

—— 研究褒賞受賞記念報告 ——

てんかん診断，発作検知を可能にする人工知能の開発

柳 澤 琢 史¹⁻³⁾，貴 島 晴 彦^{1, 2)}

要旨：脳波や脳磁図はてんかん診療に欠かせない検査であるが，その判読には専門的な知識と経験を要するため，幅広い利用の妨げとなっている。脳波脳磁図の自動診断は，てんかん診療の均てん化や質の向上に寄与すると期待される。特に近年，脳波・脳磁図のビッグデータに深層学習を用いた自動診断が高い精度を達成することが報告されている。そこで我々は，脳波・脳磁図用の深層学習モデルを作成し，当院で収集した脳磁図のビッグデータに適用し，これまでの脳波特徴を用いた場合よりも高い精度でてんかんの診断ができることを示した。さらに，深層学習が抽出した脳波特徴を同定する方法を開発し，発作検知をする深層学習モデルに適用した。その結果，これまで，経験的にわかっていた脳波特徴を高周波数帯域に拡張する形で新しい脳波特徴を得ることに成功した。てんかんの診断や発作検知を可能にする人工知能技術の開発と成果について報告する。

てんかん治療研究振興財団 研究年報 2022；33：11-18

Key Words：深層学習，自動診断，発作検知，脳波特徴，Integrated gradient

序論

てんかん専門医は脳波や脳磁図の波形を判読し，てんかん診断や発作同定に利用している。波形の判読には多くの知識と経験が必要であり，脳波・脳磁図の普及を妨げる要因ともなっている。一方で，大量のデータと深層学習モデルを使うことで，さまざまな自動診断が実現されるようになり，脳波・脳磁図についても自動診断が期待されている。深層学習などの機械学習は，多次元の情報をそのままの形で処理することができるため，人が認識しにくい脳波特徴についても，より高精度に認識できると期待される。AIを用いることで，てんかん診断の効率化や診断精度の改善が期待される。また，AIの精度はてんかんの治療にも影響する。AIによって脳波からてんかん発作を自動で検知・予測す

ることで，患者に警告したり，自動で助けを呼ぶことで，発作に伴う危険を軽減できる¹⁾。また，発作を検知した際に，脳を電気刺激して発作を抑制するResponsive Neurostimulation (RNS)²⁾は，難治性てんかんに対して，米国等で治療に利用されている。RNSは頭蓋内電極で計測した信号を体内に埋め込まれたデバイス内のAIで判別し，脳を電気刺激するため，AIの精度は治療の精度に直結する。これまでの報告では治療成績も良く，長期的な効果も期待されている³⁾。つまり，AIによるてんかん脳波の自動解析は，治療にも貢献する。

一方で，医師が脳波・脳磁図から診断を行う場合は，いくつかの脳波特徴に基づいて判読を行っていく。つまり，高次元の脳波データから低次元の特徴量を抽出する作業が不可欠となる。これは，医師が診断の根拠を示す上でも重

