

jh170036-DAH

Deep Learning を用いた医用画像診断支援に関する研究

佐藤 一誠（東京大学）

概要 本共同研究では、Deep Learning (DL) を用いた高性能なコンピュータ支援検出 (CAD) の開発を行うことを目的とする。まず、非同期並列実行型ベイズ的最適化に基づくパラメータ探索を用いた DL 学習方法を GPU クラスタ上にて構築した。構築した環境下で複数の CAD (頭部救急 CT 画像における疾患検出 (脳梗塞, 脳出血など), 肝臓造影 MR 画像を用いた結節性病変検出, 頭部造影 MR 画像を用いた転移性脳腫瘍検出など) の開発を行った。さらに、今後の学習の効率化に向けて、ChainerMN による複数ノード・複数 GPU を用いた DL 学習方法について検討した。

1. 共同研究に関する情報

(1) 共同研究を実施した拠点名

東京大学

(2) 共同研究分野

- 超大規模数値計算系応用分野
- 超大規模データ処理系応用分野
- 超大容量ネットワーク技術分野
- 超大規模情報システム関連研究分野

(3) 参加研究者の役割分担

佐藤 一誠 学習環境構築, CAD 開発
埴 敏博 学習環境構築, GPU 間通信,
ノード間通信
林 直人 臨床面での評価
花岡 昇平 CAD 開発, 臨床面での評価
野村 行弘 学習環境構築, CAD 開発
竹永 智美 学習環境構築, CAD 開発
村田 仁樹 CAD 開発
佐藤 大介 CAD 開発, 臨床面での評価
中尾 貴祐 CAD 開発, 臨床面での評価
大島 聡史 GPU 向け最適化, 自動チューニ
ング
星野 哲也 GPU 向け最適化, 自動チューニ
ング
関谷 勇司 ネットワーク運用技術
渡邊 大志 学習環境構築
亀澤 諒亮 学習環境構築, CAD 開発
磯 穰 学習環境構築, CAD 開発

2. 研究の目的と意義

CT や MRI をはじめとする医用画像を解析し、自動検出された病変の位置を提示する、コンピュータ支援検出 (computer-assisted detection, CAD) は画像診断における医師の見落とし低減を目的として古くから研究されている。しかし、日常的に臨床で使われている CAD は乳房 X 線写真に対する腫瘍検出、胸部 X 線写真もしくは胸部 CT における肺結節検出 (図 1) などに限られているのが現状である。その要因の 1 つとして、多くの CAD が病変検出性能を経験のある医師と同程度とした場合、医師に比べて拾いすぎ (偽陽性, false positive; FP) が多いことが挙げられる。このため、CAD のさらなる高性能化が必須である。

一方、CAD の主要技術である機械学習分野において、Deep Learning (DL) がこの数年注目を集めており、医用画像処理にも利用されている。DL の一手法である畳み込みニューラルネットワーク (convolutional neural network, CNN) は脳の視覚野

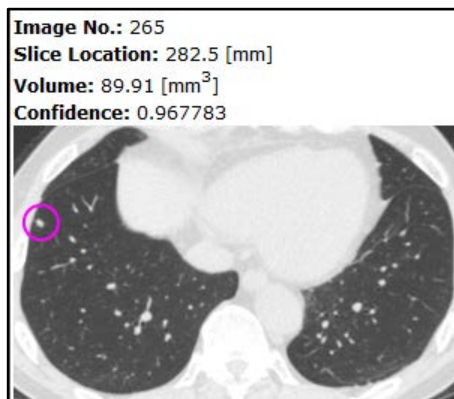


図 1 胸部 CT 肺結節自動検出例 (○: 検出された病変)

における情報処理を模したものであり、画像診断は視覚処理である。このため、医師が画像診断に用いる有用な情報を適切に CNN に学習させることができれば、CAD のさらなる高性能化の実現が期待できる。

本共同研究で DL を用いた高性能な多種の CAD が開発されることにより、臨床医学における画像診断の質的向上ならびに医師の負担軽減が期待される。CAD の開発には医師が診断した医用画像データを用いることから、データを作成した国の診断基準が CAD に反映される。そのため、国内で使用される CAD は国内で開発することが非常に重要であると考え、CAD の研究・開発ならびに臨床使用に関して多くの経験を持つ 4 名の放射線科医師（林、花岡、佐藤大介、中尾）、2 名の診療放射線技師（野村、竹永）、および 1 名の医学物理士（村田）が共同研究者として参画している。共同研究者間で密に連携を取ることによって、CAD の開発から臨床での評価までを促進させる体制を整えている。さらに、多くの臨床施設で高性能な CAD が使われるようになれば、本共同研究に賛同する臨床施設の参画による拠点連携研究へ進展することが期待される。

