深層学習を用いた樹幹からの打撃音に基づく樹高および材積の推定

美濃 羽 靖^{*・1}・和田 誠²・田中 紡²

本研究では、立木を叩いた際に発生した音を画像化し深層学習を用いて樹高、材積を推定した。立木 20 本の樹幹を 1 本につき 100 回打撃した際に発生した音を録音、0.6 秒間における各周波数の音圧を表したスペクトログラムを 10,000 枚作成し入力画像とした。深層学習システムは NNC を、深層学習アルゴリズムは出力層を回帰層とした LeNet を用いた。学習用データを 5 セットに分割し、三つの学習パターン (LP-I: 訓練事例 8 割, 未知事例 2 割, LP-II: 大中小三区分から 1 本ずつ抽出した木を未知事例, LP-III: 2 本ずつ抽出した木を未知事例) の樹高、材積を推定した。推定精度の検証には平均絶対誤差, 平均絶対パーセント誤差および決定係数を用いた。その結果, 各学習パターンの未知事例に対する R^2 値は, LP-III の樹高 (0.3672) を除き, 非常に高い値 (0.9192 から 0.9996) を示した。LP-III の樹高では, 30 m 以上が過小に, 30 m 以下が過大に推定される傾向を示した。一方, 材積はどの学習パターンにおいても全体的に偏りのない推定を行うことができたことから, 本手法は材積推定において有効であることが示唆された。

キーワード: 音の画像化, 深層学習, スペクトログラム, 打撃音, Neural Network Console (NNC)

Yasushi Minowa,^{*・1} Makoto Wada,² Tsumugi Tanaka² (2021) Estimation of Tree Height and Volume Based on the Produced Hammering Sound Using Deep Learning. J Jpn For Soc 103: 351–360 This study aimed to estimate the tree height and volume based on image data, which are obtained by conversion of the sound generated from hammering a stem using deep learning. We hammered 20 trees 100 times, recorded the hammering sound, and generated the spectrogram, which presented the sound pressure at each frequency for 0.6 s. Data comprising 10,000 images were loaded into a deep learning model. We used the Neural Network Console (NNC) as the deep learning system and the LeNet, which forms a programmed regression layer to an output layer, as the deep learning algorithm. We divided 10,000 images into 5 equal sets, and performed three learning patterns (LP-I, LP-II, and LP-III). LP-I used the four sets as training data and the remainder as test data, LP-II used three trees and LP-III used six trees, which select for each tree from three divisions (large, medium, small), as test data. A performance evaluation of the proposed model was performed using three indicators: mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). Each learning pattern provided very good estimates (R^2 values for each learning pattern for the test data ranged from 0.9192 to 0.9996), except for the height estimate by LP-III ($R^2 = 0.3672$). LP-III generated very poor height estimates with a bias and tended to underestimate by >30 m and overestimate by <30 m. However, each learning pattern provided a good estimate of the tree volume, generally without any bias. Thus, we found this method to be more effective for estimation of tree volume than tree height.

Key words: deep learning, hammering sound, Neural Network Console (NNC), sound-to-image conversion, spectrogram

I. は じ め に

森林を適切に管理する上で森林の現状を把握することは重要であり, その方法の一つとして林分調査があげられる。林分調査時には一般的に単木測定が行われ, 直径や樹高を計測し材積表等を用いて単木材積を算出する。直径の計測は輪尺や直径巻尺などを用いるため比較的容易に計測が可能であるが, 樹高の計測では, Vertex (Haggröf website) などの測高器を用いるため, 直径に比べ計測に時間がかかること, 測高器が比較的高価であること, といった問題があげられる。また, 近年では ALS (Airborne laser scanner) や TLS (Terrestrial laser scanner) を活用した単木あるいは林分の樹高や材積を推定する方法も開発されつつあるが, 現状ではレーザ計測の費用が比較的高価であり, また点群データの取り扱いや解析に技術を要するため, 利用のハードルは高い (林野庁 2018)。そこで, 簡便かつ安価に単木測定を行うことができれば, 実務レベルおよび研究レベルでの

森林調査にとって有益であると考えられる。また, 誰もが簡便に単木測定を行うことができれば, 森林・環境教育を含む様々な関心レベルを持つ人々に役立つと期待される。

そこで, 本研究では, 単木測定における簡便かつ安価な測定方法として, ハンマーなどを用いて樹幹を打撃し, その際に発生した打撃音を基に樹高や材積といった量的な立木情報の推定を試みた。そのためには, 樹幹を打撃した際に発生した音から樹高や材積を高い精度で推定する方法が必要となる。音情報を量的情報と関連付けた研究事例としては, 例えば, ヤング率の測定に用いられる縦振動法 (打撃音試験) を用いた丸太重量の測定 (小泉ら 1998), コンクリートの打音データに自己組織化マップを適用してコンクリートの欠陥の深さ方向をパターン分類した研究 (園田ら 2011) やコンクリートの打音データの音圧の時刻歴波形や周波数特性を用いてコンクリート内部の欠陥の有無や深さとの関係を解析した研究 (小池ら 2015), などがあげられる。ただし, 音情報を用いた研究では, 音の特性 (例

*連絡先著者 (Corresponding author) E-mail: sharmy@uf.kpu.ac.jp <https://orcid.org/0000-0001-7487-5692>

¹ 京都府立大学大学院生命環境科学研究科 〒606-8522 京都市左京区下鴨半木町 1-5 (Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan)

² 京都府立大学生命環境学部 〒606-8522 京都市左京区下鴨半木町 1-5 (Faculty of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan)
(2021 年 5 月 1 日受付; 2021 年 8 月 27 日受理)

©2021 一般社団法人日本森林学会: この著作はクリエイティブ・コモンズのライセンス CC BY-NC-ND (引用を表示し, 改変せず, 非営利目的に限定) の条件の下で再配布・二次利用が可能なオープンアクセスです。 <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.ja>

