

音とAIがひらく身体機能評価の最前線

～ 食べる音から働く音まで ～

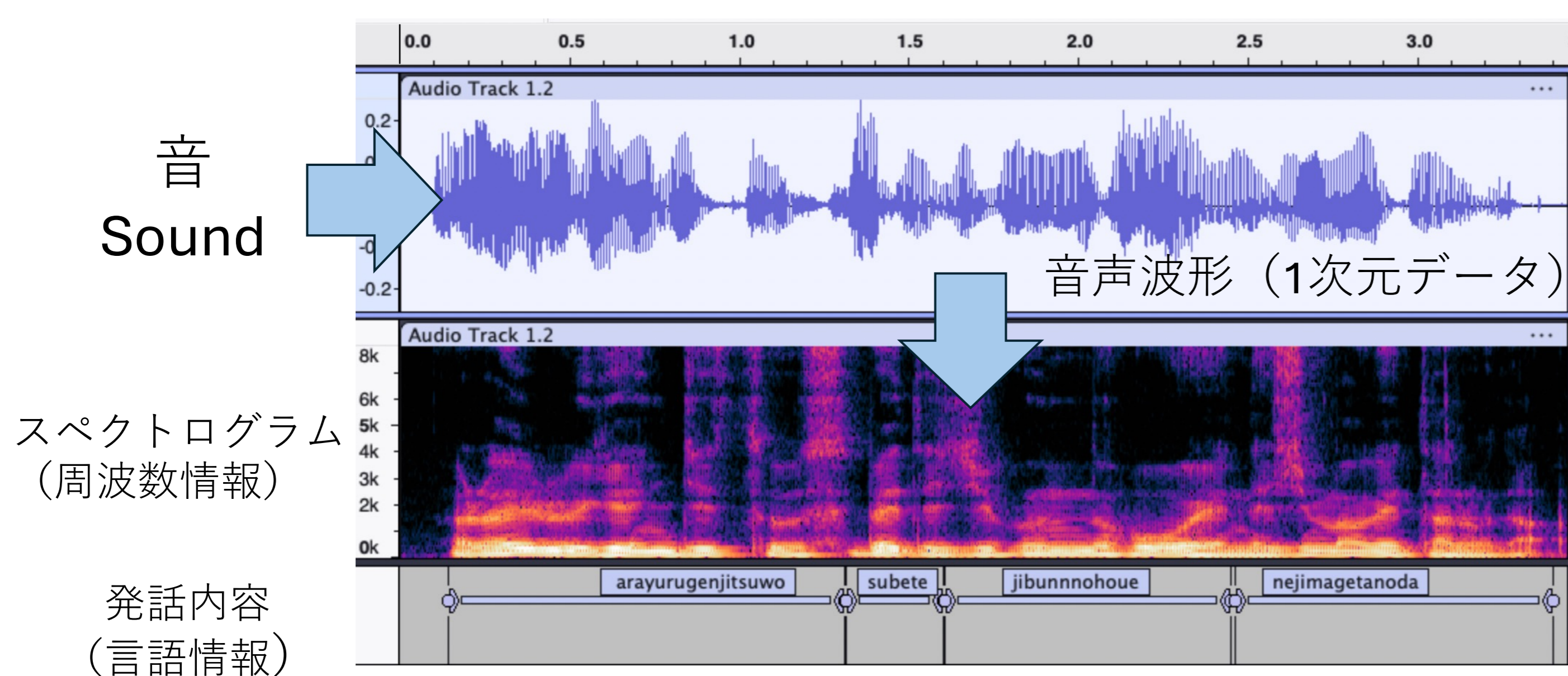
造形学部 スマートデザイン学科 西村 雅史

研究の概要

音は画像などに比べると、とてもサイズの小さい1次元のデータに過ぎませんが、その時間や周波数(スペクトログラム)には非常に多くの情報が含まれていることが知られています。

例えば音声は言語情報を伝達するだけでなく、様々な非言語情報を含んでいます。声のトーン、速度、大きさ、間の長さ、ため息など、話者の感情、態度、意図、心的状態などを反映した情報や、性別や個性といった情報も含まれます。

また、体が出すその他の生体音(呼吸音、心音、嚥下音など)も医療分野における基本情報として活用されています。最近では機械の異常診断や人の行動分析にも音情報が活用されるようになってきました。



