Fundamentals of Media Processing

Lecturer:
池畑　諭（Prof. IKEHATA　Satoshi）
児玉　和也（Prof. KODAMA　Kazuya）

Support:
佐藤　真一（Prof. SATO　Shinichi）
孟　洋（Prof. MO　Hiroshi）

Course Overview (15 classes in total)

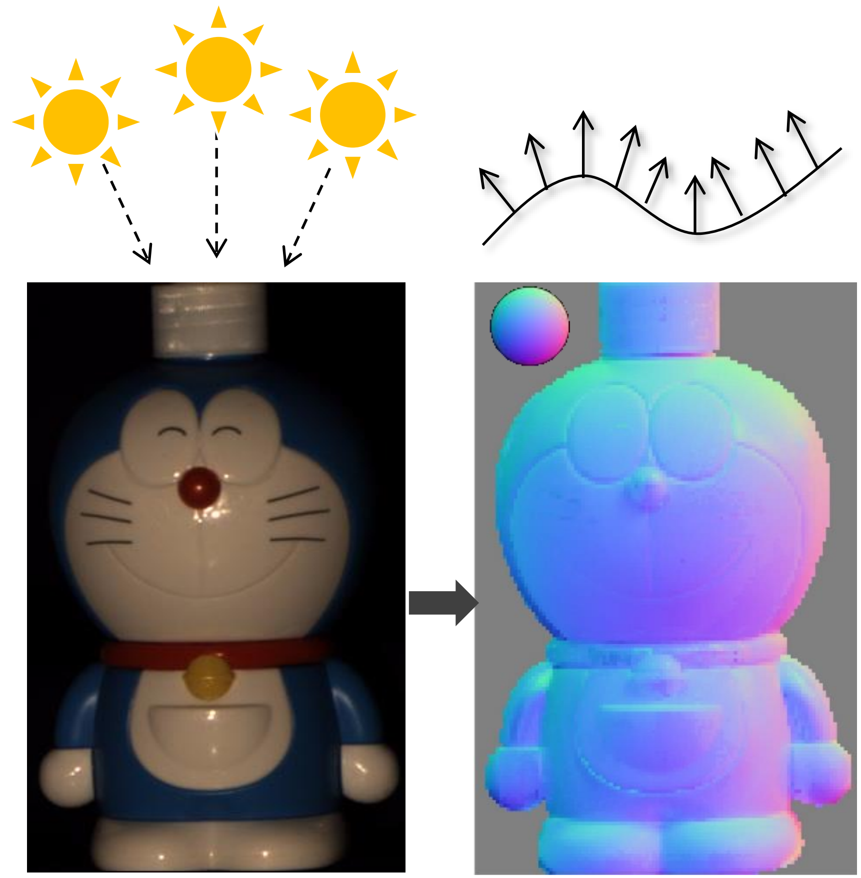
1-10 (2018) **Machine Learning by Prof. Satoshi Ikehata**

11-15 (2019) Signal Processing by Prof. Kazuya Kodama

<u>Grading will be based on the final report.</u>

About Me

- Satoshi Ikehata, Ph.D (sikehata@nii.ac.jp)
- Research Field: 3D Computer Vision
    - 3D Indoor modeling
    - Photometric Stereo



What is “Media”?

