Number recognition for use in a Sudoku solver

Ralf Vogler VOGLERR@IN.TUM.DE
TU München

Simon Grötzinger GROETZIN@IN.TUM.DE
TU München

Abstract

The goal of this work is to implement number recognition for a Soduko solver. First we will describe the generation of synthetic data, in this case labelled images of numbers. Then we compare different machine learning techniques for supervised learning using the generated data. Once the training is complete, images of real Sudokus are captured using the webcam. The different methods then try to correctly predict each number.

1. Introduction

Number and letter recognition is key in many real-world applications. Digitalizing print media, automated sorting of letters, or recognition of signs in robotics or driver assistance systems require well performing algorithms for character recognition. In the following we analyze the performance of different techniques for number recognition, namely Neural Networks (NN), Support Vector Machines (SVM) and Logistic Regression (LR), by using them in an automated Sudoku solver.

2. Overview

Figure 1 shows the basic data flow and the participating data manipulators. First of all we generate data by sampling position, rotation, size and color of a number and place it based on these values on a white 20x20 pixel image. Examples of 9 such generated images are shown in Figure 2. These images are preprocessed by an octave script and written to a matrix. The corresponding labels are defined by the name of the folder the images are stored in. The $X^{n\text{x}400}$ and the target vector y are the inputs for training NN, SVM and LR, as described in Section 4. Trained models/parameters are saved in ".mat" files that can be loaded by the number prediction. The other input for number prediction is an input matrix of the same form as X with a row entry for each Sudoku field that was detected to contain a number.

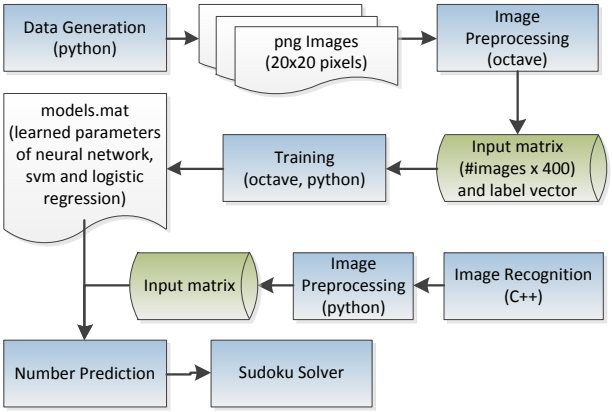

Figure 1. Overview

3. Generation of Training Data

3.1. Image Generation

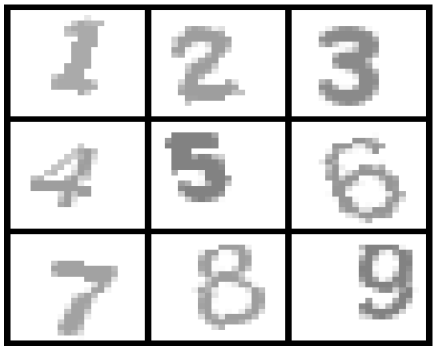

Figure 2. Examples of generated numbers

Appearing in *Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel, 2010.

3.2. Image Processing

4. Training

4.1. Neural Network

Applying a NN seems promising as it was already successfully applied to the quite similar problem of detecting handwritten digits. ref!For the NN we choose a simple architecture with 25 hidden sigmoidal units (+ bias unit), as shown in figure 3. The size of the input layer is 400 (given by the 20 x 20 pixel images) and the size of the output layer is defined by the number of classes, in this case the numbers 1 to 9. This simple architecture turned out to perform pretty well compared to more complex ones with a higher number of hidden layers. Training is performed with error backpropagation, thus minimizing the error function

$$E(w) = \sum_{n=1}^{N} E_n(w)$$
$$\text{with} \quad E_n(w) = \frac{1}{2}\sum_{k}(y_{nk} - t_{nk})$$
$$\text{and} \quad y_k = \sum_i w_{ki} x_i$$

subject to w. For NNs we test two implementations.