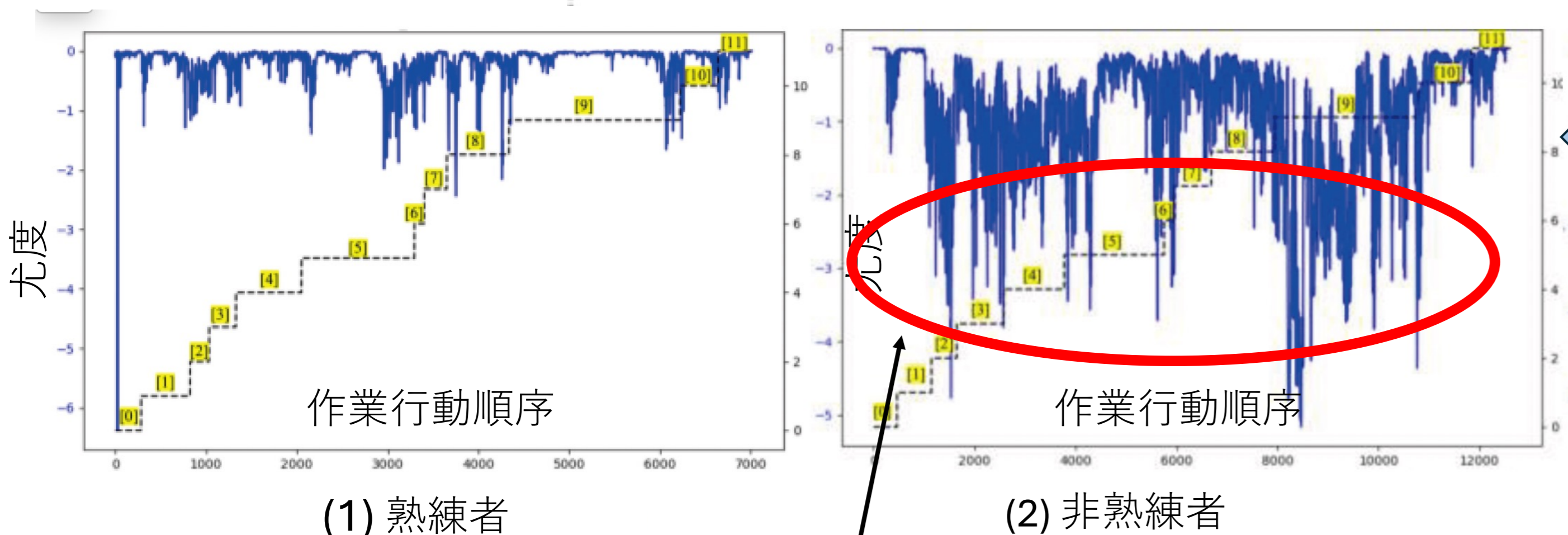
