Deep Learning for IoT Big Data and Streaming Analytics: A Survey

Presenter: Keren Ye

Mohammadi, Mehdi, et al. "Deep Learning for IoT Big Data and Streaming Analytics: A Survey." *arXiv preprint arXiv:1712.04301* (2017).

Introduction

Sensor data: big or fast/real-time data streams

Deep Learning (DL) in the IoT domain

- DL methods
- DL applications

Introduction - IoT data v.s. General big data

Large-Scale Streaming Data

- Large scale, stream data, distributed

Heterogeneity

Time and space correlation

- Sensors are attached to a specific location

High noise data

Introduction - IoT challenge: fast & accurate

Fast analytics in smaller scale platforms

- Fast decisions (e.g. autonomous cars)
- Edge computing or on-device computing

Multi-modal data processing

- Streaming data from multiple sources

Accurate recognition in real-time

Introduction - why DL

Discover feature by the model

- Avoid hand crafted and engineered feature sets
- Features that might not be discovered by human

Improved accuracy

- Learn from large amount of data

Contributions - A survey of

IoT data

State-of-the-art DL methods

- In IoT domain
- Big data & streaming data

Approaches and technologies for deploying DL

Challenges and future directions

IoT data characteristic

Be streamed continuously (real-time decisions)

- Using data parallelism, incremental processing
- Limitation of computing, storage, and power

Be accumulated as a source of big data (offline analysis)

- Mainly solved using cloud infrastructures, e.g. Apache Storm
- Lack of frameworks on fog or on device

Deep learning - Goal

Supervised methods

- Classification
- Regression

Unsupervised methods

- Clustering
- Compression

Deep learning - CNNs

Convolutional Neural Networks (CNNs)

- Assumes spatial dependency
- E.g. image or sensor data
    - Sensor data time dependency is based on the choice of window size

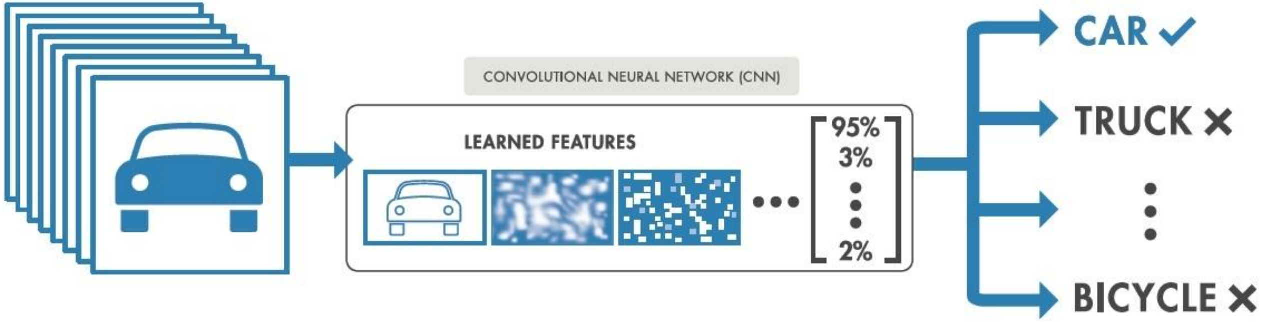


Image credit: Convolutional Neural Network
https://www.mathworks.com/discovery/convolutional-neural-network.html

Deep learning - RNNs (LSTM)

Recurrent Neural Networks (RNNs)

- Assumes long time dependency (models P(c|s0,s1,s2,...sn))
- E.g. speech signal, natural language processing

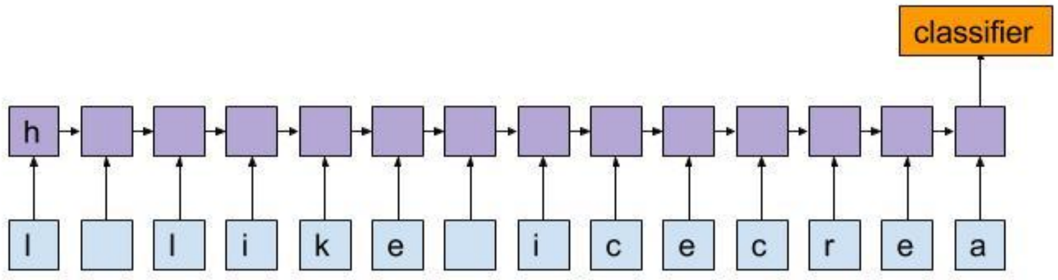


Image credit: How to read: character level deep learning
https://offbit.github.io/how-to-read/

Deep learning - AEs (VAEs)

