

# 「この画像は見たことない」と 言えるAIを目指して

九州工業大学大学院 生命体工学研究科  
人間知能システム工学専攻  
准教授 田向権  
博士研究員 鈴木章央

令和2年12月10日@JST 新技術説明会



# 発表の枠組み

---

- 導入編
  - ✓ この技術について
  - ✓ AI, ディープラーニングと画像分類
- 技術編
  - ✓ 今回の目標や技術紹介
  - ✓ 実験とその成果
- アプリケーション編
  - ✓ 防犯カメラへの導入
  - ✓ 製造ラインへの導入
  - ✓ 波形データを扱う



# 導入編



# この技術の概要

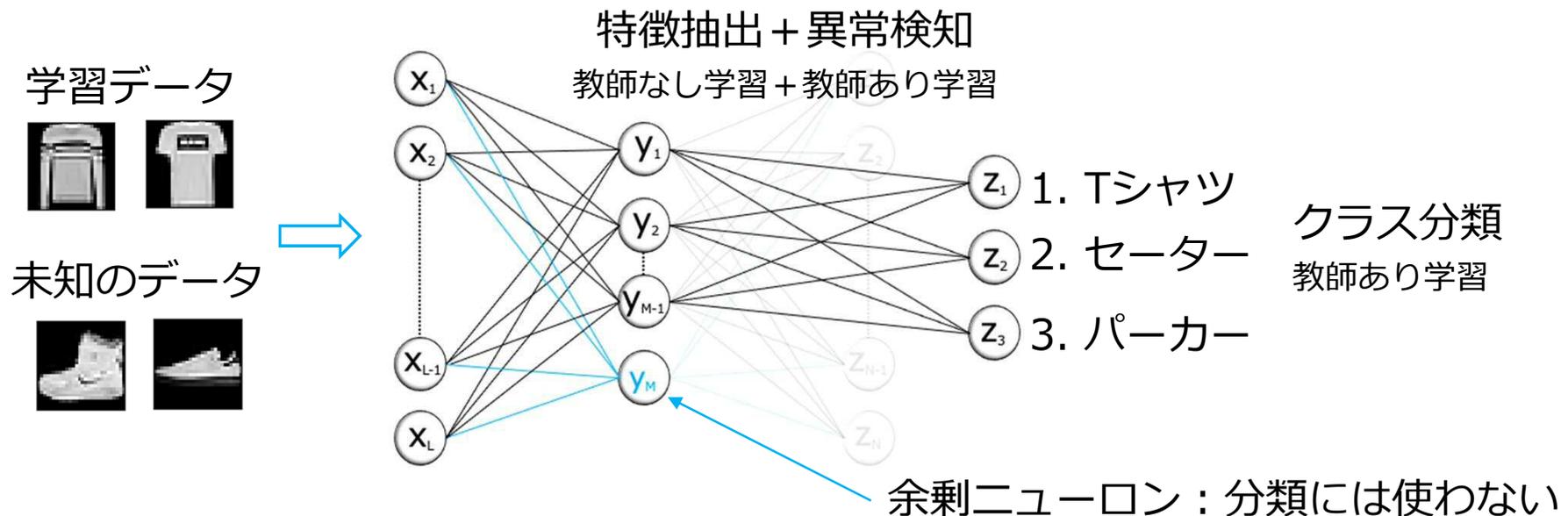
---

- 分野：画像処理  
→今回は応用先も画像を用いたものをベースに
- 手法：AIの中でもディープラーニングを扱い、  
教師なし学習と教師あり学習を混合したもの
- 新規性：AIの学習手法や構造  
→AIの基本的な部分になるので、分野も横断でき、  
理論的にも実運用的にも応用の幅が広い
- 段階：理論の検証が完了している段階  
→次の段階は精度のブラッシュアップ  
→応用先の要求精度によっては運用も可能