えば、音圧や周波数特性など）から最大周波数などの特定の特徴量を抽出し、回帰や平均値などの統計量を用いた統計解析、あるいはSVM（Support vector machine）（Vapnik and Lerner 1963；山崎 2010）や決定木（Quinlan 1993；金 2005）といった機械学習を用いた解析、などが多い（松井 2018）。しかし、解析に用いる特徴量は対象によって様々であるため簡単に特定できない。それゆえ、特徴量の抽出では、解析を実行する研究者・開発者らの経験等に大きく作用されることが多い（山下 2016）。これに対し、近年、音情報を画像化し深層学習を適用して分類・推定を行う研究が進んでいる。深層学習とは、1943年にMcCullochとPittsが形式ニューロンモデルを発表して以来、様々に発展してきた人工知能およびその応用に関する研究の中で、2000年以降、コンピュータ技術の発展とともにその成果が注目を浴びている技術である（山下 2016）。特に、1998年に発表されたLeCunらによる畳み込みニューラルネットワーク（Convolutional neural network, 以下、CNN）は、画像認識分野において威力を発揮している（LeCun *et al.* 1998；岡谷 2015；山下 2016）。深層学習は、主に画像認識分野における研究例、応用事例が数多くあるが、音声認識、音イベント検出、音響信号処理といった音情報に対しても深層学習を適用した研究が行われている（Dai *et al.* 2017；磯 2017；Lee *et al.* 2017；篠田 2017；井本 2018, 2019；坂東 2020）。音情報に深層学習を適用する際は、一次元情報である音を二次元情報の画像に変換するが、その際、本研究でも用いたスペクトログラムを用いた研究がいくつか報告されている。スペクトログラムとは、音の周波数スペクトラムを時間、周波数および振幅を色の違いや濃淡で表したグラフである。音情報に深層学習を適用した研究事例としては、例えば、会話からの音声抽出（高市ら 2019）、生き物の鳴き声からの種の識別（松井 2018）や蛇口から流れる水音の認識（佐々木・佐々木 2015）といった非音声である環境音認識、コンピュータによる医師診断支援システムの開発のための呼吸音の自動分類（南ら 2020）、打音を用いた鋳鉄の材質判定（篠原ら 2020）、などがあげられる。深層学習は、従来の画像識別方法と異なり、画像識別に用いる画像特徴量を研究者や開発者が直接与える必要がなく、コンピュータ自らが与えられた教師データから特徴量を抽出して認識を行うため、画像認識の専門家でなくても幅広く利用できる（牧野・西崎 2018）。また、深層学習を実行できるソフトウェアの多くが無料で比較的簡易に利用可能であることから、様々な分野の人々が深層学習を用いた研究等を行うことができるといった現状がある。

以上のように、音情報を画像に変換することにより、様々な分野において深層学習を適用することが可能となること示されている。そこで、筆者らは、樹幹を叩いたときに発生した音をスペクトログラムで図示・画像化し、深層学習を用いて画像認識することにより、樹高および材積を推定することを研究目的とした。なお、スペクトログラムを深層学習の入力画像とした主な理由は、樹幹を打撃した際に発生した音は、音の大きさである音圧（dB）が時間と

ともに変化すると考えられるため、音を画像化する際、画像データに時間情報が含まれるスペクトログラムが有効であると考えたからである。また、多くの画像認識では深層学習は分類型の学習モデルを構築するが、樹高、材積が連続値であることから、回帰型の学習モデルによる推定を行った。

II. 打撃音測定試験および解析方法

1. 打撃音測定試験地

本研究では、京都府立大学生命環境学部附属大野演習林内（以下、大野演習林）に位置する22林班内において打撃音の測定試験を2020年10月30日、11月16日の2日間に行った。測定試験地の面積は約0.30 ha、南向き斜面に位置し、傾斜度は約23°から36°である。試験地内には、1966年に植栽されたスギが成立しており、筆者らが2018年10月に毎木調査を行っている。毎木調査では、胸高直径は直径巻尺を、樹高はHagröf社のVertexを用いて測定した。材積は大野演習林で調整された2変数材積式（本吉ら 1964）を用いて算出した。打撃音の測定試験を行う対象木の選定手順は次の通りである。まず、2018年10月の毎木調査の結果を基に測定試験の対象となる立木を選定した。選定基準は、樹高および直径をそれぞれの最大、最小値を基に3区分（大・中・小）し、各区分にできるだけ偏りがないよう選定した。次に、選定した対象木は、打撃音の測定の対象木となることから、2020年10月に直径、樹高の再計測を行った。なお、2018年の測定時から成長したため、当初とは異なる区分に該当された立木は測定対象木から除外し、他の立木を選定した。また、2018年の毎木調査では、40年以上木材市場に関わってきた林業熟練者により、割れやキズ、腐りなどに関する立木評価を行っている。本研究では、立木評価の結果を基に樹幹内部や表面に欠点が見られないと評価された立木を測定対象木とした（表-1）。

2. 打撃音測定試験

打撃音の録音環境には、Spectra SOFT社の「SpectraPLUS-SC Sound Card Edition」（以下、SpectraPLUS）（SpectraSoft website）を搭載した測定用PCを使用した。Spectra PLUSは、音データから時間や周波数ごとの音圧レベルの表示、音響特性や残響特性を解析することができるスペクトラムアナライザーソフトである。録音用マイクには、BEHRINGER社の無指向性マイクであるECM8000（BEHRINGER website）を、またマイクアンプにはTASCAM社のUS-144MKII（TASCAM website）を使用した。上記のソフト、マイク、マイクアンプを用いた録音環境については、同録音環境を用いた先行研究もあり（宮城ら 2021）、本研究の目的に適用と考えられる。

本研究では、Spectra PLUSのRecorderモードを用いて、立木の胸高位置における樹幹表面をゴム製ハンマーで打撃したときの音を録音した。録音は1回につき、20回分の打撃音を打撃音の間隔を3~4秒程度として録音した。この操作を1本の立木に対して5回行い、最終的には立木1本当たり100回の打撃音を録音した。打撃方法については、地面から約1.3 mの高さにおいて、幹から2~3 cm程度離

表-1. 毎木調査を基にした「大・中・小」区分および学習パターンⅡ、Ⅲに用いた立木 No. とその組み合わせ

立木 No.	胸高 直径 (cm)	樹高 (m)	区分	学習パターン						立木 No.	立木 材積 (m ³)	区分	学習パターン					
				LP-II ¹⁾			LP-III ²⁾						LP-II ¹⁾			LP-III ²⁾		
				set_ h201	set_ h202	set_ h203	set_ h301	set_ h302	set_ h303				set_ v201	set_ v202	set_ v203	set_ v301	set_ v302	set_ v303
197	25.2	19.3	小					○		197	0.4548	小					○	
174	45.9	21.8					○			139	0.5228		○					○
152	29.7	22.5							○	204	0.5574				○			
139	24.3	23.2				○				152	0.7199					○		
204	24.8	23.8		○			○			121	0.8235			○			○	
173	55.2	25.4						○		120	0.9705							○
112	34.2	25.7			○				○	112	1.0695					○		
148	36.9	26.6	中			○			○	148	1.2703	中						○
134	47.0	26.9					○			136	1.3534					○		
156	52.0	27.0			○			○		174	1.4980						○	
121	28.2	27.6					○			133	1.5731			○		○		
177	45.5	28.2		○					○	177	1.9606		○					○
120	29.5	29.8						○		134	1.9705				○		○	
104	52.1	30.3		大						○	163		2.0043	大				
163	42.9	31.6				○		○		127	2.2130					○		
136	34.0	32.1	○							156	2.3647				○			○
127	44.5	32.6						○		173	2.4558	○					○	
133	36.6	32.7					○			137	2.6632		○					
137	48.5	33.6			○				○	104	2.6953							○
132	55.5	34.2					○			132	3.4447					○		
最小値	24.3	19.3								最小値	0.4548							
最大値	55.5	34.2								最大値	3.4447							
平均値	39.6	27.7								平均値	1.6293							