¹⁾ 大阪大学大学院医学系研究科 脳神経外科
〔〒565-0871 大阪府吹田市山田丘2-2〕

²⁾ 大阪大学医学部附属病院 てんかんセンター
〔〒565-0871 大阪府吹田市山田丘2-2〕

³⁾ 大阪大学 高等共創研究院
〔〒565-0871 大阪府吹田市山田丘2-2〕

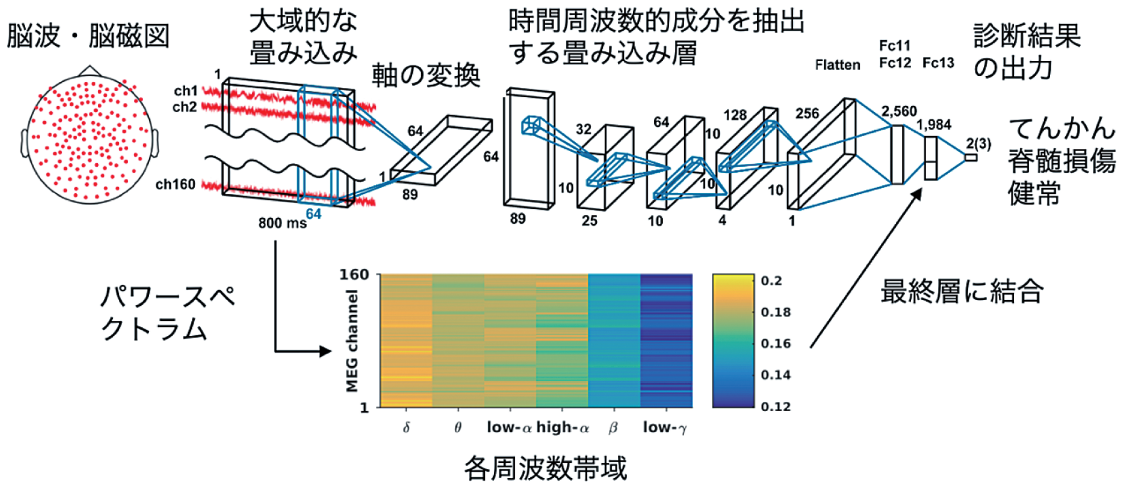


Fig. 1 脳波・脳磁図判読用のCNN (MNet) の構造

入力層へはチャンネル数分の脳波・脳磁図信号を入力した。畳み込みとプーリングの層を重ねていくが、最初の3層は時間方向の特徴量を抽出する形となる。これらを並べ直し、power spectrumを作るような形で層を作成することで、波形データから特徴量を自動抽出する。

要な作業である。実際、脳波・脳磁図の特徴量として様々な指標が提案されてきた。脳波に含まれる各周波数帯域パワー^{4, 5)}、異なる周波数間の同期的関係性 (cross-frequency coupling)⁶⁾、脳領域間での情報伝達ネットワーク (機能的結合解析およびグラフ理論)⁷⁻¹¹⁾、エントロピー¹²⁾などが挙げられる。例えば、てんかんの診断においては、パワー、Phase-locking value (PLV)、エントロピーを用いることで、90.5%の精度でてんかみを識別できたとの報告がある⁵⁾。ただし、てんかん患者は28名、健常者も14名と小規模な検討であり、その汎化性については疑問が残る。また、同様な小規模の検討はこれまでも多く報告されてきた¹³⁾。しかし、既に見つかっている特徴量が全てである保証はない。AIが多次元情報から抽出する特徴の中には、人が未だ認識していない特徴量があるかもしれない。AIによる自動診断の精度を高めるとともに、AIが“発見”した特徴量を明らかにし、人も利用できる形にすることが、てんかん治療の質を改善するために不可欠であると考え。我々は脳波・脳磁図を自動診断するための深層学習モデルを作成し、高い精度でてんかんの診断や発作検知ができることを示した。さらに、学習された深層学習モデルから脳波特徴を同定する方法

を開発し、人が使いやすい脳波特徴を作成することに成功した。

方法と結果

深層学習を用いた識別

近年、深層学習¹⁴⁾を用いた様々な自動診断が報告されている。深層学習は、機械学習の一種で、非線形な応答をする素子を層状に配置し、多層のネットワークを形成したアルゴリズムである。特に、畳み込み層とプーリング層と呼ばれるフィルター層を持つConvolutional neural network (CNN) と呼ばれるネットワークが、画像識別などに高い成績を示すことから注目された。2012年に、HintonらがCNNを使い、画像認識の精度を圧倒的に改善したことで注目を集めた¹⁵⁾。既に、脳波などの波形信号に対してCNNを用いた自動識別が報告されている¹⁶⁾。当初は比較的小規模なデータでも学習が成立する小規模なネットワークを用いた識別が主流であり¹⁷⁾、識別精度についても改善の余地があると考えられていた¹⁸⁾。しかし、近年は大規模なデータセットを用いて実用的な脳波識別を行うAIが報告されるようになった。例えばSpikeNetは9571例の頭皮脳波を用いて、発作間欠期脳波を学習することで、これまでの方法よりも有意に

