

# これから始めるディープラーニング

---

岡本 一志

電気通信大学 大学院情報理工学研究科

Head of Data Science Research Group

# 自己紹介

---

岡本 一志 (おかもと かずし)

1984年生まれ (32歳) , 高知県出身

## 学歴・職歴

- 2002 - 2006 高知工科大学 工学部
- 2006 - 2008 高知工科大学 大学院工学研究科
- 2008 - 2011 東京工業大学 大学院総合理工学研究科
- 2011 - 2015 千葉大学 アカデミック・リンク・センター 特任助教
- 現在 - 電気通信大学 大学院情報理工学研究科 助教
- 現在 - 千葉大学 アカデミック・リンク・センター 客員研究員

# 岡本研究室について

---

## コンセプト

統計学・計算機科学・機械学習および現場の知識を  
駆使し、ビッグデータから有用な知見を引き出す研究の展開

現在は、機械学習やデータ可視化の方法論をベースに、購買履歴や行動ログ等の分析に関心があり、これらのビッグデータ分析に応用可能な要素技術の開発に取り組んでいます

データ可視化

情報推薦

並列計算

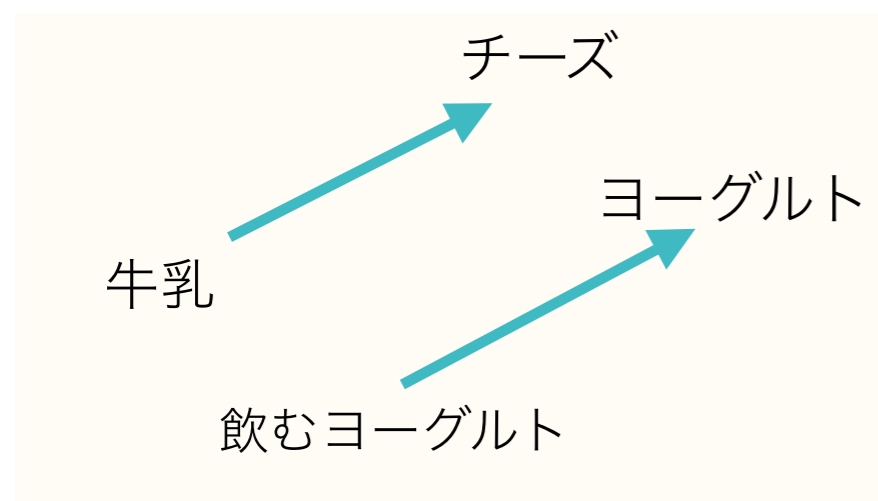
ソフト  
コンピューティング

# レシート集合から商品のベクトル表現の学習

Word2Vecでは文脈から単語ベクトルを作成

1枚のレシート = 文脈, 単語 = 商品

商品の特徴をもつベクトルから簡単に情報推薦したい



+



=



# 図書館内での情報利用行動の自動認識

注目対象の  
把握

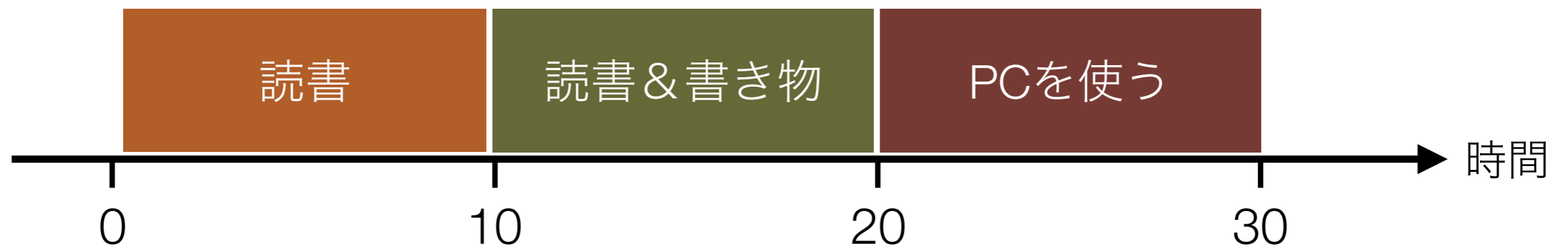


一人称視点カメラ

運動情報の  
取得



3軸加速度センサ



# この講演の目標

---

これから深層学習（deep learning）を使ってみようとお考えの方々を想定し，次の観点で概説します

- 深層学習とはこういったものか？
- 深層学習はどのような時に使えばよいのか？
- 深層学習はどのように使えばよいのか？

理論的側面では，ニューラルネットワークの基本的なところから多層化時の学習方法までの考え方についてお話します

# 深層学習とはどういったものか？

---

# 深層学習 (deep learning) とは？

---

多くの層を持ったニューラルネットワークモデルを用いた  
機械学習の総称

層の多いニューラルネットワークによって，入力データから  
本質的な情報を表現する特徴を学習 (特徴学習)

コンピュータビジョンや音声認識，自然言語処理などの分野で  
の機械学習のタスクにおける成功

- (昔) なぜ，ニューラルネットワークを使うの？
- (今) なぜ，ニューラルネットワークを使わないの？

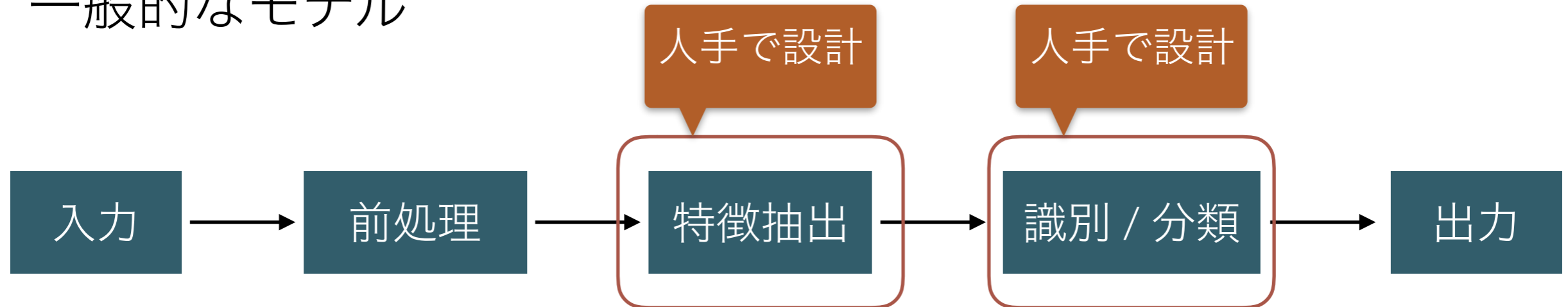


# ニューラルネットワークの歴史的経緯

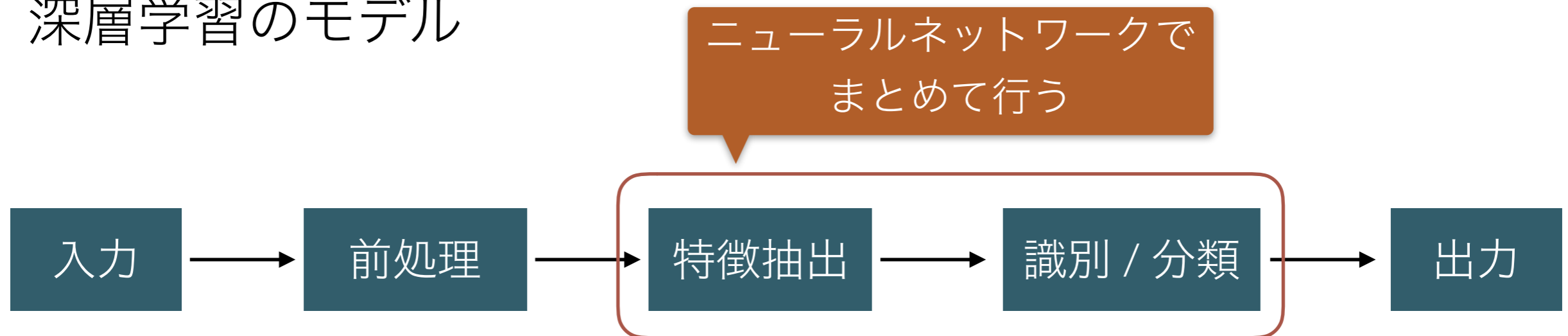
1957	F. Rosenblatt	パーセプトロン (perceptron)
1969	M. Minsky S. Papert	単純パーセプトロンの線形非分離問題の提起
1979	K. Fukushima	ネオコグニトロン (neocognitron)
1986	D. E. Rumelhart G. E. Hinton J. R. Williams	誤差逆伝播法 (backpropagation)
1989	Y. LeCun	畳込みニューラルネットワーク (convolutional neural networks)
2006	G. E. Hinton S. Osindero Y. W. Teh	Deep Belief Nets