本共同研究で構築するパラメータ自動探索を用いた DL の学習フレームワークは本共同研究のみでなく、DL を用いた様々な研究の促進に寄与することが期待される。

3. 当拠点公募型共同研究として実施した意義

DL は画像認識をはじめとする複数の領域で高い性能を発揮しているものの、パラメータ調整が難しいことが問題である。本共同研究で DL 学習の高速化およびパラメータ自動探索を用いた DL 学習方法を確立することは DL を用いた CAD 開発を進める上で必須である。このフレームワークを構築するためには、CAD 開発に係わる各分野の専門家と情報基盤センターの緊密な協力が不可欠であり、当拠点公募型共同研究として取り組む必要がある。さらに、本研究成果を広く公開することによって、各基盤センターの GPU 搭載のスーパー

コンピュータにおいても同様のことが容易に実現できるようになると考える。

本共同研究は東大単独で実施しているが、過去の研究[A]より CAD 開発用の症例データを多施設連携で収集することの重要性を認識しており、将来的には複数の臨床施設の参画による拠点連携研究へ発展させることを視野に入れている。

[A] 野村行弘, 増谷佳孝, 三木聡一郎, 他, 遠隔読影環境における多施設連携型 CAD 開発, 実運用, および継続的性能改善, MEDICAL IMAGING TECHNOLOGY, vol.32, no.2, pp.98-108, 2014.

4. 前年度までに得られた研究成果の概要

該当無し

5. 今年度の研究成果の詳細

(1) パラメータ自動探索を用いた DL 学習方法の構築 (研究成果[2],[8],[10])

図 2 に構築した DL の学習フレームワークでの処理の流れを示す。GPU クラスタ (RB-H) のログインノード上でジョブ並列スクリプト言語 (Xcrypt) [B]で記述したジョブ投入モジュールを実行し、Bayesian optimization (BO) [C]により生成されたパラメータを用いた学習ジョブを RB-H の計算ノードに繰り返し投入する。

BO は入力から出力を得るコストが高いブラックボックス関数を効率的に最適化する手法、すなわちブラックボックス関数 f の定義域を \mathcal{X} として、

$$x^* = \operatorname{argmax}_{x \in \mathcal{X}} f(x) \quad (1)$$

を求める問題である。BO では一般的にブラック

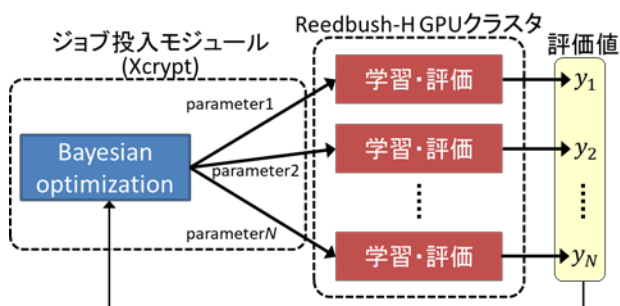


図 2 実装した DL の学習方法の流れ

表 1 構築したフレームワークの検証条件

	脳動脈瘤検出	転移性脳腫瘍
パラメータ数	9	17
学習症例数	300	214
評価症例数	50	67
partial AUC 計算範囲	~3FPs/case	~10FPs/case

ボックス関数の事前分布としてガウス過程を仮定し、観測が得られる毎に、ブラックボックス関数の事後分布を更新し、事後分布によって周辺化されたベイズ予測分布の情報を元に次に観測すべき点を決定する。DL の学習におけるパラメータ探索では、BO で生成されたパラメータを用いて学習した結果を検証用症例に対して適用した際の評価値が最大となるパラメータを探索する。構築したフレームワークは並列ジョブ数を 1 とした場合は通常の BO として動作し、2 以上の場合は非同期並列実行型 BO として動作する。また、探索開始時ならびに、処理が完了したジョブが 3 個に達するまではランダム値をパラメータとして与える。