¹⁾ 学習パターンⅡ：大中小の区分から樹木1本分ずつの画像を全て取り除いたものを訓練事例 (8,500 枚)、取り除かれた画像を未知事例 (1,500 枚) とした学習。²⁾ 学習パターンⅢ：大中小の区分から樹木2本分ずつの画像を全て取り除いたものを訓練事例 (7,000 枚)、取り除かれた画像を未知事例 (3,000 枚) とした学習。

れた地点に録音用マイクを設置し、マイクの先端付近の箇所に対しゴム製ハンマーを用いて、波形が測定できる一定の強さで打撃した。また、すべての打撃は同一人物が行った。人が叩くため、叩く強さを正確には制御できないが、できるだけ叩く強さが一定となるよう打撃した。

3. 打撃音の画像化

本研究では、Spectra PLUS で表示できる二つのグラフ「Time Series」および「Spectrogram」を用いて打撃音の解析およびその画像化を行った。Time Series は横軸が時間 (秒)、縦軸は振幅 (振幅には dB や Pascal など) を表すオシロスコープグラフであり、Spectrogram は横軸が時間、縦軸が周波数 (Hz)、振幅を色の違いや濃淡で表したスペクトログラムグラフである。次に解析に用いる画像の作成方法を述べる。図-1 に示すように、Spectra PLUS では、20 回分の打撃音を含む一つの録音ファイルを読み込むと、Time Series に各打撃音の瞬間における音圧のピークを示すグラフが表示される。そこで、本研究では、打撃音の開始・終了を正確に特定できないこと、および音の減衰の程度を踏まえ、解析に用いる音は打撃音のピークの 0.1 秒前から 0.5 秒後の 0.6 秒間とした。この 0.6 秒間について、振幅には音圧 (dB) を用い、音圧が大きい場合 (赤色) から小さい場合 (青色) にかけてカラーグラデーション表示した画像をスペクトログラムとして作成した。さらに、一度の打撃音に対し、Time Series を用いて少しずつ時間を進めることで (ここでは、0.005 秒)、5 枚の画像を作成し、最終的に 1 回の録音ファイルから 100 枚、立木 1 本から 500 枚のスペクトログラムグラフ表示した画像データを作成し

た。上述の方法で作成した画像には凡例や軸の値が含まれることから、画像処理ソフト ImageJ 1.50 (NIH website) を用いて、グラフの軸を除去した正方形画像を作成し、深層学習で用いる入力画像データとした。なお、後述する深層学習環境では、自動的に 64×64 pixel サイズに変換されることから、本研究では、変換時に画像が歪まないよう、縦横比が同じとなるよう正方形であることのみを考慮し、445×445 pixel サイズとした。

4. 畳み込みニューラルネットワークと Neural Network Console

本研究では、深層学習として畳み込みニューラルネットワークを用いた。畳み込みニューラルネットワークとは、ニューラルネットワークモデルの一つで、主に画像認識分野に用いられる代表的な深層学習である。CNN は、畳み込み層、プーリング層および全結合層からなる何段もの深い層を持ったニューラルネットワークである。畳み込み層では入力画像に対し、畳み込み演算を適用して小領域における画像の局所的な特徴量を抽出し、プーリング層では畳み込み層で抽出した特徴量を圧縮する。これを何度も繰り返した後、全結合層において最終的な処理を行いその結果を出力する (山下 2016)。本研究では、LeCun ら (1999) が開発した LeNet を基本とし、出力層には回帰層を用いたネットワーク構造を使用した (図-2)。LeNet とは LeCun により提案された CNN の元祖となるネットワークであり、基本的な構造として畳み込み層とプーリング層のセットを 2 回繰り返すのが特徴である。深層学習を行う方法には、SONY の提供する Neural Network Console (以下、NNC)

1) 1セット 20回の打撃音

3) 0.005 秒ずつずらす→5 回繰り返す

2) 1回の打撃音を含む 0.6 秒間を 1 回の音データとする

4) 正方形にトリミング後、圧縮

図-1. Spectra PLUS の Time Series, Spectrogram を用いた打撃音の画像化手順

を用いた (SONY Neural Network Console website)。NNC は GUI による操作によって、モデルの設定やパラメータの調整などを自動でかつ簡易に行うことが可能であるため、NNC を用いた応用研究は数多くある (杉浦・奥村 2018; 菊地ら 2019; 大光寺ら 2020; 鶴ら 2020)。NNC のネットワーク構造およびそのパラメータは次のようにした。ネットワークの基本構成は、図のように Convolution 層 4 層, Pooling 層 2 層からなる CNN とした。Input 層は RGB3 チャンネルの $3 \times 64 \times 64$ (チャンネル数, 縦サイズ, 横サイズ) を, Convolution 層は始めの 2 層にはフィルタの KernelShape のサイズを 5×5 とした BinaryConnect を用い, 残り 2 層には Affine を用いた。Pooling 層は近傍の入力値の最大値を出力する MaxPooling を用い, フィルタは 2×2 とした。Active 層の活性化関数は ReLU を用いた。また, 精度向上や計算の収束を早めるため, Convolution 層や Active 層の後ろに BatchNormalization (データの正規化) を挿入した。出力層の前に Sigmoid 関数を加え, 出力層の損失関数は, 連続値を出力するニューラルネットワークを最適化するため, データセットの変数との二乗誤差を最小化する SquaredError を設定した。画像解像度は 96 dpi, バッチサイズは 64 を用いた。また, 各画像に対して, 毎木調査から得られた樹高および材積とラベル付けを行う必要があるため, 樹高, 材積を正規化し入力値とした。

