

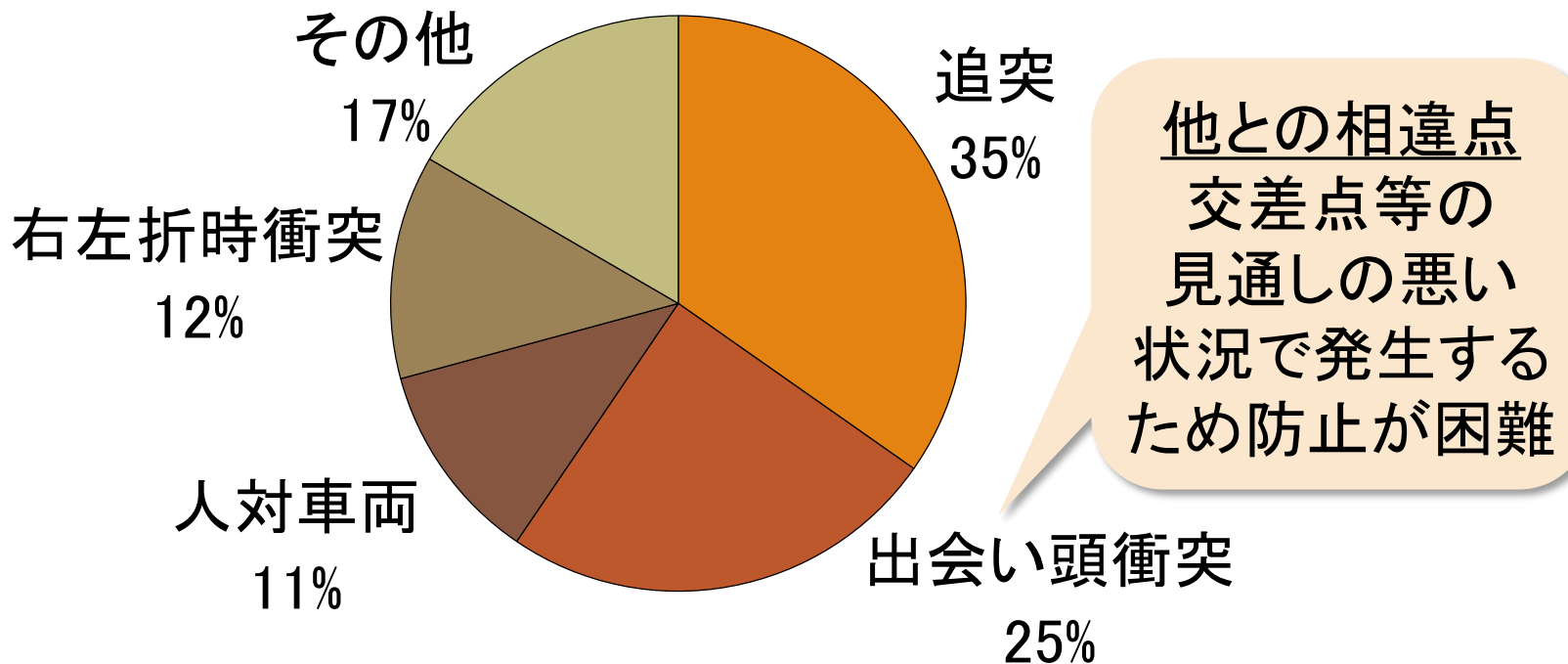
ニューラルネットワークに基づく 道路環境音を用いた 接近車両検出手法に関する研究

情報工学専攻 修士課程 旭研究室

183426017 平田秀平

背景：出会い頭事故防止の重要性

- 出会い頭衝突事故は2番目に多く、その防止は優先課題



平成30年 類型別事故件数*

目指しているシステム

接近する他車を音で検出するシステム

1. 自車にマイクを設置
2. 走行中リアルタイムに周辺環境音を録音
本研究の対象
3. **環境音から接近する他車の有無を判定**
4. 接近する他車が認識された時点で
運転者に警告

先行研究/関連研究

1. 走行音の到来方向をクラスタリングし閾値処理で接近車両の有無を判定する手法*¹
2. 走行音のレベルもしくはレベル上昇度の閾値処理で接近車両を検知する手法*²
3. 走行音の到達時間差の軌跡のフィッティングにより連続車両通過を検出する手法*³

*1 旭健作, 坂野秀樹, 山本修身, 小川明, 山田啓一, “音響による複数到来車両の検出手法の構築と評価”, 電子情報通信学会技術研究報告. ITS., vol. 2010-78, pp. 65-68, 2011.

*2 星野博之, “ドライバ支援のための走行音による接近車両検知システム”, 日本音響学会誌, vol. 62-3, pp. 265-274, 2006.

*3 石田繁巳, 三村晃平, 劉 嵩, 田頭茂明, 福田 晃, “路側設置マイクロフォンによる車両カウントシステム”, 情報処理学会論文誌, vol. 58, No. 1, pp. 89-98, 2017.