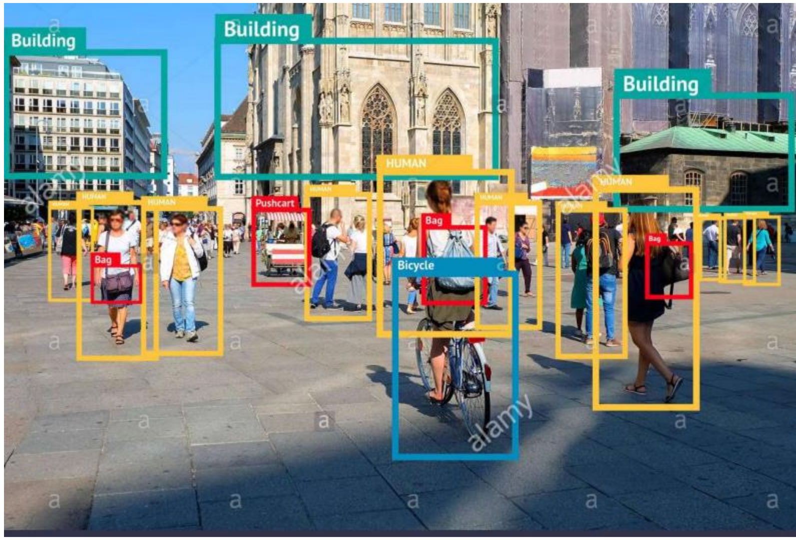
- Image
- Video
- Text
- Audio

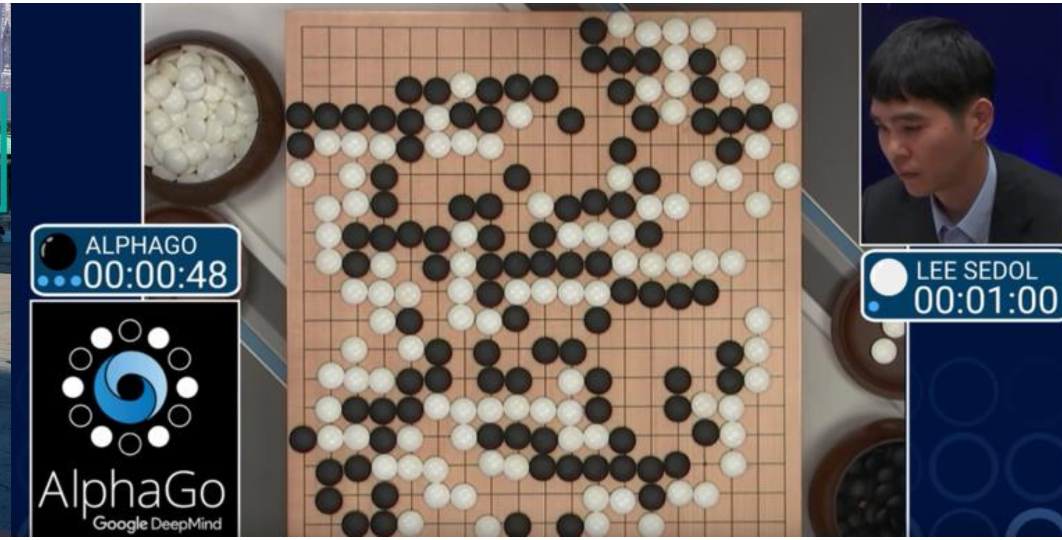
Signal (Continuous) Information

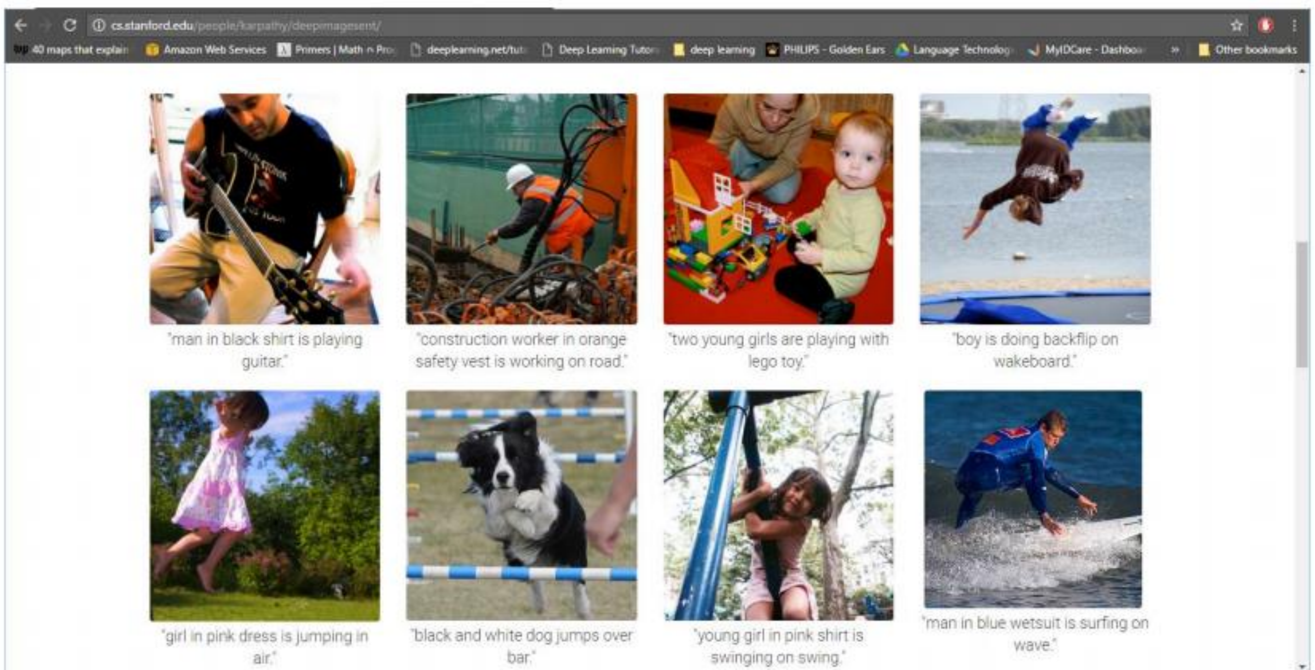
Digital (Discrete) Information

Also referred to as “Multimedia” in computer science

Artificial Intelligence (AI) is a Magical Word?