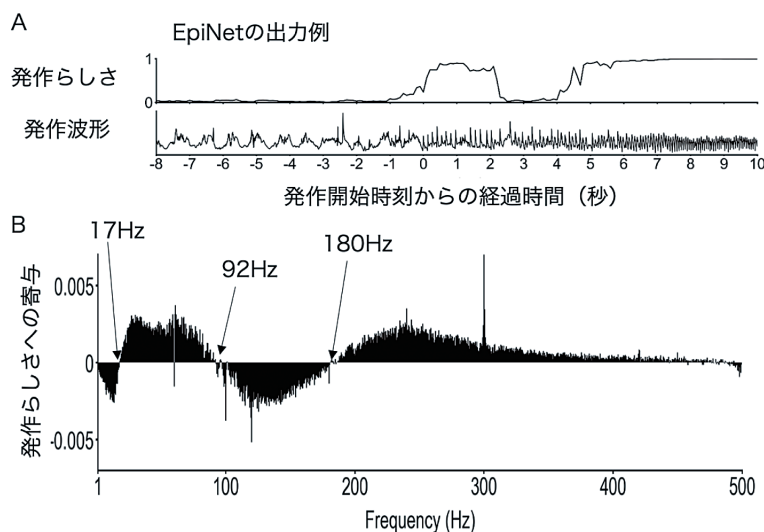


Fig. 2 EpiNetによる発作検知と特徴抽出

A. 発作時脳波（下段）にEpiNetを適用し、その出力である発作らしさをプロットした（上段）。発作開始時刻に相当する時刻0秒に一致して、発作らしさが上昇している。

B. Modified Integrated gradient法を持ちいてEpiNetが学習した発作らしさへの寄与度を周波数ごとに評価した。

高い精度で発作間欠期てんかん性放電を同定できることを示した¹⁹⁾。また、発作検知についても、大量のデータを公開され、様々なアルゴリズムが試されることで、高い精度で、てんかん発作を検知できることが示されている^{20, 21)}。

CNNを用いた波形識別は大きく2つに分類される。1つは、波形をそのままの形で入力するタイプ²²⁾、もう一つは波形をパワースペクトラムなどの特徴量に変換して識別するタイプである²³⁾。入力を周波数特徴とすることで、識別に寄与した周波数特徴を同定でき、説明可能性を高めることができる一方で、位相情報などに含まれる情報が使えない可能性がある。我々は、これらの特徴を組み合わせる形で、波形を入力として、周波数特徴に相当する特徴を抽出し、その後、画像として識別を行うモデル(MNet)を作成した(Fig. 1)²⁴⁾。

MNetは、波形信号自体を入力とし、全チャンネルにわたる時間方向の畳み込みを行い、大域的な信号特徴に対してパワースペクトラムを得るような形で設計されている。モデルの構築においては、東京大学原田研究室との共同研究のものと、環境音の自動判別のために作成された

Tokozumeらのネットワークを改変して作成した²⁵⁾。これに大阪大学で計測したてんかん(140例)、脊髓損傷(26例)及び健常者(67例)の安静時脳磁図を適用した。各症例から220秒の信号を取得し、800ms毎に切り出し入力とした。また、最終層に各電極での6つの周波数帯域(δ , θ , low- α , high- α , β , low- γ)のパワーを入力した。対照として、同じパワーを入力としてサポートベクトルマシーン(SVM)により識別した場合と比較した。その結果、3群の識別精度はMNetが70.7%で、SVMの63.4%より有意に高い精度であった。特にてんかんの鑑別は、感度 87.9%、特異度 86.0%であった。MNetが神経疾患の診断に有用であることが示唆された。

MNetは多チャンネルの時系列信号を一つの畳み込み層でまとめることで、大域的な周波数特徴を抽出することを目指して設計された。このネットワークは、脳全体としての活動状態からてんかんを識別することに寄与したと考えられる。一方、チャンネルごとの解析が重要な場合もある。難治性てんかんの焦点切除術において、てんかん源性域を同定することは、手術成績を

左右する重要な診断である。通常は、頭蓋内電極を留置し、発作時脳波を解析し、発作時脳波を最も早く同定できる電極を探索する。我々は、最も早く発作波形を呈した電極の波形に注目し、発作を検知する深層学習モデルを作成した。この場合、入力1次元の波形信号であるため、1次元CNN (EpiNet) を作成し、発作時と非発作時の脳波を識別した²⁶⁾。また、比較対象として、脳波のパワーやphase-amplitude couplingを特徴量として識別するSVMを用いた場合と比較した。その結果、EpiNetはSVMと比較して有意に高い精度で、発作と非発作時の脳波を識別した (AUCにて、EpiNet, 0.944 ± 0.067 ; SVM, 0.808 ± 0.253 , $p < 0.05$)。実際、発作時の脳波に適用すると、発作時刻に一致して、EpiNetの出力である発作らしさ (0から1の値をとり、1が最も発作らしい) は上昇することが示された (Fig. 2)。