5. シミュレーション条件およびモデルの性能評価

本研究における学習パターンは以下の 3 通りとした。学習パターン I (LP-I) は, 録音データを 5 セットに分け (1 セット = 20 本 $\times$ 100 枚 = 2,000 枚), そのうちの 4 セット (8,000 枚) を訓練事例に, 残りの 1 セット (2,000 枚) を未知事例とした学習 (set01 から set05) である。学習パターン II (LP-II) は, 樹高, 材積ともにそれぞれ降順で並び替え, 大・中・小に 3 区分した後, 各区分から立木 1 本分

図-2. 深層学習に用いたネットワークの構成

ずつ (3 本 $\times$ 500 枚 = 1,500 枚) の画像を全て取り除いたものを訓練事例, 取り除かれた画像を未知事例とした学習 (樹高: set_h201 から set_h203, 材積: set_v201 から set_v203) である。学習パターン III (LP-III) は, LP-II と同様に大・中・小区分を行った後, 樹高, 材積ともに各区分から立木 2 本分ずつ (6 本 $\times$ 500 枚 = 3,000 枚) の画像を全て取り除いたものを訓練事例, 取り除かれた画像を未知事例とした学習 (樹高: set_h301 から set_h303, 材積: set_v301 から set_v303) である。LP-II および LP-III で用いた立木の組み合わせは, セット間でできるだけ重複しないよう, かつ偏りが生じないように選択した (表-1)。学習回数である epoch 数は, 各学習パターンともに最大値を 500 epoch とした。また, 学習時には NNC のオプションの一つである構造自動探索を用いた。一般に, 深層学習では, ネットワークの構造やハイパーパラメータ (人が調整するパラメータ) は自由に設定できる反面, その組み合わせは膨大であるため,

その中から最適なモデルを見つけるのは非常に困難である。また、ハイパーパラメータは明示的に決められているわけではない。そこで、NNCでは、構造自動探索を用いることにより、解析者がハイパーパラメータを設定しなくても様々な学習モデルを実行することができる。構造自動探索では、最初に用いたネットワーク構造を基に、ネットワーク構造やそのパラメータを自動で構築・設定し、順次学習を実行する。ただし、NNCの構造自動探索オプションには明示的に終了回数等を設定することはできないため、解析者が実行を停止しない限り学習は継続される。そこで、本研究では、便宜的にすべての学習パターンに対し、50通りの学習モデルを構築した時点で終了することとした。構築した学習モデルは、樹高、材積ともにLP-Iは250、LP-IIおよびLP-IIIは150とし、樹高、材積併せて、計1,100の学習モデルを作成した。

学習モデルの精度検証には、平均絶対誤差 (Mean absolute error, 以下 MAE, 式 (1)), 平均絶対パーセント誤差 (Mean absolute percentage error, 以下 MAPE, 式 (2)) および決定係数 (R^2) の三つの指標を用いた。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{obs,i} - y_{pred,i}| \quad (1)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_{obs,i} - y_{pred,i}}{y_{obs,i}} \right| \quad (2)$$

ただし、 y_{obs} は実測値、 y_{pred} は推定値を表す。また、構造自動選択によって学習した50の学習モデルから、各学習パターンでそれぞれ一つずつ最適モデルを決定した。最適モデルを決定する方法は、以下の手順で行った。①. 50の学習モデルについて、決定係数、MAE、実測値から推定値を引いた値の標準偏差、の三つの指標を算出する。②. 各指標について50の学習モデルを1位から50位までとして順位付けを行う (ただし、決定係数は最も高い値を、MAE および実測値から推定値を引いた値の標準偏差は最も小さい値をそれぞれ上位とする)。なお、ここで用いた

三つの指標は、それぞれ異なる側面からモデルの適切さを定量化しているため、最適モデルの決定には三つの指標を総合的に評価する必要があると考えた。そこで、本研究では三つの順位を足し合わせた値を総合的な指標とし、最も小さい値を示したモデルを最適モデルとして、解析に用いることとした。

III. 結果と考察

表-2 は、LP-I から LP-III について、樹高および材積の推定精度を表したものである。また、LP-I から LP-III ともに、三つの指標は未知事例に対する推定結果を示した。

LP-I の結果より、全データの R^2 値は樹高が 0.9993、材積が 0.9996 とどちらも高い値を示した。MAE は樹高が 0.08、材積が 0.01 となり材積の値が非常に小さい値を示した。一方、MAPE は樹高が 0.30、材積が 1.44 となり、MAE とは逆に樹高の方が低い値を示した。また、樹高、材積の各セットともにどの指標に対してもほぼ同じ値を示した。LP-II の結果より、全データの R^2 値は樹高が 0.9433、材積が 0.9914 となり、樹高が若干低いものの、総じてどちらも高い値を示した。MAE は樹高が 0.67、材積が 0.05、MAPE は樹高が 2.42、材積が 4.69 となり、LP-I と同様の傾向を示した。また、3 セット行った学習結果では、樹高では set_h203 の R^2 値が他のセットの R^2 値と比べ低い値を示した。同様に、材積では、set_v202 の R^2 値が他のセットの R^2 値と比べ若干ではあるが低い値を示した。LP-III の結果より、全データの R^2 値は樹高が 0.3672、材積が 0.9192 となり、材積については LP-I や LP-II と同様、高い値を示したが、樹高の R^2 値は LP-I や LP-II と比べ非常に低い値を示した。また、樹高については、3 セット行った結果のばらつきが大きく、set_h302 の R^2 値は 0.0992 とほぼ推定できなかった結果を示した一方、set_h303 の R^2 値は 0.8525 と比較的高い値を示す結果となった。MAE は樹高が 2.61、材積が 0.17、MAPE は樹高が 10.06、材積が 12.05 となり、LP-I、LP-II と同様、MAE は材積が、MAPE は樹

表-2. 深層学習を用いた各学習パターンにおける樹高および材積の推定精度

学習パターン	データセット	樹 高			データセット	材 積		
		MAE	MAPE	R^2		MAE	MAPE	R^2
LP-I ¹⁾	全データ ⁴⁾	0.08	0.30	0.9993	全データ ⁴⁾	0.01	1.44	0.9996
	set01	0.09	0.31	0.9992	set01	0.02	1.68	0.9996
	set02	0.08	0.28	0.9994	set02	0.01	1.22	0.9997
	set03	0.09	0.32	0.9993	set03	0.01	1.31	0.9996
	set04	0.08	0.29	0.9993	set04	0.01	1.51	0.9995
	set05	0.08	0.29	0.9992	set05	0.01	1.47	0.9996
LP-II ²⁾	全データ ⁴⁾	0.67	2.42	0.9433	全データ ⁴⁾	0.05	4.69	0.9914
	set_h201	0.60	2.14	0.9746	set_v201	0.04	4.39	0.9975
	set_h202	0.45	1.59	0.9735	set_v202	0.08	5.42	0.9856
	set_h203	0.96	3.52	0.8869	set_v203	0.05	4.27	0.9947
LP-III ³⁾	全データ ⁴⁾	2.61	10.06	0.3672	全データ ⁴⁾	0.17	12.05	0.9192
	set_h301	2.49	10.12	0.4369	set_v301	0.30	18.92	0.9776
	set_h302	4.01	15.10	0.0992	set_v302	0.09	6.63	0.9735
	set_h303	1.34	4.96	0.8525	set_v303	0.11	10.60	0.9768

