From artificial intelligence to data science: Challenges and opportunities of information fusion

Chee-Yee Chong
Independent Researcher
Los Altos, California, USA

cychong@ieee.org

Presented at 20th International Conference on Information Fusion
Xi’an, China, 12 July 2017
Outline

- Artificial intelligence (AI) and data science boom
- AI and information fusion
- Data science and information fusion
- Opportunities for information fusion
Artificial Intelligence (AI) Is Red Hot

• AI has recent spectacular successes defeating humans
  • Deep Blue beat world champion in chess (1997)
  • Watson won US Jeopardy quiz show (2011)
  • AlphaGo beat world’s top player (Ke Jie) in go (2017)
• AI is in everyday applications
  • Smart phones – speech recognition, machine translation
  • Homes – smart thermostat, robotic vacuum
  • Cars – driver assistance
• Companies and governments have big investments in AI
  • Hardware: Nvidia (GPU), IBM (Watson), Intel
  • Internet and software: Microsoft, Google, Amazon, Facebook, Alibaba, Baidu, Tencent
  • Automation: automotive, manufacturing
  • Health care: drug discovery, personalized medicine
China is a Leader in AI Research and Applications

• China leads in publishes papers in deep learning (a form of AI) since 2014
• China's National Engineering Laboratory of Deep Learning Technology was established at Baidu Campus (2017)

Source: MIT Technology Review, July/Aug 2017
Data Science Has Explosive Growth

- Data science (data analytics, data mining) is growing rapidly due to
  - Large amounts of data available
  - Fast computing to exploit data
- In addition to traditional data sources, data are available from
  - Internet of Things (IoT)
  - Social media (Facebook has 2 billion users, Tencent has almost 1 billion WeChat users)
- Many applications
  - Business intelligence in retail, manufacturing, etc.
  - Ecommerce and social media: Facebook, Amazon, Netflix, Google
  - Pharmaceutical and drug discovery
  - Personalized medicine
Industries Affected by AI and Data Science

Internet and Cloud
- Image classification
- Language processing
- E-commerce tagging
- Digital personal assistants (e.g., Amazon’s Alexa and Apple’s Siri)
- Product recommendations

Health Care
- Wearable health data recognition
- Cancer cell detection
- Diabetic grading
- Drug discoveries

Media
- Video search
- Captioning
- Programming recommendations (e.g., Netflix and Comcast)
- Virtual and augmented reality

Security and Defense
- Face detection
- Video surveillance
- Geolocation
- Real-time objects and threat detection (e.g., detect explosives and match faces to criminal databases in real time)

Automation
- Store automation (e.g., Amazon’s new grab-and-go supermarket)
- Factory automation
- Drones
- Investment and insurance automation
- Self-driving automobiles

Source: Nvidia Corporation, Macquarie Research, and T. Rowe Price, 2017
AI and Data Science Pose Challenges to Information Fusion

• AI and data science address many traditional problems in information fusion
  • Low-level processing and object recognition
  • Video surveillance
  • Activity detection and behavioral analysis
  • Network and patterns of life analysis
• AI and data science are more visible because of
  • Better performance than humans in several applications
  • Use in many real systems
• Consequence
  • Sponsors turn to AI and data science to solve information fusion problems
  • AI and data science attract more students and researchers
Example: CVPR 2017, Hawaii, July 21 - 26

- Papers: 2620 submissions, 783 presentations
- Registrations: 4200, Sponsors: 127, Sponsor income: $859K
Perspectives from Silicon Valley

1980

- Government/non-profit
  - NASA Ames
  - SRI
- Large aerospace/defense
  - Lockheed
  - TRW
  - Ford Aerospace
  - Loral
  - GTE
- Small companies
  - Systems Control
  - Advanced Information & Decision Systems (AIDS)

2017

- Government/non-profit
  - NASA Ames
  - SRI
- Large aerospace/defense
  - Lockheed Martin
  - Loral
  - General Dynamics
  - BAE Systems
- Computing
  - Apple
  - Intel
  - Nvidia
  - Samsung
- Automotive
  - Tesla
  - BMW
  - Mercedes
  - Toyota
- Internet/software
  - Google
  - Facebook
  - Microsoft
  - Amazon
  - Netflix
  - Baidu
  - Palantir

