

(12) PATENT APPLICATION PUBLICATION

(21) Application No.202141031333 A

(19) INDIA

(22) Date of filing of Application :13/07/2021

(43) Publication Date : 16/07/2021

(54) Title of the invention : A SYSTEM OF COMPUTER VISION AND MACHINE LEARNING FOR OPTICAL CHARACTER RECOGNITION

(51) International classification	:G06K0009000000, G06F0040177000, G06K0009460000, G06F0016901000, B25J0009160000	(71)Name of Applicant : 1)SRIVIDHYA E Address of Applicant :D/o. S ELANGO VAN, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, FACULTY OF COMPUTING, SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY (DEEMED TO BE UNIVERSITY), JEPPIAR NAGAR, RAJIV GANDHI SALAI, CHENNAI 600119, TAMIL NADU, INDIA Tamil Nadu India 2)Dr. KISHORE VERMA S 3)SONYA A 4)GOKULAKRISHNAN S 5)MURUGANATNHAM A 6)ABRAR AHMED K 7)JAYANTHI S 8)RANJANA S 9)Dr. REVATHI A R 10)Dr. SRIDEVI D 11)MYTHREYI N R 12)Dr. MANIKANDAN J
(31) Priority Document No	:NA	(72)Name of Inventor : 1)SRIVIDHYA E 2)Dr. KISHORE VERMA S 3)SONYA A 4)GOKULAKRISHNAN S 5)MURUGANATNHAM A 6)ABRAR AHMED K 7)JAYANTHI S 8)RANJANA S 9)Dr. REVATHI A R 10)Dr. SRIDEVI D 11)MYTHREYI N R 12)Dr. MANIKANDAN J
(32) Priority Date	:NA	
(33) Name of priority country	:NA	
(86) International Application No	:NA	
Filing Date	:NA	
(87) International Publication No	: NA	
(61) Patent of Addition to Application Number	:NA	
Filing Date	:NA	
(62) Divisional to Application Number	:NA	
Filing Date	:NA	

(57) Abstract :

In the present invention, we depict an efficient and highly scalable parallel architecture to segment input images containing tabular data with and without borders into cells and reconstruct the tabular data while preserving the tabular format. Many state-of-the-art algorithms have been developed that can be used for the purpose of OCR but extracting text from images containing tables while preserving the structure of the table still remains a challenging task. The performance improvement thus made can be used to ease the tedious task of digitizing tabular data in bulk. The same architecture can be used for regular OCR applications to improve performance if the data is in huge quantities.

No. of Pages : 23 No. of Claims : 4