# 深層学習の特徴

## 一般的なモデル



## 深層学習のモデル



# なぜ、いま、深層学習なのか？

---

他の手法よりも機械学習タスクでの精度が圧倒的に良い

ブレイクスルー => 特徴抽出の自動化

- 特徴に求められる性質
  - クラス内変動に不変（不変性）
  - わずかな違いにも敏感（識別力）
- 人手で有効な特徴を設計するのは困難

学習アルゴリズムの発達

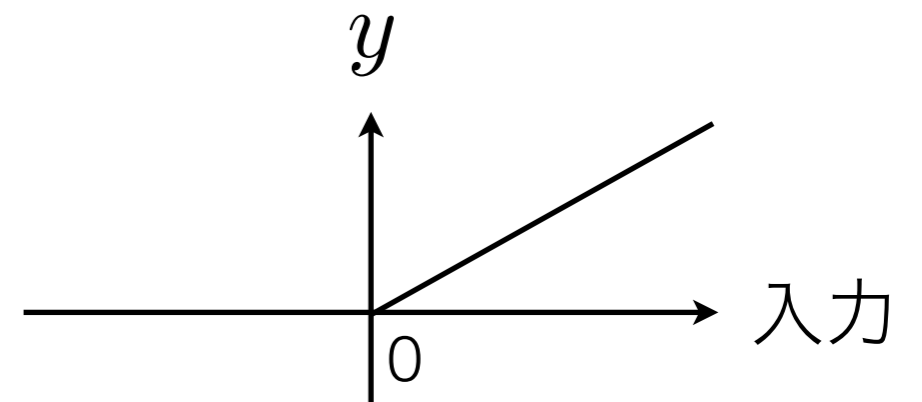
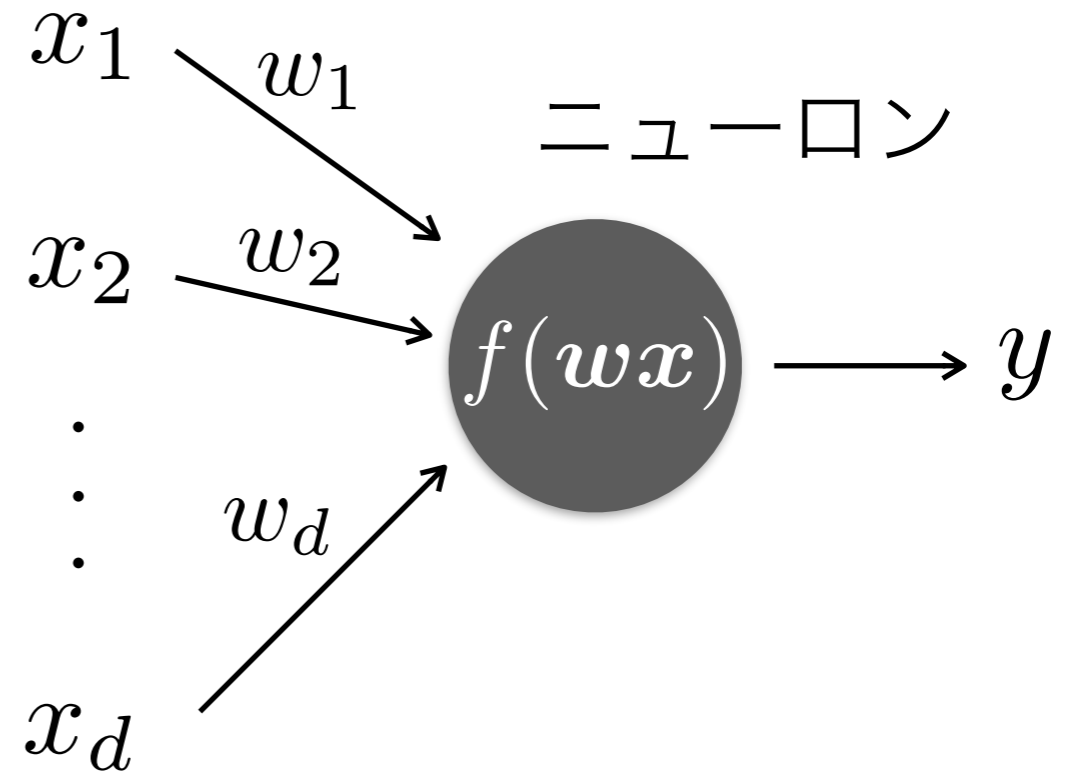
学習に利用可能なデータの増大, 計算環境の高度化

# ニューロンと活性化関数

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \in R^d$$

$$y \in R$$

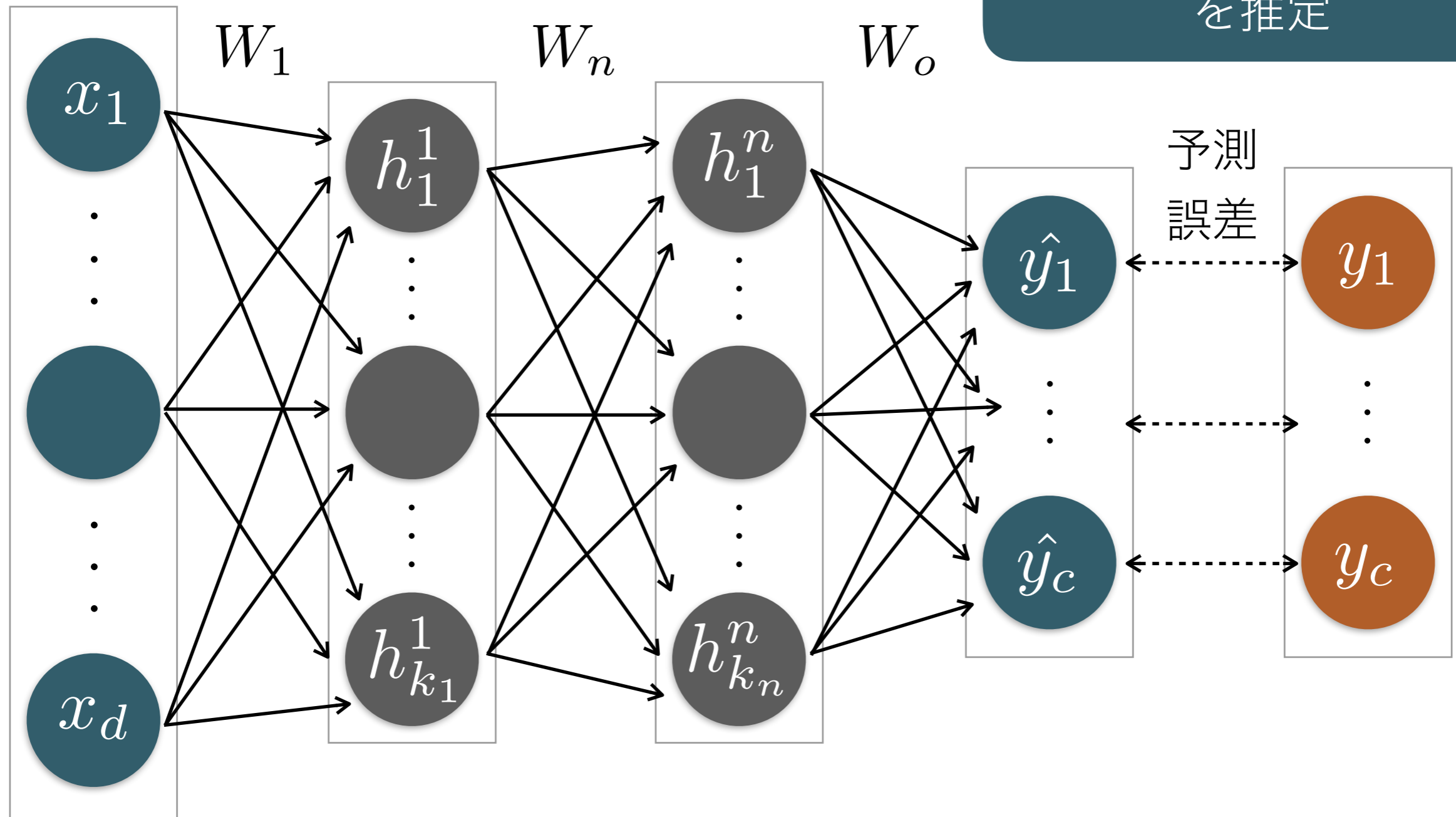
$$f : R^d \longrightarrow R \quad \text{活性化関数}$$



$$y = \max\{0, x\}$$

# ニューラルネットワークモデル

予測誤差から  
 $W^1, W^2, \dots, W^n, W^o$   
を推定



入力層

隠れ層

出力層

教師信号

# 様々なニューラルネットワーク

---

フィードフォワードニューラルネットワーク (feedforward neural network)

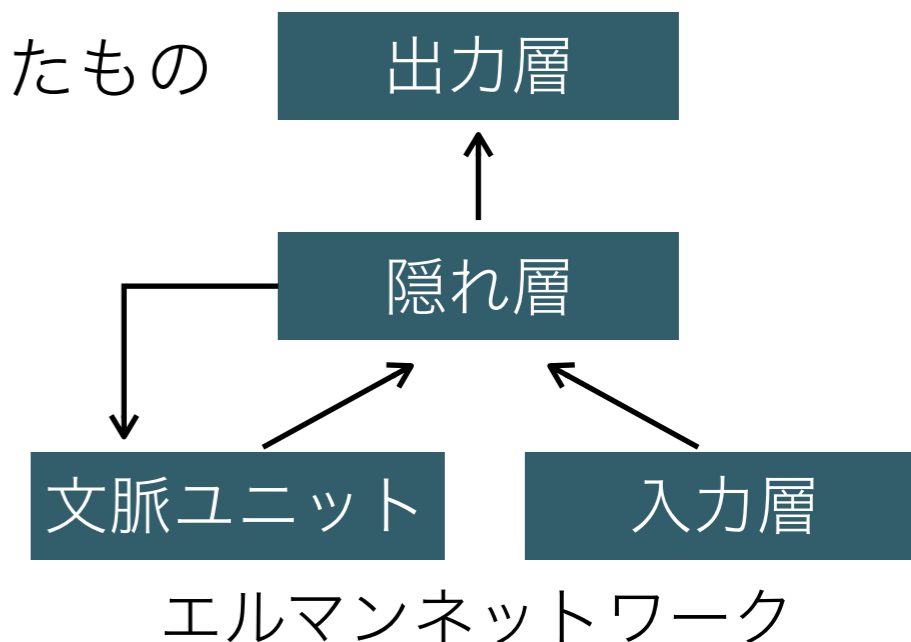
- 入力ベクトルの次元数が決まっている場合に使われる
- 誤差逆伝播法により重みを更新することで学習を行う
- 畳込みニューラルネットワーク (convolutional neural network)

再帰型ニューラルネットワーク (recurrent neural network)