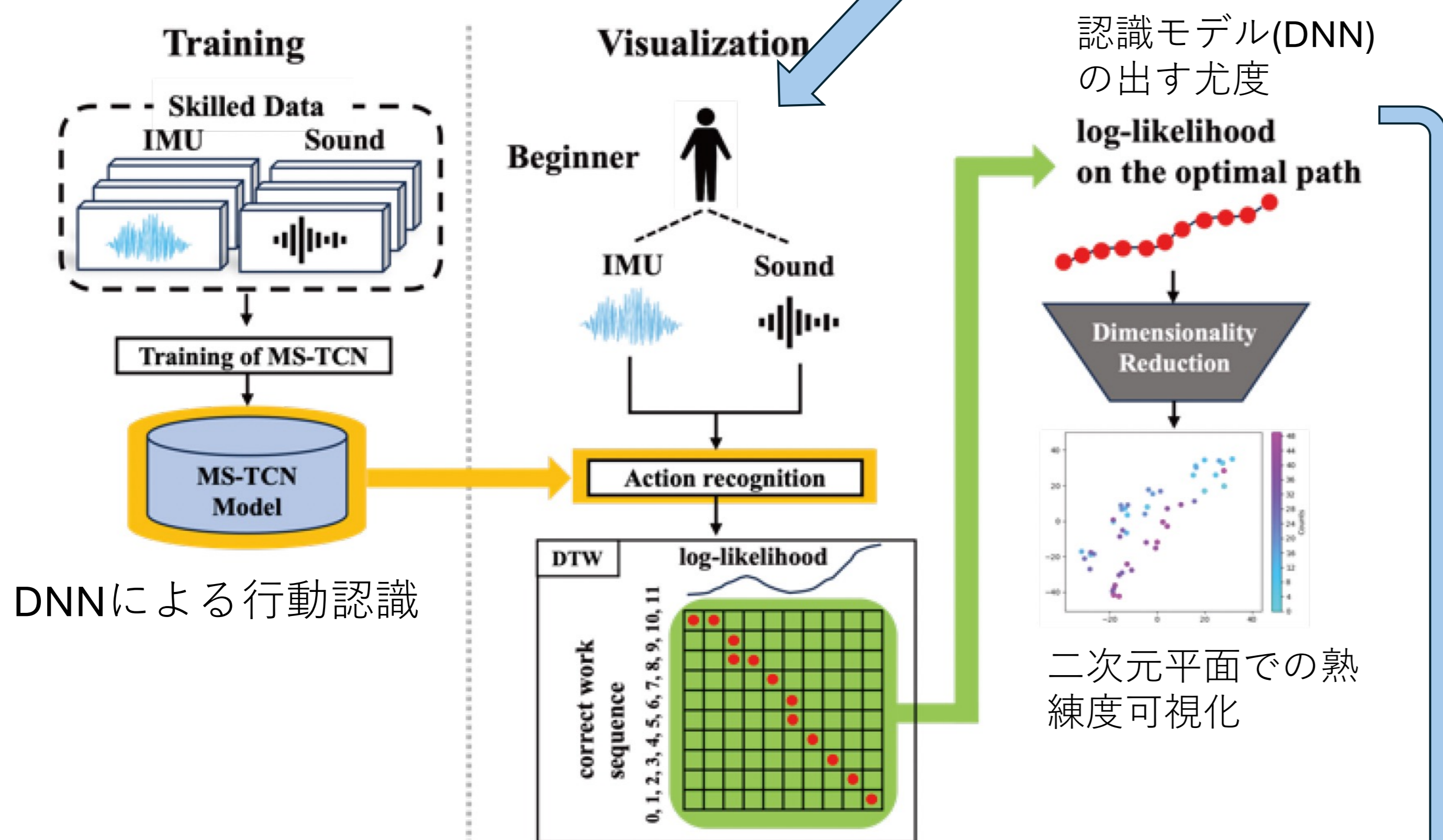