# この技術の目標

学習データには反応せずに、未知のデータには反応するような余剰ニューロンを設置し、その状態を見ることでデータの既知・未知を判定し、既知データは分類できるようにする

例) トップスには余剰ニューロンは反応しないが靴には反応する  
トップスは分類器で半袖か長袖か分類する



# この技術でできるようになること

## この画像は見たことない, という

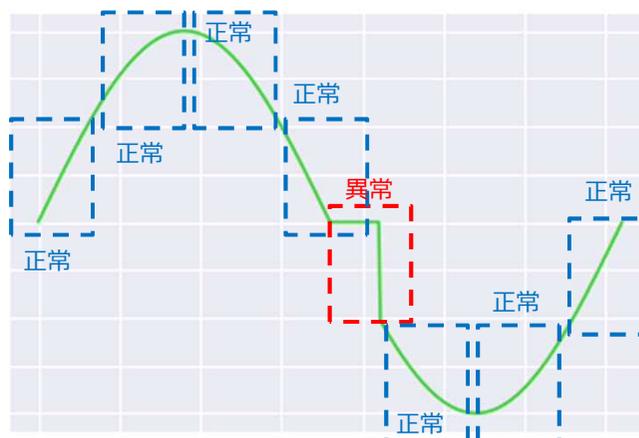
知らない人をわかるようにする



知っているものだけを分別する



異常(知らない形の領域)を見つける



主にAIと言われるのはディープラーニングだが、  
ディープラーニングだけがAIではない

- ディープラーニングが注目を集める
- 連られて機械学習や他のニューラルネットワークにも注目

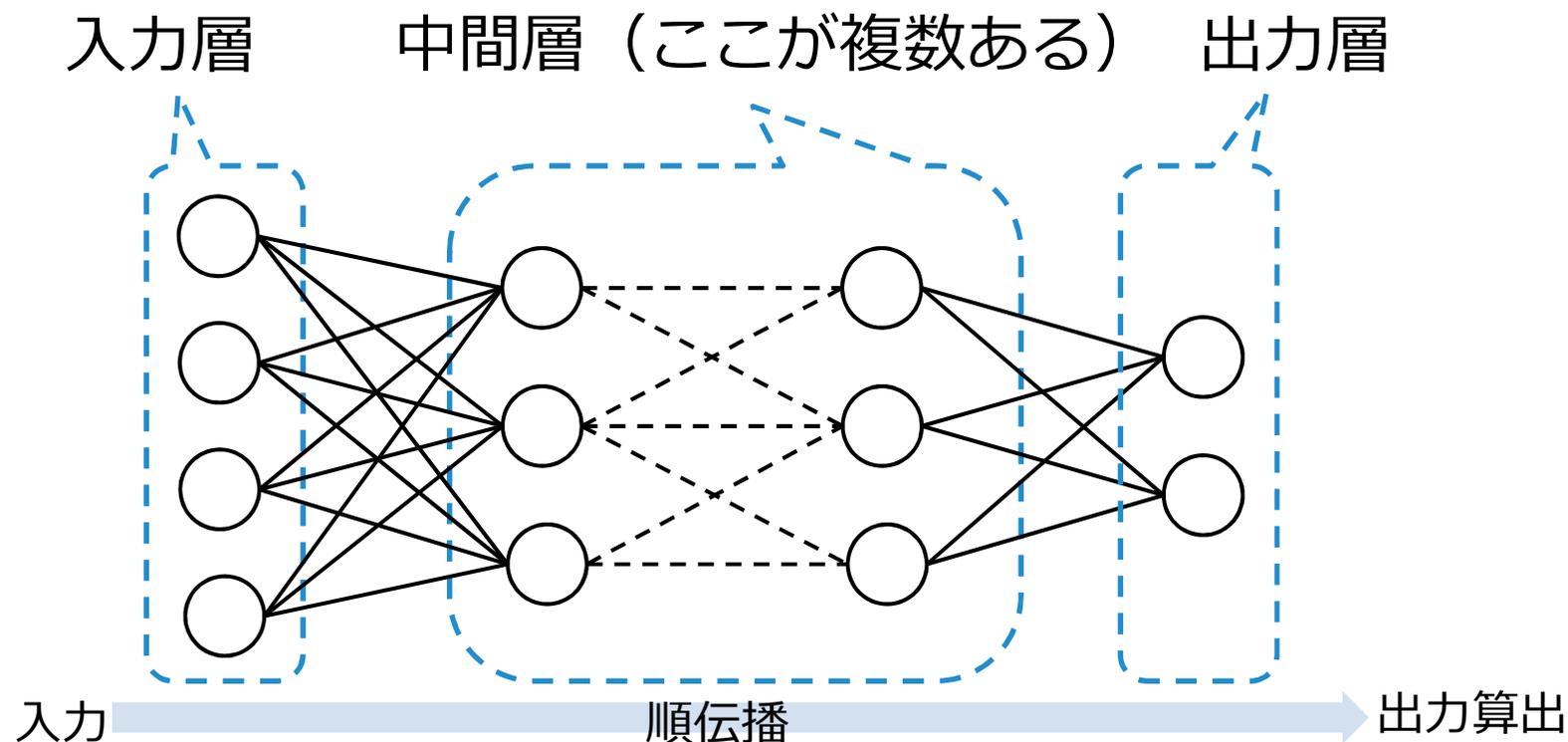
AI：人間の知的能力をコンピュータ上で実装したソフトウェアやシステム

機械学習：入力データと出力データの組み合わせを与えることで  
関係性や振る舞いを自動的に獲得していくアルゴリズム

ニューラルネットワーク：脳の神経回路をネットワーク構造で模倣し、  
数式でを使って伝達機構を表現した機械学習アルゴリズム

ディープラーニング：ニューラルネットワークの中でも  
ネットワークの層が深いものをさす

# ディープラーニングの概要



## 得意なこと

- 人が経験則で行っていること
- 大量のデータが用意できるもの
- 答えが用意されているもの

## 不得意なこと

- データが少ないもの
- 未来予測
- 知らないものを知らないという

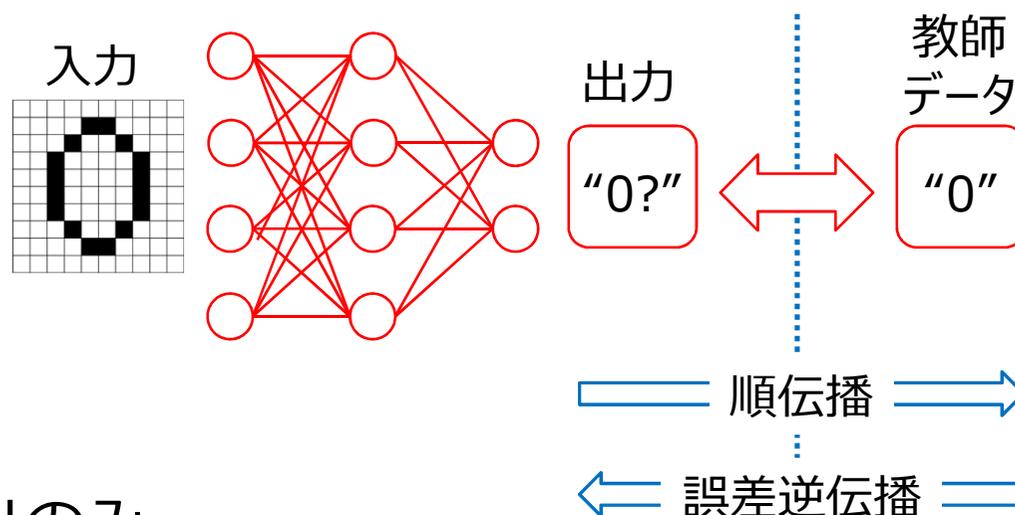
# ディープラーニングの2つの学習法

## 教師あり学習：特徴抽出＋分類器

1. 人がクラスを決める