Vax 11/750 Minicomputer
3 Mhz 120K FLOPS, 2 MB RAM
1200 baud dialup modem

IPhone 6
1.5 Ghz 115 GFLOPS, 1 GB RAM
100 Mb network

Bar-Shalom until 1977
Reid (MHT)

me
Information Fusion is a Natural Application for AI

- Humans solve information fusion problems all the time
  - Low level perception – environment, objects
  - High level understanding – situation, prediction

Traditional approach relies on humans to develop algorithm

AI allows automation of algorithm development
AI Approaches for Fusion

• Expert systems (mimic human experts)
  • Medical diagnosis
  • Signal understanding
• Probabilistic reasoning (with models)
  • Situation understanding
  • Object recognition
• Neural networks / deep learning
  • Feature extraction
  • Speech understanding
  • Object recognition
  • Video tracking

First wave of AI
Handcrafted knowledge*

Second wave of AI
Statistical learning*

Third (future) wave of AI
Contextual Adaptation*

*J. Launchbury, “A DARPA perspective on artificial intelligence,” 2017
Expert Systems for Fusion – MYCIN (~1975)

• Medical diagnosis
  • Inputs: test results
  • Outputs: infectious disease
• Rule-based system
  • Knowledge base of few hundred IF_THEN rules: IF symptom A THEN disease B
  • Inference engine by backward chaining
  • Certainty factors to represent uncertainty
  • Heuristic combination of evidence
• Performs better than many doctors
• Stimulated research on uncertainty reasoning

Expert Systems for Fusion – HASP/SIAP (1970’s)

- Signal understanding system
  - Inputs: acoustic signals from hydrophones
  - Outputs: detection, location and type of vessel
- Rule-based system
  - Hierarchy of rules for signal to symbol transformation
  - Inference by knowledge sources responsible for different levels of processing

Issues with Expert Systems for Information Fusion

- Knowledge acquisition
  - Finding experts that can articulate their reasoning; experts with good intuition may not be suitable
  - Extracting knowledge from experts
- Knowledge representation
  - Consistency and completeness of rules
  - Representation of uncertainty
- Inference engine
  - Control of inference
  - Reasoning with uncertainty
  - Processing speed
AI in 1980’s

• 1981 – Japan started “Fifth Generation Computer” project to build intelligent computers
• United States responded with “Strategic Computing Initiative” with AI as main objective, including Autonomous Land Vehicle (ALV) program
  • Demonstrated road following in 1985 demo
  • Vision system very sensitive to light and shadow
• Booming AI industry
  • Hardware companies: Symbolics, Lisp Machines, Inc., Thinking Machine Corp. (TMC)
  • Software companies: AI Corp, Carnegie Group, Intelicorp, Teknowledge
• Overhype ended in AI winter in late 1980’s

Artificial Neural Networks (1980’s)

- Motivated by physiology and function of neurons in brain
- Long history
  - McCulloh, Pitts – 1940’s
  - Widrow – 1960’s
  - Hopfield, Rumelhart, Hinton – 1980’s
- Weights learned from training data
- Excellent for low level recognition task
- Implementation issues
  - Black box approach cannot explain results
  - Performance sensitive to training data
Uncertainty Reasoning in AI

• Information fusion has to deal with uncertainty in evidence (input) and knowledge
  • Drawback of artificial neural network approach
  • Recognized very early by expert system developers
• Uncertainty reasoning approaches
  • Rule-based methods
  • Probabilistic reasoning
  • Evidence theory
    • Dempster Shafer
    • Dezert-Smarandache Theory (DSmT)
  • Fuzzy sets
• Probabilistic reasoning became very popular starting in 1980’s

Probabilistic Reasoning/Graphical Models

- Probability model expressed graphically as networks
  - Nodes are random variables
  - Weights on edges represent conditional probabilities
- Inference computes conditional probabilities given evidence using methods such as
  - Node elimination
  - Junction tree
  - MCMC
- Very natural for researchers with background in estimation theory
- Considered AI because of separation into knowledge and automatic inference
Evidence Accrual for Detection Hypothesis