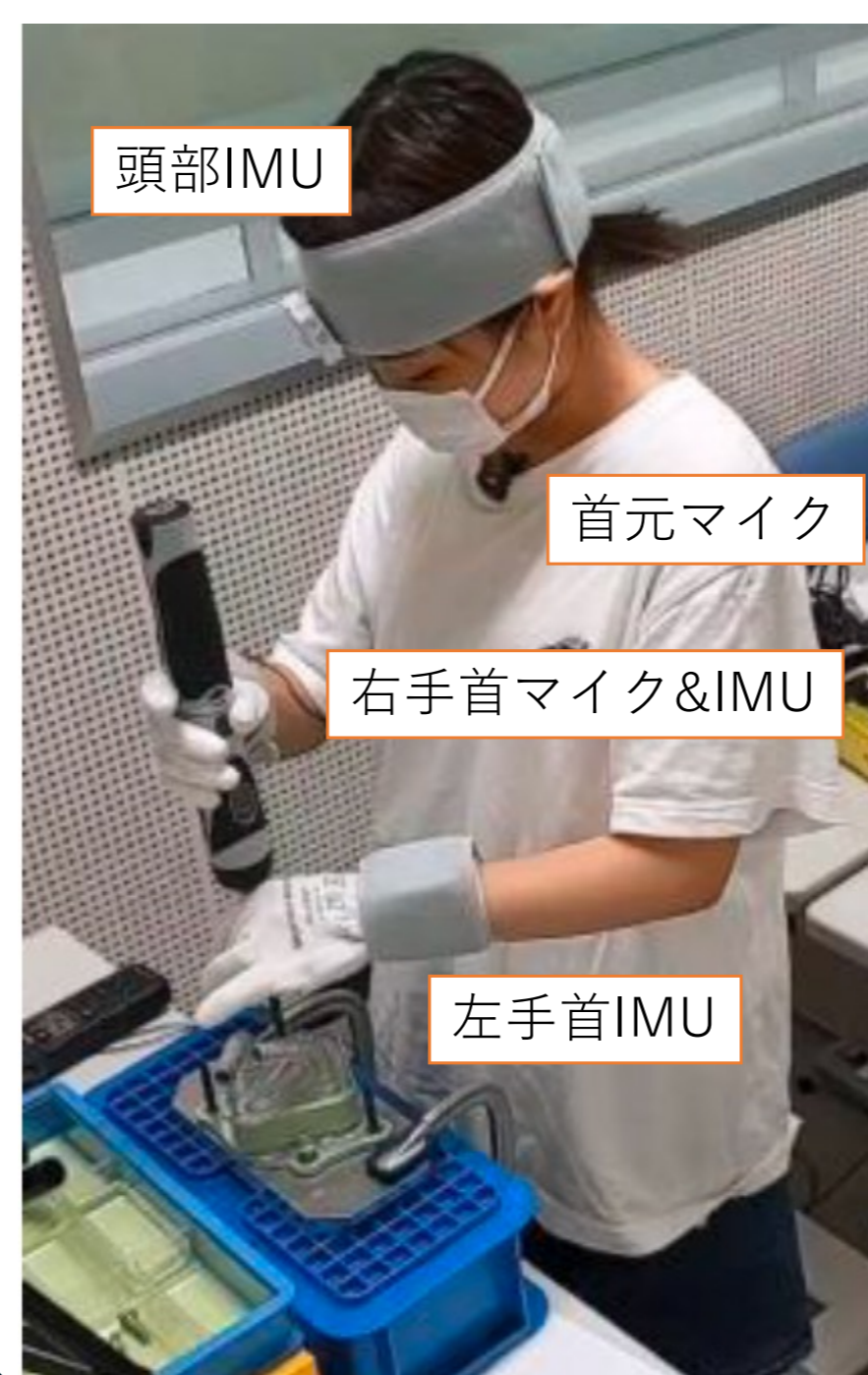
音でわかる作業の質や異常[1]

