

カメラで撮影された医療健康機器から数値を理解する Digital Symbol Understanding System (DSUS) 開発の試み

Development of Digital Symbol Understanding System (DSUS) for understand numerical values from healthcare device pictures

大槻 涼^{1*} 杉山 治² 矢野 修治¹ 松村 耕平¹ 多田 昌裕³ 野間 春生¹ 黒田 知宏²
Ryo Otsuki¹, Osamu Sugiyama², Shuji Yano¹, Kohei Matsumura¹,
Masahiro Tada³, Haruo Noma¹, Tomohiro Kuroda²

¹ 立命館大学 ² 京都大学医学部附属病院 ³ 近畿大学

¹ Ritsumeikan University, ² Kyoto University Hospital, ³ Kindai University

Abstract: In this study, we develop a prototype of the numerical record understanding system from the image on healthcare devices. We divided understanding phase into partial object recognition and understanding of what the numbers represent, and We implemented each with Single Shot Multibox-Detector and Encoder-Decoder model. Verifying the DSUS using the Levenshtein distance, the difference between the DSUS output and the correct answer data for the training data became 4, and for the test data became 6.

1 はじめに

近年、日常生活内の個人の記録を計測するための医療健康機器が普及してきた。これに伴い、服薬記録、健診記録などの個人の記録を一元管理する目的で作られた Personal Health Record (以下、PHR) という仕組みに、医療健康機器から得られた、個人の日々の運動記録などの情報を保存することも提案される [1]。このように、PHR として個人の日常記録を併せて確認できるようにすることで、自己の健康管理や健康維持に役立てようとする動きが出てきた。

このような医療健康機器を用いた PHR 収集システムを開発する中で、システム開発に負担をかける要因の一つとして、多種多様な機器への対応がある。個人の日常的な記録を収集する高価な機器には、スマートフォンや、スマートウォッチなどのウェアラブルデバイスがあり、各機器が持つ通信機能は、NFC 通信や USB 接続と多様である。しかし、歩数計など安価なデバイスの中には、通信機能を全く有さない機器も存在する。

本研究では、人は目視で機器からの情報取得を行なうという観点から、機器が通信機能を有さなくとも、数値の表示さえできれば情報を取得できる Digital Symbol Understanding System (以下、DSUS) の開発を目指す。

2 DSUS の設計

医療健康機器の表示画面から記録を取得するには、機器を一度カメラなどを用いて撮影し、その画像に対して、画像の中から数値の位置を取得する処理、個々の数値の値が何であるかを識別する処理、数値が何を表しているのか数値の意味を取得する処理を行う必要がある。本研究では、数値取得の 1、2 番目の処理を部分数字認識処理、3 番目の処理を数字意味理解処理と呼ぶ。特に、数字の意味理解処理を与えられた数字の意味を持った数値列への並べ替え問題と定義し、この並べ替え問題を解くアプローチによって意味理解を行う。

本研究で提案する DSUS は、歩数計の画面に表示される、歩数 5 桁、月 2 桁、日 2 桁、時間 4 桁を検出し、時間情報を排除した、歩数・月・日の順に並んだ 9 桁の数字列への並べ替えを行う。

DSUS の全体図を、図 1 に示す。本研究では、画面に情報が提示された医療健康機器の一例として歩数計を用いた。図 1 に示すように、DSUS は歩数計画像から数字と各数字の座標を取得する部分数字認識処理部と、数字と各数字の座標を目的的数字列へと並べ替えを行う数字意味理解処理部から成る。部分数字認識処理には Single Shot Multibox-Detector [2] (以下、SSD) を用い、数字意味理解処理部には Recurrent Neural Network で構成された Encoder-Decoder モデル [3] を用いた。

また、SSD が検出できた数字を出力確信度が高い順に、13 個取得する。この 13 という数字は、歩数計の画

*連絡先: (立命館大学)
(滋賀県草津市野路東 1 丁目 1-1)
E-mail:rotsuki@mxdlab.net