Autoencoders (AEs)

- Compress (f) and reconstruct (g) data, s.t. g(f(x))=x

Variational Autoencoders (VAEs)

- Compress (f) and regenerate (g) the distribution, s.t. data x is likely to be drawn from the distribution g(f(x))

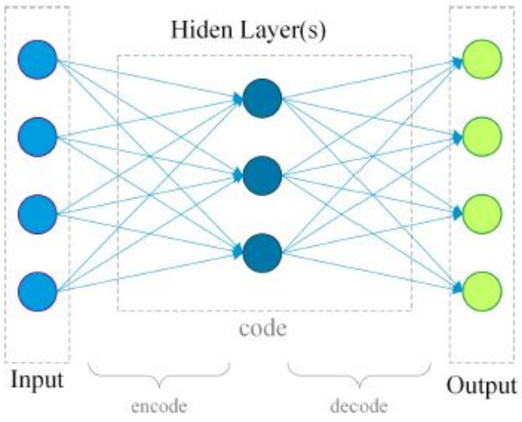


Fig. 8. Structure of an autoencoder network.

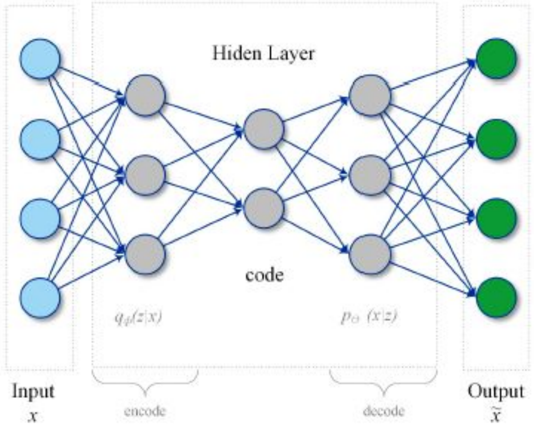


Fig. 9. Structure of a variational autoencoder network.

Deep learning - GANs

Generative Adversarial Networks (GANs)

- G: generate fake data
- D: try to find the fake data

Explanation, after training

- G can generate practical data that seems to be drawn from the data distribution
- D can distinguish the nuance from fake and real data

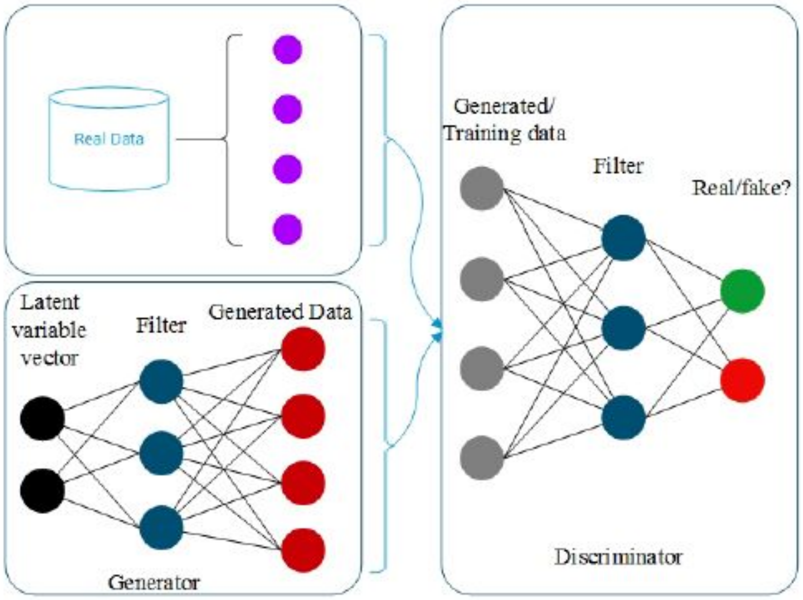
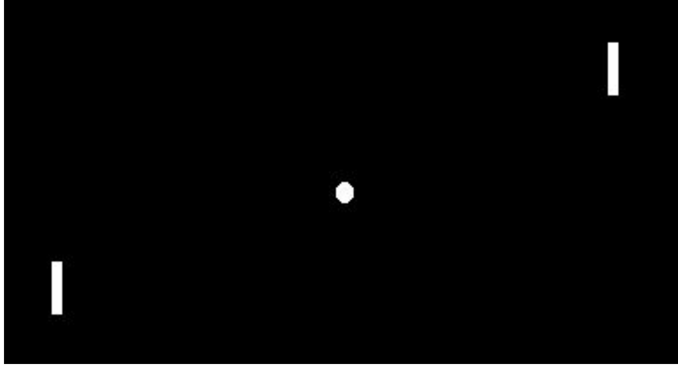


Fig. 10. Concept of a generative adversarial network.