・異常音検出:機械の異常状態を音で検出する試みは古くから行われており、特に現在ではAIによる自動化が進んでいます。

・作業音による作業の質の評価:ボルトを刺す時の音や電動ドライバーの動作音等から、作業者の熟練度を推定できることがわかっています。右図ではIMU(慣性計測装置)と併用したケースを示していますが、精度的には音情報だけでもIMUと同等の識別性能が得られています。

下の図は、DNNによる自動識別とその時に得られるモデルの尤度が作業者のスキルレベルによって大きく変化することを示しています。

(ヤマハ発動機, 静岡大との共同研究)



作業間違いがあった場所や非熟練者では尤度の大幅な低下が見られる

食べる音からわかる口腔機能の低下

日本では高齢者人口が総人口の30%以上を占めるようになりました。

活力ある社会を維持するため、高齢者の健康寿命の延伸が喫緊の課題となっています。

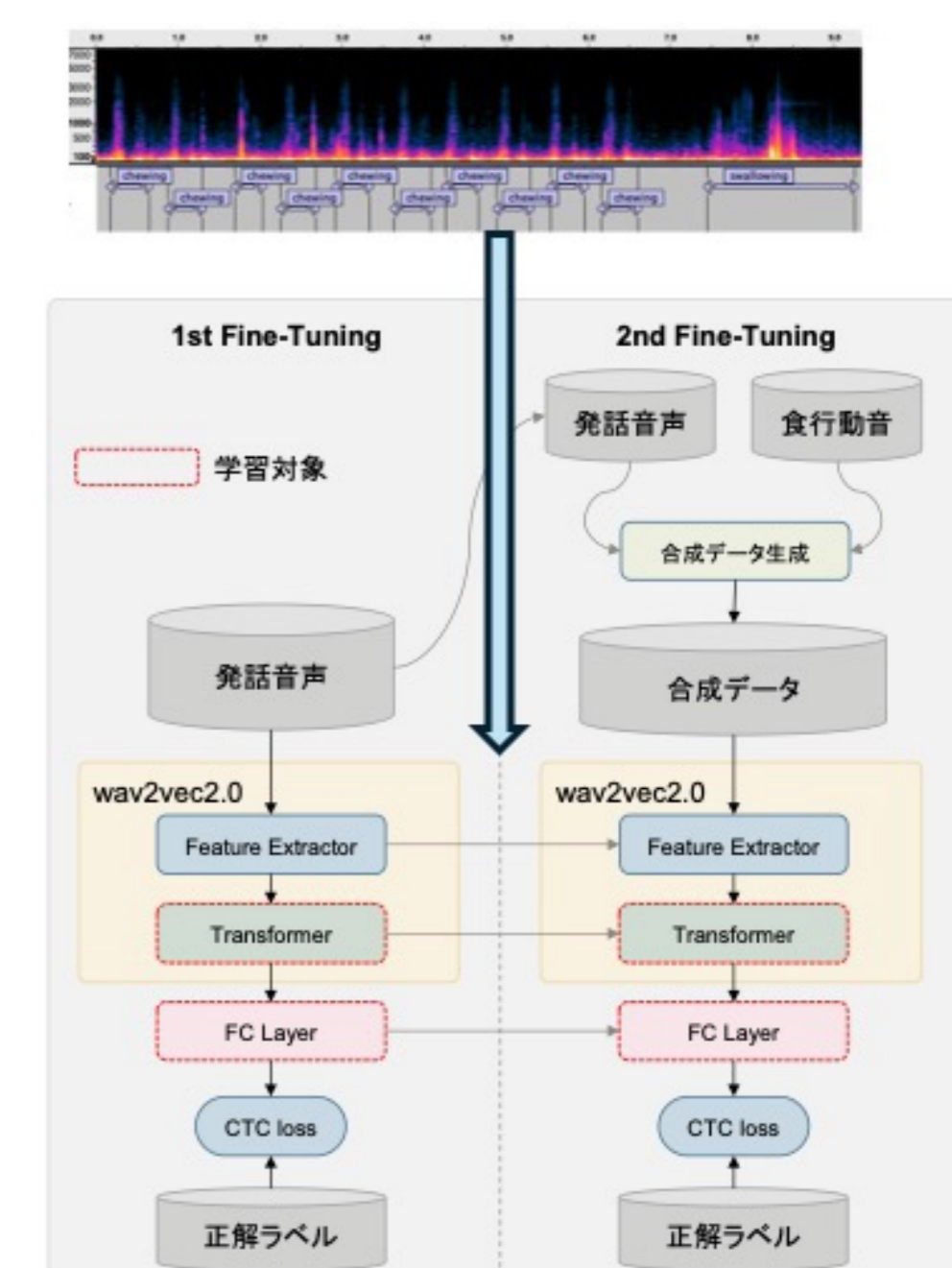
特に、「食べる」機能を正常に保ち続けることは健康寿命の延伸に直接関わる重要なテーマです。我々はこの課題に対処するため、生体音情報を利用し、口腔機能の低下を簡便に検出したり、長期的にモニタリングするためのシステムを開発しています。(千葉大, 広島大, 静岡大, 神戸大, 君津中央病院との共同研究)



生体音収録用 皮膚接触型マイク

① 食行動の自動認識[2]