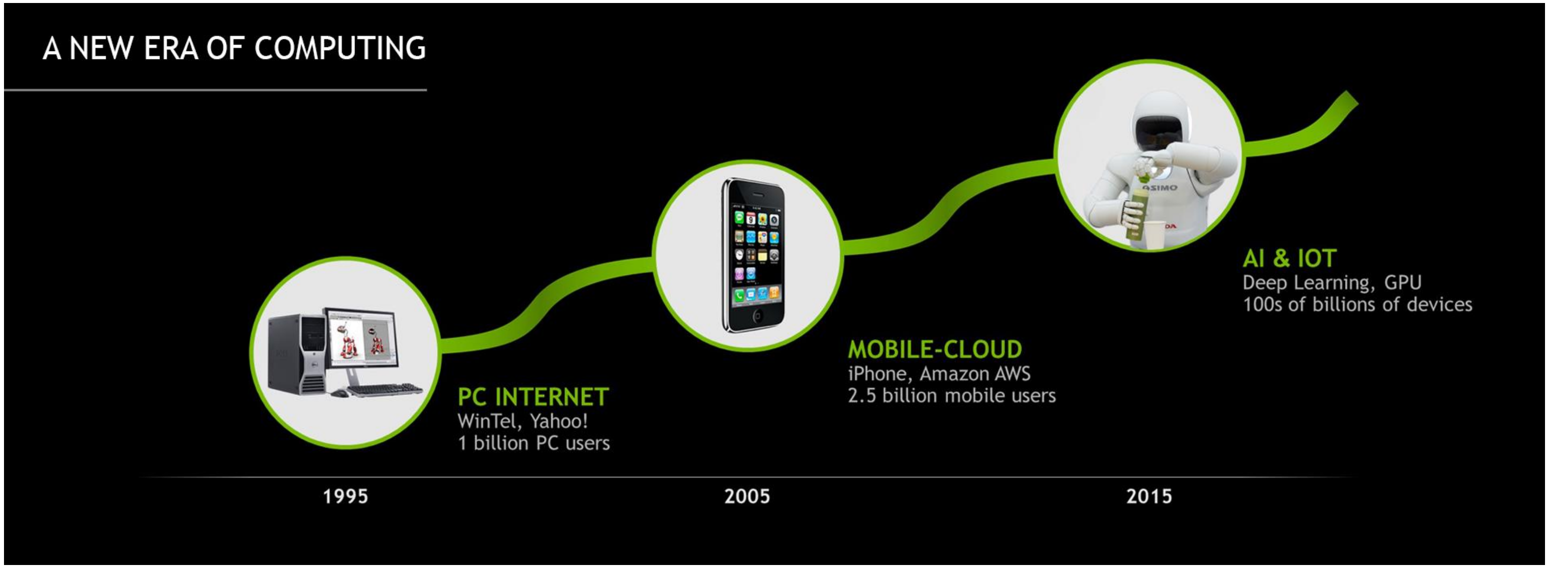




The era of AI

- Deep learning is already the “Standard” in academia and industry
- Everyone can enjoy the deep learning just by simply using public software
- However, most people (probably) do not know what is going behind it

- If you want to be an AI researcher/engineer, the basic knowledge about the machine learning is requirement



https://blogs.nvidia.co.jp/2016/10/27/intelligent-industrial-revolution/

メディア処理基礎 Fundamentals of Media Processing

科目コード(Course Number) 20DIFd02
複合科学研究科 School of Multidisciplinary Sciences 情報学専攻 Department of Informatics 情報メディア科学 Multimedia Information Science
学年(Recommended Grade) 1年 2年 3年 4年 5年
2単位(credit) 後学期 2nd semester

児玉 和也 (KODAMA Kazuya) 佐藤 真一 (SATO Shinichi) 孟 洋(MO Hiroshi) 池畑 諭(IKEHATA Satoshi)

〔授業の概要 Outline〕
メディア処理の全般に関わる基礎技術について、パターン認識理論および信号処理理論を中心に概説をおこなう。これらの理論は、情報メディアを解析し、特徴を抽出したり、望ましい形に変換するためには欠かせない技術である。必要に応じ演習の時間を設け、映像情報等を実際に処理してみることで、より理解を深める。

This course explains the overview of the basic technologies related to whole aspect of media processing especially pattern recognition theory and signal processing theory. These technologies are indispensable for media analysis, feature extraction, media conversion, and so on. Project works such as video information processing will be assigned upon necessity to deepen the understanding.

〔教育目標・目的 Aim〕
マルチメディアに関わるパターン認識並びに信号処理の基本技術の習得を目的とする。

Understanding the basic technologies of pattern recognition and signal processing for multimedia.

〔成績評価 Grading criteria〕
（1）評価方法（提出期限等を含む）
複数のレポート。最終レポートの締切は2月中旬を予定。
（2）割合
レポート（100%）
（3）評価基準
当該領域について正しくかつ十分深く理解しており、加えて自分の考えを適切に述べることができる。

(1) Evaluation
Several reports will be imposed. Final report will be due on mid February.
(2) Ratio
Report (100%)
(3) Criteria
Correct and sufficient understanding of the field and the ability to describe own thought.

