Experimental Results of Autoencoder Features With Fine-tuning

***Abstract*—**

I. Introduction

In the previous experiments (ICDAR, HIP, and DAS), Kai, Oussama, and Me used Mathias' framework to extract autoencoder features. Generally, these features were learned by autoencoders without fine-tuning (if I am correct). In our HIP paper and also my thesis, we observed that: *a significant number of autoencoder features are redundant or irrelevant for layout classification.* This observation can be explained by the fact that autoencoders (without fine-tuning) are trained for reconstruction, rather than classification.

Autoencoders with fine-tuning are a kind of supervised feature learning technique (if I am correct). The learned features are supposed to be optimal or nearly optimal for classification. In other words, the learned features are supposed to be less redundant and less irrelevant for classification, compared to features learned by autoencoders without fine-tuning. This is an important motivation of our following experiments.

II. CNN Architecture

III. Sequential Forward Selection

IV. Experimental Results

In the experiments, we investigate redundancy and irrelevance of features leaned by autoencoders with fine-tuning. We use autoencoder features to perform layout analysis, which classifies pixels on an image into five classes: out of page, background, comments, decorations, and text. Sequential Forward Selection (SFS) is used for feature selection. The results are shown in Table I.

Table I shows that for all classifiers using features learned by autoencoders with fine-tuning, generally feature dimension is decreased significantly, and comparable accuracy is obtained. So from Table I, we have the following observations. Regarding features learned with fine-tuning, a significant number of them are redundant or irrelevant for layout classification. This observation is similar with our previous observations about features learned by autoencoders without fine-tuning.

In addition, regarding redundancy and irrelevance, there is no significant difference between features learned with and without fine-tuning. In other words, regardless of whether autoencoder features are learned with or without fine-tuning, they are equally redundant and irrelevant for layout classification in general.

V. Conclusions

References

[1]

TABLE I: Investigation of redundancy and irrelevance of features leaned by autoencoders with fine-tuning. The used classifiers are Support Vector Machine (SVM), Naive Bayes (NB), a decision algorithm (C4.5), and k-Nearest Neighbors (k-NN) where k=1.

Classifier	Features	Dimension	Accuracy
SVM	Full features Selected features	46	
NB	Full features Selected features	46 7	88.10 85.81
C4.5	Full features Selected features	46 9	77.34 81.83
k-NN	Full features Selected features	46 22	77.91 77.28