

令和7年度 卒業論文概要					
所属	物理情報工学科・電子物理工学コース				
学生番号	222C3035	学生氏名	楠木寿季也	指導教員	小田部荘司
論文題目	超伝導リザバーコンピューティングによる音声および画像認識				

## 1. 諸言

近年、人工知能 (AI) の中でも機械学習の研究が盛んに行われている。物理リザバーコンピューティングは、AI ハードウェア分野で主軸となっている。現在、より低消費電力な物理リザバーの登場が望まれている。先行研究によって、超伝導体を用いたリザバーコンピューティングの時系列予測タスクへの有用性が示された<sup>[1]</sup>。しかし、より応用的なタスクである音声認識や画像処理などの分類タスクへの有用性はまだ示されていない。そこで本研究では、超伝導リザバーを用いて音声認識タスク、話者分類タスク、および画像認識タスクを行い、有用性を分類精度の観点から調査した。

## 2. 解析

Ginzburg-Landau 方程式に時定数を導入し、時間依存性を持たせた方程式を Time-Dependent Ginzburg-Landau 方程式 (以下、TDGL 方程式) という。規格化した TDGL 方程式は、

$$\gamma \frac{\partial \Psi}{\partial t} = (\nabla - i\mathbf{A})^2 \Psi - \alpha \Psi - \beta |\Psi|^2 \Psi \quad (2.1)$$

$$\tau_A \frac{\partial \mathbf{A}}{\partial t} = \text{Im}[\bar{\Psi}(\nabla - i\mathbf{A})\Psi] - \nabla \times \nabla \times \mathbf{A} \quad (2.2)$$

と記述される。式(2.1)、式(2.2)を Affine Integrator (AFI 法) ④で解くことで超伝導体内部の電界応答を数値的に算出した。

音声認識タスクおよび話者分類タスクでは、発生データから Mel-frequency cepstral coefficients<sup>[3]</sup>を用いて特徴量を抽出し、超伝導体に電流密度の時間変化として入力した。さらに、画像認識タスクでは、Modified National Institute of Standards and Technology (MNIST) データセットを利用し、0-9までの手書き数字画像を用いた。各数字の画像は28×28ピクセルのグレースケール画像で構成されており、前処理としてベクトル化および次元削減を行った後、超伝導体に同様にして入力した。このとき、超伝導体内部から50点の電界の変化を計算し、これをリザ

バーのデータとした。超伝導体のシミュレーションサイズは数百nm四方の正方形とした。

さらに、リザバー計算によって得られたデータを入力として、多クラス分類を行うためサポートベクターマシン (SVM) を用いて学習を行った。ここで、 $X$  をリザバーデータ、 $Y$  を教師ラベルとする。

SVM を用いて、リザバー出力に基づく分類器を学習した。線形 SVM の場合、以下の目的関数を最小化する。

$$\min_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (2.3)$$

ここで、 $\mathbf{w}$  は重みベクトル、 $b$  はバイアス項、 $\xi_i$  はスラック変数、 $C > 0$  は正規化パラメータである。また、非線形なクラスの境界を表現するため、カーネルトリックを用いた SVM を適用した。特に Radial Basis Function カーネル

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \quad (2.4)$$

を用いることで、入力特徴空間を高次元空間へ暗黙的に写像し、非線形分離を可能にした。

性能評価においては、テストデータをバッチ単位に分割し、各バッチ内の出力スコアを平均化した後にクラス判定を行った。得られた結果は混同行列として可視化し、正規化を施すことでクラス間の識別傾向を明らかにした。

このとき、超伝導内部のパラメータは変更しながら音声認識タスク、話者分類タスク、および画像認識タスクを行った。具体的には、ピン力に関する $\alpha$ を $-70 - 0$ 、ピンの数を $0 - 60$ まで変更し、各パラメータのときの精度を比較した。

## 3. 結果および考察

超伝導リザバーを用いた数字の音声分類タスクの混同行列を図1に示す。このとき、パラメータは $\alpha$ は $-10$ 、ピンの数は $10$ である。混同行列では、縦軸はテストデータの実際のラベルで横軸はモデルが予測したラベルを表す。実

際のクラスに対し、どのクラスがどの程度出力されたかを正規化した割合であり、対角成分は正しく分類された割合を示す。図1から音声数字認識では、一部の数字において、高精度で認識ができたが、発音の類似性を持つ数字間では、誤分類が残り、全体として中程度の分類精度となっている。