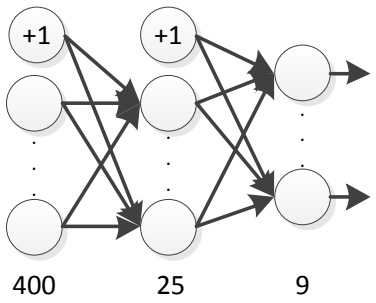

Figure 3. Neural Network with one hidden layer

On the one hand we use a partly self-developed octave code for two-layer NNs which was written as part of the stanford machine learning course[1], on the other hand we include PyBrain (Schaul et al., 2010) for comparison and easier evaluation of more sophisticated network architectures. To find a good value for the regularization parameter λ we perform a grid search with k-fold cross-validation which gives us a good performance when choosing $\lambda = 1$.

[1] http://www.ml-class.org

4.2. Support Vector Machines

The next technique we apply to the discussed problem are Support Vector Machines (Bishop, 2006) which try to solve the following optimization problem:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\xi_i \\ \text{subject to} \quad & y_i(w^T\phi(x_i) + b) - 1 + \xi_i \geq 0 \\ \text{and} \quad & \xi_i \geq 0. \end{aligned}$$

Using a radial basis function (RBF) as kernel we perform a grid search over C and γ to find those parameters that provide highest accuracy on the test set. We use the open source library "libsvm" (Chang & Lin, 2011) wrapped by some octave scripts for setting up data sets and performing grid search with k-fold cross-validation. Grid search with cross validation resulted in accuracies illustrated in Figure 4. From the best parameter pairs we choose the one that performs best on the validation set, which gives us C =? and λ =?.

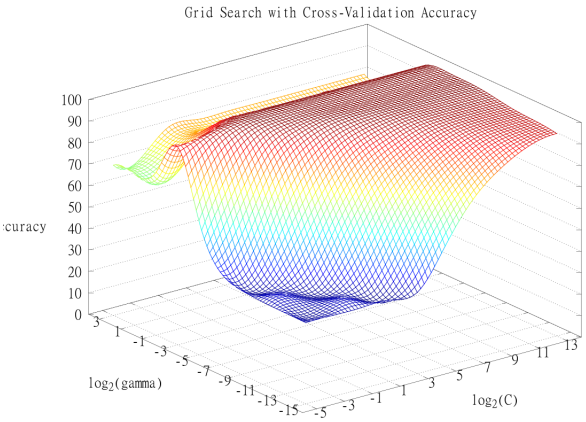

Figure 4.

4.3. Logistic Regression

Finally we apply Logistic Regression to the datasets. Again we use a (vectorized) octave implementation for learning and prediction. Therefore we minimize the (regularized) error function

$$E(w) = \frac{1}{n}(-y^T ln(\Phi w) + (1 - y^T) ln(1 - \Phi w)) + \frac{\lambda}{2n} w^T w$$

with respect to w. Minimization is performed by an algorithm[2] that was provided during the stanford machine learning course. Grid search and k-fold cross-validation recommend a value for the regularization parameter λ.

5. Image Recognition and Preprocessing

Image recognition is written in C by extensive use of the open source library OpenCV[3]. Images captured by the webcam are continously analyzed to find a Sudoku square. Images are transformed to greyscale, then an adaptive threshold algorithm is used to preprocess the image for contour detection by OpenCV. OpenCV returns a list of contours which are further checked for certain criteria like having four corners or being of certain size. Since the contours are not detected precise enough lines are fitted through six checkpoints on each edge in order to achieve subpixel accuracy. The content enclosed by the corners is perspectively warped to fit a predefined square size. Finally, the thresholded Sudoku field is saved to disk. A python script is then used to split the whole field into 9x9 subfields with each subfield being flood filled around the corners for noise reduction. The format of the subfields now equals to the one provided by the generated images which allows us to apply the trained NN/SVM/LR models and, finally, apply a Sudoku solver.

6. Evaluation

For evaluation of the different techniques we define the following datasets:

- d1: 4500 generated images
- d2: $\sim$ 1000 gathered images
- d3: D1 $\cup$ D2
- dVal: images of a captured Sudoku

7. Future Work

- include information if a sudoku is unsolvable
- online learning

[2]http://sprinkler.googlecode.com/svn/trunk/regression/fmincg.m

[3]http://opencv.willowgarage.com/wiki/

Table 1. Classification accuracies for Neural Network, SVM and Logistic Regression on different data sets.