MAE, 平均絶対誤差; MAPE, 平均絶対パーセント誤差; R^2 , 決定係数; LP-I は未知事例に対する 5 回の平均値, LP-II および LP-III は未知事例に対する 3 回の平均値, MAE, MAPE, R^2 は 50 パターン作成した学習モデルのうち, MAE が最も小さかったモデルの値, LP-II および LP-III のデータセット名は, 表-1 で示した組み合わせで用いたセット名と同じ。¹⁾ 訓練事例 (8,000 枚), 未知事例 (2,000 枚) とした学習。²⁾ 大中小の区分から樹木 1 本分ずつの画像を全て取り除いたものを訓練事例 (8,500 枚), 取り除かれた画像を未知事例 (1,500 枚) とした学習。³⁾ 大中小の区分から樹木 2 本分ずつの画像を全て取り除いたものを訓練事例 (7,000 枚), 取り除かれた画像を未知事例 (3,000 枚) とした学習。⁴⁾ 各セットを合わせて計算した値。

高が低い値を示した。

LP-Iの結果から、全データの20%を未知事例とした推定であっても、同一立木に対する打撃音のスペクトログラムが訓練事例に含まれている場合は樹高および材積ともに高い精度で推定することが可能であることが示唆された。LP-IIおよびLP-IIIの結果より、推定対象である立木に対する打撃音の波形画像をすべて未知事例とした場合は、LP-Iに比べると推定精度は低下する結果を示したが、LP-IIの結果より、各区分からの抽出本数が1本の場合では、訓練事例に全く用いなかった画像に対しても、かなり高い精度で推定することが可能であることが示唆された。一方、LP-IIIの結果より、各区分からの抽出本数が2本だと、訓練事例に含まれない画像パターンが多くなることから、深層学習を用いた推定であっても推定精度は低下し、特に樹高に対する推定精度は非常に低くなる傾向を示した。また、樹高については、セット間でのばらつきが非常に大きいことから、未知事例として用いられた樹高の組み合わせ方の違いが推定精度に大きく影響していると考えられた。ただし、材積に対する推定精度は極端に低下していなかったこと、一方、樹高に対しては、セット間でばらつきが大きかったことから、打撃音の波形画像を画像化し深層学習を用いた解析においては、樹高のような1次元の情報より、材積のような3次元の情報に対して有効ではないと考えられた。

図-3から図-8は、毎木調査で得られた実測値と深層学習による推定値を示した散布図である。横軸は実測値を、縦軸は推定値を表している。LP-Iでは検証を行った5回の結果をすべて示しており、一つの立木に対する推定値は500個(1セット(100枚/本)×5セット)、全部で10,000個(500個×20本)の推定値が示されている。また、LP-IIおよびLP-IIIでは、LP-Iと同様に一つの立木に対する推定値は500個、LP-IIは4,500個(500個×3本×3セット)、LP-IIIは9,000個(500個×6本×3セット)の推定値が示されている。散布図内の実線は推定値と実測値が一致する値を示し、白抜きの凡例は推定値の平均値である。

図-3および図-4より、LP-Iの樹高、材積ともに、平均値はほぼ実線上に分布していること、平均値以外の推定値のばらつきも小さいことから全体的に偏りなく高い精度で推定されていると考えられた。図-5より、LP-Iの結果に比べ、全体的に各立木の推定結果にばらつきが多く見られた。特徴としては、比較的樹高が高い木の推定値が過小評価される傾向を示した。図-6より、材積の場合は、LP-Iの材積と比べ、表-2で示したように、全体的に絶対誤差が大きくなっているためMAEやMAPEは大きな値を示したが、図から推定値にはそれほど偏りは生じていない結果となった。しかしながら、比較的材積が大きい立木における推定値のばらつきは大きい結果となった。図-7より、推定結果および推定値のばらつきは全体的に大きく、樹高30m付近を境に、30m以上の立木を過小評価、30m以下の立木を過大評価する傾向を示した。図-8より、LP-IIIの樹高と比べると、全体的に高い推定精度を示したが、最も材積が大きかったNo.132の立木における推定結果は他の結果と比べて、実測値よりもかなり過小な値として推定さ

れる結果となった。また、材積が小さい立木に対しては、若干ではあるが過大に推定される結果を示した。今回の学習に用いたデータセットに材積が 3 m^3 を超えるものがNo.132の 3.4447 m^3 と1本のみであり、その次に大きな立木がNo.104の 2.6953 m^3 であったこと、すなわち 3 m^3 前後のデータが存在しなかったことにより推定精度が低く見積もられたのではないかと考えられた。一方、材積が小さい立木では、極端に離れたサンプルが存在していないため、それほど推定精度は低くならなかったのではないかと考えられた。図-5から図-8のLP-IIおよびLP-IIIの結果より、未知事例に対しては、樹高と材積では、推定精度のみならず誤差の傾向においても違いが見られた。

LP-IIIでは、解析対象とした立木20本のうち、6本(全体の3割)が未知事例となっている。樹高、材積、それぞれ3セットを作成したが、その際、同じ立木が何度も選ばれないよう、また偏りが無いよう選定した。ただし、3割を未知事例としたことから、未知事例の中には選択された立木の樹高や材積の値が訓練事例の最大、最小値の範囲には含まれていない場合が存在する。また、偏りが無いよう選定する際、例えば、樹高については樹高の中での偏りが無いよう選定したが、その選定された立木を材積で見た場合に偏りが生じているかどうかは考慮していない。そこで、表-3より、LP-IIIにおける訓練事例、未知事例に用いた樹高および材積の値の範囲について検討した。表中の樹高および材積の値は、NNCの入力値として用いた正規化した値である。表中の未知事例は、各学習パターンにおいて未知事例として用いた立木の樹高および材積の範囲について示している。また、分散は散らばりの指標であるが、ここでは分散が小さいほど同じような大きさの立木が選ばれていると考えて検討する。表-3 1)より、樹高については、set_h302の未知事例の樹高の最小値が訓練事例外であること、および材積の分散が小さいこと、また、set_h303では未知事例の樹高および材積の分散が小さいこと、といった点があげられる。set_h303は、他のセットに比べ推定精度が高かったが、その要因として、訓練事例は解析対象とした立木の樹高をすべて含み、かつ分散が大きいこと、すなわち各区分から偏りなく選定されていたことが考えられる。一方、set_h302の推定精度は低い結果を示したが、その要因として、set_h302の未知事例には訓練事例の最大、最小値の範囲には含まれていない立木が存在し、かつ未知事例に比べ訓練事例の分散が小さいこと、すなわち訓練事例に偏りが生じたことが考えられる。表-3 2)より、材積については、set_v301の未知事例の材積の最大値が訓練事例外となったが、 R^2 値は他のセットとほとんど変わらない結果を示した。ただし、set_v301のMAEやMAPEは他のセットに比べ大きな値を示したことから、 R^2 値は高いものの他のセットに比べ誤差が大きくなったと考えられる。