〔授業計画 Lecture plan〕
1. 概要説明
2. ベイズの定理／確率分布／正規分布
3. ランダムベクトル／線形代数基礎／直交展開／主成分分析
4. パラメトリック分布の推定／ノンパラメトリック分布の推定
5. 線形識別／クラスタリング
6. 主成分分析による顔の検出
7. 主成分分析による顔の検出(演習)
8. 主成分分析による顔の認識
9. 主成分分析による顔の認識(演習)
10. 信号変換
11. 信号変換に基づくフィルタ処理
12. 多次元信号処理
13. 適応的信号処理
14. 総合演習
15. 総合討論

1. Introduction
2. Bayes decision theory, probability distribution, normal distribution
3. Random vector, linear algebra, orthogonal expansions, principal component analysis(PCA)
4. Parametric density estimation, nonparametric density estimation
5. Linear discriminant analysis, clustering
6. Face detection by PCA
7. Face detection by PCA (project)
8. Face recognition by PCA
9. Face recognition by PCA (project)
10. Signal transformation
11. Filtering technologies based on signal transformation
12. Multi-dimensional signal processing
13. Adaptive signal processing
14. Wrap-up project
15. Discussion

〔実施場所 Location〕
国立情報学研究所(NII)：講義室1（12階1212号室）
NII: Lecture Room 1(12F, 1212)

〔使用言語 Language〕
日本語または英語

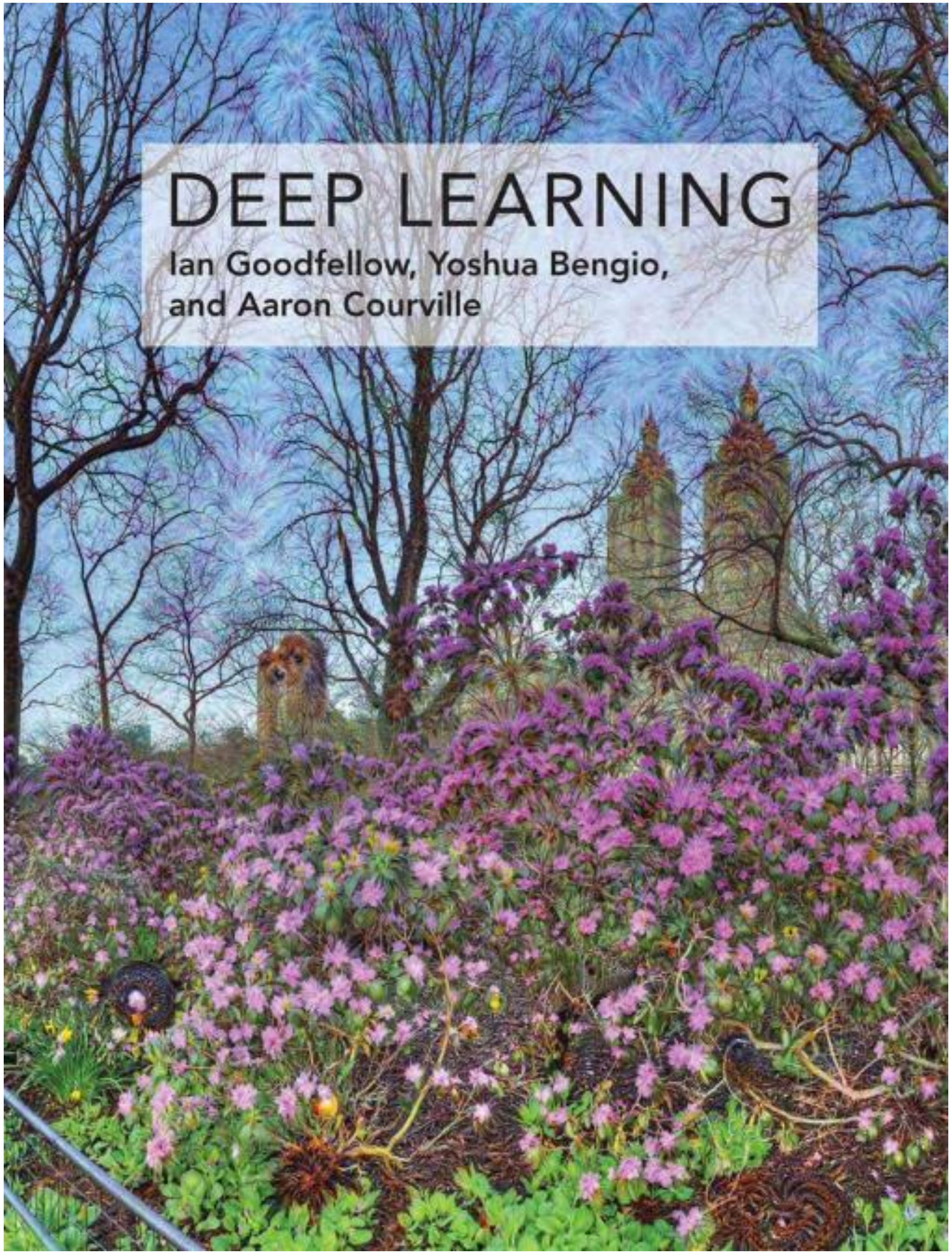
〔教科書・参考図書 Textbooks and references〕
必要に応じてプリントを配布する。
Handouts will be provided if necessary.
講義内で適宜、紹介する。
Textbooks will be introduced upon necessity.

〔関連URL Related URL〕
URL:

〔上記URLの説明 Explanatory Note on above URL〕

〔備考・キーワード Others/Keyword〕
とくになし
N/A
線形代数の基礎知識を有すること。
Basic knowledge of linear algebra is required.