2. クラスに沿った特徴を獲得



分類は得意だが  
未知データの対応は難しい  
→必ず何かに分類してしまう



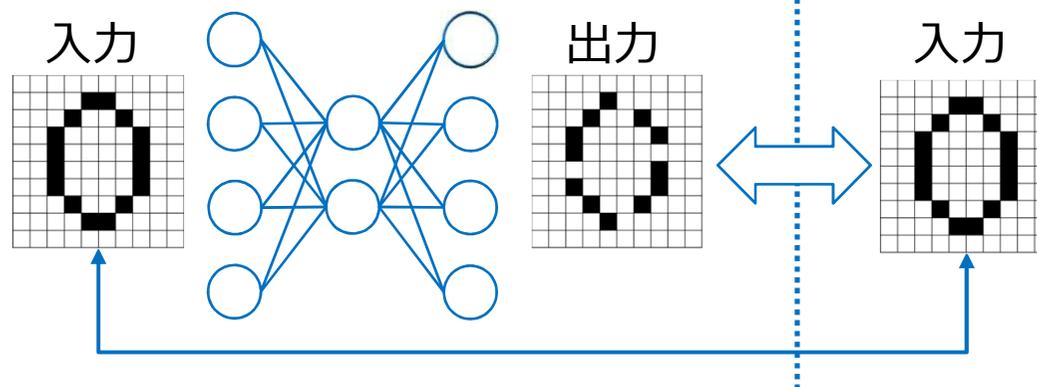
目標出力

## 教師なし学習：特徴抽出のみ

1. 人がクラスを決めない
2. 自動的に特徴を獲得

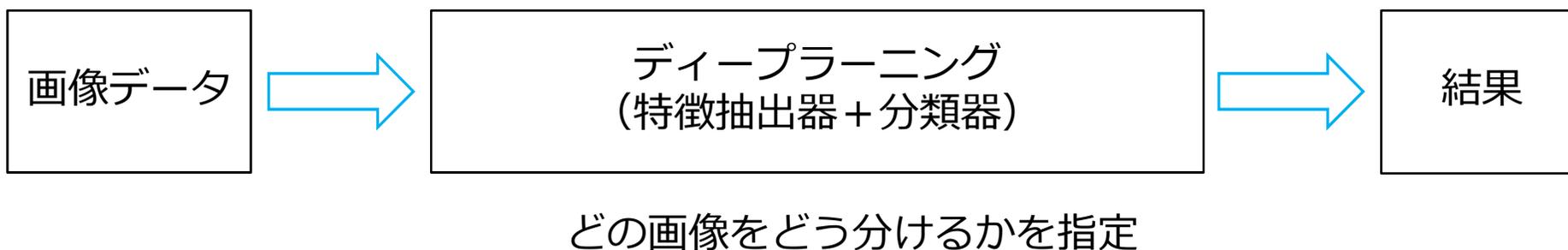
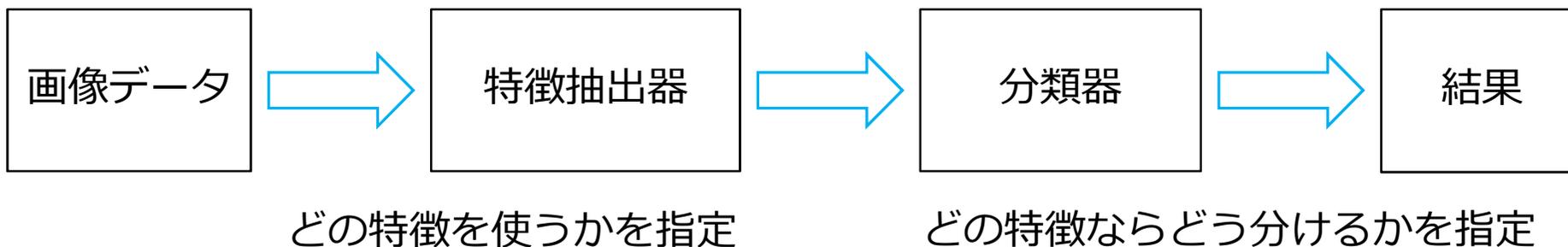


教師あり学習よりも  
未知データに対応しやすい



## 特徴抽出とクラス分類を行う

特徴：形や色，模様など，クラス：区分や仕分け方法のこと



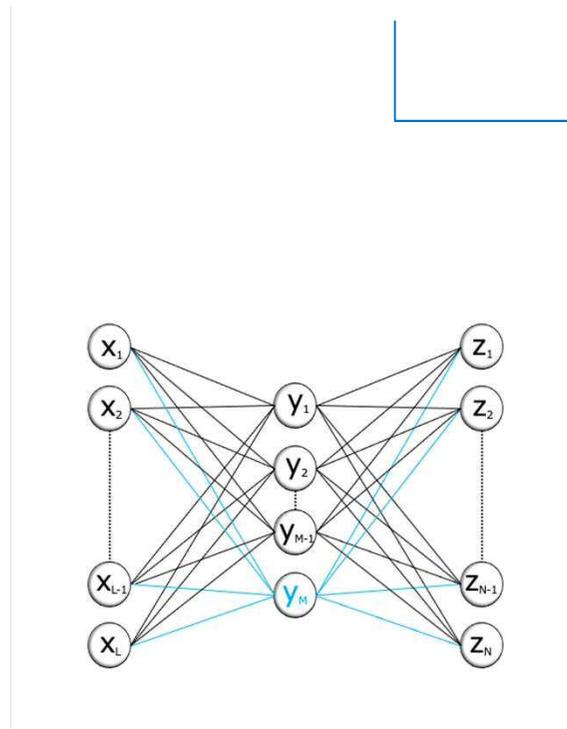


# 技術編

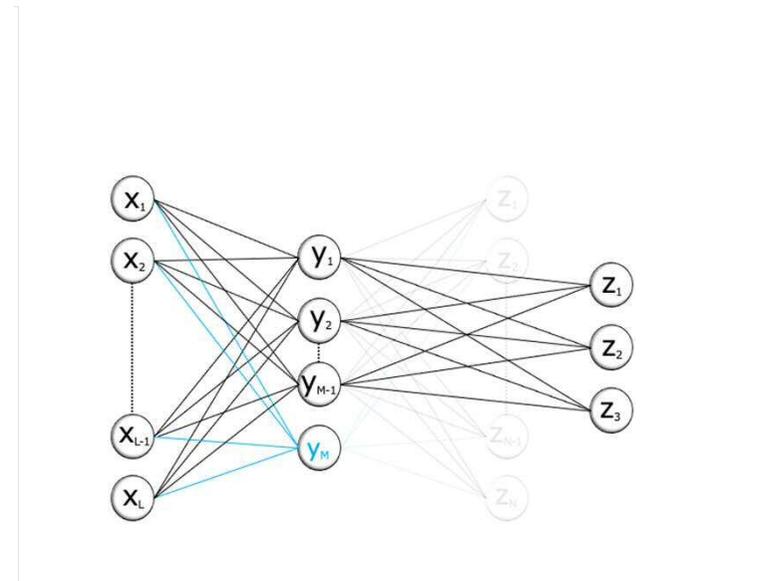


## 異常検知と分類, それぞれの実装と検証を分けて行う

### 特徴抽出 + 異常検知の実装



### 異常検知 + 分類の検証



## 異常検知の理論, 実装とその精度検証

### 特徴抽出 + 異常検知の実装

課題：理論と精度の検証

### 異常検知 + 分類の検証

課題：余剰ニューロンを除いても分類が可能かどうか

理論の組み立て

実装

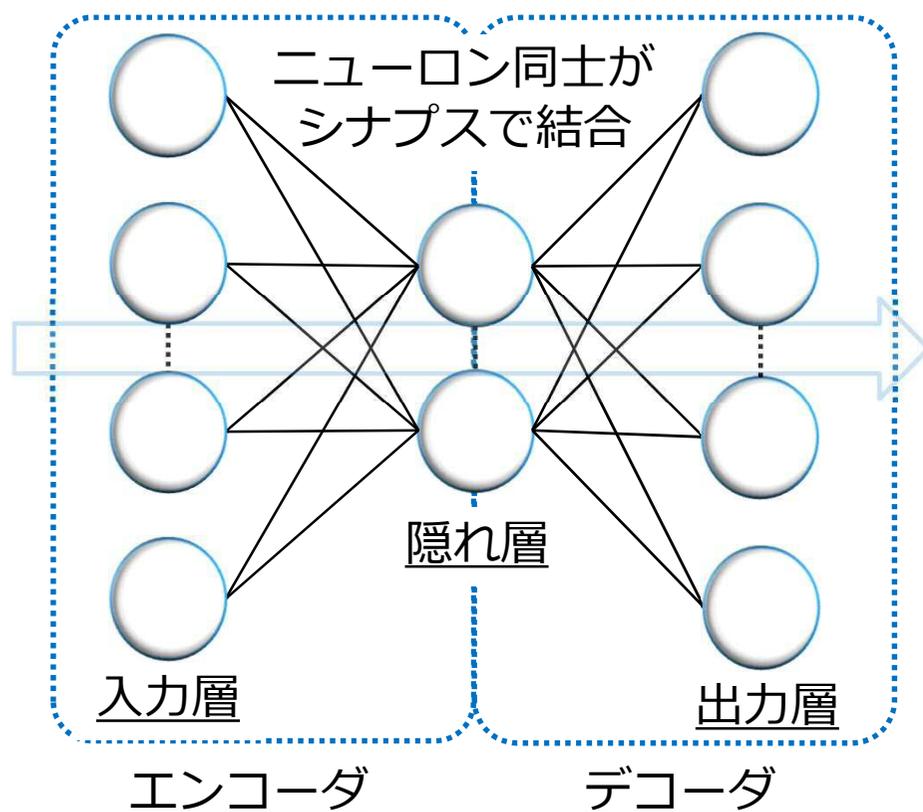
検証

精度向上

次のフェーズへ

# 今回用いるAI：自己符号化器

1. 3層構造のニューラルネットワーク
2. 入力が出力を再構成するように学習



## 自己符号化器 [1]の特徴と利点

1. 形状  
→砂時計型の構造のため、データを圧縮するように学習し、自動的に特徴獲得  
→処理の切り離しが可能
2. 教師なし学習  
→扱うデータが人手による分類に依存しないためどんなデータも学習可能