本研究の概要

- 道路環境音のスペクトログラムから以下の3状態を判別する判定器の構築・評価
 - 判定器は畳み込みニューラルネットワークを用いる

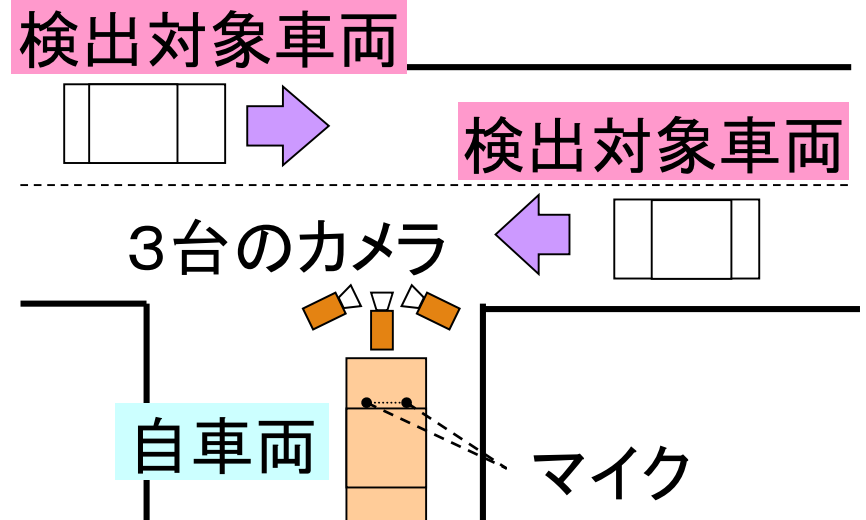
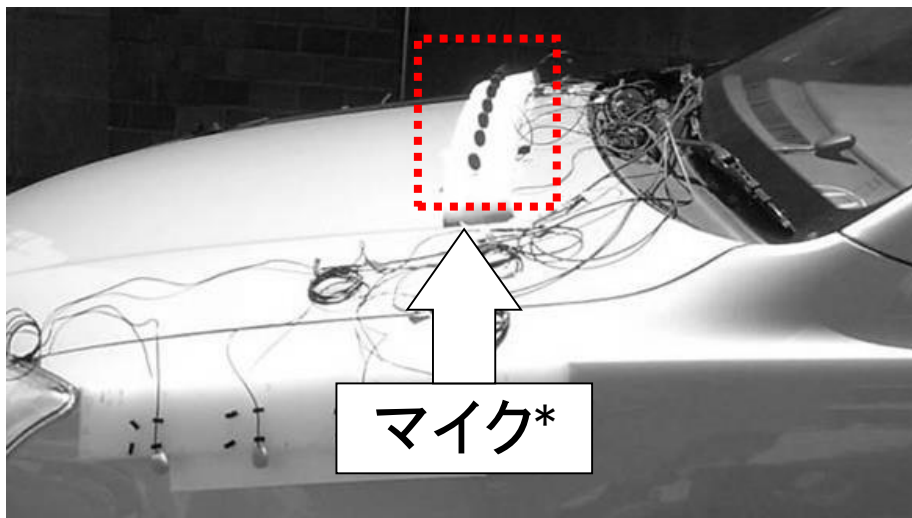
接近	他車が自車に近づいている状態
離反	他車が自車から遠ざかっている状態
なし	他車が自車付近に存在しない状態

道路環境音とは

- 道路上で収録された音のこと
 - 路上では常に風の音や歩行者の歩く音等が発生
 - 車両が通過すると上記に加えて走行音が発生
- 車両走行音の構成要素
 - **ロードノイズ(タイヤと路面の摩擦による音)**
 - エンジン音
 - 風切り音
 - その他

道路環境音の収録

- マイクとビデオカメラで道路環境音と周辺映像を収録
 - 2車線直線道路の脇にセダン車を停車し、そのボンネット上にマイクを設置
 - 走行状態を模擬するためセダン車はアイドリング



* マイクは複数あるが、今回使用したのは中央のマイク1つのみ

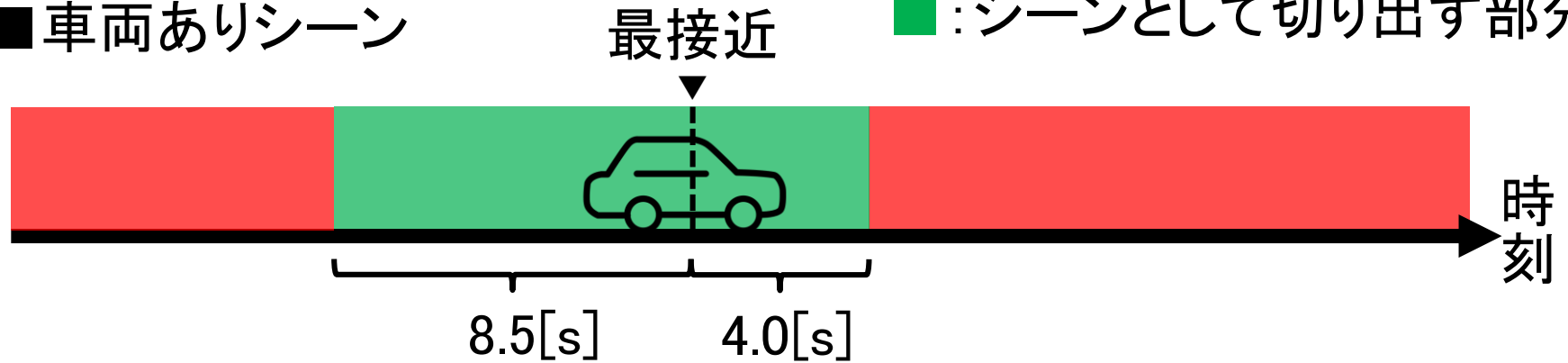
シーンの切り出し(1)

- 道路環境音は車種などのメタデータと共にシーンという単位で収録
 1. 車両ありシーン
 - 収録音全体から車両が走行している部分を切り出したもの
 2. 車両なしシーン
 - 収録音全体から路上に車両が存在しない部分を切り出したもの
 - 風の音などが含まれる

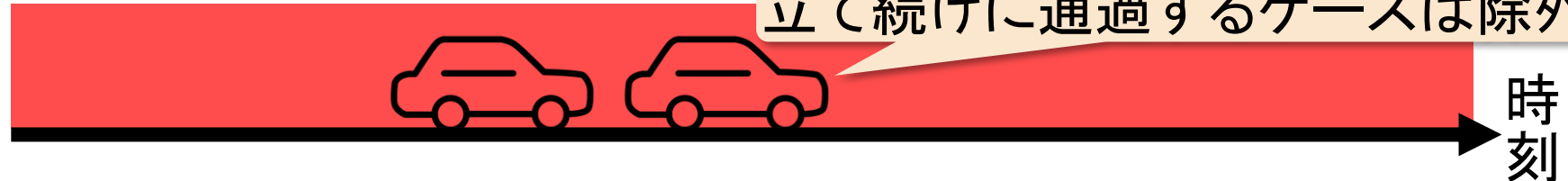
シーンの切り出し(2)