- 時系列情報や文脈依存の情報を扱えるようにしたもの

ボルツマンマシン (Boltzmann machine)

- 確率的に動作するニューラルネットワーク
- 入力データの確率分布を近似



# ニューラルネットワークの学習

---

確率的勾配降下法 (stochastic gradient descent)

- 全データを一度にではなく1つずつデータを使って重みを更新
- いわゆるオンライン学習

誤差逆伝播法 (backpropagation)

多層化時の工夫

- プレトレーニング
- ドロップアウト

# 確率的勾配降下法 (stochastic gradient descent)

---

$$W \leftarrow W - \eta \frac{\partial}{\partial W} E(f(W, \mathbf{x}_i), \mathbf{y}_i)$$

学習データと教師データのペア

$$(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_n, \mathbf{y}_n)$$

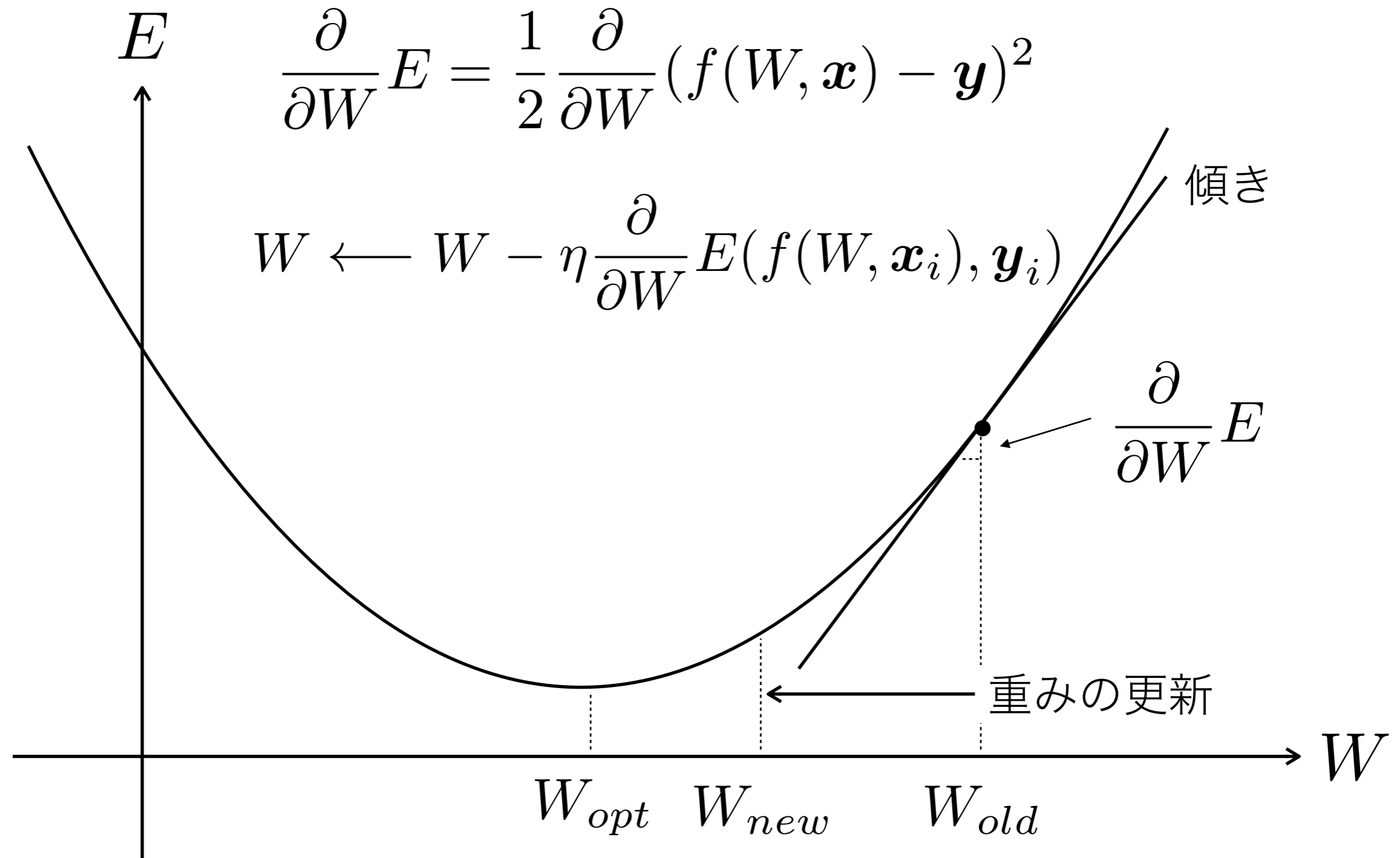
$$\mathbf{x}_i = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \in R^d$$

誤差関数 (2乗誤差)

$$E(f(W, \mathbf{x}), \mathbf{y}) = \frac{1}{2} \|f(W, \mathbf{x}) - \mathbf{y}\|^2$$



# 確率的勾配降下法 (stochastic gradient descent)



# 誤差逆伝播法 (backpropagation)

---

## コスト関数の勾配の計算法

- 勾配の計算方法を指す場合と学習ルールを指す場合がある
- コスト関数の偏微分の計算則に相当

D. E. Rumelhart & G. E. Hinton & J. R. Williamsにより提案

勾配は出力信号と教師信号から求めることができる

層の数が増えるにしたがって勾配が消失 (ゼロに近づく)

