
力覚と視覚のマルチモーダル深層学習を用いた食品評価に関する研究

(¹ 大阪大, ² (株)J-オイルミルズ)

○加藤優侑¹, 橋本和紀², 堀金智貴², 長畑雄也², 東森充¹

【目的】ヒトは咀嚼において、歯や舌を用いて食品の粉碎・混合・まとめを繰り返し、巧みに食塊を形成していく。このような咀嚼の工学的理解と再現は、食品開発に多大に貢献するものと期待され、従来、人工咀嚼装置に関する研究開発が行われてきた。ただし、食品が固形状からペースト状へと遷移していくようなヒトの食塊形成過程を再現するには至っていなかった。これに対し、著者らのグループは、食品に対する粉碎・混合・まとめ動作を繰り返し行うことで、ヒトのような食塊形成を実現する咀嚼ロボットシミュレータを開発した。本研究では、この装置を用いて食塊形成過程の力覚および視覚データを測定し、これらを入力としたマルチモーダル深層学習によって食品を高精度に評価する手法を提案する。

【方法】咀嚼ロボットシミュレータの上歯部に6軸力覚センサを設置し、粉碎時の鉛直方向反力である咬合力を計測した。下歯部の下方にはカメラを設置し、粉碎時の食塊画像を撮影した。20種類の市販ハンバーグを試験食品とし、それぞれについて12サンプルの人工咀嚼実験を行った(計 $20 \times 12 = 240$ 試行)。1回の人工咀嚼において、粉碎・混合・まとめから構成されるサイクルを10回繰り返し、食塊を形成した。この間の咬合力(力覚)データおよび食塊画像(視覚)データを入力とするマルチモーダル学習モデルを構築し、試験食品の識別実験を行った。学習にはサイクル数 $n = \{1, 5, 10\}$ のデータを使用し、入力条件を変えた際の識別正解率 a を比較することで、提案手法の有効性を検証した。

【結果】入力に咬合力データのみを用いた場合、識別正解率は $a = 0.69$ となった。また、入力に食塊画像データのみを用いた場合、識別正解率は $a = 0.75$ となった。さらに、咬合力データと食塊画像データの両方を用いた場合、識別正解率は $a = 0.79$ となり、最も高い値となった。以上より、咀嚼ロボットシミュレータによって測定した力覚データと視覚データのマルチモーダル深層学習により、高精度の食品評価が実現できる可能性が示唆された。食塊形成過程の力学的特性と幾何学的特性を経時的に評価できるため、食品評価の新機軸となることが期待できる。