■ 車両ありシーン

■ : シーンとして切り出す部分

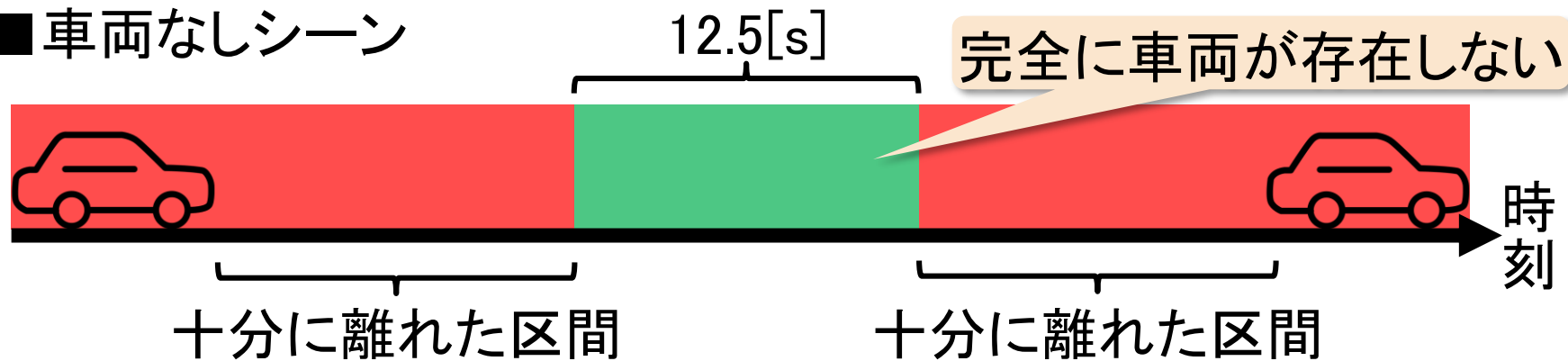


立て続けに通過するケースは除外



■ 車両なしシーン

完全に車両が存在しない

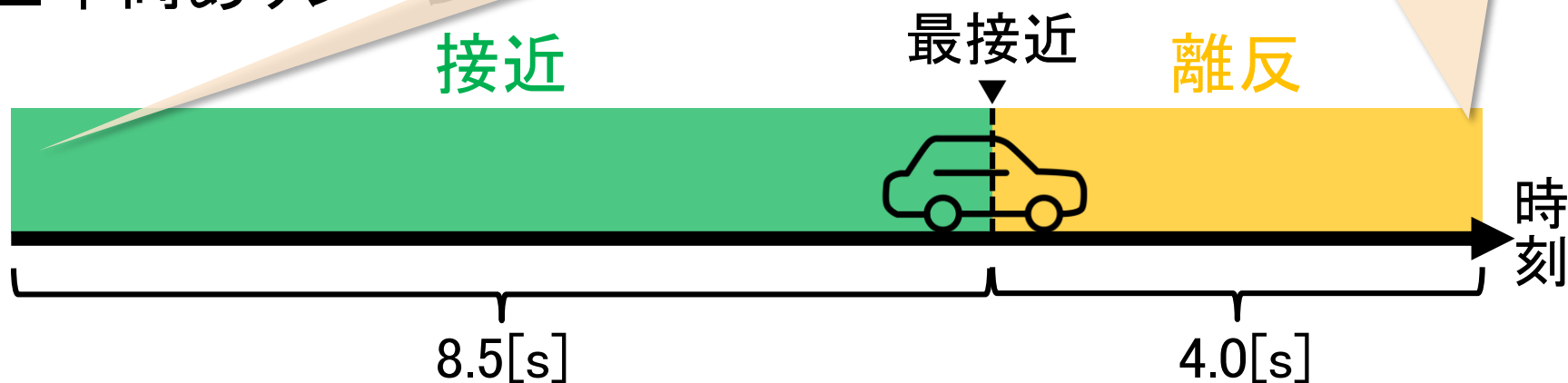


データのラベリング

接近, 離反, なしの3ラベルを付与

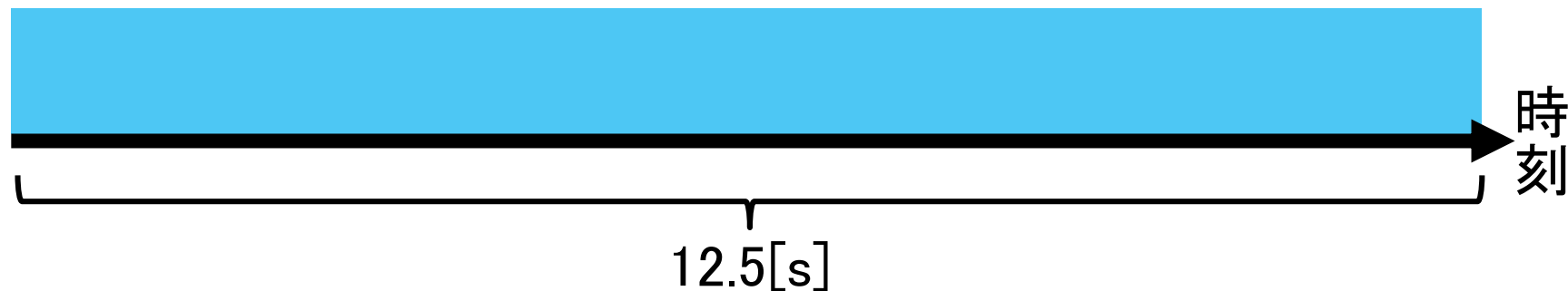
最接近遠方でも, 実際に車両の接近/離反を認識できるか否かに関わらず一律にラベル付け