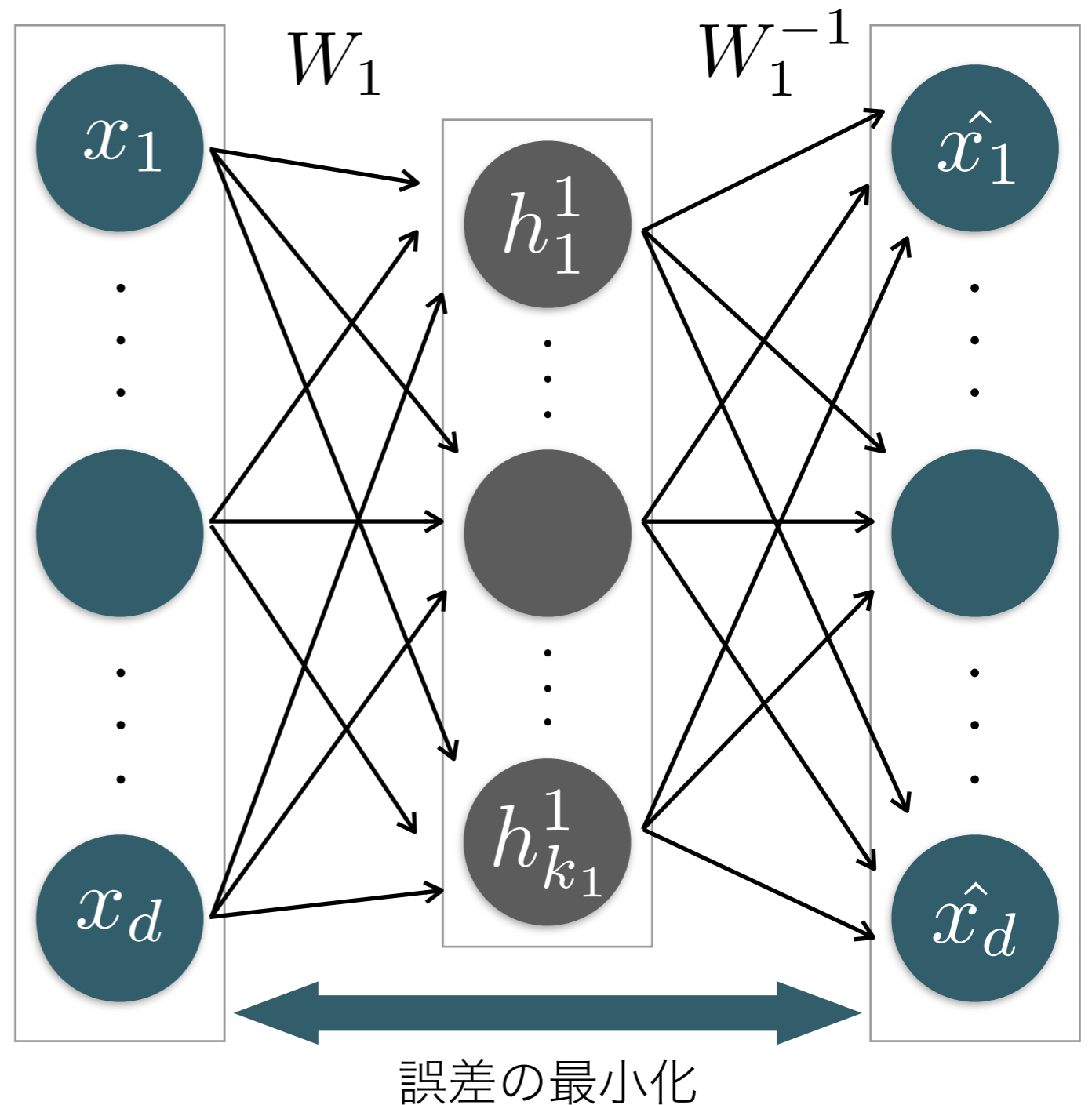
- 活性化関数の変更が近年のブレイクスルーのひとつ

# 多層化時の学習法（プレトレーニング）

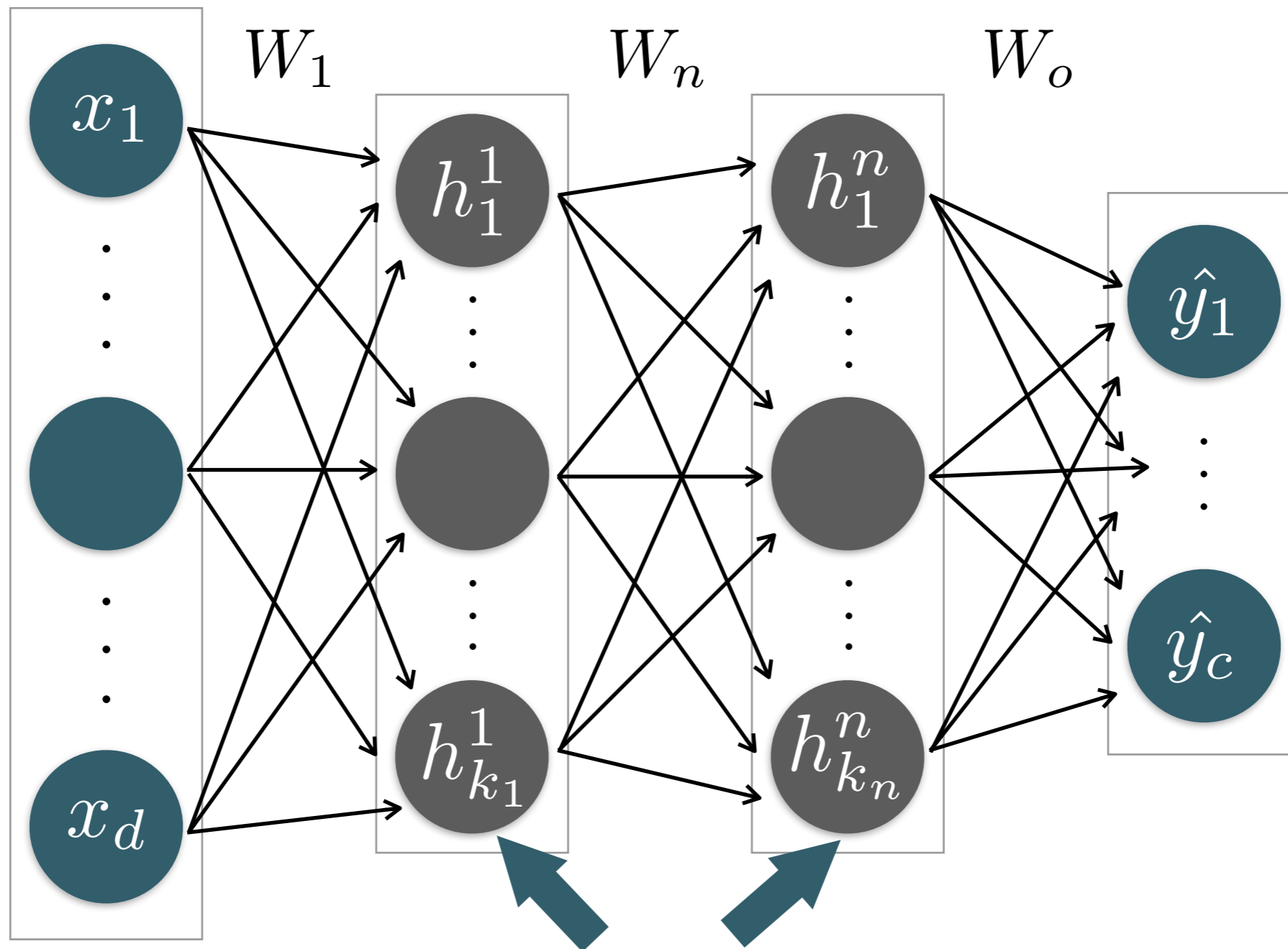
入力（学習データや隠れ層の出力）をできるだけ再現するように $W$ を学習

各層でチューニング

最終的に学習データと教師信号を使ってファインチューニングを行う



# 多層化時の学習法（ドロップアウト）



一定の確率でニューロンを一時的に無かったことにする

深層学習はどのような時に使えばよいのか？

---

# 想定されるタスク

---

画像認識・音声認識・自然言語処理

## 教師あり学習

- ・ 学習用データと正解ラベルのペアが揃っている
- ・ テストデータのラベルを推定する

## 教師なし学習

- ・ 学習用データのみ用意されている
- ・ クラスタリングや異常検知

## 強化学習

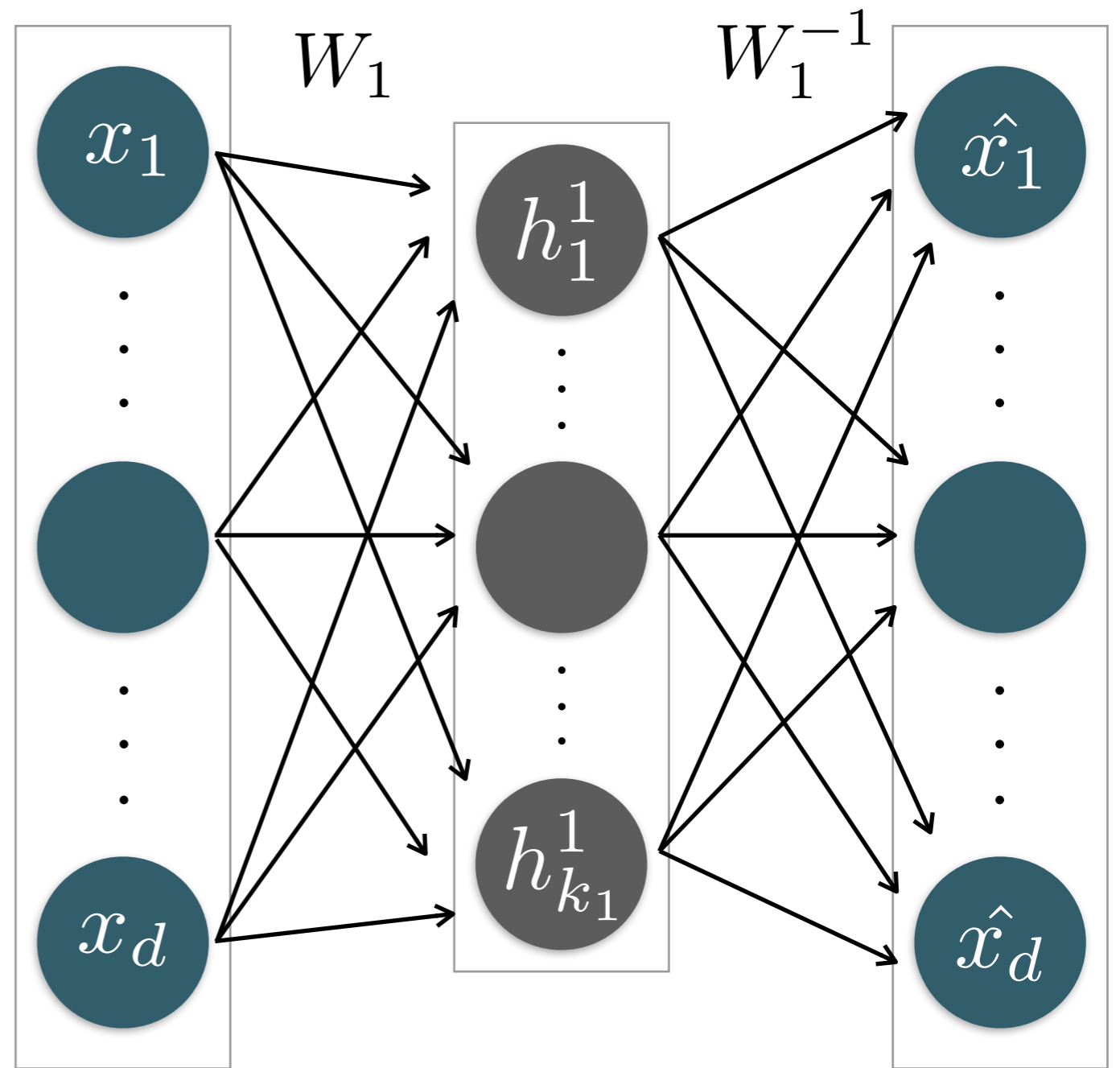
# 特徴表現の学習

入力をできるだけ再現  
する変換の学習

層が増えていくに従って  
得られる特徴が抽象化さ  
れていく傾向

次元削減

機械学習のタスクに有効  
な特徴とは何か？の理解



# 深層学習の画像応用について

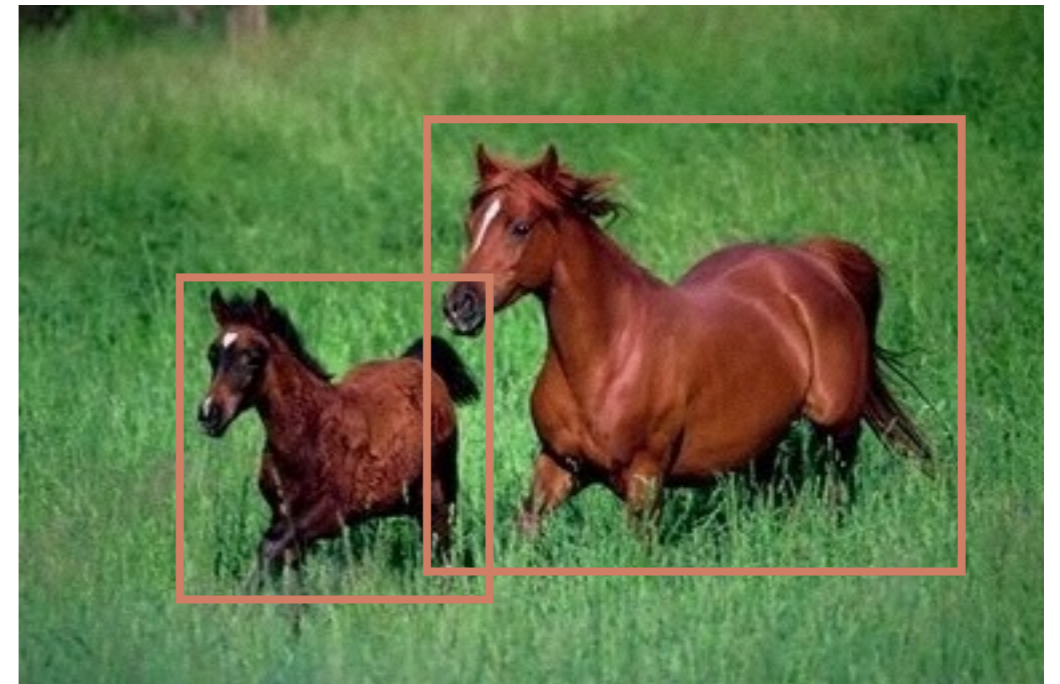
---

## 物体カテゴリ認識

畳込みニューラルネットワークに基づく手法が主流

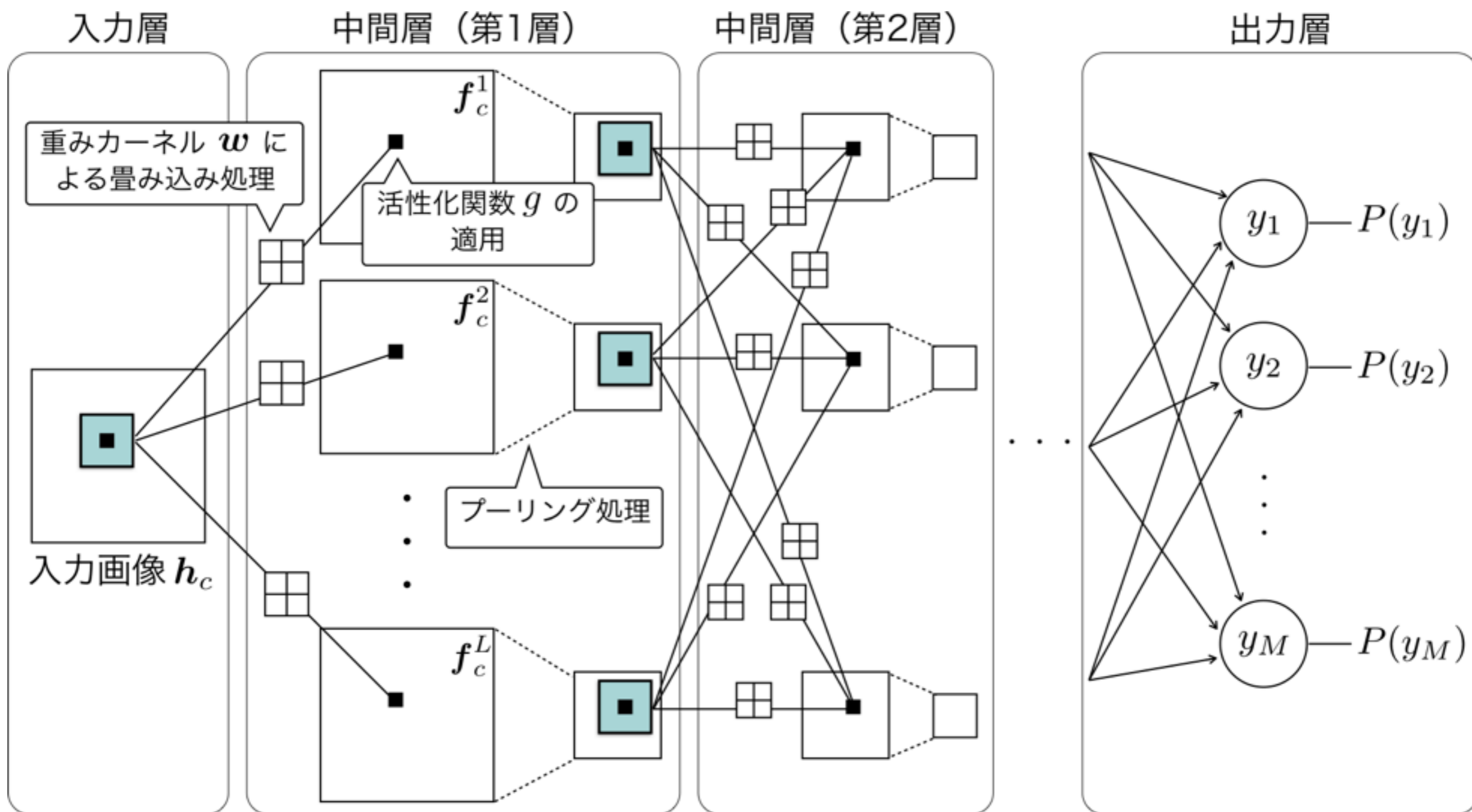
## 苦手としているタスク

- 物体検出（位置決め）
- 動画像認識（時間情報の利用）





# 畳込みニューラルネットワーク (CNN)



# 深層学習はどのように使えばよいのか？

---

# 開発フレームワークの利用

---

1から実装はスキル・ノウハウが必要で、ちょっと試してみようと思うと敷居が高い

現在の開発フレームワークはほとんどはPythonが読み書きできれば利用可能

高速な処理を考えると計算機に関する知識はあった方が良い

- Linux
- GPGPU  
(General-Purpose Computing on Graphics Processing Units)

# 開発フレームワーク (1)

## TensorFlow

- Google発
- 専用のハードウェアまである
- ネットワークの可視化ツールが付随
- ドキュメントが豊富 (英語)
- Python / C++で記述

## Chainer

- Preferred Networks発
