

CNN を用いた木材破壊音検知における環境音の影響

東京理科大学 学生会員 ○董 子鳴

東京理科大学 正会員 佐伯 昌之

1. 研究の背景と目的

地震が発生した際、構造部材の破壊音を検知することができれば、構造被害を迅速に把握するための有益な情報になりえる。そのため、著者らは木材を対象とし、その破壊音を検知することを試みている。

これまでに、木材破壊音のスペクトログラムを CNN (Convolutional Neural Network) モデルに学習させ、純粋な木材破壊音であれば 95% 以上の精度で分類できることが確認されている¹⁾。また、実際の地震時を想定して、様々な環境で木材破壊音が録音され、ノイズとして環境音を混合したデータセットが生成された。そして、学習とテストで異なる録音環境のデータセットを用いて CNN による分類精度が調べられ、accuracy は 79.9%、recall は 62.7% となり、非常に低い判定精度が得られた²⁾。そこで、本研究では、この精度低下の原因を調べることにした。

2. 本研究で使ったデータ

使用する木材破壊音は、先行研究で用いたものであり、松、杉、ホワイトウッドの集成材、樺 (ツガ)、ヒノキの 5 種類の木材を曲げ破壊および継ぎ手破壊した際の音を録音したものである。録音には 6 種類のスマートフォン (iphone12, iphone7, ipad mini, Xperia, AQUOS, Rakuten Hand) が使用され、デバイスは次の 6 つの状況に置かれた。

状況 A : 箱の中、木材から 1.5 m、直置き。

状況 B : 箱の中、木材から 1.5 m、布に包む。

状況 C : 箱の外、木材から 1.5 m、直置き。

状況 D : 箱の外、木材から 1.5 m、布に包む。

状況 E : 部屋の外、直置き。

状況 F : 部屋の隅、直置き。

また、ノイズとなる環境音としては日常生活で頻繁に発生する環境音が収録されている K. J. Piczak の ESC-50 データセット³⁾が用いられた。

以上の録音データは wav 形式で保存されており、サンプリング周波数は 44.1 kHz である。

3. スペクトログラム作成

純粋な木材破壊音の wav ファイルから、木材破壊音の最大振幅を中心に、データ長 1.5 秒で切り出した後、スペクトログラムに変換した。ESC-50 データセットに保存されている環境音は 4 秒のデータであるため、本研究用に最大振幅を中心に 1.5 秒で切り出し、スペクトログラムに変換した。

混合音は、木材破壊音の最大振幅を中心にデータ長 1.5 秒で切り出したデータに、環境音の最大振幅を一致させて混合した後、スペクトログラムに変換した。図 1 に例を示す。

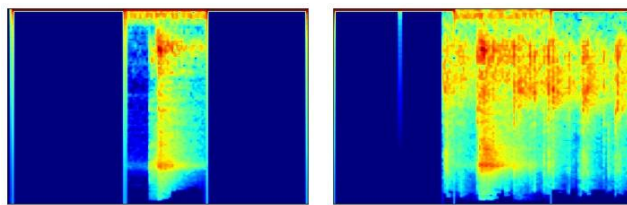


図 1 木材破壊音 (左) と clapping をノイズとして混合した木材破壊音 (右) のスペクトログラム

4. 使用した CNN モデル

本研究では、著者らの先行研究と同じ CNN のネットワーク構造を用いた。すなわち、畳み込み層は 3 層 (Conv.1, Conv.2, Conv.3)、プーリング層は 2 層 (Pool.1, Pool.2)、その後 3 層からなる全結合層 (Fully Connected Layer, FC) で構成されている。

CNN モデルの学習データのうち、正解として状況 C の音源に環境音をランダムに選択して混合した 679 個の混合音データを用いた。また、不正解として、ESC-50 に収録されている環境音データ 1500 個を用いた。これ等を、エポック数 300、バッチサイズ 32 で学習した。

テストでは、録音環境 C 以外の木材破壊音に環境音をランダムに混合した 2,440 個の木材破壊音データと、学習にも用いた 1500 個の環境音データを用いた。

キーワード CNN, 機械学習, 木材破壊音, 環境音, スペクトログラム

連絡先 〒278-0022 千葉県野田市山崎 2641 東京理科大学野田キャンパス 構造工学研究室 E-mail: 7622529@ed.tus.ac.jp

5. 複数回の再現実験

表1 テスト結果

	loss	acc	recall	prec	f1
Cmix0	0.319	0.975	0.969	0.991	0.980
Cmix1	0.671	0.958	0.939	0.992	0.965
Cmix2	0.977	0.945	0.935	0.976	0.955
Cmix3	0.662	0.965	0.953	0.989	0.971
Cmix4	1.422	0.933	0.913	0.977	0.944
Cmix5	1.598	0.891	0.858	0.962	0.907
Cmix6	3.370	0.805	0.732	0.940	0.823