■ 車両ありシーン



■ 車両なしシーン

なし



車両ありシーン(普通車)の例

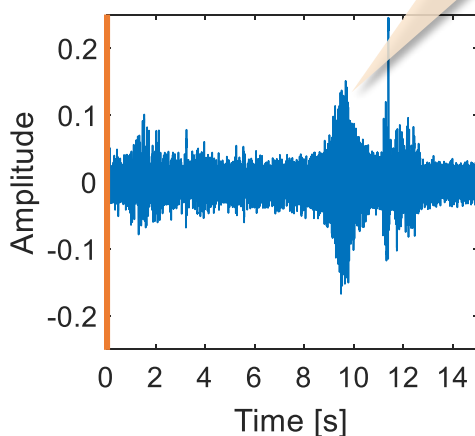


最接近付近で
ピーク

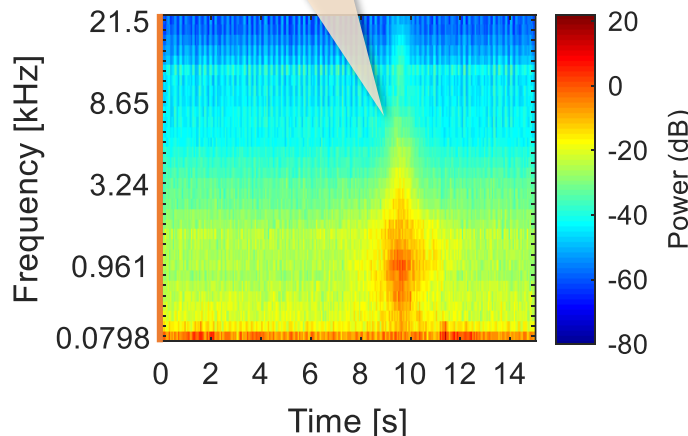
左方カメラ

正面カメラ

右方カメラ



音波形



メルスペクトログラム

フレーム分析条件

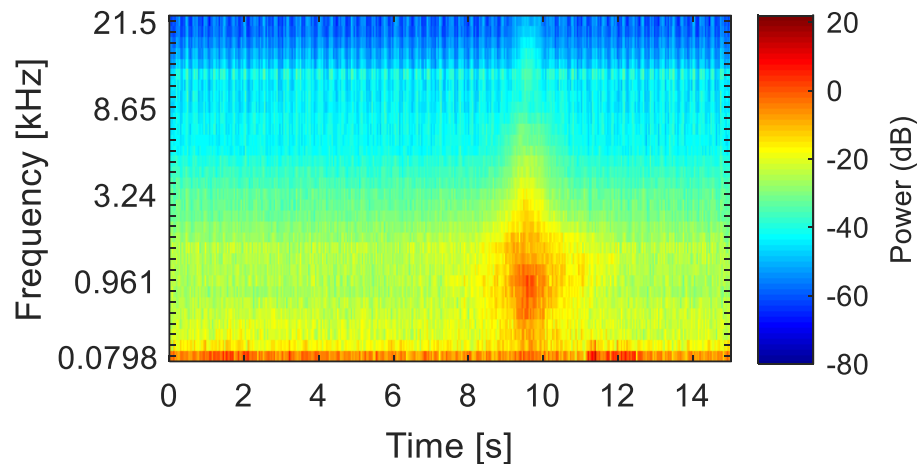
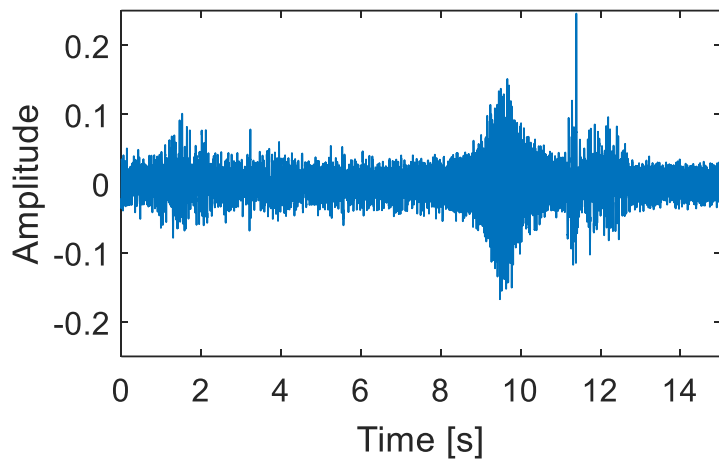
サンプリング 周波数[kHz]	48
フレーム幅[ms]	30
フレームシフト[ms]	15
窓	ハミング
メルバンド数	32

