

ICTイノベーションフォーラム2018

検診結果に基づく深層学習による予測システムの開発と ひろしま健康長寿ネットワークの構築

(SCOPE課題番号：162308002, H28~29年度)

研究代表者：市村匠¹⁾

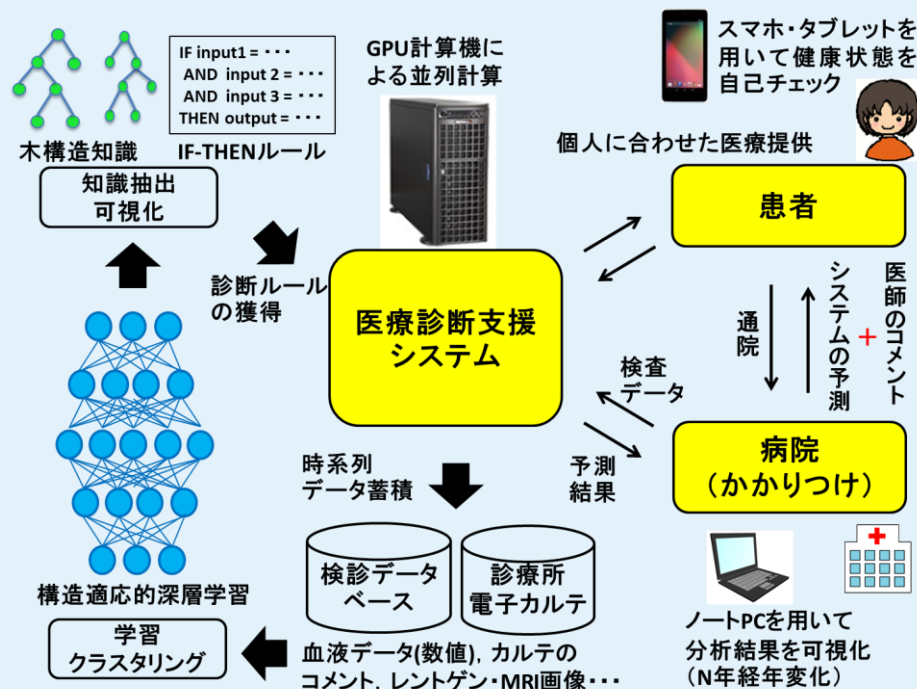
研究分担者：田村慶一²⁾, 原田俊英³⁾, 飯田忠行³⁾

- 1) 県立広島大学 高度人工知能プロジェクト研究センター長,
地域連携センター長, 経営情報学部 経営情報学科
- 2) 広島市立大学大学院 情報科学研究科
- 3) 県立広島大学 保健福祉学部 理学療法学科

研究開発の内容

- 近年，生活習慣病等による死亡のリスクの上昇，特定健診の受診率の低下，中山間地や島しょ部における専門医不足が問題になっている。
- 本研究課題では，深層学習(Deep Learning)による理論的な学習手法を開発するとともに，検診機関，診療所と連携し，血液検査や問診データ，医療画像等から構成されるマルチモーダルな医療検診データから，医療に当たる医師の診断を支援する医療診断支援システム構築した。
- 開発したシステムを広島県の複数の病院に導入し，継続的に検査データの経年変化をモニタリングし，日常的な健康管理を行い，健康寿命を増加させ，N年経年変化の予測の正答率を向上させた。
- このために，
 - 独自の深層学習手法
 - クラスタリング手法
 - 知識獲得手法
 - N年経年変化の予測手法
 - 医療診断支援システムの構築に関する研究開発を行った。

患者自身がデータに基づいた健康管理の重要性を認識できれば，検診率が増加し，癌の早期発見につながる。



開発するシステムの概要

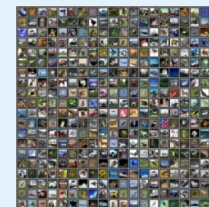
研究開発の成果 (1)

①構造適応型Deep Learningの学習手法の開発

- 構造適応型学Deep Belief Network(DBN)を開発した[1]
 - DBNは、深層学習の代表的な手法の一つで、事前学習を行った複数のRestricted Boltzmann Machine (RBM)を階層的に組み合わせる手法である。
 - 学習中に最適なネットワーク構造を自動決定できる。
- 通常、高い分類精度を持ったネットワークの構築するためには試行錯誤的なパラメタ調整が必要。
- 画像ベンチマーク(CIFAR-10等)に対し既存の深層学習手法より高い分類精度を示した。胸部X線画像CXR8の8種類の疾病に対して99%の分類精度を示した。

CIFAR-10とCIFAR-100[2]に対する分類精度

- 50,000 学習用画像
- 10,000 検証用画像
- 10 カテゴリ (CIFAR-10), 100 カテゴリ(CIFAR-100)



| 手法 | CIFAR-10 | CIFAR-100 |
|-----------------|----------|-----------|
| 従来手法 (CNN) | 96.5% | 75.7% |
| 提案手法 (構造適応型DBN) | 97.4% | 84.5% |

