- *How does data access differ between the U.S. and its top competitors?*

Technology Advancements

- *What are the potential breakthroughs in developing low-power, low-processing deep learning systems?*
- *What are the potential breakthroughs in developing systems that can learn from first-principles in order to enable generalizability?*
- *What are the advancements in automatic data generation and labelling?*
- *How do the capabilities of deep learning systems compare with those of other methods of machine learning?*

Impacts

- *What are the common weaknesses or biases of deep learning systems as opposed to humans when evaluating data?*
- *How quickly does research transition from laboratory experimentation to real-world applications?*
- *What are the newest problems to which deep learning is being applied?*

VII. Further Reading (Most Cited Work)

Rank	Citations	Article Title	Source	Year
1	918	Long short-term memory	NEURAL COMPUTATION	1997
2	912	Deep Neural Networks for Acoustic Modeling in Speech Recognition	IEEE SIGNAL PROCESSING MAGAZINE	2012
3	704	Deep learning	NATURE	2015
4	589	Representation Learning: A Review and New Perspectives	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	2013
5	575	Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition	IEEE TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING	2012
6	510	Dropout: A Simple Way to Prevent Neural Networks from Overfitting	JOURNAL OF MACHINE LEARNING RESEARCH	2014
7	484	Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion	JOURNAL OF MACHINE LEARNING RESEARCH	2010
8	434	Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles	PROTEINS-STRUCTURE FUNCTION AND GENETICS	2002
9	348	Acoustic Modeling Using Deep Belief Networks	IEEE TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING	2012
10	343	Deep learning in neural networks: An overview	NEURAL NETWORKS	2015

Rank	Citations	Patent Title	Assignee	Year
1	475	Computerized method for evaluating glycosylated hemoglobin of patient based on blood glucose data, involves computing weighted deviation of blood glucose and estimating using predetermined mathematical relation	UNIV VIRGINIA PATENT FOUND	2000
2	91	System for feature detection with Z-factors for artificial intelligence, has set of detected feature stored on primary memory unit and Z-factor corresponding to set of detected features is determined based on energy measure and baseline	-	2012
3	72	Software application development environment adds number of additional metal data to metadata of class to facilitate development of software application, and organizes additional metadata and metadata in a tabular fashion	TABLECODE SOFTWARE CORP	2003
4	69	Representing set of reference face images corresponding to set of first vectors in input space of first dimension by projecting first vectors to high dimensional feature space of a second dimension using projection function	HONDA MOTOR CO LTD	2001
5	69	Wind-power method for operating a wind energy installation adjusts the angle of a rotor blade while reducing rotor speed	SENVION SE, REPOWER SYSTEMS AG, SENVION GMBH, REPOWER SYSTEMS SE	2006
6	55	Computer readable apparatus for implementing generalized learning rules in stochastic systems in computerized spiking network, has set of instructions for operating group in accordance with learning rule based on input signal	BRAIN CORP	2012
7	48	Stochastic signal processing system to implement task-specific learning, has control parameter that is provided in accordance with task, and adjustment signal that is provided to modify parameter based on input and output signals	BRAIN CORP	2012
8	42	New stabilized immune modulatory RNA compound for Toll-like receptor 7 and 8, useful for treating or preventing cancer, autoimmune disorder, inflammatory disorders, infectious disease, skin disorders, allergy, or asthma	HYBRIDON INC	2006
9	42	Posteriori probabilities of phoneme symbols calculating apparatus extracts speech feature parameters from speech signals of uttered speech sentence of input character series, calculates posteriori probability of speech symbol phoneme symbol by using bidirectional recurrent neural network	ATR INTERPRETING TELECOM RES LAB, ATR ONSEI HONYAKU TSUSHIN KENKYUSHO KK	1997
10	41	Collaborative filtering method for use in e.g. music distribution system, involves determining recommendation rating for selected item based on neighborhood model and selectively recommending selected item to user in response to rating	AT & T CORP	2008

Figure 5 The following list shows **(Top)** the most cited articles in the deep learning field from the Web of Science and **(Bottom)** most cited patents from the Derwent Patent Index.