構築したフレームワークを用いた CAD 学習について検証を行った。対象とした CAD は文献[D] の MRA 画像の脳動脈瘤検出、および(2-3)の造影 MR 画像の転移性脳腫瘍検出の 2 種類である。脳動脈瘤検出ではネットワーク構成を固定しており、転移性脳腫瘍検出では convolution 層の数などのネットワーク構成を決定するためのパラメータも探索対象とした。並列数は $N=\{1, 2, 4, 8\}$ の 4 種類とし、それぞれ 50 試行 (ジョブ) の学習を 3 回実施した。比較としてランダムサーチを 150 試行 (50 試行 \times 3 回) 実施した。評価には free-response receiver operating characteristic 曲線における partial area under the curve (partial AUC) を使用した。その他の条件は表 1 に示す。

図 3 および図 4 に各 CAD においてパラメータ探索を行った際の当該試行までの評価値の最大値推移 (3 回分の平均値) を示す。これらの結果より、構築したフレームワークを用いたパラメータ探索ではランダムサーチと同等以上の性能が得られていることが示された。

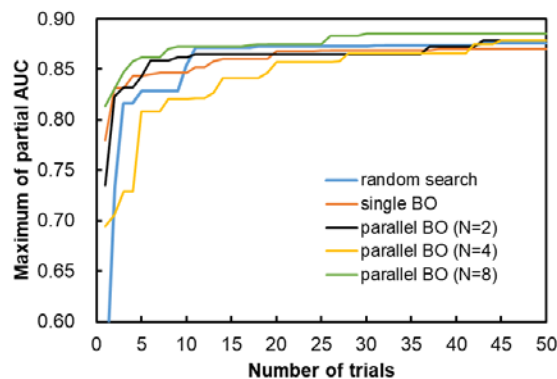


図 3 構築したフレームワークによるパラメータ探索結果 (脳動脈瘤検出)

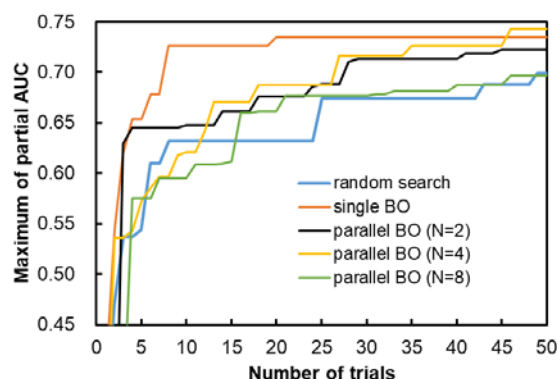


図 4 構築したフレームワークによるパラメータ探索結果 (転移性脳腫瘍検出)

(2) DL を用いた CAD の開発

(2-a) 4D-CNN を用いた EOB-MR 画像における肝結節病変検出 (研究成果[3],[6])

肝結節性病変の検出において EOB-MRI 検査は他の検査と比較し有意に検出感度が高い。しかしながら 5 時相の造影検査であるため、画像診断における医師の労力が多大である。本研究では、5 時相の 3 次元画像を入力とした 4D-CNN を用いて、ボクセルベースで肝結節性病変の初期検出を行った。使用症例は静磁場強度やメーカーが異なる 4 装置より収集した 165 症例であり、内訳は学習用 84 症例、validation 用 36 症例、テスト用 45 症例とした。テスト用症例での性能評価の結果、検出感度 50%における症例あたりの FP 数は 13.1 個であった。図 5 に 4D-CNN による肝結節病変検出結果の一例を示す。

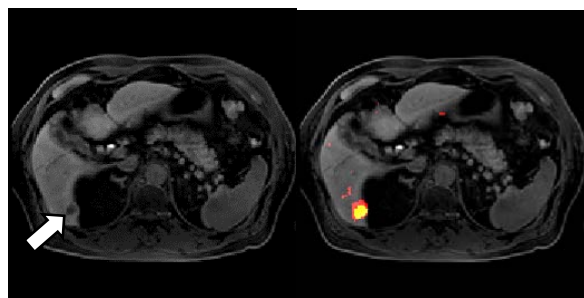


図 5: 4D-CNN を用いた肝結節性病変検出結果 (黄: 正しく検出できた領域, 赤: 誤検出領域, 矢印: 肝転移)

(2-b) AutoEncoder を用いた頭部救急 CT 画像の異常検知 (研究成果[1],[7])

