IS IT NECESSARY TO IMPROVE THE COLORECTAL POLYPS DATABASES FOR DETECTION CAD SYSTEMS BASED ON DEEP LEARNING?

L.F. Sánchez-Peralta[1], F.M. Sánchez Margallo, J. Bote Chacón, F. Soria Gálvez,
E. Morcillo Martín, A. Picón Ruiz, J. B. Pagador

[1]Jesús Usón Minimally Invasive Surgery Centre, Cáceres, Spain

Correspondence to: L.F. Sánchez-Peralta (lfsanchez@ccmijesususon.com)

Grants: 732111 – H2020 program

Objectives

Early detection of colorectal cancer is essential. To train and validate CAD systems for automatic detection of polyps using deep learning techniques, it is necessary a proper database of (still) frames that represent the as much clinical variability as possible. The objective of this work is to analyze the clinical representability of the currently available polyp databases.

Methods

4 publicly available databases have been analyzed, both for images (*CVC-ColonDB* y *CVC-ClinicDB*) and video (*ASU-Mayo Clinic polyp database* y *CVC-VideoClinicDB*) to determine the information they provide.

Results

Image databases (*CVC-ColonDB*: 380 images; *CVC-ClinicDB*: 612 images) are smaller than video datasets (*ASU-Mayo Clinic polyp database*: 8.591 frames; *CVC-VideoClinicDB*: 11.954 frames), providing in all cases a binary mask as ground truth for automatic detection. Only *CVC-ClinicDB* and *CVC-VideoClinicDB* provide clinical information. Both indicate Paris classification and *CVC-ClinicDB* also add histologic information (adenomatous/hyperplastic).

Comparing presence of polyps types accordingly to Paris classification with their incidence, 0-I (Ip+Is) polyps appear in greater way in databases (62.42% vs 79.26% vs 57%), while 0-II (IIa+IIb) similarly or under expected (37.58% vs 20.74% vs 39%). None presents 0-IIc polyps.

Conclusions

CVC-ClinicDB and *CVC-VideoClinicDB* provide an added value with clinical information. Nevertheless, including all types of polyps and further clinical information (Kudo classification, NICE classification, location and histology, among other) might provide greater intelligence to CAD systems for automatic detection based on deep learning, helping to endoscopically identify the invasion of the submucosal layer for in-situ decision making about lesion management.