Notes on English Alphabet Engine

The current English alphabet engine is capable of identifying all the characters of English language (both upper and lower case alphabets). The engine can handle only one character at a time. The code structure of the C++ program is as follows :

Initially, a txt file containing the co-ordinate points of a character is taken and stored into a structure called FILE_STROKEINFO. The co-ordinate points are stored in batches of strokes.

1. SVM parameters and the RuleList are loaded into the program.
    I. Algorithm for recognizing the character is coded in the function Recognize_English () which returns the corresponding ASCII code value of the character.
    II. Recognize_English() first checks whether the current stroke corresponds to a straight line or a curve or just a dot.
    III. If it is not a straight line, the points are normalized based on height of the stroke followed by interpolation to a fixed number of points. In our case, it is 200 based on convenience.
    IV. Curve length based normalization is done later followed by smoothening and finally the stroke points are again interpolated to 32 points.
    V. The stroke id is then recognized based on SVM. The nature of a straight line stroke and that of dot is recognized based on it's aspect ratio.
    VI. Then, the English alphabet is identified after comparison with the Rule_List.