テストデータは固定し、学習する混合音に含まれる環境音がランダムに選択される条件で、学習+テストを7回行った。その結果を表1に示す。テストの結果を見ると、accuracyは80.5～97.5%, recallは73.2～96.9%となり、大きな差があることが分かる。そこで、Cmix0とCmix6で使用された混合音の中に、どの環境音が選択されたのかを調べることにした。

表2はCmix0とCmix6において、学習に使用された混合音に含まれる環境音の数を示している。これによると、50種類の環境音の中で学習された回数に5以上の差がある環境音の種類は27種であった。特に赤色のbrushing_teeth, thunderstorm, vacuum_cleaner, coughingは、Cmix6に含まれる数がCmix0より多かった。

表2 学習データの混合音に構成する環境音割合
(前 Cmix0, 後 Cmix6)

hen	12, 14	crickets	14, 13	clock_alarm	9, 13
mouse_click	19, 15	pouring_water	14, 8	door_knock	12, 8
sheep	17, 11	laughing	14, 15	car_dorn	8, 13
snoring	11, 7	water_drops	19, 14	crow	18, 15
toilet_flush	15, 8	insect(flying)	13, 7	engine	17, 20
crackling_fire	18, 10	pig	12, 19	breathing	12, 12
chainsaw	17, 15	keyboard_typing	13, 19	rooster	11, 17
crying_baby	9, 15	brushing_teeth	6, 24	cow	12, 13
airplane	14, 8	helicopter	10, 17	train	11, 14
church_bells	20, 15	clapping	13, 11	fireworks	13, 9
sea_wave	20, 14	sneezing	13, 11	vacuum_cleaner	9, 19
cat	14, 20	chirping_birds	17, 11	clock_tick	17, 12
can_opening	17, 13	rain	17, 9	frog	15, 16
glass_breaking	15, 8	thunderstorm	5, 17	hand_saw	12, 11
wind	19, 14	dog	17, 15	siren	15, 11
wood_creaks	10, 16	washing_machine	13, 13	drinking_sipping	16, 19
footsteps	10, 17	coughing	5, 14		

6. 環境音の影響実験

学習する環境音が精度に与える影響を調べるため、状況Cの木材破壊音に環境音を1種類だけ混合した

音を生成し、これをCNNモデルに学習させた。これをESC-50に含まれる50種類すべての環境音について行い、精度を調べた。テストデータは先の再現実験と同じである。

解析結果のうち、表2で色を付けた環境音を混合させた場合を表3に示す。赤色と水色を比べると、水色の方が赤色よりも精度が高いことが分かる。6章の再現実験では、Cmix6のモデルに学習に対して不利なデータ（赤色）が多く含まれたため、モデルの学習が上手くいかず、精度が低下したと考えられる。

表3 ノイズとして混合する環境音による精度の違い

順位	C+	loss	acc	recall	prec	f1
5	toilet_flush	57.3	0.71	0.54	0.98	0.70
6	crackling_fire	10.8	0.69	0.51	0.97	0.67
18	brushing_teeth	14.6	0.59	0.36	0.94	0.53
34	glass_breaking	49.3	0.52	0.24	0.94	0.38
40	coughing	27.5	0.51	0.22	0.95	0.36
43	rain	76.5	0.49	0.19	0.97	0.31
44	thunderstorm	157.5	0.49	0.18	0.96	0.30
46	vacuum_cleaner	73.1	0.47	0.17	0.90	0.29

7. まとめ

本研究では、木材破壊音のノイズとして混合する環境音の違いが、CNNモデルの学習効果に与える影響を調べた。その結果、学習に適したデータとそうでないデータが存在する可能性を示した。

今後は環境音と木材破壊音の音圧のデシベルを調べ、環境音の音圧を調整し、木材破壊音認識に与える影響を検討する。

参考文献

1) 竹森司, 佐伯昌之：木材破壊音検出を目的としたNNとCNNの比較検討, 土木学会論文集 A2 (応用力学), Vol.77, No.2(応用力学論文集 Vol.24), I_535-I_542, 2021.

2) 小野真由子, 竹森司, 佐伯昌之：CNNによる木材破壊音検出のためのデータ拡張手法の検討, 土木学会 第25回応用力学シンポジウム, 2E01-05-03, オンライン, 2022

3) K. J. Piczak :ESC :Dataset for Environmental sound Classification, Proceeding of the 23rd Annual ACM Conference on Multimedia, Brisbane, Australia, pp. 1015-1018, 2015