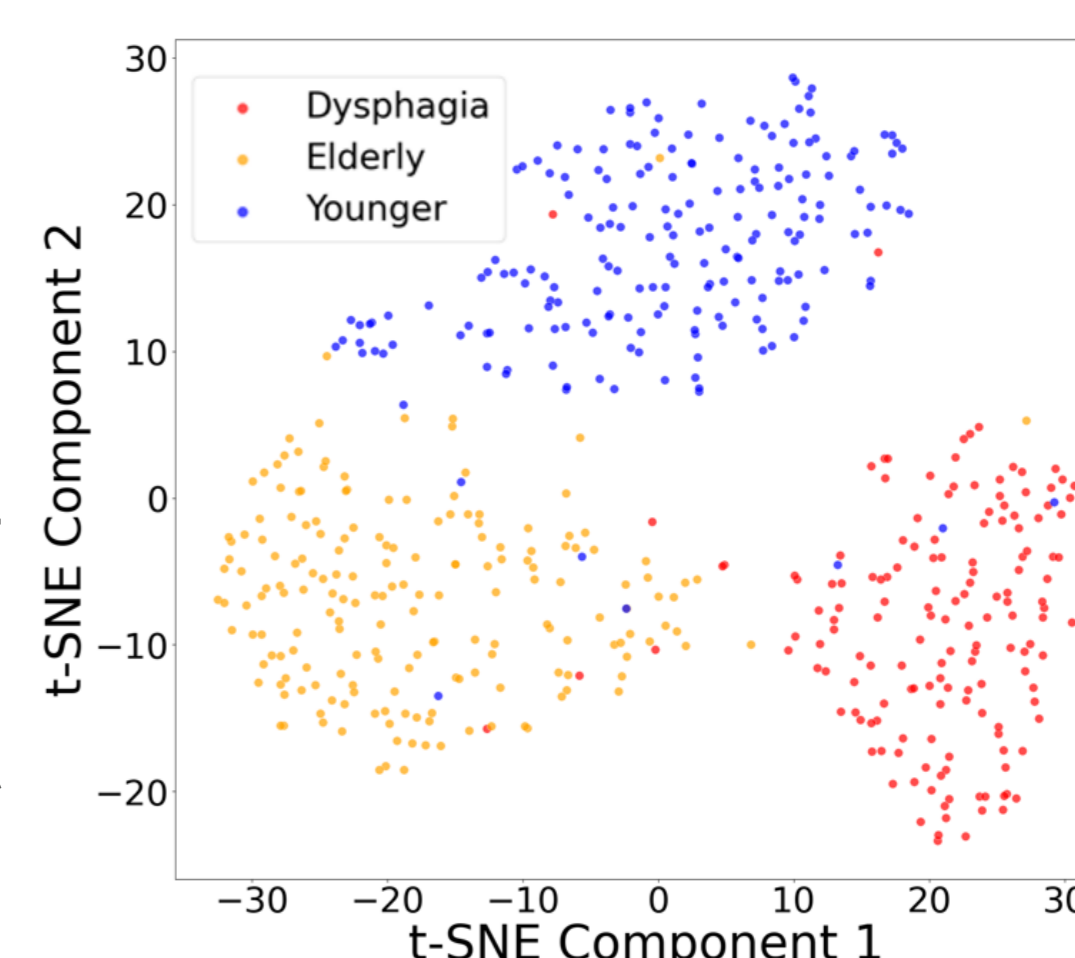
SSL(自己教師あり学習モデル)と呼ばれる、数万時間単位の音声データで学習されたDNNをベースモデルとし、ターゲットとする生体音によるFine-Tuningを繰り返し適用することで、咀嚼、嚥下、音声といったイベントの検出性能は90%を遥かに超え、既に十分な認識性能が得られています。



食行動認識モデルの学習フロー

② 嚥下機能の自動評価[3]