学習用データに対しては100%

[1] S.Kamada, T.Ichimura, A.Hara, and K.J. Mackin, "Adaptive Structure Learning Method of Deep Belief Network using Neuron Generation-Annihilation and Layer Generation", Neural Computing and Applications (2018)

[2] <https://www.cs.toronto.edu/~kriz/cifar.html>

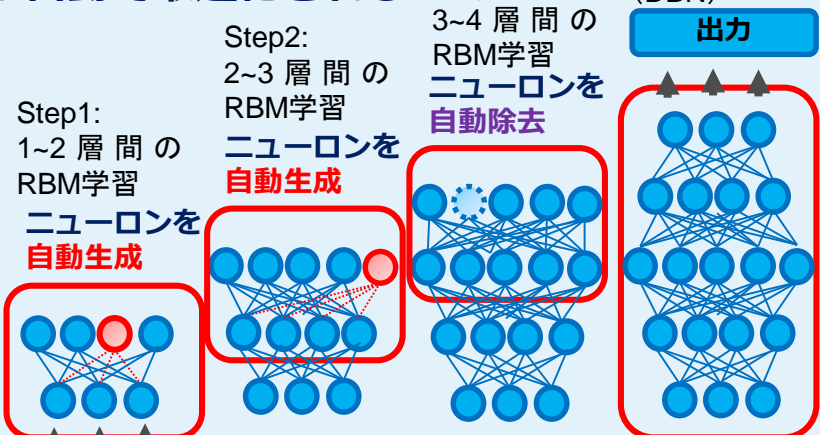
入力データに応じて、ニューロンの数、層の数
が自動で最適化される

Step1:
1~2層間のRBM学習
ニューロンを自動生成

Step2:
2~3層間のRBM学習
ニューロンを自動生成

Step3:
3~4層間のRBM学習
ニューロンを自動除去

Step4:
出力層を加えて教師あり学習(DBN)



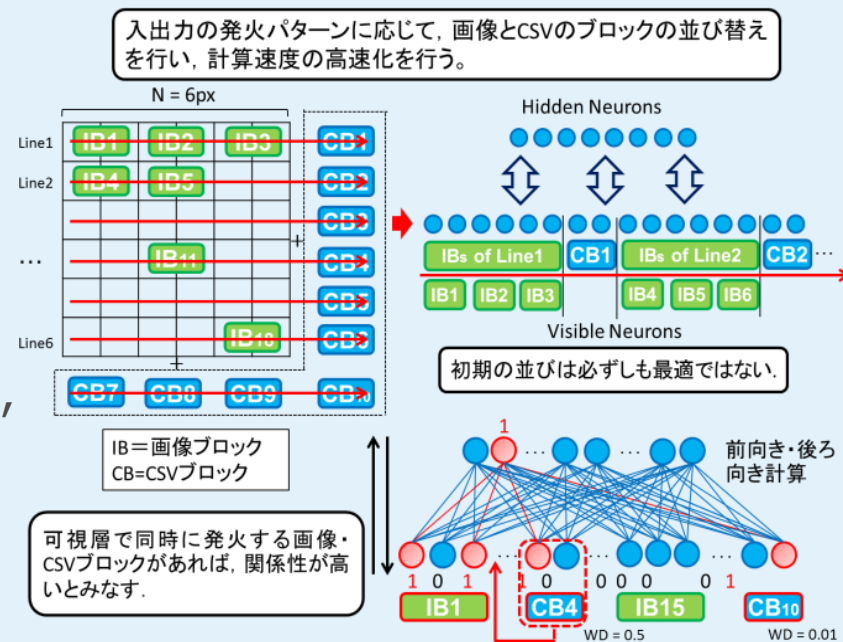
構造適応型DBN

最適化されたネットワーク構造

研究開発の成果（2）

②医療マルチモーダルデータのクラスタリング並列システムにおけるクラスタリングとRBM学習法の開発

- 医療検診マルチモーダルデータを高速に学習する手法を開発した(特許出願[1])
 - 医療データは血液検査, 問診, 医療画像(X線, CT, MRI等)から構成される。
 - ①の学習中の信号の流れから画像とその他の検診データの関連性を探索しながら最適に並び替えて配置する手法を開発した。
 - 過去の検診データに対して, 分類精度を維持したまま**学習時間を約30%削減**できた。
- 医療マルチモーダルデータのクラスタリング手法を開発した。
 - Deep Learningの構造とニューロンの活性化状態やパラメタをもとにユニットのクラスタリングを行う手法を開発し, クラスを分類するためのユニット群を発見した。
 - セルベースのDBSCANアルゴリズムにおいて, データ空間を複数の領域に分割し, 各領域においてマルチコアCPUによるクラスタリング処理を行うことで, 並列計算による計算速度の高速化を実現した。



