# Abstract

*Optical Character Recognition (OCR) is a mechanical or electronic conversion of scanned images handwritten, typed, or printed text into encoded to computer. Studies on lately OCR provide a method with good accuracy and computational speed is getting lighter. Computing is getting lighter allows the engine with lower specifications such as a smart phone capable of working on OCR.*

*In this final project, OCR methods applied using Contour Analysis with complex number as contour representation and feature extraction methods using Intercorelation and Autocorrelation function. For feature preselection it use the longest common subsequence algorithm (LCS). For contour smoothing Rames-Douglas-Peucker (RDP) algorithm is used.*

*Test results shows the contour smoothing using RDP with optimal epsilon value configuration by 2.0 and 32 contour points, improves 4% accuracy than does not using it. Feature preselection using LCS is more effective than the Euclidean distance. LCS with optimal tolerance value configuration, which is at 0.1, produces 68 contour as result of preselection from total 676 contour. Euclidean distance with threshold value of 0.3 produces 158 contour as result of preselection process from total 676 contour to match.*

**Keywords**: OCR, *feature preselection*, Rames-Douglas-Peucker *algorithm*, *longest common subsequence*, *contour analisis.*