Converting Printed Arabic Text Images into Editable Text

An efficient method to automatically convert printed Arabic text images to editable text.

The Invention

The invention is a methodology that can be used to accurately recognize printed Arabic text. It can also be used for other foreign languages. The methodology has proven its accuracy and efficiency. It is based on a machine learning technique. Some of the benefits of converting scanned documents into digital files are: avoid retyping, electronic search, edit text, Save storage and ease of Access.

Market Need

When we have a hardcopy of an Arabic document that needs updating, a printed Arabic magazine article that we need to create an editable text version of it, or when we receive an electronic image of an Arabic brochure and we need to update the text, we can convert all of these documents into a digital format with fully editable text in a matter of minutes using an optical character recognition software that uses the proposed methodology efficiently and accurately.

Moreover, such methodology can be tuned for recognizing Arabic car plates with an accuracy approaching 100% for clear photos of the plates.
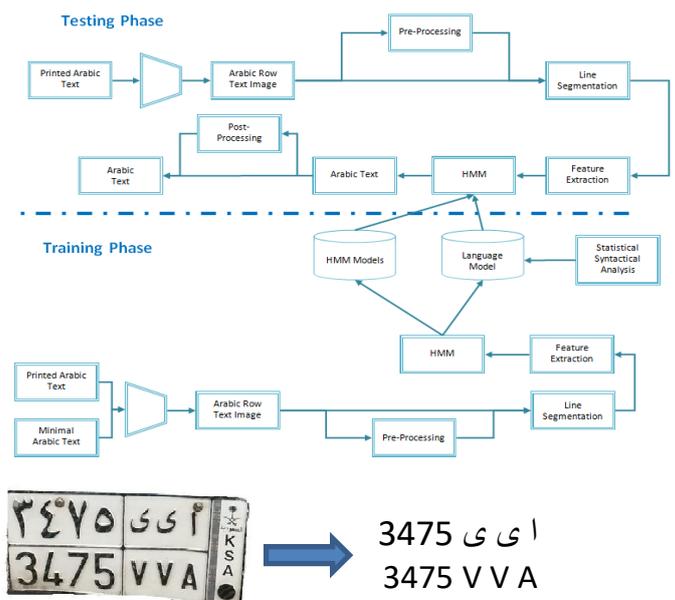
Competitive Advantage

Several aspects of our technique resulted in the high recognition rates. Our technique is based on a novel hierarchical sliding window technique with overlapping and non-overlapping windows which is reported for the first time. The achieved high recognition rates using the methodology indicate the effectiveness of the used technique. The achieved accuracy has reached as high as 99.9% for single font recognition.

Readiness for Market / Looking for a Development Partner

The methodology has been successfully implemented through setting up a machine learning prototype for training and testing. The analysis of the results has shown high performance with accuracies that never been reported.

The next step needed is to implement the method in a stand alone system that can be used through the internet.

One we to commercialize the expected product is to provide it as an online service that can be used by different users via the internet.

ا ى ى 3475
3475 V V A

Patent Protection

This technology is protected by the issued US patent number US8150160.

For further information please contact:
Email: ip-license@kfupm.edu.sa
Telephone: +966-13-860 7685