- 日本発なので日本語ドキュメントが豊富
- 国内の研究室で広く使われている印象
- Pythonで記述



Tensor Processing Unit board

<https://cloudplatform.googleblog.com/2016/05/Google-supercharges-machine-learning-tasks-with-custom-chip.html>

# 開発フレームワーク (2)

---

## Caffe

- Berkeley Vision and Learning Center 発
- 画像認識でよく使われている
- 学習済みのモデルを配布するフレームワークがある
- C++で実装, Python & Matlabで記述

## Pylearn2

- Theano ベース
- ネットワーク構造を設定ファイルで記述
- Python で記述

# 開発フレームワーク (3)

---

## cuDNN

- Nvidia製
- GPUで動作, CUDAに関して知識が必要
- TensorFlowやChainerなどのバックエンドで利用
- 深層学習用のフレームワークの中ではHW寄り

## Theano

- Université de Montréal発
- 数式そのものを記述
- 微分を解析的に実行できる
- Pythonで記述

# ニューラルネットワークの計算モデル

---

行列演算がほぼ全て

マルチスレッド化・マルチコア化

- GPGPUがあると学習時間が圧倒的に削減される
- 1命令で多数のコアが同時に同じ計算ができると効率が良い

深層学習における計算資源

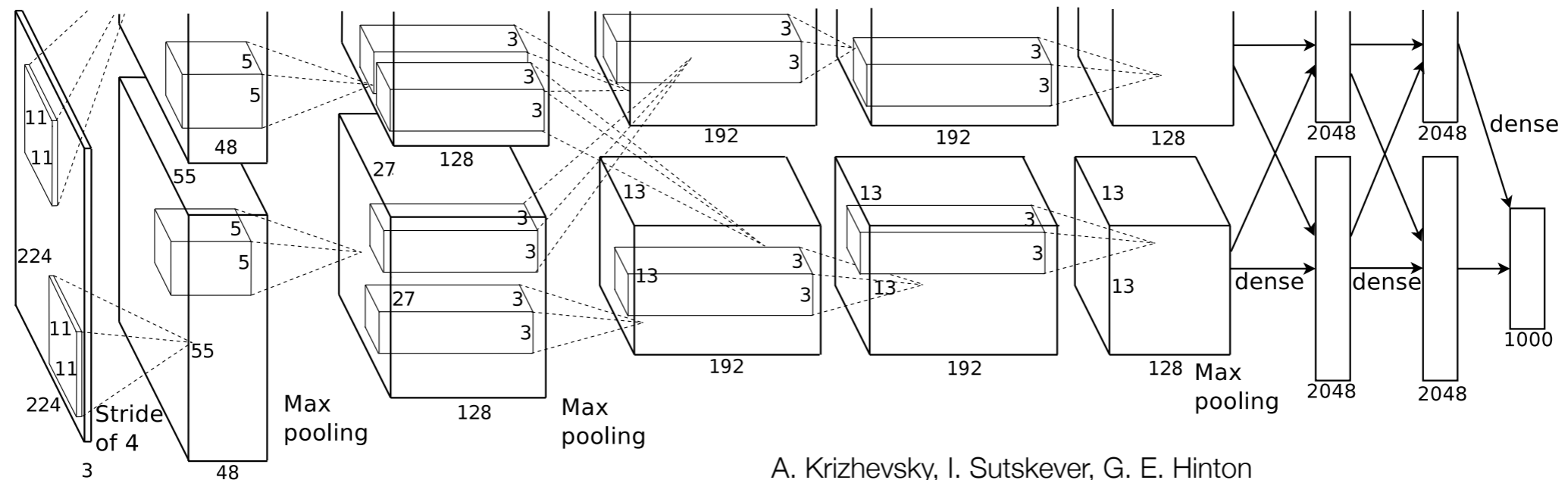
- とにかく計算資源と学習用データを持っている人たちが強い
- 大学の1研究室レベルだととてもじゃないけど追いつけない

# 学習済みのモデルの活用

## AlexNet

- ImageNet Large Scale Visual Recognition Challenge 2012で優勝

最終層を目的に併せて変更. 基礎的な視覚能力は獲得しているので、手元にあるデータのみでファインチューニング



A. Krizhevsky, I. Sutskever, G. E. Hinton  
ImageNet Classification with Deep Convolutional Neural Networks  
Advances in Neural Information Processing Systems 25, pp.1097-1105, 2012



