



コンピュータ支援検出(CAD*)

* computer-assisted detection

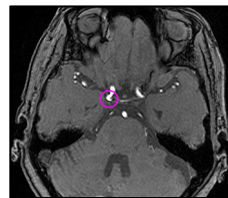
- コンピュータ上で医用画像を解析し、自動検出された病変の位置を提示
⇒ 医師の病変見落とし低減が目的



CAD処理結果例



胸部CT肺結節検出 (○: 検出病変)

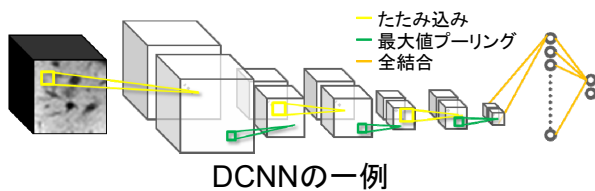


頭部MRA脳動脈瘤検出 (○: 検出病変)

CADへのDeep Learningの適用

Deep convolutional neural network (DCNN)

- 脳の視覚野における情報処理を模したもの
- 医師が画像診断に用いる有用な情報を適切に学習させることができれば、CADの高性能化が期待できる



DCNN使用上の課題

- ハイパーパラメータが多数
 - ネットワークの構成
 - 最適化のパラメータなど
- 多数症例による学習が望ましい
 - 数千~数万症例規模を想定 (データサイズ: 数百GB~数TB)

大規模データに対応し、かつ効率の良い学習が実行可能なHPC環境が望まれる

パラメータ自動探索モジュールの構築

ベイズ的最適化(BO**) [Mockus, 1974]

** Bayesian optimization

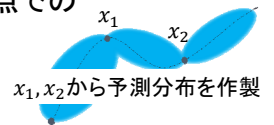
- ブラックボックス関数の最大化を解く方法

- 関数にGaussian Processを仮定

- 関数からサンプルした点が同時ガウス分布 $f(x_1), f(x_2) \dots f(x_n) \sim \mathcal{N}(\mu_X, \Sigma_X)$

- 観測点 $\{X, f(X)\}$ から任意の点での予測分布を生成

- 値の高くなりそうな点を次に観測



ハイパーパラメータチューニングへの利用

[Snoek+, 2012]

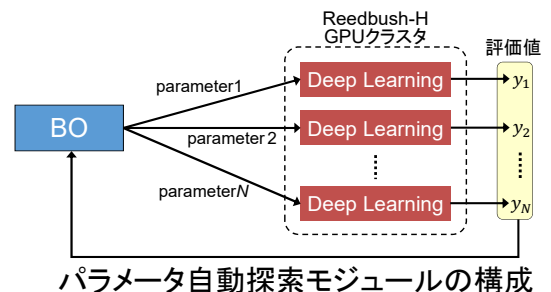
- 評価値が最大となるパラメータを探索 (逐次的)
 - 関数の値: CADの性能評価値
 - 引数: ハイパーパラメータ



GPUクラスタでのパラメータ探索実装に向けて 非同期並列実行可能なアルゴリズムを確立

Reedbush-H GPUクラスタへの実装

- GPUクラスタ (東大・Reedbush-H) の各ノードにBOで生成したパラメータを用いた学習ジョブを繰り返し投入するためのモジュール群を構築
 - 非同期並列実行型BO計算モジュール
 - ジョブ投入モジュール
 - 並列スクリプト言語 Xcrypt を使用
 - CAD性能評価モジュール
- ⇒ 他のGPUクラスタでも再利用可能なモジュール構築を目指す



文献

- Moćkus, J. "On Bayesian methods for seeking the extremum." *Optimization Techniques IFIP Technical Conference*. Springer Berlin Heidelberg, 1974.
- Snoek, Jasper, Hugo Larochelle, and Ryan P. Adams. "Practical Bayesian optimization of machine learning algorithms." *Advances in Neural Information Processing Systems*. 2012