# 自己符号化器の学習方法

## 教師あり学習：特徴抽出 + 分類器

1. 人がクラスを決める
2. クラスに沿った特徴を獲得



分類は得意だが  
未知データの対応は難しい  
→必ず何かに分類してしまう



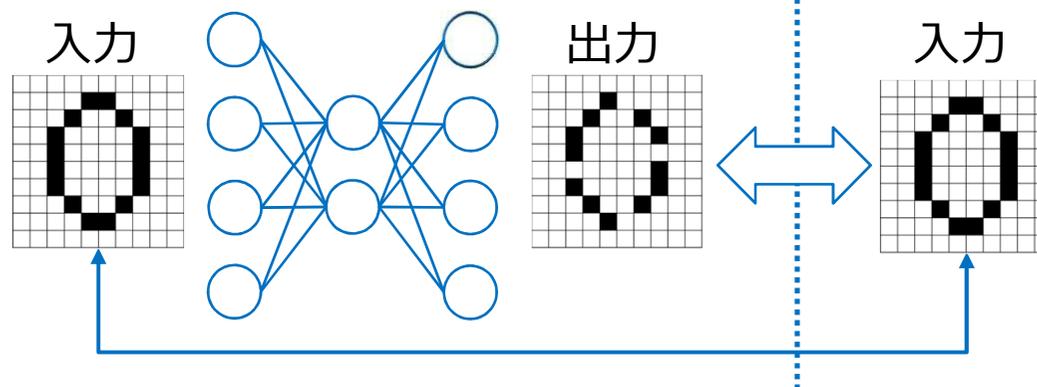
目標出力

## 教師なし学習：特徴抽出のみ

1. 人がクラスを決めない
2. 自動的に特徴を獲得

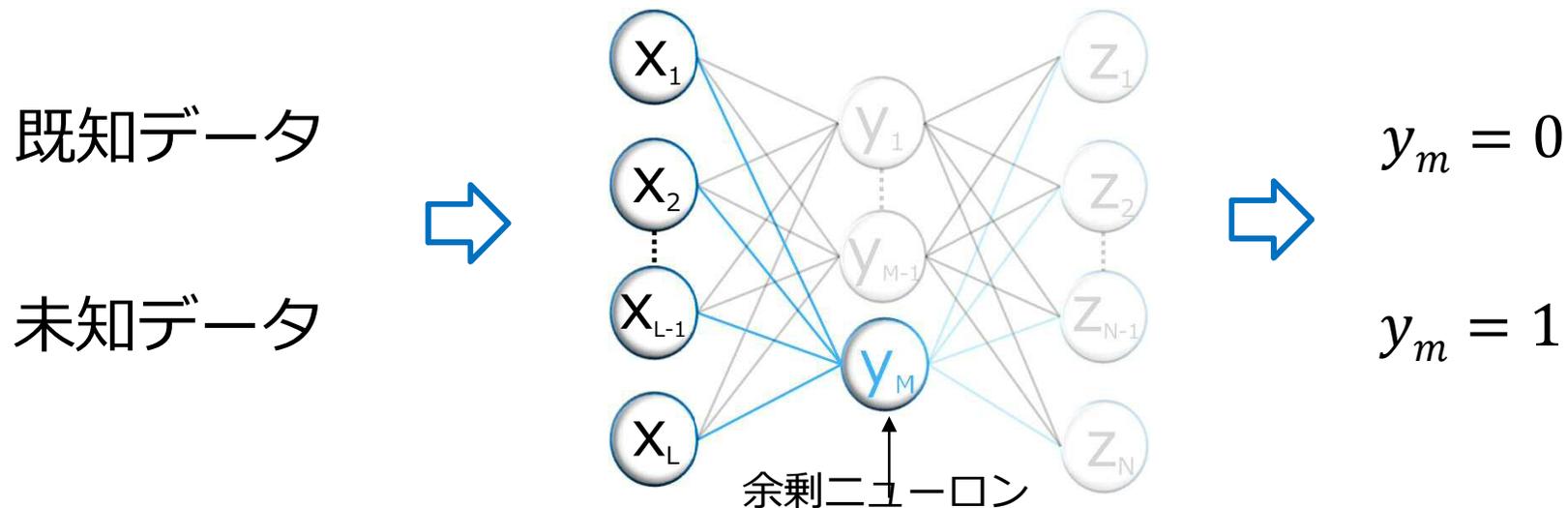


教師あり学習よりも  
未知データに対応しやすい



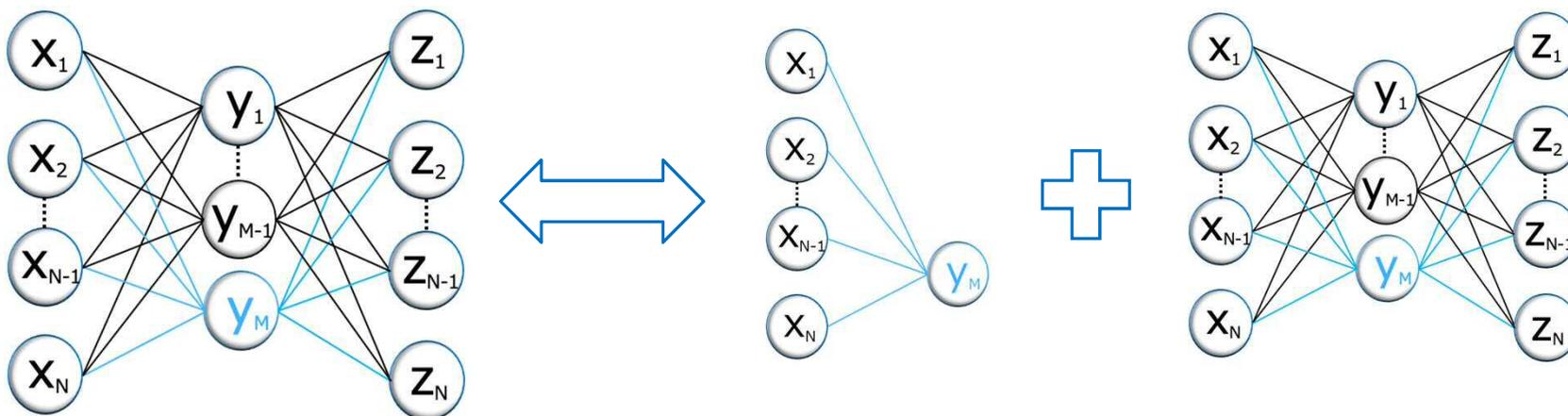
# 余剰ニューロンの理屈

1. 既知データと未知データで潜在表現は全く異なると仮定
2. 潜在表現空間に, 既知データの表現に不要な次元を用意  
→中間層に余剰ニューロン
3. 余剰ニューロンは既知データの学習では反応させない
4. 既知データには反応しなくても未知データには反応する