頭部救急 CT の対象疾患は梗塞, 脳内出血, くも膜下出血, 硬膜下出血, 硬膜外出血, 腫瘍など多岐にわたり, さらに, 複数の疾患が同時に存在することもある. 本研究では, 頭部救急 CT 画像における 3D convolutional autoencoder (3D-CAE)[E]を用いた異常検知手法を構築した. 3D-CAE は正常症例から抽出した 3D パッチを用いて学習を行い, 評価時には 3D パッチごとの異常度として平均二乗誤差を用い, 症例ごとのパッチの異常度の最大値をその症例の異常度とした. 50 例で学習した 3D-CAE に対して 38 例の頭部救急 CT (異常 22 例および正常 16 例) を用いて評価した結果, 感度 68%, 特異度 88%, AUC は 0.87 に達し, 頭部救急 CT での異常検知において有用である可能性を示した. 図 6 は 3D-CAE によって生成された異常度マップの一例であり, 病変部分とマップの異常度の高い部分 (白い部分) が概ね一致している.

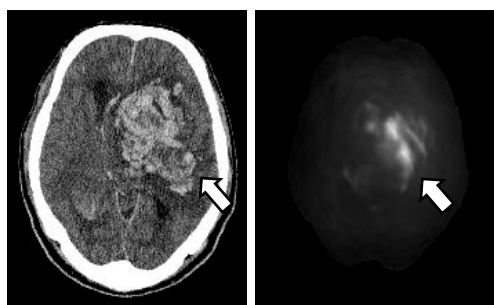


図 6: 3D-CAE によって生成された異常度マップ (左: 元画像, 右: 異常度マップ. 矢印: 脳内出血)

(2-c) 頭部造影 MR 画像の転移性脳腫瘍検出 (研究成果[11])

転移性脳腫瘍は, その存在の有無や個数によって治療方針が大きく変わるため, 正確な診断が必要となる. しかし, 微小な腫瘍, 脳血管と隣接する腫瘍, コントラストの低い腫瘍の検出は困難であり見落される可能性が高い. 本研究では, 3D-CNN を用いた頭部造影 MR 画像の転移性脳腫瘍検出手法を構築した. 使用症例はメーカーが異なる 3 装置より収集した 348 症例であり, 内訳は学習用 214 症例, validation 用 67 症例, テスト用 67 症例とした. convolution 層の数などのネットワーク構成を決定するためのパラメータを含めたパラメータ探索を行い, validation 用症例で最良の性能が得られた学習結果をテスト用症例で評価した結果, 検出感度 70%における症例あたりの FP 数は 8.4 個であった. 図 7 に 3D-CNN による転移性脳腫瘍検出結果の一例を示す.

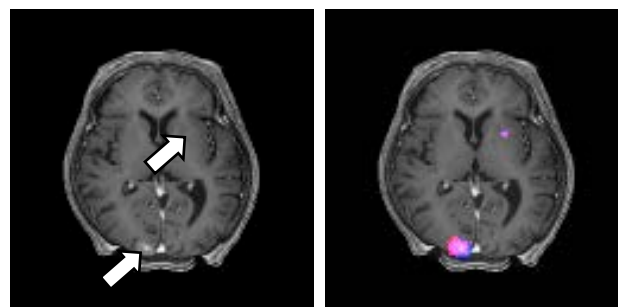


図 7: 3D-CNN を用いた転移性脳腫瘍検出結果 (マゼンダ: 正しく検出できた領域, 赤: 誤検出領域, 青, 未検出領域, 矢印: 転移性脳腫瘍)

(3) DL を用いた学習のさらなる効率化に向けて (研究成果[5],[9])

今後の学習の効率化に向けて, 複数ノード・複数 GPU を用いてデータ並列によって学習を高速化する方法について検討を行った. 具体的には, 分散 DL 学習ライブラリの ChainerMN[F]について RB-H を用いて評価を行った.

ChainerMN は, 複数ノードの GPU を用いてデータ並列に学習を行う, Chainer の追加パッケージである. ChainerMN を使用するためには, Chainer で用いる DL 向けライブラリ cuDNN や

