

部 門	競 技 部 門	No.1 登録番号	
-----	---------	-----------	--

No.2

1) 予定開発期間：5ヶ月

2) 予定開発人数：3人

	4月	5月	6月	7月	8月	9月	10月
問題分析							
設計							
実装							
試用・トレーニング							

No.3	<p>実現方法</p> <p>1) 音声の解析アルゴリズム</p> <p>全体の方針 機械学習を行い、合成された音声分離する。</p> <p>・音声データの前処理 合成された音声データ（波形データ）をSTFT（短時間フーリエ変換）を用いてメル周波数スペクトログラムに変換する。本来必要であった時系列ごとの処理や、入力データである48,000*t(s)個のデータを、画像化することによって、最大縦*横（ピクセル）のデータ数まで削減することができ、時系列ごと（リカレントニューラルネットワーク）の処理が不要になり、CNN（畳み込みニューラルネットワーク）での処理が可能になる。</p> <div> </div> <p>問題データ 教師データ作成</p> <p>Sample</p> <p>・機械学習の方針 問題データは単体のサンプルデータ（88種）をランダム（3～20種）で合成して作成し、分割やずらしなども表現する。その際にどの音声合成されたかを表す教師データも作成する。 スペクトログラム化された問題データを入力データとして渡し、CNNで画像の機械学習を行い、どの音声が使われたのかを表す0or1の88個の配列を出力する。 問題データを作成した際に生成された教師データと出力データとの差で誤差をとり、学習を行う。</p> <div> </div> <p>教師データ</p> <p>・本番（本戦）での出力 CNNの出力は、使われた音声は1、使われていない音声は0になっている88種類の配列である。 取り札が44種類のため、日本語・英語のどちらかが1だと、その音声に対応した札を1として出力する。</p>
	<p>2) その他（独創的なところ）</p> <p>・予め、どこまでダウンサンプリングを行うとスペクトログラムの品質が落ちるのかを調べておく。 これを行うことで、CNNの処理を高速化でき、かつあまり劣化していないスペクトログラムを使うことができる。 ・スペクトログラムにメル周波数スペクトログラムを用いることで、人間の聞き取れる周波数をスペクトログラム化する。 これにより、音声の特徴などを見つけやすくなるのではないかと考えた。</p>

No.4	開発言語：Python 開発環境：Anaconda, Visual Studio Code, Neural Network Console, Git Hub
------	---