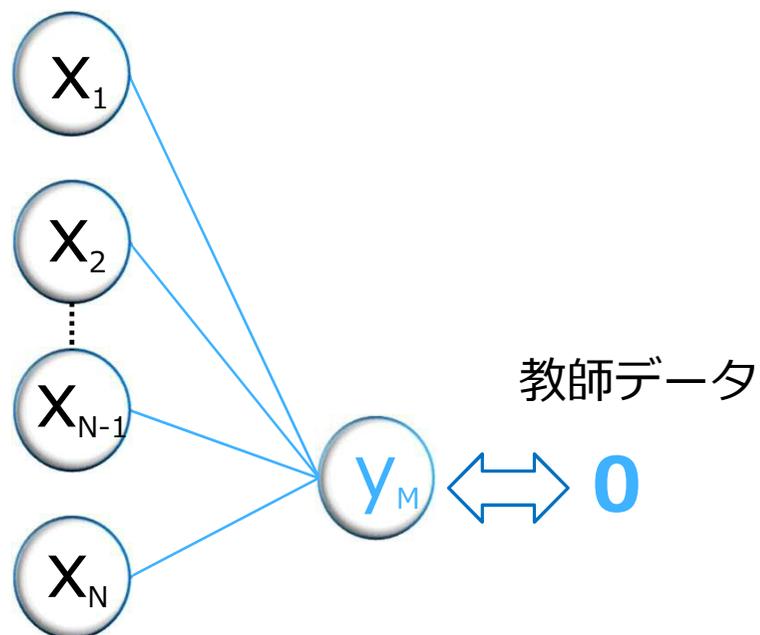
- 自己符号化器を分割して学習

→ エンコーダの学習で余剰ニューロンの反応を制御

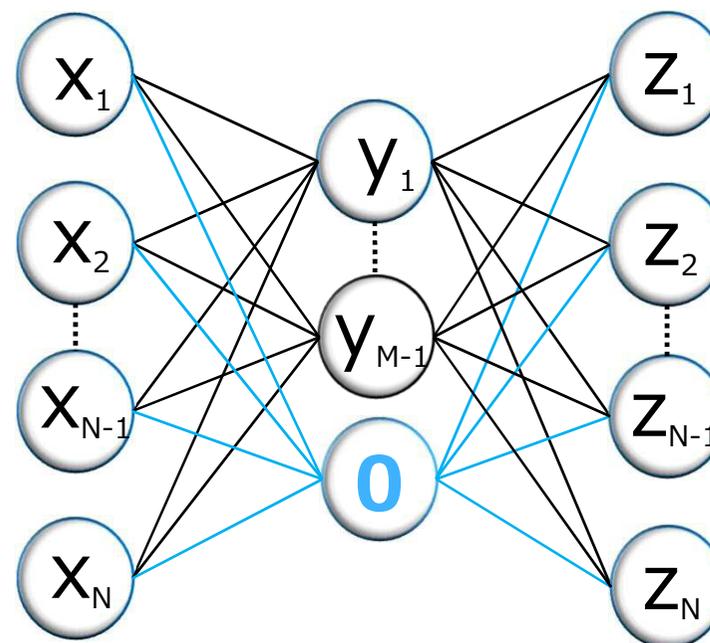


- 同時に小数のノイズを学習

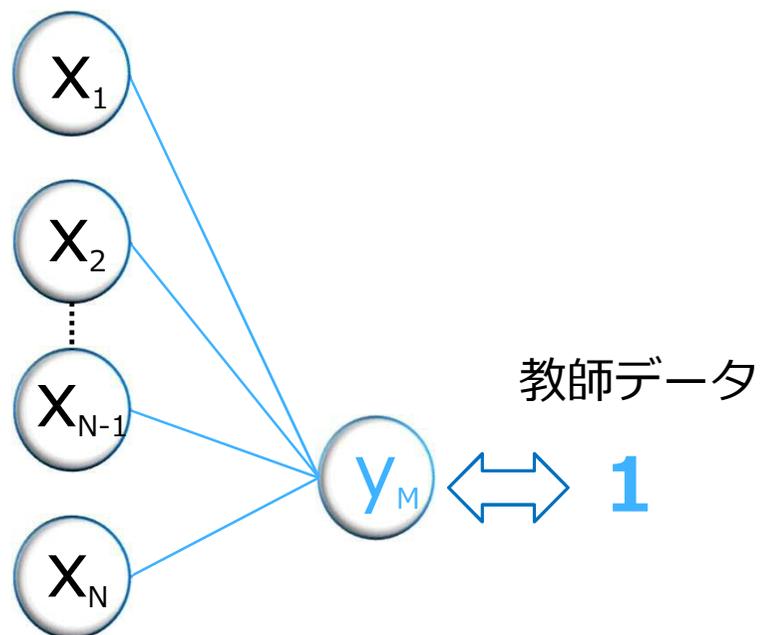
→ 既知データ以外には反応するようにしておく



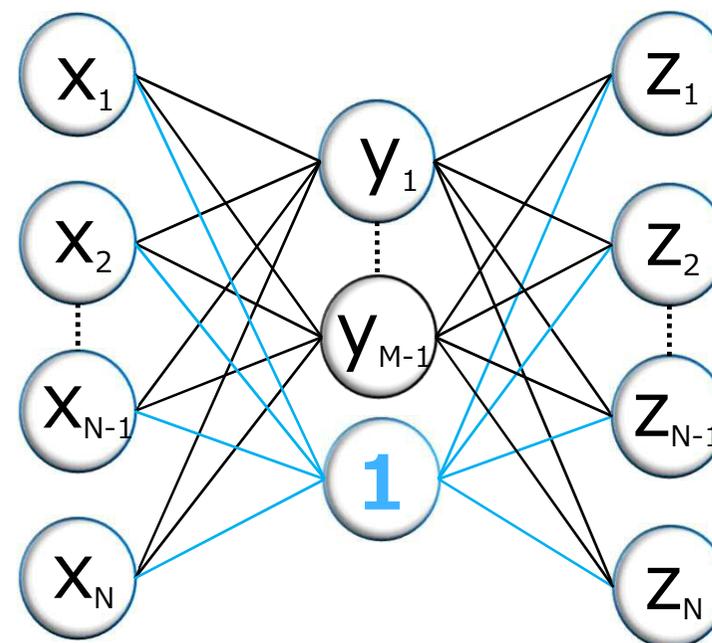
エンコーダを一部切り離し  
教師データとして0を与えて  
教師あり学習する



余剰ニューロンに0ベクトルを  
与えてAEとして学習する



エンコーダを一部切り離し  
教師データとして1を与えて  
教師あり学習する



余剰ニューロンに1ベクトルを  
与えてAEとして学習する

# 検証実験

1. 既知データのみでの学習(余剰ニューロンの学習の検証)
2. 分類対象外データの同時学習
3. ノイズの同時学習

→それぞれ、学習が可能か、余剰ニューロンの反応を検証

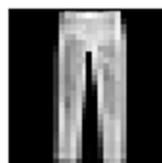
- AEの構成：784-100；うち余剰ニューロン10
- 活性化関数：中間層はReLU関数，出力層は恒等関数
- エンコーダ，自己符号化器全体のどちらを学習するときも，平均二乗誤差をとって回帰させる
- エポック数：50