Data set	NN	SVM	LR	Winner
D1	95.9± 0.2	96.7± 0.2	96.7± 0.2	?
D2	83.3± 0.6	80.0± 0.6	?	?
D3	61.9± 1.4	83.8± 0.7	?	?
DVal	74.8± 0.5	78.3± 0.6	?	?

8. References

References

Bishop, Christopher M. *(Pattern Recognition and Machine Learning) By Bishop, Christopher M. (Author) Hardcover on (08 , 2006)*. Springer, 8 2006.

Chang, C.C. and Lin, C.J. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., Rückstieß, T., and Schmidhuber, J. Pybrain. *The Journal of Machine Learning Research*, 11:743–746, 2010.

9. Electronic Submission

As in the past few years, ICML 2010 will rely exclusively on electronic formats for submission and review. We assume that all authors will have access to standard software for word processing, electronic mail, and web file transfer. Authors who do not have such access should send email with their concerns to `icml10.programchairs@gmail.com`.

9.1. Templates for Papers

Electronic templates for producing papers for submission are available for LaTeX and Microsoft Word. Templates are accessible on the World Wide Web at: `http://www.icml2010.org/`

Send questions about these electronic templates to `icml10.programchairs@gmail.com`.

If you are using the templates, the formatting instructions below will already be enforced. The content rules you must follow are:

- The maximum paper length is 8 pages.
- Do not include author information or acknowledgments in your initial submission.

- Place figure captions *under* the figure (and omit titles from inside the graphic file itself). Place table captions *over* the table.
- References must include page numbers whenever possible and be as complete as possible. Place multiple citations in chronological order.

Please see below for details on each of these items.

9.2. Submitting Papers

Submission to ICML 2010 will be entirely electronic, via a web site (not email). The URL and information about the submission process are available on the conference web site at

`http://www.icml2010.org/sub.html`

Paper Deadline: The deadline for paper submission to ICML 2010 is Monday, February 1, 2010, at 11:59 p.m. Samoa time. If your full submission does not reach us by this date, it will not be considered for publication. There is no separate abstract submission this year.

Anonymous Submission: To facilitate blind review, no identifying author information should appear on the title page or in the paper itself. Section 10.3 will explain the details of how to format this.

Simultaneous Submission: ICML will not accept any paper which, at the time of submission, is under review for another conference or has already been published. This policy also applies to papers that overlap substantially in technical content with conference papers under review or previously published. ICML submissions must not be submitted to other conferences during ICML's review period. Authors may submit to ICML substantially different versions of journal papers that are currently under review by the journal, but not yet accepted at the time of submission. Informal publications, such as technical reports or papers in workshop proceedings which do not appear in print, do not fall under these restrictions.

To ensure our ability to print submissions, authors must provide their manuscripts in **PDF** format. Furthermore, please make sure that files contain only Type-1 fonts (e.g., using the program `pdffonts` in linux or using File/DocumentProperties/Fonts in Acrobat). Other fonts (like Type-3) might come from graphics files imported into the document.

Authors using **Word** must convert their document to PDF. Most of the latest versions of Word have the facility to do this automatically. Submissions will not be accepted in Word format or any format other than PDF. Really. We're not joking. Don't send Word.

Those who use **LaTeX** to format their accepted papers need to pay close attention to the typefaces used. Specifically, when producing the PDF by first converting the dvi output of LaTeX to Postscript the default behavior is to use non-scalable Type-3 PostScript bitmap fonts to represent the standard LaTeX fonts. The resulting document is difficult to read in electronic form; the type appears fuzzy. To avoid this problem, dvips must be instructed to use an alternative font map. This can be achieved with something like the following commands:

dvips -Ppdf -tletter -G0 -o paper.ps paper.dvi
ps2pdf paper.ps

Note that it is a zero following the "-G". This tells dvips to use the config.pdf file (and this file refers to a better font mapping).