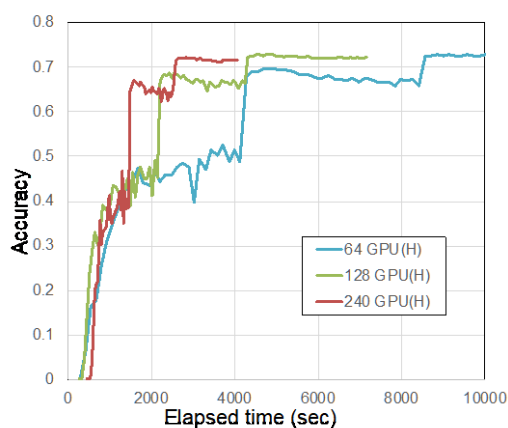


図 8: ImageNet における学習曲線
(64 GPU は 14300 秒まで)

numpy の GPU 版実装である cupy に加えて, GPU 間通信を最適化する NCCL (NVIDIA Collective Communication Library), 複数ノードでのプロセス起動やセットアップを行うための MPI ライブラリの連携が必要であり, これらを RB-H に導入して ChainerMN の実行環境を構築した. その上で, CNN の例として ImageNet の画像分類の training, recurrent neural network (RNN) の例として Seq2seq による機械翻訳の training を RB-H の全系まで用いて実施し性能評価を行った.

図 8 に, ImageNet の場合の各エポックにおける実行時間と training 結果の validation accuracy を示す. 100 エポックまで実行した結果, 64 GPU に対して 128 GPU では 1.77 倍, 240 GPU では 3.54 倍の性能が得られ, 良好な強スケーリング性を示した. 一方で training の結果得られた精度はいずれも約 72% であり, GPU 数が増えることによるバッチサイズの増加による training 結果への悪影響を防ぐことができている.

図 9 に, Seq2seq の場合の各エポックにおける実行時間と機械翻訳の評価指標である BLEU (BiLingual Evaluation Understudy) を示す. 64 ノード 128 GPU では想定した 15 エポックまでの学習を 13.6 時間で終えることができたが, 32 ノード 64 GPU の実行の場合, 12 エポックまでで時間制限 (この場合は 24 時間) により打ち切りとなった. いずれも BLEU は 23.5% に達しており妥当な結果である.

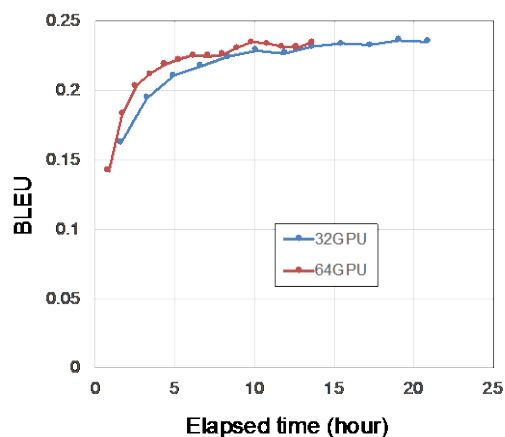


図 9: Seq2seq における学習曲線
(32 GPU は時間制限により途中で打ち切り)

なお, 本項の成果の一部は, 東京大学情報基盤センターによる「Reedbush-H 大規模 HPC チャレンジ」プロジェクトに採択され実施した内容を含む.

- [B] Hiraishi T, Abe T, Iwashita T, Nakashima H, Xcrypt: A Perl extension for job level parallel programming, Second International Workshop on High-performance Infrastructure for Scalable Tools WHIST 2012 (held as part of ICS'12), 2012.
- [C] Snoek J, Larochelle H, Adams RP, Practical Bayesian optimization of machine learning algorithms, Advances in Neural Information Processing Systems 2012;2951-2959.
- [D] Nakao T, Hanaoka S, Nomura Y, Sato I, et al., Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography, Journal of Magnetic Resonance Imaging, 2018;47:948-953.
- [E] Hosseini-Asl E, Keynton R, El-Baz A, Alzheimer's disease diagnostics by adaptation of 3d convolutional network, Proc. of ICIP 2016, no. MA.L5.4, pp.126-130, Arizona, USA, Sept. 2016.
- [F] Akiba T, Fukuda K, Suzuki S, ChainerMN: scalable distributed deep learning framework, arXiv: 1710.11351, 2017

6. 今年度の進捗状況と今後の展望

本共同研究全体では概ね順調に進捗した。パラメータ自動探索を用いた DL 学習方法は当初の計画通り Xcrypt を用いたジョブ投入モジュール, および非同期並列実行型 BO によるパラメータ生成モジュールを含めたフレームワークが構築できた。本共同研究以外のユーザへの提供については利便性向上を含めた改善が必要であると考えており, 次年度の課題実施と平行して進めていく予定である。

各 CAD の開発では, 3 種類の CAD において初期段階の数十～数百症例を用いたアルゴリズムの検討を進めると同時に, より大規模なデータ (数百～数千症例) を用いた学習に向けたデータの整備を進めてきた。今後, 大規模データを用いた学習を進めるにあたり, 小規模なデータセットに対して試行したパラメータ探索の結果を大規模なデータセット学習に反映させる非同期並列実行型 BO の構築を進めていく。