VIII. About this Publication

Referenced work in this publication does not constitute endorsement by the United States Department of Defense (DoD) of the linked web sites, nor the information, products or services contained therein. In addition, the content featured does not necessarily reflect DoD's views or priorities. To provide feedback or ask questions, contact us at asdre-st-bulletin-reply@sainc.com. This publication is authored and distributed by:

- **Ryan Zelnio, Ph.D.**, Associate Director - Tech Watch/Horizon Scans, Office of Net Technical Assessments, OSD AT&L/OASD(R&E)
- **Artjay Javier, Ph.D.**, ONTA S&T Analyst
- **David Byrd**, ONTA S&T Analyst
- **Danielle Paynter**, Detailee to ONTA

IX. References

Allison, I. (2017, August 29). *AI, Algorithms and the Future of Finance: Companies Use Deep Learning to "Read the Tea Leaves" in Market Data.* Retrieved from Newsweek: http://www.newsweek.com/ai-algorithms-and-future-finance-companies-use-deep-learning-read-tea-leaves-656366

Deng, L., & Yu, D. (2014). *Deep Learning: Methods and Applications.* NOW Publishers. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Deep_learning

DWPI Classification System. (n.d.). Retrieved from http://ip-science.thomsonreuters.com/support/patents/dwpiref/reftools/classification/

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Chapter 1: Introduction. In *Deep Learning* (pp. 1-26). MIT Press, http://www.deeplearningbook.org/.

Hadad, Y. (2017, March 16). *30 Amazing Applications of Deep Learning.* Retrieved from Yaron Hadad: http://www.yaronhadad.com/deep-learning-most-amazing-applications/

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, 1527-1554.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, p.436-444.

Lecun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). *Backpropagation Applied to Handwritten Zip Code Recognition.* Holmdel, NJ: AT&T Bell Laboratories.

McCulloch, W. S., & Pitts, W. (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics, 5*, 115-133.

Meade, B. (2017, April 21). *Neural Networks & Earthquakes*. Retrieved from MGHPCC: https://www.mghpcc.org/neural-networks-earthquakes/

OECD Cateogry Scheme. (n.d.). Retrieved from http://ipscience-help.thomsonreuters.com/incitesLive/globalComparisonsGroup/globalComparisons/subjAreaSchemesGroup/oecd.html

Parloff, R. (2016, September 28). *Why Deep Learning is Suddenly Changing Your Life*. Retrieved from Fortune: http://fortune.com/ai-artificial-intelligence-deep-machine-learning/

Peck, M. (2016). Talking Hypersonics. *Aerospace America, 54*(4), 8-11.

Perez, C. (2016, November 17). *40 Ways Deep Learning is Eating the World.* Retrieved from Medium: https://medium.com/intuitionmachine/the-ultimate-deep-learning-applications-list-434d1425da1d

Raina, R., Madhavan, A., & Ng, A. Y. (2009). Large-Scale Deep Unsupervised Learning using Graphics Processors. *Proceedings of the 26th International Conference on Machine Learning.* Montreal, Canada.

Rosenblatt, F. (1958). The Perceptron: A Probabalistic Model for Information Storage and Organziation in the brain. *Psychological Review, 65*(6), 386-408.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol 1* (p. 3180362). Cambridge, MA: MIT Press.

Schmisseur, J. (2015). Hypersonics into the 21st century: A perspective on AFOSR-sponsored research in aerothermodynamics. *Progress in Aerospace Sciences, 72*, 3-16.

Shrivastava, P. (2017, July 19). *Challenges in Deep Learning*. Retrieved from Parallel Dots: http://blog.paralleldots.com/technology/deep-learning/challenges-in-deep-learning/

Silver, D., Schrittwieser, J., Simonyan, K., Antonglou, I., Huang, A., Guez, . . . al, e. (2017). Mastering the Game of Go without Human Knowledge. *Nature*, 354-359.

Trost, J. (2016, December 30). *Collection of Deep Learning Cyber Security Research Papers.* Retrieved from Medium: https://medium.com/@jason_trost/collection-of-deep-learning-cyber-security-research-papers-e1f856f71042

X. APPENDIX

A. Scientometric Methodology

TechSight is an open-resource, cloud-based ecosystem developed and maintained by ONTA. As of this writing, it consists of an ElasticSearch database infrastructure with a Kibana front-end and some commercial and custom-written plugins. The databases used for this analysis were global scientific publications using the Web of Science and global patent applications using the Derwent Patent Index, both provided by Clarivate, Inc. All of the figures generated in this report come from visualizations generated by Techsight.

More dashboards for this specific report are available on the ONTA TechSight system and contain additional visualization elements not included in this report since their dynamic nature is not compatible with static reporting. These include network visualizations that allow for a finer-grained analysis and allow the user to delve into specific information on top performing universities, companies, authors and inventors. To access it, sign up for a free account at:

https://registration.761link.net/accountRequest-ZoneB/accountRequest/techsight.

You must be either a DoD employee, or a contractor supporting DoD, and register using your .mil e-mail address.

A search query is manually developed by an analyst to capture best capture the field of this report. Development of this query is directed at improving precision (by eliminating non-relevant documents from the results) and recall (by collecting as many relevant documents as possible) through the use of Boolean operators and unique terms. Since ElasticSearch is being used, differences in term suffixes are automatically accounted for and require no additional specification.