以上の結果を踏まえ、本研究で行った深層学習に用いた入力および出力情報について考察する。学習用データ数について、深層学習を適用するためには数多くの学習用デー

図-3. 深層学習による推定結果 —学習パターンⅠ, 樹高—

図-4. 深層学習による推定結果 —学習パターンⅠ, 材積—

図-5. 深層学習による推定結果 —学習パターンⅡ, 樹高—

図-6. 深層学習による推定結果 —学習パターンⅡ, 材積—

図-7. 深層学習による推定結果 —学習パターンⅢ, 樹高—

図-8. 深層学習による推定結果 —学習パターンⅢ, 材積—

表-3. 学習パターンⅢにおける訓練事例、未知事例に用いたデータの範囲

1) 樹高					2) 材積				
学習パターン		訓練事例		未知事例	学習パターン		訓練事例		未知事例
LP-III ¹⁾		樹高		材積	LP-III ¹⁾		材積		樹高
set_h301	最小	0.00	0.17	0.03	set_v301	最小	0.00	0.09	0.21
	最大	0.96	1.00	1.00		最大	0.75	1.00	0.90
	分散	0.08	0.09	0.10		分散	0.07	0.09	0.08
set_h302	最小	0.17	0.00	0.00	set_v302	最小	0.02	0.00	0.00
	最大	1.00	0.89	0.67		最大	1.00	0.67	0.83
	分散	0.07	0.09	0.06		分散	0.08	0.05	0.07
set_h303	最小	0.00	0.21	0.09	set_v303	最小	0.00	0.02	0.26
	最大	1.00	0.96	0.75		最大	1.00	0.75	0.70
	分散	0.09	0.06	0.07		分散	0.08	0.07	0.02

表中の値は正規化した値。¹⁾ 学習パターンⅢ：大中小の区分から樹木2本分ずつの画像を全て取り除いたものを訓練事例（7,000枚）、取り除かれた画像を未知事例（3,000枚）とした学習。

タを用意する必要がある、また、データの加工等に時間を要することが多いといったことから、解析対象から大量の学習用データを収集できるのかどうかの検討、あるいは効率良くデータを収集する方法を構築するにはどうしたら良いかといった課題がある。深層学習では、機械学習と同様、様々な条件に適用したモデルを構築するために大量の学習用データを用意し学習を行う必要がある。しかし、大量のデータを収集するためには労力を要することから、深層学習では、少ないデータで効率良く学習する方法がいくつか提案されている。その代表的な方法として、水増し（Data augmentation）や転移学習（Transfer learning）といった方法がある（山下 2016；南ら 2020）。特に、転移学習は、別の対象や分野、あるいはまったく異なる分野において学習した学習済みモデルを用いることにより、識別したい対象について、大量のデータを用意することができなくても効率良く学習モデルを構築することが可能であるため、学習用データ数を少なくするための非常に有効な手段である。本研究では、深層学習を適用するため、立木 20 本について 10,000 枚の画像データを作成したが、かなりの労力を要したこと、また実用時にはさらに多くのパターンを学習する必要があることから、転移学習などを適用して、より効率よく学習を行うことが必要である。

学習アルゴリズムについて、本研究では NNC 上で LeNet を用いた解析を行ったが、LeNet は層数が少ないため、比較的速く計算することが可能であり、多くの学習パターン、学習モデルを構築、検証するには有効である。しかしながら、一般に、深層学習では層数を多くする方向で開発が進められ、また層数の多いモデルの方がより認識性能が高い結果を示している。よって、今後は LeNet 以外の深層学習アルゴリズムを用いた検証も必要である。また、本研究では、樹高、材積ともに同じ学習アルゴリズムであったが、推定精度を向上させる上では、必ずしも同じアルゴリズムである必要はないため、それぞれに対して異なる学習アルゴリズムを適用することも考えられる。

樹幹からの打撃音について、本研究では人が立木をハンマーで打撃した音を用いた。これは、将来的には人が叩いた音を現地にてスマートフォンなどで録音し、その場で推定を行うことを考慮しているためである。また、深層学習

では、多少ノイズが含まれているような画像であっても、大量に用意し学習することにより、認識を可能としていることから、人が叩いた音がそれほど正確な打撃でなくても問題ないと考えられる。しかしながら、大量の学習用データを取得する上では、人が叩く方法は制限が多いため、今後は人が叩く以外の方法について検討を行う必要がある。

立木の内部情報について、本研究では測定対象木として毎木調査の際に行った立木評価の結果から、割れやキズ、腐れなどの欠点がないと思われる立木を選定した。よって、本研究の結果には、腐れなどの影響はほとんどないと考えられる。ただし、立木評価を行ったのが 2 年前であること、また打撃した立木を伐倒し腐れなどを確認したわけではないことから、完全に影響がないとは断定できない。また、打撃しただけではわからない程度に立木内部が腐朽している場合などについても、打撃音は異なると考えられる。打撃音を測定した立木をすべて伐倒し腐れなどの調査を行うのは大変な労力を要する作業であり、また、深層学習に適用するためには大量の学習用データが必要であるといった観点から、すべての測定対象木を伐倒するのは現実的ではないと考えられる。しかしながら、立木の内部状況の違いは音情報における重要な要因の一つになると考えられることから、今後、十分に検討する必要がある。また、立木の内部情報と音との関連性については、音圧や周波数特性といった音の特性を用いた音響解析から、例えば、木製ハンマーで樹幹を打撃し、そのときの音の共振周波数から樹幹内部を診断する方法である横打撃共振法（小玉ら 1999；釜口ら 2001）、超音波を用いた立木の材質測定（中村 1995）、あるいは打音による樹内腐朽簡易診断装置「ぼん太」（陶山ら 2012；ワールド測量設計 2014）、といった音を利用して立木の内部構造を調べる研究、実用事例がある。本研究の成果は、材積や樹高といった量的な情報ではあるが、音響解析から立木の内部情報が得られることから、質的な情報についても検討する必要があると考えられる。