深層学習が学習した特徴量の同定

EpiNetにより、これまで同定されている脳波特徴を用いた場合よりも高い精度で、発作時と非発作時を識別できることが示された。これは、深層学習が、脳波・脳磁図の特徴をデータから学習することで、これまでの特徴量以上に発作を特徴づける特徴量を同定したことを示唆する。そこで我々は、Integrated gradient法を改良したmodified integrated gradient法により、発作時脳波を発作として認識するために、EpiNetがどの周波数帯に注目したかを定量的に評価した。その結果、17Hz以下および92-180Hzのパワーの増加は発作らしくない特徴に対応し、17-92Hzおよび180Hz以上のパワーの増加は発作らしい特徴に対応することが示された (Fig. 2)。これまでも高周波数帯域の増加は発作らしさを表すことが知られていたが、92-180Hzは正常な脳活動に伴うhigh- γ 帯域の活動とも重なり、てんかん発作時の脳波としては特徴的ではないことが示された。

これまでに、経験則として知られていたEpileptogenicity ratio (ER) は、 $((\beta \text{ パワー}) + (\gamma \text{ パワー})) / ((\theta \text{ パワー}) + (\alpha \text{ パワー}))$ で表され、ERが高い状態は発作を示唆していた。本

研究でも17Hz以下の θ と α 帯域のパワー増加は発作らしさを減らすことに寄与し、17-92Hzに含まれる β と γ 帯域のパワー増加は発作らしさを増すことに寄与しているため、矛盾のない結果といえる。そこで我々は、経験的に導かれたERを深層学習の学習結果から得た特徴で拡張し新たな特徴量を作成した。つまり、各周波数帯パワーに発作らしさへの寄与度を累乗し、掛け合わせたものをdata-driven epileptogenicity index (d-EI) とした。このd-EIを用いて発作時と非発作時の識別を行ったところ、各周波数帯パワーやPACなどよりも有意に高い精度で識別ができることが示された。さらに、発作前後でのd-EIの変化量を、留置された電極全てに求め、焦点切除術の予後との関係を検証したところ、発作時にd-EIが増加する電極を切除範囲に含めた場合は、含めない場合よりも焦点切除術の予後が良いことがわかった。これらから、d-EIが患者によらない発作時脳波の特徴を捉えていることが示唆された。

考察

これまでの機械学習では、人が特徴量を設計し、その良し悪しが識別の精度を左右していた。一方、深層学習では、ネットワークがデータから特徴量を学習するため、人が気づかない特徴についても抽出され、高い精度の識別ができると考えられている。例えば、大量の心電図信号を深層学習で学習することで、専門医が見ても異常が指摘できない心電図から、1年以内の死亡の可能性を高い精度で推定することに成功している²⁷⁾。また、病理画像から予後を予測するように深層学習モデルを学習させることで、学習されたモデルから新しい病理分類を提案した研究も報告されている²⁸⁾。今後、脳波・脳磁図に対しても、その特徴をデータから学習し識別精度を高めることで、予後や症状をより良く説明するための、新しい特徴量の発見が期待される。ただし、深層学習を臨床応用する際には、学習用のデータをどの様に準備するか、どの様に汎化性能を担保するか、ネットワークの出す診断に対する説明可能性を確保するかなどに注意する必要がある。

一般に、深層学習モデルを学習するには大量のデータが必要となる。例えば、様々な皮膚病変画像の学習では約13万枚の画像を学習して皮膚科専門医と遜色のない診断精度を持つAIが開発された²⁹⁾。大量のデータを必要とする理由はモデルが大量のパラメータを持っているため、データが少ないと、AIが学習データに過学習してしまい、新たなデータを正しく識別できなくなるためである³⁰⁾。一方、パラメータ数を減らして小さいモデルを使うと、精度が低下する傾向がある¹⁸⁾。大量で高品質なビッグデータを使い、大規模なネットワークで学習することで、高精度な推定が期待される。今後、多施設共同で多疾患の脳波・脳磁図データが収集され共有されることで、高精度なAIによるてんかん診療の改善が期待される。