• Additive rule for evidence accrual (Stein and Winter)\(^1\)

\[
P(H|E_O \cup E_N) = \frac{P(H)}{P(E_O \cup E_N)} [P(E_O|H) + P(E_N|H) (1 - P(E_O|H))] 
\]

• Justified as Probability Hypothesis Density (Mahler)\(^2\)

\[
D_{k|k}(x_k | Z_{1:k}) = \sum_{z \in z_k} \frac{p_{D,k}(x_k)g_k(z | x_k)}{\lambda_k c_k(z) + \psi(z | Z_{1:k-1})} + (1 - p_{D,k}(x_k)) D_{k|k-1}(x_k | Z_{1:k-1}) 
\]

\[
N_{k|k} = \int D_{k|k}(x_k | Z_{1:k})dx_k
\]

Expected number of objects

Model-Based Object Recognition

Model Driven ATR

Focus Attention

Index

Image & Hypothesis Space Reduction

Search

Hypothesize & Test

On-Line Models

Predict Features

Match Features

Extract Features

Target Reports

Context & Collated Information

Image

Sensor, Image Formation, and Acquisition Parameters
Moving and Stationary Target Acquisition and Recognition (MSTAR)

• Program manager said “neural network is not allowed”
AI for Information Fusion from 1990’s to 2010

• Probabilistic graphical models become main AI approach for fusion
  • Rigorous treatment of uncertainty
  • Model-based approach is explainable
  • Many Inference techniques
  • Models can be learned from data
  • Can be extended to handle evidence theory, e.g., valuation networks
• Mathematic framework is similar to that of object tracking
  • Predict features from model
  • Match extracted features with prediction (association problem)
  • Update probabilities
• Meanwhile, artificial neural networks are used for many low level functions where modeling is difficult and data is available for training
• Then, computers become more powerful and massive amounts of data are available
Deep Learning

- Deep neural network uses multiple hidden layers between input and output layers to model complex nonlinear relationships
- Input layers can be images or audio signals instead of features
What Makes Deep Learning Possible

- Deep learning is possible due to advances in computing hardware

Largest networks

10 layers
1B parameters
10M images
30 GPU days
Deep Learning is Attracting Much Attention from Information Fusion

- Deep learning has been successfully used in
  - Video surveillance
  - Object and threat detection
  - Driverless vehicles
  - Cyber security (where modeling is very difficult and data is plentiful)
- There is interest in application to more traditional problems
  - Information Fusion journal has Call for papers for a special issue on “Deep Learning for Information Fusion”
    https://www.journals.elsevier.com/information-fusion/call-for-papers/call-for-papers-for-a-special-issue-on-deep-learning-for-inf
  - US Air Force has solicitation for Target Tracking via Deep Learning
    https://www.sbir.gov/sbirsearch/detail/1208333
Data Science and Big Data Analytics

• Extracting insight from data has been around for a long time
  • Statistical analysis
  • Data mining
• Over the past twenty years, digitization and networking have collected massive amounts of heterogeneous data from
  • Enterprise data
  • Sensor networks
  • Internet sites
• Tools have been developed to handle BIG data
  • Preparation
  • Storage
  • Integration
  • Analysis: statistics, text analytics, predictive analytics, machine learning
Example: Large Synoptic Survey Telescope (LSST)

- Science goals
  - Understanding mysterious dark matter and dark energy
  - Hazardous asteroids and remote solar system
  - Transient optical sky
  - Formation and structure of Milky Way

- Big data problem
  - 15 terabytes (TB) \((10^{12})\) raw data each night
  - 60 petabytes (PB) \((10^{15})\) over 10 years

**Ultimate tracking and fusion (with radio telescopes)**

Telescope in Chile
- Three-mirror, three-lens optical assembly
- 8.4 meter primary mirror
- FOV 40 times size of full moon
- Image area: 64 cm diameter, 3.2 gigapixels
Data Science Is Widely Used for Risk Analysis

- Credit scoring (country, company, individuals)
- Insurance (personal, auto, home)
- Criminal justice system
  - Example: COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) for forecasting which criminals are most likely to reoffend
    - Has bias – false positives and negatives
    - Black box approach does not provide explanations