車両あり, 車両なしシーンの例

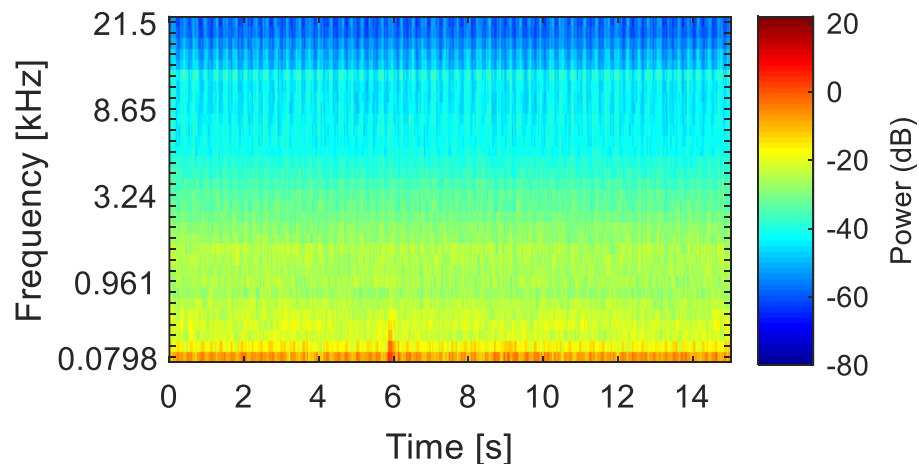
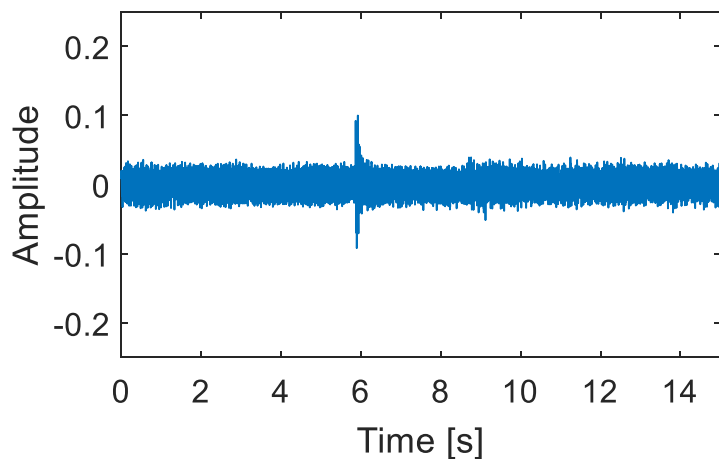
音波形