Sorry, this shrubs is not accurate…



Chapter 1-9 (out of 20)

An introduction to a broad range of topics in deep learning, covering mathematical and conceptual background, deep learning techniques used in industry, and research perspectives.

- Due to my background, I will mainly talk about “image”
- I will introduce some applications beyond this book

Deep Learning

An MIT Press book in preparation

Ian Goodfellow, Yoshua Bengio and Aaron Courville

Book Exercises External Links

Lectures

We plan to offer lecture slides accompanying all chapters of this book. We currently offer slides for only some chapters. If you are a course instructor and have your own lecture slides that are relevant, feel free to contact us if you would like to have your slides linked or mirrored from this site.

1. Introduction
   - Presentation of Chapter 1, based on figures from the book [.key] [.pdf]
   - **Video** of lecture by Ian and discussion of Chapter 1 at a reading group in San Francisco organized by Alena Kruchkova
2. Linear Algebra [.key][.pdf]
3. Probability and Information Theory [.key][.pdf]
4. Numerical Computation [.key] [.pdf] [youtube]
5. Machine Learning Basics [.key] [.pdf]
6. Deep Feedforward Networks [.key] [.pdf]
   - Video (.flv) of a presentation by Ian and a group discussion at a reading group at Google organized by Chintan Kaur.
7. Regularization for Deep Learning [.pdf] [.key]
8. Optimization for Training Deep Models
   - **Gradient Descent and Structure of Neural Network Cost Functions** [.key] [.pdf] These slides describe how gradient descent behaves on different kinds of cost function surfaces. Intuition for the structure of the cost function can be built by examining a second-order Taylor series approximation of the cost function. This quadratic function can give rise to issues such as poor conditioning and saddle points. Visualization of neural network cost functions shows how these and some other geometric features of neural

Free copy of the book and useful materials are available at

https://www.deeplearningbook.org/lecture_slides.html

10/16 (Today) Introduction Chap. 1

Basic of Machine Learning (Maybe for beginners)

10/23 Basic mathematics (1) (Linear algebra, probability, numerical computation) Chap. 2,3,4

10/30 Basic mathematics (2) (Linear algebra, probability, numerical computation) Chap. 2,3,4

11/6 Machine Learning Basics (1) Chap. 5

11/13 Machine Learning Basics (2) Chap. 5

Basic of Deep Learning

11/20 Deep Feedforward Networks Chap. 6

11/27 Regularization and Deep Learning Chap. 7

12/4 Optimization for Training Deep Models Chap. 8

CNN and its Application

12/11 Convolutional Neural Networks and Its Application (1) Chap. 9 and more

12/18 Convolutional Neural Networks and Its Application (2) Chap. 9 and more

Class material is available at

https://satoshi-ikehata.github.io/mediaprocessing.html

Fundamentals of Media Processing (Deep Learning Part)

Fall 2018, 13:00 to 14:30
Instructor: Satoshi Ikehata

Textbook

"Deep Learning" by Ian Goodfellow. The book is available for free online or available for purchase.

Syllabus

Class Date	Topic	Slides
Tue, Oct. 16	Introduction	pdf, pptx
	Basic of Machine Learning	
Tue, Oct. 23	Basic mathematics (1) (Linear algebra, probability, numerical computation	pdf
Tue, Oct. 30	Basic mathematics (2) (Linear algebra, probability, numerical computation	pdf
Tue, Nov. 6	Machine Learning Basics (1)	pdf
Tue, Nov. 13	Machine Learning Basics (2)	pdf
	Basic of Deep Learning	
Tue, Nov. 20	Deep Feedforward Networks	pdf
Tue, Nov. 27	Regularization and Deep Learning	pdf
Tue, Dec. 4	Optimization for Training Deep Models	pdf
	CNN and its Application	
Tue, Dec. 11	Convolutional Neural Networks and Its Application (1)	pdf
Tue, Dec. 18	Convolutional Neural Networks and Its Application (2)	pdf

Comments, questions to >Satoshi Ikeahta (sikehata@nii.ac.jp).

10/23

Basic mathematics (1) (Linear algebra, probability, numerical computation)

- Basic of Scalars, Vectors, Matrices and Tensors
- Norms (e.g., l_2-norm, l_∞-norm)
- Eigen Decomposition
- Singular Value Decomposition
- Solving a Homogeneous Equation (Ax=0)
- Probability Distributions
- Marginal Probability, Conditional Probability
- Expectation, Variance and Covariance

=

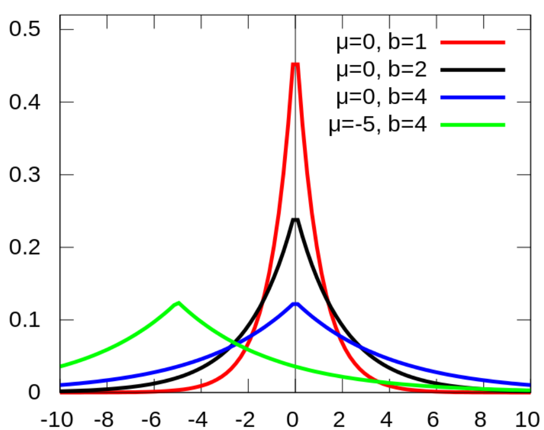
41	56	62	150	23	26	251	50	39
197	186	219	111	209	45	14	237	145
57	109	122	200	147	29	109	4	29
113	179	104	87	222	63	138	125	15
156	97	174	199	200	112	244	64	217
230	166	19	148	119	253	254	246	243
144	14	129	71	12	90	161	175	198
207	68	69	94	248	162	43	26	198
168	117	171	250	62	18	79	79	161
151	112	166	212	95	239	184	52	67