B. How large is a research field or area of innovation? (Frequency Analysis)

The size of a research field can be estimated in terms of total aggregated knowledge, for which the metric cumulative document counts (x) is a suitable proxy. Under the assumption that every article is unique and therefore constitutes a single unit of knowledge, the sum of all the articles in a research field approximates the total knowledge accumulated in the field. Another suitable metric is total community size, for which the number of unique authors is a suitable metric since these are the workers that generate knowledge. A larger workforce tends to correlate to a greater capacity to produce knowledge and therefore grows proportionately with aggregated knowledge. Some fields exhibit differing productivity (i.e. documents per unique worker) depending on ease of publishing, difficulty in carrying out experiments and field-dependent variables. Fields like particle physics and clinical medicine tend to have articles with a large number of authors due to the difficulty of

the experiments. Fields such as nanoscience and nanotechnology tend to have higher productivity due to the ease of publishing new results. Fields like mathematics tend to have only a single author due to the nature of the work, and fields in computer science tend to have generally low publication rates relative to their research production. Similar factors affect patent indicators and are notably shaped by key differences between the two corpora, such as motivations for publishing versus patenting, the differences between peer review and patent examination, and the choice of technical terminology. Field sizes and influence are based on analyst observations and experience in a semi-quantitative rough order of magnitude sense: very small fields $0 \leq x < 100$ articles, small fields $100 \leq x < 1{,}000$ articles, medium fields $1{,}000 \leq x < 10{,}000$ articles, large fields $10{,}000 \leq x < 100{,}000$ articles, very large fields $x > 100{,}000$ articles. For influence: poorly cited <1 citation/article, medium citation rate ~1 citation/article, high citation rate >10 citations/article.

C. How influential is a research field? (Citation Analysis)

Scientific articles contain a list of references that cite previously published articles. The number of times an article has been referenced by other articles is called its citation count. Over time, an article's citation count tends to increase as subsequently published articles cite that article. Citation count tends to correlate with an article's influence, indicating the article's content has influenced other articles. Citation is also a suitable proxy for quality, as more articles describing the first reports of original work tend to have higher citations. An exception to this rule are review articles which tend to have very large citation counts and contain no original work but are cited typically to point new readers to a compact source for their further education in the field. Despite this exception, it is not inappropriate to include review articles in a citation analysis because the articles tend to be more widely read, and are a demand signal that a field has aggregated enough knowledge that a convenient repository for that knowledge is desirable. Since citation counts provide a usable proxy for "quality", this analysis provides a counterbalance to the "quantity" metric of document counts.

D. How fast is the research field or area of innovation growing? (Trend Analysis)

Scientific fields grow over time as researchers publish related articles, building on early seminal works. Emerging and potentially disruptive research areas typically display rapid, exponential-like growth early in their lifecycle.

E. What are the key areas of research, development and innovation? (Semantic Analysis)

The content of a research field can be understood from a hierarchical framework. Understanding the parentage of the field creates awareness of the nature and character of the field relative to the context of current scientific organization. As a proxy, we use the Web of Science's Subject Categories field, which are inspired by OECD's Field of Science (FOS) categories (OECD Cateogry Scheme, n.d.). While a field tends to localize around a specific section of this hierarchy, outliers sometimes exist arising from relevant articles in unrelated research fields indicating this field has influenced work or been adopted by these other fields. Similarly, patents in the Derwent Patent Index (DWPI Classification System, n.d.) are inspired by the WIPO classification and section scheme and lend themselves to similar visualization schemes. Research topics can often be conceptually subdivided into sub-topics. These sub-topics are often differentiated by specific keywords which are indicative of the content of these subtopics and represent segments of research focusing on research drivers such as key questions or specific innovations. Quantitatively tracking these keywords indicates the relative popularity of these sub-topics.

F. What are the leading countries? (Country Cross-Analysis)

Authors and inventors are affiliated with organizations whose addresses are in specific countries. By subdividing the data according to country, we can produce analyses at the national level that broadly indicate a country's participation level in a research field. Top 10 Countries by Publications are determined by the address of the affiliation of the author in the Web of Science. Note that an author can have multiple affiliations, thus belong to multiple countries. Top 10 Countries by Patent Application are determined by the address of the affiliation of the assignee in the Derwent Patent Index. An alternative approach is to use the inventor affiliation, which results in larger country counts since a patent can have multiple inventors, but only one assignee. Patent protection can be granted by applying to nation-specific authorities (i.e. U.S. Patent and Trademark Office), regional authorities (i.e. European Patent Office) or international authorities (World Intellectual Property Organization). It is often useful to compare which countries patents in a specific technology are granted and comparing that to where those companies are affiliated as it indicates whether one country is seeking IP protection in another country, or worldwide, for its products.