Another alternative is to use the **pdflatex** program instead of straight LaTeX. This program avoids the Type-3 font problem, however you must ensure that all of the fonts are embedded (use `pdffonts`). If they are not, you need to configure pdflatex to use a font map file that specifies that the fonts be embedded. Also you should ensure that images are not downsampled or otherwise compressed in a lossy way.

Note that the new 2010 style files use the `hyperref` package to make clickable links in documents. If this causes problems for you, add `nohyperref` as one of the options to the `icml2010` usepackage statement.

9.3. Reacting to Reviews

We will continue the ICML tradition in which the authors are given the option of providing a short reaction to the initial reviews. These reactions will be taken into account in the discussion among the reviewers and area chairs.

9.4. Submitting Final Camera-Ready Copy

The final versions of papers accepted for publication should follow the same format and naming convention as initial submissions, except of course that the normal author information (names and affiliations) should be given. See Section 10.3.2 for details of how to format this.

The footnote, "Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute." must be modified to "Appearing in *Proceedings of the 27^{th} International Conference on Machine Learning*, Haifa, Israel, 2010.

For those using the LaTeX style file, simply change `\usepackage{icml2010}` to

```
\usepackage[accepted]{icml2010}
```

Authors using **Word** must edit the footnote on the first page of the document themselves.

Camera-ready copies should have the title of the paper as running head on each page except the first one. The running title consists of a single line centered above a horizontal rule which is 1 point thick. The running head should be centered, bold and in 9 point type. The rule should be 10 points above the main text. For those using the LaTeX style file, the original title is automatically set as running head using the `fancyhdr` package which is included in the ICML 2010 style file package. In case that the original title exceeds the size restrictions, a shorter form can be supplied by using

```
\icmltitlerunning{...}
```

just before `\begin{document}`. Authors using **Word** must edit the header of the document themselves.

10. Format of the Paper

All submissions must follow the same format to ensure the printer can reproduce them without problems and to let readers more easily find the information that they desire.

10.1. Length and Dimensions

Papers must not exceed eight (8) pages, including all figures, tables, references, and appendices. Any submission that exceeds this page limit or that diverges significantly from the format specified herein will be rejected without review.

The text of the paper should be formatted in two columns, with an overall width of 6.75 inches, height of 9.0 inches, and 0.25 inches between the columns. The left margin should be 0.75 inches and the top margin 1.0 inch (2.54 cm). The right and bottom margins will depend on whether you print on US letter or A4 paper, but all final versions must be produced for US letter size.

The paper body should be set in 10 point type with a vertical spacing of 11 points. Please use Times Roman typeface throughout the text.

10.2. Title

The paper title should be set in 14 point bold type and centered between two horizontal rules that are 1 point thick, with 1.0 inch between the top rule and the top edge of the page. Capitalize the first letter of content words and put the rest of the title in lower case.

10.3. Author Information for Submission

To facilitate blind review, author information must not appear. If you are using LaTeX and the `icml2010.sty` file, you may use `\icmlauthor{...}` to specify authors. The author information will simply not be printed until `accepted` is an argument to the style file. Submissions that include the author information will not be reviewed.

10.3.1. Self-Citations

If your are citing published papers for which you are an author, refer to yourself in the third person. In particular, do not use phrases that reveal your identity (e.g., "in previous work (**?**), we have shown ...").

Do not anonymize citations in the reference section by removing or blacking out author names. The only exception are manuscripts that are not yet published (e.g. under submission). If you choose to refer to such unpublished manuscripts (**?**), anonymized copies have to be submitted as Supplementary Material via CMT. However, keep in mind that an ICML paper should be self contained and should contain sufficient detail for the reviewers to evaluate the work. In particular, reviewers are not required to look a the Supplementary Material when writing their review.

10.3.2. Camera-Ready Author Information

If a paper is accepted, a final camera-ready copy must be prepared. For camera-ready papers, author information should start 0.3 inches below the bottom rule surrounding the title. The authors' names should appear in 10 point bold type, electronic mail addresses in 10 point small capitals, and physical addresses in ordinary 10 point type. Each author's name should be flush left, whereas the email address should be flush right on the same line. The author's physical address should appear flush left on the ensuing line, on a single line if possible. If successive authors have the same affiliation, then give their physical address only once.

