Stichting NIOC en de NIOC kennisbank

Stichting NIOC (www.nioc.nl) stelt zich conform zijn statuten tot doel: het realiseren van congressen over informatica onderwijs en voorts al hetgeen met een en ander rechtstreeks of zijdelings verband houdt of daartoe bevorderlijk kan zijn, alles in de ruimste zin des woords.

De stichting NIOC neemt de archivering van de resultaten van de congressen voor zijn rekening. De website www.nioc.nl ontsluit onder "Eerdere congressen" de gearchiveerde websites van eerdere congressen. De vele afzonderlijke congresbijdragen zijn opgenomen in een kennisbank die via dezelfde website onder "NIOC kennisbank" ontsloten wordt.

Op dit moment bevat de NIOC kennisbank alle bijdragen, incl. die van het laatste congres (NIOC2023, gehouden op donderdag 30 maart 2023 jl. en georganiseerd door NHL Stenden Hogeschool). Bij elkaar bijna 1500 bijdragen!

We roepen je op, na het lezen van het document dat door jou is gedownload, de auteur(s) feedback te geven. Dit kan door je te registreren als gebruiker van de NIOC kennisbank. Na registratie krijg je bericht hoe in te loggen op de NIOC kennisbank.

Het eerstvolgende NIOC vindt plaats op donderdag 27 maart 2025 in Zwolle en wordt dan georganiseerd door Hogeschool Windesheim. Kijk op www.nioc2025.nl voor meer informatie.

Wil je op de hoogte blijven van de ontwikkeling rond Stichting NIOC en de NIOC kennisbank, schrijf je dan in op de nieuwsbrief via

www.nioc.nl/nioc-kennisbank/aanmelden nieuwsbrief

Reacties over de NIOC kennisbank en de inhoud daarvan kun je richten aan de beheerder:

R. Smedinga kennisbank@nioc.nl.

Vermeld bij reacties jouw naam en telefoonnummer voor nader contact.

Artikel

Trainable Vision Systems

Ground-Truth Creation for End-User Trainable Machine Vision Systems Applied to Antibiotic Disk Print Reading

Applied to Antibiotic Disc Print Reading.

Door: Klaas Dijkstra, NHL Hogeschool.

Trefwoorden: end-user software engineering, kunstmatige intelligentie, genetische algoritmes, geometrische patroonherkenning.

With increasing microbial resistance, quickly determining antibiotic resistance is important. For automating this task, detailed knowledge about a task is known primarily by the lab technician (domain expert). In this research an automatic optimization system which can be trained by an end-user is used (figure 1). Collecting an accurate ground-truth is challenging in trainable systems. Surprise-explain-reward is used as a design methodology for the ground-truth collection software. Optimized results using the ground-truth collected by end-users show excellent performance.

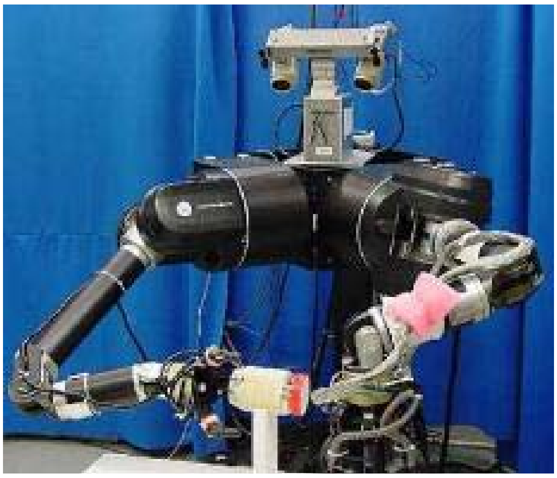

Figure 1. Trainable Machine Vision System.

Introduction

With increasing microbial resistance [1] a method for quickly determining antibiotic susceptibility is required. An analysis which is performed on a regular basis in a microbial laboratory is Antibiotic Susceptibility Testing by Disc Diffusion (ASTDD). This method is used to determine the susceptibility of bacteria to a certain antibiotic found in a sample from a patient. This information is used by the physician to determine which antibiotic to prescribe. With ASTDD a Petri-dish containing agar (which is a bacteria growth medium) is inoculated with sample material from a patient. After this, discs of about 6 mm diameter are placed on the inoculated Petri-dish, where each disc contains a printed abbreviation of the antibiotic contained in the disc. The antibiotic contained in the disc flows into the agar, and at the concentration of antibiotic in the agar is decreasing with the distance to the disc. The antibiotic is diffused into the agar by the disc. The dish is incubated a predetermined number of hours to stimulate bacterial growth. During the incubation process the bacteria starts to grow on the agar at locations for which it still can resist the antibiotic concentration. After incubation the agar contains bacterial growth all over the Petri-dish except for circular areas around the discs. This research focusses mainly on reading or classification of the disc prints.

Problem

In the field of microbiology detailed knowledge about the task is known primarily by the lab technician (domain expert). The technology expert primarily knows about the details of the algorithms used. Typically the gap of cross-domain knowledge is filled by either providing the domain expert with convenient methods for configuring parameter settings, or by educating the technology expert with domain knowledge. In this research the proposed strategy for achieving end-user trainability is to define three roles in the automation of the tasks. First the technology expert limits the possible solutions by choosing algorithms potentially suitable for automating the task. These algorithms can be configured for a specific task by exposing parameters. The domain expert, who usually manually performs the task, defines it by creating a ground-truth using specially designed software. Artificial intelligence is utilized to automatically optimize parameter settings to let the chosen algorithm mimic the ground-truth as close as possible.