メルスペクトログラム

車両あり

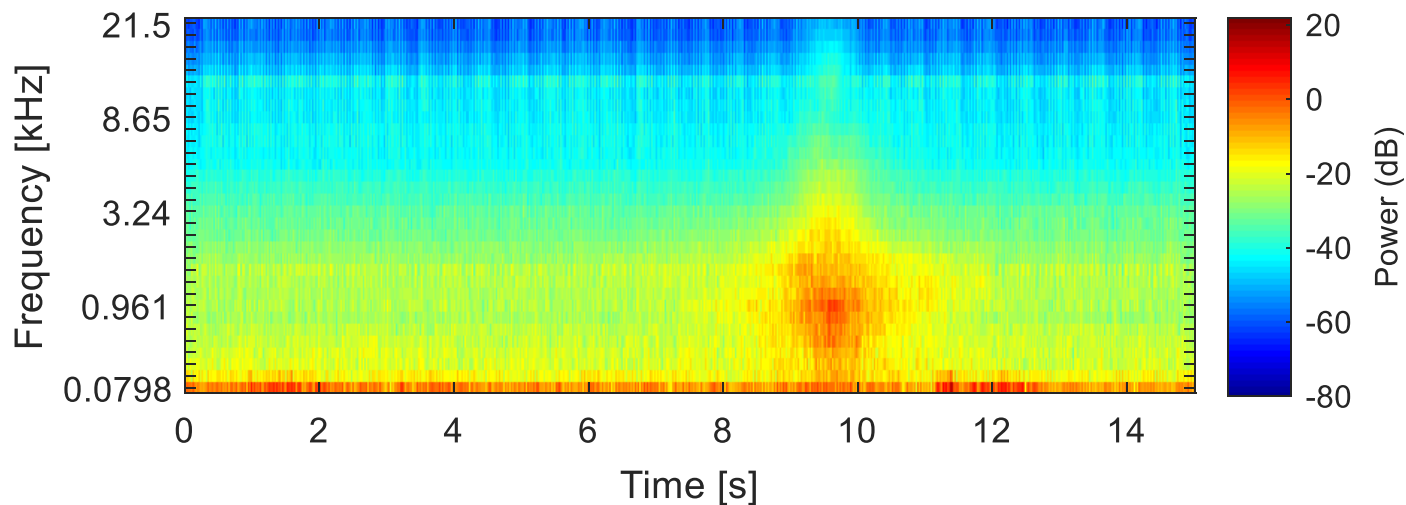


車両なし

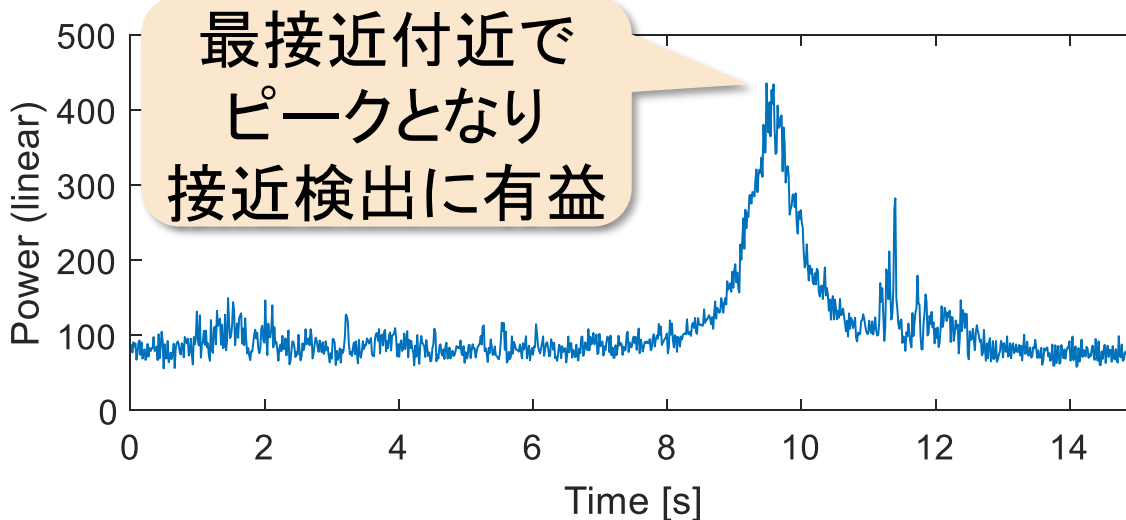


パワー情報

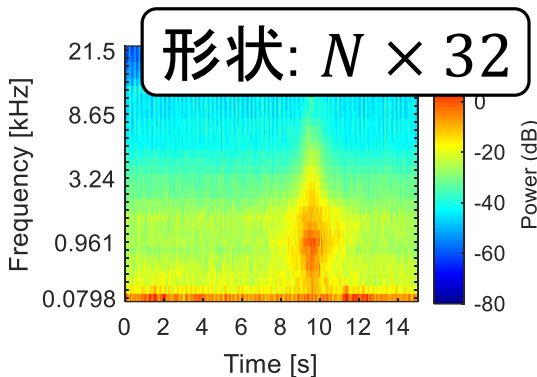
スペクトログラム
メル



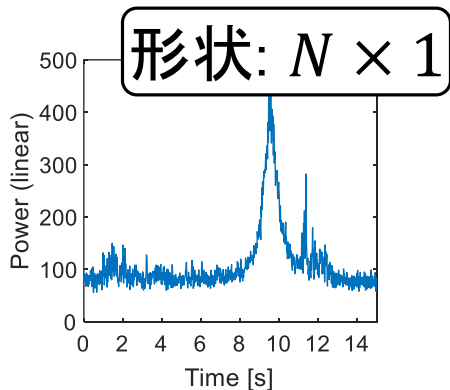
(線形スケール)
パワー



判定器の入カデータ形式



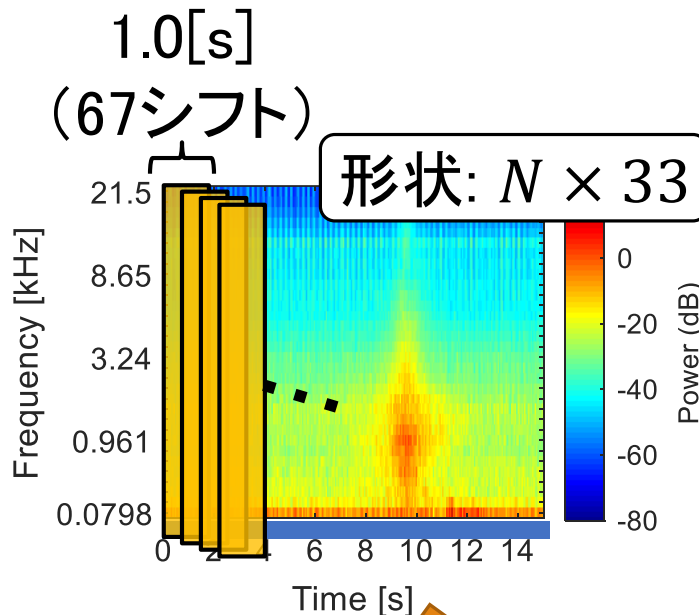
メルスペクトログラム



パワー(線形スケール)

N : シーンのスペクトログラムの時間軸方向のサイズ

連結



中心時刻を
入力データの
時刻とする

1.0[s]分を15[ms]
シフトで取り出す

形状: 67×33

判定器のネットワーク構造

以下の構造の畳み込みニューラルネットワーク

層	種類	数	サイズ	活性化関数	ドロップアウト率
1	Conv2D	32	3×3	ReLU	—
2	Conv2D	32	3×3	ReLU	—
3	Max-Pool	—	2×2	—	0.25
4	Conv2D	64	3×3	ReLU	—
5	Conv2D	64	3×3	ReLU	—
6	Max-Pool	—	2×2	—	0.25
7	Dense	512	—	ReLU	0.50
8	Dense	3	—	Softmax	—