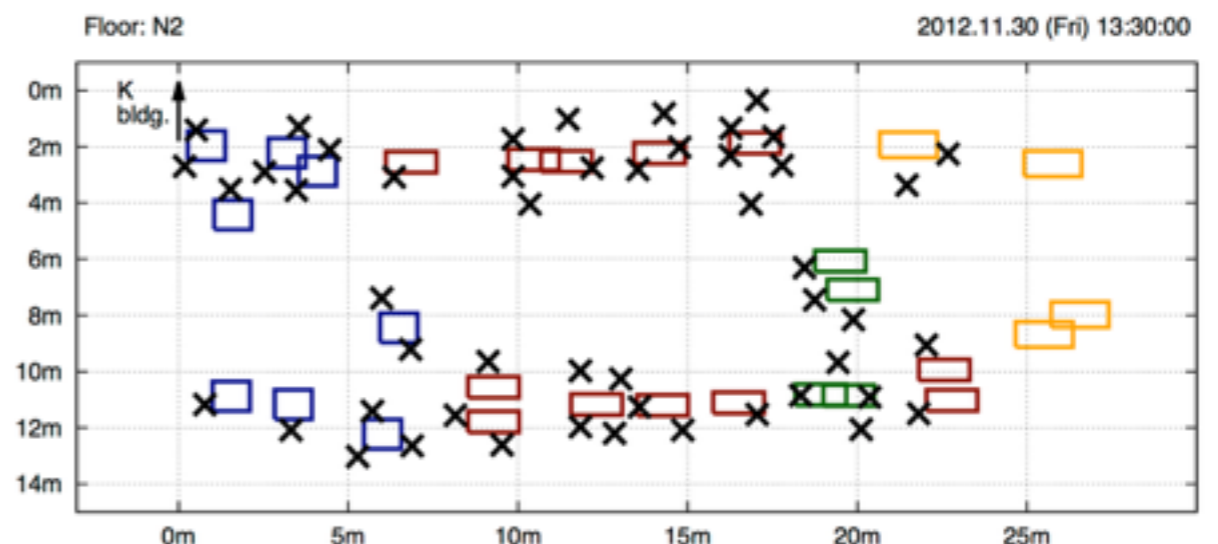
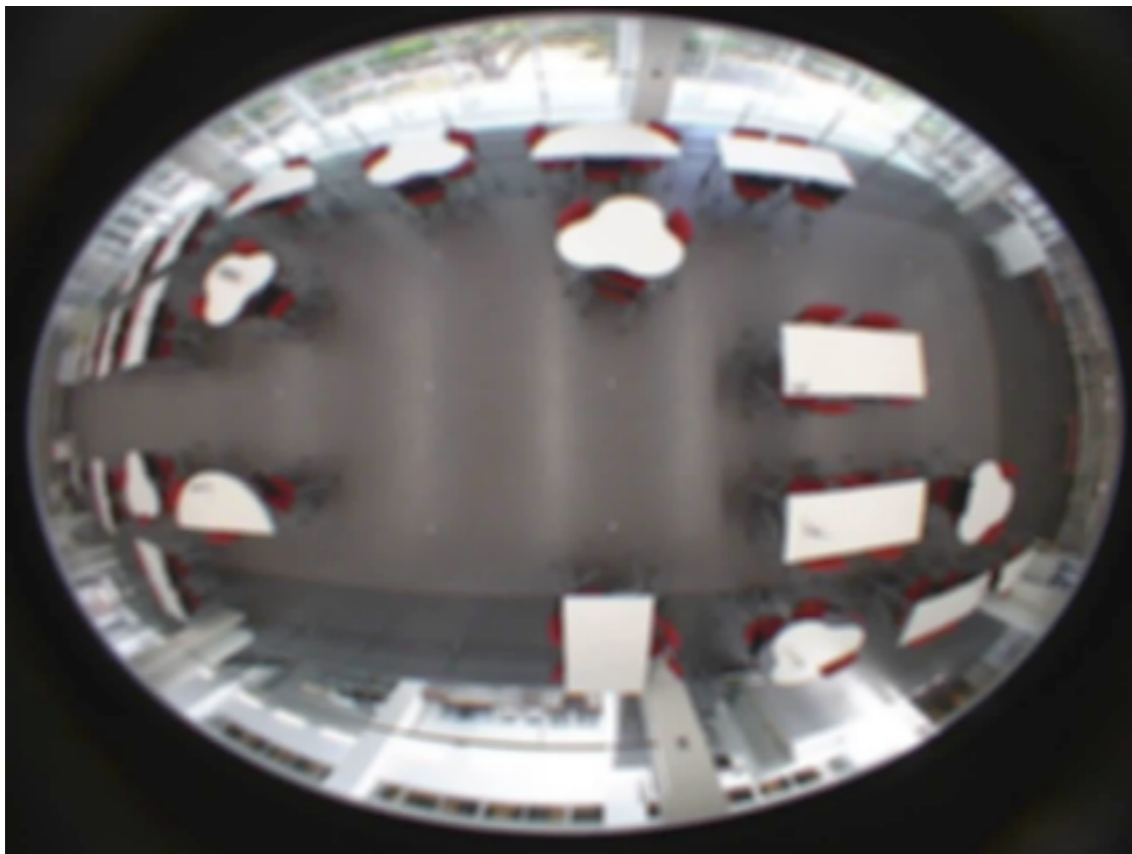
# 深層学習の応用例の紹介

---

# 定点カメラ映像を用いた アクティヴ・ラーニング・スペースの利用状況調査

千葉大学 アカデミック・リンク・センターの情報利用行動定点観測  
プロジェクトのひとつの調査として実施

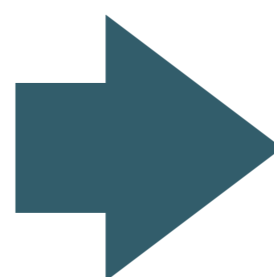
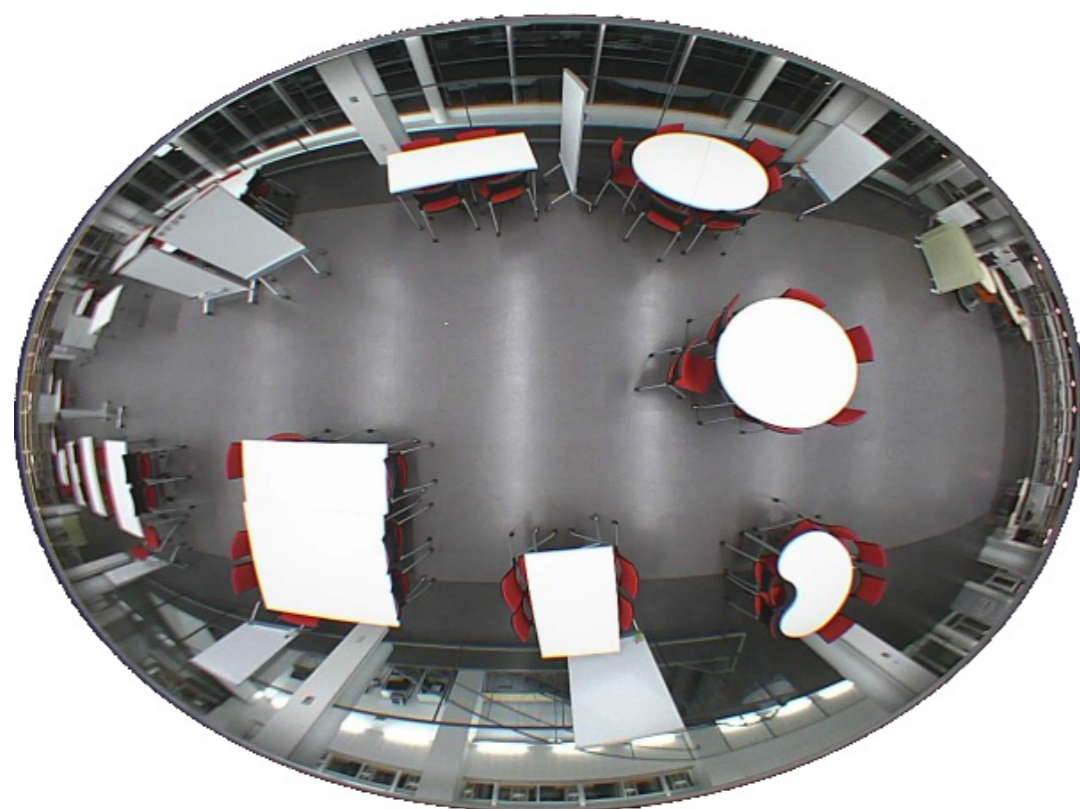
- 学生の情報利用行動および学習・生活空間の利用状況の分析を通じて、  
アカデミック・リンクが学生の情報利用行動・学習行動・学習成果に与える  
影響と効果を継続的・横断的に検証