音情報の画像変換について、本研究では音といった一次元信号から二次元画像を作成し、単純フーリエ変換などから作成される三次元グラフであるスペクトログラムを深層学習の入力画像として用いた。本研究では、予備調査としてスペクトログラムではなく横軸を周波数、縦軸を音圧と

したスペクトラムを用い、本研究と同様に画像化し深層学習を適用して樹高や材積の推定を行った。その結果、訓練事例に対しては非常に高い精度で推定することが可能であったが、未知事例に対しては高い推定精度は得られなかった。その要因の一つとして、スペクトラムには時間情報が含まれていないが、立木を叩いた音は時間と共に減衰し、減衰の仕方は叩いた木の大きさ等にも影響するため、スペクトラムのような時間情報が欠如したグラフでは減衰の違いを表すことができないためであると考えられる。また、三次元グラフ画像としては、スペクトログラム以外にも連続ウェーブレット変換から作成されるスカログラムがあり、深層学習の入力値としてスカログラムを用いた研究事例もある（南ら 2020）。南ら（2020）では、スペクトログラム、スカログラムをそれぞれ単独で用いて学習を行った場合、スペクトログラムの推定値の方がスカログラムの推定値より高い精度を示したが、これら二つを併用すると単独で用いるよりさらに高い推定精度を示したことから、音情報からより精度の高い立木測定を行うためにはスカログラムについて検討する必要があると考える。また、黒澤ら（2020）によれば、人間が視覚できる物体などを画像化した視覚的画像と音をスペクトログラムなどにより画像化した聴覚的 2 次元特徴量は画像として異なるとしている。すなわち、視覚的画像は画像中のどの部分についても形状に違いがなく、対象とする物体がどこにあるかに依存されない。よって、深層学習を用いた視覚的画像の識別では、対象画像の全領域を走査する必要がある。一方、聴覚的 2 次元特徴量は周波数帯の違いが波形や画像の濃淡に影響するため、深層学習を適用する際、視覚的画像とは異なり、全領域ではなく特定の領域のみを走査するだけでよい。本研究では、スペクトログラム画像の全領域に対して深層学習を適用したが、黒澤ら（2020）が指摘するように、スペクトログラム画像は聴覚的 2 次元特徴量から作成された画像であることから、全領域ではなく特定の領域のみを用いた解析も検討する必要がある。

深層学習から得られた出力値について、一般に機械学習を用いた推定では、学習が不十分なためモデルの推定精度が十分でないといった未学習（underfitting）や訓練事例を学習し過ぎたため未知事例に対する推定精度が低くなるといった過学習（overfitting）といった問題が生じることが多いため（Raschka and Mirjalili 2018）、学習モデルを作成する際に留意する必要がある。特に、学習モデルにおけるバイアス（Bias：実測値と推定値との誤差の平均）とバリエーション（Variance：推定値の散らばりの度合い）に関しては、バイアスが高いと未学習、バリエーションが高いと過学習となるため、これらができるだけ低くなるようモデルを構築することが重要であるが、一般にバイアスとバリエーションはトレードオフの関係にあるため、両者を同時に低くするのは難しい（Raschka and Mirjalili 2018）。未学習であれば学習モデルの学習回数を増やすといった対策が、過学習であれば学習に用いたデータ数が少ないことや構築した学習モデルが複雑であることが主な要因であるため、学習に用いるサンプル数を増やす、複雑なモデルを単純化する、といっ

た対策が考えられる。本研究では、多くの学習モデルにおいて、未知事例の推定精度は非常に高い値を示した。また、NNC の構造自動探索から各学習パターンにおいて最適モデルとされた epoch 数の平均が 366 であったことから、epoch 数の上限を 500 とした学習条件で十分な学習が実行できたのではないかと考えられた。なお、過学習、未学習については評価指標がなく、訓練事例や検証事例の学習曲線を見て判断することとなる。よって、異なる epoch 数を用いたシミュレーションを行う必要があり時間を要することとなるが、NNC では構造自動探索によって最適な epoch 数の結果を出力することができるため、予め epoch 数を大きく設定すれば、比較的容易に十分学習したモデルを構築することができると思われる。

IV. お わ り に

本研究では、立木を打撃した際に発する音を録音し、録音された音を画像化し深層学習を適用することにより、樹高や材積といった立木情報を推定することが可能であることが示唆された。本研究では、打撃音を画像化する場合、時間情報を含んだグラフとしてスペクトログラムを用いたが、その結果、未知事例に対しても非常に高い推定精度が得られたことから、スペクトログラムによる音の画像化は有効であるといえる。なお、本研究では、一つの林分における試験調査のみであり、林内の環境条件や季節の違いについては検討していない。また、同一立木であっても樹体内の水分状況や枝葉の有無により、録音された音が違う可能性は十分に考えられる。より汎用性を高めるためには、様々な状況における打撃音を収集し、本研究の手法を適用することで推定可能であるかどうかを検証する必要があると考える。筆者らは、将来的には測定対象木を数回打撃、その場でスマートフォンに録音するだけで樹高や材積が推定可能なシステムの構築を目的としている。また、樹幹からの打撃音のみで樹高や材積が推定できるのであれば、全測定対象木を打撃することにより簡便かつ安価に毎木調査が実行できると考えられる。一方、深層学習からの推定値に対し、実際に測高器で測定した値で補正することより推定精度の向上が見込めるのであれば、例えば、全測定対象木については打撃音を測定し深層学習を用いて推定した後、測定対象木の何割かは従来通りの方法（測高器を用いた計測）で計測するといった方法も考えられる。林分内のすべての立木の樹高を測高器で計測するのは労力を要するが、何本か、あるいは何割かの立木であれば比較的容易に計測できると考えられることから、深層学習の推定値に対し、測高器で測定した値で補正する方法についても検討する必要がある。

謝 辞

本研究を行うにあたり、原稿改訂に関するご意見を多数いただいた 2 名の審査者および編集委員の方々、2018 年の毎木調査において立木評価を行っていただいた森林組合連合会コーディネーターの前田清二氏、毎木調査にご協力いただいた本学森林科学科流域情報学研究室および砂防学研究室の方々にお礼を申し上げます。