Deep learning - RL



Deep reinforcement learning

- Assumes the feedback is delayed (compare to classification)
- E.g. HVAC control

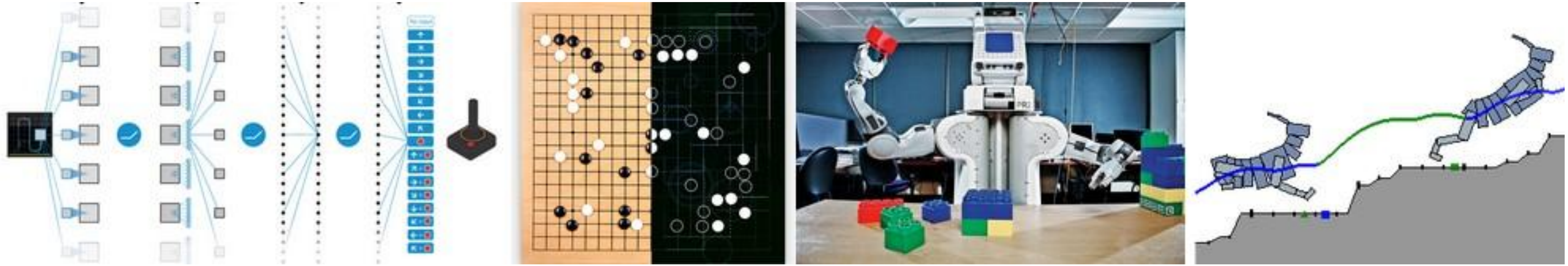


Image credit: Deep Reinforcement Learning: Pong from Pixels
http://karpathy.github.io/2016/05/31/rl/

Deep learning - and so on

Restricted Boltzmann Machine (RBMs)

Deep Belief Network (DBNs)

...

Deep learning - frameworks

Frameworks

- H2O, Tensorflow, Torch, Theano, Caffe, Neori

IoT (Mobile device) support

- Tensorflow Lite

DL Applications in IoT

A. **Foundational services**
B. Applications

DL Applications in IoT - Foundational services

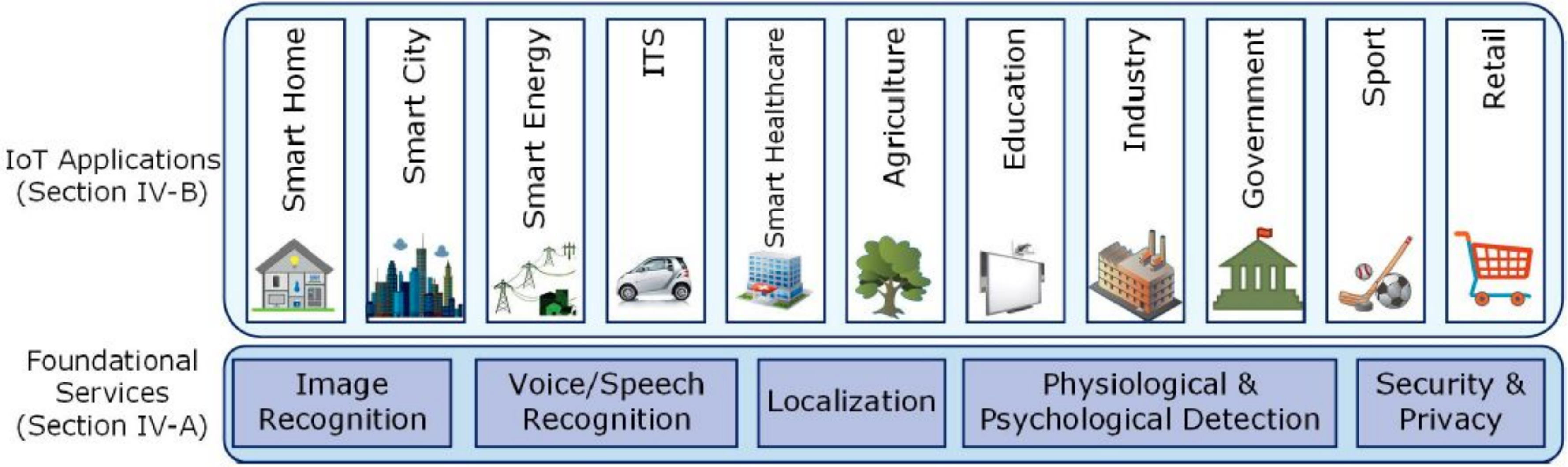


Fig. 15. IoT applications and the foundational services.