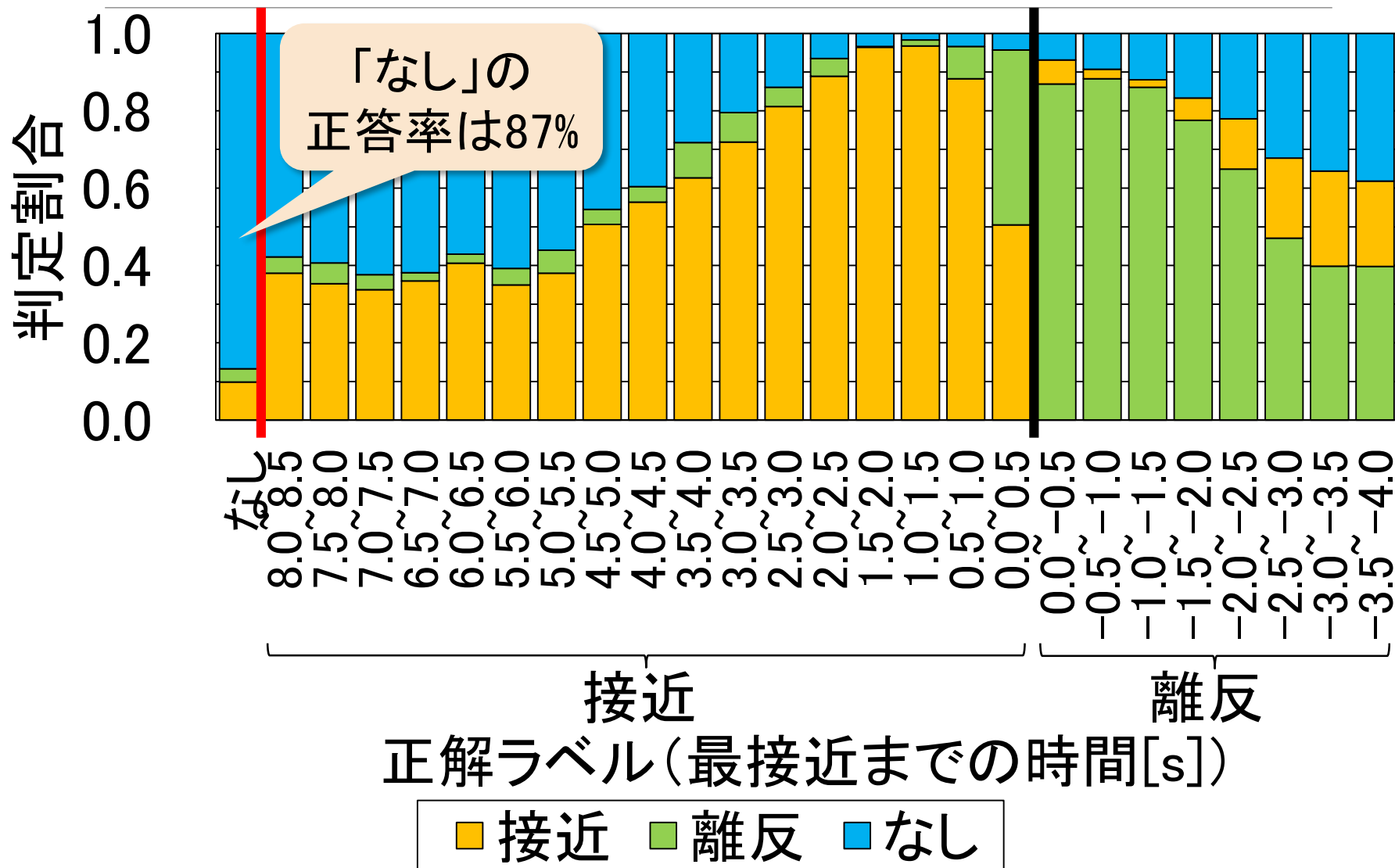
訓練と評価

- 訓練方法
 - 下表の訓練条件で3分割交差検証
- 評価方法
 1. 検証用データを0.5[s]ごとの時間区間に分けて判定器へ入力
 2. それぞれの区間での判定割合を求める

訓練条件

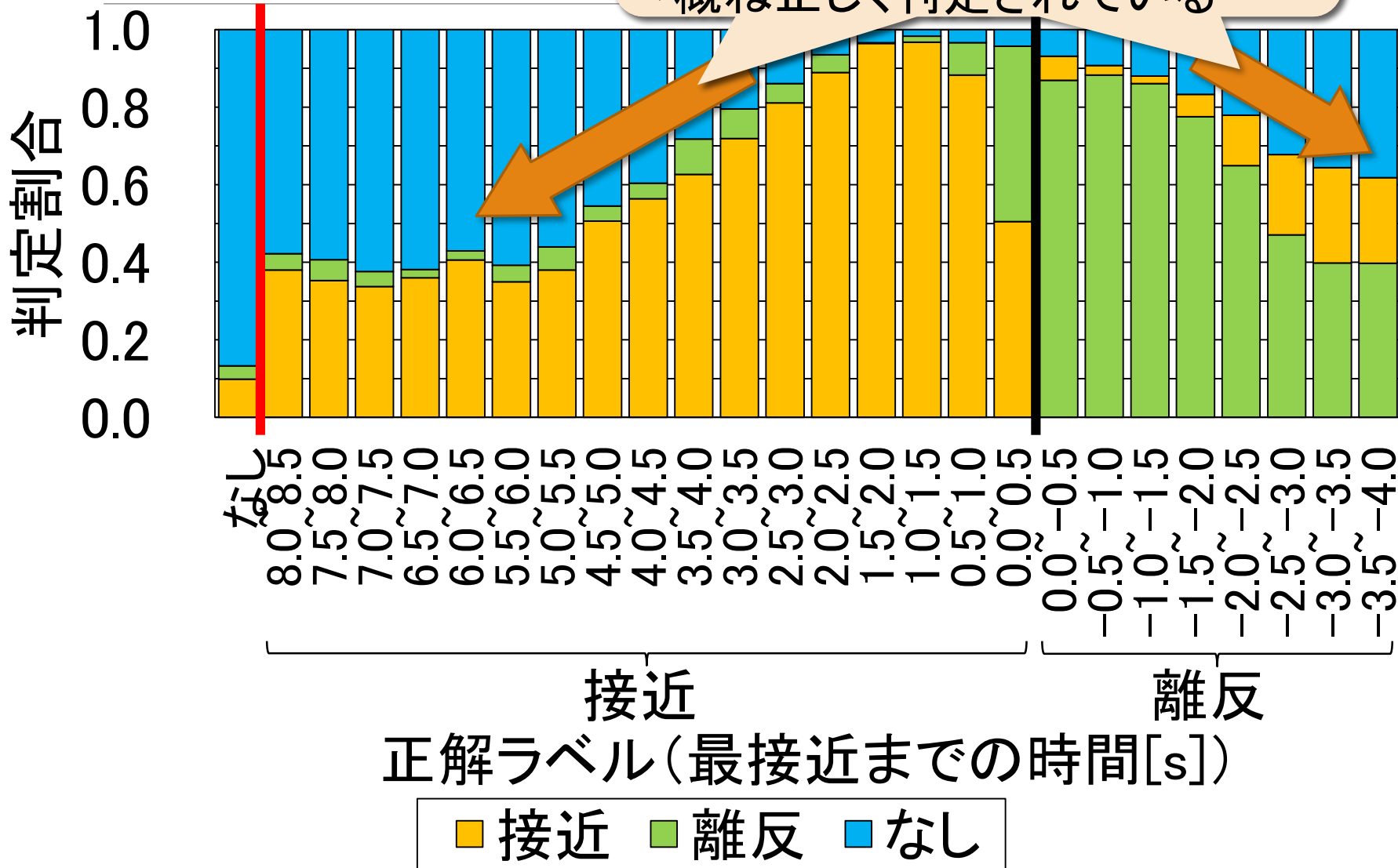
検証用データ割合	33%
最適化手法	Adam
学習率	1.0×10^{-6}
損失関数	多クラス交差エントロピー