# 全方位カメラからの人位置検出

360度の視野を持つカメラ

球面座標系への投影や歪み補正が必要（要カメラパラメータ）



切り出し

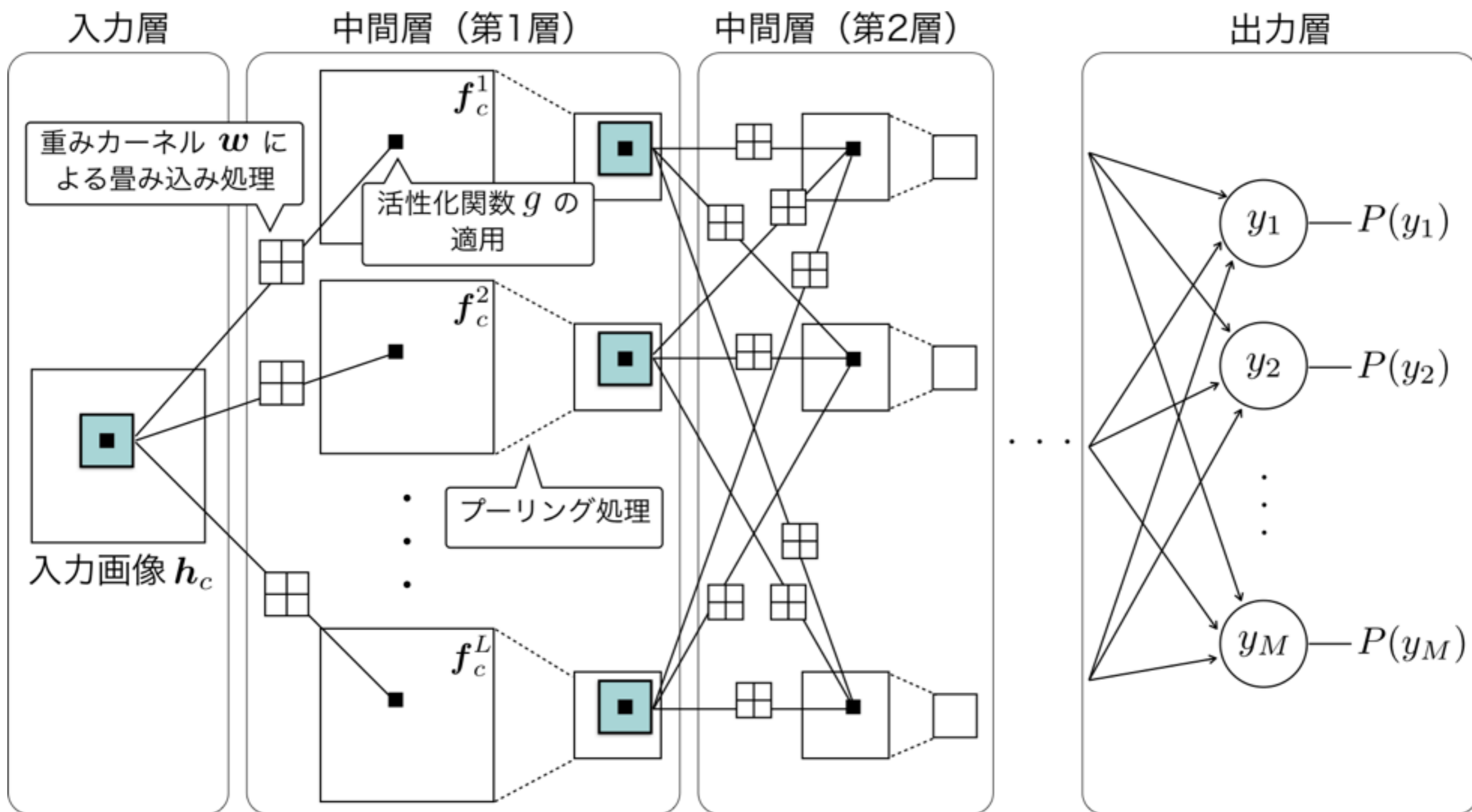
検出ウィンドウ



千葉大学 アカデミック・リンク・センターのプロジェクトの一環でデータ収集

インタラクティブ情報アクセスと可視化マイニング研究会

# 畳込みニューラルネットワーク (CNN)



# 学習用データ

---

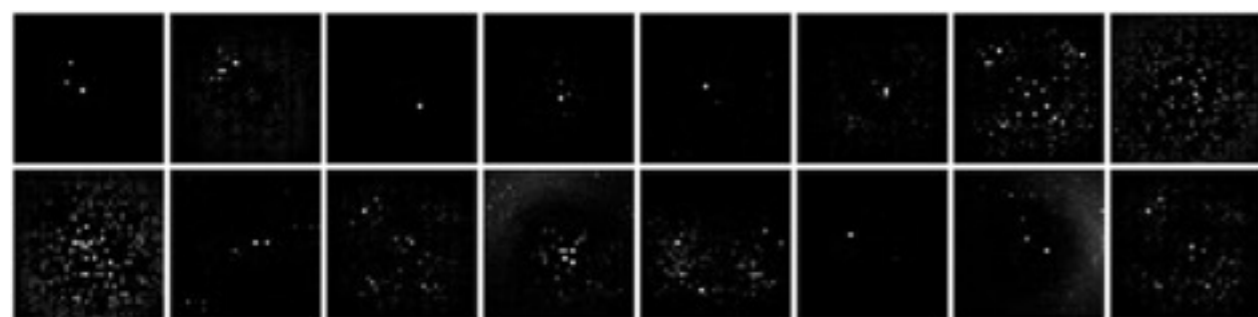
200枚の全方位画像から

- 20,000件のポジティブサンプル
- 20,000件のネガティブサンプル

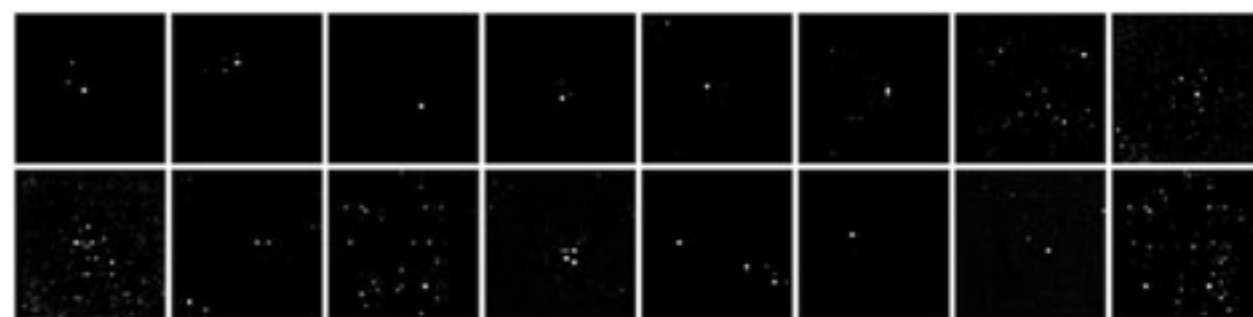
を作成（頭部画像は50枚）

合成画像の作成では，頭部画像にランダムな変形を加えて合成

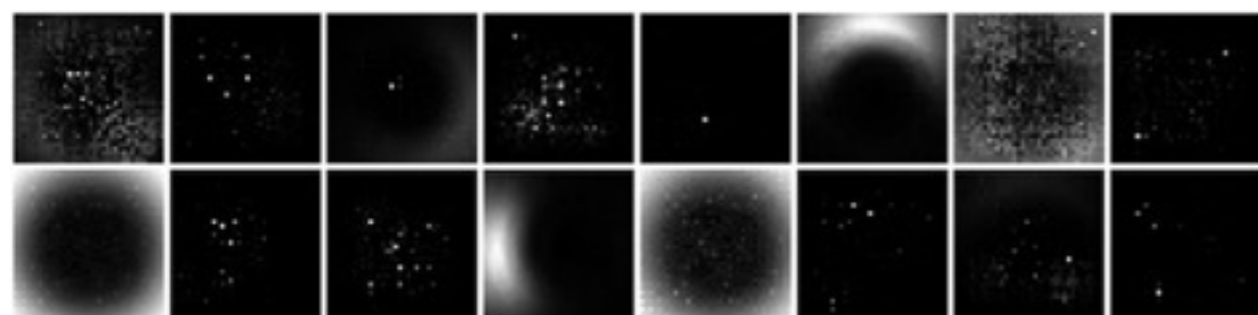
# 中間層第1層における畳込み処理後の特徴マップ



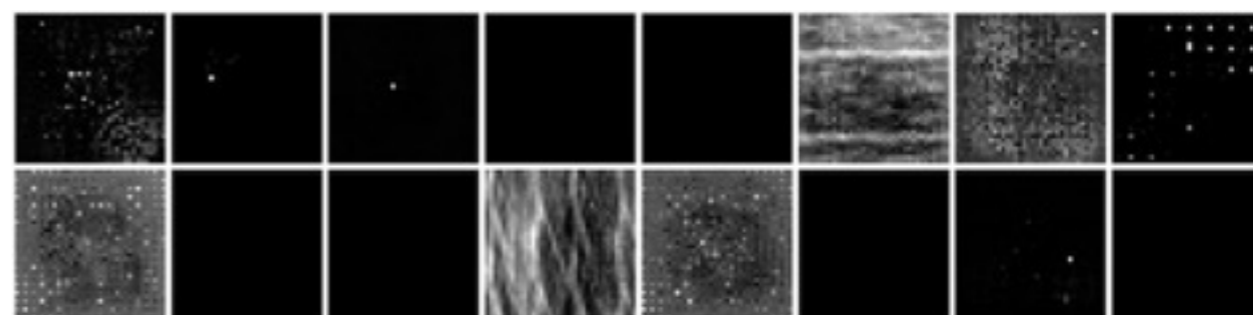
ポジティブサンプル, カーネルサイズ3x3



