Single character OCR using Support Vector Machine

Olli Jarva & Jarno Rantanen

Aalto University School of Science

olli@jarva.fi & jarno@jrw.fi

Abstract

This paper describes optical character recognition for bitmap characters using Support Vector Machine with RBF kernel, including parameter search and bitmap normalization. Error rate with k-fold cross validation (k=20, n=42152) was 11%.

KEYWORDS: SVM, Support Vector Machine, RBF, OCR, Character Recognition

1 Dataset description

Training dataset:

- n=42152
- 16 x 8 black and white bitmaps
- Lowercase characters, n=26 (a-z)

Testing dataset:

- n=10000
- Same format

2 Character preprocessing

- Minimize noise by moving characters to bottom left corner. 0.5% improvement

3 What is SVM?

- http://www.csie.ntu.edu.tw/ cjlin/papers/guide/guide.pdf
- http://www.ivanciuc.org/Files/Reprint/Ivanciuc_SVM_CCR_2007_23_291.pdf
- RBF kernel: $K(x_i, x_j) = exp(-\gamma||x_i - x_j||^2), \gamma > 0$

Optimization problem $(x_i, y_i), i = 1, ..., l$ where x_i is ... ([**?**]):

minimize w, b, ξ: $\frac{1}{2}w^T w + C \sum_{i=1}^{l} \xi_i$

subject to $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$

4 RBF kernel parameter search

γ and C

- Initial search space 2**x for x in range(-15, 15)
- Select best area for next round
- Validate by taking final arguments and calculating error rates for +- few percent for both variables.

5 Results and performance

- k-fold cross validation: k=20, error rate 11%
- k-fold cross validation: k=5, error rate 11.5%
- One iteration with training set n=40000 and validation set n=2152 about 17 min with 2.1GHz Xeon (single thread)
- about 300MB of memory for training set n=42152
- Predicting one character: about 2 milliseconds

6 Quick comparison to other algorithms

- kNN (+PCA/LDA)
- ...?

[1]

References

[1] D. Albanese, R. Visintainer, S. Merler, S. Riccadonna, G. Jurman, and C. Furlanello. mlpy: Machine learning python, 2012.