AIは高精度な診断に寄与すると期待されるが、AIはあくまで医師の診断や診療を補助する技術として今後の発展が期待される。AIの判断はあくまで参考的な所見であるが、見落としを予防し、医師に新たな発見を促すことで、診療の質を向上できると期待される。今後、高齢化社会に向かって、てんかんや認知症など機能的診断が不可欠な疾患が増加していく。そのような状況に置いて、神経機能診断がどこでも誰でも受けられるシステムの構築は日本の医療においても重要な課題である。日本が誇る均質で高品質な医療を背景としたビッグデータを形成し、人工知能技術を開発することで、てんかんでなく神経疾患診療の質が向上すると期待される。

文献

- 1) Cook, M.J., O'Brien, T.J., Berkovic, S.F., Murphy, M., Morokoff, A., Fabinyi, G., D'Souza, W., Yerra, R., Archer, J., Litewka, L., et al. (2013). Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy : a first-in-man study. *Lancet Neurol* **12**, 563-571.
- 2) Sun, F.T., and Morrell, M.J. (2014). The RNS System : responsive cortical stimulation for the treatment of refractory partial epilepsy. *Expert Rev Med Devices* **11**, 563-572.
- 3) Nair, D.R., Laxer, K.D., Weber, P.B., Murro, A.M., Park, Y.D., Barkley, G.L., Smith, B.J., Gwinn, R.P., Doherty, M.J., Noe, K.H., et al. (2020). Nine-year prospective efficacy and safety of brain-responsive neurostimulation for focal epilepsy. *Neurology* **95**, e1244-e1256.
- 4) Kerr, W.T., Anderson, A., Lau, E.P., Cho, A.Y., Xia, H., Bramen, J., Douglas, P.K., Braun, E.S., Stern, J.M., and Cohen, M.S. (2012). Automated diagnosis of epilepsy using EEG power spectrum. *Epilepsia* **53**, e189-192.
- 5) Soriano, M.C., Niso, G., Clements, J., Ortin, S., Carrasco, S., Gudin, M., Mirasso, C.R., and Pereda, E. (2017). Automated Detection of Epileptic Biomarkers in Resting-State Interictal MEG Data. *Front Neuroinform* **11**, 43.
- 6) Edakawa, K., Yanagisawa, T., Kishima, H., Fukuma, R., Oshino, S., Khoo, H.M., Kobayashi, M., Tanaka, M., and Yoshimine, T. (2016). Detection of Epileptic Seizures Using Phase-Amplitude Coupling in Intracranial Electroencephalography. *Sci Rep* **6**, 25422.
- 7) Nissen, I.A., Stam, C.J., Reijneveld, J.C., van Straaten, I.E., Hendriks, E.J., Baayen, J.C., De Witt Hamer, P.C., Idema, S., and Hillebrand, A. (2017). Identifying the epileptogenic zone in interictal resting-state MEG source-space networks. *Epilepsia* **58**, 137-148.
- 8) Douw, L., de Groot, M., van Dellen, E., Heimans, J.J., Ronner, H.E., Stam, C.J., and Reijneveld, J.C. (2010). 'Functional connectivity' is a sensitive predictor of epilepsy diagnosis after the first seizure. *PLoS One* **5**, e10839.
- 9) Verhoeven, T., Coito, A., Plomp, G., Thomschewski, A., Pittau, F., Trinka, E., Wiest, R., Schaller, K., Michel, C., Seeck, M., et al. (2018). Automated diagnosis of temporal lobe epilepsy in the absence of interictal spikes. *Neuroimage Clin* **17**, 10-15.
- 10) Lopes, M.A., Krzeminski, D., Hamandi, K., Singh, K.D., Masuda, N., Terry, J.R., and Zhang, J. (2021). A computational biomarker of juvenile myoclonic epilepsy from resting-state MEG.