M. Spielkamp, “Inspecting algorithms for bias,” *MIT Technology Review*, July/Aug 2017
Data Analytics Has Always Been Key Part of Information Fusion

Issues of Deep Learning for Information Fusion

• Performance is only as good as data
  • Large amounts of data are needed
• Training data for rare events are sparse
  • Inherent flaws can be exploited
• Results are hard to explain
  • Black box provides no visibility
  • Research on explainable AI is still ongoing
• Thus few machine learning systems are used in critical missions or making life death decisions
## AI and Data Analytics Can Help Information Fusion

### Information Fusion Challenges vs. AI / Big Data Techniques

<table>
<thead>
<tr>
<th>Information Fusion Challenges</th>
<th>AI / Big Data Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big data</td>
<td>MapReduce, Hadoop, etc.</td>
</tr>
<tr>
<td>Hard soft heterogeneous data</td>
<td>Natural language understanding</td>
</tr>
<tr>
<td>Tracking in high confusion situations</td>
<td>Learning, graph techniques</td>
</tr>
<tr>
<td>Activity / behavior analysis</td>
<td>Learning</td>
</tr>
<tr>
<td>Network analysis</td>
<td>Graph techniques</td>
</tr>
<tr>
<td>Context aware processing</td>
<td>Learning</td>
</tr>
</tbody>
</table>

- AI and data science techniques can help information fusion when
  - Modeling is difficult
  - Training data are available (real or simulated)
Large Scale Target Tracking: Challenges and Approaches

• Challenges
  • Handling measurement and track data from multiple sources with different characteristics
  • Coping with sensing gaps and large numbers of false alarms
  • Maintaining track and ID for long durations despite target maneuvers and background traffic
  • Handling large numbers of targets and large volumes of data
  • Producing high confidence results

• Approaches
  • Feature-aided tracking with learning
  • Graph based association
Direct Learning of Sensor Input to Object Tracks

- Learns $P(y_t | x_{1:t})$
- No system or sensor model

Association Graph for Tracking

- Association graph provides efficient representation for possible associations
  - Nodes: measurements or tracklets
  - Edges: possible associations
  - Paths: tracks
  - Path cover: association hypothesis
- Track likelihood is sum of pairwise association likelihoods under Markov assumption
  \[ P(y_{i+1} | y_1, \ldots, y_i) = P(y_{i+1} | y_i) \]
- Best association hypothesis can be computed in polynomial time by
  - Bipartite matching
  - Minimum cost network flow

Learning Association Costs in Association Graph

- Pairwise association costs can be computed
  - Attributes of detections
  - Measurement models
- Alternatively, they can be learned from data using back propagation

Learning Features for Feature-Aided Tracking

- Before confusion: learn features from high confidence tracks
- After confusion: use features to resolve ambiguity

Computing Association Hypothesis Probability

• Hypothesis probability is important for assessing tracking confidence
• Probability of association hypothesis
  \[ P(H \mid Y) = C^{-1} \prod_{T \in H} L(T) \]
• Computing normalizing factor requires counting hypotheses, e.g., Murty’s algorithm\(^1\)
• For association problems that can formulated as bipartite matching, normalization factor is partition function\(^2\)
  \[ Z(G) = \prod_{1 \leq k \leq |V|} P^{-1}_{G_k} (v_k \notin M) \]
  recursive estimation of probability that a node is not in a matching, based on correlation decay

Conclusions

- Information fusion researchers have love hate relationship with computers
  - **Love** to develop algorithms for computers
  - **Hate** when computers become too smart to take over their job
- Users do not care about specific approach as long as it provides solution
  (Deng Xiaopeng: *It doesn't matter whether a cat is white or black, as long as it catches mice*)
- AI and deep learning still cannot replace algorithm developers for mission critical systems where trusted results are required
- Researchers should **move out of our comfort zone**
  - Use fusion in non-traditional applications
  - Exploit full potential of computers (e.g., MCMC for nonlinear filtering)
- We can also learn from computer vision community by providing challenge data sets (www.votchallenge.net, www.mot.challenge.net) to stimulate research and evaluate algorithms
Thank You
謝謝