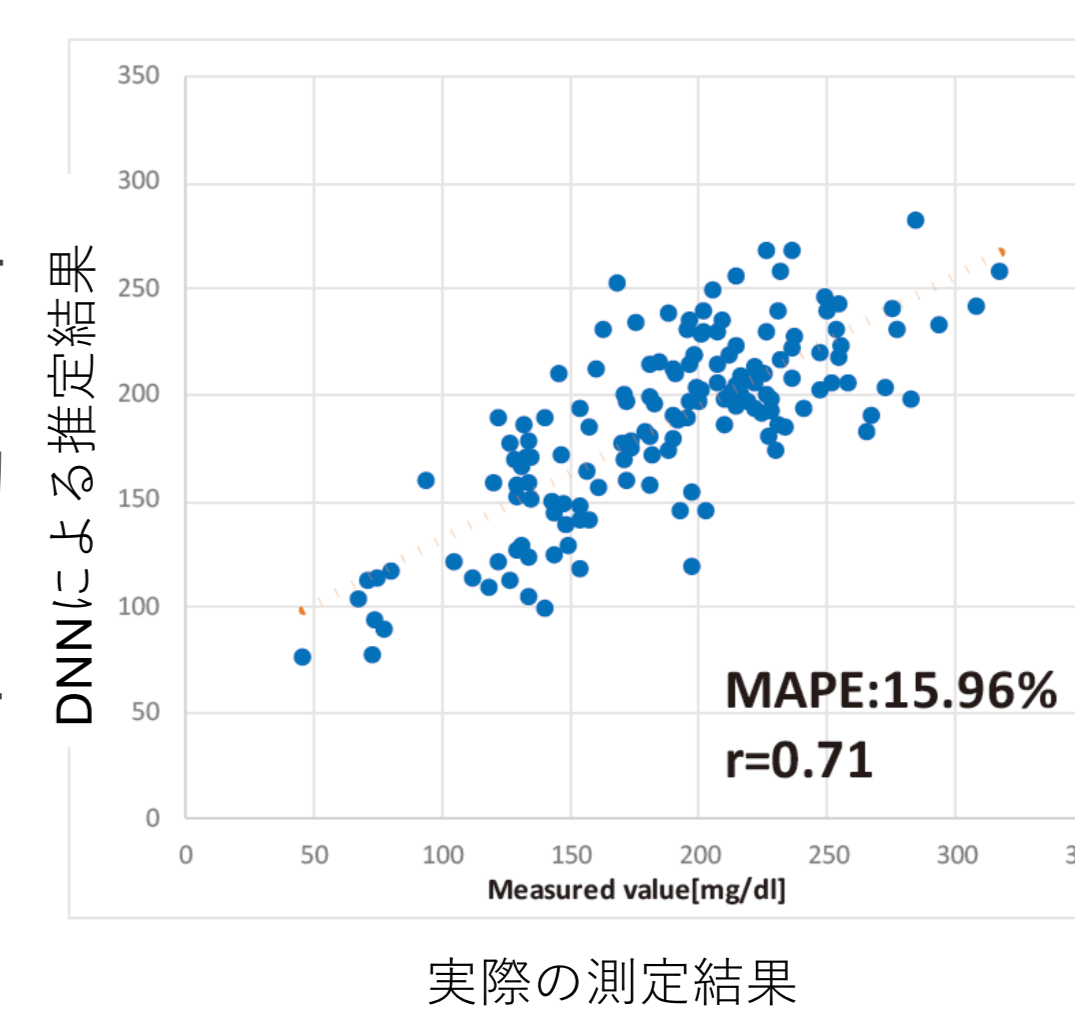
医療の現場でも嚥下音を聴診することによる嚥下障害の診断が行われています。ただ、診断には熟練を要するため、その機能を自動化する試みを行なっています。右図は嚥下障害者と健常者の嚥下音の特徴をDNNで抽出し、可視化したものですが、それぞれの違いを明確に捉えていることがわかります。



若年者・健常高齢者・嚥下障害者の嚥下音の可視化結果

③ 咀嚼機能の自動評価[4]

認知機能の低下した高齢者に対して、咀嚼能力を正確に測定することは難しいとされ、ここでも咀嚼音に基づく、機能自動評価の検討を行っています。右図は音とDNNによる自動評価と、既存の評価方法との相関を表しており、咀嚼音だけで、かなり正確な推定が可能になりつつあることがわかります。



医療機器による推定結果と嚥下音とDNNによる推定結果の相関

話す声からわかるストレスや軽度認知症[5]

声にはストレスといった人の心的な変化が現れることが知られています。我々はうつ病性障害患者に対し、音声のスペクトル包絡、特にその低次周波数成分にその影響が現れることを見出しました。

また、発見が困難とされてきた軽度認知症についても声にその特徴が現れることが分かってきました。タブレットを用いた認知症検査時の音声や、対話ロボットとの雑談等でもその特徴が現れることを確認し、学会で報告しています。(筑波大, IBMとの共同研究)

[1] T.Hirade et al., "Visualization of the Proficiency Level of Assembly Work Using Inertial and Sound Sensors," Journal of Signal Processing, Vol.28, No.4 (2024.7)
[2] T.Tsukagoshi et al., Simultaneous Speech and Eating Behavior Recognition Using Data Augmentation and Two-Stage Fine-Tuning, Sensors, Vol.25, No.5 (2025.3)
[3] R.Ushijima et al., Visualization and Classification of Swallowing Difficulties Using Scalogram-Based CNN, Proceedings of NCSP'25 (2025.2)
[4] M.Takeuchi et al., Automatic Chewing Ability Assessment Using a Self-Supervised Learning Model, Journal of Signal Processing, Vol.29, No.4 (2025.7)
[5] K. Yoshi et al., Screening of Mild Cognitive Impairment Through Conversations With Humanoid Robots: Exploratory Pilot Study, JMIR formative research (2023.1)