引用文献

- 坂東宜昭 (2020) 私のブックマーク：深層ベイズ学習と音環境理解. 人工知能 35: 855-861
- BEHRINGER <https://www.behringer.com/> (Accessed on 30 January, 2020)
- Dai W, Dai C, Qu S, Li J, Das S (2017) Very deep convolutional neural networks for raw waveforms. International Conference on Acoustics, Speech and Signal Processing
- 大光寺岳・鎌倉智之・北畠尚子 (2020) 静止気象衛星画像と深層畳み込みニューラルネットワークを用いた北西太平洋海域の台風の強度分類. 天気 67: 109-115
- Hagröf <https://www.haglof.jp/> (Accessed on 17 February, 2020)
- 井本桂右 (2018) 音響イベントと音響シーンの分析. 音響誌 74: 198-207
- 井本桂右 (2019) 環境音分析の研究動向. 音響誌 75: 512-518
- 磯 健一 (2017) 音声認識における Deep Learning の活用. 日本神経回路学会誌 24: 27-38
- 釜口明子・中尾哲也・中井毅尚 (2001) 横打撃共振法を用いた生立木の非破壊的内部診断. 樹木医学研究 5: 59-63
- 菊地輝行・崎田晃基・秦野輝儀・吉川 慶・西山 哲・大西有三 (2019) 深層学習による崩壊・非崩壊地の自動判読手法の開発. 日本地すべり学会誌 56: 255-263
- 金 明哲 (2005) WEKA と樹木モデル. ESTRELA 132: 64-69
- 小玉泰義・釜口明子・中尾哲也・原田照太 (1999) 横打撃共振法による樹幹内空洞の検出. 樹木医学研究 3: 5-9
- 小池耕太郎・井山徹郎・野内彩可・村上祐貴 (2015) コンクリート内部欠陥の深さが打音特性に及ぼす影響. コンクリート工学年次論文集 37: 1849-1854
- 小泉章夫・佐々木貴信・福井敬二・佐々木正吾 (1998) 縦振動法による丸太重量の測定. 東北森林科学会誌 3: 17-20
- 黒澤義明・目良和也・竹澤寿幸 (2020) 画像変換手法による音声強調のためのスペクトログラム変換. 人工知能学会全国大会論文集 34: 1-4
- LeCun Y, Bottou L, Benjio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86: 2278-2324
- LeCun Y, Haffner P, Bottou L, Bengio Y (1999) Object Recognition with Gradient-Based Learning. In: Shape, Contour and Grouping in Computer Vision, Lecture Notes in Computer Science. Springer Nature Switzerland AG, 319-345
- Lee J, Park J, Kim KL, Nam J (2017) Sample-level deep convolutional neural networks for music auto-tagging using raw waveforms. Sound and Music Computing Conference
- 牧野浩二・西崎博光 (2018) 算数&ラズパイから始めるディープ・ラーニング. CQ 出版社
- 松井孝典 (2018) 研究事例：深層学習技術による環境音識別. 音響誌 74: 407-408
- 南 弘毅・陸 慧敏・金 亨燮・平野 靖・間普真吾・木戸尚治 (2020) 時間・周波数解析と畳み込みニューラルネットワークを用いた呼吸音の自動分類. Med Imag Tech 38: 40-47
- 宮城昭博・美濃羽靖・森 有花・金田 茜・川村真央 (2021) 森林空間がもたらす騒音に対する音の物理的減衰効果. 森林計画誌 54: 93-103
- 本吉瑠璃夫・本城尚正・妹尾俊夫 (1964) 大野演習林におけるスギ人工林材積表の調製. 京府大演報 8: 33-49
- 中村 昇 (1995) 超音波による立木の材質測定—トドマツ・カラマツ類の場合—. 日林北支論 43: 163-165
- NIH <https://imagej.nih.gov/ij/> (Accessed on 30 January, 2020)
- 岡谷貴之 (2015) 深層学習. 講談社
- Quinlan JR (1993) C4.5: programs for machine learning. Morgan Kaufmann Publishers
- Raschka S, Mirjalili V (2018) Python 機械学習プログラミング 達人データサイエンティストによる理論と実践. 福島真太郎 監訳, インプレス.
- 林野庁 (2018) 高精度な森林情報の整備・活用のためのリモートセンシング技術やその利用方法等に関する手引き. 林野庁
- 佐々木逸士・佐々木健 (2015) 非音声の環境音認識に関する研究. 精密工学会学術講演会講演論文集: 973-974
- 篠田浩一 (2017) 音声言語処理における深層学習：総説. 音響誌 73: 25-30
- 篠原美月・内田 希・岩見祐貴・平本雄一・加藤雅也・菅野利猛 (2020) 人工知能を用いた打音試験による片状及び球状黒鉛鑄鉄の材質判定. J Comput Chem Jpn: 164-166
- 園田佳巨・川端健太・別府万寿博・福井雄貴 (2011) 打音データを用いたコンクリート内部欠陥の評価に関する基礎的考察. 構造工学論文集 A 57: 802-811
- SONY Neural Network Console <https://dl.sony.com/ja/> (Accessed on 10 March, 2021)
- SpectraSoft <http://spectrasoft.jp/> (Accessed on 30 January, 2020)
- 杉浦太郎・奥村圭二 (2018) 炭材存在下での石灰石熱分解挙動—深層学習を用いた反応性評価—. 鉄と鋼 104: 758-765
- 陶山大志・永石憲道・坂越浩一 (2012) 横打撃共振法によるクロマツ樹幹内部の腐朽・空洞面積率の推定. 日林誌 94: 161-166
- 高市晃佑・片上敬雄・黒澤義明・目良和也・竹澤寿幸 (2019) 深層学習を用いた画像変換に基づく会話からの音声抽出. 人工知能学会全国大会論文集 33: 1-3
- TASKAM <https://tascam.jp/jp/> (Accessed on 30 January, 2020)
- the National Institutes of Health <https://imagej.nih.gov/ij/> (Accessed on 30 January, 2020)
- 鶴 鉄雄・菅原正博・西村治彦 (2020) 深層ニューラルネットによるファッション画像のシルエット識別の実現に向けた試み. 日本感性工学会論文誌 9: 117-126
- ワールド測量設計 (2014) 打撃音樹内腐朽簡易診断装置「ぼん太」 <http://www.world-ss.co.jp/HP2/ponta2/> (Accessed on 10 March, 2021)
- Vapnik V, Lerner A (1963) Pattern recognition using generalized portrait method. Automat Remote Cont 24: 774-780
- 山崎俊彦 (2010) 画像の特徴抽出: Histogram of Oriented Gradients (HOG). 映像情報メディア学会誌 64: 1322-329
- 山下隆義 (2016) イラストで学ぶディープラーニング. 講談社