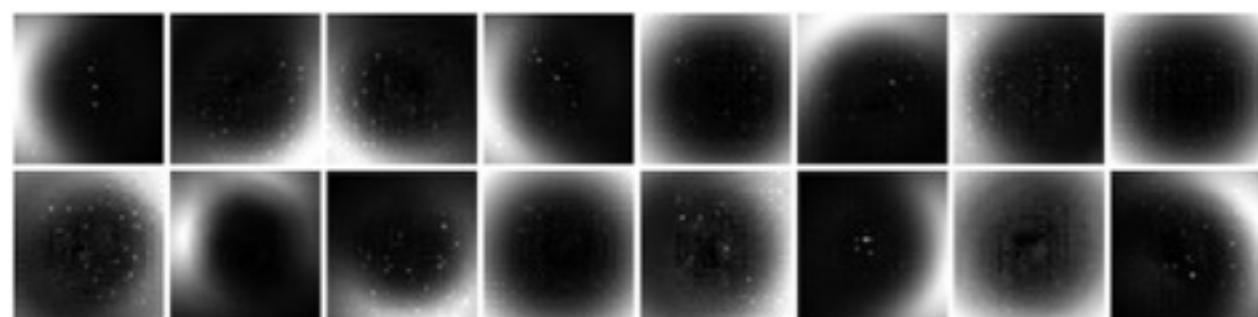
ネガティブサンプル, カーネルサイズ3x3



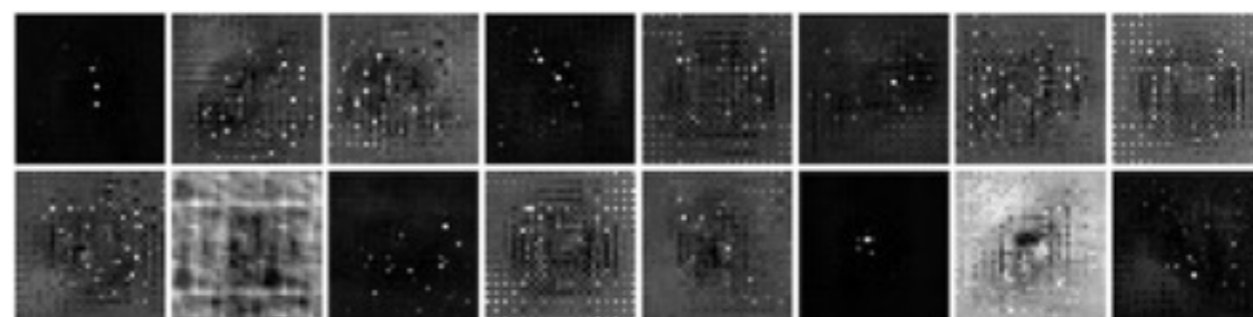
ポジティブサンプル, カーネルサイズ5x5



ネガティブサンプル, カーネルサイズ5x5

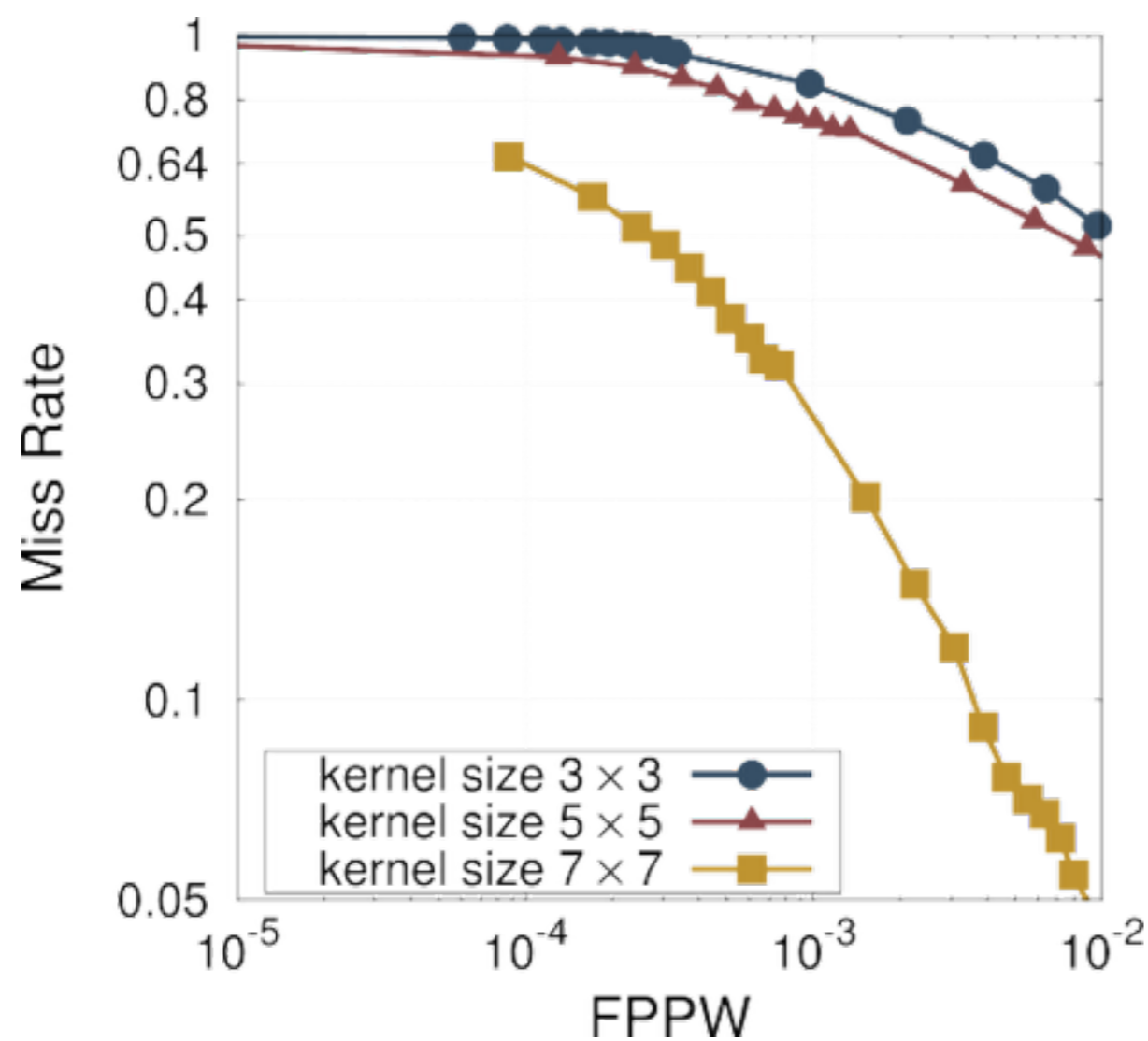


ポジティブサンプル, カーネルサイズ7x7

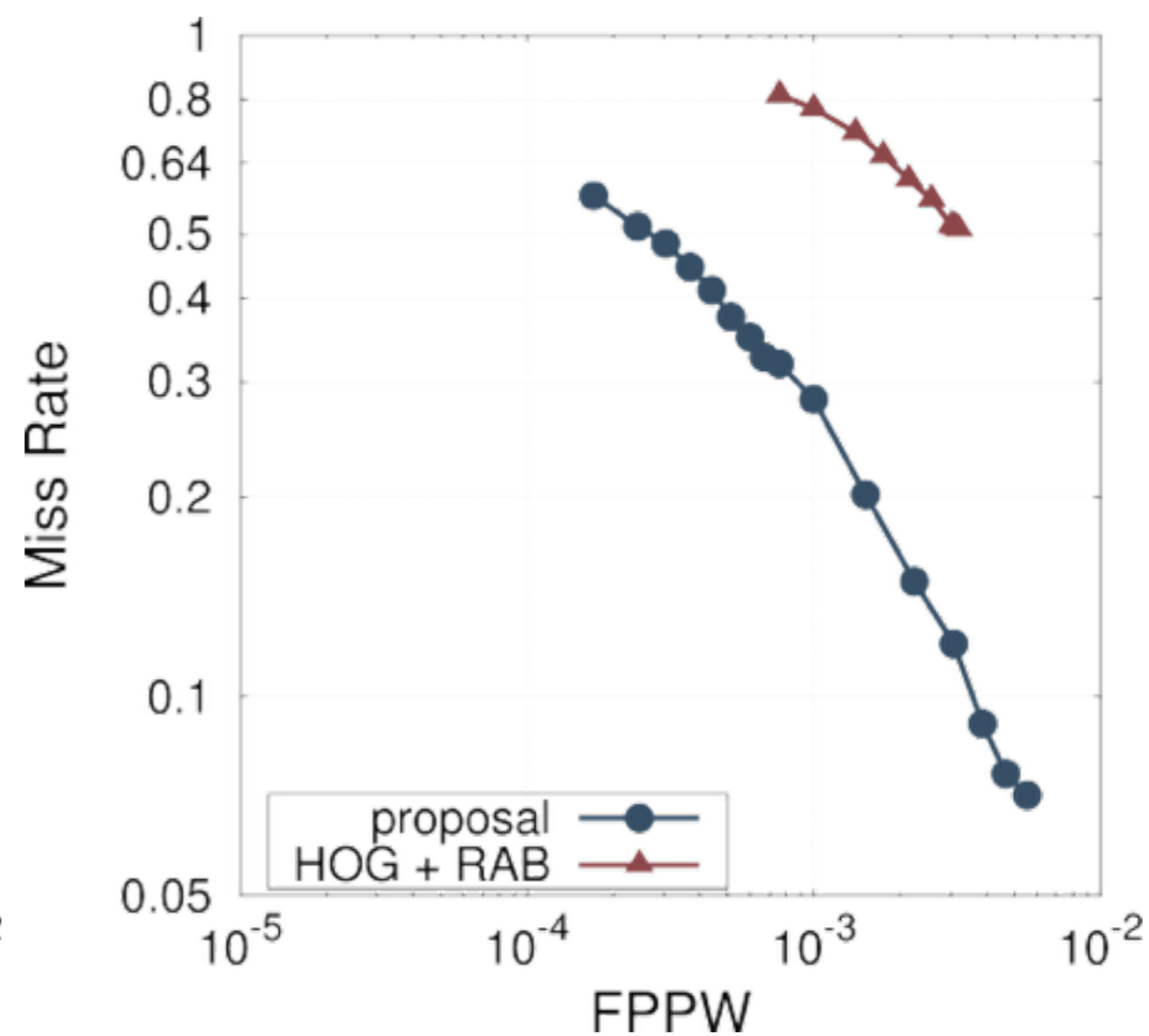


ネガティブサンプル, カーネルサイズ7x7

# 検出性能



特徴マップ数16におけるDET曲線



既存手法と提案法のDET曲線

# おわりに

---

これから深層学習を使ってみようとお考えの方々を想定し、次の観点で概説しました

- 深層学習とはどのようなものか？
- 深層学習はどのような時に使えばよいのか？
- 深層学習はどのように使えばよいのか？

理論的面では、ニューラルネットワークの基本的なところから多層化時の学習方法までの考え方についてお話ししました

深層学習で大事な点は、入力データの特徴を学習するところにあります。どのような特徴を学習するか？の観点で研究を進められると新しい発見があるかもしれません