判定結果 (3分割分の平均)



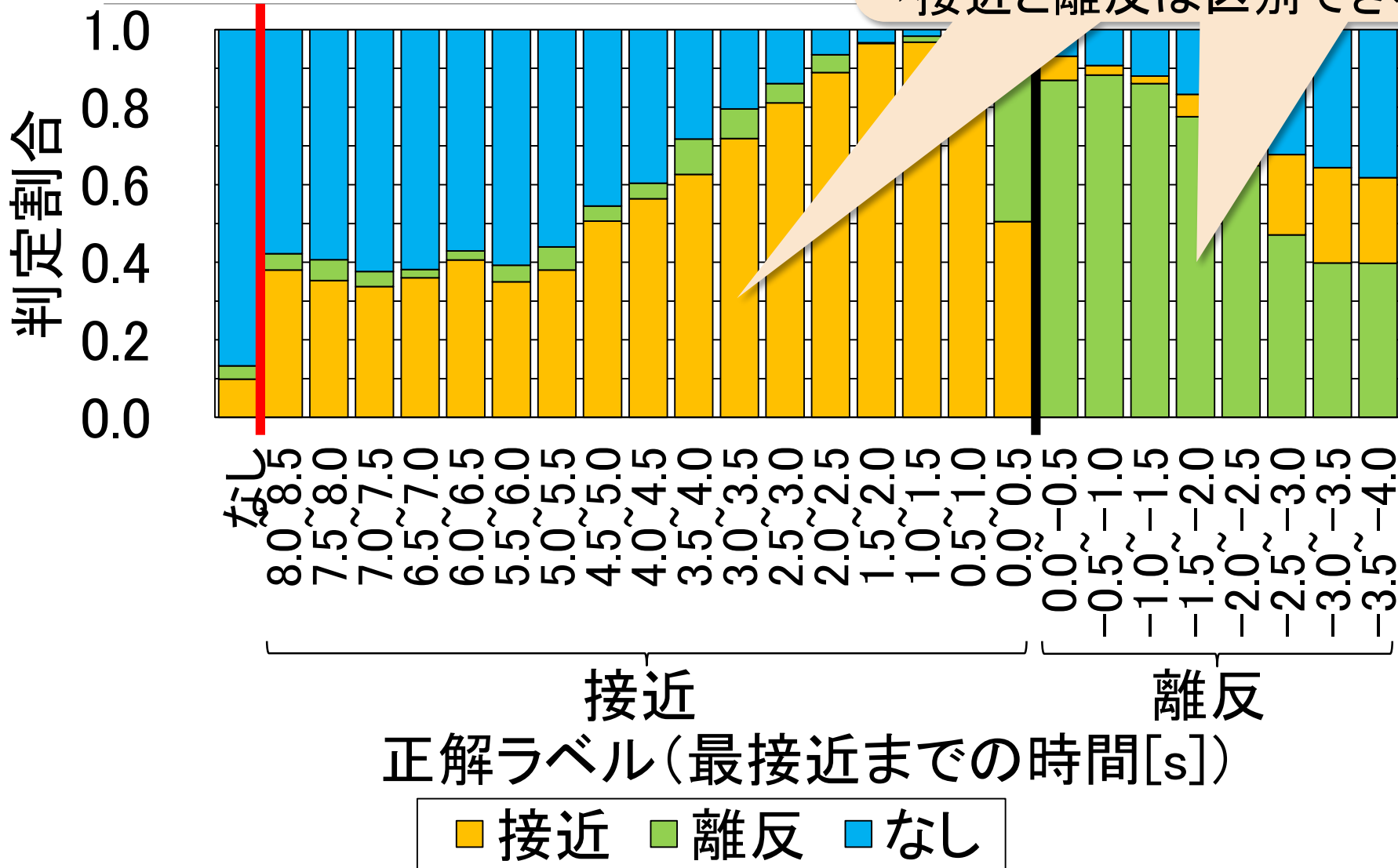
判定結果 (3分割)

最接近から遠ざかるにつれて
「なし」に誤る割合が増加している
→概ね正しく判定されている



判定結果 (3分割分の平均)

概ね接近は接近, 離反は
離反に判定されている
→ 接近と離反は区別できる



先行研究との比較

- 先行研究よりも検出率が高いことが示された

		先行研究1	本研究
用いる情報		到来方向 (CSP係数)	スペクトログラム とパワー
判定方法		クラスタリング と閾値判定	CNN
接近離反の区別		不可	可
検出率[%]	0[秒前]	約93	96*
	1[秒前]	約90	98*
	2[秒前]	約75	94*
	5[秒前]	約61	44*

提案手法による衝突回避の有効性

走行速度 [km/h]	必要な 停車時間[s] ※	左記時間以内に 接近と判定される 割合[%]
7	1.0	97
20	1.5	96
32	2.0	89
44	2.5	81
57		72
69		63

徐行の一般的な速度

※ 交差点等では車は

徐行していることが多い

※ 次式で反応時

(乾燥路面)として

走行速度[m/s]

$$\text{停止時間[s]} = \text{反応時間[s]} + \frac{\text{走行速度[m/s]}}{\text{重力加速度[m/s}^2] \times \text{摩擦係数}}$$

まとめ

- 畳み込みニューラルネットワークにより, 道路環境音から車両接近, 車両離反, 車両なしのいずれかを判別する手法を検討した
- 他車との衝突回避のために必要な停車時間よりも十分な余裕をもって他車の接近を検出できることを示した
- 車両接近と車両離反を音響上正しく区別できることを示した