Image is a matrix

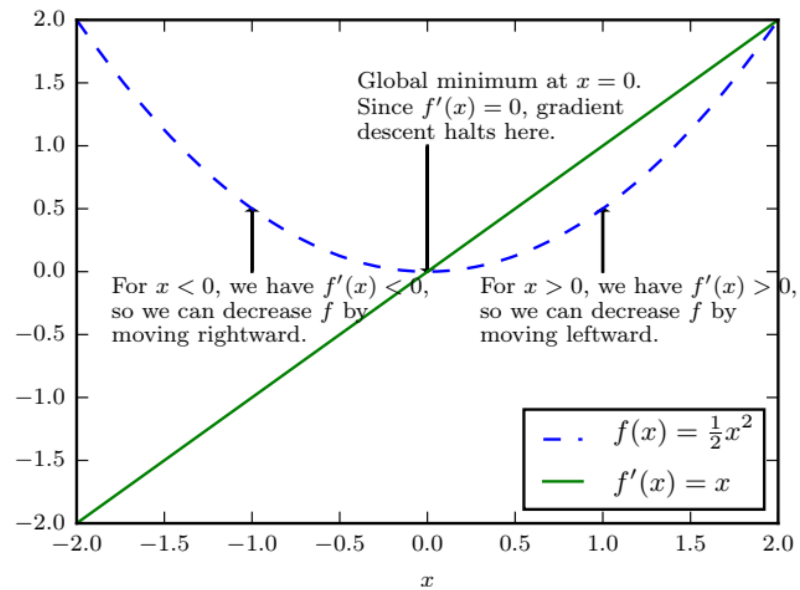
10/30

Basic mathematics (2) (Linear algebra, probability, numerical computation)

- Probabilistic Distribution (e.g., Bernoulli, Gaussian, Laplace)
- Mixture of Distribution (e.g., Gaussian Mixture Model)
- Bayes' rule
- Information Theory (Shanon entropy, KL divergence, cross entropy)
- Structured Probabilistic Models
- Numerical Computation (Overflow, Underflow)
- Gradient-based Optimization
- Jacobian and Hessian Matrix
- Constrained Optimization
- Linear least-squares



Laplace Distribution



Gradient descent

11/06

Machine Learning Basics (1)

- Machine Learning Tasks (E.g., Classification, Regression, translation…)
- Classification of Machine Learning Algorithms (supervised, semisupervised, unsupervised)
- Linear Regression ($\boldsymbol{y} = \boldsymbol{\omega}^T \boldsymbol{x}$)
- Capacity, Overfitting and Underfitting
- The No Free Lunch Theorem
- Regularization, Cross Validation (Training and Validation)
- Estimators, Bias and Variance
- Maximum Likelihood Estimation (MLE)
- Bayesian Statistics (↔ frequent statistics)
- Maximum A Posteriori (MAP) Estimation

Bayesian versus Frequentism

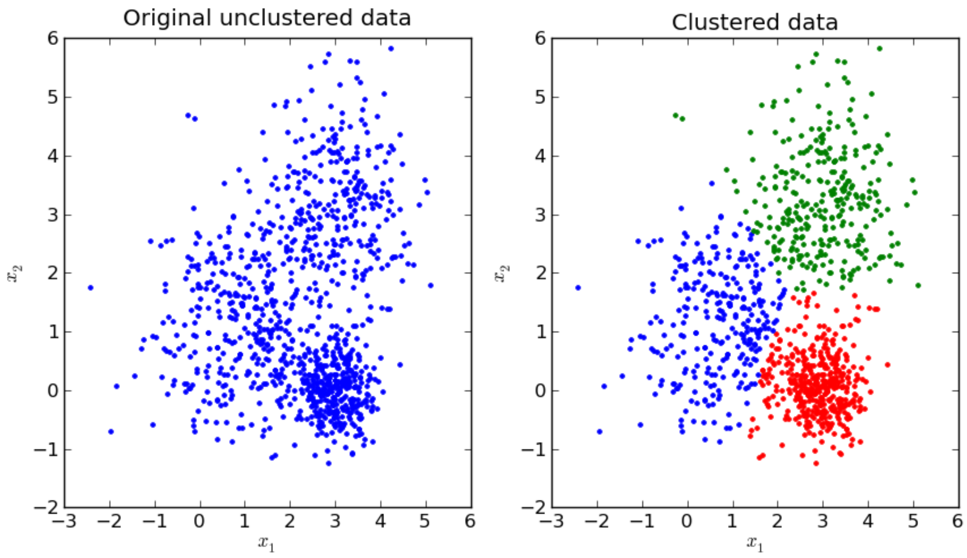
	Bayesian	Frequentist
Basis of method	Bayes Theorem → Posterior probability distribution	Uses pdf for data, for fixed parameters
Meaning of probability	Degree of belief	Frequentist definition
Prob of parameters?	Yes	Anathema
Needs prior?	Yes	No
Choice of interval?	Yes	Yes (except F+C)
Data considered	Only data you have	….+ other possible data
Likelihood principle?	Yes	No