A sample file (in PDF) with author names is included in the ICML2010 style file package.

10.4. Abstract

The paper abstract should begin in the left column, 0.4 inches below the final address. The heading 'Abstract' should be centered, bold, and in 11 point type.

The abstract body should use 10 point type, with a vertical spacing of 11 points, and should be indented 0.25 inches more than normal on left-hand and right-hand margins. Insert 0.4 inches of blank space after the body. Keep your abstract brief, limiting it to one paragraph and no more than six or seven sentences.

10.5. Partitioning the Text

You should organize your paper into sections and paragraphs to help readers place a structure on the material and understand its contributions.

10.5.1. Sections and Subsections

Section headings should be numbered, flush left, and set in 11 pt bold type with the content words capitalized. Leave 0.25 inches of space before the heading and 0.15 inches after the heading.

Similarly, subsection headings should be numbered, flush left, and set in 10 pt bold type with the content words capitalized. Leave 0.2 inches of space before the heading and 0.13 inches afterward.

Finally, subsubsection headings should be numbered, flush left, and set in 10 pt small caps with the content words capitalized. Leave 0.18 inches of space before the heading and 0.1 inches after the heading.

Please use no more than three levels of headings.

10.5.2. Paragraphs and Footnotes

Within each section or subsection, you should further partition the paper into paragraphs. Do not indent the first line of a given paragraph, but insert a blank line between succeeding ones.

You can use footnotes[4] to provide readers with additional information about a topic without interrupting the flow of the paper. Indicate footnotes with a number in the text where the point is most relevant. Place the footnote in 9 point type at the bottom of the column in which it appears. Precede the first footnote in a column with a horizontal rule of 0.8 inches.[5]

10.6. Figures

You may want to include figures in the paper to help readers visualize your approach and your results. Such artwork should be centered, legible, and separated from the text. Lines should be dark and at least 0.5 points thick for purposes of reproduction, and text should not appear on a gray background.

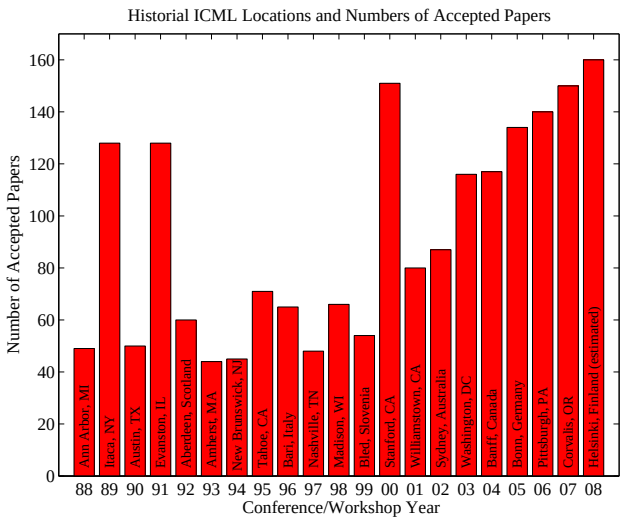

Figure 5. Historical locations and number of accepted papers for International Machine Learning Conferences (ICML 1993 – ICML 2008) and International Workshops on Machine Learning (ML 1988 – ML 1992). At the time this figure was produced, the number of accepted papers for ICML 2008 was unknown and instead estimated.

Label all distinct components of each figure. If the figure takes the form of a graph, then give a name for each axis and include a legend that briefly describes each curve. Do not include a title inside the figure; instead, the caption should serve this function.

Number figures sequentially, placing the figure number and caption *after* the graphics, with at least 0.1 inches of space before the caption and 0.1 inches after it, as in Figure 5. The figure caption should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left. You may float figures to the top or bottom of a column, and you may set wide figures across both columns (use the environment `figure*` in LaTeX), but always place two-column figures at the top or bottom of the page.

10.7. Algorithms

If you are using LaTeX, please use the "algorithm" and "algorithmic" environments to format pseudocode. These require the corresponding stylefiles, algorithm.sty and algorithmic.sty, which are supplied with this package. Algorithm 1 shows an example.