DL Applications in IoT - image / speech recognition

DL Applications in IoT - Indoor localization

Localizing user position (DeepFi, Wang et al.) - using fingerprinting WiFi channel state information

- Offline training: train all the weights based on the previously stored channel state information fingerprints
- Online inference: predict user positions

Localizing soccer robots (Lu et al.) - using LSTM

- Data from Inertia Navigation System (INS) and vision perceptions

DL Applications in IoT - Physiological and Psychological State Detection

Using DL model such as CNN and RNN to predict

- Human poses
- Human activity, using sensor data
- Gestures, using video frames
- Emotion, using video frames

DL Applications in IoT - Security

False Data Injection (FDI)

- Use Conditional DBM (He et al.)
  - to extract FDI attack features from historical data
  - Use these features for attack detection in real time
- Anomaly detection method (Yuan et al.)

DL Applications in IoT - Privacy

Preserve privacy in the training data

- Distributed learning (SGD)
- Performed in a parallel and asynchronous way

DL Applications in IoT

A. Foundational services
B. Applications

DL Applications in IoT - Smart homes

Home supplies (Microsoft and Liebherr) - information from refrigerator

- A better control on the home supplies and expenses
- Monitoring and predicting health trends

Energy efficiency (Manic et al.)

- Using DL models to predict energy consumption
- Comparison among DL models

DL Applications in IoT - Smart homes

Home care - Fall detection (Feng et al.)

- Use RBMs and DBNs
- Normal postures in such environment are standing, sitting, bending, and lying. Lying on the floor longer than a threshold is considered as a fallen posture.

DL Applications in IoT - Smart city

Multi-domains

- Transportation, energy, agriculture

Predicting crowd movements patterns (Song et al.)

- four-layer LSTM neural network to learn from GPS
- People’s mobility, transportation model as two separated tasks

Real-time crowd density prediction (Liang et al.)

- leverages the mobile phone users’ telecommunication data

DL Applications in IoT - Smart city

Waste management and garbage classification

- A vision based system

Empty spots in parking lots (Amato et al.)

- Camera sensor and deep CNNs
- Raspberry Pi 2 odel

DL Applications in IoT - Energy

Local energy consumption patterns

- make decisions based on real-time analytics
- Use RBM to predict energy flexibility in real-time (Mocanu et al.)

Smart grid

- forecasting the power from solar, wind, or other types of natural sustainable sources
- Performance of different prediction models (Gensler et al.)

DL Applications in IoT - transportation

Predict traffic congestion evolution (Ma et al.)

- RBM and RNN, long time dependency and short time dependency

Traffic sign detection and recognition

- For autonomous driving, driver assistance systems
- Real-time system using GPU (Lim et al.)

Pedestrians detection

DL Applications in IoT - Healthcare and wellbeing

Analysis of medical images

Time series medical data in conjunction with RNN based models for early diagnosis

DL Applications in IoT - Agriculture

Plant disease recognition (Sladojevic et al.)

- Based on leave images

Land and crop detection

- Remote sensing, automated monitoring
- CNN image processing

Prediction and detection tasks for automatic farming

- Obstacle detection

DL Applications in IoT - Industry

IoT and cyber-physical systems

Visual inspection

- For car production (Luckow et al.)

Fault detection (FDC)

- Auto-encoder based model (Shao et al.)

DL on IoT Devices

Trade-off

- Improve performance
- Reduce model size and inference time

Methods

- Network compression
- Approximate computing
- Accelerators - specific hardware and circuits

Fog and cloud-centric DL for IoT

Cloud computing

- may not be ideal (security, time constraints)

On Device

- Not easy to aggregating several sources of IoT data

Fog computing - a trade-off

- Avoid of transmitting large amounts of data to cloud
- Real-time since fog is hosted locally, close to data

Fog and cloud-centric DL for IoT

Intelligent IoT gateway (Al-Fuqaha et al.) - Edge

- Equipping IoT gateways and edge nodes with efficient DL algorithms can localize many complex analytical tasks that are currently performed on the cloud

Google's Tensor Processing Unit (TPU) - Cloud

- based on GPU servers

Fog for IoT - challenges

DL service discovery (protocols)

DL model and task distribution

- Distribute tasks and aggregating results - time sensitive

Design factors

Mobile edge

- Dynamically join and leave the system
- Energy management

Fog for IoT - attempts (proof of concept)

Qaisar et al.

- Deploying CNN models on fog nodes for machine health prognosis
- Find free nodes to delegate analytic tasks

Li et al.

- Running CNN models
- Leverages the collaboration of mobile and edge devices

Thanks