

題目：LSTM 構造をもつ VGG16 による時系列情報を考慮した呼吸音の自動分類

氏名：浅谷 尚希

呼吸器疾患は、2016 年の世界保健機関(WHO：World Health Organization) の調査によると、世界の死因上位 10 位のうち 4 つを占める重大な疾患で、世界中で死亡者数は 800 万人を超える。現在呼吸器疾患の診察方法は聴診器を用いた呼吸音の聴診が主に行われている。聴診は簡便で、安全安価な診断方法で長年用いられてきたが、精度の高い診断を行うためには、多くの異常呼吸音のパターンを覚える必要があり、また医師の熟練度により診断結果が左右する問題がある。そのため、呼吸音を定量的に分類し「第二の意見」として結果を出力する、コンピュータ支援診断(CAD：Computer Aided Diagnosis)システムが必要である。

ところで 20 世紀後半から、音響工学および分析技術などの発展により、聴診を研究対象とする学者が増加した。近年では呼吸音分類の分野で深層学習を用いた手法が注目されており、CNN(Convolutional Neural Network)を用いた手法が音声分類の分野でも注目を集めている。CNN は画像認識の分野で優れた性能を発揮する技術であるが、画像認識の分野のみならず、音声認識の分野でも高精度の結果を示すことが分かってきた。その手法の流れは、1 次元の音声信号を 2 次元の時間一周波数領域に変換することにより、音声を画像として扱うものである。そして、画像に変換された音声を CNN に入力として与えることにより、特徴を自動抽出し、分類を行う方法である。しかし、CNN のみを用いた音声分類には、音声信号がもつ時系列情報を考慮することが困難であるという問題点が存在する。

このような背景から本論文では、呼吸音データに対し、短時間フーリエ変換を用いたスペクトログラム、Constant-Q 変換を用いた対数周波数スペクトログラム、さらに連続ウェーブレット変換を用いたスカログラムの 3 種類の画像を CNN に入力として与え、得られた特徴の時系列情報を RNN(Recurrent Neural Network)を用い、呼吸音の分類を行った。提案手法を呼吸音 26 症例に適用した結果、分類性能として accuracy = 90[%]が得られ、その有効性が確認できた。