今年度は DL を用いた学習のさらなる効率を図るために ChainerMN に関する基礎的な検討を進めたものの, 実際の CAD の学習への適用までは至らなかった。解決策の 1 つとして, HPC 向けコンテナ (singularity[G]など) に ChainerMN 実行環境を構築し, 東大病院等にある初期開発環境でも容易に利用できるようにする予定である。

[G] <https://singularity.lbl.gov/index.html>

7. 研究成果リスト

(1) 学術論文

該当無し

(2) 国際会議プロシーディングス

[1] D. Sato, S. Hanaoka, Y. Nomura, T. Takenaga, S. Miki, T. Yoshikawa, N. Hayashi, O. Abe, A primitive study on unsupervised anomaly detection with an autoencoder in emergency head CT volumes, Proc. of SPIE, vol.10575, 105751P, Feb. 2018.

(3) 国際会議発表

[2] Y. Nomura, I. Sato, T. Hanawa, S. Hanaoka, T. Nakao, T. Takenaga, D. Sato, T. Hoshino, Y. Sekiya, S. Ohshima, N. Hayashi, O. Abe, Preliminary development of training environment for deep learning on supercomputer system, CARS 2018, June 2018 (accepted).

[3] T. Takenaga, S. Hanaoka, Y. Nomura, T. Nakao, M. Nemoto, M. Murata, S. Miki, T. Yoshikawa, N. Hayashi, O. Abe, A preliminary study of the computerized detection of nodular liver lesion in Gd-EOB-DTPA-enhanced magnetic resonance images with 4D CNN, CARS 2018, June 2018 (accepted).

(4) 国内会議発表

[4] 埜敏博, 星野哲也, 中島研吾, 大島聡史, 伊田明宏, GPU 搭載スーパーコンピュータ Reedbush-H の性能評価, 情報処理学会研究報告.HPC, 2017-HPC-159(9), pp.1-6, 2017 年 4 月

[5] 埜敏博, 野村行弘, 医用画像処理への HPC 環境利用に向けて, 第 1 回人工知能応用医用画像研究会 (シンポジウム), 2017 年 6 月.

[6] 竹永智美, 花岡昇平, 根本充貴, 野村行弘, 中尾貴祐, 三木聡一郎, 吉川 健啓, 林直人, 阿部修, 3D-DCNN を用いた EOB-MR 画像における肝結節病変自動検出法の開発, 第 36 回日本医用画像工学会抄録, OP4-3, 2017 年 7 月.

[7] 佐藤大介, 花岡昇平, 野村行弘, 竹永智美, 三木聡一郎, 吉川健啓, 林直人, 阿部修, AutoEncoder を用いた頭部救急 CT 画像の異常検知に関する初期的検討, 電子情報通信学会技術研究報告, vol.117, no.220, MI2017-46, pp.25-26, 2017 年 9 月.

[8] 野村行弘, 佐藤一誠, 埜敏博, 花岡昇平, 中尾貴祐, 竹永智美, 佐藤大介, 星野哲也, 関谷勇司, 大島聡史, 林直人, 阿部修, スーパーコンピュータ上での Deep Learning 学習環境の初

期構築, 電子情報通信学会技術研究報告,
vol.117, no. 281, MI2017-47, pp.1-2, 2017 年 11
月.

- [9] 埜敏博, Reedbush システムを用いたディープ
ラーニング, GTC Japan, 2017 年 12 月.
- [10] 野村行弘, 佐藤一誠, 埜敏博, 花岡昇平, 中
尾貴祐, 村田仁樹, 竹永智美, 佐藤大介, 星野
哲也, 関谷勇司, 林直人, 阿部修, 非同期並列
実行型ベイズ的最適化を用いた Deep Learning
学習環境の構築, 第 2 回人工知能応用医用画
像研究会, 2018 年 6 月. (発表予定) .
- [11] 村田仁樹, 花岡昇平, 野村行弘, 竹永智美,
中尾貴祐, 高橋渉, 名和要武, 吉川健啓, 林直
人, 阿部修, MRI における脳転移検出に対す
る Convolutional Neural Network の応用: 線量
分布を正解データとして利用, 第 37 回日本医
用画像工学会, 2018 年 7 月 (発表予定) .
- (5) その他 (特許, プレス発表, 著書等)
該当無し