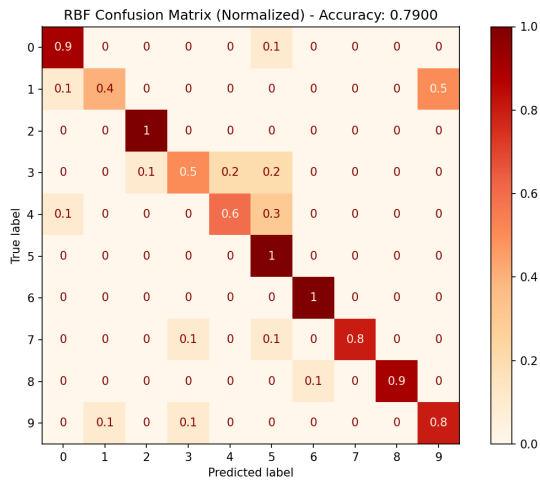


図1 音声数字分類モデルにおける混同行列

次に、話者分類タスクを行った結果を図2に示す。このとき、話者に数字1と発音させ、このとき、パラメータは $\alpha$ は-20、ピンの数は40である。図2から話者分類タスクでは多くの話者を高精度に識別できており、リザバーが話者固有の特徴を効果的に表現できていることが示された。

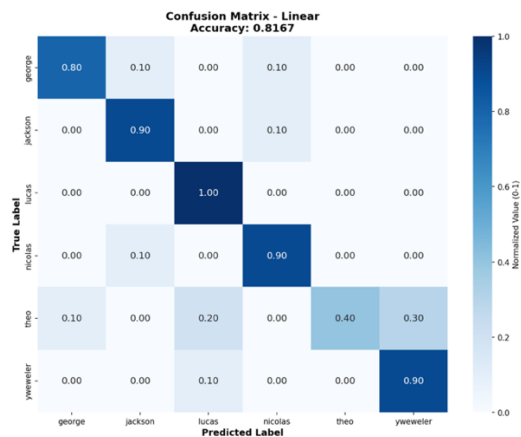


図2 話者分類モデルにおける混同行列

さらに、手書き文字認識を行った結果を図3に示す。このとき、パラメータは $\alpha$ は-20、ピンの数は40である。図3から数字によっては、高い認識精度が得られた一方で、形状が類似する数字間では誤分類が確認された。

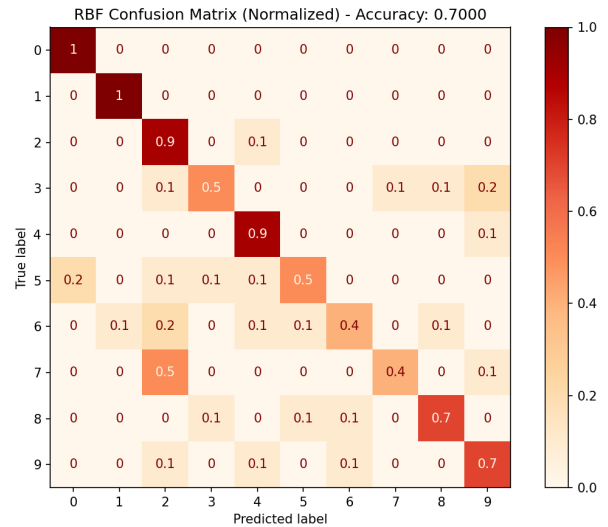


図3 手書き文字認識モデルにおける混同行列

#### 4. 結言

本研究では、超伝導体をリザバーとして用いた物理リザバーコンピューティングにより、音声認識、話者分類、手書き文字認識の各タスクを実施した。TDGL方程式をAFI法で数値的に解き、得られたリザバー出力をSVMにより学習した結果、各タスクにおいて良好な分類性能が得られた。

超伝導リザバーが入力信号に含まれる時間的・非線形的特徴を内部状態として表現できている。一方で、混同行列の解析から、クラス間の類似性が高い場合には誤分類が生じやすいことが確認された。

以上より、超伝導体を用いたリザバーは分類タスクにおいても有効であることが定量的に示された。

#### 参考文献

- [1] K. Arita *et al.*, *Physica C: Superconductivity and its Applications*, **662**, 1354522(2024),
- [2] T. Matsuno *et al.*, *J. Phys. Soc. Jpn.*, **89**(5), 054006 (2020).
- [3] A. Karacali *et al.*, *NOLTA, IEICE*, **15**(4), 871–882 (2024).

#### 研究実績

楠木寿季也 他、「超伝導体を用いたリザバーコンピューティングによる音声認識」、2025年度電気学会若手研究発表会、オンライン、令和7年11月5日