[4]For the sake of readability, footnotes should be complete sentences.

[5]Multiple footnotes can appear in each column, in the same order as they appear in the text, but spread them across columns and pages if possible.

Algorithm 1 Bubble Sort

Input: data x_i, size m
repeat
  Initialize $noChange = true$.
  for $i = 1$ **to** $m - 1$ **do**
    if $x_i > x_{i+1}$ **then**
      Swap x_i and x_{i+1}
      $noChange = false$
    end if
  end for
until $noChange$ is $true$

Table 2. Classification accuracies for naive Bayes and flexible Bayes on various data sets.

Data set	Naive	Flexible	Better?
Breast	95.9± 0.2	96.7± 0.2	√
Cleveland	83.3± 0.6	80.0± 0.6	×
Glass2	61.9± 1.4	83.8± 0.7	√
Credit	74.8± 0.5	78.3± 0.6	
Horse	73.3± 0.9	69.7± 1.0	×
Meta	67.1± 0.6	76.5± 0.5	√
Pima	75.1± 0.6	73.9± 0.5	
Vehicle	44.9± 0.6	61.5± 0.4	√

10.8. Tables

You may also want to include tables that summarize material. Like figures, these should be centered, legible, and numbered consecutively. However, place the title *above* the table with at least 0.1 inches of space before the title and the same after it, as in Table 2. The table title should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left.

Tables contain textual material that can be typeset, as contrasted with figures, which contain graphical material that must be drawn. Specify the contents of each row and column in the table's topmost row. Again, you may float tables to a column's top or bottom, and set wide tables across both columns, but place two-column tables at the top or bottom of the page.

10.9. Citations and References

Please use APA reference format regardless of your formatter or word processor. If you rely on the LaTeX bibliographic facility, use `natbib.sty` and `icml2010.bst` included in the style-file package to obtain this format.

Citations within the text should include the authors' last names and year. If the authors' names are included in the sentence, place only the year in parentheses, for example when referencing Arthur Samuel's pioneering work (**?**). Otherwise place the entire reference in parentheses with the authors and year separated by a comma (**?**). List multiple references separated by semicolons (**???**). Use the 'et al.' construct only for citations with three or more authors or after listing all authors to a publication in an earlier reference (**?**).

Authors should cite their own work in the third person in the initial version of their paper submitted for blind review. Please refer to Section 10.3 for detailed instructions on how to cite your own papers.

Use an unnumbered first-level section heading for the references, and use a hanging indent style, with the first line of the reference flush against the left margin and subsequent lines indented by 10 points. The references at the end of this document give examples for journal articles (**?**), conference publications (**?**), book chapters (**?**), books (**?**), edited volumes (**?**), technical reports (**?**), and dissertations (**?**).

Alphabetize references by the surnames of the first authors, with single author entries preceding multiple author entries. Order references for the same authors by year of publication, with the earliest first. Make sure that each reference includes all relevant information (e.g., page numbers).

10.10. Software and Data

We strongly encourage the publication of software and data with the camera-ready version of the paper whenever appropriate. This can be done by including a URL in the camera-ready copy. However, do not include URLs that reveal your institution or identity in your submission for review. Instead, provide an anonymous URL or upload the material as "Supplementary Material" into the CMT reviewing system. Note that reviewers are not required to look a this material when writing their review.

Acknowledgments

Do not include acknowledgements in the initial version of the paper submitted for blind review.

If a paper is accepted, the final camera-ready version can (and probably should) include acknowledgements. In this case, please place such acknowledgements in an unnumbered section at the end of the paper. Typically, this will include thanks to reviewers who gave useful comments, to colleagues who contributed to the ideas, and to funding agencies and corporate sponsors that provided financial support.

References

Bishop, Christopher M. *(Pattern Recognition and Machine Learning) By Bishop, Christopher M. (Author) Hardcover on (08 , 2006)*. Springer, 8 2006.

Chang, C.C. and Lin, C.J. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., Rückstieß, T., and Schmidhuber, J. Pybrain. *The Journal of Machine Learning Research*, 11:743–746, 2010.