- Clin Neurophysiol **132**, 922-927.
- 11) van Mierlo, P., Holler, Y., Focke, N.K., and Vulliemoz, S. (2019). Network Perspectives on Epilepsy Using EEG/MEG Source Connectivity. *Frontiers in neurology* **10**, 721.
 - 12) Kannathal, N., Choo, M.L., Acharya, U.R., and Sadasivan, P.K. (2005). Entropies for detection of epilepsy in EEG. *Comput Methods Programs Biomed* **80**, 187-194.
 - 13) Bao, F.S., Gao, J.M., Hu, J., Lie, D.Y., Zhang, Y., and Oommen, K.J. (2009). Automated epilepsy diagnosis using interictal scalp EEG. *Annu Int Conf IEEE Eng Med Biol Soc* **2009**, 6603-6607.
 - 14) LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* **521**, 436-444.
 - 15) Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Commun Acn* **60**, 84-90.
 - 16) De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., Visentin, D., et al. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* **24**, 1342-1350.
 - 17) Lawhern, V.J., Solon, A.J., Waytowich, N.R., Gordon, S.M., Hung, C.P., and Lance, B.J. (2018). EEGNet : a compact convolutional neural network for EEG-based brain-computer interfaces. *J Neural Eng* **15**, 056013.
 - 18) Alfredo Canziani, A.P., Eugenio Culurciello (2017). An Analysis of Deep Neural Network Models for Practical Applications. *arXiv* **1605.07678**.
 - 19) Jing, J., Sun, H., Kim, J.A., Herlopian, A., Karakis, I., Ng, M., Halford, J.J., Maus, D., Chan, F., Dolatshahi, M., et al. (2020). Development of Expert-Level Automated Detection of Epileptiform Discharges During Electroencephalogram Interpretation. *JAMA Neurol* **77**, 103-108.
 - 20) Kuhlmann, L., Karoly, P., Freestone, D.R., Brinkmann, B.H., Temko, A., Barachant, A., Li, F., Titericz, G., Jr., Lang, B.W., Lavery, D., et al. (2018). Epilepsycosystem.org : crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain* **141**, 2619-2630.
 - 21) Baldassano, S.N., Brinkmann, B.H., Ung, H., Blevins, T., Conrad, E.C., Leyde, K., Cook, M.J., Khambhati, A.N., Wagenaar, J.B., Worrell, G.A., et al. (2017). Crowdsourcing seizure detection : algorithm development and validation on human implanted device recordings. *Brain* **140**, 1680-1691.
 - 22) Emami, A., Kunii, N., Matsuo, T., Shinozaki, T., Kawai, K., and Takahashi, H. (2019). Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images. *Neuroimage Clin* **22**, 101684.
 - 23) Frey, M., Tanni, S., Perrodin, C., O'Leary, A., Nau, M., Kelly, J., Banino, A., Doeller, C.F., and Barry, C. (2019). Deepinsight : a general framework for interpreting wide-band neural activity. *bioRxiv*.
 - 24) Aoe, J., Fukuma, R., Yanagisawa, T., Harada, T., Tanaka, M., Kobayashi, M., Inoue, Y., Yamamoto, S., Ohnishi, Y., and Kishima, H. (2019). Automatic diagnosis of neurological diseases using MEG signals with a deep neural network. *Sci Rep* **9**, 5057.
 - 25) Tokozume, Y., and Harada, T. (2017). Learning Environmental Sounds with End-to-End Convolutional Neural Network. *Int Conf Acoust Spee*, 2721-2725.
 - 26) Yamamoto, S., Yanagisawa, T., Fukuma, R., Oshino, S., Tani, N., Khoo, H.M., Edakawa, K., Kobayashi, M., Tanaka, M., Fujita, Y., et al. (2021). Data-driven electrophysiological feature based on deep learning to detect epileptic seizures. *J Neural Eng* **18**.
 - 27) Raghunath, S., Ulloa Cerna, A.E., Jing, L., vanMaanen, D.P., Stough, J., Hartzel, D.N., Leader, J.B., Kirchner, H.L., Stumpe, M.C., Hafez, A., et al. (2020). Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med* **26**, 886-891.
 - 28) Courtiol, P., Maussion, C., Moarii, M., Pronier, E., Pilcer, S., Sefta, M., Manceron, P., Toldo, S., Zaslavskiy, M., Le Stang, N., et al. (2019). Deep

- learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med* **25**, 1519-1525.
- 29) Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118.
- 30) Bishop, C.M. (2006). *Pattern recognition and machine learning*. (New York : Springer).