# Fashion-MNISTデータセット

- 縦横28ピクセル, グレースケール, 10クラス
- 学習用 : 6,000枚/クラス, テスト用 : 1,000枚/クラス



0:Tシャツ



1:ズボン



2:プルオーバー



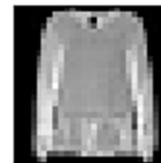
3:ドレス



4:コート



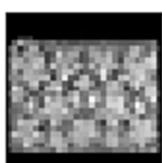
5:サンダル



6:シャツ



7:スニーカー



8:バッグ

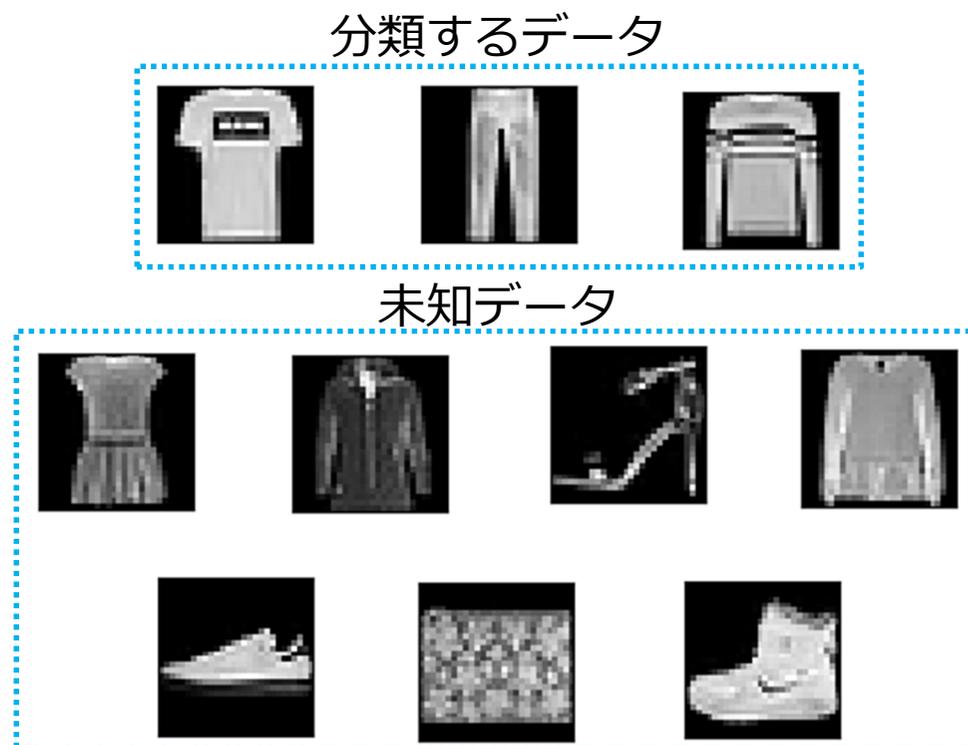


9:アングルブーツ

# 実験1. 既知データのみでの学習

## 特定のニューロンを反応させない学習法の検証

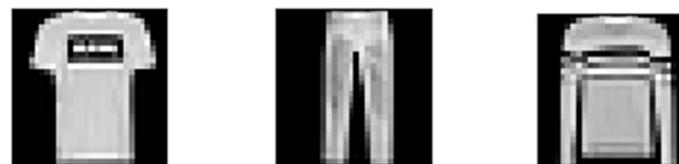
- 既知データ(分類するデータ)の学習のみ  
→ラベル0~2のデータ
- データ数：18,000枚
- ミニバッチ：200
- 分類するデータ,  
未知データで検証



## 未知データに対して余剰ニューロンを反応させる

- 分類するデータの学習  
→ラベル0~2のデータ  
→18,000枚
- 分類対象外データの学習  
→ラベル7~9のデータ  
→450枚(ランダム)
- ミニバッチ：200, 5
- 分類するデータ,  
分類対象外データ,  
未知データで検証

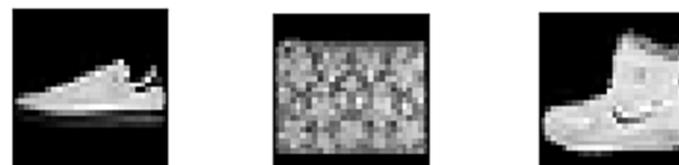
分類するデータ



未知データ



分類対象外データ

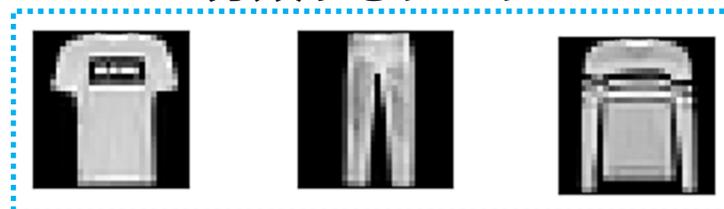


# 実験3. ノイズの同時学習

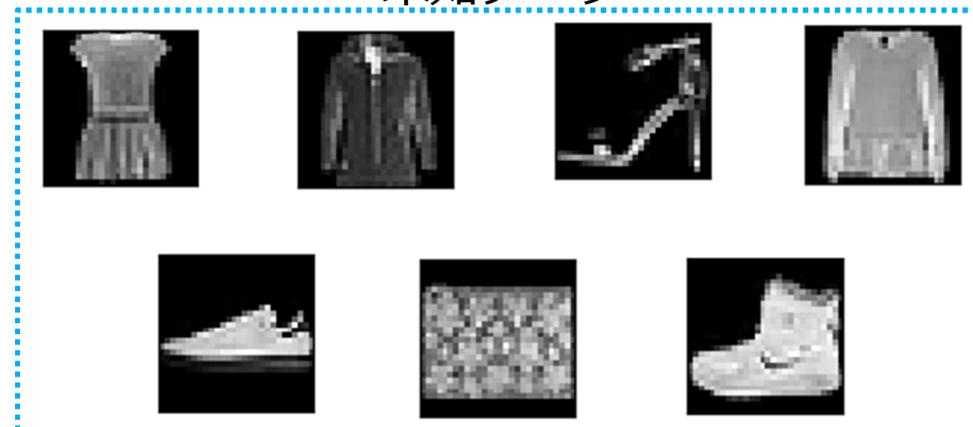
ノイズに余剰ニューロンが反応するようにして、より一般的な入力に反応できるようにさせる