52

11/13

Machine Learning Basics (2)

- Supervised Learning (Support Vector Machine, Decision Tree)
- Unsupervised Learning (Principle Component Analysis, k-means)
- Stochastic Gradient Descent (SGD) Algorithm
- Curse of Dimensionality
- Local Constancy Smoothness Regularization
- Manifold Learning



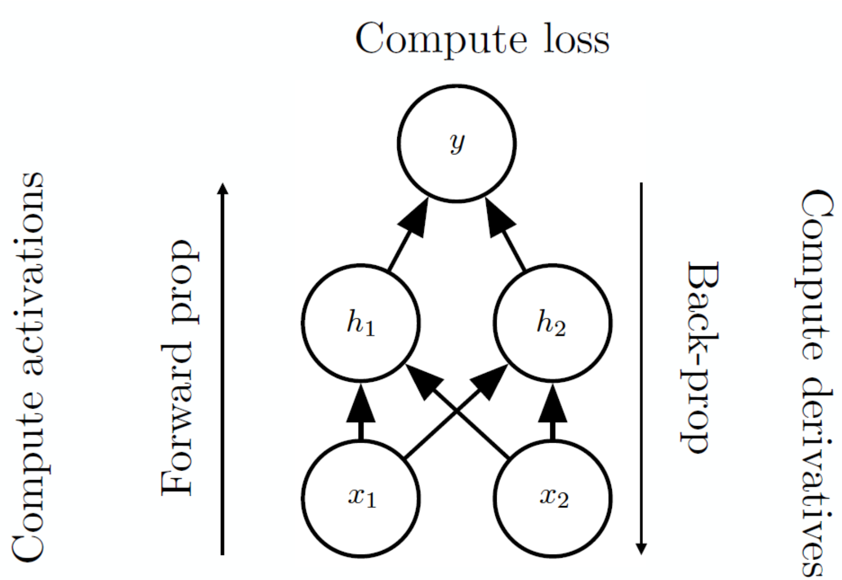
Example of K-means clustering

Deep Feedforward Networks (Feedforward Neural Networks, Multilayer Perceptron)

- Basic of Neural Networks
- Gradient-based Learning (non-linearity of Neural Network)
- Loss Function (Cost Function)
- Output Layer (Linear Unit, Sigmoid Unit, Softmax Unit)
- Activation Layer (Rectified Linear Units:ReLu, Logistic Sygmoid, Tangent)
- University Approximation Theorem (A simple neural network can represent complex function)
- Back Propagation

Output Type	Output Distribution	Output Layer	Cost Function
Binary	Bernoulli	Sigmoid	Binary cross-entropy
Discrete	Multinoulli	Softmax	Discrete cross-entropy
Continuous	Gaussian	Linear	Gaussian cross-entropy (MSE)
Continuous	Mixture of Gaussian	Mixture Density	Cross-entropy
Continuous	Arbitrary	See part III: GAN, VAE, FVBN	Various

Output Type

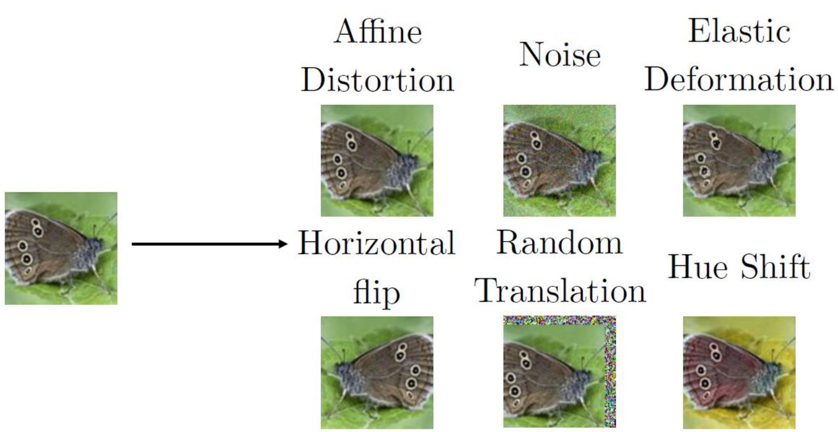


Back Propagation

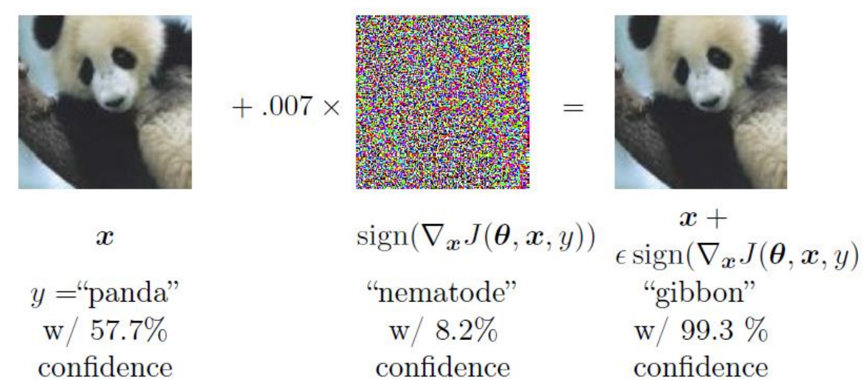
11/27

Regularization for Deep Learning (Making an algorithm perform on new input)