[1] 市村匠, 鎌田真, 深層学習自動学習システム, クライアント装置およびサーバ装置, 特願2017-200830

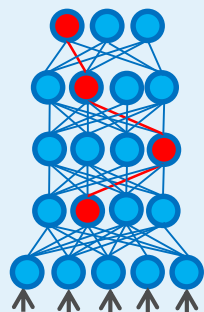
研究開発の成果（3）

③ Deep Learningから知識獲得法の開発

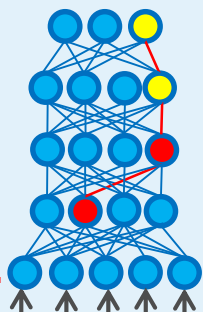
- ①の手法の学習後ネットワーク構造から知識獲得手法を開発した。
 - 入出力パターンから誤った出力を生じる信号の流れを発見し、正しく出力するように信号の流れを変換(調整)するFine Tuning法を開発した。
 - 過去検診データに対し、①の手法は95.5%の分類精度を示したが、③のFine Tuning法により、その分類精度は97.6%に向上した。

④ 構造適応型リカレントDeep Learningの構築

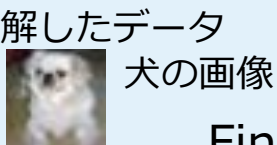
- 患者の経年変化を捉えた生存率の推定のために、時系列データに対するDeep Learning学習アルゴリズムを①の手法に基づいて開発した。
 - ベンチマークデータ(音素列, モーションキャプチャ)に対し高い予測精度を示した。
 - 過去の検診データに対し、4年目の患者の状態を90%以上の精度で予測できた。



信号のパターンを分析し、誤って出力した信号を正しく変換する知識を抽出した。



不正解したデータ
犬の画像
(猫に類似)



Fine Tuning法

ベンチマークテストに対する予測精度

Nottingham[1]: 音素列時系列データ

CMU[2]: モーションキャプチャ

| 手法 | Nottingham | CMU |
|---------------|------------|-------|
| リカレントDBN | 75.8% | 70.8% |
| 構造適応型リカレントDBN | 99.4% | 91.4% |

研究開発の成果（４）

⑤ 医療診断支援システムのプロトタイプ開発とその評価

- ①～④の手法による推論結果及び検査結果をタブレット端末上で表示する医療診断支援システムのプロトタイプを開発した。

⑥ 病院と連携した医療診断支援ネットワークの構築

- 広島県の医療機関と協働し、実証実験を行った。タブレット端末を通じてシステムの検査結果を送受信できる体制を構築した。

- 収集されたデータを用いた分析した結果、

- 血液検査項目の中で、癌に関連があると考えられる項目の特徴がいくつか見られた。癌の判定において、今後の経過観察で注意すべき特徴であり、癌の早期発見に至る可能性のある知見が得られた。
- 検診データの医師の判定結果の中には、ヒューマンエラーによる誤判定の可能性が考えられるケースが見つかった。

今後、医師らと共にケーススタディを行い、オーバーエスティメイトを含む誤判定がなくなるような体制を検診機関で構築する。

入力データに医療画像がない場合でも、血液検査データのみから推論可能

医療診断支援システム

脂質検査の結果、悪玉コレステロールが境界域(高い人と正常の人の間)でした。これ以上高くないよう飽和脂肪酸が多い動物性の脂肪を控え、多価不飽和脂肪酸が多い植物油や魚をとるようにしましょう。また、卵などコレステロールの多い食品も控え目にし、禁煙しましょう。ただし、もしあなたが糖尿病、慢性腎臓病、心血管病(心臓や血管の病気)などを持っている場合は、動脈硬化が進行している可能性が高く、心筋梗塞などになりやすい状態と考えられますので、医療機関でこれらの病気についての検査をお勧めします。引き続きご自身の身体の状態を確認するために、これからも健診を受診しましょう。

画像診断結果


胸部X線画像

| 項目名 | 疾病確率 |
|----------|-------|
| 異常なし | 2.6% |
| 経過観察 | 43.2% |
| がんの可能性あり | 20.3% |
| がんの疑いあり | 10.9% |
| がんの疑いが強い | 3.9% |



胸部CT画像

| 項目名 | 疾病確率 |
|----------|-------|
| 異常なし | 39.9% |
| 経過観察 | 40.2% |
| がんの可能性あり | 5.8% |
| がんの疑いあり | 8.5% |
| がんの疑いが強い | 6.2% |



医療診断支援システムの出力画面

今後の研究開発成果の展開及び波及効果創出への取り組み

- 深層学習やクラスタリング手法を開発し、医療診断支援システムを開発した。
- 広島県の複数の医療機関、島しょ部の検査機関等を開発したシステムを導入し、検査データの収集を行いながら継続的なシステムの運用を行い、健康寿命を増加させるための日常的な健康管理を行うことが必要である。
- システムを通じて患者がデータに基づいた健康管理を自ら行う重要性を認識できるようになれば、健康管理に必要な検診の受診につながることを期待できる。
- 専門医不足の問題が指摘されているにシステムを導入し、医師がシステムを利用することで、初期対応の迅速化が高まることを期待できる。
- 県立広島大学では、包括連携協定を締結している広島県環境保健協会と協働し、実用化を検討している。