- 分類するデータの学習  
→ラベル0~2のデータ  
→18,000枚
- ノイズの学習  
→450枚(一様乱数)
- ミニバッチ：200, 5
- 分類するデータ,  
未知データで検証

分類するデータ



未知データ



ノイズ



# 実験結果の比較：未知データへの反応の割合

分類するデータは反応せず，未知データには反応させたい  
→分類：100%，未知：0%が理想

学習の種類	既知データのみ		分類対象外データを同時		ノイズを同時	
	枚数	割合(%)	枚数	割合(%)	枚数	割合(%)
データの属性						
分類する	3,000/3,000	100	2,844/3,000	94.8	2,831/3,000	94.4
未知	6,829/7,000	97.6	2,345/4,000	58.6	3,319/7,000	47.4
分類対象外			14/3,000	0.47		

# 実験結果の比較：未知データへの反応の割合

分類するデータは反応せず，未知データには反応させたい  
→分類：100%，未知：0%が理想

学習の種類	既知データのみ		分類対象外データを同時		ノイズを同時	
データの属性	枚数	割合(%)	枚数	割合(%)	枚数	割合(%)
分類する	3,000/3,000	100	2,844/3,000	94.8	2,831/3,000	94.4
未知	6,829/7,000	97.6	2,345/4,000	58.6	3,319/7,000	47.4
分類対象外			14/3,000	0.47		

既知データのみでの学習だと  
既知・未知データどちらも反応しにくい

# 実験結果の比較：未知データへの反応の割合

分類するデータは反応せず，未知データには反応させたい  
→分類：100%，未知：0%が理想

学習の種類	既知データのみ		分類対象外データを同時		ノイズを同時	
	枚数	割合(%)	枚数	割合(%)	枚数	割合(%)
データの属性						
分類する	3,000/3,000	100	2,844/3,000	94.8	2,831/3,000	94.4
未知	6,829/7,000	97.6	2,345/4,000	58.6	3,319/7,000	47.4
分類対象外			14/3,000	0.47		

同時に何かを学習させると  
既知データの入力でも反応するようになる

# 実験結果の比較：未知データへの反応の割合

分類するデータは反応せず，未知データには反応させたい  
→分類：100%，未知：0%が理想

学習の種類	既知データのみ		分類対象外データを同時		ノイズを同時	
	枚数	割合(%)	枚数	割合(%)	枚数	割合(%)
データの属性						
分類する	3,000/3,000	100	2,844/3,000	94.8	2,831/3,000	94.4
未知	6,829/7,000	97.6	2,345/4,000	58.6	3,319/7,000	47.4
分類対象外			14/3,000	0.47		

意味のあるデータ(衣料品)でもノイズでも反応するようになる

これは0になった割合なので，ノイズを学習した方が反応を示している

# 実験のまとめと精度向上にむけて

- 実験結果より、余剰ニューロンを持つ自己符号化器の学習が可能であるとわかった
- データセットは分類対象データとノイズだけでよいが、今回の実験設定だと反応が小さくなる
- 情報量の多いデータセットと畳み込み層の利用  
→ エッジや線などの低次元の特徴から物体などの高次元の特徴まで獲得できる
- この理論の適用にふさわしい課題を見つける

## 課題に合わせて異常検知の精度向上と分類の精度検証

### 特徴抽出 + 異常検知の実装

課題：理論と精度の検証

理論の組み立て

実装

検証

精度向上

### 異常検知 + 分類の検証

課題：余剰ニューロンを除いても分類が可能かどうか

既存の理論

実装

検証

精度向上



# アプリケーション編

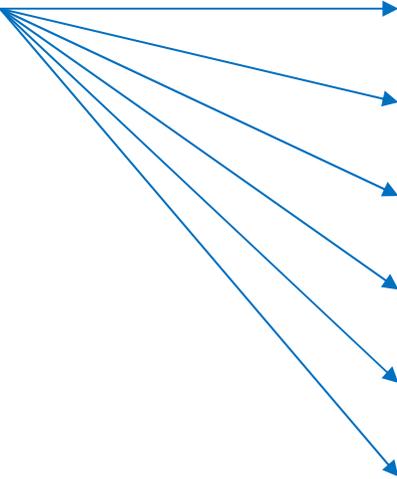


# 改めてこの技術で何ができるか考える

この画像は見たことない, という

【あるべきでない】 【あるはずでない】 と視認するものを  
対象とシーンから考えてみる

## 対象の種類や状態 + シーン

- 人物の外観や行動
  - 日用品, 食べ物の状態
  - 天候の状態や推移
  - スマホの画面
  - 書類
  - 取引先の顔色
- 
- 家屋・建屋
  - 道路
  - 工場
  - 公園
  - 学校 (敷地込みの建屋)
  - 公の空間

## 人物の外観 + 家屋

顔の領域を画像から切り出し→提案手法で検知

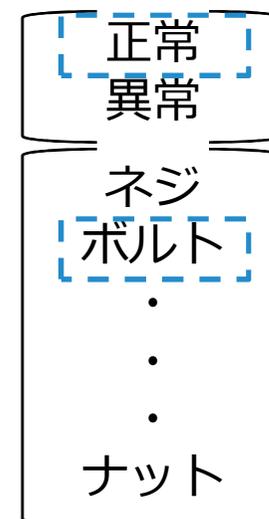
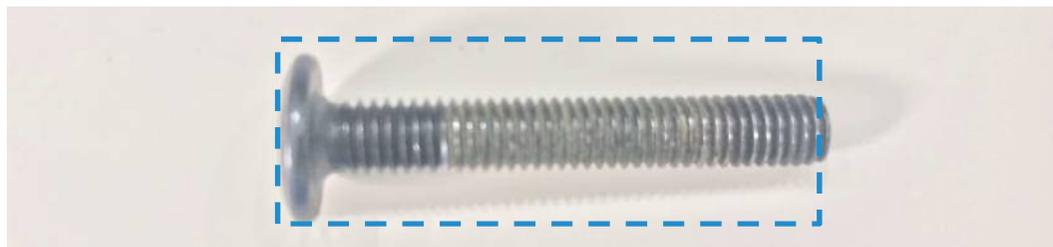
- メリット：顔の画像を登録しなくてもよい，顔が少し変わっても大丈夫  
※登録をした方が運用までのリードタイムは短くなる
- デメリット：競合が多く，精度を競う必要がある



## ものの外観 + 工場

コンベア上の画像から物体を切り出し→提案手法で検知

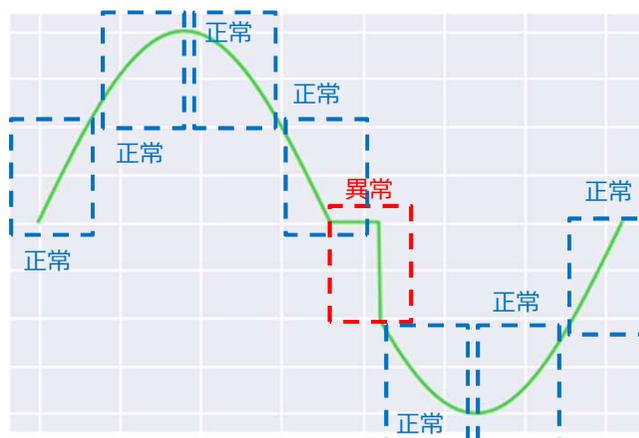
- メリット：異物がどうかに加えて，似ている部品もわかる  
※流れる部品を全て洗い出す必要がある
- デメリット：AIが異物とした根拠が掴めない  
※なぜ，なにが異物として混入したかを解決する必要がある