- Parameter Norm Penalties (weight-decay(ℓ_2), sparse(ℓ_1))
- Dataset Augmentation (Rotation, Translation, Injection of noise)
- Semi-Supervised Learning (labeled + unlabeled training data)
- Multitask Learning (Task specific parameters + Generic parameters)
- Early Stopping
- Parameter Tying/Sharing
- Sparse Representations, Bagging, Dropout
- Adversarial Training



Data augmentation

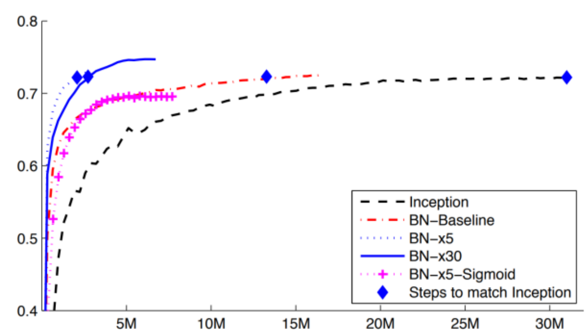


Adversarial example

12/4

Optimization for Training Deep Models

- Empirical Risk Minimization (Optimization on the training dataset)
- Surrogate Loss Function (e.g., learning 0 or 1 by loglikelihood function)
- Stochastic (per one sample), Batch (per all data) and Minibatch (per subset)
- Theoretical Limitation in Optimization: Local Minima, Plateaus, Saddle Points
- Stochastic Gradient Descent (SGD) for learning
- Momentum (Update parameters using x_t and x_{t-1})
- Parameter Initialization Strategies
- Algorithms with Adaptive Learning Rates (AdaGrad, RMSProp, Adam)
- Approximate Second-Order Methods (Newton's, Conjugate Gradient, BFGS)
- Batch Normalization
- Coordinate Decent
- Supervised Pretraining (e.g., Using ImageNet weights for other tasks)
- Continuation Methods and Curriculum Learning



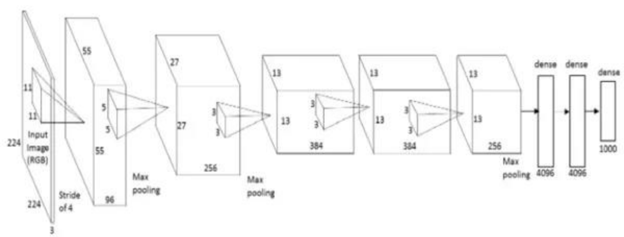
"Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Ioffe and Szegedy 2015

12/11, 18

Convolutional Neural Networks (1)

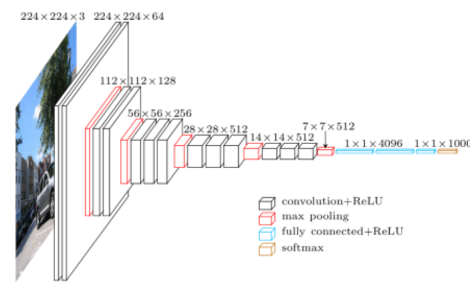
- History and Basics of CNN
- Pooling (Max pooling, Average pooling)
- Stride
- Unshared Convolution, Tiled Convolution
- Network Architectures (LeNet, AlexNet, ResNet, DenseNet…)

TBD (Applications of Deep Learaning in Conputer Vision)



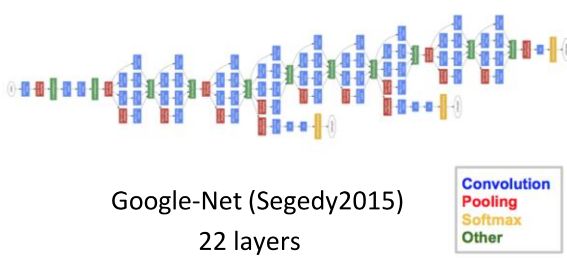
Alex-Net (Krizhevsky2012)

8 layers



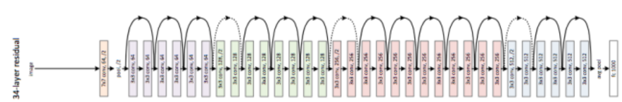
VGG-Net (Simonyan2013)

16 layers



Google-Net (Segedy2015)

22 layers



Res-Net (He2015)

125 layers

Course Website

https://satoshi-ikehata.github.io/mediaprocessing.html

Contact: sikehata@nii.ac.jp