This research of end-user trainable systems is different from regular trainable systems because the goal is that no manual parameter configuration is required by neither the domain expert nor the technology expert. Parameter settings are automatically estimated by an optimizer using only the ground-truth. This makes having an accurate ground-truth of hallmark importance.

Creating a ground-truth

Creating a ground-truth from a set of images seems a trivial task, but in practice many aspects make ground-truth creation problematic. The most prominent are: Tedious work; Human error; Software errors; Weak definitions; Inter-operator bias; Intra-operator bias.

The most important problem of specifying a ground-truth is that it has to be specified. This is often tedious, subjective work mostly performed by humans. The other problems mentioned here (except for software errors) are in one way or another result of this first remark. Humans make mistakes in specifying a ground-truth. Software can make mistakes if not properly tested. Another curse of ground-truth collection it that there may be no ground-truth available or definable in the problem context. The ground-truth may be defined for many extreme cases, but not for the more fringe cases ('gray-area') for which there is no definition in the given context. Inter-operator bias means that one end-user can think of a ground-truth differently than another end-user. Intra-operator means that today an end-user will specify a different ground-truth than, over a year, a month or even a minute.

Special software is designed to collect samples of disc prints. This software uses the Surprise-Explain-Reward methodology from the field of end-user software engineering [2]. This methodology alleviates several problems encountered in ground-truth creation. The end-user starts by selecting the labels for disc prints found by the system, and annotates the disc prints. The current partial ground-truth is then used to propose classification for disc prints to the end-user. This creates the surprise. The surprise is explained by telling the end-user that the partial ground-truth is used to help the annotation process. Whenever the classification proposed by the ground-truth tool is wrong, the end-user corrects the annotation and a better model for the discs in the wrongly classified class is searched for automatically. The end-user is rewarded because the proposed classification is improved. Potentially this has many advantages:

- Work becomes less tedious: work will be done by using the already specified partial ground-truth;

- Intra-operator bias is reduced: pre-classification produces identical answers for identical images;
- users get more confidence in the method: direct feedback is realized by the proposed classification.

For the remaining problems: inter-operator bias is not solved using this method and software errors should be avoided by software testing. Finally, disk prints are clearly defined and no gray-area exists.

Experiments

A general computer vision method has been developed to find the disc prints. The first step is using a circular Hough transform to find the round disc, and then a local automatic threshold selection algorithm is used to find the disc print. This algorithm was able to find disc prints on all used images. No additional configuration of parameter settings where necessary.

Experiments for automating disc print reading have been performed using three datasets. Rosco and Oxoid are both disc print vendors. A mixed set has been created by an end-user using the surprise-explain-reward software in a microbial laboratory. The Oxoid set contains 5620 discs divided in 37 classes, the Rosco set contains 1168 tablets divided in 29 classes. The mixed set contains 390 images divided in 36 classes, from which 19 are Oxoid and 17 are Rosco.

As classification algorithm the Blob Matcher was chosen [3]. This is a type of classifier specific for matching distinct geometrical objects with low intra-class variance. A single disc print model is used to classify disc prints by geometric matching. In a sense the Blob Matcher compares single known models to a disc print which has to be classified. Clearly this assumes prior domain knowledge about the nature of disc print reading, because disc prints are distinct geometric objects with low intra-class variance. The remaining parameters of the Blob Matcher are optimized using a Genetic Algorithm [4]. After optimization the classification of disc prints yields 98.5%, 98.7% and 90.7% on the Rosco, Oxoid and Mixed set respectively. This is better than the non-optimized approach: 96%, 93.8% and 87.2%.

Conclusion and future work

Ground-truth is created by end-users using software utilizing surprise-explain-reward methodology. Results show using this ground-truth with an end-user trainable system outperforms manual configuration. Classification results show excellent performance on all three sets. This research focused on disc print reading. Future work could investigate other automation tasks in microbiology.

Literature

[1] European Antimicrobial Resistance Surveillance Network, 2012. *Report for Quarter 4 2012.* Health Protection Surveillance Center, Available at: <http://www.hpsc.ie/hpsc/AZ/MicrobiologyAntimicrobialResistance/EuropeanAntimicrobialResistanceSurveillanceSystemEARSS/EARSSSurveillanceReports/2012Reports/File,13921,en.pdf> [Accessed 3 April 2013].

[2] Ko, A. J., Abraham, R., Beckwith, L., Blackwell, A., Burnett, M., Erwig, M., Scaffidi, C., Lawrance, J., Lieberman, H., Myers, B., Rosson, M. B., Rothermel, G., Shaw, M., and Wiedenbeck, S. The state of the art in end-user software engineering. Computing Surveys 43 (2011).

[3] van de Loosdrecht, J. (2013), `Course Computer Vision', www.vlmv.nl/course [Accessed 11 March 2013].

[4] Eiben, A. E. & Smith, J. E. (2007), Introduction to Evolutionary Computing, 2nd edn, Springer.

Wilt u reageren op dit artikel of de presentatie? Neem dan contact op met:
Klaas Dijkstra; onderzoeker; Kenniscentrum Computer Vision; NHL Hogeschool.
K.Dijkstra@nhl.nl