## センサー値 + 工場

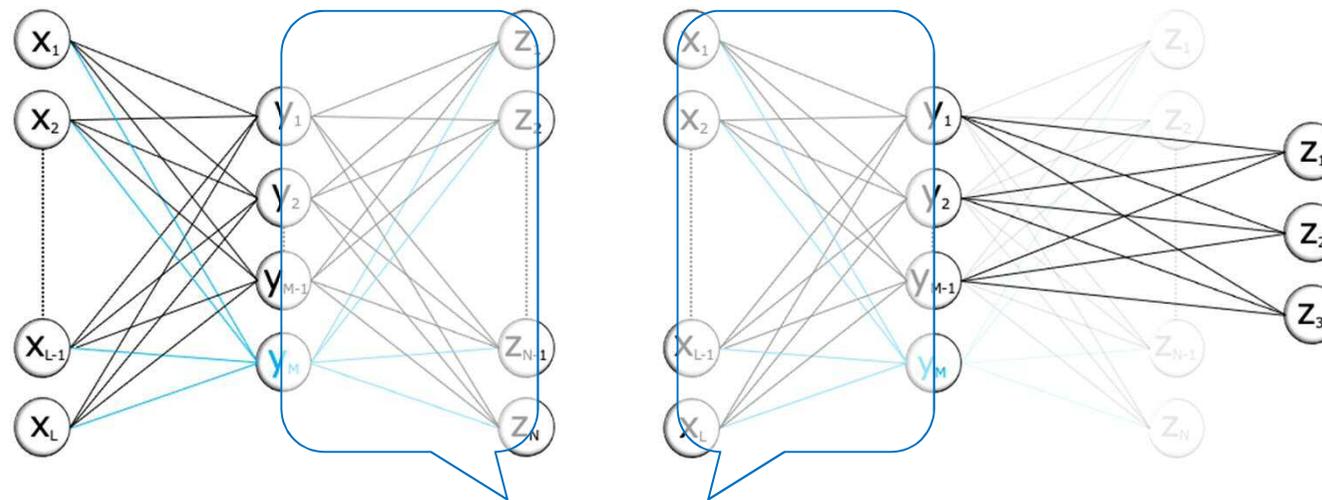
センサー値を図に描画→提案手法で検知

- メリット：点検回数を減らせる，故障予測にも期待できる
- デメリット：複雑な波形は扱いにくい，リアルタイムで判定できない



# 他の分野への拡張

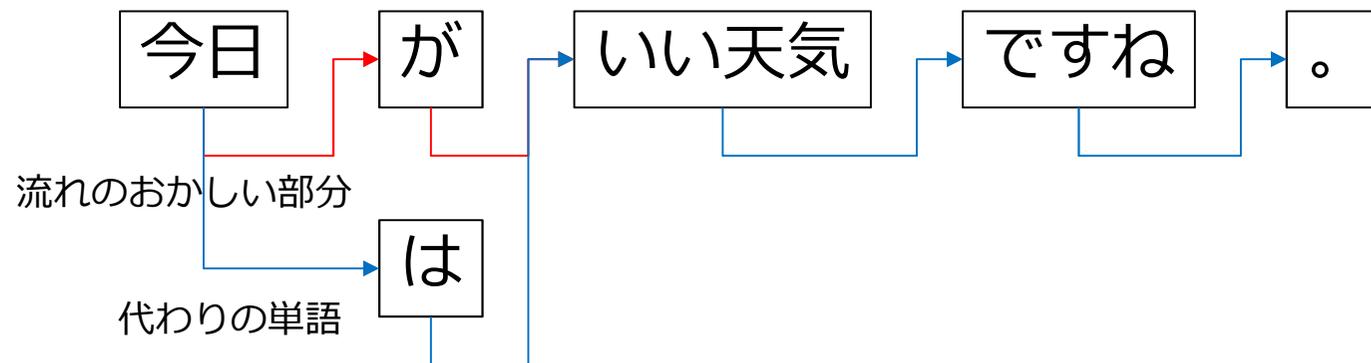
- この技術の特色： AIの学習手法や構造
- 今回使っていないAIのアルゴリズム
  - ✓ 画像に強い： 畳み込みニューラルネットワーク
  - ✓ 時系列に強い： リカレントニューラルネットワーク



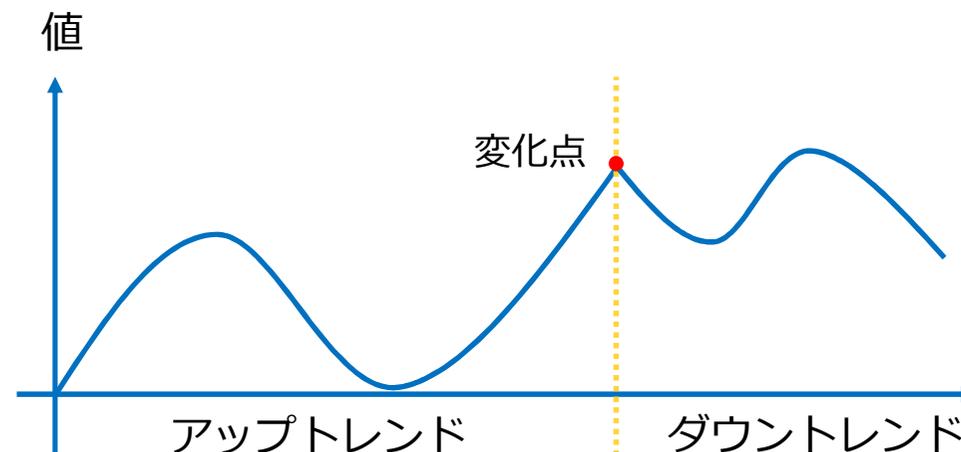
層と層の間の結合方法を他の部品に変えることができる

# 時系列分野への転用

- 文章の流れのおかしさを検知, 代わりにの単語を表示  
→文章の流れを学習して, 単語を予測+提案手法



- 株の終値を分類してトレンドと変化点を検知  
→株の流れを学習して, ティレンドを分類+提案手法



# 問合せ

- 九州工業大学大学院 生命体工学研究科  
人間知能システム工学専攻  
田向研究室 田向権, 鈴木章央
- Tel : 093-695-6143  
Fax : 同上
- e-mail : tamukoh@brain.kyutech.ac.jp  
[suzuki@brain.kyutech.ac.jp](mailto:suzuki@brain.kyutech.ac.jp)
- 九州工業大学 オープンイノベーション推進機構 産学官連携本部  
  
小柳 嗣雄  
  
TEL : 093-884-3498  
  
FAX : 093-884-3531  
  
e-mail : koyanagi-t@ccr.kyutech.ac.jp